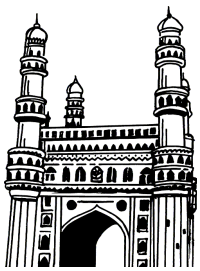
Rahul's ✔
Topper's Voice



M.B.A.

IV Semester

(Osmania University)

Latest 2022 Edition

BUSINESS INTELLIGENCE

- ☞ **Study Manual**
- ☞ **FAQ's and Important Questions**
- ☞ **Short Question & Answers**
- ☞ **Multiple Choice Questions**
- ☞ **Fill in the blanks**
- ☞ **Solved Previous Question Papers**
- ☞ **Solved Model Paper**

Price ~~249-00~~ 199-00

- by -

WELL EXPERIENCED LECTURER

Rahul Publications™

Hyderabad. Ph : 66550071, 9391018098

M.B.A.

IV Semester

(Osmania University)

BUSINESS INTELLIGENCE

Inspite of many efforts taken to present this book without errors, some errors might have crept in. Therefore we do not take any legal responsibility for such errors and omissions. However, if they are brought to our notice, they will be corrected in the next edition.

Sole Distributors :

☎ : 66550071, Cell : 9391018098

VASU BOOK CENTRE

Shop No. 2, Beside Gokul Chat, Koti, Hyderabad.

Maternity Hospital Opp. Lane, Narayan Naik Complex, Koti, Hyderabad.
Near Andhra Bank, Subway, Sultan Bazar, Koti, Hyderabad -195.

BUSINESS INTELLIGENCE

CONTENTS

STUDY MANUAL

PREVIOUS SOLVED QUESTION PAPERS

SOLVED MODEL PAPER

SYLLABUS

UNIT - I

Introduction to Business Intelligence (BI)

Definition, History and Evolution, Styles of Business Intelligence, Benefits if Business Intelligence, Real-time Business Intelligence, Business Intelligence Value Chain, Architecture Business Intelligence.

UNIT - II

Data Warehousing and Data Mining

a) **Date Ware Housing (DWH):** Definition, Characteristic, types, Date ware housing frame work, DwH 3 tier architecture, Alternative Architectures, Data ware housing Integration, Data ware housing- Development Approaches, Real time Data ware housing.

b) **Data Mining :** Definition, Characteristic, Benefits, Date Mining Functions, Data Mining Applications, Data Mining techniques and tools. Text Mining, Web Mining.

UNIT - III

Business Performance Measurement (BPM)

Definition, BPM vs BI, Summary of BPM Process, Performance Measurement, BPM Methodologies, BPM Architecture and Applications, Performance Dash boards.

UNIT - IV

Business Analytics and Data Visualization

(a) Business Analytics - Definitions, Tools and techniques of BA, Advanced Business Analytics Business Analytics and Web, Usage, Benefits and Success of Business Analytics.

(b) Data Visualization: Definition, New Direction in Data Visualization, GIS, GIS vs GPS

UNIT - V

Business Intelligence Implementation

(a) Implementing Business Intelligence – Implemental Factors, Critical Success factors of Business Implemental, Managerial Issues related to BI Implementation. Business Intelligence and Integration Implementation – Types, Need, Level of Business Intelligence Integration.

(b) Emerging trends in Business Intelligence Implementation- Social Net works and Business Intelligence, Collaborative Decision Making, RFID and Business Intelligence, Reality Mining.

Contents

Frequently Asked & Important Questions

UNIT - I

1. Briefly explain the various benefits of Business Intelligence.

Ans : **(June-19, Imp.)**

Refer Unit-I, Q.No. 7.

---

2. Elucidate the concept of Real-time Business Intelligence.

Ans : **(Oct.-20, June-19, Imp.)**

Refer Unit-I, Q.No. 13.

---

3. Explain the history and evolution of business intelligence.

Ans : **(June-18)**

Refer Unit-I, Q.No. 4.

---

4. Explain the role of business intelligence in organizational development.

Ans : **(June-18)**

Refer Unit-I, Q.No. 12.

---

5. What are the characteristics and need of business intelligence?

Ans : **(Imp.)**

Refer Unit-I, Q.No. 2.

UNIT - II

1. Discuss about data warehousing frame work.

Ans : **(June-18)**

Refer Unit-II, Q.No. 6.

---

2. Discuss about 3 Tier Architectures of Data Ware Housing.

Ans : **(Oct.-20)**

Refer Unit-II, Q.No. 7

---

3. What are the types of Data warehousing?

Ans : **(May-19)**

Refer Unit-II, Q.No. 3.

4. Discuss about data mining applications.

Ans : **(Oct.-20, Imp.)**

Refer Unit-II, Q.No. 22.

5. What are the Characteristics of Data Mining?

Ans : **(June-18)**

Refer Unit-II, Q.No. 18.

UNIT - III

1. Compare and Contrast Business Performance measurement and business intelligence.

Ans : **(June-19)**

Refer Unit-III, Q. No. 3.

2. Outline the summary of Business Process Management Process.

Ans : **(Oct.-20)**

Refer Unit-III, Q. No. 5.

3. Explain briefly about Performance Dash Boards.

Ans : **(Oct.-20, Imp.)**

Refer Unit-III, Q. No. 23.

4. Elucidate the Architecture of Business Performance Measurement.

Ans : **(June-19, Imp.)**

Refer Unit-III, Q. No. 20.

5. Explain the various applications of business performance measurement.

Ans : **(June-18, Imp.)**

Refer Unit-III, Q. No. 21.

UNIT - IV

1. What are the various Tools and Techniques of Business Analytics ?

Ans : **(June-19, Imp.)**

Refer Unit-IV, Q.No. 3.

2. What are the Usage, Benefits and Success of Business Analytics ?

Ans : **(Oct.-20, June-18, Imp.)**

Refer Unit-IV, Q.No. 8.

3. Explain about GIS (Geographical Information System).

Ans : **(Imp.)**

Refer Unit-IV, Q.No. 16.

4. What are the techniques of Data Visualization?

Ans : **(June-19, Imp.)**

Refer Unit-IV, Q.No. 15.

5. Compare and contrast GIS Vs GPS.

Ans : **(June-18)**

Refer Unit-IV, Q.No. 18.

UNIT - V

1. Elucidate Critical Success factors of Business Implementation ?

Ans : **(Oct.-20, June-18)**

Refer Unit-V, Q.No. 2.

2. Explain the Managerial Issues Related to Business Intelligence Implementation?

Ans : **(June-18)**

Refer Unit-V, Q.No. 3.

3. Describe briefly about Collaborative Decision Making.

Ans : **(June-19)**

Refer Unit-V, Q.No. 13.

4. Describe briefly about Mobile Social Networking.

Ans : **(June-19, June-18, Imp.)**

Refer Unit-V, Q.No. 11.

5. What do you understand by reality mining ?

Ans : **(Imp.)**

Refer Unit-V, Q.No. 16.

UNIT I

Introduction to Business Intelligence (BI)

Definition, History and Evolution, Styles of Business Intelligence, Benefits if Business Intelligence, Real-time Business Intelligence, Business Intelligence Value Chain, Architecture Business Intelligence.

1.1 Introduction to Business Intelligence

1.1.1 Definition

Q1. Define Business Intelligence.

Ans : **(Imp.)**

Introduction

Business Intelligence (BI) is a set of ideas, methodologies, processes, architectures, and technologies that change raw data into significant and useful data for business purpose. Business Intelligence can handle large amounts of data to help identify and evolve new opportunities for the business. Making use of these new opportunities and applying a productive scheme on it can provide a comparable market benefit and long-term stability

Business Intelligence (BI) technologies provide chronicled, present and predictive view of business operations. Common functions of enterprise Intelligence technologies are reporting, online analytical processing, analytics, data excavation, process excavation, business performance management, benchmarking, text mining, predictive analytics and prescriptive analytics.

Meaning of Business Intelligence

Business Intelligence is a set of processes, architectures, and technologies that convert raw data into meaningful information that drives profitable business actions. It is a suite of software and services to transform data into actionable intelligence and knowledge.

Business Intelligence has a direct impact on organization's strategic, tactical and operational business decisions. Business Intelligence supports fact-based decision making using historical data rather than assumptions and gut feeling.

Business Intelligence tools perform data analysis and create reports, summaries, dashboards, maps, graphs, and charts to provide users with detailed intelligence about the nature of the business.

Q2. What are the Characteristics and need of business intelligence?

Ans : **(Imp.)**

Characteristics of Business Intelligence

1. It is created by procuring data and information for use in decision-making.
2. It is a combination of skills, processes, technologies, applications and practices.
3. It contains background data along with the reporting tools.
4. It is a combination of a set of concepts and methods strengthened by fact-based support systems.
5. It is an extension of Executive Support System or Executive Information System.
6. It collects, integrates, stores, analyzes, and provides access to business information
7. It is an environment in which business users get reliable, secure, consistent, comprehensible, easily manipulated and timely information.
8. It provides business insights that lead to better, faster, more relevant decisions.

Need of Business Intelligence

1. Analyzing a large amount of data to provide historical, current and predictive views of the business operations.

2. To support business decisions from operations to strategic.
3. To process real-time data and develop effective business strategies.
4. Easy interpretation of unstructured data and using it to support decisions.

Q3. What are the various stages of Business Intelligence ?

Ans :

The various stages of Business Intelligence are as follows :

1. Data Sourcing

Defining the data to be loaded into the system. Usually BI applications gathers data from a data warehouse (Data marts, OLTP or OLAP).

2. ETL (Extract Transform Load)

Extracting the source data and transforming per business rules and loading into the Data Warehouses.

3. Data Warehousing

Storing transformed data into various Data warehouses types and making it available for business analysis.

4. Data Analysis

Applying various techniques like data mining, text mining, Process mining to identify trends and patterns in business operations.

5. Decision Making

Based on the reports, dashboards and alerts from previous stage, making valuable business decisions and bench marking future growth.

1.1.2 History and Evolution of Business Intelligence

Q4. Explain the history and evolution of business intelligence.

Ans : **(June-18)**

History Business Intelligence

The term BI was coined by the Gartner Group in the mid-1990s. However, the concept is much older; it has its roots in the Management Information Systems (MIS) reporting systems of the 1970s. During that period, reporting systems were static and two- dimensional and had no analytical capabilities. In the early 1980s, the concept of executive information systems (EIS) emerged. This concept expanded the computerized support to top-level managers and executives.

Some of the capabilities introduced were dynamic multidimensional (ad hoc or on-demand) reporting, forecasting and prediction, trend analysis, drill down to details, status access, and critical success factors (CSFs). These features appeared in dozens of commercial products until the mid-1990s. Then the same capabilities and some new ones appeared under the name BI.

Evolution of Business Intelligence

Today, a good BI-based enterprise information system contains all the information executives need. So, the original concept of EIS was transformed into BI. By 2005, BI systems started to include artificial intelligence capabilities as well as powerful analytical capabilities. Figure illustrates the various tools and

techniques that may be included in a BI system. It illustrates the evolution of BI as well. The tools shown in Figure provide the capabilities of BI. The most sophisticated BI products include most of these capabilities; others specialize in only some of them.

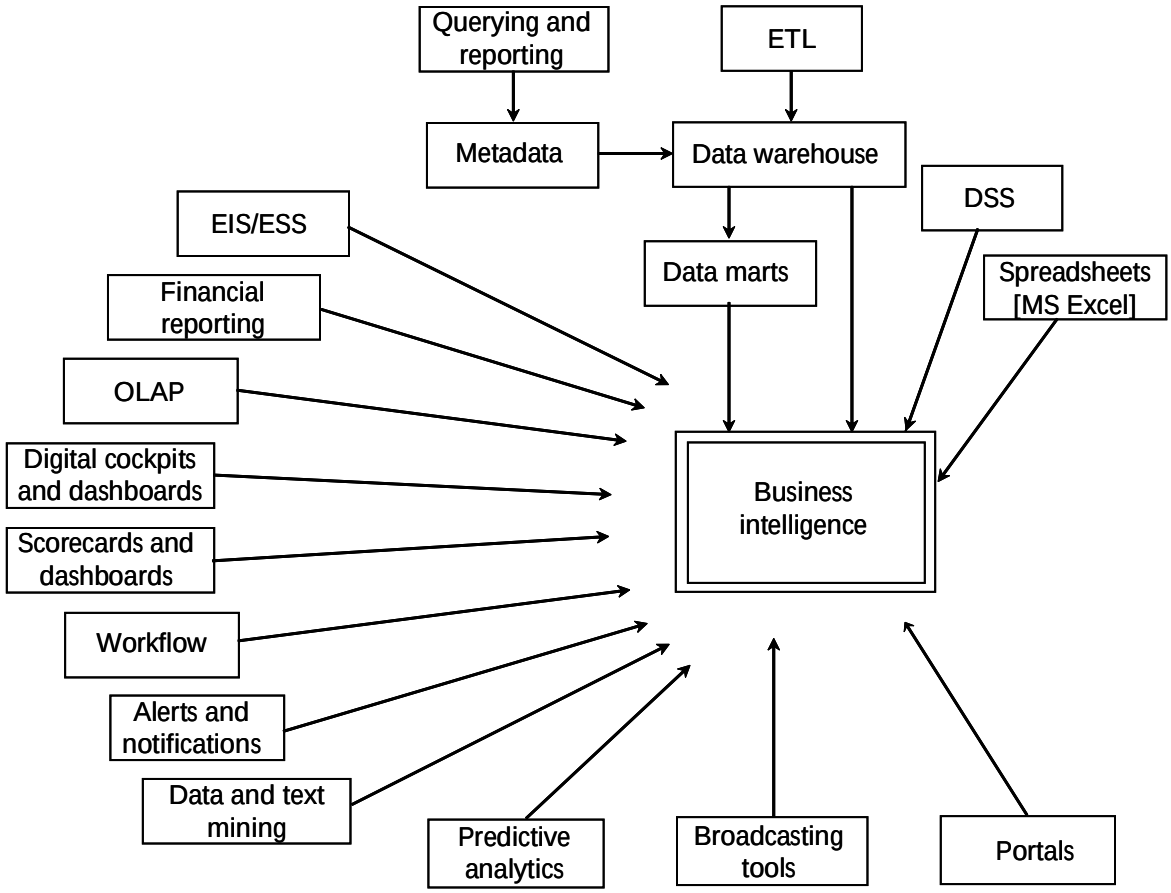


Figure : Evolution of BI

Q5. What are the components of business intelligence?

Ans :

The Major Components of Business Intelligence (BI)

Depending on the business type, some Business Intelligence strategies will be more useful than others.

i) OLAP (Online Analytical Processing)

This component of BI allows executives to sort and select aggregates of data for strategic monitoring. With the help of specific software products, a certification in business intelligence helps business owners can use data to make adjustments to overall business processes.

ii) Advanced Analytics (or) Corporate Performance Management (CPM)

This set of tools allows business leaders to look at the statistics of certain products or services. For instance, a fast food chain may analyze the sale of certain items and make local, regional and national modifications on menu board offerings as a result. The data could also be used to predict in which markets a new product may have the best success.

iii) Realtime Business Intelligence

In a mobile society, this particular component of Business Intelligence is becoming increasingly popular. Using software applications, a business can respond to real-time trends in email, messaging systems or even digital displays. Because it is all in real-time, an entrepreneur can announce special offers that take advantage of what's going on in the immediate. Marketing professionals can use data to craft creative limited-time specials such as a coupon for hot soup on a cold day. CEO's may be interested in tracking the time of day and location of customers as they interact with a website so marketing can offer special promotions in real-time while the client is engaged on the website.

iv) Data Warehousing

Data warehousing lets business leaders sift through subsets of data and examine interrelated components that can help drive business. Looking at sales data over several years can help improve product development or tailor seasonal offerings. Data warehousing can also be used to look at the statistics of business processes including how they relate to one another. For instance business owners can compare shipping times in different facilities to look at which processes and teams work most efficiently. Data warehousing also involves storing huge amounts of data in ways that are beneficial to different divisions within the company.

v) Data Sources

This component of Business Intelligence involves various forms of stored data. It is about taking the raw data and using software applications to create meaningful data sources that each division can use to positively impact business. Business Intelligence analysts using this strategy may create data tools that allow data to be put into a large cache of spreadsheets, pie charts, tables or graphs that can be used for a variety of business purposes. For example, data can be used to create presentations that help to structure attainable team goals. Looking at the strategic aspect of data sources can also help organizations make fact-driven decisions that take into account a more holistic view of the needs of the company.

vi) Business Intelligence Tools

Under the umbrella term of Business Intelligence, there are many tools that are used to analyze the various components of BI and construct them into actual problem-solving actions. Today's ubiquitous use of the Internet and the great entrepreneurial spirit of our free-market economy have fostered niche markets and start-ups, as well as, consulting firms and other business ventures that have helped build many BI tools. These specific business tools can help leaders look at components of their business in more depth and detail. The most common tools in use today include business and data analytics, predictive analytics, cloud technology, mobile Business Intelligence, Big Data consultation and visual analytics.

1.1.3 Styles of Business Intelligence

Q6. List out the various styles of Business Intelligence.

Ans :

1. Enterprise Reporting

Broadly deployed pixel-perfect report formats for operational reporting and score cards/dashboards targeted at information consumers and executives.

2. Cube Analysis

OLAP slice-and-dice analysis of limited data sets, targeted at managers and others who need a safe and simple environment for basic data exploration within a limited range of data.

3. Ad Hoc Query and Analysis

Full investigative query into all data, as well as automated slice and-dice OLAP analysis of the entire database - down to the transaction level of detail if necessary. Targeted at information explorers and power users.

4. Statistical Analysis and Data Mining

Full mathematical, financial, and statistical treatment of data for purposes of correlation analysis, trend analysis, financial analysis and projections. Targeted at the professional information analysts.

5. Alerting and Report Delivery

Proactive report delivery and alerting to very large populations based on schedules or event triggers in the database. Targeted at very large user populations of information consumers, both internal and external to the enterprise.

1.1.4 Benefits of Business Intelligence

Q7. Briefly explain the various benefits of Business Intelligence.

Ans : **(June-19)**

Main Benefits of Business Intelligence

1. Time Savings

One of the key advantages of Business Intelligence is that most business processes are automated, which generates important savings both in time and costs, and in turn contributes to increasing overall productivity levels.

2. Easier and quicker access to information

It is clear that, over the last few years, the amount of business data has propelled. Therefore, it is important that companies focus their efforts on digitizing and collecting their data via document management software. However, it is equally important that Business Intelligence tools offer easily accessible information that clearly shows the evolution of the data and, this way, allows companies to anticipate future events.

3. Correct and relevant decisions

In order to stand apart from the competition, reduce costs, and increase profits, a company must make intelligent decisions. To do this, these decisions must logically be based on trustworthy and relevant data, and this is exactly where traditional methods begin to fail.

Q8. What are the various modules of business intelligence ?

Ans :

Modules of business intelligence

1. Dashboards
2. Key Performance Indicators
3. Graphical OLAP
4. Forecasting and Predictive analysis
5. Graphical Reporting

1. Dashboards

Business Intelligence dashboards can provide a customized snapshot of daily operations, and assist the user in identifying problems and the source of those problems, as well as providing valuable, up-to-date information about financial results, sales and other critical information all in one place.

2. Key Performance Indicators

Business Intelligence provides simplified KPI management and tracking with powerful features, formulae and expressions, and flexible frequency, and threshold levels. This module enables clear, concise definition and tracking of performance indicators for a period, and measures performance as compared to a previous period. Intuitive, color highlighters ensure that users can see these indicators in a clear manner and accurately present information to management and team members. Users can further analyse performance with easy-to-use features like drill down, drill through, slice and dice and graphical data mining.

3. Graphical OLAP

Graphical Business Intelligence (GBI) OLAP technology makes it easy for your users to find, filter and analyse data, going beyond numbers, and allowing users to visualize the information with eye-catching, stunning displays, and valuable indicators and gauges, charts, and a variety of graph types from which to choose.

4. Forecasting and Predictive Analysis

Our predictive analysis uses historical product, sales, pricing, financial, budget and other data, and forecasts the measures with numerous time series options, e.g., year, quarter, month, week, day, hour or even second to improve your planning process.

5. Graphical Reporting

BI Reports delivers web-based BI reports to anyone (or everyone) in the organization within minutes! The BI suite is simple to use, practical to implement and affordable for every organization. With our BI reporting and performance reporting module, you just point- and-click and drag-and-drop and you can instantly create a report to summarize your performance metrics, or operational data.

Q9. List out the various disadvantages of business intelligence.

Ans :

Some of the major Business Intelligence disadvantages are:

(i) Piling of Historical Data

The major objective of Business intelligence system is to stockpile past data about a firm's deals and reveal it in such a way that it permits professionals in decision making. On the flip side, this information generally amounts to a small portion of what the firms actually require to function, besides its restrained worth. While in other situations, the user may not have interest in historical data as many markets that the company regulates are in frequent alteration.

(ii) Cost

Business intelligence at times can be a little too much for small as well as for medium sized enterprises. The use of such system can be expensive for basic business transactions.

(iii) Complexity

Another disadvantage of business intelligence could be its complexity in implementation of data. It can be so intricate that it can make business techniques rigid to deal with. In the view of such premise, many business experts have predicted that these intricacies can ultimately throttle any business.

(iv) Muddling of commercial settings

Business Intelligence can cause commercial settings to turn out to be much more muddled.

(v) Limited use

Like all improved technologies, business intelligence was first established keeping in consideration the buying competence of affluent firms. Even today BI system cannot be afforded by most of the companies. Although, traders in the past few years have started modifying their services towards medium and small sized industries, but the fact is that many of such firms does not consider them to be highly essential, for its complexity.

(vi) Time Consuming Implementation

Many firms in today's fast paced industrial scenario are not patient enough to wait for the execution of Business intelligence in their organization. It takes around 18 months for data warehousing system to completely implement the system.

Q10. Explain the future trends of business intelligence.

Ans :

3 Future Trends of Business Intelligence

In simple words, business intelligence tools need to be – Simplified, Specialized and Personalized.

1. Simplification

Every product needs to end up in the hands of its final user. BI's real users are the decision makers in a company, statuesquely called the "Executives". BI has successfully transitioned from the obscure backrooms of IT into the realms of analysts. However, the journey will end when the Executives (most of whom are not very tech-savvy) can use appropriate tools to aid their primary job – making decisions.

2. Specialization

As humans we seek specialist medical opinion for anything more than the flu as we want

expert analysis on matters critical to us as individuals. Similarly, each industry and each firm is as unique as would be their data and practices. The era of generic, domain-agnostic systems will give way to systems that have deep understanding of the client's business vertical. Specialist systems built on emerging technologies such as machine learning, knowledge engineering, artificial intelligence and the like will provide invaluable insights to the Executives on matters specific to their industry vertical to enable better decision making.

3. Personalization

Most of the current solutions are too hard to use and too complicated to implement and maintain. They provide limited utility to the bottom-line focussed Executives. Smart BI, with sophisticated technology that is almost transparent, resulting in personal digital assistants will unlock the vast under-tapped potential of the BI promise.

The combination of these three factors will result in the next wave of explosion in business intelligence – penetrating the mass-market / Small and Medium Businesses (SMBs). The wealth at the bottom of the pyramid!

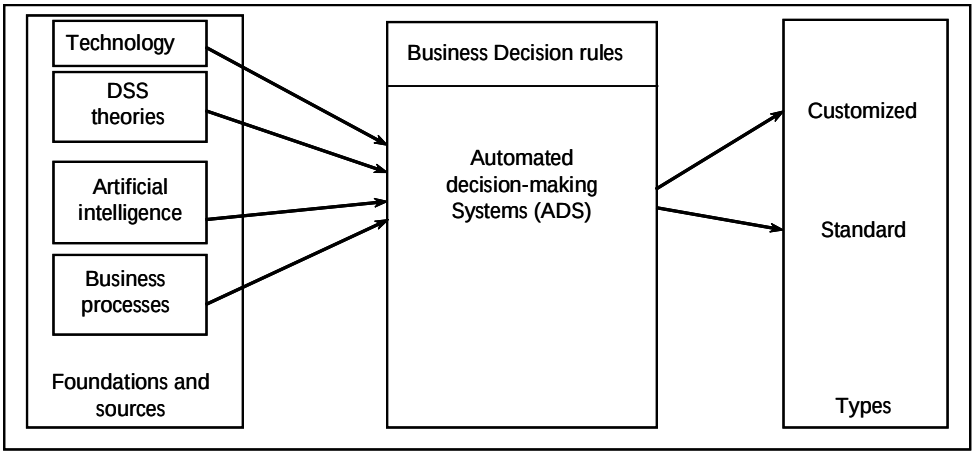
At Einsights, we have been diligently applying these three principles to the world of analytics – simplification, specialization and personalization. It's our raison d'etre. We are striving to make the decision making process easier through the intelligent use of technology. Decisions are not made in a vacuum; they are context-sensitive and so should be the analytics that lead to the decision making moment.

Q11. Briefly explain Automated decision systems.

Ans :

Automated Decision System (ADS) is a new rule-based system that is introduced to support decision making process, so, it Walso called as Decision Automated System (DAS). It provides solution to various problems such as one functional area to a certain repetitive managerial problem, industry problem (i.e. pricing of the product and services).

The following figure describes the framework of automated decision-making system.



Initially, the automated decision systems are used in airline industries. In airlines industries, ADS are used to price the tickets dynamically depending up on the actual demand. They provides rule-based solutions rather than the various management science approaches that provides model based solutions to generic structured problems. Hence, they are known as revenue management or revenue optimization systems.

Some of the examples of business rules included in this rule-based systems are as follows,

(i) It provides a discount of 'p' to the non-business passengers only when 80% of seats in the flight from source (i.e. Sydney) to destination (i.e., Colambia) are sold out before 4 days to departure.

(ii) A user can be provided with $20,000 cred it line, when the user has own house and earns more than $2,00,000 per year.

(iii) If the cost of a product is more than $2,000 then the company or organization owns that product only once in a year, and the purchasing agent does not requires any specific approval.

Q12. Explain the role of business intelligence in organizational development.

Ans : **(June-18)**

There are many reasons why companies choose business intelligence solutions. If you're contemplating BI software for your business, you may be wondering if it's worth the time, effort, and expense to add it to your existing software suite.

BI-Tools was curious about whether BI software lived up to its hype and whether companies felt they got the maximum benefit from it. The results may help solidify your decision whether or not to add BI to your software suite.

- **Better planning and analysis:** Companies felt that BI systems helped them the most with faster reporting, planning, and analysis. 64% of responding companies ranked their ability to report, plan and analyze data as "good" after implementing a business intelligence suite.
- **Increased accuracy:** Among the companies surveyed, 56% felt that business intelligence data increased the accuracy of their business analysis and planning.
- **Helped considerably with sales forecasting:** Among the many tasks that companies felt that business intelligence data helped with, 57% ranked sales forecasting and planning as the area receiving the most benefit from BI data. Other areas where they felt that BI date provided assistance was in customer behavior analysis (40%) and a unified view of customers (32%).
- **Improved pricing and offers:** Pricing and offer optimization benefited somewhat from the implementation of a BI system. 27% of respondents felt that the additional data derived from their BI system helped them improve their pricing structure to become more competitive, as well as improve the attractiveness of their offers.

1.2 Real-time Business Intelligence

Q13. Elucidate the concept of Real-time Business Intelligence.

Ans : **(Oct.-20, June-19)**

Business users (Application managers, role managers and risk managers etc.) are progressively demanding access to the real-time data (or unstructured data) that includes the contents of data warehouse. For instance, the buses in foreign countries like Texas, Houston are designed with several facilities like data gathering devices, traffic controllers that are capable of accessing and changing the traffic signal/lights interval. Thus, making the system more efficient and reliable. An organization is said to be successful (or survive) only when it deals with the real-time data updates. For instance, the cities nearer to the sea and with low elevation distance make use of the real-time data collection and analysis. By using this equipment the officials will get to know the present situations of the cities, like the working of the flood-controlling pumps which takes an active participation in saving the people when the floods occurs. Interestingly, it is also used in several parts of the world like Netherlands and New Orleans.

Earlier, Business Intelligence tools and dataware housing mainly focuses on assisting managers in preparing tactical and strategic decisions with the help of the historical data. In the year 2003, a shift has started using real-time data warehousing for making operational decisions. Some of the IT vendors like Microsoft, Oracle, IBM and other are preferring to actively use these technologies.

Hewlett-packard has turned towards an adaptive enterprise strategy to deliver the on-demand computing.

Today, there is a faster development in the business intelligence (BI) software which generates the real-time data for the real-time business analysis and decision making. This growth includes acquiring the right information to operational and tactical personnel who are responsible for managing the short-term features to run an organization. Thus, they can use new business analysis and up-to-the-minute results for taking decisions.

Now a days, the demand of the current data is more. So, most of the IT manager have queries like how the real-time business analysis can be conducted, etc. Many of the real-time projects are in development and deployment stages only and the demand towards the real-time application will continues to increase. For instance, the growth of business rules management and ADS, enforces to implement most of the automated business processes in data ware house environment. The reply of a real-time system is mandatory when the processes need immediate data updates then it required to answer the analytical queries. Thus, the response time for data mining, OLAP and queries is approximately zero.

Real-time business analysis can be achieved by more frequently updating the data warehouses. In the year 2003, the daily basis updation of data warehouses became famous and the time interval is continues to decrease. Moreover, business analysis applications are being deployed apart from real-time queries. For instance, such application can quickly recognize, customer buying patterns on the basis of store displays, provides an instantaneous modifications for replacement or display itself. Some of the other applications are fraud detection, revenue management, call-centre support transportation and other financial transactions.

Real-time requirement includes various queries that updates the data. Hence, these requirements can also change the design view of a database, OLAP, data warehouses and several data mining tools. Alternatively, substantial business value demonstration can be adopted for a business process. Some of the examples of web-based and real-time business intelligence softwares are as follows,

(i) Live Business (Data mirror)

(ii) Live sheet for Excel (KnowNow)

(iii) NetIQ corporation

(iv) Supply chain analytics and Business Intelligence series 8(cognos).

(v) Web intelligence (business objects)

(vi) DB2 intelligent minor scoring (IMS; IBM)

(vii) Databeacon (Databeacon Inc.)

(viii) SonicMQ (Sonic Software)

(ix) Supply chain Intelligence Suite (SAS)

(x) Power Center standard and Advanced versions (Informatica).

Q14. What are the benefits of Real-time Business Intelligence ?

Ans :

Benefits of Real-time Business Intelligence:

1. Optimizing Asset Performance

Identifying potential problems in field assets, grids and wells, utility distribution equipment can help avoid unplanned service interruptions in advance. Also, any delay in responding to an equipment's failure and maintenance issues can lead to operational inefficiency. This inefficiency can be regulated by monitoring the performance and health of assets to predict decay points or equipment failure. Hence, analyzing real-time machine data or operational data can help solve these challenges. Real-time insights into asset health, peak periods, supply and demand analysis, and abnormal conditions can help improve asset performance.

Data analytics-driven business intelligence solutions that provide analyses via rich visuals and statistics are what utilities need to adopt. The capabilities provided by BI systems can help enhance the reliability, capacity and availability of the assets, which in turn improves asset performance. This can help ensure smoother operations at peak periods, eliminating downtime occurrence. It also allows them to trigger maintenance schedule operations at regular intervals to receive notifications on asset health.

2. Enhancing Customer Experience

Intelligence from assets, smart grids and SCADA enriched with customer data can provide critical insights into a customer's utility usage. Utilities can build systems of intelligence around the consumption pattern to empower consumers with usage insights and influence their usage behavior. Segmentation at a micro level to develop a 360-degree view of customer's usage behavior can help bring the best possible ways to optimize their usage. Utilities can thus provide cost-effective plans delivering personalized experience, thereby improving loyalty, customer lifetime value and reducing customer churn. They can now provide customers with self-service intelligence capabilities to view and manage their utility usage, consumption pattern, historical usage, billing, payment and abnormal usage conditions.

As a good instructive example, find out how a team of asset data analytics consultants implemented BI capabilities for a giant water utility company to open a channel for water conservation and to deliver differentiated customer service.

3. Improving Operational Efficiency

Real-time business intelligence systems bring forth an interaction process for utility operators to operatively view and analyze operational data. BI dashboards enable them to monitor everything from meter reading to uptime for every single minute. It helps gain high visibility of historical comparison of asset performance, access real-time data and promptly react to make better decisions in a timely fashion. It also helps find opportunities to improve operational efficiency as and when an event occurs.

Dashboards also generate comprehensive performance reports to help utility operators develop an optimal course of action depending on specific key performance indicators. It also displays the overall performance of the assets in real time on a periodic basis, i.e., year-to-year, monthly, weekly, daily or hourly. These capabilities of business intelligence automate and help optimize asset performance, reduce risk and operational costs, and enhance business responsiveness.

4. Dynamic Forecasting and load Management

Business intelligence acts as a modelling platform to determine patterns, spot trends and understand predictable behavior in order to ramp up overall efficiency. Real-time and accurate utility demand forecasting is crucial to successfully managing supply and demand operations in a timely fashion. Utilities have to leverage real-time or historical data of consumption, usage, supply and weather constraints to perform complex analysis. This would help them to identify key performance indicators and patterns for dependable forecasting of asset failures and performance. They can also determine correlation between grid performance and grid conditions for maintenance.

Workflow improvements gained by integrating business intelligence solutions with existing automation processes can help avoid asset failures, thus resulting in asset longevity and higher uptime. Getting actionable insights, prescriptions and foresight is necessary for better decision making towards preventive maintenance, quality of service and outage management. The most efficient way to build systems of intelligence is to take advantage of an IoT and data analytics partner's expertise to leverage Azure IoT and Cortana Intelligence Suite, which brings forth the powerful capabilities to transform data into intelligent actions.

5. Prevention of Utility Loss and Fraud Management

As enormous volumes of data streams driven from smart meters, grids and sensors become readily available, utilities can perform real-time and predictive analytics to gain critical insights. This can help operators continuously monitor the vital signs of utility distribution systems, reliably and quickly assess system

integrity, and gain hidden insights. This in turn helps them in outage and fault detection, and to determine anomalies on supply-distribution lines, thereby optimizing utility distribution. Thus minimizing utility loss and providing early detection of supply, theft, leaks, consumption and over usage.

6. The Bottom Line

The benefits of real-time business intelligence are numerous, tangible and evident, which allows utilities to make faster and smarter decisions, whether it's sending notifications to customers within minutes of an outage or responding to spikes in demand even before they happen. Thus, empowering utilities with the power of facts, transforming the way they do business forever. With business intelligence solutions, utilities can optimize asset performance, create visibility and boost operational efficiency that supports changing business processes. Utilities that infuse real-time intelligence systems into their business operations can secure their unique position in today's market.

1.3 BUSINESS INTELLIGENCE VALUE CHAIN

Q15. Explain briefly about Business Intelligence Value Chain.

Ans :

Business Intelligence Value Chain

In competitive Marketplace, it is vital for every business enterprise whether small or big to cope with the pace of the market growth. Intelligence is the ability to learn and understand new situations by updating yourself with current information. This is why organizations are becoming more dependent on data or information for improving development of product and services that can outsell their competitors. So, they are primarily relying on analytical database solutions like data warehouses. However, in the entire process these information and later applying them in action describes a chain of methods called value chain. Value chain is the process that converts data into information and then applies that knowledge in taking productive business decision, and the outcome is the ultimate value.

i) Data resources

The first step of the value chain is the collection, management and analysis of data to produce information that can help in taking different types of business decision. Data warehouse generally contains various kinds of data both related and unrelated but here BI tries to co-relate, analyze and explore valuable information from those data resource. However, it becomes essential to first list out similar or related data that can assist in identifying healthy business opportunities. These can be descriptive, quantitative or qualitative.

ii) Information engineering

Here, information engineering refers to the kind of information needed and way for its presentation that should meet current and future demand of business. Data is generally available in raw forms and is base for extracting information that sustains the knowledge environment for developing business strategies.

iii) Knowledge Gain

It's the process of knowing something through experience (experiment), instinct, beliefs that make you aware about different things and help to act intelligently. Knowledge is the core aspect to understand what is relevant and significant to business issues and it can be used for further advantage. So interpreting available information can assists in gaining knowledge that leads to business intelligence.

iv) Action taken

Data analysis and information exchange create knowledge that are used in resolving complex problems, innovating new ideas and taking various decisions like strategic, tactic for specific business purposes. So combining knowledge and information used to take action; it includes the sharing and accumulating knowledge for gaining successful business ideas.

v) Outcome

Information and knowledge offer an insight to the kind of action is to be taken and is essential to integrate knowledge and business ideas to positive business outcomes. A well-planned business strategy with adequate information generally supports to reduce cost, save time, optimize resources, increase revenue, satisfy customer, and over all to reach the desired business goals.

Hence the method or process that includes data collection, management, information extraction, and knowledge application in developing business strategies collectively forms the value chain in Business Intelligence. However, the entire business value creation depends on various business domains including strategic planning, financial management, R&D, human resource, marketing, sales, corporate governance and lastly the most important information system.

The data source can be added with a new value in every step of the business intelligence value chain. The process involved in the business intelligence value chain is,

Data resources $\rightarrow$ Information $\rightarrow$ knowledge worker $\rightarrow$ Business intelligence $\rightarrow$ Business strategies $\rightarrow$ Business goals

1. Data resource helps in generating the information using the information using the information engineering process.
2. Information helps in generating the knowledge worker in a knowledge environment.
3. Knowledge worker helps in generating business intelligence in the intelligent learning enterprise.
4. Business intelligence helps in generating the business strategies.
5. Business strategies helps in the generation of the goals of an enterprise or organization. Which is the last ster in business intelligence value chain.

1.4 Architecture of Business Intelligence

Q16. Explain the Architecture of Business Intelligence.

Ans : **(Oct.-20)**

The architecture of Business Intelligence (BI) system consists of four major components. They are,

1. Data Warehouse
2. Business Analytics
3. Business Performance Management (BPM)
4. User Interface.

The following figure describes the high-level architecture of business intelligence (BI) and also the relationship among the four major parts of business intelligence.

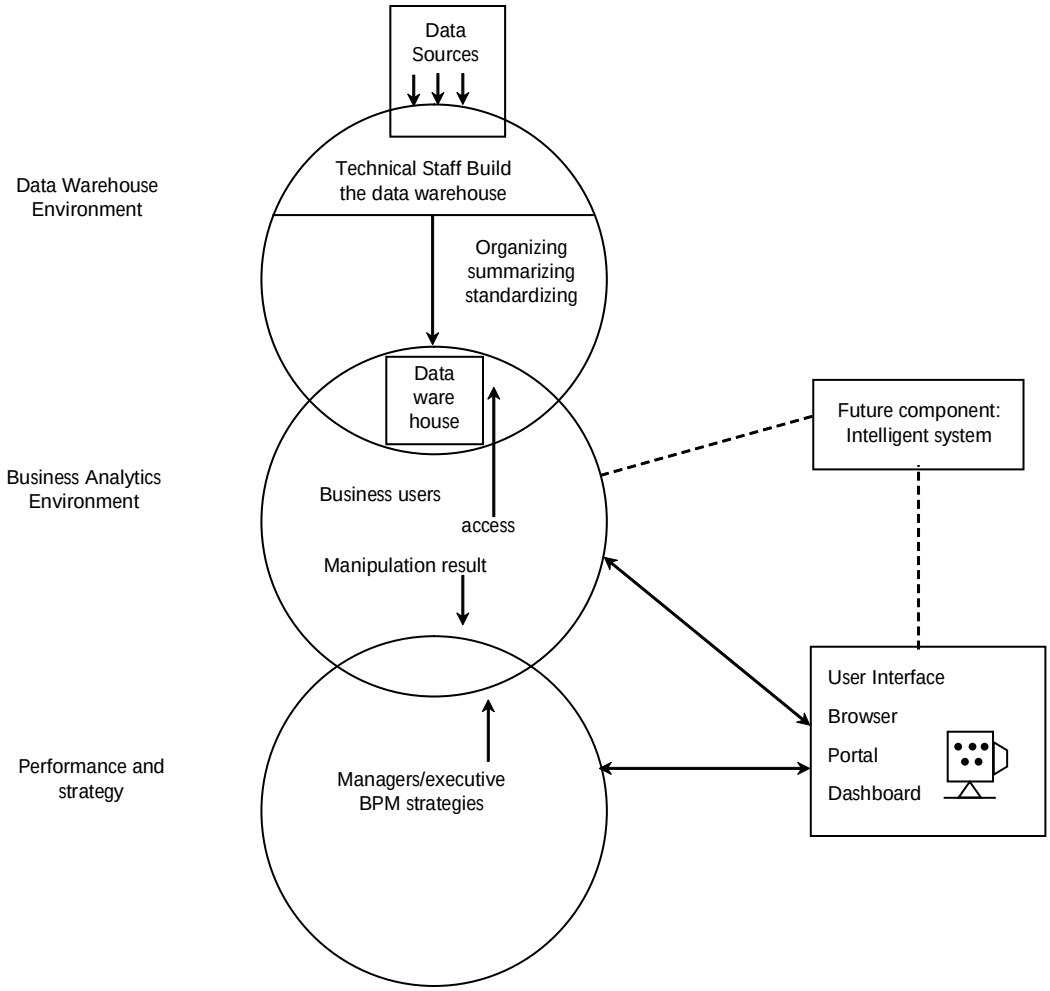


Figure: High-level Architecture of Business Intelligence (BI)

1. Data Warehouse

The data warehousing is mainly deals with the technical staff. Data warehouse and its elements are known to be as the foundation of medium-to-large business intelligence system. Generally, it consists of organized & outlined historical data that helps the end users to access the data and information easily. Now a days, some of the data warehouses include historical data as well as the current data, there by offering the real-time decision support.

2. Business Analytics

The data or information in a data warehouse can be accessed by the end users through various tools and techniques. These tools and techniques are divided in to two categories. They are,

(i) Data, Text, Web Mining and other Sophisticated Mathematical and Statistical Tools: The main purpose of data mining is to discover the relation that connects different database entities. The process of data mining can be done on textual data and also on web data. Data mining can be done by using various tools and techniques such as predictive analysis-techniques, neural computing, artificial intelligence techniques, statistical and mathematical techniques.

(ii) Reports and Queries: Business analytics consists of various reports and queries such as, static and dynamic reporting, multi-dimensional view, drill down to details, discovery of information and so on.

3. Business Performance Management (BPM)

Business performance management is also called as corporate performance management. BPM includes the evolution and architecture of Business Intelligence.

BPM introduced a new concept called as management and feedback that extends the measuring, monitoring and comparison of sales, cost, profit and profitability. It involves in various processes Such as planning, forecasting and budgeting. The traditional Decision Support Systems (DSS), Executive Information System (EIS), Business Intelligence (BI) helps in the bottom-up extraction of information from data where as BPM provides/offers a top-down application of corporate wide strategy. Moreover, it is basically combines with the balanced scorecard methodology as well as dashboards.

4. User Interface

User interface includes dashboards and other information broadcasting tools such as dash boards, portal and browser. Using these tools, a user can connect to user interface. Dashboards are meaningful groups of corporate/marketing performance measures (key performance indicators), exceptions and trends. They are used to combine the information collected from various business areas and also provides a graph that shows the comparison between the actual performance and the required metrics. Hence, it shows view of the organization's health. In addition to this, user interface includes some other information broadcasting tools such as digital cockpits, corporate portals and visualization tools that ranges from multidimensional cube presentation to virtual reality are known to be the integral parts of business intelligence systems. Since, business intelligence is derived from EIS, business intelligence acquires many visual aids for executives are converted into business intelligence software. However, Geographical Information system (GIS) also plays an incremental role in DSS.

Short Question and Answers

1. Define Business Intelligence.

Ans :

Business Intelligence (BI) is a set of ideas, methodologies, processes, architectures, and technologies that change raw data into significant and useful data for business purpose. Business Intelligence can handle large amounts of data to help identify and evolve new opportunities for the business. Making use of these new opportunities and applying a productive scheme on it can provide a comparable market benefit and long-term stability

Business Intelligence (BI) technologies provide chronicled, present and predictive view of business operations. Common functions of enterprise Intelligence technologies are reporting, online analytical processing, analytics, data excavation, process excavation, business performance management, benchmarking, text mining, predictive analytics and prescriptive analytics.

Meaning of Business Intelligence

Business Intelligence is a set of processes, architectures, and technologies that convert raw data into meaningful information that drives profitable business actions. It is a suite of software and services to transform data into actionable intelligence and knowledge.

Business Intelligence has a direct impact on organization's strategic, tactical and operational business decisions. Business Intelligence supports fact-based decision making using historical data rather than assumptions and gut feeling.

Business Intelligence tools perform data analysis and create reports, summaries, dashboards, maps, graphs, and charts to provide users with detailed intelligence about the nature of the business.

2. Characteristics of Business Intelligence

Ans :

i) It is created by procuring data and information for use in decision-making.

ii) It is a combination of skills, processes, technologies, applications and practices.

iii) It contains background data along with the reporting tools.

iv) It is a combination of a set of concepts and methods strengthened by fact-based support systems.

v) It is an extension of Executive Support System or Executive Information System.

vi) It collects, integrates, stores, analyzes, and provides access to business information

vii) It is an environment in which business users get reliable, secure, consistent, comprehensible, easily manipulated and timely information.

viii) It provides business insights that lead to better, faster, more relevant decisions.

3. Various stages of business intelligence.

Ans :

The various stages of Business Intelligence are as follows :

i) Data Sourcing

Defining the data to be loaded into the system. Usually BI applications gathers data from a data warehouse (Data marts, OLTP or OLAP).

ii) ETL (Extract Transform Load)

Extracting the source data and transforming per business rules and loading into the Data Warehouses.

iii) Data Warehousing

Storing transformed data into various Data warehouses types and making it available for business analysis.

iv) Data Analysis

Applying various techniques like data mining, text mining, Process mining to identify trends and patterns in business operations.

v) Decision Making

Based on the reports, dashboards and alerts from previous stage, making valuable business decisions and bench marking future growth.

4. Real-time Business Intelligence.

Ans :

Business users (Application managers, role managers and risk managers etc.) are progressively demanding access to the real-time data (or unstructured data) that includes the contents of data warehouse. For instance, the buses in foreign countries like Texas, Houston are designed with several facilities like data gathering devices, traffic controllers that are capable of accessing and changing the traffic signal/lights interval. Thus, making the system more efficient and reliable. An organization is said to be successful (or survive) only when it deals with the real-time data updates. For instance, the cities nearer to the sea and with low elevation distance make use of the real-time data collection and analysis. By using this equipment the officials will get to know the present situations of the cities, like the working of the flood-controlling pumps which takes an active participation in saving the people when the floods occurs. Interestingly, it is also used in several parts of the world like Netherlands and New Orleans.

5. Value Chain.

Ans :

In competitive Marketplace, it is vital for every business enterprise whether small or big to cope with the pace of the market growth. Intelligence is the ability to learn and understand new situations by updating yourself with current information. This is why organizations are becoming more dependent on data or information for improving development of product and services that can outsell their competitors. So, they are primarily relying on analytical database solutions like data warehouses. However, in the entire process these information and later applying them in action describes a chain of methods called value chain. Value chain is the process that converts data into information and then applies that knowledge in taking productive business decision, and the outcome is the ultimate value.

6. Architecture of Business Intelligence.

Ans :

The architecture of Business Intelligence (BI) system consists of four major components. They are,

1. Data Warehouse
2. Business Analytics
3. Business Performance Management (BPM)
4. User Interface.

The following figure describes the high-level architecture of business intelligence (BI) and also the relationship among the four major parts of business intelligence.

7. Data Warehouse

Ans :

The data warehousing is mainly deals with the technical staff. Data warehouse and its elements are known to be as the foundation of medium-to-large business intelligence system. Generally, it consists of organized & outlined historical data that helps the end users to access the data and information easily. Now a days, some of the data warehouses include historical data as well as the current data, there by offering the real-time decision support.

8. Business Analytics

Ans :

The data or information in a data warehouse can be accessed by the end users through various tools and techniques. These tools and techniques are divided in to two categories. They are,

(i) Data, Text, Web Mining and other Sophisticated Mathematical and Statistical Tools: The main purpose of data mining is to discover the relation that connects different database entities. The process of data mining can be done on textual data and also on web data. Data mining can be done by using various tools and techniques such as predictive analysis-techniques, neural computing, artificial intelligence techniques, statistical and mathematical techniques.

(ii) Reports and Queries: Business analytics consists of various reports and queries such as, static and dynamic reporting, multi-dimensional view, drill down to details, discovery of information and so on.

9. Business Performance Management (BPM)

Business performance management is also called as corporate performance management. BPM includes the evolution and architecture of Business Intelligence.

BPM introduced a new concept called as management and feedback that extends the measuring, monitoring and comparison of sales, cost, profit and profitability. It involves in various processes Such as planning, forecasting and budgeting. The traditional Decision Support Systems (DSS), Executive Information System (EIS), Business Intelligence (BI) helps in the bottom-up extraction of information from data where as BPM provides/offers a top-down application of corporate wide strategy. Moreover, it is basically combines with the balanced scorecard methodology as well as dashboards.

10. User Interface

User interface includes dashboards and other information broadcasting tools such as dash boards, portal and browser. Using these tools, a user can connect to user interface. Dashboards are meaningful groups of corporate/marketing performance measures (key performance indicators), exceptions and trends. They are used to combine the information collected from various business areas and also provides a graph that shows the comparison between the actual performance and the required metrics. Hence, it shows view of the organization's health. In addition to this, user interface includes some other information broadcasting tools such as digital cockpits, corporate portals and visualization tools that ranges from multidimensional cube presentation to virtual reality are known to be the integral parts of business intelligence systems. Since, business intelligence is derived from EIS, business intelligence acquires many visual aids for executives are converted into business intelligence software. However, Geographical Information system (GIS) also plays an incremental role in DSS.

Choose the Correct Answers

1. _____ is a market for long-term funds (or) securities with a maturity period of above one year. [a]

   (a) Capital market (b) Secondary market

   (c) Money market (d) Primary market

2. _____ is a financial intermediary that pools the savings of investors and invests into a diversified portfolio of securities. [b]

   (a) Commercial Bank (b) Mutual funds

   (c) Insurance Companies (d) None

3. Commercial banks performs four major functions. They are Traditional functions, Incidental functions, Core functions and _____. [a]

   (a) Ancillary functions (b) Letter of credit

   (c) Promotional functions (d) All

4. To control the liquidity risk through fundamental approach two methods are used. They are: Asset management _____. [c]

   (a) Technical management (b) Deposits

   (c) Liability management (d) None

5. Simulation method is one of the technique used to measure _____. [d]

   (a) Liquidity risk (b) Foreign exchange risk

   (c) Credit risk (d) ALM Interest Rate Risk

6. _____ affects the incomes/expenses, value of assets/liabilities, and the market value. [c]

   (a) Reinvestment risk (b) Basis risk

   (c) Real interest rate risk (d) Call/Put risk

7. In der micro-level approach, for credit-risk management three major principles are followed. They axe, Pricing, monitoring and _____. [a]

   (a) Evaluation (b) Operational Efficiency

   (c) Netting (d) None

8. _____ is the default risk wherein, any one party involved in the contract with other party facts to meet his financial obligations. [d]

   (a) Liquidity risk (b) Contingency risk

   (c) Volatility risk (d) Credit risk

9. When a mismatch exists between the benchmark rates, _____ take place/occurs. [c]

(a) Call/Put risk (b) Credit risk

(c) Basis risk (d) Prepayment risk

10. 'Asset Quality' is one of the letter of the _____ abbreviation. [b]

(a) Key Performance Indications (b) CAMELS

(c) Asset Liability Management (d) None

Fill in the Blanks

1. The "banker-customer" relationship was traditionally termed as _____ relationship.
2. _____ is referred as an entity (or) an individual carrying-out transactions with the banks.
3. The market which deals with the monetary short-term debt instruments is termed as _____.
4. As per the Banking regulations Act, 1949 banks are classified under three sections. They are public sector, private sector and _____.
5. _____ plays a major role in the development of an economy.
6. The turn CAMELS stands for _____.
7. Non-deposit funding sources are also termed as _____.
8. 'Offering trust services' is one of the _____ services offered by banks.
9. Selling retirement plans is one of the modern services offered by _____.
10. The various sources of funds for banks are equity, reserves and _____.

ANSWER

1. Debtor and Creditor Relationship
2. Customer
3. Money market
4. Foreign sector
5. Commercial Banks
6. Capital Adequacy, [CAMELS]
   Asset Quality
   Management Quality
   Earnings
   Liquidity
   Sensitivity to market risk
7. Wholesale funding sources
8. Traditional services
9. Banks
10. Borrowings

UNIT II

Data Warehousing and Data Mining

a) Date Ware Housing (DWH): Definition, Characteristic, types, Date ware housing frame work, DwH 3 tier architecture, Alternative Architectures, Data ware housing Integration, Data ware housing- Development Approaches, Real time Data ware housing.

b) Data Mining : Definition, Characteristic, Benefits, Date Mining Functions, Data Mining Applications, Data Mining techniques and tools. Text Mining, Web Mining.

2.1 DATA WAREHOUSING (DWH)

2.1.1 Definition of Data Warehousing

Q1. Define Data Warehousing.

Ans : **(June-18)**

A data warehouse is a subject-oriented, integrated, time-variant and non-volatile collection of data in support of management's decision making process.

i) Subject-Oriented

A data warehouse can be used to analyze a particular subject area. For example, "sales" can be a particular subject.

ii) Integrated

A data warehouse integrates data from multiple data sources. For example, source A and source B may have different ways of identifying a product, but in a data warehouse, there will be only a single way of identifying a product.

iii) Time-Variant

Historical data is kept in a data warehouse. For example, one can retrieve data from 3 months, 6 months, 12 months, or even older data from a data warehouse. This contrasts with a transactions system, where often only the most recent data is kept.

For example, a transaction system may hold the most recent address of a customer, where a data warehouse can hold all addresses associated with a customer.

iv) Non-volatile

Once data is in the data warehouse, it will not change. So, historical data in a data warehouse should never be altered.

2.1.2 Characteristics of Data Ware Housing

Q2. What are the characteristics of data ware housing?

Ans :

Characteristics of Data Warehousing

A common way of introducing data ware-housing is to refer to its fundamental characteristics.

1. Subject Oriented

Data are organized by detailed subject, such as sales, products, or customers, containing only information relevant for decision support. Subject orientation enables users to determine not only how their business is performing, but why. A data warehouse differs from an operational database in that most operational databases have a product orientation and are tuned to handle transactions that update the database. Subject orientation provides a more comprehensive view of the organization.

2. Integrated

Integration is closely related to subject orientation. Data warehouses must place data from different sources into a consistent format. To do so, they must deal with naming conflicts and discrepancies among units of measure. A data ware-house is presumed to be totally integrated.

3. Time Variant (time series)

A warehouse maintains historical data. The data do not necessarily provide current status (except in real-time systems). They detect trends, deviations, and long-term relationships for forecasting and comparisons, leading to decision making. Every data warehouse has a temporal quality. Time is the one important dimension that all data warehouses must support. Data for analysis from multiple sources contains multiple time points (e.g., daily, weekly, monthly views).

4. Nonvolatile

After data are entered into a data warehouse, users cannot change or update the data. Obsolete data are discarded, and changes are recorded as new data.

These characteristics enable data warehouses to be tuned almost exclusively for data access. Some additional characteristics may include the following:

5. Web Based

Data warehouses are typically designed to provide an efficient computing environment for Web-based applications.

6. Relational/Multidimensional

A data warehouse uses either a relational structure or a multidimensional structure. A recent survey on multidimensional structures can be found in Romero and Abello (2009).

7. Client/Server

A data warehouse uses the client/server architecture to provide easy access for end users.

8. Real Time

Newer data warehouses provide real-time, or active, data access and analysis capabilities (Basu, 2003; and Bonde and Kuckuk, 2004).

9. Metadata

A data warehouse contains metadata (data about data) about how the data are organized and how to effectively use them.

2.1.3 Types of Data warehousing

Q3. What are the types of Data warehousing?

Ans : **(June-19)**

Whereas a data warehouse is a repository of data, data warehousing is literally the entire process (Watson, 2002). Data warehousing is a discipline that results in applications that provide decision support capability, allows ready access to business information, and creates business insight. The three main types of data warehouses are data marts, operational data stores (ODSs), and enterprise data warehouses (EDWs). In addition to discussing these three types of warehouses next, we also discuss metadata.

1. Data Marts

Whereas a data warehouse combines databases across an entire enterprise, a data mart is usually smaller and focuses on a particular subject or department.

- A data mart is a sub-set of a data warehouse, typically consisting of a single subject area (e.g., marketing, operations). A data mart can be either dependent or independent.
- A dependent data mart is a subset that is created directly from the data warehouse. It has the advantages of using a consistent data model and providing quality data.
- Dependent data marts support the concept of a single enterprise-wide data model, but the data warehouse must be constructed first. A dependent data mart ensures that the end user is viewing the same version of the data that is accessed by all other data warehouse users.
- The high cost of data warehouses limits their use to large companies.
- As an alternative, many firms use a lower-cost, scaled-down version of a data warehouse referred to as an independent data mart. An independent data mart is a small warehouse designed for a strategic business unit or a department, but its source is not an EDW.

2. Operational Data Stores (ODS)

An operational data store (ODS) provides a fairly recent form of customer information file.

- This type of database is often used as an interim staging area for a data warehouse.
- Unlike the static contents of a data warehouse, the contents of an ODS are updated throughout the course of business operations.
- An ODS is used for short-term decisions involving mission-critical applications rather than for the medium and long-term decisions associated with an EDW.
- An ODS is similar to short-term memory in that it stores only very recent information. In comparison, a data warehouse is like long-term memory because it stores permanent information.
- An ODS consolidates data from multiple source systems and provides a near-real-time, integrated view of volatile, current data.
- The ETL processes for an ODS are identical to those for a data warehouse. Finally, open marts are created when operational data need to be analyzed multidimensional. The data for an open mart come from an ODS.

3. Enterprise Data Warehouses (EDWs)

An enterprise data warehouse (EDW) is a large-scale data warehouse that is used across the enterprise for decision support.

- It is the type of data warehouse that Directive' developed, as described in the opening vignette.
- The large-scale nature provides integration of data from many sources into a standard format for effective BI and decision support applications.
- EDWs are used to provide data for many types of DSS, including customer relationship management (CRM), supply chain management (SCM), business performance management (BPM), business activity monitoring (BAM), product life cycle management (PLM), revenue management, and sometimes even knowledge management systems (KMS).

Q4. Explain briefly about Data Warehouse model.

Ans :

There are three data warehouse models.

1. Enterprise Warehouse

An enterprise warehouse collects all of the information about subjects spanning the entire organization.

It provides corporate-wide data integration, usually from one or more operational systems or external information providers, and is cross-functional in scope.

It typically contains detailed data aswell as summarized data, and can range in size from a few gigabytes to hundreds of gigabytes, terabytes, or beyond.

An enterprise data warehouse may be implemented on traditional mainframes, computer super servers, or parallel architecture platforms. It requires extensive business modeling and may take years to design and build.

2. Data Mart

A data mart contains a subset of corporate-wide data that is of value to specific group of users. The scope is confined to specific selected subjects. For example, a marketing data mart may confine its subjects to customer, item, and sales. The data contained in data marts tend to be summarized.

Data marts are usually implemented on low-cost departmental servers that are Unix/Linux - or Windows-based. The implementation cycle of a data mart is more likely to be measured in weeks rather than months or years. However, it may involve complex integration in the long run if its design and planning were not enterprise-wide.

3. Virtual Warehouse

A virtual warehouse is a set of views over operational databases. For efficient query processing, only some of the possible summary views may be materialized.

A virtual warehouse is easy to build but requires excess capacity on operational database servers.

Q5. What is metadata ?

(or)

Define metadata

Ans :

Metadata are data about data. When used in a data warehouse, metadata are the data that define warehouse objects. Metadata are created for the data names and definitions of the given warehouse. Additional metadata are created and captured for time stamping any extracted data, the source of the extracted data, and missing fields that have been added by data cleaning or integration processes. A metadata repository should contain the following:

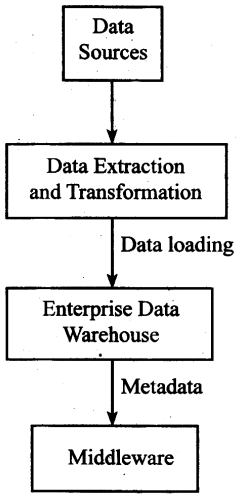
- A description of the structure of the data warehouse, which includes the warehouse schema, view, dimensions, hierarchies, and derived data definitions, as well as data mart locations and contents.
- Operational metadata, which include data lineage (history of migrated data and the sequence of transformations applied to it), currency of data (active, archived, or purged), and monitoring information (warehouse usage statistics, error reports, and audit trails).
- The algorithms used for summarization, which include measure and dimension definition algorithms, data on granularity, partitions, subject areas, aggregation, summarization, and predefined queries and reports.
- The mapping from the operational environment to the data warehouse, which includes source databases and their contents, gateway descriptions, data partitions, data extraction, cleaning, transformation rules and defaults, data refresh and purging rules, and security (user authorization and access control).
- Data related to system performance, which include indices and profiles that improve data access and retrieval performance, in addition to rules for the timing and scheduling of refresh, update, and replication cycles.
- Business metadata, which include business terms and definitions, data owner ship information, and charging policies.

2.1.4 Data Warehousing Frame Work

Q6. Discuss about data warehousing frame work.

Ans : **(June-18)**

Data warehousing is important for most or all the organizations to manage large amount of data for assisting decision support system. Typical framework of data warehousing process is shown in the below figure.



The following are the components involved in data warehousing process framework.

1. Data Sources

The data to be stored in a data warehouse can be extracted from many data sources such as ERP system, OLTP, web/internet, legacy system or any other source.

2. Data Extraction and Transformation

The data is extracted from the data sources and it is converted into appropriate format using ETL (Extraction, Transformation and Loading) software.

3. Data Loading and Enterprise Data Warehouse

The converted data is loaded into a comprehensive database called Enterprise Data Warehouse (EDW) which provides detailed information of data to support decision making.

4. Metadata

Metadata is also included for users and data administrators to efficiently access the data stored in the warehouse. This metadata provides rules followed in arranging and storing the database so that the required data can be easily retrieved even with web tools.

5. Middleware

Middleware offers tools which can be used for accessing the data. These tools can vary for different types of users. Some of the tools are data/text mining, OLAP, visualization tools. System analysts can access the data using Structured Query Language (SQL).

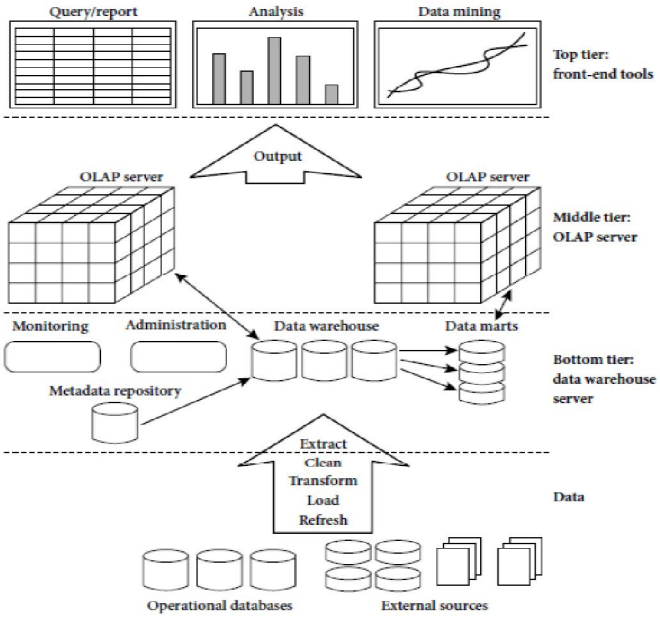
This process is followed in order to manage the complexities involved in maintaining and using massive amount : data present in a warehouse. It continuously provides reliability and availability even if the number of users increases anonymously.

2.1.5 3-Tier Architectures of Data Ware Housing

Q7. Discuss about 3 Tier Architectures of Data Ware Housing.

Ans : **(Oct.-20)**

A three tier data warehouse architecture:



Tier-1

The bottom tier is a warehouse database server that is almost always a relational database system. Back-end tools and utilities are used to feed data into the bottom tier from operational databases or other external sources (such as customer profile information provided by external consultants). These tools and utilities perform data extraction, cleaning, and transformation (e.g., to merge similar data from different sources into a unified format), as well as load and refresh functions to update the data warehouse.

The data are extracted using application program interfaces known as gateways. A gateway is supported by the underlying DBMS and allows client programs to generate SQL code to be executed at a server. Examples of gateways include ODBC (Open Database Connection) and OLEDB (Open Linking and Embedding for Databases) by Microsoft and JDBC (Java Database Connection). This tier also contains a metadata repository, which stores information about the data warehouse and its contents.

Tier-2

The middle tier is an OLAP server that is typically implemented using either a relational OLAP (ROLAP) model or a multidimensional OLAP.

OLAP model is an extended relational DBMS that maps operations on multidimensional data to standard relational operations.

A multidimensional OLAP (MOLAP) model, that is, a special-purpose server that directly implements multidimensional data and operations.

Tier-3

The top tier is a front-end client layer, which contains query and reporting tools, analysis tools, and/or data mining tools (e.g., trend analysis, prediction, and so on).

Q8. Explain the Issues in Selecting among Data Ware Housing Architecture.

Ans :

Several issues must be considered when deciding which architecture to use. Among them are the following :

- **Which database management system (DBMS) should be used**

  Most data warehouses are built using relational database management systems (RDBMS). Oracle SQL Server and DB2 are the ones most commonly used. Each of these products supports both client/server and Web-based architectures.

- **Parallel processing and/or partitioning be used**

  Parallel processing enables multiple CPUs to process data warehouse query requests simultaneously and provides scalability. Data warehouse designers need to decide whether the database tables will be partitioned (i.e., split into smaller tables) for access efficiency and what the criteria will be. This is an important consideration that is necessitated by the large amounts of data contained in a typical data warehouse. A recent survey on parallel and distributed data warehouses can be found in Furtado (2009). Teradata has successfully adopted and often commented on its novel implementation of this approach.

- **Data migration tools be used to load the data warehouse**

  Moving data from an existing system into a data warehouse is a tedious and laborious task. Depending on the diversity and the location of the data assets, migration may be a relatively simple procedure or (in contrary) a months-long project. The results of a thorough assessment of the existing data assets should be used to determine whether to use migration tools, and if so, what capabilities to seek in those commercial tools.

- **Tools will be used to support data retrieval and analysis**

  Often it is necessary to use specialized tools to periodically locate, access, analyze, extract, transform, and load necessary data into a data warehouse. A decision has to be made on (i) developing the migration tools in-house, (ii) purchasing them from a third-party provider, or (iii) using the ones provided with the data warehouse system. Overly complex, real-time migrations warrant specialized third-party ETL tools.

2.1.6 Alternative of Data Ware Housing

Q9. What are the alternative of Data Ware Housing?

Ans :

The alternative architectures of data warehousing are as follows,

1. Independent Data Marts Architecture

In this architecture individual data marts are created which works without depending on each other. These marts focus on individual goals instead of fulfilling the overall organizational goals. It is the simplest and cost effective architecture among all the other architectures. The drawback of this architecture is that there exist data inconsistencies leading to complexities in analyzing data present in different data marts.

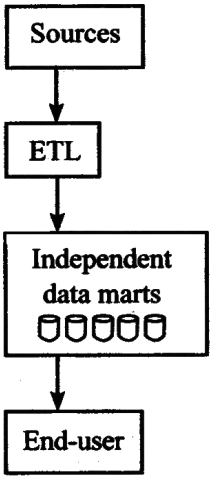


Fig.: Architecture of Independent Data Marts

2. Data Mart Bus Architecture

This architecture is an extension of independent data marts architecture. In this architecture, the data marts are connected together through a middleware. Using such an architecture data consistency can be achieved. It can process complex queries but does not fulfil the performance requirements.

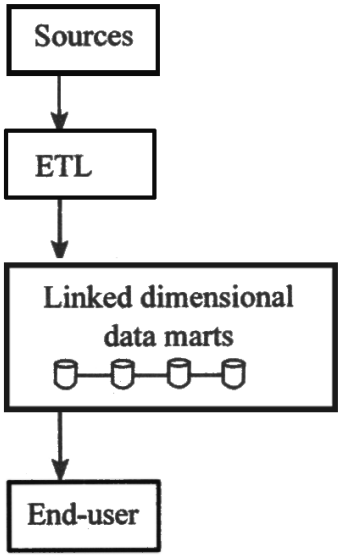


Fig.: Architecture of Data Mart Bus

3. Hub-and-Spoke Architecture

The main objective of this architecture is to achieve scalability and maintainability within the infrastructure. To achieve this, different subject areas are considered one after the other in an iterative way. It carries a centralized warehouse and multiple data marts which are interdependent. The advantage of this architecture is its simplicity and customization support for interfaces and reports. However, the drawbacks of this architecture includes redundancy and latency of data.

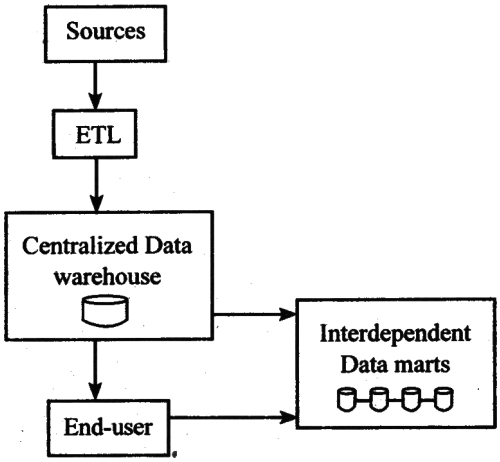


Fig.: Architecture of Hub-and-spoke Data Warehousing

4. Centralized Data Warehouse Architecture

In this architecture, a single data warehouse is used without any data marts. Such an architecture eliminates the restrictions on data which are imposed by data marts. In this way, the burden on technical team to classify or modify data according to different data marts gets eliminated.

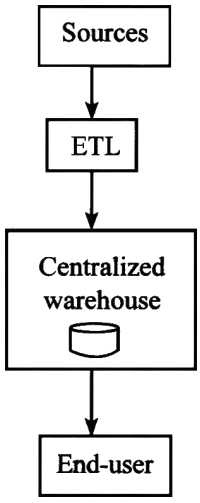


Figure: Architecture of Centralized Data Warehousing

5. Federated Architecture

In this architecture, all the ways in which analytical resources can be integrated to fulfill the evolving business requirements are determined. This architecture integrates systems which are of different kinds and cannot be compared. It is typically employed by middleware tool vendors that support join and distributed queries. These tools are based on XML(Extensible Markup Language) which provides the overall view of distributed data sources. These architectures are used to avoid data warehouse replacement.

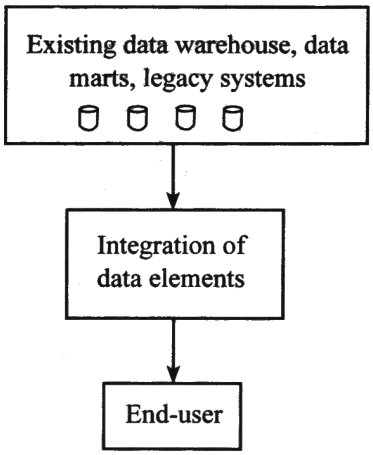


Figure: Federated Architecture

Q10. Explain the design process of Data Ware Housing?

Ans :

A data warehouse can be built using a top-down approach, a bottom-up approach, or a combination of both.

The top-down approach starts with the overall design and planning. It is useful in cases where the technology is mature and well known, and where the business problems that must be solved are clear and well understood.

The bottom-up approach starts with experiments and prototypes. This is useful in the early stage of business modeling and technology development. It allows an organization to move forward at considerably less expense and to evaluate the benefits of the technology before making significant commitments.

In the combined approach, an organization can exploit the planned and strategic nature of the top-down approach while retaining the rapid implementation and opportunistic application of the bottom-up approach.

The warehouse design process consists of the following steps:

Choose a business process to model, for example, orders, invoices, shipments, inventory, account administration, sales, or the general ledger. If the business process is organizational and involves multiple complex object collections, a data warehouse model should be followed. However, if the process is departmental and focuses on the analysis of one kind of business process, a data mart model should be chosen.

Choose the grain of the business process. The grain is the fundamental, atomic level of data to be represented in the fact table for this process, for example, individual transactions, individual daily snapshots, and so on.

Choose the dimensions that will apply to each fact table record. Typical dimensions are time, item, customer, supplier, warehouse, transaction type, and status.

Choose the measures that will populate each fact table record. Typical measures are numeric additive quantities like dollars sold and units sold.

2.1.7 Data Ware Housing Integration

Q11. Define Data Integration.

Ans :

Data integration refers to the process of combining data from different data sources such as databases, flat files etc. It involves three processes which enable users to access data and also to array of ETL, analysis tools and data warehousing environment to access the data. These three processes are as follows,

1. Data Access

This process allows data to be accessed and extracted from any type of sources.

2. Data Federation

This process allows business views to be integrated with different data stores.

3. Change Capture

This process captures the modifications done with respect to identification, capture and delivery of data sources. Integration of data and metadata can be performed using the following technologies.

(i) Enterprise Application Integration (EAI)

This technology helps in storing the data in the warehouse from various source systems. Integration is performed with respect to the functionality of application. Here, the functionality is distributed among multiple system to achieve reuse and flexibility. EAI can be used for different purposes such as forwarding decisions to OLTP, performing data acquisition on near-realtime data warehouse.

(ii) Service-Oriented Architecture (SOA)

This technology helps in integrating business processes by dividing large application into smaller modules or services. It is used to maintain proper data flow in data warehouse which allows creation of a new function for each data element.

(iii) Enterprise Information Integration (EII): This technology uses a set of tools to ensure real-time data integration over a set of sources. Source can be a relational data base, multi dimensional database or web service. The tools offered by Eli makes the data available to the users in the form of relational data by making use of views. The most critical component of this approach is XML, which provides various tags to generate knowledge at any stage.

4. Extraction, Transformation and Load (ETL)

ETL is a set of Extraction, Transformation and Loading processes which are performed in the data staging environment. It typically converts the irrelevant data extracted from several operational systems to the relevant/ useful information. This information thus be stored into the data warehouse environment. The warehouse cannot be implemented efficiently if the data is not collected, cleansed and integrated in an appropriate way. Hence, the ETL process must be efficiently designed and implemented for optimized development of data warehouse.

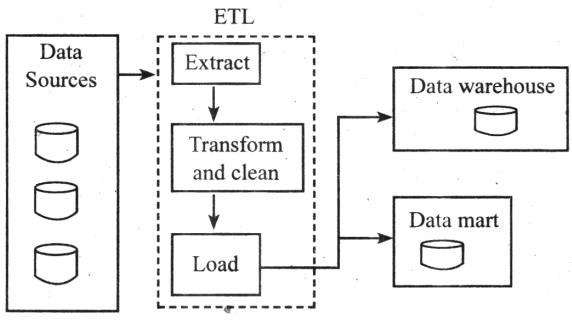


Figure: ETL Process

(a) Data Extraction

The function of the data extraction is to find and extract the data from various relational and non-relational data sources for future use. Data extraction process includes,

(i) Determining the operational systems from which the data can be collected.

(ii) Capture the appropriate data from data sources to make required data analysis.

(iii) Identify the extraction frequency of data i.e., daily, weekly or monthly to keep the data warehouse up to date.

(b) Data Transformation

Data transformation is a process in which the data collected from many dissimilar source systems are transformed in accordance to a standardized format. The main purpose of performing a data transformation is to enhance the quality of the data before providing it as input in the data warehouse. Before transformation process, reliability, consistency and validity of the data are checked.

Data transformation is carried out by the following steps,

- Data elements should be identified that are to be moved to data warehouse.
- Calculated and derived data values should be identified.
- Next step is to perform data cleaning.
- Define the rules for transformation, data mapping criteria and metadata updation.
- In this step, data should be restructured.
- Data fields from various sources should be combined into one entity.
- For the captured data, data values, data types and field lengths should be standardized.
- Data integrity should be maintained within the data warehouse.
- Change the data according to the structure and format of data warehouse.
- Analysis of data warehouse can be made easy by simplifying and re-arranging the individual fields.

Some of the transformation processes are aggregation, normalization, union, join, enforcing keys etc.

3. Data Loading

Data loading is a process of loading the extracted and transformed data into the data warehouse repository.

The data loading usually, requires the data warehouse to be kept off-line such that no users can access the warehouse at the time of loading. Thus, a specified time must be assigned for loading the data. This can be done by fragmenting the data loading process into smaller number of portions such that, only few tables can be loaded concurrently.

Once the loading process has been completed, the loaded data need to be tested so as to ensure whether these loads have been correctly done in the data warehouse repository. In addition to this, a plan is designed to test the quality of the loaded records.

Q12. What are the techniques of data Integration ?

Ans :

There are many sophisticated ways the unified view of data can be created today. No more ETL is the only way to achieve the goal and that is a new level of complexity in the field of Data Integration.

There are several organizational levels on which the Data Integration can be performed and let's discuss them briefly.

1. Manual Data Integration

Technically speaking, this is really not a Data Integration. In this approach, a web based user interface or an application is created for users of the system to show them all the relevant information by accessing all the source systems directly. There is no unification of data in reality.

2. Middleware Data Integration

A middleware data integration solution is essentially a layer between two disparate systems allowing them to communicate. Middleware integration can act like a glue that holds together multiple legacy applications, making seamless connectivity possible without requiring the two applications to communi-cate directly.

3. Data Virtualization Integration Approach

Data Virtualization allows us to leave data in the source systems while allowing to create a new set of unified views. This provides a way for users to access the unified view of disparate source system's data across whole enterprise.

A lot of organizations today prefer this approach because of the benefits and technologies that exist today to support this approach. The main benefit of the virtual integration approach is near real time view of data from the source systems. It eliminates a need for separate data store for the consolidated unified data.

However, that doesn't mean it's the best way to do Data Integration although it certainly has a short term benefit. The drawbacks of this approach include limited possibility of data's history availability or data version management and extra load on the source systems involved which may have an adverse effect on the performance of the source systems.

4. Data Warehouse Approach Of Data Integration

This is the most commonly known approach to Data Integration you may already know if you have read Ralf Kimball and/or Bill Inmon.

This approach requires creation of a new Data Warehouse (of Data Marts) which stores a unified version of data extracted from all the source systems involved and manage it independent of the original source systems.

The benefits of this approach include ability to easily manage history of data (or data versioning), ability to combine data from very disparate sources (mainframes, databases, flat files, etc.) and to store them in a central repository of data.

2.1.8 Data Ware House Development Approaches

Q13. What are the development approaches of Data Warehouse?

Ans :

1. Data Warehouse Development

A data warehousing project is a major undertaking for any organization and is more complicated than a simple, mainframe selection and implementation project because it comprises and influences many departments and many input and output interfaces and it can be part of a CRM business strategy. A data warehouse provides several benefits that can be classified as direct and indirect.

A) Direct Benefits

1. End users can perform extensive analysis in numerous ways.
2. A consolidated view of corporate data (i.e., a single version of the truth) is possible.
3. Better and more-timely information is possible. A data warehouse permits information processing to be relieved from costly operational systems onto low-cost servers; therefore, many more end-user information requests can be processed more quickly.

4. Enhanced system performance can result. A data warehouse frees production processing because some operational system reporting requirements are moved to DSS.
5. Data access is simplified.

B) Indirect benefits

Indirect benefits result from end users using these direct benefits. On the whole, these benefits enhance business knowledge, present competitive advantage, improve customer service and satisfaction, facilitate decision making, and help in reforming business processes, and therefore, they are the strongest contributions to competitive advantage.

Given the potential benefits that a data warehouse can provide and the substantial investments in time and money that such a project requires, it is critical that an organization structure its data warehouse project to maximize the chances of success. In addition, the organization must, obviously, take costs into consideration. Kelly (2001) described a ROI approach that considers benefits in the categories of keepers (i.e., money saved by improving traditional decision support functions); gatherers (i.e., money saved due to automated collection and dissemination of information); and users (i.e., money saved or gained from decisions made using the data warehouse).

Costs include those related to hardware, software, network band-width, internal development, internal support, training, and external consulting. The net present value (NPV) is calculated over the expected life of the data warehouse. Because the benefits are broken down approximately as 20 percent for keepers, 30 percent for gatherers, and 50 percent for users, Kelly indicated that users should be involved in the development process, a success factor typically mentioned as critical for systems that imply change in an organization.

Data Warehouse Development Approaches

Data warehouse can be developed using either of the following two approaches.

1. **The inmon model :** The edw approach Inmon's approach emphasizes top-down development, employing established database development methodologies and tools, such as entity-relationship diagrams (ERD), and an adjustment of the spiral development approach. The EDW approach does not preclude the creation of data marts. The EDW is the ideal in this approach because it provides a consistent and comprehensive view of the enterprise. Murtaza (1998) presented a framework for developing EDW.

2. The Kimball Model

The Data Mart Approach Kimball's data mart strategy is a "plan big, build small" approach. A data mart is a subject-oriented or department-oriented data warehouse. It is a scaled-down version of a data warehouse that focuses on the requests of a specific department such as marketing or sales. This model applies dimensional data modeling, which starts with tables. Kimball advocated a development methodology that entails a bottom-up approach, which in the case of data warehouses means building one data mart at a time.

There is no one-size-fits-all strategy to data warehousing. An enterprise's data warehousing strategy can evolve from a simple data mart to a complex data warehouse in response to user demands, the enterprise's business requirements, and the enterprise's maturity in managing its data resources. For many enterprises, a data mart is frequently a convenient first step to acquiring experience in constructing and managing a data warehouse while presenting business users with the benefits of better access to their data; in addition, a data mart commonly indicates the business value of data warehousing. Ultimately, obtaining an EDW is ideal. However, the development of individual data marts can often.

Q14. Compare and contrast Inmon and Kimball's Approaches.

Ans :

S.No.	Inmon's Approach	S.No.	Kimball's Approach
1.	It follows a top-down approach.	1.	It follows a bottom-up approach.
2.	Data warehouse architecture considers the overall organization.	2.	Data marts are used which consider various departments of the organization.
3.	It follows a complex method.	3.	It follows a simple method.
4.	The origin of this approach is spiral method.	4.	The origin of this approach is RDBMS.
5.	The physical design of this approach is thorough.	5.	The physical design of this approach is light.
6.	This approach include tools like ERD, DFD etc.	6.	This approach include dimensional modeling tools.
7.	It offers very low accessibility to the end-users.	7.	It offers higher accessibility to the end-users.
8.	The objective of this approach is to offer a technical solution which follows database methods.	8.	The objective of this approach is to simplify the data access so that end -users can directly access the database.

2.1.9 Real Time Data Ware Housing

Q15. Define about Real Time Data Ware Housing. Explain the need for real time warehousing.

Ans :

Real-time data warehousing (RDW), also known as active data warehousing (ADW), is the process of loading and providing data via the data warehouse as they become available.

- It evolved from the EDW concept. The active traits of an RDW/ADW supplement and expand traditional data warehouse functions into the realm of tactical decision making.
- People throughout the organization who interact directly with customers and suppliers will be empowered with information-based decision making at their fingertips.
- Even further leverage results when an ADW provides information directly to customers and suppliers.
- The reach and impact of information access for decision making can positively affect almost all aspects of customer service, SCM, logistics, and beyond. E-business has become a major catalyst in the demand for active data warehousing.

Need for Real Time Warehousing

According to Basu (2003), the most distinctive difference between a traditional data warehouse and an RDW is the shift in the data acquisition paradigm. Some of the business cases and enterprise requirements that led to the need for data in real time include the following:

- A business often cannot afford to wait a whole day for its operational data to load into the data warehouse for analysis.
- Until now, data warehouses have captured snapshots of an organization's fixed states instead of incremental real-time data showing every change and analogous patterns over time.

- With a traditional hub-and-spoke architecture, keeping the metadata in sync is difficult. It is also costly to develop, maintain, and secure many systems as opposed to one huge data warehouse so that data are centralized for BI/BA tools.
- In cases of huge nightly batch loads, the necessary ETL setup and processing power for large nightly data warehouse loading might be very high, and the processes might take too long. An EAI with real-time data collection can reduce or eliminate the nightly batch processes.

Despite the benefits of an RDW, developing one can create its own set of issues. These problems relate to architecture, data modeling, physical database design, storage and scalability, and maintainability. In addition, depending on exactly when data are accessed, even down to the microsecond, different versions of the truth may be extracted and created, which can confuse team members.

Real-time solutions present a remarkable set of challenges to BI activities. Although it is not ideal for all solutions, real-time data warehousing may be successful if the organization develops a sound methodology to handle project risks, incorporates proper planning, and focuses on quality assurance activities. Understanding the common challenges and applying best practices can reduce the extent of the problems that are often a part of implementing complex data warehousing systems that incorporate BI/BA methods.

Q16. Compare traditional data warehousing and active data warehousing.

Ans :

S.No.	Traditional Data Warehousing	S.No.	Active Data Warehousing
1.	Traditional warehousing can only support strategic decisions.	1.	Active warehousing can support both strategic and tactical decisions.
2.	Generation of results is difficult.	2.	Generation of results is done with operations.
3.	Data can be updated daily, weekly or monthly.	3.	Data needs to be updated instantly (within minutes).
4.	Few users can generate queries to the system concurrently.	4.	Thousands of users can generate queries to the system concurrently.
5.	It imposes several restrictions on reporting for verifying patterns and processes.	5.	It provides flexibility with ad hoc reporting and machine based modelling for identifying relationships.
6.	Users involved in this environment include knowledge workers, internal users and power users.	6.	Users involved in this environment include external users, operational staff and call centre.

2.2 DATA MINING

2.2.1 Definition of Data Mining

Q17. Define Datamining.

Ans :

Data mining refers to extracting or mining knowledge from large amounts of data. The term is actually a misnomer. Thus, data miningshould have been more appropriately named as knowledge mining which emphasis on mining from large amounts of data.

It is the computational process of discovering patterns in large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems. The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use.

- The key properties of data mining are
- Automatic discovery of patterns
- Prediction of likely outcomes
- Creation of actionable information
- Focus on large datasets and databases

The Scope of Data Mining

Data mining derives its name from the similarities between searching for valuable business information in a large database.

For example, finding linked products in gigabytes of store scanner data — and mining a mountain for a vein of valuable ore. Both processes require either sifting through an immense amount of material, or intelligently probing it to find exactly where the value resides. Given databases of sufficient size and quality, data mining technology can generate new business opportunities by providing these capabilities:

i) Automated prediction of trends and behaviors

Data mining automates the process of finding predictive information in large databases. Questions that traditionally required extensive hands-on analysis can now be answered directly from the data - quickly. A typical example of a predictive problem is targeted marketing. Data mining uses data on past promotional mailings to identify the targets most likely to maximize return on investment in future mailings. Other predictive problems include forecasting bankruptcy and other forms of default, and identifying segments of a population likely to respond similarly to given events.

ii) Automated discovery of previously unknown patterns

Data mining tools sweep through databases and identify previously hidden patterns in one step. An example of pattern discovery is the analysis of retail sales data to identify seemingly unrelated products that are often purchased together.

Other pattern discovery problems include detecting fraudulent credit card transactions and identifying anomalous data that could represent data entry keying errors.

2.2.2 Characteristics of Data Mining

Q18. What are the Characteristics of Data Mining?

Ans : **(June-19)**

Data mining service is an easy form of information gathering methodology where in which all the relevant information goes through some sort of identification process.

And eventually at the end of this process one can determine all the characteristic of the data mining process.

1. Increased Quantities of Data

In earlier days, data mining system can be determined with the help of their clients and customers, but in today's date, one can acquire any number of information without the help of those clients.

Moreover, after this kind of revolution in the mining system, it also added one more problem and that is large quantities of work.

With the help of these information technology, one can acquire a large number of information without any extra burden or trouble.

2. Provides Incomplete Data

Most of the people provide incomplete information about themselves in some of the survey conducted with the help of data mining systems.

Therefore, people ignore the value of their information and that is why they provide incomplete information about themselves in those surveys conducted for the benefit of the mining systems.

Moreover, these mining systems changed the perspective of people and because of that, people fear the exchange of their personal information.

3. Complicated Data Structure

Data mining is a form where in which all the information is gathered and incorporated with the help of information collection techniques. These information collecting techniques are more of manual and rest are technological.

Therefore, most of the understanding and determination of these mining can be a bit complicated than other structure of information technology.

Q19. Explain the steps involved in data mining.

Ans :

Steps

1. Problem Definition

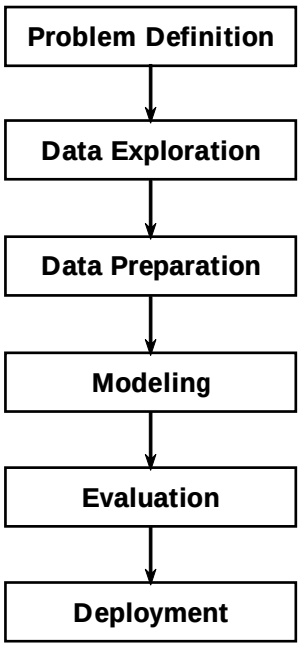
A data mining project starts with the understanding of the business problem. Data mining experts, business experts, and domain experts work closely together to define the project objectives and the requirements from a business perspective. The project objective is then translated into a data mining problem definition.

In the problem definition phase, data mining tools are not yet required.

2. Data Exploration

Domain experts understand the meaning of the metadata. They collect, describe, and explore the data. They also identify quality problems of the data. A frequent exchange with the data mining experts and the business experts from the problem definition phase is vital.

In the data exploration phase, traditional data analysis tools, for example, statistics, are used to explore the data.



3. Data Preparation

Domain experts build the data model for the modeling process. They collect, cleanse, and format the data because some of the mining functions accept data only in a certain format. They also create new derived attributes, for example, an average value.

In the data preparation phase, data is tweaked multiple times in no prescribed order. Preparing the data for the modeling tool by selecting tables, records, and attributes, are typical tasks in this phase. The meaning of the data is not changed.

4. Modeling

Data mining experts select and apply various mining functions because you can use different mining functions for the same type of data mining problem. Some of the mining functions require specific data types. The data mining experts must assess each model.

In the modeling phase, a frequent exchange with the domain experts from the data preparation phase is required.

The modeling phase and the evaluation phase are coupled. They can be repeated several times to change parameters until optimal values are achieved. When the final modeling phase is completed, a model of high quality has been built.

5. Evaluation

Data mining experts evaluate the model. If the model does not satisfy their expectations, they go back to the modeling phase and rebuild the model by changing its parameters until optimal values are achieved. When they are finally satisfied with the model, they can extract business explanations and evaluate the following questions:

- Does the model achieve the business objective?
- Have all business issues been considered?

At the end of the evaluation phase, the data mining experts decide how to use the data mining results.

6. Deployment

Data mining experts use the mining results by exporting the results into database tables or into other applications, for example, spreadsheets.

The Intelligent Miner products assist you to follow this process. You can apply the functions of the Intelligent Miner products independently, iteratively, or in combination.

The following figure shows the phases of the Cross Industry Standard Process for data mining (CRISP DM) process model.

IM Modeling helps you to select the input data, explore the data, transform the data, and mine the data. With IM Visualization you can display the data mining results to analyze and interpret them. With IM Scoring, you can apply the model that you have created with IM Modeling.

2.2.3 Functions of Data Mining

Q20. What are the Functions of Data Mining?

Ans : **(Oct.-20)**

Data mining deals with the kind of patterns that can be mined. On the basis of the kind of data to be mined, there are two categories of functions involved in Data Mining -

1. Descriptive functions
2. Classification and Prediction

1. Descriptive Function

The descriptive function deals with the general properties of data in the database. Here is the list of descriptive functions.

- Class/Concept Description
- Mining of Frequent Patterns
- Mining of Associations
- Mining of Correlations
- Mining of Clusters

i) Class/Concept Description

Class/Concept refers to the data to be associated with the classes or concepts. For example, in a company, the classes of items for sales include computer and printers, and concepts of customers include big spenders and budget spenders. Such descriptions of a class or a concept are called class/concept descriptions. These descriptions can be derived by the following two ways.

a) Data Characterization

This refers to summarizing data of class under study. This class under study is called as Target Class.

b) Data Discrimination

It refers to the mapping or classification of a class with some predefined group or class.

ii) Mining of Frequent Patterns

Frequent patterns are those patterns that occur frequently in transactional data. Here is the list of kind of frequent patterns.

a) Frequent Item Set

It refers to a set of items that frequently appear together, for example, milk and bread.

b) Frequent Subsequence

A sequence of patterns that occur frequently such as purchasing a camera is followed by memory card.

c) Frequent Sub Structure

Substructure refers to different structural forms, such as graphs, trees, or lattices, which may be combined with item-sets or subsequences.

iii) Mining of Association

Associations are used in retail sales to identify patterns that are frequently purchased together. This process refers to the process of uncovering the relationship among data and determining association rules.

For example, a retailer generates an association rule that shows that 70% of time milk is sold with bread and only 30% of times biscuits are sold with bread.

iv) Mining of Correlations

It is a kind of additional analysis performed to uncover interesting statistical correlations between associated-attribute-value pairs or between two item sets to analyze that if they have positive, negative or no effect on each other.

v) Mining of Clusters

Cluster refers to a group of similar kind of objects. Cluster analysis refers to forming group of objects that are very similar to each other but are highly different from the objects in other clusters.

2 Classification and Prediction

Classification is the process of finding a model that describes the data classes or concepts. The purpose is to be able to use this model to predict the class of objects whose class label is unknown. This derived model is based on the analysis of sets of training data. The derived model can be presented in the following forms "

- Classification (IF-THEN) Rules
- Decision Trees
- Mathematical Formulae
- Neural Networks

2.2.4 Benefits/Advantages of Data Mining

Q21. What are the benefits/advantages of data mining.

Ans :

Data mining is an important part of knowledge discovery process that we can analyze an enormous set of data and get hidden and useful knowledge. Data mining is applied effectively not only in the business environment but also in other fields such as weather forecast, medicine, transportation, healthcare, insurance, government etc. Data mining has a lot of advantages when using in a specific industry. Besides those advantages, data mining also has its own disadvantages e.g., privacy, security, and misuse of information. We will examine those advantages and disadvantages of data mining in different industries in a greater detail.

Advantages of Data Mining

1. Marketing / Retail

Data mining helps marketing companies build models based on historical data to predict who will respond to the new marketing

campaigns such as direct mail, online marketing campaigc etc. Through the results, marketers will have an appropriate approach to selling profitable products to targeted customers.

Data mining brings a lot of benefits to retail companies in the same way as marketing. Through market basket analysis, a store can have an appropriate production arrangement in a way that customers can buy frequent buying products together with pleasant. In addition, it also helps the retail companies offer certain discounts for particular products that will attract more customers.

2. Finance / Banking

Data mining gives financial institutions information about loan information and credit reporting. By building a model from historical customer's data, the bank, and financial institution can determine good and bad loans. In addition, data mining helps banks detect fraudulent credit card transactions to protect credit card's owner.

3. Manufacturing

By applying data mining in operational engineering data, manufacturers can detect faulty equipment and determine optimal control parameters. For example, semiconductor manufacturers have a challenge that even the conditions of manufacturing environments at different wafer production plants are similar, the quality of wafer are a lot the same and some for unknown reasons even has defects. Data mining has been applying to determine the ranges of control parameters that lead to the production of the golden wafer. Then those optimal control parameters are used to manufacture wafers with desired quality.

4. Governments

Data mining helps government agency by digging and analyzing records of the financial transaction to build patterns that can detect money laundering or criminal activities.

2.2.5 Data Mining Applications

Q22. Discuss about data mining applications.

Ans :

Data mining has become a popular tool in addressing many complex businesses issues. It has been proven to be very successful and helpful in many areas, some of which are shown by the following representative examples. The goal of many of these business data mining applications is to solve a pressing problem or to explore an emerging business opportunity in order to create a sustainable competitive advantage.

1. Customer Relationship Management

Customer relationship management (CRM) is the new and emerging extension of traditional marketing. The goal of CRM is to create one-on-one relationships with customers by developing an intimate understanding of their needs and wants. As businesses build relationships with their customers over time through a variety of transactions (e.g., product inquiries, sales, service requests, warranty calls), they accumulate tremendous amount of data. When combined with demographic and socioeconomic attributes, this information-rich data can be used to

i) Identify most likely responders/ buyers of new products/services (i.e., customer profiling);

ii) Understand the root causes of customer attrition in order to improve customer retention (i.e., churn analysis);

iii) Discover time-variant associations between products and services to maximize sales and customer value; and

iv) Identify the most profitable customers and their preferential needs to strengthen relationships and to maximize sales.

2. Banking

Data mining can help banks with the following:

i) Automating the loan application process by accurately predicting the most probable defaulters;

ii) Detecting fraudulent credit card and online-banking transactions;

iii) Identifying ways to maximize customer value by selling them products and services that they are most likely to buy; and

iv) Optimizing the cash return by accurately forecasting the cash flow on banking entities (e.g., ATM machines, banking branches).

3. Retailing and logistics

In the retailing industry, data mining can be used to

i) Predict accurate sales volumes at specific retail locations in order to determine correct inventory levels;

ii) Identify sales relationships between different products (with market-basket analysis) to improve the store layout and optimize sales promotions;

iii) Forecast consumption levels of different product types (based on sea-sonal and environmental conditions) to optimize logistics and hence maximize sales; and

iv) Discover interesting patterns in the movement of products (especially for the products that have a limited shelf life because they are prone to expiration, perishability, and contamination) in a supply chain by analyzing sensory and RFID data.

4. Manufacturing and production

Manufacturers can use data mining to

i) predict machinery failures before they occur through the use of sensory data (enabling what is called condition-based maintenance).

ii) identify anomalies and commonalities in production systems to optimize manufacturing capacity; and

iii) discover novel patterns to identify and improve product quality.

5. Brokerage and Securities Trading

Brokers and traders use data mining to

i) predict when and how much certain bond prices will change;

ii) forecast the range and direction of stock fluctuations;

iii) assess the effect of particular issues and events on overall market movements; and

iv) identify and prevent fraudulent activities in securities trading.

6. Insurance

The insurance industry uses data mining techniques to

i) forecast claim amounts for property and medical coverage costs for better business planning;

ii) determine optimal rate plans based on the analysis of claims and customer data;

iii) predict which customers are more likely to buy new policies with special features; and

iv) identify and prevent incorrect claim payments and fraudulent activities.

7. Computer Hardware and Software

Data mining can be used to

i) predict disk drive failures well before they actually occur;

ii) identify and filter unwanted Web content and e-mail messages;

iii) detect and prevent computer network security bridges; and

iv) identify potentially unsecure software products.

8. Government and Defense

Data mining also has a number of military applications. It can be used to

i) forecast the cost of moving military personnel and equipment;

ii) predict an adversary's moves to develop more successful strategies for military engagements;

iii) predict resource consumption for better planning and budgeting; and

iv) Identify classes of unique experiences, strategies, and lessons learned from military operations for better knowledge sharing throughout the organization.

9. Travel industry (airlines, hotels/resorts, rental car companies)

Data mining has a variety of uses in the travel industry. It is successfully used to

i) predict sales of different services (seat types in airplanes, room types in hotels/ resorts, car types in rental car companies) in order to optimally price services to maximize revenues as a function of time-varying transactions (commonly referred to as yield management)

ii) forecast demand at different locations to better allocate limited organizational resources;

iii) identify the most profitable customers and provide them with personalized services to maintain their repeat business.

iv) retain valuable employees by identifying and acting on the root causes for attrition.

10. Health Care

Data mining has a number of health care applications. It can be used to,

i) identify people without health insurance and the factors underlying this undesired phenomenon.

ii) identify novel cost-benefit relationships between different treatments to develop more effective strategies;

iii) forecast the level and the time of demand at different service locations to optimally allocate organizational resources; and

iv) understand the underlying reasons for customer and employee attrition.

11. Medicine

Use of data mining in medicine should be viewed as an invaluable complement to traditional medical research, which is mainly clinical and biological in nature. Data mining analyses can

i) identify novel patterns to improve survivability of patients with cancer;

ii) predict success rates of organ transplantation patients to develop better donor-organ matching policies;

iii) identify the functions of different genes in the human **chromosome** (known as genomics); and

iv) discover the relationships between symptoms and illnesses (as well as illnesses and successful treatments) to help medical professionals make informed and correct decisions in a timely manner.

12. Entertainment Industry

Data mining is successfully used by the entertainment industry to,

i) Analyze viewer data to decide what programs to show during prime time and how to maximize returns by knowing where to insert advertisements;

ii) Predict the financial success of movies before they are produced to make investment decisions and to optimize the returns;

iii) Forecast the demand at different locations and different times to better schedule entertainment events and to optimally allocate resources; and

iv) Develop optimal pricing policies to maximize revenues.

13. Homeland Security and Law Enforcement

Data mining has a number of home-land security and law enforcement applications. Data mining is often used to

i) identify patterns of terrorist behaviors

ii) discover crime patterns (e.g., locations, timings, criminal behaviors, and other related attributes) to help solve criminal cases in a timely manner;

iii) predict and eliminate potential biological and chemical attacks to a nation's critical infrastructure by analyzing special-purpose sensory data; and

ivi) identify and stop malicious attacks on critical information infrastructures (often called information warfare').

14. Sports

Data mining was used to improve the performance of National Basketball Association (NBA) teams in the United States. The NBA developed Advanced Scout, a PC-based data mining application that coaching staff use to discover interesting patterns in basketball game data. The pattern interpretation is facilitated by allowing the user to relate patterns to videotape.

2.2.6 Data Mining Techniques & Tools

Q23. Discuss about Data Mining Techniques & Tools.

Ans : **(June-18)**

There are several major data mining techniques have been developing and using in data mining projects recently including association, classification, clustering, prediction, sequential patterns and decision tree. We will briefly examine those data mining techniques in the following sections.

1. Association

Association is one of the best-known data mining technique. In association, a pattern is discovered based on a relationship between items in the same transaction. That's is the reason why association technique is also known as relation technique. The association technique is used in market basket analysis to identify a set of products that customers frequently purchase together.

Retailers are using association technique to research customer's buying habits. Based on historical sale data, retailers might find out that customers always buy crisps when they buy beers, and, therefore, they can put beers and crisps next to each other to save time for the customer and increase sales.

2. Classification

Classification is a classic data mining technique based on machine learning. Basically, classification is used to classify each item in a set of data into one of a predefined set of classes or groups. Classification method makes use of mathematical techniques such as decision trees, linear programming, neural network, and statistics. In classification, we develop the software that can learn how to classify the data items into groups. For example, we can apply classification in the application that "given all records of employees who left the company, predict who will probably leave the company in a future period." In this case, we divide the records of employees into two groups that named "leave" and "stay". And then we can ask our data mining software to classify the employees into separate groups.

3. Clustering

Clustering is a data mining technique that makes a meaningful or useful cluster of objects which have similar characteristics using the automatic technique. The clustering technique defines the classes and puts objects in each class, while in the classification techniques, objects are assigned into predefined classes. To make the concept clearer, we can take book management in the library as an example. In a library, there is a wide range of books on various topics available. The challenge is how to keep those books in a way that readers can take several books on a particular topic without hassle.

By using the clustering technique, we can keep books that have some kinds of similarities in one cluster or one shelf and label it with a meaningful name. If readers want to grab books in that topic, they would only have to go to that shelf instead of looking for the entire library.

4. Prediction

The prediction, as its name implied, is one of a data mining techniques that discovers the relationship between independent variables and relationship between dependent and independent variables. For instance, the prediction analysis technique can be used in the sale to predict profit for the future if we consider the sale is an independent variable, profit could be a dependent variable. Then based on the historical sale and profit data, we can draw a fitted regression curve that is used for profit prediction.

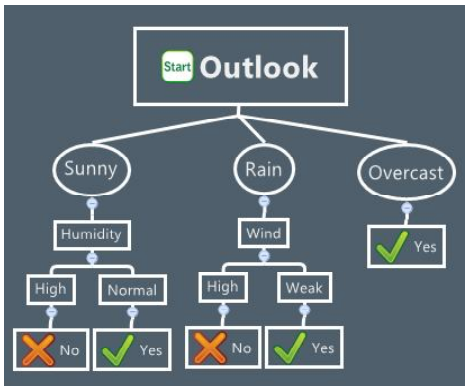
5. Sequential Patterns

Sequential patterns analysis is one of data mining technique that seeks to discover or identify similar patterns, regular events or trends in transaction data over a business period.

In sales, with historical transaction data, businesses can identify a set of items that customers buy together different times in a year. Then businesses can use this information to recommend customers buy it with better deals based on their purchasing frequency in the past.

6. Decision Trees

The A decision tree is one of the most commonly used data mining techniques because its model is easy to understand for users. In decision tree technique, the root of the decision tree is a simple question or condition that has multiple answers. Each answer then leads to a set of questions or conditions that help us determine the data so that we can make the final decision based on it. For example, We use the following decision tree to determine whether or not to play tennis:



Starting at the root node, if the outlook is overcast then we should definitely play tennis. If it is rainy, we should only play tennis if the wind is the week. And if it is sunny then we should play tennis in case the humidity is normal.

We often combine two or more of those data mining techniques together to form an appropriate process that meets the business needs.

Software Tools of Data Mining

Some of the software tools of data mining are,

1. IBM Intelligent Miner
2. SPSS Clementine
3. Oracle Data Mining (ODM)
4. Classification and Regression Tree (CART)
5. Weka.

1. IBM Intelligent Miner

IBM intelligent miner offers many data mining functions such as classification, association mining, predictive modeling, clustering, deviation detection, sequential pattern analysis and regression.

This tool also offers an application tool kit that contains statistical method, data visualization tools neural network al-gorithms and data preparation tools. It is highly scalable with mining algorithms and can be tightly coupled with IBM's DB2 database system.

2. SPSS Clementine

SPSS clementine offers an integrated development environment for developers

and end users to perform data mining. It performs various functions including prediction, clustering, association mining and classification. It is also equipped with object oriented module interface through which users can add their utilities to visual programing environment.

3. Oracle Data Mining (ODM)

Oracle Data Mining (ODM) tool is an add on feature of Oracle log Enterprise Edition database. This tool performs various data mining functions such as classification, prediction, sequence similarity search and analysis, association mining and clustering.

4. Classification and Regression Trees (CART)

Classification and Regression Trees (CART) is a data mining tool from Salford systems. This tool generates regression trees for prediction and decision trees for classification. It improves the accuracy of data mining by employing boosting mechanism.

5. Weka

Weka is an open-source data mining tool developed on Java platform in the university of Waikato, Newzealand. It is capable to perform various data mining functions such as regression, classification, association mining and data preprocessing.

Q24. Define

(i) Cluster Analysis

(ii) Association Rule Mining

Ans :

(i) Cluster Analysis

The process of collecting homogeneous data objects within the same group called cluster and heterogeneous to the objects in other groups (clusters) is called as cluster analysis.

Clustering is a technique of combining a group of physical objects into classes of homogeneous objects. Clustering of data objects into single class is equivalent to data compression.

Various fields such as data analysis, market research, image processing, pattern recognition, business, statistics and biology uses the concept of cluster analysis.

On the basis of 'similarity', huge data sets are partitioned into groups. This process is known as 'data segmentation' which is equivalent to clustering. Outlier detection applications use clustering, these include the detection of credit card fraud and the monitoring of criminal activities in electronic commerce.

Cluster analysis is a function of data mining which may be used as a unique tool to attain knowledge about the distribution of data, to note the features of every cluster and to carry out further analysis by paying attention on a particular group of clusters. Cluster analysis can also be used as a preprocessing step for characterization, attribute subset selection and classification algorithms. Clustering can also be used to help in classifying documents on web for detecting information.

There is a strong development in clustering of data. Statistics, spatial database technology, marketing, biology, data mining and machine learning fields are contributing a lot towards the research of cluster analysis.

(ii) Association Rule Mining

Association Rules are specified over both categorical and numerical attributes. These comprise of two statistical measures-'utility' and 'certainty'. These measures represent the usefulness and confidence of the extracted association rule. There are two threshold conditions set by experts, which must be satisfied in order to generate relevant association rules.

Association rule can be mined by considering the following steps,

(a) Search for Item Sets that Occur most Frequently

Frequent Item Sets are defined as a set of items, that occur as frequently as, predefined minimum support threshold items.

(b) Produce Strong Association Rule from the Resultant Item Set

Association Rule, must satisfy both Minimum Support Threshold, as well as Minimum Confidence threshold conditions.

Q25. List out various software tools of data mining.

Ans :

Some of the software tools of data mining are,

1. IBM Intelligent Miner
2. SPSS Clementine
3. Oracle Data Mining (ODM)
4. Classification and Regression Tree (CART)
5. Weka.

1. IBM Intelligent Miner

IBM intelligent miner offers many data mining functions such as classification, association mining, predictive modeling, clustering, deviation detection, sequential pattern analysis and regression.

This tool also offers an application tool kit that contains statistical method, data visualization tools neural network al-gorithms and data preparation tools. It is highly scalable with mining algorithms and can be tightly coupled with IBM's DB2 database system.

2. SPSS Clementine

SPSS clementine offers an integrated development environment for developers and end users to perform data mining. It performs various functions including prediction, clustering, association mining and classification. It is also equipped with object oriented module interface through which users can add their utilities to visual programing environment.

3. Oracle Data Mining (ODM)

Oracle Data Mining (ODM) tool is an add on feature of Oracle log Enterprise Edition database. This tool performs various data mining functions such as classification, prediction, sequence similarity search and analysis, association mining and clustering.

4. Classification and Regression Trees (CART)

Classification and Regression Trees (CART) is a data mining tool from Salford systems. This tool generates regression trees for prediction and decision trees for classification. It improves the accuracy of data mining by employing boosting mechanism.

5. Weka

Weka is an open-source data mining tool developed on Java platform in the university of Waikato, Newzealand. It is capable to perform various data mining functions such as regression, classification, association mining and data preprocessing.

2.2.7 Text Mining

Q26. Define text mining. Explain the applications of text mining.

Ans :

Text mining (also known as text data mining or knowledge discovery in textual databases') is the semiautomated process of extracting patterns (useful information and knowledge) from large amounts of unstructured data sources. Remember that data mining is the process of identifying valid, novel, potentially useful, and ultimately under-standable patterns in data stored in structured databases, where the data are organized in records structured by categorical, ordinal, or continuous variables.

Text mining is the same as data mining in that it has the same purpose and uses the same processes; but with text mining, the input to the process is a collection of unstructured (or less structured) data files such as Word documents, PDF files, text excerpts, XML files, and so on.

In essence, text mining can be thought of as a process (with two main steps) that starts with imposing structure to the text-based data sources followed by extracting relevant information and knowledge from this structured text-based data using data mining techniques and tools.

Applications of Text Mining

Following are among the most popular application areas of text mining:

1. Information Extraction

Identification of key phrases and relationships within text by looking for predefined sequences in text via pattern matching.

2. Topic Tracking

Based on a user profile and documents that a user views, text mining can predict other documents of interest to the user.

3. Summarization

Summarizing a document to save time on the part of the reader.

4. Categorization

Identifying the main themes of a document and then placing the document into a predefined set of categories based on those themes.

5. Clustering

Grouping similar documents without having a predefined set of categories.

6. Concept Linking

Connects related documents by identifying their shared concepts and, by doing so, helps users find information that they perhaps would not have found using traditional search methods.

7. Question Answering

Finding the best answer to a given question through knowledge-driven pattern matching.

8. Text Mining Lingo

The following list describes some commonly used text mining terms:

9. Unstructured data (versus structured data)

Structured data have a predetermined format. They are usually organized into records with simple data values (categorical, ordinal, and continuous variables) and stored in databases. In contrast, unstructured data do not have a predetermined format and are stored in the form of textual documents. In essence, the structured: data are for the computers to process while the unstructured data are for humans to process and understand.

10. Corpus

In linguistics, a **corpus** (plural corpora) is a large and structured set of texts (now usually stored and processed electronically) prepared for the purpose of conducting knowledge discovery.

11. Terms

A term is a single word or multiword phrase extracted directly from the corpus of a specific domain by means of natural language processing (NLP) methods.

12. Concepts

Concepts are features generated from a collection of documents by means of manual, statistical, rule-based, or hybrid categorization methodology. Compared to terms, concepts are the result of higher level abstraction.

13. Stemming

The process of reducing inflected words to their stem (or base or root) form. For instance, stemmer, stemming, stemmed are all based on the root stem.

14. Stop Words

Stop words (or noise words) are words that are filtered out prior to or after processing of natural language data (i.e., text). Even though there is no universally accepted list of stop words, most natural language processing tools use a list that includes articles (a, am, the, of, etc.), auxiliary verbs (is, are, was, were, etc.), and context-specific words that are deemed not to have differentiating value.

15. Synonyms and Polysemes

Synonyms are syntactically different words (i.e., spelled differently) with identical or at least similar meanings (e.g., movie, film, and motion picture). In contrast, polysemes, which are also called homonyms, are syntactically identical words (i.e., spelled

exactly the same) with different meanings (e.g., bow can mean "to bend for-ward," "the front of the ship," "the weapon that shoots arrows," or "a kind of tied ribbon").

16. Tokenizing

A token is a categorized block of text in a sentence. The block of text corresponding to the token is categorized according to the function it performs. This assignment of meaning to blocks of text is known as tokenizing. A token can look like anything; it just needs to be a useful part of the structured text.

17. Term Dictionary

A collection of terms specific to a narrow field that can be used to restrict the extracted terms within a corpus.

18. Word Frequency

The number of times a word is found in a specific document.

19. Part-of-speech Tagging

The process of marking up the words in a text as corresponding to a particular part of speech (such as nouns, verbs, adjectives, and adverbs) based on a word's definition and the context in which it is used.

20. Morphology

A branch of the field of linguistics and a part of natural language processing that studies the internal structure of words (patterns of word formation within a language or across languages).

21. Tetmt-by-document matrix (occurrence matrix or term-document matrix)

A common representation schema of the frequency-based relationship between the terms and documents in tabular format where terms are listed in rows, documents are listed in columns, and the frequency between the terms and documents is listed in cells as integer values.

22. Singular-value decomposition (latent semantic indexing)

A dimensionality reduction method used to transform the term-by-document matrix to a manageable size by generating an intermediate representation of the frequencies using a matrix manipulation method similar to principle component analysis.

Q27. What are the different types of Text Mining Applications ?

Ans :

As the amount of unstructured data collected by organizations increases, so does the value proposition and popularity of text mining tools. Many organizations are now realizing the importance of extracting knowledge from their document-based data repositories through the use of text mining tools. Following are only a small subset of the exemplary application categories of text mining.

1. Marketing Applications

Text mining can be used to increase cross-selling and up-selling by analyzing the unstructured data generated by call centers.

- Text generated by call-center notes as well as transcriptions of voice conversations with customers-can be analyzed by text mining algorithms to extract novel, actionable information about customers' perceptions toward a company's products and services.
- Additionally, blogs, user reviews of products at independent Web sites, and discussion board postings are a gold mine of customer sentiments.
- This rich collection of information, once properly analyzed, can be used to increase satisfaction and the overall lifetime value of the customer.
- Text mining has become invaluable for customer relationship management. Companies can use text mining to analyze rich sets of unstructured text data, combined with the relevant structured data extracted from

organizational databases, to predict customer perceptions and subsequent purchasing behavior.

- Coussement and Van den Poel (2009) successfully applied text mining to significantly improve the ability of a model to predict customer churn (i.e., customer attrition) so that those customers identified as most likely to leave a company are accurately identified for retention tactics.
- Ghani et al. (2006) used text mining to develop a system capable of inferring implicit and explicit attributes of products to enhance retailers' ability to analyze product data-bases. Treating products as sets of attribute-value pairs rather than as atomic entities can potentially boost the effectiveness of many business applications, including demand fore-casting, assortment optimization, product recommendations, assortment comparison across retailers and manufacturers, and product supplier selection.
- The proposed system allows a business to represent its products in terms of attributes and attribute values with-out much manual effort. The system learns these attributes by applying supervised and semisuper vised learning techniques to product descriptions found on retailers' Web sites.

2. Security Applications

One of the largest and most prominent text mining applications in the security domain is probably the highly classified ECHELON surveillance system.

- As rumor has it, ECHELON is assumed to be capable of identifying the content of telephone calls, faxes, e-mails, and other types of data, intercepting information sent via satellites, public switched telephone networks, and microwave links.
- In 2007, EUROPOL developed an integrated system capable of accessing, storing, and analyzing vast amounts of structured and unstructured data sources in order to track transactional organized crime.
- Called the Overall Analysis System for Intelligence Support, this system aims to integrate the most advanced data and text mining technologies available on today's market.
- The system has enabled EUROPOL to make significant progress in supporting its law enforcement objectives at the international level (EUROPOL, 2007).
- The U.S. Federal Bureau of Investigation (FBI) and the Central Intelligence Agency (CIA), under the direction of the Department for Homeland Security, are jointly developing a superconductor data and text mining system.
- The system is expected to create a gigantic data warehouse along with a variety of data and text mining modules to meet the knowledge-discovery needs of federal, state, and local law enforcement agencies. Prior to this project, the FBI and CIA each had its own separate databases, with little or no inter-connection.

3. Biomedical Applications

Text mining holds great potential for the medical field in general and biomedicine in particular for several reasons.

- First, the published literature and publication outlets (especially with the advent of the open source journals) in the field are expanding at an exponential rate.
- Second, compared to most other fields, the medical literature is more standardized and orderly, making it a more "minable" information source.
- Finally, the terminology used in this literature is relatively constant, having a fairly standardized ontology. What follows are a few exemplary studies where text mining techniques were successfully used in extracting novel patterns from biomedical literature.

- Experimental techniques such as DNA microarray analysis, serial analysis of gene expression (SAGE), and mass spectrometry proteomics, among others, are generating large amounts of data related to genes and proteins.
- As in any other experimental approach, it is necessary to analyze this vast amount of data in the context of previously known information about the biological entities under study.
- The literature is a particularly valuable source of information for experiment validation and interpretation.
- Therefore, the development of automated text mining tools to assist in such interpretation is one of the main challenges in current bioinformatics research.

4. Academic Applications

The issue of text mining is of great importance to publishers who hold large data-bases of information requiring indexing for better retrieval.

This is particularly true in scientific disciplines, in which highly specific information is often contained within written text. Initiatives have been launched, such as Nature's proposal for an Open Text Mining Interface and the National Institutes of Health's common Journal Publishing Document Type Definition (DTD), that would provide semantic cues to machines to answer specific queries contained within text without removing publisher barriers to public access.

Q28. Outline the process of Text Mining.

Ans :

The three major steps involved in text mining process are,

1. Establishing the Corpus
2. Developing the Term-Document matrix
3. Extracting the Knowledge

The typical process of text mining is illustrated in the figure below,

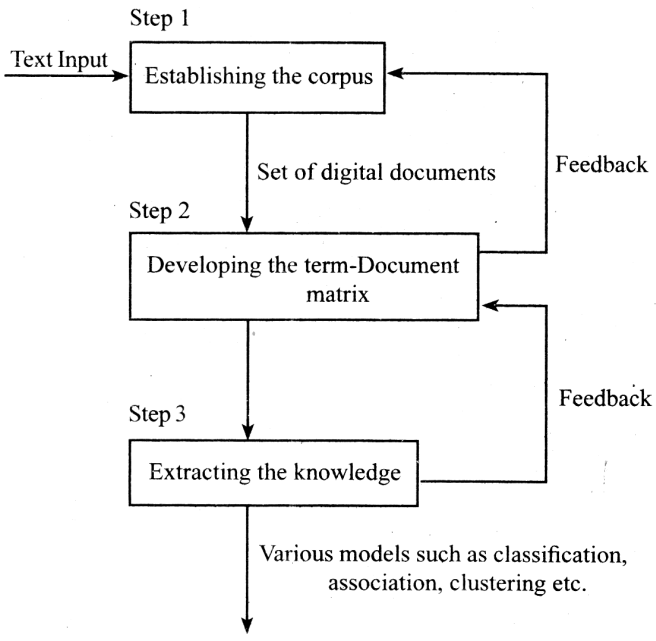


Figure: Task-mining Process

1. Establishing the Corpus

In this step, the extracted text documents such as word documents, HTML files, XML files etc., including voice recordings are added to the text collection. These documents are arranged and transferred into a specific format so that it can be further processed by the computer. The text documents in the digitized format are forwarded to the second step as its input.

2. Developing the Term-document Matrix

In this step, a term-document matrix is developed from the digitized documents provided as input to this step. This matrix contains a set of rows and columns where rows represent documents and columns represent the terms as shown below.

	Risk of Investment	Software Engineering	SAP	Development
DOC1	1		1	
DOC2		2		
DOC3	1		1	2
DOC4		3		

The cells between rows and columns of the matrix carry indices. The values of indices are nothing but the frequency of the terms occurring in the documents. The common terms such as articles, auxiliary verbs are excluded and the terms considered in the matrix are called stop terms. Synonyms can also be included to make the search appropriate.

To perform proper Alteration in indices, stemming method can be used. This method considers all the forms of a verb as a single word thereby minimizing the number of words. An example of stemming is word pick which considers all its verb forms i.e., picked, picking.

3. Extracting the Knowledge

In this step, the terms-document matrix generated from step2 is integrated with knowledge extraction methods such as classification, association, clustering etc., to generate knowledge from the TDM.

2.2.8 Web Mining

Q29. Define Web mining. What are the different areas of web mining ?

Ans :

Web mining (or Web data mining) is the process of discovering intrinsic relation-ships (i.e., interesting and useful information) from Web data, which are expressed in the form of textual, linkage, or usage information. The term Web mining was first used by Etzioni (1996); today, many conferences, journals, and books focus on Web data mining. It is a continually evolving area of technology and business practice. Figure presents the three main areas of Web mining: Web content, Web structure, and Web usage mining.

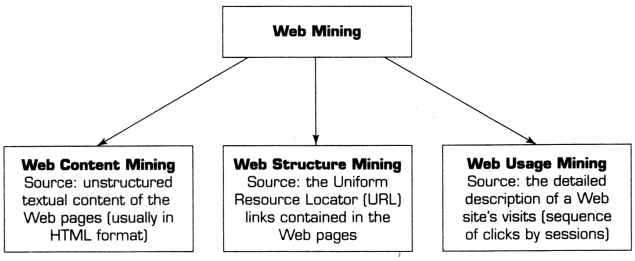


Figure : The Three Main Areas of Web Mining

Areas of Web Mining

The three main areas of web mining are,

1. Web content mining
2. Web structure mining
3. Web usage mining

1. Web Content Mining

Web content mining is a process of extracting relevant information from web contents. Basically, web content comprises not only textual information but also graphical information, real-time information and hyperlinks. The textual information in web content data is a combination of unstructured (free text), semi-structured (HTML pages) and highly structured (database generated HTML pages) data. However, web content data is unstructured due to which text mining techniques can be used for performing web content mining.

2. WebStructure Mining

Web structure mining is the process of generating summary of webpages of websites. Web structuring checks link structure i.e., hyperlinks among various websites and classifies the webpages on the basis of hyperlinks found. It discovers the relationship of webpages and the pages to which the links are found. The relationship is determined on the basis of synonyms or similar contents found on the webpages.

It also determines the network in a certain domain. This determination makes the process of querying more easier and efficient.

3. Web Usage Mining

The Web usage mining simply refers to the process of searching the Weblog records in order to determine the procedures in which users can access the Web pages.

If the Weblog records database are examined carefully, it.

a) Determines all the possible e-commerce customers.

b) Efficiently delivers the qualitative Internet Information services to the user.

c) Enhances the performance of the Web server.

The Web server will record a new Weblog entry every time the Web page is accessed. This web log entry contains the information about the address of the web page (i.e., URL), IP address of the requester, and a timestamp.

The most frequently referred Websites and Web-based e-commerce servers need to maintain a large Weblog database which consists of information about millions of Weblog records. This Weblog database can be accessed with the help of Weblog mining techniques.

Short Question and Answers

1. Define Data Warehousing.

Ans :

A data warehouse is a subject-oriented, integrated, time-variant and non-volatile collection of data in support of management's decision making process.

i) Subject-Oriented

A data warehouse can be used to analyze a particular subject area. For example, "sales" can be a particular subject.

ii) Integrated

A data warehouse integrates data from multiple data sources. For example, source A and source B may have different ways of identifying a product, but in a data warehouse, there will be only a single way of identifying a product.

iii) Time-Variant

Historical data is kept in a data warehouse. For example, one can retrieve data from 3 months, 6 months, 12 months, or even older data from a data warehouse. This contrasts with a transactions system, where often only the most recent data is kept.

For example, a transaction system may hold the most recent address of a customer, where a data warehouse can hold all addresses associated with a customer.

2. What is metadata ?

Ans :

Metadata are data about data. When used in a data warehouse, metadata are the data that define warehouse objects. Metadata are created for the data names and definitions of the given warehouse. Additional metadata are created and captured for time stamping any extracted data, the source of the extracted data, and missing fields that have been added by data cleaning or integration processes. A metadata repository should contain the following:

- A description of the structure of the data warehouse, which includes the warehouse schema, view, dimensions, hierarchies, and derived data definitions, as well as data mart locations and contents.
- Operational metadata, which include data lineage (history of migrated data and the sequence of transformations applied to it), currency of data (active, archived, or purged), and monitoring information (warehouse usage statistics, error reports, and audit trails).
- The algorithms used for summarization, which include measure and dimension definition algorithms, data on granularity, partitions, subject areas, aggregation, summarization, and predefined queries and reports.
- The mapping from the operational environment to the data warehouse, which includes source databases and their contents, gateway descriptions, data partitions, data extraction, cleaning, transformation rules and defaults, data refresh and purging rules, and security (user authorization and access control).

3. Define Data Integration.

Ans :

Data integration refers to the process of combining data from different data sources such as databases, flat files etc. It involves three processes which enable users to access data and also to array of ETL, analysis tools and data warehousing environment to access the data. These three processes are as follows,

1. Data Access

This process allows data to be accessed and extracted from any type of sources.

2. Data Federation

This process allows business views to be integrated with different data stores.

3. Change Capture

This process captures the modifications done with respect to identification, capture and delivery of data sources. Integration of data and metadata can be performed using the following technologies.

4. Data Extraction

Ans :

The function of the data extraction is to find and extract the data from various relational and non-relational data sources for future use. Data extraction process includes,

(i) Determining the operational systems from which the data can be collected.

(ii) Capture the appropriate data from data sources to make required data analysis.

(iii) Identify the extraction frequency of data i.e., daily, weekly or monthly to keep the data warehouse up to date.

5. Define Data Mining.

Ans :

Data mining refers to extracting or mining knowledge from large amounts of data. The term is actually a misnomer. Thus, data miningshould have been more appropriately named as knowledge mining which emphasis on mining from large amounts of data.

It is the computational process of discovering patterns in large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems. The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use.

- The key properties of data mining are
- Automatic discovery of patterns
- Prediction of likely outcomes
- Creation of actionable information
- Focus on large datasets and databases.

6. Functions of Data Mining

Ans :

1. Descriptive Function

The descriptive function deals with the general properties of data in the database. Here is the list of descriptive functions.

- Class/Concept Description
- Mining of Frequent Patterns
- Mining of Associations
- Mining of Correlations
- Mining of Clusters

i) Class/Concept Description

Class/Concept refers to the data to be associated with the classes or concepts. For example, in a company, the classes of items for sales include computer and printers, and concepts of customers include big spenders and budget spenders. Such descriptions of a class or a concept are called class/concept descriptions. These descriptions can be derived by the following two ways.

a) Data Characterization

This refers to summarizing data of class under study. This class under study is called as Target Class.

b) Data Discrimination

It refers to the mapping or classification of a class with some predefined group or class.

ii) Mining of Frequent Patterns

Frequent patterns are those patterns that occur frequently in transactional data. Here is the list of kind of frequent patterns.

a) Frequent Item Set

It refers to a set of items that frequently appear together, for example, milk and bread.

b) Frequent Subsequence

A sequence of patterns that occur frequently such as purchasing a camera is followed by memory card.

c) Frequent Sub Structure

Substructure refers to different structural forms, such as graphs, trees, or lattices, which may be combined with item-sets or subsequences.

7. Cluster Analysis

Ans :

The process of collecting homogeneous data objects within the same group called cluster and heterogeneous to the objects in other groups (clusters) is called as cluster analysis.

Clustering is a technique of combining a group of physical objects into classes of homogeneous objects. Clustering of data objects into single class is equivalent to data compression.

Various fields such as data analysis, market research, image processing, pattern recognition, business, statistics and biology uses the concept of cluster analysis.

On the basis of 'similarity', huge data sets are partitioned into groups. This process is known as 'data segmentation' which is equivalent to clustering. Outlier detection applications use clustering, these include the detection of credit card fraud and the monitoring of criminal activities in electronic commerce.

Cluster analysis is a function of data mining which may be used as a unique tool to attain knowledge about the distribution of data, to note the features of every cluster and to carry out further analysis by paying attention on a particular group of clusters. Cluster analysis can also be used as a preprocessing step for characterization, attribute subset selection and classification algorithms. Clustering can also be used to help in classifying documents on web for detecting information.

8. Association Rule Mining

Ans :

Association Rules are specified over both categorical and numerical attributes. These comprise of two statistical measures-'utility' and 'certainty'. These measures represent the usefulness and confidence of the extracted association rule. There are two threshold conditions set by experts, which must be satisfied in order to generate relevant association rules.

Association rule can be mined by considering the following steps,

(a) Search for Item Sets that Occur most Frequently

Frequent Item Sets are defined as a set of items, that occur as frequently as, predefined minimum support threshold items.

(b) Produce Strong Association Rule from the Resultant Item Set

Association Rule, must satisfy both Minimum Support Threshold, as well as Minimum Confidence threshold conditions.

9. Applications of Text Mining

Ans :

Following are among the most popular application areas of text mining:

i) **Information Extraction:** Identification of key phrases and relationships within text by looking for predefined sequences in text via pattern matching.

ii) **Topic Tracking:** Based on a user profile and documents that a user views, text mining can predict other documents of interest to the user.

iii) **Summarization:** Summarizing a document to save time on the part of the reader.

iv) **Categorization:** Identifying the main themes of a document and then placing the document into a predefined set of categories based on those themes.

v) **Clustering:** Grouping similar documents without having a predefined set of categories.

vi) **Concept Linking:** Connects related documents by identifying their shared concepts and, by doing so, helps users find information that they perhaps would not have found using traditional search methods.

10. Define Web mining.

Ans :

Web mining (or Web data mining) is the process of discovering intrinsic relation-ships (i.e., interesting and useful information) from Web data, which are expressed in the form of textual, linkage, or usage information. The term Web mining was first used by Etzioni (1996); today, many conferences, journals, and books focus on Web data mining. It is a continually evolving area of technology and business practice. Figure presents the three main areas of Web mining: Web content, Web structure, and Web usage mining.

Choose the Correct Answers

1. What is type of dataware house [d]
   (a) Data mart (b) OLAP
   (c) OLTP (d) All
2. Feature of data ware housing [d]
   (a) Sub oriented (b) Meta data
   (c) Server (d) All
3. Alternative dataware housing is _____ [a]
   (a) independent data mart (b) Social data mart
   (c) Subjective data mart (d) Informative data mart
4. _____ mining tasks characterize the general properties of the data in the databases. [a]
   (a) Descriptive (b) Natural
   (c) Real time (d) All
5. Technique of data integration [c]
   (a) Middle ware data integration (b) Manual data integration
   (c) a & b (d) None of these
6. The of an itemset identifies how frequency a rule is applied to the provided dataset. [b]
   (a) Count (b) Support
   (c) Confidence (d) Rule
7. Data ware house approach is [a]
   (a) The inmon model (b) EDW
   (c) ERD (d) All
8. Example of the type of data warehouse [a]
   (a) Data mart (b) BCD
   (c) Real time (d) All the above
9. Tier-3 of data warehouse architecture represents [c]
   (a) Front-end client layer (b) Application server
   (c) Warehouse server (d) None of the above
10. Application of data mining is [d]
   (a) CRM (b) Retailing
   (c) Insurance (d) All

Fill in the Blanks

1. OLAP stand for _____ .
2. OLTP stand for _____ .
3. _____ is a process of extraction, transformation and loading.
4. Application of text mining _____.
5. Area of web mining _____ .
6. Classification is one of the method _____.
7. Function data mining _____ .
8. ODS stands for _____.
9. Technique of data mining _____.
10. The Inmon model is _____ approach.

ANSWER

1. Online analycatical process
2. One line transaction processing
3. ETL
4. Clustering
5. Web content mining
6. Data analysis
7. Descrietive function
8. Operational datastores
9. Association
10. Dataware house approach

UNIT III

Business Performance Measurement (BPM)

Definition, BPM vs BI, Summary of BPM Process, Performance Measurement, BPM Methodologies, BPM Architecture and Applications, Performance Dash boards.

3.1 Business Performance Measurement - Definition

Q1. What is Business Performance ?

Ans :

Business performance evaluates the overall organizational performance in significant areas like customer satisfaction, service and product performance, operational and financial performance, cost leadership, human resources, human capital. Performance of the company is also assessed in relation to that of the rival companies.

An environment from which business takes inputs, processes them and produces its outputs is usually referred as business environment. An environment comprises various external factors which influence business performance. They are political and legal factors, social and cultural factors, technological factors, global factors and natural factors.

Q2. Define business performance measure-ment. Explain the features and components of business performance measurement.

Ans :

BPM may be defined as "Different business tools, techniques, methodologies and metrics employed by an organization to measure, monitor and manage the performance of the business enterprise". There is always a confusion between the term Performance Management and Business Performance Management (BPM). However both are interlinked with each other and in most cases used as synonyms.

Definitions

i) **According to Dr. T.V Rao,** "performance management involves, thinking through various, facts of performance, identifying critical dimensions performance, planning, reviewing, developing and enhancing performance and related competencies".

ii) **According to Reward Management Associates** "performance management may be defined as managerial process which consist of planning performance, managing performance through observation and feedback, appraising performance and rewarding performance".

iii) In simple words we any define performance management as "a continuous process of identifying, measuring and developing the performance levels of individual is and aligning it with the strategic goals of the organization.

Features of Business Performance Measurement

Some key features related to business performance management are,

(a) BPM is known by other names such as Corporate Performance Management (CPM), Enterprise Performance Management (EPM), Strategic Performance Management (SPM).

(b) The term CPM was coined by the market analytics firm Gartner.com. Whereas the term BPM was first used by the BPM standards group.

(c) BPM employs a wide range of tools, techniques and metrics to assess the actual performance of the business organization.

Components of Business Performance Measurements

According to business analytics researcher colbert (2009) following are the major components of BPM,

(i) A key component of BPM is a set of closely linked management and analytical processes. These processes are fully integrated with technology. The primary focus area of these close looped management processes is to address and assess the financial and operational activities of the organization.

(ii) BPM components include all the business tools and techniques which are used for defining, formulating and measuring the strategic goals of the organization. These tools and techniques are used for not only formulating strategic goals, but also for measuring the actual organizational performance with the previously set goals.

(iii) Other important components of BPM includes tools and techniques of financial and operational planning, metrics employed for the measurement of key performance indications and so on.

Thus, it may be stated that business organizations operate in a highly competitive environment. Thus, they need to adapt, adjust and align themselves on a continuous basis as per the changing business environment. Researcher Axson (2017) had stated that "An organization's performance management processes are the principle mechanism for assessing the impact of change and turning the business in order to survive and prosper.

3.2 BPM Vs. BI

Q3. Compare and Contrast Business Performance measurement and business intelligence.

Ans : **(June-19)**

Business Performance Measurement

- BPM is promoted and sold by the same companies that market and sell the BI tools and suites.
- BI has evolved so that many of the original differences between the two no longer exist (e.g., BI used to be focused on departmental rather than on enterprise-wide projects).
- BI is a crucial element of BPM.

Business Intelligence

The term BI now describes the technology used to access, analyze, and report on data relevant to an enterprise. It encompasses a wide spectrum of software, including ad hoc querying, reporting, online analytical processing (OLAP), dashboards, scorecards, search, visualization, and more. These software products started as stand-alone tools, but BI software vendors have incorporated them into their BI suites.

The following are the key differences between BPM and BI :

S. No.	Area	Business Performance Measurement	Business Intelligence
1.	**Definition/ Meaning**	BPM refers to the different business tools, methods and metrics employed by an organization to measure, moniter and manage business performance.	BI refers to the art and science of analysing any data or information, with a specific business objective.
2.	**Scope**	The scope of BPM is wider. Impact BPM can be considered as an outgrowth of BI.	The scope of BI is comparatively narrower.
3.	**Tools**	BPM employs multiple tools, methods and metrics, which includes all the BI tools as well.	BI employs a wide variety of tools and techniques such as OLP adhoc querying, dashboards, score-cards etc.
4.	**Processes**	It involves multiple processes including BI+, planning, cycle of plan, monitoring and analysing of business performance.	BI processes are limited only upto analysing of business data.
5.	**Nature**	BPM is highly flexible in nature and is actively integrated with the ongoing organizational projects. It would be aligned as per the changing business needs.	BI is static in nature. Its primary focus is upon measuring the performance of already completed business projects.

Q4. Outline the process of business performance Measurement.

Ans : **(June-18**

The various steps involved in the process of business performance measurement are as follows,

Step-1: Define 'What to Measure'

The first step in the process of business performance measurement is defining what aspects of the organization need to be measured. Doing so involves the following,

(i) Decision about specific outcomes to be measured.

(ii) Design measures which give best proof of such outcomes.

(iii) Define measures of performance.

Setp-2: Collect Performance Data

After defining performance measures, focus should be laid on collecting performance related data. In this step, what data to be collected, from where such data to be collected and how such data to be collected is decided. Data collection system must be designed, improved and executed to maximize availability of data.

Step-3: Store the Collected Data

The data collected in step-2 must be stored properly so that stored data can be easily accessed when ever required.

Data referencing model can be used to minimize the cost involved in data management and to facilitate cross-functional usage of data.

Step-4: Convert the Raw Data into Information

The raw data collected is converted into information through analysis. The easiest analysis approach must be adopted to produce information in the required form that can resolve all the queries. Analysis step involves selecting analysis methods and application of analysis method to raw data.

Step-5: Communication of Information

Communicate performance information in such a manner that the information provide easy, trust worthy, relevant and visual solutions to the queries. This step involves activities like - designing graphs which helps in interpretation and decision taking, designing performance reports for owners and audience of performance measures and designing and executing performance reporting processes.

Step-6: Interpretation

In this step, performance information is interpreted to draw conclusions on what is actually happening. The conclusions drawn with respect to performance outcomes helps in taking decision, whether to take correct measures or not.

Step-7: Action

In this step, decisions are taken regarding what aspects need to be improved, to what extent improvement is required and how such improvement is possible. Decision making processes are designed for effective utilization of performance measures.

3.3 Summary of BPM Process

Q5. Outline the summary of Business Process Management Process.

Ans : **(Oct.-20)**

The Business Process Management (BPM) is not an individual tool or methodology. It is a set of repetitive processes which are dynamic in nature and needs to be carried out on a regular basis, adjusting the various processes, as per the organizational requirements. The following figure shows the BPM process or cycle.

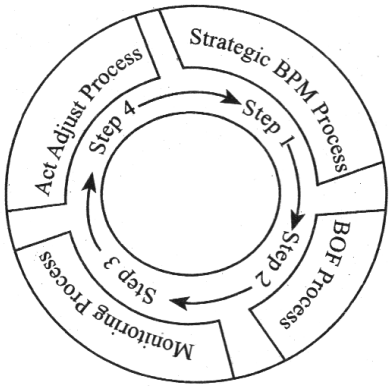


Fig. : BPM Cycle

Step 1 : Strategic BPM Process

This is the first step of the BPM cycle under which the organizations initiates a detailed analysis of the strategic gap that exists in the organization and formulates a detailed strategic planning. It includes the following steps.

(a) The firm conducts a current situation analysis to determine its current position.

(b) A SWOT (Strength, Weakness, Opportunities, Threats) analysis is conducted.

(c) The organizations Critical Success Factors (CFS's) are identified.

(d) The creation of strategic vision and the development of the business level strategy.

(e) The organizations strategic goals and objectives are identified.

Step-2 : BOF Process

Under the Budgeting, Operations and Financial (BOF) process, the strategic goals and objectives of the organization are converted into well defined tactical initiative at the sub process levels.

Operational planning comprises of tactics and initiatives which lead to the desired organizational level. A good operational plan could be either centered around the budget or tactic as per the organizational requirements.

Financial planning and budgeting play a vital role in the overall success of the organization, because the resources of score and needs to be allocated prudently and to extract maximum benefit out of them. If the focus is upon the tactic, then the organization may prepare a tactic based financial budge.

Step-3 : Monitoring Process

Under this step of the BPM cycle, a good and effective performance monitoring mechanism is set in place to ensure that all organizational processes are working towards the desired organizational mission/goal. The various tools techniques and models of performance monitoring clearly showcases the deviations from the desired performance.

A diagnostic control system is a cybernetic system, that is, it intakes inputs, process the information and provide outputs. This output is then compared with pre-established standard performance benchmarks and feedback is provided stating whether there had been any variance from the expected performance.

All organization, irrespective whether they employ BPM system or not, would use some type of diagnostic control system. The key elements of this system are,

(i) Inputs : In the form of actual performance.

(ii) Monitoring Elements : This include any type of tool, technique, metric or measure used by the organization for processing the information. Example: Balanced Score Card, Dashboards, Human Resource System, Financial management system etc.

(iii) Output : This element comprises of any type of end result produced by the monitoring system.

Step-4 : Act Adjust Process

The final step of the BPM cycle involves measuring the business performance by using a variety of tools, techniques, metrics and models and take corrective action.

Q6. What is 'Strategy Gap'? What are the reasons/sources of strategy gap?

Ans :

Strategy gap refers to the difference between the organizations strategic plan and the actual execution of that plan. Thus, this difference between the strategic plan and the actually accomplished strategy refers to strategy gap.

(a) Lack of Communication

The most common reason for strategy gap is lack of communication between the top management and the organizational workforce, with regard to the organizations strategy. According to the researching firm Palladium Group, less than 10% of the workforce is made aware of the strategic plan. Farther the strategic plan which is communicated generally lacks clarity.

(b) Sync between Rewards and Incentives

It is a major reason for strategic gap. According to research Norton (2007) 70% of organizations do not align the incentives of the middle level management to the success of the strategic plan. Many firms wrongly link pay to short term performance.

(c) Resource Crunch

Any grand strategic plan need equally strong funds and resources both in terms of finance and non-financial resources. According to researcher Norton, any good strategic plan, which is not funded sufficiently is doomed to fail. Statistical information had concluded that only 30% of the organizations sufficiently finance their strategic plans.

(d) Time Resource

Researcher Norton had proved that 90% of the top management spend less then 1 hour per day to discuss and focus upon the organization's strategic plan, this intum would lead to strategy gap. Time is a key component/resource that need to be spent on a regular basis upon the organization's strategic plan.

Q7. Describe the steps involved in Harrah's closed-loop marketing model.

Ans :

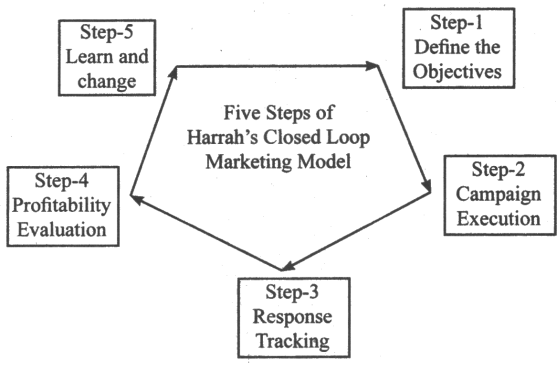
Harrah's closed loop marketing model was presented by Watson and Volonino in the journal "Harnessing customer information for strategic advantage, technical challenges and business solution" (2001). The main aim of this model is to analyze the customer responses and adjust the business offering accordingly to run the business successfully.

Research studies had established that the probability of failure of any business is between 60 percent to 80 percent. For example, the probability of failure for an IT business is 70% whereas, it is 90% in case of pharmaceutical products.

The reasons for the failure of the business may be the result of no proper competitor analysis, insufficient demand analysis, changing customer tasks and preferences and so on.

Harrah's Closed Loop Marketing System Model

As depicted in the figure below the Harrah's model comprises of five sequential steps,



Step-1: Define the Objectives

It is the first step of the Harrah's closed loop marketing model. Under this step, the quantifiable objectives of the marketing campaign are defined or procedures are tested in the form of excepted values utilizing the customers in the test group verses the customers in the control group.

Step-2: Campaign Execution

Under this step the campaign is executed. The campaign is designed and executed in detail such that the optimum mix of offer and message is provided at the right time. The customers are selected after a careful review.

Step-3: Response Tracking

The next step involves tracking of the customers response. Other metrics are also employed to gauge the type of response provided by the customer, what type of incentive is generating a positive response and in how much monetary terms, profitability per customer, the quantity and at what price the customer had purchased the product and so on. This step involves a detailed tracking of the consumers response.

Step-4: Profitability Evaluation

Under this step the marketing campaign is evaluated in terms of the net profitability generated by the campaign and a comparative analysis is also conducted in relation to the response generated by other similar marketing campaign.

Step-5: Learn and Change

Under this final step of Harrah's closed loop marketing model, the information and knowledge gathered from the marketing campaign is analyzed, lessons learned and changes made accordingly, to boost and enhance the profitability and sales of the business organization.

3.4 Performance Measurement

Q8. Define Performance Measurement. What are the characteristics of performance measurement system?

Ans :

Definitions

i) **According to Simons (2002),** performances measurement systems :

Assist managers in tracking the implementations of business strategy by comparing actual results against strategic goals and objectives. A performance measurement system typically comprises systematic methods of setting business goals together with periodic feedback reports that indicate progress against goals.

Performance measurement has been widely defined in the wealth of literature on this subject. Two definitions are used to define the concept in relation to management accounting.

ii) **According to CIMA Official Terminology, 2005.** 'The process of assessing the proficiency with which a reporting entity succeeds, by the economic acquisition of resources and their efficient and effective deployment, in achieving its objectives. Performance measures may be based on non-financial as well as on financial information.'

iii) **According to Interpretability Clearing house Glossary of Terms, 2005.** The process of developing measurable indicators that can be systematically tracked to assess progress made in achieving predetermined goals and using such indicators to assess progress in achieving these goals.'

Characteristic Features of an Effective PMS

(a) A good PMS would facilitate the management to align the top level strategic objectives with the grass, root level initiatives.

(b) It would assist the enterprise to identify potential opportunities and threats before hand and help them to take proactive action accordingly.

(c) It would guide the management to enhance the various organizational processes and procedures as per the need of the changing business environment.

(d) It would assist the planners with forecasting and planning of the organizations strategic long term and short term goals in a timely manner.

(e) It would provide a clear road map about the priority areas of the organization that needs to be financed on a priority basis.

(f) It would facilitate the managers to delineate organizational responsibilities and prepare an objective-reward and compensation plan which is in alignment with the roles and responsibilities of the different jobs.

(g) A important characteristic feature of a good PMS is that it would make it easy for the managers to measure the actual performance of the employees with the expected performance. Thus, make them understand the extent of deviation and initiate corrective action.

Therefore, it may be stated that a holistic PMS is an integral part of any successful business organization. The different types of PMS being used by organizations since the last 40 years include. Activity Based Costing (ABC), Balanced Score Card (BSC) etc.

Q9. Explain the steps in development of an effective PMS.

Ans :

1. The performance measurement system must be integrated with the overall strategy of the business.

2. There must be a system of regular feedback and review of actual results against the original plan and the performance measures themselves.

3. The performance measurement system must be comprehensive. It needs to include the range of factors that contribute to the organisation's success such as competitive performance, quality of service and innovation. This requires a range of financial and nonfinancial indicators.

4. The system must be owned and supported throughout the organisation. The implementation must be top-down so that individuals setting strategy can determine the objectives and develop appropriate top-level measures. These should filter down to the rest of the organisation. Other levels throughout the organisation should set their own measures in consultation with the level above and these must be consistent with the top-level measures.

5. Measures need to be fair and achievable. Where performance measures are used to reward managers' performance, the evaluation should include only the elements they have direct control over.

6. The system and results reporting need to be simple, clear and understandable, particularly to non-finance professionals. There is a need to priorities and focus so that only the key performance indicators for the business in strategic terms are measured.

Q10. What do you understand key performance indicators kpI?

Ans :

There is a difference between a "run of the mill" metric and a "strategically aligned" metric. The term key performance indicator (KPI) is often used to denote the latter. A KPI rep-resents a strategic objective and measures performance against a goal.

Definition

According to Eckerson (2009), KPIs are multidimensional. Loosely translated, this means that KPIs have a variety of distinguishing features, including:

- **Strategy:** KPIs embody a strategic objective.
- **Targets:** KPIs measure performance against specific targets. Targets are defined in strategy, planning, or budget sessions and can take different forms (e.g., achievement targets, reduction targets, absolute targets).
- **Ranges:** Targets have performance ranges (e.g., above, on, or below target).
- **Encodings:** Ranges are encoded in software, enabling the visual display of performance (e.g., green, yellow, red). Encodings can be based on percentages or more complex rules.
- **Time frames:** Targets are assigned time frames by which they must be accomplished. A time frame is often divided into smaller intervals to provide performance mileposts.
- **Benchmarks:** Targets are measured against a baseline or benchmark. The previous year's results often serve as a benchmark, but arbitrary numbers or external benchmarks may also be used.

A distinction is sometimes made between KPIs that are "outcomes" and those that are "drivers." Outcome KPIs—sometimes known as lagging indicators—measure the output of past activity (e.g., revenues). They are often financial in nature, but not always. Driver KPIs—sometimes known as leading indicators or value drivers—measure activities that have a significant impact on outcome KPIs (e.g., sales leads).

In some circles, driver KPIs are sometimes called operational KPIs, which is a bit of an oxymoron (Hatch, 2008). Most organizations collect a wide range of operational metrics. As the name implies, these metrics deal with the operational activities and performance of a company. The following list of examples illustrates the variety of operational areas covered by these metrics:

- **Customer performance:** Metrics for customer satisfaction, speed and accuracy of issue resolution, and customer retention.
- **Service performance:** Metrics for service-call resolution rates, service renewal rates, service level agreements, delivery performance, and return rates.
- **Sales operations:** New pipeline accounts, sales meetings secured, conversion of inquiries to leads, and average call closure time.
- **Sales plan/forecast:** Metrics for price-to-purchase accuracy, purchase order-to-fulfillment ratio, quantity earned, forecast-to-plan ratio, and total closed contracts.

Whether an operational metric is strategic or not depends on the company and its use of the measure. In many instances, these metrics represent critical drivers of strategic outcomes. For instance, Hatch (2008) recalls the case of a mid-tier wine distributor that was being squeezed upstream by the consolidation of suppliers and downstream by the consolidation of retailers.

In response, it decided to focus on four operational measures: on-hand/on-time inventory availability, outstanding "open" order value, net-new accounts, and promotion costs and return on marketing investment. The net result of its efforts

was a 12 percent increase in revenues in 1 year. Obviously, these operational metrics were key drivers. However, as described in the following section, in many cases, companies simply measure what is convenient with minimal consideration as to why the data are being collected. The result is a significant waste of time, effort, and money.

Q11. State the problems with existing performance measurement systems.

Ans :

Problems/Drawbacks with Existing PMS

Following are the major problems/drawbacks with the existing performance system,

(a) Most companies claim to have a good performance measurement system as opposed to performance management system.

(b) A research survey had found that 50% to 90% of the organizations use some variation of the Balance Score Card (BSC) for PMS. However when the managers were questioned as to what constitutes "Balance" in their BSC, they were unclear and confused.

(c) A major limitation is that more than 75% of the organizations focus upon performance measures which are financial in nature. Further their area of focus is primarily internal performance rather than external performance.

(d) Using only financial data for performance measurement suffers from the following limitations,

  (i) It does not tell about future events.

  (ii) No clarity of the processes is provided by it.

  (iii) These focus upon short term achievement rather than long term goals.

(e) Many organizations state that they are tracking more than 200 performance measures at the organizational level. However, this would only complicate and confuse the end users rather than assisting in PMS.

(f) In case of many measures the management lacks direct control over the measure. For example, an organization can only monitor and analyze the extent of customer satisfaction, earnings per share, share price etc. The only areas of performance measure which can be fully controlled are employees actions. However, such control does not have a substantial impact upon the corporate level strategy or business level strategy. These would be effective only when the corporate goals are channeled down until the bottom level of the organizational hierarchy.

Ingredients of a Good Collection of Performance Measure

A good performance measure would include the following ingredients,

(i) The performance measure need to focus upon key factors.

(ii) The measure need to be balanced in outlook, taking into consideration the past, present and future information.

(iii) The measures need to flow in a the top down hierarchy from the top till the bottom level of the organization.

(iv) The performance measure targets need to be formulated, based on scientific research and objectivity rather than arbitrary.

(v) The performance measures need to balance the need of all the stakeholder's such as the shareholders employees, suppliers, and other key stakeholders.

Q12. What are the benefits of performance measurement system ?

Ans :

1. Enhanced Decision Making and Control

It is impossible to make the right decisions without a good understanding of an organization's performance. A multidimensional performance measurement framework means that decision support can be enhanced at all levels of the organization. This ranges from decisions concerning employee performance to strategic decision making at board level.

2. Supported Strategic Planning and Target Setting

The ability to measure performance and progress gives meaning to the process of developing strategic plans and goals. An effective performance measurement system should emphasize the link from the corporate level to management and operational levels. In this way, decision-making and resulting actions and control are in line with strategy.

3. Improved Communication

Involvement in setting targets and results reporting can enhance staff and stakeholder understanding and support of strategies and decisions. It also provides a common language which encourages interdepartmental knowledge sharing.

4. Accountability

Measuring and reporting performance gives decision makers a significant tool to achieve accountability at employee and organizational level. These relationships become clearer when outcomes and outputs are measured by a commonly accepted standard. For public sector organizations, this also includes government and public accountability for the public funds used.

3.5 BPM METHODOLOGIES

Q13. Describe the various methodologies of BPM.

Ans :

There is more to performance measurement than simply keeping score. An effective performance measurement system should help do the following:

- Align top-level strategic objectives and bottom-level initiatives.
- Identify opportunities and problems in a timely fashion.
- Determine priorities and allocate resources based on those priorities.
- Change measurements when the underlying processes and strategies change.
- Delineate responsibilities, understand actual performance relative to responsibilities, and reward and recognize accomplishments.
- Take action to improve processes and procedures when the data warrant it.
- Plan and forecast in a more reliable and timely fashion.

A holistic or systematic performance measurement framework is required to accomplish these aims, as well as others. Over the past 40 or more years, various systems have been proposed. Some of them, such as activity-based costing (ABC) or activity-based management, are financially focused. Others, such as total quality management, are process oriented.

1. Balance Score Card (BSC)

The concept of balanced scorecard (BSC) was primarily developed by Kaplan and Norton in an article in the Harvard Business Review. The balanced score card - measures the drive performance'. It developed great levels of interest for the senior business managers which resulted in the further development. The attention was basically diverted from short-term measurement to creation of growth, learning and value-added services to the consumers.

Several organizations are making of the balanced scorecard as a supporting or underlying structure for making vital process decisions. The performance measurement process has been converted into strategic management system due to the recent development of balanced scorecard.

Balanced scorecard is defined as a conceptual framework which is used for converting the strategic objectives of an organization into a group/collection of performance indicators which are divided into four areas i.e., financial, customer, internal business processes and learning and growth.

Purpose of Balance Scorecard

Kaplan and Norton have found that balanced scorecard is used in many organizations in order to,

(i) Update and clarify the existing strategy.

(ii) Recognize and arrange strategic initiatives.

(iii) Communicate the strategy all over the firm.

(iv) Coordinate both unit as well as individual goals with strategy.

(v) Connect strategic objectives to annual budgets and long term targets.

(vi) Organize periodic performance reviews so as to learn and improve the strategy further.

2. Six Sigma

Six sigma is a process improvement method which was developed by Motorola in 1980's. If a business is capable of producing only 3.4 defects out of every million opportunities, then the business is said to have met the "six sigma criteria".

In other words, six sigma is a statistical control limit given to the process of completing a job or task either completely or partially with a confidence level of 99.9997%.

Features of Six Sigma

1. Six sigma aims at continuous improvement by reducing the cost of production and increasing customer satisfaction and return on investment.
2. One of the features of six sigma is to produce the output defectless or error free.
3. It ensures quality products.
4. Six-sigma's philosophy is "Do it right the first time and every time".
5. Six sigma's principles are applied in manufacturing and service industries.
6. Six sigma is a statistical process control technique which is applied to gain complete confidence in the company's product and services and also the management.
7. It is helpful to solve problems in an organized manner.
8. The six sigma emphasizes on the following areas,
   (a) Independent variable to the process.
   (b) Eliminates the root cause of any problem and tries to prevent it.
   (c) Input is given much importance than output.
   (d) Emphasizes problem and not the cause.
   (e) Concentrates on controlling the problem and not on monitoring.

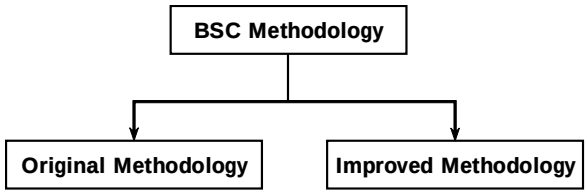
Q14. Explain different methodologies used in Balanced Scorecard.

Ans :

"The balanced scorecard provides managers with the instrumentation they need to navigate to future competitive success".

Balanced scorecard follows two methodologies, namely,

1. Original methodology and
2. Improved methodology.



1. Original Methodology

Initially the balanced scorecards consisted of simple tables which were divided into four parts/ perspectives, namely,

1. Financial
2. Customer
3. Internal business process and
4. Learning and growth.

Balanced scorecards were designed by taking five or six good measures of each perspective. Several authors had been giving different headings to these perspectives and have also recommended few additional perspectives. This earlier/initial balanced scorecard faced a major difficulty of justifying the selection of measures made i.e., "of all the measures one could have chosen, why did one select these".

If in case the user is not confident or sure about the selected measure within the balanced scorecard, then they will be even not confident about the information they provide. Eventhough these BSCs are not so common but still they are being used and designed even today.

Thus, the early-style balanced scorecards are difficult to be designed in such a manner that builds confidence that these BSCs are designed well. Due to this reason, most of the BSCs are abandoned as soon as they are accomplished.

2. Improved/Enhanced Methodology

This method was started in the mid 1990's. Which was an enhanced version of the early-style balanced scorecard. This method selects 'die measures on the basis of "strategic objectives" set and plots them on the "strategic linkage model" or "strategy map". This model has divided the strategic objectives across the same set of "perspectives".

The managers identifies five to six goals of each perspective and then shows interconnection between these goals by plotting causal links on the map/diagram. Balanced scorecard is then devised by selecting suitable measures for each objective by reaching at some conclusion about each objective.

This type of balanced scorecard was used from 1996 onwards which provided greater contextual justification for the selected measures. It was easy also.

During 1990's again there were few changes made in the design approach. The "2nd generation" balanced scorecard faced a problem of plotting causal links between twenty or medium term strategic goals, but still it remained the abstract activity.

The third generation balanced scorecard represented the present state-of-the-art in scorecard design.

Kaplan and Norton has concluded that the companies use balanced scorecard for the following purpose,

- Executing strategy
- Clarifies strategy and makes it functional
- Identifies and aligns strategic initiatives
- Associating budget with strategy.
- Improving strategy by conducting timely strategic performance reviews.

In 1997, it was found out by Kurtzman that 64% companies who were questioned started using the balanced scorecard to measure performance.

Government agencies, military units, business units, corporations, nonprofit organizations and schools implement BSC. Different organizations use balanced scorecard differently. Many theorists recommend not to adapt any organization's BSC, as the advantages of BSC can be derived from its implementation.

Most of the 1st generation or the early-state BSC failed due to this reason only that they adapted other organization's BSC, which were prepared by the consultants who had inadequate or poor knowledge about the organization and management's responsibility.

Q15. Explain the importance of Balanced Scorecard ?

Ans :

It highlights the importance of non-financial drivers of performance, and clarifies vision throughout the organization. Specifically the benefits are :

1. Greater Participation

After constructing the Balanced Scorecard leaders cascade strategy down to business units, divisions and support functions. In doing so, top management does not dictate strategy to operating units. Instead it recognizes that operating units have a better 'feel' for local conditions, operating technologies, and competition. So it encourages the latter to define their own strategy so as to dovetail with the organization's overall strategy.

2. Improved Communication

Top management explains the Balanced Scorecard and the strategy map through newsletters, brochures, bulletin boards, meetings, training, and education programs. This is reinforced by the top management's conviction in the efficacy of the Balanced Scorecard. This conviction gets translated into their personal behavior and their commitment is reflected in the initiatives taken to accomplish their strategy.

3. Improved Understanding of Mission

Departments performing line functions (such as the production department in a manufacturing concern) as well as those in staff functions such as human resources, purchase, finance and internal audit, are asked to define what services ought to be provided, as also the quality of these services and their cost. They examine and understand the interrelationships between each staff function and line functions. In doing so they realize what each function (whether line or staff), contributes to overall organizational success. The concept of partnerships with each other and with the corporate parent is grasped, enabling the emergence of corporate-level synergies.

4. Empowerment

Employees get a clearer understanding of what the organization wishes to accomplish, and how they can contribute to these accomplishments. They realize that their work can and does make a difference to the organization, and this increases their intrinsic motivation. They come to work with enthusiasm, creativity and initiatives, actively searching for better ways in which they can help the organization succeed. Personal objectives arc linked to team objectives. Typically a new incentive plan is drawn up that enables employees to benefit financially when targets for strategic measures are achieved.

5. Style of management

Chief executive officers (CEOs) understand that they cannot implement corporate strategy alone. They need contribution, cooperation and ideas from everyone in the organization. Persons far from corporate headquarters whether they are employees, vendors, or distributors can suggest new ways of doing business. So CEOs need to alter their style of management from autocratic to participative. They discard top down direction and embrace top down communication, helping the 'partners' learn how they can contribute to successful implementation of corporate strategy.

6. Improved Training

Top management is aware that employee skills enhance internal processes. The strategy map reveals the strategic chain of cause and effect relationships that link greater investment in improving employee skills to improved financial performance. Top management realizes the importance of skill upgradation and the specific skills that should be enhanced. The Balanced Scorecard shows people how improving of capabilities leads to the organization achieving it's long term financial goals. The organization implements the learning and growth objectives, by focusing on sustained investment in training.

7. Selection of Suppliers

In choosing strategic partners, the organization pays as much attention to low cost supplies as quality of supplies, lead time, on time delivery performance of suppliers, defect free supplies, and whether suppliers are electronically connected to the organization. This last factor allows the online ordering of supplies as well as online payment. Attention to cost of supplies alone is dangerous. Low cost suppliers may turn out to be extremely high cost if they fail to deliver supplies on schedule, or the deliveries do not conform to quality standards. The organization will have to spend considerable time, effort and cost, in monitoring quality, arranging for replacement of defective goods, sourcing supplies from other vendors at short notice, and stocking up of inventory when the supply is known to be erratic. By incorporating measures for superior supplier

relationships into the Balanced Scorecard, employees understand the value of forging long term relationships with key suppliers. This recognition generates initiatives from the employees and feedback on the most critical elements of the supplier relationship.

8. Enhancement of value in the community

Organizations cannot afford to ignore the community within which they operate. They may wish to move beyond mere compliance, setting industry standards in areas such as corporate governance, community concern, and environmental protection. One chemical company Stated - 'Our strategy is to be seen by the community within which Mr operate, ax not only a law abiding corporate citizen, but as an outstanding. corporate citizen, measured both environmentally and by creating well paying, safe, and productive jobs. If regulations are tightened, though other firms may get affected, we expect to have earned the right to continue operations.' In recognition of this, environmental and community performance was a key part of its strategy. Such recognition enables an organization to successfully operate in markets that differ in terms of cultural, linguistic, and economic bases and geographic locations.-

Thus the balanced scorecard provides a framework to align the organization, focuses teams and individuals on strategic priorities, provides a structure for multiple initiatives, drives this capital and resource allocation process, integrates strategic management across the organization, and integrates the entire supply chain.

Q16. Write down about the criticisms of Balanced Scorecard.

Ans :

BSC inspite of being successful suffers from few criticisms. The first criticism is that the scores are not on the basis of any proven economic or financial theory and thus has no basis in the decision making. The balanced scorecard is a subjective process and does not make any provision to evaluate quantities (i.e., risk and economic) in an economically sound way/manner.

BSC was also criticized for not providing a bottom line score or a unified view with clear recommendations. BSC was criticized for simply being a metric.

Few people have also criticised BSC by viewing that BSC receives/gets positive feedback due to placebo effect, due to the absence of empirical studies relating to the BSC's use in making better decisions for enhancing the financial performance of the companies.

The balanced scorecard had always attracted criticisms from different sources. Most of these criticisms were from the academic community who hated the empirical nature of the structure. Kaplan and Norton failed to include any citation of prior art in their initial topics.

Few criticisms focused upon the technical defects in the methods and design of the original balanced scorecard and has over time driven the evolution of the device through its several generations.

Pundits and consultants have criticised that the balanced scorecard does not provide a bottom line score or a unified view with clear recommendations. It is simply of list of metrics.

Q17. What are the objectives of Six Sigma ?

Ans :

Six Sigma encompasses a set of practices designed to identify, analyze and remedy causes of defect within a process or product. Defect is not simply a problem with a product; it is anything that causes a company to operate less cost-effectively, at a slower pace and with an end result of lowered customer satisfaction. Acquiring and constantly pleasing customers is Six Sigma's ultimate goal.

1. Overall Business Improvement

Six Sigma methodology focuses on business improvement. Beyond reducing the number of defects present in any given number of products, a business employing Six Sigma methods must seek improvement through any means available. That means identifying and remedying problems wherever they occur. Six Sigma calls anything that damages business functionality in a way that increases

defects, raises costs, slows productivity or reduces customer satisfaction a source of pain. The elimination or remediation of these sources of pain leads to overall business improvement.

2. Remedy Defects/Variability

Any business seeking improved numbers must reduce the number of defective products or services it produces. Defective products can irrevocably harm customer satisfaction levels, as each customer ending up with a defective product becomes a potential lost costumer—and because the displeased customer will tend to pass the word about this defective product along. Then you've got to fix the defects, which can increase research and production costs dramatically.

3. Reduce Costs

Reduced costs equal increased profits. A company implementing Six Sigma principles has to look to reduce costs wherever it possibly can—without reducing quality. Cost reduction potential exists throughout a company. Acquire cheaper raw materials of equal or comparable value; reduce transportation costs via alternate shipping methods; streamline production and quality control processes with automation or improved equipment technology; cut personnel costs with outsourcing, downsizing or other methods; or reduce rent payments by moving production or sales facilities to different locations.

Even the adoption of greener business practices can lead to reduced costs, as powered-down electronics, recycled paper and reduced wastage can have significant impact. No change is too small to consider.

4. Improve Cycle Time

Any reduction in the amount of time it takes to produce a product or perform a service means money saved, both in maintenance costs and personnel wages. Additionally, customer satisfaction improves when both retailers and end users receive products sooner than expected. The company that can get a product to its customer faster may win her business, regardless of questions of quality or cost. There's a reason fast food was the definitive concept in food service during the 20th century.

5. Increase Customer Satisfaction

The sources of pain that Six Sigma methodologies seek to remedy interrelate. Customer satisfaction depends upon successful resolution of all Six Sigma's other objectives. But customer satisfaction is an objective all its own. Every aspect of a business' self-representation, from marketing strategies to sales personnel performance, can have a positive or negative affect on customer satisfaction. Seek positive customer response to these self-representations, and customer satisfaction will improve.

Q18. Write about DMAIC performance model and lean six sigma.

Ans :

DMAIC Method

1. D - Define Phase

The define phase helps in defining the problem clearly as and when they occur. This is the first step in six sigma methodology. Define phase identifies those factors which are to be measured, analyzed, improved and controlled for greater revenue and needs improvement to carry out the project successfully.

Define phase should be stated clearly so that the problem can be solved effectively.

2. M - Measure Phase

Measure phase is used to identify the critical internal processes which directly affects the CTQ measurements. Once the problem is defined, then the measure phase starts wherein it measures or evaluates the problems in the process which affects the CTQ standard.

The important factors of six sigma are, defect, variation, Critical-to-Quality (CTQ), process capability and Design for Six Sigma (DFSS).

After measuring the defects the impact of these defects should be measured because defects are measurable factors of a process. Once the defects are identified and measured the cost involved in them can be saved.

The defects in the process are analyzed by black belt which affects the CTQ characteristic. The black-belt uses measurement systems analysis which consist of gauge studies and a complete evaluation of the capability of the process.

Gauge

The primary objective of gauge is repeatability and reproducibility study (gauge R&R). It is used to know whether measurement systems are functioning properly or not.

(i) It helps in finding out the defects in the defined measurements and in the initial stages itself and helps in correcting them at the right time.

(ii) It provides information which forms the basis for determining the credibility and the areas to minimize defects. Under the following four criteria gauge R&R repeats measurements,

(a) Accuracy: To know how accurate the measurement is.

(b) Repeatability: If the same item is measured many times will it give the same result?

(c) Reproducibility: If the same items are measured by other people/other pieces of equipment, what will be the result?

(d) Stability: To know the stability of accuracy, repeatability and reproducibility.

3. A - Analyze Phase

Under this phase the defects which are affecting the Critical-to-Quality (CTQ) are analyzed. Hypothesis and statistical test are used to determine the factors which are critical to the outcome. Analyze phase acts as a cycle as it passes through various steps.

Step 1: Develop hypothesis about the reason(s).

Step 2: Analyze process and/or data.

Step 3: If the hypothesis is correct, add cause(s) to the list of vital few. If the hypothesis is incorrect refine it and again go to step 2 or reject it and go to step (1).

Hypothesis testing is used to compute the probability of those identified factors which have an impact on critical- to-quality outcomes. Soon after this, statistical conclusions are made for arriving at the corrective actions.

4. I - Improve Phase

Once the factors which affect the variables are measured and analyzed then improve phase begins to rectify the defects. In improve phase, all those factors are improved which affect the critical-to-quality outcomes. Thus, the maximum acceptable range of each variable is identified and variations are measured using the measurement system.

5. C – Control Phase

Control phase is used to maintain continuous improvement and in some cases control phase does not exist because mostly the problem gets eliminated in the above phases. Control phase ensures quality, productivity and improvement in the processes continuously. The scope of control mechanisms is wider.

Lean Six Sigma

In recent years organizations had been working towards integrating six sigma system with manufacturing this alingment is popularly known as lean manufacturing/lean production.

Lean Production

Lean production is also known as lean manufacturing. It is a systematic approach that identifies and eliminates the waste through continuous improvement by flowing the product as per the customer demand. Lean production aims at reducing the time, resources and efforts without making any compromise on the quality of the product.

The concept of lean corresponds to "no-waste", wherein waste refers to anything that is added to the cost of the product but not to the value of the product. There exist three types of works in any process. They are,

(a) Value-Added: These are the process steps that produce value for the customer and for which the customer is ready to pay. In short, these are essential to meet the requirements of customer.

(b) Non-Value Added: These process steps are not essentially needed to produce value for the customer but are required for other reasons such as fulfilling company mandates, legal requirements and regulatory requirements.

(c) Waste: The process steps that produce no value for the customer but consumes resources.

The company 'Toyota' originated the concept of lean production. It identified eight types of waste,

1. **Defects:** Time and money wasted for determining mistakes and fixing the defects.
2. **Over-production:** Production more than what is needed.
3. **Waiting:** Work delayed due to waiting machines, people and material.
4. **Transportation:** Time and efforts involved in movement of materials, products, information and people.
5. **Motion:** Unnecessary work-hand offs or transfer.
6. **Extra-processing:** Doing more work on improving the quality than needed.
7. **Inventory:** Holding additional material or the resources that are not required.

Q19. Compare and contrast Six Sigma and Balance Score Card.

Ans :

Sl. No.	Six Sigma	Sl.No.	Balance Score Card
1.	Six sigma possesses higher quality of products or services which holds the probability of only 3.4 defects per million opportunities (DPMO).	1.	Balance score card relates to the strategic goals and objectives of the business organization.
2.	Six sigma is a problem solving methodology which helps in reducing defects and its related costs.	2.	Balance scorecard has been designed and developed to balance the organizations strategic goals. It focuses towards vision and values.
3.	Six sigma considers the process management, improvement, measurement as daily activities for improving the process management.	3.	Balance score card considers the strategic management system measures.
4.	Six sigma not only produces defect-free goods or services or continuous improvement but also helps in reinvention/ reinnovation of the business.	4.	Balance score card not only identifies measurements around the organizations vision and value, but also highlights targets for each measure.
5.	Six sigma's implementation in the company brings standard improvement in the performance. such as organizational vision, strategy, plan, feedback etc.	5.	Balance score card implementation helps with identifying measures for key management processes
6.	Six sigma uses 'critical-to-quality'aspect which is most crucial from the customer's point of view.	6.	Balance score card uses elements such as organization's internal operations, strategic goals, customers and so on.
7.	Six sigma also solves the customer's problems effectively.	7.	Balance scorecard helps towards balancing strategic organizational elements, such as customers, internal operations etc.
8.	In six sigma, Champions, black belts, master black belts and green belts are used for successful implementation of six sigma projects.	8.	Balance score card employs strategy maps, strategic themes and other tools for successful implementation of its mission in the organization.
9.	The ISO, CMM and six sigma are integrated together in the system in order to improve the efficiency of the services which helps in solving the problems.	9.	Balance score card is an integration of performance management elements and t managemen methodologies to facilitate the enterprise to achieve targets related to its finances internal processes, customers and learning and development.
10.	Six sigma follows, "do it right the first time" philosophy.	10.	Balance score card follows the goal of translating vision and strategy into clear objectives, measures, targets and initiatives.

3.6 BPM ARCHITECTURE

Q20. Elucidate the Architecture of Business Performance Measurement.

Ans : **(June-19)**

Business Performance Measurement as an umbrella term covering the processes, methodologies, metrics, and technologies used by enterprises to measure, mon-itor, and manage business performance.

BPM Architecture

The term **system architecture** refers to both the logical and physical design of a system. The logical design entails the functional elements of a system and their interactions. The physical design specifies how the logical design is actually implemented and deployed across a specific set of technologies such as Web browsers, application servers, communication protocols, databases, and the like. From a physical standpoint, any particular. BPM solution or implementation is likely to be quite complex. From a logical standpoint, they are usually quite simple.

Elements of BPM Architecture

Logically speaking, a BPM system consists of three basic parts or layers (see Figure). Included are:

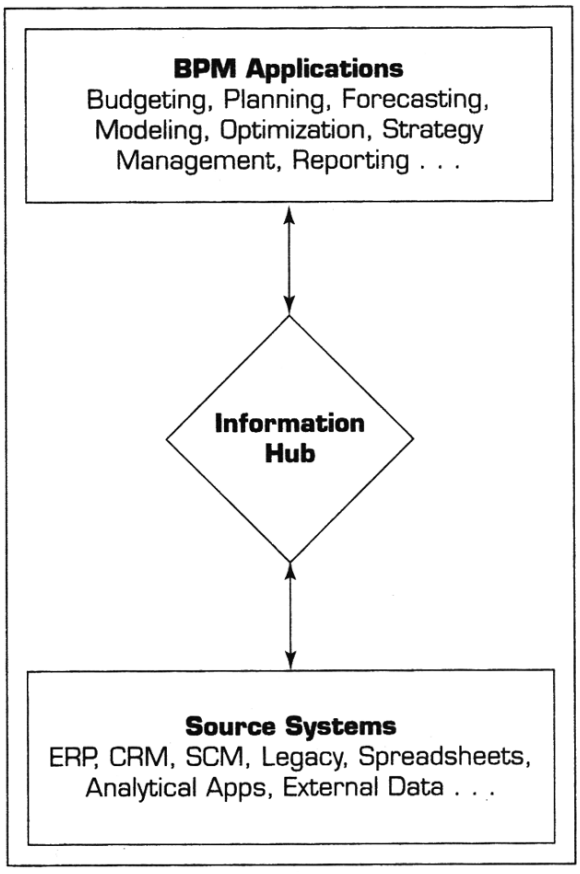


Fig. : BPM Logical System Architecture

A) BPM applications

This layer supports the BPM processes used to transform user interactions and source data into budgets, plans, forecasts, reports, analysis, and the like. The particular applications used from one BPM implementation to the next will vary from organization to organization, depending on their specific needs and strategic focus.

Any BPM solution should be flexible and extensible enough to allow an organization to find its own path, including decisions about which applications to include and when to roll them out. Practically speaking, however, there are some BPM applications that are used quite frequently. These applications are discussed momentarily.

B) Information Hub

Most BPM systems require data and information from a variety of source systems (e.g., ERP or CRM systems). The data and information can be accessed in a number of ways. However, in a well-designed BPM system, the data from these systems are usually mapped to and stored at a central location, typically a data warehouse or data mart.

C) Data Source Systems

This layer represents all of the data sources containing information fed into the BPM information hub. For most large enterprises, this will include financial and other operational data from a variety of enterprise systems. Complete solutions will also access key external information, such as industry trends and competitor intelligence, to provide deeper context and insight into company performance. Rarely are source data accessed directly by the BPM applications. Typically, an extraction, transformation, and load (ETL) application, an enterprise application integration system, or Web Services are used to move or connect the data to the information hub.

3.7 BPM APPLICATIONS

Q21. Explain the various applications of business performance measurement.

Ans : **(June-18)**

In BPM, a wide variety of applications are needed to cover the closed-loop processes running from strategic planning to operational planning and budgeting to monitoring to adjustment and action. Despite the breadth of the processes, the industry analyst group Gartner contends that the majority of the processes can be handled by the following applications.

1. Strategy management

Strategy management applications provide a packaged approach to support strategic planning, modeling, and monitoring to improve corporate performance, accelerate management decision making, and facilitate collaboration. These solutions are usually tied to strategy maps or methodologies such as the balanced scorecard. Strategy management can encompass capabilities for the following:

- Creation and evaluation of high-level business plans using a "base case plus" or "initiative-based" approach, along with scenario modeling.
- Initiative/goal management using project-management-like tools to enable responsible managers to execute specific tasks related to a strategy.
- Scorecards and strategy maps to record strategies, objectives, and tasks; measure performance; and provide a collaborative environment for effective, enterprise wide communication.
- Dashboards (or cockpits) to aggregate and display metrics and KPIs so they can be examined at a glance before further exploration via additional BI tools.

BPM suites should, at the very least, provide dashboard capabilities to help display performance information in a way that is easily understood by users. However, more-

sophisticated organizations are implementing strategy maps (linked frameworks of KPIs) using scorecard software to link BPM to other aspects of per-formance management. Strategy management is, therefore, becoming an increasingly important aspect of BPM suites.

2. Budgeting, planning and forecasting

These applications support the develop-ment of all aspects of budgets, plans, and forecasts. They encompass short-term financially focused budgets, longer-term plans, and high-level strategic plans. These applications should deliver workflow capabilities to manage budget/plan creation, submission, and approval, and they provide the facility to dynamically create fore-casts and scenarios. They should also support the development of an enterprise- wide planning model that links operational plans to financial budgets. In addition, they must be capable of sharing data with domain-specific applications such as sup-ply chain planning.

3. Financial consolidation

This type of application enables organizations to rec-oncile, consolidate, summarize, and aggregate financial data based on different accounting standards and federal regulations. These applications are a fundamen-tal part of BPM because they create the audited, enterprise-level view of financial information that must be shared with other BPM applications to analyze variance from targets.

4. Profitability modeling and optimization

These applications include activity-based costing (ABC) applications that determine and allocate costs at a highly granular level and activity-based management applications that provide capabilities to enable users to model the impact on profitability of different cost and resource allocation strategies. Some applications have moved beyond the traditional ABC focus to enable revenue to be allocated in addition to costs for model packaging, bundling, pricing, and channel strategies.

5. Financial, statutory, and management reporting

BPM applications require specialized reporting tools that can format output as structured financial state-ments. They may also need to support specific generally accepted accounting principles (GAAP) presentation rules, such as U.S. GAAP or international financial reporting standards. They also include visualization techniques that are specifically designed to support the analysis of variance from budgets or targets, such as hyperbolic trees.

Q22. What changes are observed in the BPM market in the last 3 to 4 years? What are the categories in gartners magic quadrant? Who are some of the vendors in the BPM leaders quadrant?

Ans :

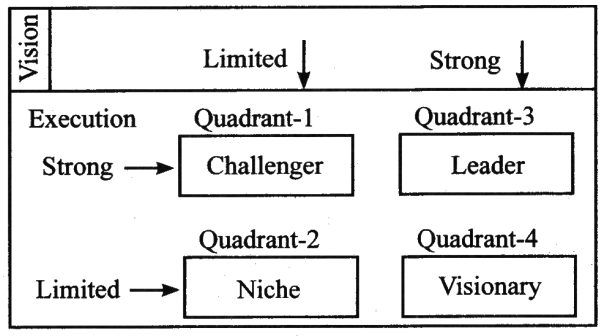
The BPM market is highly dynamic and competitive in nature. It comprises of software enterprises which provide core BPM applications such as budgeting, planning, forecasting, score carding etc. There had been many changes observed in the BPM market during the past 3 to 4 years as discussed below,

1. The biggest change was the consolidation of the BPM market vendors. For example, previously there were multiple vendors such as hyperion, SAS, cognogs, etc. However, these smaller to BPM vendor had been acquired by bigger software firms, such as,
   (a) Cognos vendor was acquired by IBM.
   (b) Hyperion vendor was acquired by oracle incorporation.
2. According to gartner, the value of commercial BPM market had increased to 2 billion dollar by the year 2007 and the estimates predicted a very high increase over the next few years.
3. The main driver for this growth is that most firms are replacing spreadsheet based applications with more advanced BPM softwares.
4. There had been a very big consolidation in the BPM market, for instance. IBM, orace, SAP and SAS together account for 78% of the total global BPM market.

5. Gartner had formulated a magic quadrant for positioning BPM companies based upon various parameters.

Categories in Gartners Magic Quadrant

Gartner had categorized BPM vendors based upon their ability to successfully execute tasks and the strength of the vendor organizations vision. Figure Gamters magic quadrant.



Thus, based upon the above shown gartners magic quadrant, vendor firms would be categorized as challenger, niche, leader or visionary.

Some of the BPM vendors which fit under the leader category are oracle incorporations, hyperion, SAP business objects and IBM's cognos. The reason for these vendors to be categorized as leaders is the sheer size of these major vendors.

3.8 Performance Dash Boards

Q23. Explain briefly about Performance Dash Boards.

Ans : **(Oct.-20)**

A performance dashboard is a business management tool that will allow a business to keep track of all aspects of their daily activity.

For example, the sales performance dashboard is able to track your sales on the east coast and compare them with your sales on the west coast.

When one of those sales unexpectedly rises or falls, you are able to investigate the reasons behind this and make the necessary changes to keep your business profitable. By converting your businesses numerical data into graphical/visual representations, performance dashboards allow the users to gain better insight to the more intricate parts of their business that they previously did not understand.

Various Types of Performance Dashboard

Performance dashboards should allow businesses to do the following:

1. Monitor critical business processes and activities using metrics of business performance that trigger alerts when potential problems arise and when goals are met.
2. Analyze the root cause of problems by exploring relevant and timely information from multiple perspectives and at various levels of detail.
3. Manage people and processes to improve decisions, optimize performance and steer the organization in the right direction.

Q24. What are the differences between Dashboards Vs. Scorecards ?

Ans :

In most BPM journals and magazines, dashboards and scorecards are used interchangeably because both consist of similar features. However, the main point of difference between the two is that, dashboards are used by managers for making operational and tactical decisions, sometimes on a weekly, daily or hourly basis, whereas scorecards are employed to monitor the, strategic alignment and to measure the extent of success achieved with regard to the strategic objectives and the targets.

Dashboard Design

Dashboards are not a new concept. Their roots can be traced at least to the EIS of the 1980s. Today, dashboards are ubiquitous. For example, a few years back, Forrester Research estimated that over 40 percent of the largest 2,000 companies in the world use the technology.

According to Eckerson (2006), a well-known expert on BI in general and dashboards in particular, the most distinctive feature of a dashboard is its three layers of information:

1. **Monitoring:** Graphical, abstracted data to monitor key performance metrics.
2. **Analysis:** Summarized dimensional data to analyze the root cause of problems.
3. **Management:** Detailed operational data that identify what actions to take to resolve a problem.

Q25. Explain the common characteristics of Dashboards and Scorecards.

Ans :

Although performance dashboards and standard performance scorecards differ, they do share some of the same characteristics.

- First, they both fit within the larger BPM or per-formance measurement system. This means that their underlying architecture is the BI or performance management architecture of the larger system.
- Second, all well-designed dashboards and scorecards possess the following characteristics:
- They use visual components (e.g., charts, performance bars, sparklines, gauges, meters, stoplights) to highlight, at a glance, the data and exceptions that require action.
- They are transparent to the user, meaning that they require minimal training and are extremely easy to use.
- They combine data from a variety of systems into a single, summarized, unified view of the business.
- They enable drill-down or drill-through to underlying data sources or reports, pro-viding more detail about the underlying comparative and evaluative context.
- They present a dynamic, real-world view with timely data refreshes, enabling the end user to stay up to date with any recent changes in the business.
- They require little, if any, customized coding to implement, deploy, and maintain.

Short Question and Answers

1. What is Business Performance ?

Ans :

Business performance evaluates the overall organizational performance in significant areas like customer satisfaction, service and product performance, operational and financial performance, cost leadership, human resources, human capital. Performance of the company is also assessed in relation to that of the rival companies.

An environment from which business takes inputs, processes them and produces its outputs is usually referred as business environment. An environment comprises various external factors which influence business performance. They are political and legal factors, social and cultural factors, technological factors, global factors and natural factors.

2. Define business performance measurement.

Ans :

BPM may be defined as "Different business tools, techniques, methodologies and metrics employed by an organization to measure, monitor and manage the performance of the business enterprise". There is always a confusion between the term Performance Management and Business Performance Management (BPM). However both are interlinked with each other and in most cases used as synonyms.

According to Dr. T.V Rao, "performance management involves, thinking through various, facts of performance, identifying critical dimensions performance, planning, reviewing, developing and enhancing performance and related competencies".

According to Reward Management Associates "performance management may be defined as managerial process which consist of planning performance, managing performance through observation and feedback, appraising performance and rewarding performance".

In simple words we any define performance management as "a continuous process of identifying, measuring and developing the performance levels of individual is and aligning it with the strategic goals of the organization.

3. Components of Business Performance Measurements

Ans :

According to business analytics researcher colbert (2009) following are the major components of BPM,

(i) A key component of BPM is a set of closely linked management and analytical processes. These processes are fully integrated with technology. The primary focus area of these close looped management processes is to address and assess the financial and operational activities of the organization.

(ii) BPM components include all the business tools and techniques which are used for defining, formulating and measuring the strategic goals of the organization. These tools and techniques are used for not only formulating strategic goals, but also for measuring the actual organizational performance with the previously set goals.

(iii) Other important components of BPM includes tools and techniques of financial and operational planning, metrics employed for the measurement of key performance indications and so on.

Thus, it may be stated that business organizations operate in a highly competitive environment. Thus, they need to adapt, adjust and align themselves on a continuous basis as per the changing business environment.

Researcher Axson (2017) had stated that "An organization's performance management processes are the principle mechanism for assessing the impact of change and turning the business in order to survive and prosper.

4. What is 'Strategy Gap.

Ans :

Strategy gap refers to the difference between the organizations strategic plan and the actual execution of that plan. Thus, this difference between the strategic plan and the actually accomplished strategy refers to strategy gap.

5. Define Performance Measurement.

Ans :

According to Simons (2002), performances measurement systems :

Assist managers in tracking the implementations of business strategy by comparing actual results against strategic goals and objectives. A performance measurement system typically comprises systematic methods of setting business goals together with periodic feedback reports that indicate progress against goals.

6. Balance Score Card (BSC)

Ans :

The concept of balanced scorecard (BSC) was primarily developed by Kaplan and Norton in an article in the Harvard Business Review. The balanced score card - measures the drive performance'. It developed great levels of interest for the senior business managers which resulted in the further development. The attention was basically diverted from short-term measurement to creation of growth, learning and value-added services to the consumers.

Several organizations are making of the balanced scorecard as a supporting or underlying structure for making vital process decisions. The performance measurement process has been converted into strategic management system due to the recent development of balanced scorecard.

Balanced scorecard is defined as a conceptual framework which is used for converting the strategic objectives of an organization into a group/collection of performance indicators which are divided into four areas i.e., financial, customer, internal business processes and learning and growth.

7. Features of Six Sigma

Ans :

i) Six sigma aims at continuous improvement by reducing the cost of production and increasing customer satisfaction and return on investment.

ii) One of the features of six sigma is to produce the output defectless or error free.

iii) It ensures quality products.

iv) Six-sigma's philosophy is "Do it right the first time and every time".

v) Six sigma's principles are applied in manufacturing and service industries.

vi) Six sigma is a statistical process control technique which is applied to gain complete confidence in the company's product and services and also the management.

vii) It is helpful to solve problems in an organized manner.

viii) The six sigma emphasizes on the following areas,

(a) Independent variable to the process.

(b) Eliminates the root cause of any problem and tries to prevent it.

(c) Input is given much importance than output.

(d) Emphasizes problem and not the cause.

(e) Concentrates on controlling the problem and not on monitoring.

8. Objectives of Six Sigma

Ans :

i) Overall Business Improvement

Six Sigma methodology focuses on business improvement. Beyond reducing the number of defects present in any given number of products, a business employing Six Sigma methods must seek improvement through any means available. That means identifying and remedying problems wherever they occur. Six Sigma calls anything that damages business functionality in a way that increases

defects, raises costs, slows productivity or reduces customer satisfaction a source of pain. The elimination or remediation of these sources of pain leads to overall business improvement.

ii) Remedy Defects/Variability

Any business seeking improved numbers must reduce the number of defective products or services it produces. Defective products can irrevocably harm customer satisfaction levels, as each customer ending up with a defective product becomes a potential lost costumer—and because the displeased customer will tend to pass the word about this defective product along. Then you've got to fix the defects, which can increase research and production costs dramatically.

iii) Reduce Costs

Reduced costs equal increased profits. A company implementing Six Sigma principles has to look to reduce costs wherever it possibly can—without reducing quality. Cost reduction potential exists throughout a company. Acquire cheaper raw materials of equal or comparable value; reduce transportation costs via alternate shipping methods; streamline production and quality control processes with automation or improved equipment technology; cut personnel costs with outsourcing, downsizing or other methods; or reduce rent payments by moving production or sales facilities to different locations.

9. Business Performance Measurement Architecture.

Ans :

Business Performance Measurement as an umbrella term covering the processes, methodologies, metrics, and technologies used by enterprises to measure, monitor, and manage business performance.

BPM Architecture

The term **system architecture** refers to both the logical and physical design of a system. The logical design entails the functional elements of a system and their interactions. The physical design specifies how the logical design is actually implemented and deployed across a specific set of technologies such as Web browsers, application servers, communication protocols, databases, and the like. From a physical standpoint, any particular. BPM solution or implementation is likely to be quite complex. From a logical standpoint, they are usually quite simple.

10. Performance Dash Boards.

Ans :

A performance dashboard is a business management tool that will allow a business to keep track of all aspects of their daily activity.

For example, the sales performance dashboard is able to track your sales on the east coast and compare them with your sales on the west coast.

When one of those sales unexpectedly rises or falls, you are able to investigate the reasons behind this and make the necessary changes to keep your business profitable. By converting your businesses numerical data into graphical/visual representations, performance dashboards allow the users to gain better insight to the more intricate parts of their business that they previously did not understand.

Choose the Correct Answer

1. BPM stands for [c]
   (a) Business Performance Management (b) Business Poor Management
   (c) Business Performance Measurement (d) All
2. Component of BPM Process. [a]
   (a) Strategic (b) API
   (c) TCI (d) All
3. Nature of BPM [a]
   (a) Narrow (b) Dynamic
   (c) (a) and (b) (d) All
4. Methodology of BPM [a]
   (a) BCG (b) ACG
   (c) BSI (d) All
5. Element of BPM Architecture. [b]
   (a) Strategy Management (b) BPM Application
   (c) Financial Strategy (d) All
6. BSC Stands for [a]
   (a) Balanced Score Card (b) Benefits Score Card
   (c) Bombay Score Card (d) All
7. Feature of six sigma [a]
   (a) Ensure quality products (b) High cost
   (c) Low performance (d) All
8. CPM stands for [a]
   (a) Corporate performance management (b) Corporate poor management
   (c) Corporate good management (d) All
9. Lean manufacturing is also known as [a]
   (a) Lean production (b) Lean quality
   (c) Lean cost (d) All
10. ERP stands for [a]
   (a) Enterprise resource planning (b) Enterprise route planning
   (c) Enterprise routing planning (d) All

Fill in the blanks

1. KPI stands for _____.
2. OLAP stand for _____.
3. ABC stand for _____.
4. Style of BI _____.
5. Methodology of BPM _____.
6. Application of BPM _____.
7. The example of data sources system _____.
8. Features of balance score card _____.
9. BI stands for _____.
10. Component of BPM cycle.

ANSWER

1. Key performance indicator
2. Online analytical process
3. Activity Based Costing
4. Cube analysis
5. Six Sigma
6. Financial consolidation
7. ERP
8. Defects per million opportunities
9. Business Intelligence
10. Monitoring process

UNIT IV

Business Analytics and Data Visualization

(a) Business Analytics - Definitions, Tools and techniques of BA, Advanced Business Analytics Business Analytics and Web, Usage, Benefits and Success of Business Analytics.

(b) Data Visualization: Definition, New Direction in Data Visualization, GIS, GIS vs GPS

4.1 Business Analytics

4.1.1 Definition

Q1. Define is business analytics ?

Ans : **(June-19)**

Business analytics is the practice of iterative, methodical exploration of an organization's data, with an emphasis on statistical analysis. Business analytics is used by companies committed to data-driven decision-making.

Definitions of business analytics

i) **According to Schaer** (2018) - "allows your business to make predictive analysis rather than reacting to changes in data".

ii) **According to Gabelli School of Business** (2018)- "involves applying models, methods, and tools to data, producing insights that lead to informed business decisions".

iii) **According to Wells** (2008) - "the application of logic and mental processes to find meaning in data".

iv) **According to Lynda** (2018) - "allows us to learn from the past and make better predictions for the future".

Business analytics (BA) refers to the skills, technologies, practices for continuous iterative exploration and investigation of past business performance to gain insight and drive business planning. Business analytics focuses on developing new insights and understanding of business performance based on data and statistical methods. In contrast, business intelligence traditionally focuses on using a consistent set of metrics to both measure past performance and guide business planning, which is also based on data and statistical methods. Business analytics makes extensive use of statistical analysis, including explanatory and predictive modeling, and fact-based management to drive decision making. It is therefore closely related to management science. Analytics may be used as input for human decisions or may drive fully automated decisions. Business intelligence is querying, reporting, online analytical processing (OLAP), and "alerts."

In other words, querying, reporting, OLAP, and alert tools can answer questions such as what happened, how many, how often, where the problem is, and what actions are needed. Business analytics can answer questions like why is this happening, what if these trends continue, what will happen next (predict), and what is the best outcome that can happen (optimize).

Business analytics. Abbreviated as BA, business analytics is the combination of skills, technologies, applications and processes used by organizations to gain insight in to their business based on data and statistics to drive business planning.

Q2. Explain the pre-requisties for effective business analytics?

Ans :

The prerequisites for effective business analysis are as follows,

(a) Goal Statement

A goal/mission statement is the most important prerequisite for BA to be effective. Every organization (big or small) need to have a clear and precise mission statement, which would motivate and guide its members to

strive and work towards the achievement of the specific mission statement. For example, the mission statement of I.B.M states "we strive to lead in the invention, development and manufacture of industry's, most advanced information technologies, including computer systems, software storage systems and micro electronics".

(b) Top Management Support

It is necessary for the lull support of the top management for the effective and result oriented implementation of BA.

(c) Team Work

Business analytics team need to work in sync with other departments of the organizations for collecting the required data and information which would be used for preparing BA reports.

4.1.2 Tools and Techniques of Business Analytics

Q3. What are the various Tools and Techniques of Business Analytics ?

Ans : **(June-19)**

Business analysis utilizes a large number of tools and techniques. These can be sub divided into,

a) Information and knowledge tools
b) Visualization tools
c) Vendor classification tools

a) Information and Knowledge Tools: These include the following tools,

i) OLAP : Online Analytical Processing (OLAP) refers to different variety of activities which are performed by the end user over the Internet. These includes generation of queries, answering of queries, requests for ad hoc reports and graphs, building of visual presentations and so on.

ii) Reports : Codd's rules state that reporting tools need to be uniform and flexible. The two most popular form of report tools are,

a) Routine Reports: These BA tools are automatically generated and distributed on a periodic basis. For example, the business analytics software would generate a weekly sales report of the store which is auto generated and mailed to the concerned managers.

b) Ad-Hoc Reports. These BA reports are created on demand for a specific user. For example, a report request for providing the list of customers who had purchased $ 10,000 or more worth of electronic item during the last three years.

iii) Data Mining: It is one of the most widely used B.A tool. The data mining tool mines the database and extract hidden, predictive information. This would be helpful in key decision making.

b) Visualization Tools: These BA tools are employed to have a balanced view of the organization. These includes the following,

i) Balanced Scorecard : The concept of Balanced Score Card (BSC) was primarily developed by Kaplan and Norton in an article in the Harvard Business Review. The balanced score card - measures the drive performance'. It developed great levels of interest for the senior business managers which resulted in the further development. The attention was basically diverted from short-term measurement to creation of growth, learning and value-added services to the consumers.

Several organizations are making of the balanced scorecard as a supporting or underlying structure for making vital process decisions. The performance measurement process has been converted into strategic management system due to the recent development of balanced scorecard.

Balanced score card is defined as a conceptual framework which is used for converting the strategic objectives of an organization into a group/collection of performance indicators which are divided into four areas i.e., financial, customer, internal business processes and learning and growth.

ii) **Dashboards :** These are visual display tools which displays raw data and visual graphics. This BA tool helps in measuring the performance.

c) **Vendor Classification Tools:** Micro Strategy divides its products into five categories. These are the five styles :

1. **Enterprise Reporting:** Enterprise reporting products are used to generate highly formatted static reports destined for broad distribution to many people. They are pixel-perfect report formats for operational reporting and dashboards.
2. **Cube analysis:** Cube-based BI tools are used to provide simple OLAP multi-dimensional slice-and-dice analytical capabilities to business managers in a limited-range environment.
3. **Ad hoc querying and analysis:** Relational OLAP tools are used to allow power users to query a database for any answer, slice-and-dice the entire database, and drill down to the lowest level of transactional information. This investigative querying is targeted to information explorers and power users.
4. **Statistical analysis and data mining:** Statistical, mathematical, and data mining tools are used to perform predictive analysis or to discover the cause-and-effect correlation between two metrics. Financial analyses and forecasts are also performed.
5. **Report delivery and alerting:** Report distribution engines are used proactively to send full reports or alerts to large user populations (internal and external), based on subscriptions, schedules, or threshold events in the databases.

Q4. Explain the categories of business analytics tools.

Ans :

Many BI activities evolved from two tools:

1. Executive information systems (EIS)

An EIS is a computer-based system that serves the information needs of top executives. It provides rapid access to timely and relevant information, to aid in monitoring an organization's performance by directly accessing management reports and to improve managerial growth and learning. An EIS is very user friendly, is supported by graphics, and provides the capabilities of exception reporting (i.e., reporting only the results that deviate from a set standard) and **drill-down** (i.e., investigating information in increasing detail). An EIS is also easily connected with online information services and e-mail.

2. Executive support systems (ESS)

An ESS is a comprehensive support system that goes beyond EIS to include analysis support, communications, office automation, and intelligence support.

4.1.3 Advanced Business Analytics

Q5. Why are advanced tools necessary for businesses?

Ans : **(Oct.-20)**

Need for Advanced Analytics Tools

The need for advanced analytical tool is required as a result of total integration of business and commerce with information technology. Every activity of any business organization is partially or wholly dependent upon the employment of information technology. Thus, it has become mandatory for organization to employ highly sophisticated business analytic tools to survive, remain competitive and increase their market share in today's highly competitive business world. Data mining tools and predictive analysis are two of the advanced business analytics tools used by global organizations.)

Data Mining and Predictive Analysis

A major step in managerial decision making is forecasting or estimating the results of different alternative courses of action. To do so, we can use one or more methods. The simplest methods involve only two variables, such as *sales* over *time* or *price* over *time.* Indeed, many statistical formulas are available to conduct such analyses (e.g., regression and correlation, trend analysis). The formulas for such methods appear as Excel or MicroStrategy functions.

However, in many cases, a prediction is more complex, involving more than two variables. For such cases, we can use more complex statistical methods, such as multiple regression analysis, or special forecasting and prediction methods. Two such methods are described next: data mining and predictive analysis.

i) Data Mining

Data mining tools automatically extract hidden, predictive information from databases. They also search for the patterns in large transaction databases. These tools are usually driven by complex statistical formulas. The easiest way to distinguish data mining from the various forms of OLAP is that OLAP can only answer questions you are certain to ask, whereas data mining answers questions you don't necessarily know you should ask.

ii) Predictive Analysis

Predictive analysis tools help determine the probable future outcome for an event or the likelihood of a situation occurring. They also identify relationships and patterns.

According to Fogarty (2004), predictive analysis uses sophisticated algorithms designed to sift through a data warehouse and identify patterns of behavior that suggest, for example, which offers your customers might respond to in the future or which customers you might be in danger of losing. For instance, when sifting through a bank's data warehouse, predictive analysis might recognize that customers who cancel an automatic bill payment or automatic deposit often move to another bank within a certain period of time.

Predictive analysis can analyze patterns of activity in accounts to identify geographic locations or changes in accounts. Then, it can correlate those factors to find patterns related to accounts that were closed. This approach would not only identify that a cancelled automatic payment was a danger sign but that location was not a relevant factor.

4.1.4 Business Analytics and Web

Q6. Define Web Analytics. Describe the use of Web in Business Analytics.

Ans : **(Imp.)**

Internet has revolutionized the way business is carried out globally. In the present day business world every aspect of the business is fully integrated with the web/ internet. The different tools, techniques skills and softwares of business analytics also need to be integrated with the web to provide real time information to the executives.

Definition of Web Analytics

It refers to the integration and applicability of business analytics activities with web based processes. In simple tenns web analysis refers to the interface between BA and the internet. It also includes the integration of business analytics along with e-commerce tools and techniques.

According to BA researcher Schlegel (2003), the methodology and tools of web analytics are highly visual in nature. He proposed the term click stream analysis.

Uses of Web in Business Analysis

A large proportion of business analytics tools and software technologies are related to the web. For example, Global heavy electronics manufacturer Hitachi incorporation (Japan) employs a wide range of web based e-commerce tools which are fully integrated with business analytics. Software applications such as ERP system, data warehouses etc. Integration of B.A with the web helps to create a synergy between both the elements. It enriches the business analytics data by providing real time inputs gains through the various e-commerce and web technology softwares.

Many multinational organizations are coming up with special tools and softwares which are being

integrated with web technologies to facilitate prudent decision making by the enterprise managers. For example, IBM uses a BA software known as decision edge software, which offers the end users OLAP capabilities on the Internet through the use of multiple web based search engines, browsers and other web technologies.

Another case in point is the BA software tool of Oracle incorporation. Its BA tools "Financial Analysers and Sales Analyses and Hummingbird BI can be early integrated with the latest web based technologies.

The scope of web analytics is huge and still contains untapped potential to be explored for the welfare of the organization.

The major benefits/uses of click stream analysis are,

a) It provides the business enterprise an idea about the way the web user visit of different web pages.

b) The interest of the potential customers can be forecasted by analysing the sequence in which the web user clicked upon different web pages. Further, the time duration spent upon each web page also indicates the interest of the end user (potential customer).

c) It would help the enterprise to measure the effectiveness of its web based promotional ads.

d) Based on the click stream analysis results, the firm may either continue with the existing web promotional campaign or it may create a new, more attractive web advertisement.

The data and information that can be gathered using the click stream analysis is huge. For instance, web analytic researchers Wemer and Abramson described a method of sorting and aggregation using which the software engineers can process 1 billion web records per day and collect useful business analytical information.

However, successful performance of click stream analysis require the following,

a) The raw click stream data need to be sourced from multiple sources.

b) The analysers need to employ cookies and session identification numbers in the URL's for the identification of amount of the data session.

c) A big proportion of the projects resources need to be allocated for date preparation.

Q7. Explain briefly about Vendor Support for Web Analytics.

Ans :

Many vendors such as Google incorporation, Cognos, web trends, informatics etc., provide extensive web support options to business enterprises to integrate their business analytics technologies with web technologies. These are discussed as follows,

1. Google

This popular vendor provides a software. "Google Analytics" to small and medium enterprises for free of cost. This web technology can be used by business organizations for click stream analysis of web logs.

2. Business Objects

This vendor provides its client organizations options to query and analyze web content through its software "Web Intelligence". It facilitates the client enterprises to track manage and store information in multiple web based data sources which can be later integrated as per the organizations conveniences.

3. Cognos

This vendor provides its client organization its web software named "Cognos 8" which helps its client customers to develop softwares easily.

4. Informatica

This vendor provides its client organization "The informatical Analytics Delivery Platform" which helps business enterprises to closely track their business performance. It provides the "Power Centre 8 BI Platform" though which organization can gather detailed business performance metrics.

5. Web Trends

This vendor provides business organizations to analyse the internet traffic on a real time basis. Employing its, software products, organizations can track customer purchasing trends, revenues, effectiveness of sales, promol and other interesting and useful details, about the online customers.

6. Advisor Solutions

It offers the client organizations an interactive chart library that provides useful information about business display needs, on demand real time analysis, point and click self services etc.

7. Angoss Knowledge Studio

This vendor provides good web mining tools for client organization.

4.1.5 Usage, Benefits and Success of Business Analytics

Q8. What are the Usage, Benefits and Success of Business Analytics ?

Ans : **(OCt.-20, June-18)**

Usage of Business Analytics

In the present business scenario, the usage of Business Analytics increasing day by day. Most of the medium and large scale Organization successfully utilizing the benefits of business analytics.

Majority managers and executives of organizations using Business Analytics as an effective system of managers business transaction. However, some managers and executives consider it as a complicated system to use.

Now a days/every business organization wants to use the system of business analytics for proper functioning of business operations but distribution of analytics tools through out the organization is a very difficult task. Because, various issues and challenges may arise during the adoption of new technologies. Some of the challenges maybe related to culture, people, processes and so on.

In addition to this, the major issue or challenge which may arise is the application of Business Analytics as per the business needs. The tools of Business Analytics has been used as to identify white-color theft in organizations. These can be effectively used in identifying the inflated invoices, customer impersonation and similar offences.

Benefits of Business Analytics

The various benefits of Business Analytics (BA) are as follows,

1. It helps the organizations in measuring how much the mission statement is accomplished.
2. It provide accessibility to essential information data through which organizations can make smart, proper and accurate business decisions.
3. It provides clear and appropriate future insights to the organization through visualization.
4. It improves the efficiency of business organizations. As the Business Analytics is helpful in gathering large amount of information data at faster rate with visual representation, organization can fonnulate and design plans or decisions for achieving goals.
5. It helps in synchronizing the strategies related to financial and operational transactions.
6. It helps in increasing revenues and improving competitiveness.
7. It helps in maintaining effective collaboration and co-ordination between different strategies.
8. It helps in responding to the needs of customers on time by using available information of customer needs and requirements.

Q9. Write about the success of Business Analytics.

Ans :

The success rate of Business Analytics in the present business environment is increasing with a good speed. According to the report of IDC, the organizations which successfully implemented and utilized Business Analytics observed high returns between 17% to 2000% with a median Return on Investment (ROI) of 122%.

The main purpose of implementing Business Analytics is to identify the white-color theft in organizations. Through, business analytics, organization will be able to identify not only internal frauds but also external frauds. BI helps to integrate

internal data with data warehouses for fraud analysis. When internal data compared with external data, the patterns and anomalies become identifiable. Moreover, suspicious activities also can be isolated, measured and tracked through fraud analysis.

Example

Williams-Sonoma corporation is an example which achieve success through business analytic techniques. It save millions of rupees with targeted marketing and multi-channel branding by using SAS data mining software.

The success of Business Analytics in recent years is summarized in the following points,

1. The usage of Business Intelligence and Analytics is relatively high and growing.
2. Large organizations are using BA more consistently compare to small organizations.
3. In 2002, most of the successful companies invested 50% income on Business Intelligence Technologies compare to Un-successful companies.
4. Government organization more frequently uses Business Analytics tools compare to any other organization or sector of economy.
5. Few companies are happy their competitive intelligence practices.
6. The companies which have used Business Analytics tools are confident on the information which they have gathered regarding customers.

Q10. Explain the various challenges in business analytics.

Ans :

- **Executive Ownership:** Business Analytics requires buy-in from senior leadership and a clear corporate strategy for integrating predictive models.
- **IT Involvement:** Technology infrastructure and tools must be able to handle the data and Business Analytics processes.
- **Available Production Data vs. Cleansed Modeling Data :** Watch for technology infrastructure that restrict available data for historical modeling, and know the difference between historical data for model development and real-time data in production.
- **Project Management Office (PMO)** : The correct project management structure must be in place in order to implement predictive models and adopt an agile approach.
- **End user Involvement and Buy-In :** End users should be involved in adopting Business Analytics and have a stake in the predictive model.
- **Change Management :** Organizations should be prepared for the changes that Business Analytics bring to current business and technology operations.
- **Explain ability vs. the "Perfect Lift"**: Balance building precise statistical models with being able to explain the model and how it will produce results.

4.2 DATA VISUALIZATION

4.2.1 Definition

Q11. What is Data visualization? Explain the importance of data visualization?

Ans : **(Imp.)**

It is a general term that describes any effort to help people understand the significance of data by placing it in a visual context. Patterns, trends and correlations that might go undetected in text-based data can be exposed and recognized easier with data visualization software.

- Today's data visualization tools go beyond the standard charts and graphs used in Microsoft Excel spreadsheets, displaying data in more sophisticated ways such as info graphics, dials and gauges, geographic maps, spark lines, heat maps, and detailed bar, pie and fever charts. The images may include interactive capabilities, enabling users to manipulate them or drill into the data for querying and analysis. Indicators designed to alert users when data has been updated or predefined conditions occur can also be included.

Importance of data visualization

- Data visualization has become the de facto standard for modern business intelligence (BI). The success of the two leading vendors in the BI space, Tableau and Qlike - both of which heavily emphasize visualization - has moved other vendors toward a more visual approach in their software. Virtually all BI software has strong data visualization functionality.
- Data visualization tools have been important in democratizing data and analytics and making data-driven insights available to workers throughout an organization. They are typically easier to operate than traditional statistical analysis software or earlier versions of BI software. This has led to a rise in lines of business implementing data visualization tools on their own, without support from IT.
- Data visualization software also plays an important role in big data and advanced analytics projects. As businesses accumulated massive troves of data during the early years of the big data trend, they needed a way to quickly and easily get an overview of their data. Visualization tools were a natural fit.
- Visualization is central to advanced analytics for similar reasons. When a data scientist is writing advanced predictive analytics or machine learning algorithms, it becomes important to visualize the outputs to monitor results and ensure that models are performing as intended. This is because visualizations of complex algorithms are generally easier to interpret than numerical outputs.

Examples of data visualization

Data visualization tools can be used in a variety of ways. The most common use today is as a BI reporting tool. Users can set up visualization tools to generate automatic dash boards that track company performance across key performance indicators and visually interpret the results.

Many business departments implement data visualization software to track their own initiatives. For example, a marketing team might implement the software to monitor the performance of an email campaign, tracking metrics like open rate, click-through rate and conversion rate.

Q12. How data visualization works?

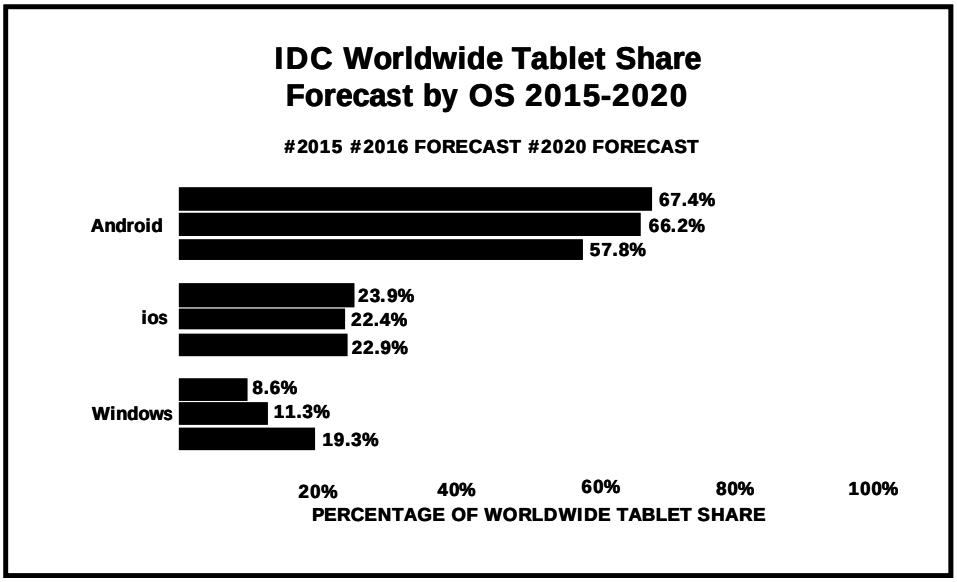
Ans :

Most of today's data visualization tools come with connectors to popular data sources, including the most common relational databases, Hadoop and a variety of cloud storage platforms. The visualization software pulls in data from these sources and applies a graphic type to the data.

Data visualization software allows the user to select the best way of presenting the data, but, increasingly, software automates this step. Some tools automatically interpret the shape of the data and detect correlations between certain variables and then place these discoveries into the chart type that the software determines is optimal.

Typically, data visualization software has a dashboard component that allows users to pull multiple visualizations of analyses into a single interface, generally a web portal.

Below is a chart forecasting tablet sales by operating system.



Q13. Explain the uses of data visualization.

Ans :

- By using data visualization, it became easier for business owners to understand their large data in a simple format.
- The visualization method is also time saving. So, businesses does not have to spend much time to make a report or solve a query. They can easily do it in a less time and in a more appealing way.
- Visual analytics offers a story to the viewers. By using charts and graphs or images, a person can easily exposure the whole concept as well the viewers will be able to understand the whole thing in an easy way.
- The most complicated data will look easy when it gets through the process of visualization. Complicated data report gets converted into a simple format. And it helps people to understand the concept in an easy way.
- With the visualization process, it gets easier to the business owners to understand their product growth and market competition in a better way.

4.2.2 New Directions in Data Visualisation

Q14. What are the New Directions in Data Visualisation ?

Ans :

Since the late 1990s, data visualization has moved both into mainstream computing, where it is integrated with decision support tools and applications, and into intelligent visualization, which includes data (information) interpretation. For interesting areas and trends, see Online File W6.3.

Major OLAP vendors provide three-dimensional visualization tools with their decision support tools. For example, Forest Tree 6.0 is a Web-enabled development tool that has a three-dimensional visualization version that enables users to visualize and easily manage multiple dimensions of data in a single view. New visual tools are continually being developed to analyze Web site performance. ADVIZOR Solutions is one such tool.

1. Dashboards and Scorecards

Visualization has been proven to be extremely important for busy executives. The EIS of the 1990s were loaded with charts, graphs, and tables. These evolved to management cockpit products

2. Visual Analysis

Today, enterprise data analysis can be done by nontechnical users who glean valuable information from business data. VizQL (from Tableau Software) is a visual database query language that powers Hyperion's Visual Explorer. Several other companies pro-vide tools for visual analysis (e.g., Analytical Endeca. Visual analysis can be done interactively

3. Financial Data Visualisation

One of the common applications of BI data visualization is in the area of finance. To prevent systems from automatically identifying meaningless patterns in data, chief financial officers (CFOs) want to make sure that the processing power of a computer is always tempered with the insight of a human being. One way to do this is through data visualization.

Q15. What are the techniques of Data Visualization?

Ans : **(June-19)**

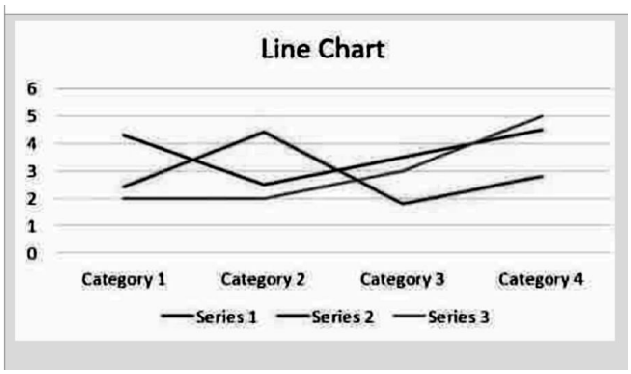
The data visualization techniques are Diagrams, charts, graphs.

Most widely used forms of data visualization are presented below:

1. Pie Chart

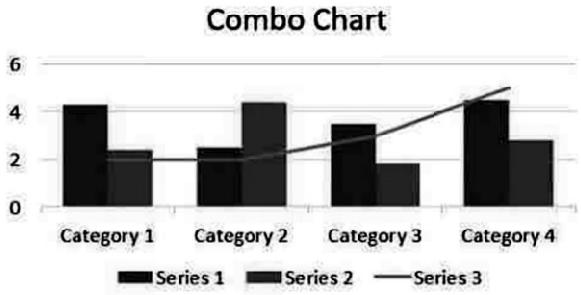
Pie Charts : Pie Charts are one of the common popular techniques. It also comes under data visualization techniques in excel. However, to some people, it can be hard to understand the chart while comparing to the line and bar type chart.

2. Line Chart



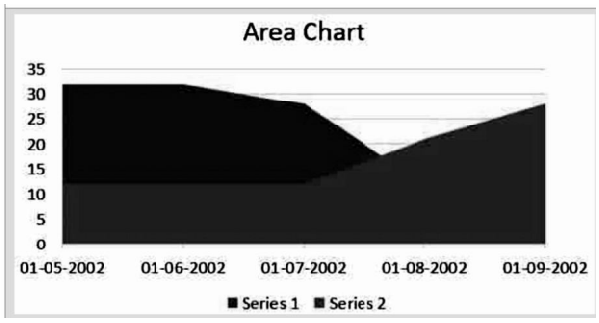
To make your data simple and more appealing you can simply use the line charts technique. Line chart basically displays the relationship between two patterns. Also, it is one of the most used techniques world wide.

3. Combo Chart



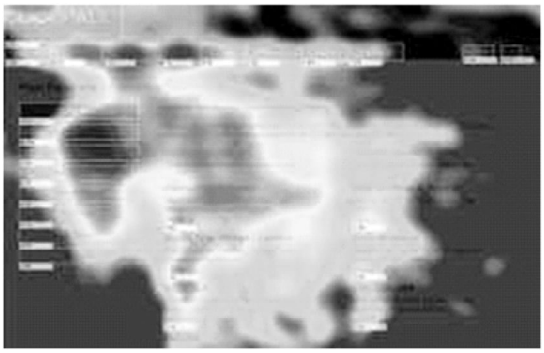
Bars charts are also one of the most commonly used techniques when it comes to comparing two different patterns. The bar charts can display the data in a horizontal way or in a vertical way. It all depends on your needs.

4. Area Chart



An area chart or area graph is similar to a line chart but provides graphically quantitative data. The areas can be filled with colour, hatch, pattern. This chart is generally used when comparing quantities which is depicted by area.

5. Heat Map

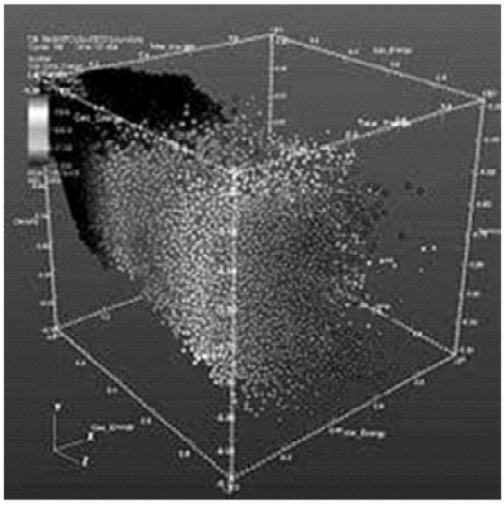


This type of chart is widely used by websites, mobile application makers, research institutes etc. These maps shows the concentration of activity/entity over a particular area.

6. Network Diagram

This is a powerful tool for finding out connections & correlations. It highlights and bridges the gaps. Shows strongly one activity is connected to other.

7. Scattered 3 D Plot



As the image shows it shows the distribution of entity in a 3 dimensional nature. It can be considered as showing location and concentration of gases in a box with different colours assigned to each gas.

4.2.3 GIS (Geographical Information System)

Q16. Explain about GIS (Geographical Information System).

Ans:

A geographical information system (GIS) is a computer-based system for capturing, storing, modeling, retrieving, checking, integrating, manipulating, analyzing, and displaying geographically referenced data by using digitized maps.

The most distinguishing characteristic of GIS is that every record or digital object has an identified geographic location. By integrating maps with spatially oriented (i.e., geographic location) databases (called *geocoding*) and other databases, users can generate information for planning, problem solving, and decision making, thereby increasing their productivity and the quality of their decisions. Areas as diverse as retailing, banking, transportation, agriculture, natural resource management, public administration, NASA, the military, emergency preparedness, and urban planning have all successfully used GIS since the beginning of the 1970s.

A Comprehensive geographical information system requires a method for,

1. Gathering data input from various sources like maps, aerial photos, satellites.
2. Performing data storage, retrieval and query process.
3. Performing data transformation, analysis and modeling.
4. Delivering the data in the form of maps, reports and plans.

GIS Applications

As GIS tools become increasingly sophis-ticated and affordable, they help more companies and governments understand precisely where their trucks, workers, and resources are; where they need to go to service a customer; and the best way to get from here to there. The areas of targeted marketing are growing rapidly, and organizations can easily segment a population by using GIS.

Applications of GIS

1. GIS supplies huge amount of efficient information that is helpful in analyzing and making decisions.
2. The graphical representation of GIS helps the managers in visualizing the data without any difficulty.
3. GIS is used to enhance the decision-making process in both public and private sectors.
4. It is used in local government for mapping as well as in decision making applications.
5. It helps the organization in gaining competitive advantage by providing the organization with the information about the areas where marketing is growing rapidly.
6. In banks, GIS is used to support activities like finding out the branch and ATM location. Determining customer's residence address, age, income level, account number.

In addition to the above applications, certain examples of successful GIS applications are as follows.

Organization		Application
1.	Sun Micro Systems	It manages leased properties at various places through entire world.
2.	Toyota and other automobile manufacturers.	It makes use of combination of GIS and GPS as navigation tool for direct the drivers to their destinations using best routes.
3.	Western Auto	It combines the data with GIS inorder to create a demographic profile of a neighbour of any store in detail and generate best product mix for the store.
4.	Pepsi, Cola, Inc., Super Value, Acodia, Inc.	It makes use of GIS for selecting the sites of new Taco Bell and Pizza hut restaurants.
5.	Sears, Roebuck & Co/Kmart	It uses GIS for supporting truck routes planning.
6.	Wood Personnel Services	It maps the neighbours where temporary workers stay for knowing the location of marketing and recruiting cities.

Many banks use GIS to support activities such as :

- Determining branch and ATM locations
- Analyzing customer demographics (e.g., residence, age, income level) for each of a bank's products
- Analyzing volume and traffic patterns of business activities
- Analyzing the geographic area served by each branch
- Determining the market potential for banking activities
- Evaluating strengths and weaknesses against those of the competition
- Evaluating branch performance.

Q17. Define Global positioning system (GPS).

Ans :

Global positioning system is a space-based weather radio navigation system by which the time, location and velocity of an object can be determined any where on the globe at any time. With the help of proper receiving equipment this system allows exact, continuous, three-dimensional location and velocity information to the user. It uses the satellite signals to produce distances in order to triangulate location anywhere on earth.

Global positioning system consists of three segments, as listed below,

1. Space segment
2. Operational control segment
3. User equipment.

1. Space Segment

It is a satellite constellation segment in which user equipment receives data messages and ranging signals through satellites present in orbit.

2. Operational Control Segment (OCS)

Operational control segment detect and maintain the orbital configuration of the satellite in space. It reviews the satellite clock corrections and other essential parameters required for finding the location of user, velocity and time.

3. User Equipment Segment

User equipment receives the Z,-band signals transmitted from the satellite and find the location of user, velocity and time.

This global position system determine the position of moving object with the help of various survey techniques.

4.2.4 GIS vs GPS

Q18. Compare and contrast GIS Vs GPS.

Ans : **(June-18)**

i) Global Positioning System (GPS)

GPS (completely known as Global Positioning System) is one of the ways to precisely pinpoint specific locations in almost any place on the planet. Simply, it is a network of satellites that determines specific coordinates on earth. This network usually operates by the transmission of certain radio signals from the said satellites into the GPS receivers on the planet using the process known as trilateration. Using the U.S. government's advanced tracking satellite technologies, the GPS system can also locate coordinates with the combination of multiple GPS satellites, receivers and oftentimes a processing circuit for data.

This technology was first developed for the purpose of the U.S. military, but little did they know that this invention could one day revolutionize how people look for what they are trying to find. The GPS has now been used in practical day to day applications and makes life easier. At present, this system is being used universally to navigate, to map an area and even to survey a specific geographical location.

ii) Geographic Information System (GIS)

On the other hand, GIS is frequently confused with GPS because it is a more generic acronym (Geographic Information System) used to describe a more complex mapping technology that is connected to a particular database. Because it's generic, it is a broader term than the GPS in its technical sense. Thus, GIS is a computer program or application that is utilized to view and handle data about geographic locations and spatial correlations among others. It simply gives the user a framework to obtain information.

Overall, the difference between a GPS and a GIS can be summed up:

1. The GPS (Global Position System) is a network that locates certain places here on earth whereas the GIS (Geographic Information System) is a computer program that process data linked to certain places or locations.
2. The GIS is a more generic framework compared to the specific GPS network.

Short Question and Answers

1. Define is business analytics

Ans :

Business analytics is the practice of iterative, methodical exploration of an organization's data, with an emphasis on statistical analysis. Business analytics is used by companies committed to data-driven decision-making.

Definitions of business analytics

i) **According to Schaer** (2018) - "allows your business to make predictive analysis rather than reacting to changes in data".

ii) **According to Gabelli School of Business** (2018)- "involves applying models, methods, and tools to data, producing insights that lead to informed business decisions".

iii) **According to Wells** (2008) - "the application of logic and mental processes to find meaning in data".

iv) **According to Lynda** (2018) - "allows us to learn from the past and make better predictions for the future".

2. Executive information systems (EIS)

Ans :

An EIS is a computer-based system that serves the information needs of top executives. It provides rapid access to timely and relevant information, to aid in monitoring an organization's performance by directly accessing management reports and to improve managerial growth and learning. An EIS is very user friendly, is supported by graphics, and provides the capabilities of exception reporting (i.e., reporting only the results that deviate from a set standard) and drill-down (i.e., investigating information in increasing detail). An EIS is also easily connected with online information services and e-mail.

3. Define Web Analytics.

Ans :

Internet has revolutionized the way business is carried out globally. In the present day business world every aspect of the business is fully integrated with the web/ internet. The different tools, techniques skills and softwares of business analytics also need to be integrated with the web to provide real time information to the executives.

Definition of Web Analytics

It refers to the integration and applicability of business analytics activities with web based processes. In simple tenns web analysis refers to the interface between BA and the internet. It also includes the integration of business analytics along with e-commerce tools and techniques.

According to BA researcher Schlegel (2003), the methodology and tools of web analytics are highly visual in nature. He proposed the term click stream analysis.

4. Benefits of Business Analytics

Ans :

The various benefits of Business Analytics (BA) are as follows,

1. It helps the organizations in measuring how much the mission statement is accomplished.
2. It provide accessibility to essential information data through which organizations can make smart, proper and accurate business decisions.
3. It provides clear and appropriate future insights to the organization through visualization.
4. It improves the efficiency of business organizations. As the Business Analytics is helpful in gathering large amount of information data at faster rate with visual representation, organization can fonnulate and design plans or decisions for achieving goals.
5. It helps in synchronizing the strategies related to financial and operational transactions.

5. Data visualization.

Ans :

It is a general term that describes any effort to help people understand the significance of data by placing it in a visual context. Patterns, trends and correlations that might go undetected in text-based data can be exposed and recognized easier with data visualization software.

- Today's data visualization tools go beyond the standard charts and graphs used in Microsoft Excel spreadsheets, displaying data in more sophisticated ways such as info graphics, dials and gauges, geographic maps, spark lines, heat maps, and detailed bar, pie and fever charts. The images may include interactive capabilities, enabling users to manipulate them or drill into the data for querying and analysis. Indicators designed to alert users when data has been updated or predefined conditions occur can also be included.

6. Uses of data visualization.

Ans :

- By using data visualization, it became easier for business owners to understand their large data in a simple format.
- The visualization method is also time saving. So, businesses does not have to spend much time to make a report or solve a query. They can easily do it in a less time and in a more appealing way.
- Visual analytics offers a story to the viewers. By using charts and graphs or images, a person can easily exposure the whole concept as well the viewers will be able to understand the whole thing in an easy way.
- The most complicated data will look easy when it gets through the process of visualization. Complicated data report gets converted into a simple format. And it helps people to understand the concept in an easy way.
- With the visualization process, it gets easier to the business owners to understand their product growth and market competition in a better way.

7. Geographical Information System

Ans:

A geographical information system (GIS) is a computer-based system for capturing, storing, modeling, retrieving, checking, integrating, manipulating, analyzing, and displaying geographically referenced data by using digitized maps.

The most distinguishing characteristic of GIS is that every record or digital object has an identified geographic location. By integrating maps with spatially oriented (i.e., geographic location) databases (called *geocoding*) and other databases, users can generate information for planning, problem solving, and decision making, thereby increasing their productivity and the quality of their decisions. Areas as diverse as retailing, banking, transportation, agriculture, natural resource management, public administration, NASA, the military, emergency preparedness, and urban planning have all successfully used GIS since the beginning of the 1970s.

A Comprehensive geographical information system requires a method for,

1. Gathering data input from various sources like maps, aerial photos, satellites.
2. Performing data storage, retrieval and query process.
3. Performing data transformation, analysis and modeling.
4. Delivering the data in the form of maps, reports and plans.

8. Global Positioning System (GPS)

Ans :

GPS (completely known as Global Positioning System) is one of the ways to precisely pinpoint specific locations in almost any place on the planet. Simply, it is a network of satellites that determines specific coordinates on earth. This network usually operates by the transmission of certain radio signals from the said satellites into the GPS receivers on the planet using the process known as trilateration. Using the U.S. government's advanced tracking satellite technologies, the GPS system can also locate coordinates with the combination of multiple GPS satellites, receivers and oftentimes a processing circuit for data.

This technology was first developed for the purpose of the U.S. military, but little did they know that this invention could one day revolutionize how people look for what they are trying to find. The GPS has now been used in practical day to day applications and makes life easier. At present, this system is being used universally to navigate, to map an area and even to survey a specific geographical location.

Choose the Correct Answer

1. BA stand for _____ [a]
   (a) Business analytics (b) Business intelligence
   (c) BPR (d) None of the above
2. Tool of BA [d]
   (a) OLAP (b) Data Mining
   (c) EIS (d) All
3. Example of Data Visualisation [b]
   (a) Business analytics (b) Risk management
   (c) Business intelligence (d) Data mining
4. _____ can be defined as the art and science of analysing any data or information. [d]
   (a) Business intelligence (b) Data mining
   (c) Visualizations tools (d) Analytics
5. New Direction in Data Visualisation [a]
   (a) Dashboard (b) Finance
   (c) Cyberspace (d) None of the above
6. EIS stand for [d]
   (a) Microsoft incorporation (b) ESS
   (c) ILOG (d) Executive Information System
7. Trends can be identified by data visualization [b]
   (a) Methods (b) Tools
   (c) Techniques (d) Types
8. GPS stand for [a]
   (a) Global Positioning System (b) Global Product System
   (c) Global Price System (d) All
9. _____ delivers data in the form of maps, reports and plans. [a]
   (a) GIS (b) GPS
   (c) Visualization (d) Spreadsheets
10. Web site performance can be analyzed with . [d]
    (a) Visualization (b) GIS
    (c) GPS (d) Visual tools

Fill in the blanks

1. Vendor classification tool _____
2 ESS stand for _____
3. Need of Advanced Analytics Tool _____
4. _____ refers to different variety of activities which are performed by the end uses over the internet.
5. _____ tools look similar to OLAP.
6. _____ allows BA through the use of web based tools.
7. New Direction of Data Visualisation _____
8. _____ allows to store, analyze and display the data geographically from base location.
9. GIS is a _____.
10. _____ is a space based weather radio navigation system.

ANSWER

1. Tube Analysis
2. Executive Support System
3. Data Mining
4. Online Analytical Processing (OLAP)
5. Data mining
6. Data visualization
7. Dash Board
8. GIS
9. Geographical Information System
10. GPS

UNIT V

Business Intelligence Implementation

(a) Implementing Business Intelligence – Implemental Factors, Critical Success factors of Business Implemental, Managerial Issues related to BI Implementation. Business Intelligence and Integration Implementation – Types, Need, Level of Business Intelligence Integration.

(b) Emerging trends in Business Intelligence Implementation- Social Net works and Business Intelligence, Collaborative Decision Making, RFID and Business Intelligence, Reality Mining.

5.1 Implementing Business Intelligence

5.1.1 Implemental Factors

Q1. Define Business Intelligence Implementation. What are the factors that affect the implementation of business intelligence?

Ans : **(Imp.)**

Introduction

Implementing BI systems can be very complex. In addition to typical issues in information system implementation, such as conducting appropriate cost-benefit analysis with intangible variables to justify the system and dealing with resistance to change, there are some complex issues related to integration, security, scalability of the system, and handling the construction of a data warehouse, analytics, and dashboards.

Factors

A large number of factors may influence BI implementation. These factors are technology-ical, administrative, behavioral, and so on. Many of these are generic to most information systems and have been researched extensively in the information systems literature.

According to Asif (2009), the following are the major factors that affect the decision making process of BI implementation.

1. Reporting and Analysis Tools

a) Features and functionality

b) Scalability and deployability

c) Usability and manageability

d) Ability to customize applications

2. Database

a) Scalability and performance

b) Manageability and availability

c) Security and customization

d) Ability to write back

3. Extraction, Transformation, and Load (ETL) Tools

a) Ability to read any source

b) Efficiency and productivity

c) Cross platform support

4. Costs Involved

a) Hardware costs (actual or opportunity)

b) Costs of software (ETL, database, applications, and front end)

c) Internal development costs

d) External developments costs

e) Internal training

f) Ongoing maintenance

5. Benefits

a) Time savings and operational efficiencies

b) Lower cost of operations

c) Improved customer service and satisfaction

d) Improved operational and strategic decision making

e) Improved employee communications and satisfaction

f) Improved knowledge sharing

These factors need to be analyzed both quantitatively and qualitatively.

5.1.2 Critical Success factors of Business Implemental

Q2. Elucidate Critical Success factors of Business Implementation ?

Ans : **(Oct.-20, June-19)**

Although there could be many factors that could affect the implementation process of a BI system, a report from Vodapalli (2009) as cited on Wikipedia shows that the following are the critical success factors for a business intelligence implementations.

a) Business driven methodology and project management.

b) Clear vision and planning.

c) Committed management support and sponsorship.

d) Data management and quality issues.

e) Mapping the solutions to the user requirements.

f) Performance considerations of the BI system.

g) Robust and extensible framework.

5.1.3 Managerial Issues Related to Business Intelligence Implementation

Q3. Explain the Managerial Issues Related to Business Intelligence Implementation ?

Ans : **(June-18)**

Managerial Issues Related to BI Implementation

Many managerial issues are related to BI implementation. Illustrative topics are:

1. System development and the need for integration

Developing an effective BI application is complex. For this reason, most BI vendors offer highly integrated collections of applications, including connection to enterprise resource planning (ERP) and customer relationship management. Notable are Oracle, Business Objects, MicroStrategy, IBM, and Microsoft. Most BI vendors pro-vide for application integration, usually Web enabled.

2. Cost-benefit issues and justification

Some BI solutions discussed in this book are very expensive and are justifiable only in large corporations. Smaller organizations can make the solutions cost effective if they leverage existing databases rather than create new ones. One solution is on-demand BI. Nevertheless a careful cost-benefit analysis must be undertaken before any commitment to BI is made.

3. Legal issues and privacy

BI analysis may suggest that a company send elec-tronic or printed catalogs or promotions to only one age group or one gender. A man sued Victoria's Secret (a brand of Limitedbrands) because his female neigh-bor received a mail order catalog with deeply discounted items and he received only the regular catalog (the discount was actually given for volume purchasing). Settling discrimination charges can be very expensive. Some data mining may result in the invasion of individual privacy.

4. BI and BPM today and tomorrow

The quality and timeliness of business information for an organization is not the choice between profit and loss - it may be a question of survival. No enterprise can deny the inevitable benefits of BI and BPM. Recent industry analyst reports show that in the coming years, millions of people will use BPM dashboards and business analytics (BA) every day.

Enterprises are get-ting more value from BI by extending information to many types of employees, maximizing the use of existing data assets. Visualization tools including dashboards are used by producers, retailers, government, and special agencies. Industry-specific analytical tools will flood the market to support analysis and informed decision making from top level to user level. BI takes advantage of existing IT technologies to help companies leverage their IT investments and use their legacy and real-time data. Thus a planned, careful, proactive approach to BI implementation is becoming a competitive necessity.

5. Cost justification

Intangible benefits. While enterprise systems provide tangible benefits, it is difficult to quantify their intangible benefits. In a down-turned economy with high energy costs, mortgage crises, and political unrest, IT investments must be economically justified.

6. Documenting and securing support systems

Many employees develop their own decision support or BI modules to increase their productivity and the quality of their work. It is advisable to have an inventory of these ad hoc systems and make certain that appropriate documentation and security measures exist, so that if the employee is away or leaves the organization, the productivity tool remains. Taking appropriate security measures is a must. End users who build their own BI applications are not professional systems builders. For this reason, there could be problems with data integrity and the security of the systems developed.

7. Ethical issues

BI and predictive analytics can lead to serious ethical issues such as privacy and accountability. In addition, mistakes can cause harm to oth-ers as well as the company.

For example, a company developed a decision sup-port system (DSS) to help people compute the financial implications of early retirement. However, the DSS developer did not include the tax implications, which resulted in incorrect retirement decisions.

Another important ethical issue is human judgment, which is frequently a key factor in decision making.

Human judgment may be subjective or corrupt, and therefore, it may lead to unethical decision making. Companies should provide an ethical code for system builders. Also, the possibility of automating managers' jobs may lead to massive layoffs.

There are ethical issues related to the implementation of expert systems and other intelligent systems. The actions performed by an expert system can be unethical, or even illegal.

For example, the expert system may advise you to do something that will hurt someone or will invade the privacy of certain individuals.

An example is the behavior of robots and the possibility that the robots will not behave the way that they were programmed to. There have been many industrial accidents caused by robots that resulted in injuries and even deaths.

8. BI Project failures

There have been many cases of failures of all types of BI projects. There are multiple reasons for such failures, ranging from human factors to software glitches. Here are some examples:

a) Failure to recognize BI projects as enterprise-wide business'initiatives and that they differ from typical stand-alone solutions.

b) Lack of business sponsors with the ability to insure funding

c) Lack of cooperation by business representatives from the functional areas

d) Lack of qualified and available staff

e) No appreciation of the negative impact of "dirty data" on business profitability

f) Too much reliance on vendors.

5.1.4 Business Intelligence and Integration Implementation

Q4. Expalin the different types of integration.

Ans :

Integrating information systems is widely practiced in enterprises, increasing the efficiency and/or effectiveness of the tasks supported.

Implementing BI almost always requires one or several integration steps. However, integration is not simple, as described in this section and the following section.

Types of Integration

Computer-based systems can be integrated so that the constituent parts of the system func-tion as one entity, as opposed to each being used separately. Integration can be at the development level or at the application system level (known as *application integration,* our main area of interest). Integration has been considered a top issue of importance for years (Spangler, 2005). There are several types of integration: integration of data, applications, methods, and processes. Integration can also be viewed from two other characteristics: functional and physical.

1. Functional integration

Functional integration implies that different applications are provided as a single system. For example, working with e-mail, using a spreadsheet, communicating with external databases, creating graphical representations, and storing and manipulating data can all be accomplished at the same workstation. Similarly, working with a business analytics tool and a dashboard is done from one interface, with one menu, resulting in one output.

2. Physical integration

Physical integration refers to packaging the hardware, software, and communication features required to accomplish functional integration. The discussion in this chapter deals primarily with functional-application integration, which can be done in two ways :

- Integration of two or more decision-support applications, creating a unified application
- Integration of one or more BI tools with other information systems such as blogs, **knowledge management,** databases, or a financial system.

Integration can occur within a company (internal integration) or between systems of different companies (external integration).

Q5. Explain nees of business intelligence software.

Ans :

There are several major objectives for BI software integration:

1. Implementing BI

For BI systems to operate, they usually need to be connected to data sources, utilities, other applications, and so on. Such connections must be accomplished effectively and efficiently.

2. Increasing the capabilities of the BI applications

Several BI development tools may complement each other. Each tool performs the subtasks at which it is the best. For example, BA can be used to recommend an optimal resource-allocation plan, and an attached dashboard can provide the control system that will alert management to deviations from the plan. The opening vignette demonstrated how social software made the BI system working better.

3. Enabling real-time decision support

By having tight integration, it is possible to support decision making in a real-time environment. An example is a transportation system that uses wireless communication and Web services to foster data flow.

4. Enabling more powerful applications

An example is using intelligent systems to provide real-time capabilities.

5. Facilitating system development

Tighter integration allows faster application development and communication among system components.

6. Enhancing support activities

Several support activities can improve the operations of BI applications. For example, blogs, Twitter, wikis, and RSS feeds provide communication and collaboration support as shown in the opening vignette.

BI integration may also result in enhanced capabilities that are not possible other-wise. For strategies for successful integration.

Q6. What are the levels of BI Integration.

Ans :

i) Functional Integration

Functional integration, discussed previously, can be considered at two different levels: across different BI and within BI. Integration of BI at these levels is appropriate for systems that can be used to solve repetitive and/or sequential decision problems. BI can also be used to facilitate integration by assisting in the transformation of the outputs of one system as inputs to another system.

ii) Physical Integration

Physical integration refers to the process of combining hardware, software and other characteristics of communication that are required for achieving functional integration.

5.2 EMERGING TRENDS IN BUSINESS INTELLIGENCE IMPLEMENTATION

Q7. Describe the Developments predicted by Gartner in Business Intelligence Market.

Ans : **(Oct.-20, Imp.)**

The following are the Gartner's predictions for the development of BI market,

1. Minimum 40% of the entire budget will be managed by business units by the year 2012.
2. 20% of the companies would employ industry specific analytic applications as most reliable component of their business intelligence portfolio untill the year 2010.
3. Collaborative decision making would surge(rise) as a new product by collaborating social software and business intelligence platform by the year 2009.
4. 20% of the analytic application corresponding to business processes might be transferred by using coarse grained application mashups by the year 2012.
5. By the year 2012 about 35% of top companies might not succeed in making an appropriate decisions on the changes existing in business and markets due to insufficient data, processes and tools.

Q8. Describe briefly about Web 2.0 as a emerging trends in business intelligence.

Ans :

While traditional technologies for supporting BI such as data warehousing, OLAP, data mining, etc. query data from the inside of the organization, novel trends of BI (BI 2.0) that focus on the analysis of external data have emerged. The outcome of having a wider data pool is that the analysis is more comprehensive, and provides better platform and support for decision making.

For example, a retail company that would typically focus its BI strategies internally, can now have access to prices of competitors, and can also have information to how the customers receive their products; this will in turn affect their strategy for the release of another product.

The internet has been the driving force for data and information dispersion over the last 10 years and will continue to be for a long time. One important thing to note is that the new trend of BI 2.0 is bidirectional: because BI applications receive and process data from the web, BI applications are starting to evolve to become pervasively web driven technologies.. Some of the BI driven applications today are Social Networks, Cloud Computing, Interactive WebApps, SaaS (Software as a Service), Semantic Web Search, Word Clouds, Wiki's, Collaborative Networks, e.t.c.

Real-time analysis: BI 2.0 has to focus on data analysis that is on-line and real time. Because of the new found volatile nature of data, BI 2.0 has to be super flexible and super reactive in nature to process and analyze data.

Intuitive and interactive analysis: Unlike traditional data processing that typically involves homogenous data types and 1-directional analysis, BI 2.0 supports interactive analysis that acts based on a user's preference, and that can be altered and re-analyzed if certain variables are changed.

Collaboration between decision-makers: A key concept in BI 2.0 is the collaboration that should exist between decision-makers in the manipulation of analyzed data.

Linking and enriching data: The concept of Business Intelligence itself is that of the linking of different data to present better information for decision making, and BI 2.0 takes it a notch higher by enriching this data to be more self-explanatory and therefore increases the utility of the information eventually provided.

Q9. Describe the concept of virtual worlds.

Ans :

Virtual world is a technology that creates an artificial environment that responds to, and is controlled by the behavior of the user. It is created by using interactive graphics software which is so closed to reality that people feel this artificial reality really does exist. The same computer-generated artificial environment can be shared by more than one person and even a large group. In many virtual reality systems, the user wears stereo goggles and headset to see and hear the environment. It wears a computerized display and gloves that contains hand position sensors to immersed in the virtual system. All these equipments accept and transmit the real time information back to the computer.

Virtual world has many applications in education, entertainment, scientific and business work. For example, Virtual world helps surgeons to reduce bleeding and trauma during brain tumors surgery. The three-dimensional modeling also helps financial decision makers in analyzing the data in a better way by using visual, spatial and aural immersion, virtual systems. For example, they can visualize the stock futures with different colors, hue and intensity which indicates the deviation from current share prices. Other information such as current trends or debt/equity ratio can be conveyed using sound effects.

5.2.1 Social Networks and Business Intelligence

Q10. Define :

(a) Social Network

(b) Social Network Analysis Software

Ans :

(a) Social Network

Social networking is built on the idea that there is structure to how people know each other and interact. The basic premise is that social networking gives people the power to share, making the world more open and connected. Although social networking is usually practiced in social networks such as MySpace and Facebook, aspects of it are also found in Wikipedia and YouTube.

Definition of Social Network

A social network is a place where people create their own space, or homepage, on which they write blogs (Web logs); post pictures, videos, or music; share ideas; and link to other Web locations they find interesting. In addition, members of social networks can tag the content they create and post it with key words they choose themselves, which makes the content searchable. The mass adoption of social networking Web sites points to an evo-lution in human social interaction.

The Size of Social Network Sites

Social network sites are growing rapidly, with some having over 100 million members. The typical annual growth of a successful site is 40 to 50 percent in the first few years and 15 to 25 percent thereafter. For a list of the major sites, including user counts.

(b) Social Network Analysis Software

Social network analysis software is used to identify, represent, analyze, visualize, or simulate network nodes (e.g., agents, organi-zations, or knowledge) and edges (relationships) from various types of input data (relational and nonrelational), including mathematical models of social networks. Various input and output file formats exist.

Network analysis tools enable researchers to investigate representations of networks of different

forms and different sizes, from small (e.g., families, project teams) to very large. Visual representations of social networks are popular and important to understand network data and to convey the results of the analysis.

Some of the representative tools that enable such presentations are:

- Business-oriented social network tools such as InFlow and NetMiner
- Social Networks Visualizer, or SocNetV, which is a Linux-based open-source package

Social networking is strongly related to mobile devices and networks.

Q11. Describe briefly about Mobile Social Networking.

Ans : **(June-19, June-18, Imp.)**

Mobile social networking refers to social networking where members converse and connect with one another using cell phones or other mobile devices. The current trend for social networking Web sites such as MySpace and Facebook is to offer mobile services. Some social networking sites offer mobile only services (e.g., Brightkite and Fonll).

There are two basic types of mobile social networks. The first type is companies that partner with wireless carriers to distribute their communities via the default start pages on cell phone browsers. For example, users can access My Space via AT&T's wire-less network. The second type is companies that do not have such carrier relationships (also known as "off deck") and rely on other methods to attract users. Examples of this second type include MocoSpace.

Windows Live Spaces Mobile can be viewed on mobile devices with limited screen size and slow data connections. It allows users to browse and add photos, blog entries, and comments directly from their mobile devices. However, it has also introduced several other features to improve the user experience with handheld devices,

Mobile social networking is much more popular in Japan, South Korea, and China than it is in the West, generally due to better mobile networks and data pricing (flat rates are widespread in Japan). The explosion of mobile Web 2.0 services and companies means that many social networks can be based from cell phones and other portable devices, extending the reach of such networks to the millions of people who lack regular or easy access to computers.

With the current software that is available, interactions within mobile social networks are not limited to exchanging simple text messages on a one-to-one basis. In many cases, they are evolving toward the sophisticated interactions of Internet virtual communities.

Mobile Enterprise Networks

Several companies have developed (or fully sponsor) mobile-based social networks. For example, in 2007, Coca-Cola created a social network that could be accessed only by cell phones in an attempt to lure young people to its sodas and other products.

Major Social Network Services: Facebook and Orkut

Now that you are familiar with social network services, let's take a closer look at some of the most popular ones.

A) Facebook : The Network Effect

Facebook (facebook.com), which was launched in 2004 by former Harvard student Mark Zuckerberg, is the second-largest social network service in the world, with more than 200 million active users worldwide as of April 2009. When Zuckerberg first created Facebook, he had very strong social ambitions and wanted to help people connect to others on the Web.

A primary reason why Facebook has expanded so rapidly is the network effect - more users means more value. As more users become involved in the social space, more people are available to connect with. Initially, Facebook was an online social space for college and high school students that automatically connected students to other students at the same school. However, Facebook realized that it could only keep college and uni-versity users for 4 years. In 2006, Facebook opened its doors to anyone aged 13 or older with a valid e-mail address. Expanding to a global audience has enabled Facebook to compete directly with MySpace.

Today, Facebook has a number of applications that support photos, groups, events, marketplaces, posted items, and notes. Facebook also has an application called "People You May Know," which helps users connect with people they might know. More applica-tions are being added constantly. A special feature on Facebook is the News Feed, which enables users to track the activities of friends in their social circles. For example, when a user changes his or her profile, the updates are broadcast to others who subscribe to the feed. Users can also develop their own applications or use any of the millions of Facebook applications that have been developed by other users.

B) Orkut : Exploring the very Nature of Social Networking Sites

Orkut **(orkut .com)** was the brainchild of a Turkish Google programmer of the same name. Orkut was to be Google's homegrown answer to MySpace and Facebook. Orkut follows a format similar to that of other major social networking sites; a homepage where users can dis-play every facet of their personal life they desire using various multimedia applications.

A major highlight of Orkut is the individual power afforded to those who create their own groups and forums, which are called "communities." Who can join and how posts are edited and controlled lie solely in the hands of the creator of each community. Moderating an Orkut community is comparable to moderating one's own Web site, given the authority the creator possesses with regard to design and control of content. Orkut users gain substantial experience with Web 2.0 tools, creating an enormous wave of online proficiency, which is sure to contribute to the development of the online environment.

Q12. What are the Implications of Business and Enterprise Social Networks?

Ans :

Although advertising and sales are the major electronic commerce activities in public social networks, there are emerging possibilities for commercial activities in business- oriented networks such as Linkedln and in enterprise social networks.

Recognizing the opportunities, many software vendors are developing Web tools and applications to support enterprise social networking. For example, IBM Lotus is encouraging its 5,000-plus solution providers who are working with Notes/ Domino, Sametime, and other Lotus software to add Lotus Connections to their product lineups, building applications based on social networking technology.

Representative areas and examples of enterprise social networking follow.

1. Finding And Recruiting Workers

Most of the public social networks, especially the business-oriented ones, facilitate recruiting and job finding. For example, recruiting is a major activity at Linkedln and was the driver for the site's development. To be competitive, companies must look at the global market for talent, and they can use global social networking sites to find it. Large companies are using their in-house social networks to find in-house talent for vacant positions.

2. Management Activities and Support

Applications in this category are related to supporting managerial decision making based on analysis of data collected in social net-works. Some typical examples include identifying key performers, locating experts and finding paths to access them, soliciting ideas and possible solutions to complex prob-lems, and finding and analyzing candidates for management succession planning.

3. Training

Several companies use enterprise social networking, and virtual worlds in particular, for training purposes. For example, Cisco is trying to use its virtual campus in Second Life for product training and executive briefings. IBM runs management and cus-tomer interaction training sessions in Second Life too.

4. Knowledge Management and Expert Location

Applications in this category include activities such as knowledge discovery, creation, maintenance, sharing, transfer, and dissemination.

An elaborate discussion on the role of discussion forums, blogs, and wikis for conversational knowledge management can be found in Wagner and Bolloju. Other examples of these applications include expert discovery and mapping com-munities of expertise.

5. Enhancing Collaboration

Collaboration in social networking is done both inter-nally, among employees from different units working in **virtual teams** for example, and externally, when working with suppliers, customers, and other business partners. Collaboration is done mostly in forums and other types of groups and by using wikis and blogs.

Using Blogs and Wikis Within the Enterprise

The use of these tools is expanding rapidly. Jefferies (2008) reports on a study that 71 percent of the best-in-class companies use blogs and 64 percent use wikis for the following applications:

- Project collaboration and communication (63%)
- Process and procedure document (63%)
- FAQs (61%).
- E-learning and training (46%)
- Forums for new ideas (41%)
- Corporate-specific dynamic glossary and terminology (38%)
- Collaboration with customers (24%).

5.2.2 Collaborative Decision Making

Q13. Describe briefly about Collaborative Decision Making.

Ans : **(June-19)**

Introduction

Collaborative Business Intelligence (BI) is a relatively new concept in which the two technologies of business intelligence and collaboration are beginning to merge in support of an improved decision- making environment. We define it as:

"The combination of business intelligence and collaborative computing processes and technologies with the objective of enhancing business decision-making. This is achieved by improving user communication and information sharing, adding business knowledge to BI results, and making collaborative decisions."

From this definition, we see that collaborative BI fulfills three main business requirements.

- **Collaborative Interaction:** Business people want to interact with each other to improve overall communication. They must be able to easily discover the BI analytics available to them and then share them with their team. Finally they must be able to discuss the meaning of the analytics and the decision options to improve, reduce or maintain the analytic indicators. The ease of discovery and the documentation of the ensuing discussions are major features of truly collaborative BI environments.
- **Information enhancement:** There is a need to enhance the information resulting from BI processing. To do this, business people must be able to add their own knowledge to the BI results. They can provide information about why events are unfolding the way they are, discover related content to the work at hand, and provide the necessary and important business context surrounding the BI results. For example, a trend may be heading down. Is this good news or bad? Adding team expertise and knowledge about the trend makes it clear that the trend is good – customer churn is lower than the same time last year. Published BI results can be enhanced through feedback mechanisms such as ratings, comments and tagging, and in some cases, blogs and microblogs.
- **Collaborative decision-making:** The last step is of course to decide on the appropriate action(s) to be taken based on the two previous activities. The collaborative BI environment must be able to track these decisions and later be able to analyze the accuracy or impact of them. This analysis produces feedback to the teams in terms of how they can improve upon the entire process. It enables best practices to be recorded and allows information workers to

track the types of information that provide useful content for decision making. Collaborative decision-making should also include the analysis of social networks and other informal teams to determine influencers of communities, who has needed expertise, and who makes the decisions in these informal networks.

(a) The Rise of Collaborative Decision Making

In a major report, Gartner Inc.'s researchers (Schlegel et ah, 2009) describe that a new style of decision support, collaborative decision making (CDM), which combines social software and BI, is emerging, and it may dramatically improve the quality of decision making by directly linking the information contained in BI systems with collaborative input gleaned through the use of social software.

The key findings of this report are :

- CDM is a category of decision-support system for nonroutine, complex decisions that require iterative human interactions.
- Ad hoc tagging regarding value, relevance, credibility, and decision context can substantially enrich both the decision process and the content that contributes to the decisions.
- Tying BI to decisions and outcomes that can be measured will enable organizations to better demonstrate the business value of BI.

Despite unprecedented information availa-bility, the past decade suffered from several imperfect decisions made in both the public and the private sectors. It is not enough to provide voluminous access to information and expect good decisions to be made as a result. Numerous social, cultural, and educational factors influence how well individuals and organizations are able to improve their decision-making ability, and need to be included in the analysis. CDM can rectify the deficiency in decision making by adding these missing factors.

(b) Collaboration in Virtual Teams' Decision Making

Restrictions on travel because of the difficult economy force many companies to find a new way to work, collaborate, and make decisions. Gartner Inc. researchers believe that the information technology market responds to the need of collaboration in virtual teams by creating systems that foster a CDM process using social software. The business application of social software techniques pioneered by consumer-driven social networks services, such as Facebook and Myspace, is well under way.

Organizations already use collaborative social software to keep informed about where colleagues are and what they are doing and thinking, and to mobilize them for urgent meetings to address problems. Designing collaborative environments that enable decision makers to discuss an issue, brainstorm options, evaluate their pros and cons, and agree on a course of action is a natural evolution of the above trend. Adding social software elements (such as tagging, recommendations, ratings, and contextual information) enriches the collaborative environment, making it (and the results derived from it) more useful.

CDM enables BI systems to tie information modeled in BI systems directly to the decision made in the collaborative environment. BI systems have traditionally been overtly disconnected from the business process. As a result, it is sometimes difficult to see the business value of BI, even in the most insightful reports and analyses. In addition, decision making has been considered an unstructured process that is not repeatable, so there is a lack of tools that facilitate the work of the decision makers. Figure 6.5 shows how CDM tools support the decision-making process.

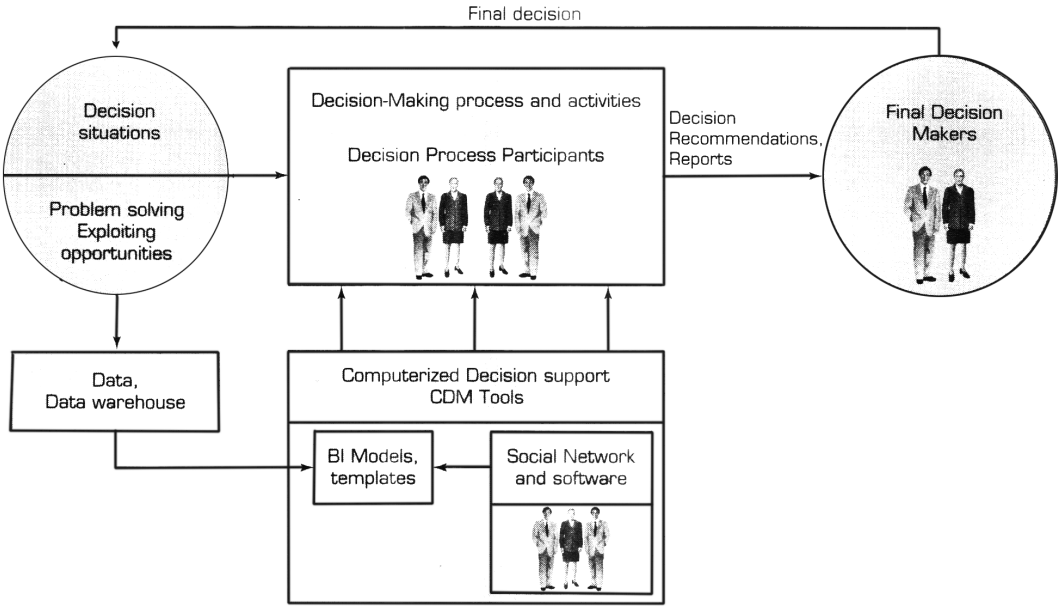


Fig. : The Framework for Collabortive Decision Making

CDM Working Mechanism

Combining BI with collaboration and social software provides a direct way to show the value of BI because the analytical insights and measures are linked to business decisions and framed in the social context. For example, in making an investment decision, users can rate their assumptions about future revenue, expense, or interest rate and compare the results of those predictions with a key per-formance indicator (KPI) that measures sales and profit.

The BI platform could then update the prediction model with actual interim results on the appropriate KPIs, help-ing the users (participants in the KPI) exceed a critical threshold, requiring a rethink-ing of a decision. Today, collaboration on decisions that are more strategic in nature than ever before—decisions that may transform the business—involves nonroutine activities such as brainstorming discovering, innovating, creating and leading teams, learning, and relating.

The outcome of manual processes of such decision can easily be lost or become part of corporate folklore in the form of anecdotes, with no formal process for decision audit, assessment, and closed-loop learning. Clearly, this is an area in need of an information system to facilitate this manual process, and CDM can be an ideal mechanism.

5.2.3 RFID and Business Intelligence

Q14. Define RFID. Explain the working mechanism of RFID.

Ans :

RFID is a generic technology that refers to the use of radio frequency waves to identify objects. Fundamentally, RFID is one example of a family of automatic identifica-tion technologies, which also include the ubiquitous bar codes and magnetic strips. Since the mid-1970s, the retail supply chain (and many other areas) has used bar codes as the primary form of automatic identification. The potential advantages of RFID have prompted many companies (led by large retailers such as Wal-Mart, Target, and Albertson's) to aggressively pursue this technology as a way to improve their supply chain and thus reduce costs and increase sales.

Working Mechanism of RFID

In its simplest form, an RFID system consists of a tag (attached to the product to be identified), an interrogator (i.e., reader), one or more antennae attached to the reader, and a computer (to control the reader and capture the data). At present, the retail supply chain has primarily been interested in using pas-sive RFID tags. *Passive tags* receive energy from the electromagnetic field created by the interrogator (e.g., a reader) and backscatter information only when it is requested. The passive tag will remain energized only while it is within the interrogator's magnetic field.

In contrast, active tags have a battery on board to energize them. Because active tags have their own power source, they don't need a reader to energize them; instead they can initiate the data transmission process on their own. On the positive side, active tags have a longer read range, better accuracy, more complex rewritable information stor-age, and richer processing capabilities.

On the negative side, due to the battery, active tags have a limited lifespan, are larger in size than passive tags, and are more expensive. Currently, most retail applications are designed and oper-ated with passive tags. Active tags are most frequently found in defense or military sys-tems, yet they also appear in technologies such as EZ Pass, where tags are linked to a prepaid account, enabling drivers to pay tolls by driving past a reader rather than stopping to pay at a tollbooth.

The most commonly used data representation for RFID technology is the Electronic Product Code (EPC), which is viewed by many in the industry as the next generation of the Universal Product Code (UPC) (most often represented by a bar code). Like the UPC, the EPC consists of a series of numbers that identifies product types and manufacturers across the supply chain. The EPC code also includes an extra set of digits to uniquely identify items.

Currently, most RFID tags contain 96 bits of data in the form of serialized global trade identification numbers (SGTIN) for identifying cases or serialized shipping con-tainer codes (SSCC) for identifying pallets (although SGTINs can also be used to identify pallets). The complete guide to tag data standards can be found on EPCglobal's Web site.

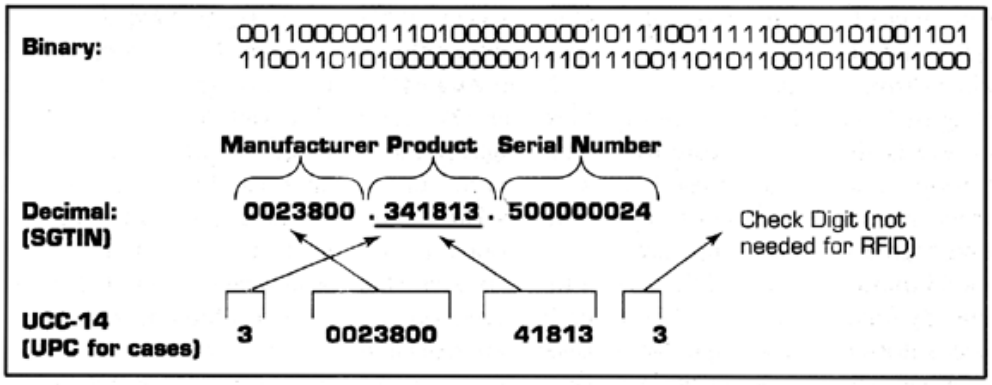


Fig.: Sample RFID Tag Data

EPCglobal, Inc. is a subscriber-driven organization of industry lead-ers and organizations focused on creating global standards for the EPC to support the use of RFID.

The various application of RFID technology are,

1. Inventory management plays an important role in retail business and this task can be performed effectively by RFID system. RFID provides alarm facility to the store keeper which helps them in maintaining appropriate stock and this results in higher customer satisfaction and huge business profits.

2. RFID system can be used in blood banks to avoid general human mistakes. Usually, in blood banks most of the blood bottles consist of similar labels which may cause confusion to the hospital management. In order to avoid this risk, RFID tags can be attached to the blood packets which helps the medical staff in identifying correct blood group.
3. RFID tags are used in stores/malls to avoid theft of some expensive items. RFID tags can be attached to the items or can be kept near the exist door. RFID system provide alarm facility through which one can identify the items missing in the store and can also avoid theft.
4. RFID technology can be used in hospitals to locate patient, doctors, nurses or other staff.
5. Apart from the above RFID applications, it is also used in animals. A small microchip is inserted into the animals skin which gives information related to animal and also its owner.

Q15. What are the advantages and disadvantages of RIFD?

Ans :

Advantages of RFID

1. RFID is a flexible identification device can be moved easily.
2. RFID tags can be read-only or read/write. The technology of RFID is versatile in nature.
3. Implementation of RFID technology results in accuracy and speed of data collection.
4. RFID tags are robust and can be used in any environments and temperatures.
5. RFID enables transmission and reception of radio signals. Since, it uses radio frequency tags and reader.
6. RFID tags can be used in stores to prevent theft, there by providing the facility of protecting expensive items in the store.
7. RFID tags have greater storage capacity.
8. RFID enables good tracking system.

Disadvantages of RFID

1. RFID technology is highly expensive.
2. If RFID tags are in liquids and metal products then, it is highly difficult to read.
3. Range of frequency differs from one country to another and it is difficult to known the working pattern of other countries where RFID tags perform their activities. It takes a lot of time.
4. At times RFID radio waves are disturbed.
5. RFID technology is considered as invasive technology.

5.2.4 Reality Mining

Q16. What do you understand by reality mining ?

Ans : **(June-19)**

- A new emerging technology is Reality Mining, fastly gearing up the phase which provides sense of data. It is a massive data source similar to RFID for developing data streams and for making analysis using business intelligence. This eases the process of decision making.

- Ideally, Reality mining can be specified as a collection and examination of data resulted from machine-sensed environment.
- This data is related to human social behavior for determining predictable patterns of behavior.
- The Reality mining develops a precise picture of the where about of people, keep tracks of their contacts by using Wireless devices such as mobile phones and GPS systems.
- Interestingly, the Big data is incorporated in Reality mining for boosting the process of research and analysis of people's interaction with technology. This effectively forms a positive transition from an individual to global community.
- Therefore, they advances there data analytics to a different type of data which includes real-time behavior of people working in groups which could be movement and spatial positioning.
- Many businesses and customers effectively implements location-enabled devices which spreads location information through the cars, buses, taxis, mobile phones, cameras and personal navigation devices.
- These devices operates upon network-connected positioning technologies which are GPS, wi-fi and cell-tower triangulation.
- The growth corresponding to location-enabled services produces huge database related to historical and real-time streaming location information.
- So, subsequently keeping in mind all above, reality mining is constructed on the concept of such datasets which facilitates real-time insight with in the human-activities.
- These large scale patterns serves as a basis for understanding and analyzing various classes of behaviors with in the particular context which is called as 'tribes'.
- In particular, the macrosense is an application created by sense Networks. This accepts the input from all mobile devices and classifies the received data streams based on the type of customer and client. This is fulfilled once proprietary clustering algorithms are applied. Consecutively, it supports the investors in better understanding the customer patterns.
- A new application called citysense is coming into picture which is adapted by sense Networks. This technology is basically used for customers in searching people who keeps similar interests.
- The citysense outlook can be seen in sensenetworks.com /citysense.php. The layout is in the form of black and white image displaying the patterns and roaming locations of the people in and out of the city. Also, the macrosense, which is a sense
- Networks's primary analytical platform examines the complete information present in the citysense for combining users and identifying tribes. It does this by performing the sampling in distribution of tribes at any point of time.
- This could be further explain with the supporting idea where clubs and hip-hop clubs own their unique tribal distribution.
- So, the moment user checks in, the microsense initiates the selected tribe distribution from the time spent in the club. Future enhancement could be the inclusion of hotspots based on the distribution and complete activity information.
- Once, the algorithms are applied, the dimensionality of the location of data gets minimized and reality mining initiates the characterization of locations based on the user activity and movement starting from huge amounts of high-dimensional location data.
- The algorithm unleashes the styles, meaning and relationships for delivering representation in human understandable form. This eventually makes the data suitable for intelligent predictions and also determining significant matches and similarities among places and people.
- On the other hand, these technologies presents lot of problems in terms of privacy. That is, any intruder can crack the location of the cell phone. Thus, compromising the privacy of the user.

Short Question and Answers

1. Define Business Intelligence.

Ans :

Implementing BI systems can be very complex. In addition to typical issues in information system implementation, such as conducting appropriate cost-benefit analysis with intangible variables to justify the system and dealing with resistance to change, there are some complex issues related to integration, security, scalability of the system, and handling the construction of a data warehouse, analytics, and dashboards.

2. Critical Success factors of Business Implementation.

Ans :

Although there could be many factors that could affect the implementation process of a BI system, a report from Vodapalli (2009) as cited on Wikipedia shows that the following are the critical success factors for a business intelligence implementations.

a) Business driven methodology and project management.

b) Clear vision and planning.

c) Committed management support and sponsorship.

d) Data management and quality issues.

e) Mapping the solutions to the user requirements.

f) Performance considerations of the BI system.

g) Robust and extensible framework.

3. Virtual worlds

Ans :

Virtual world is a technology that creates an artificial environment that responds to, and is controlled by the behavior of the user. It is created by using interactive graphics software which is so closed to reality that people feel this artificial reality really does exist. The same computer-generated artificial environment can be shared by more than one person and even a large group. In many virtual reality systems, the user wears stereo goggles and headset to see and hear the environment. It wears a computerized display and gloves that contains hand position sensors to immersed in the virtual system. All these equipments accept and transmit the real time information back to the computer.

Virtual world has many applications in education, entertainment, scientific and business work. For example, Virtual world helps surgeons to reduce bleeding and trauma during brain tumors surgery. The three-dimensional modeling also helps financial decision makers in analyzing the data in a better way by using visual, spatial and aural immersion, virtual systems. For example, they can visualize the stock futures with different colors, hue and intensity which indicates the deviation from current share prices. Other information such as current trends or debt/equity ratio can be conveyed using sound effects.

4. Social Network

Ans :

Social networking is built on the idea that there is structure to how people know each other and interact. The basic premise is that social networking gives people the power to share, making the world more open and connected. Although social networking is usually practiced in social networks such as MySpace and Facebook, aspects of it are also found in Wikipedia and YouTube.

Definition of Social Network

A social network is a place where people create their own space, or homepage, on which they write blogs (Web logs); post pictures, videos, or music; share ideas; and link to other Web locations they find interesting. In addition, members of social networks can tag the content they create and post it with key words they choose themselves, which makes the content searchable. The mass adoption of social networking Web sites points to an evo-lution in human social interaction.

5. Social Network Analysis Software

Ans :

Social network analysis software is used to identify, represent, analyze, visualize, or simulate network nodes (e.g., agents, organi-zations, or knowledge) and edges (relationships) from various types of input data (relational and nonrelational), including mathematical models of social networks. Various input and output file formats exist.

Network analysis tools enable researchers to investigate representations of networks of different forms and different sizes, from small (e.g., families, project teams) to very large. Visual representations of social networks are popular and important to understand network data and to convey the results of the analysis.

Some of the representative tools that enable such presentations are:

- Business-oriented social network tools such as InFlow and NetMiner
- Social Networks Visualizer, or SocNetV, which is a Linux-based open-source package

Social networking is strongly related to mobile devices and networks.

6. Collaborative Decision Making.

Ans :

Collaborative Business Intelligence (BI) is a relatively new concept in which the two technologies of business intelligence and collaboration are beginning to merge in support of an improved decision- making environment. We define it as:

"The combination of business intelligence and collaborative computing processes and technologies with the objective of enhancing business decision-making. This is achieved by improving user communication and information sharing, adding business knowledge to BI results, and making collaborative decisions."

7. RFID

Ans :

RFID is a generic technology that refers to the use of radio frequency waves to identify objects. Fundamentally, RFID is one example of a family of automatic identifica-tion technologies, which also include the ubiquitous bar codes and magnetic strips. Since the mid-1970s, the retail supply chain (and many other areas) has used bar codes as the primary form of automatic identification. The potential advantages of RFID have prompted many companies (led by large retailers such as Wal-Mart, Target, and Albertson's) to aggressively pursue this technology as a way to improve their supply chain and thus reduce costs and increase sales.

8. Application of RFID

Ans :

1. Inventory management plays an important role in retail business and this task can be performed effectively by RFID system. RFID provides alarm facility to the store keeper which helps them in maintaining appropriate stock and this results in higher customer satisfaction and huge business profits.
2. RFID system can be used in blood banks to avoid general human mistakes. Usually, in blood banks most of the blood bottles consist of similar labels which may cause confusion to the hospital management. In order to avoid this risk, RFID tags can be attached to the blood packets which helps the medical staff in identifying correct blood group.
3. RFID tags are used in stores/malls to avoid theft of some expensive items. RFID tags can be attached to the items or can be kept near the exist door. RFID system provide alarm facility through which one can identify the items missing in the store and can also avoid theft.

9. Advantages of RFID

Ans :

1. RFID is a flexible identification device can be moved easily.
2. RFID tags can be read-only or read/write. The technology of RFID is versatile in nature.
3. Implementation of RFID technology results in accuracy and speed of data collection.
4. RFID tags are robust and can be used in any environments and temperatures.

5. RFID enables transmission and reception of radio signals. Since, it uses radio frequency tags and reader.

6. RFID tags can be used in stores to prevent theft, there by providing the facility of protecting expensive items in the store.

7. RFID tags have greater storage capacity.

8. RFID enables good tracking system.

10. Reality mining

Ans :

A new emerging technology is Reality Mining, fastly gearing up the phase which provides sense of data. It is a massive data source similar to RFID for developing data streams and for making analysis using business intelligence. This eases the process of decision making.

Ideally, Reality mining can be specified as a collection and examination of data resulted from machine-sensed environment.

This data is related to human social behavior for determining predictable patterns of behavior.

The Reality mining develops a precise picture of the where about of people, keep tracks of their contacts by using Wireless devices such as mobile phones and GPS systems.

Interestingly, the Big data is incorporated in Reality mining for boosting the process of research and analysis of people's interaction with technology. This effectively forms a positive transition from an individual to global community.

Therefore, they advances there data analytics to a different type of data which includes real-time behavior of people working in groups which could be movement and spatial positioning.

Choose the Correct Answers

1. Factor of BI implementation _____ [d]

   (a) Database (b) ETL

   (c) Cost involved (d) All

2. Managerial issue of BI implementation _____ [a]

   (a) Cost Justification (b) Web 2.0

   (c) Social Networks (d) All

3. Example of mobile enterprise networks _____ [a]

   (a) Facebook (b) Twitter

   (c) Instagram (d) All

4. Disadvantage of RFID _____ [a]

   (a) High exepensive (b) Social benefits

   (c) Reality Mining (d) Second Life

5. The technology that enables a user to perform interaction with computer simulated environment irrespective of its physical presence is [a]

   (a) Virtual World (b) BI

   (c) Software Integration (d) Web 2.0

6. The BI implementation of _____ saves the time and improves the operational efficiency. [d]

   (a) Physical Integration (b) Social Network

   (c) Business Analytics (d) Business Intelligence

7. Revolution of B I [a]

   (a) Web 2.0 (b) Web 2.1

   (c) Web 2.2 (d) Web 2.2

8. _____ is a social networking which allows the users with common interest to collaborate and converse using mobile phones. [a]

   (a) Mobile Social Networking (b) Social Networking

   (c) Mobile Community (d) Mobile Enterprise Networks

9. CDM Stand for _____ [a]

(a) Collaborative decision making (b) Software Integration

(c) Support Activity (d) All

10. Need of Integration in BI _____ [a]

(a) Real time decision support (b) Bar Coding

(c) Data Mining (d) All

Fill in the blanks

1. RFID Stand for _____
2. What is the type of BI Integration _____
3. _____ provides various light weight programming techniques.
4. _____ has helped many people in conducting business and earning profits.
5. Emerging Trend in BI implementation _____
6. CDM stand for _____
7. The two RFID tags are _____ and _____.
8. What is the critical success factor of BI implementaion _____
9. ETL stand for _____
10. _____ software basically provides quantitative and qualitative analysis.

ANSWER

1. Radio Frequency Identification
2. Functional Integration
3. Web 2.0
4. Second Life
5. Advanced Analytics
6. Collaborative Decision Making
7. Active tag, Passive tag
8. Data Management and Quality Issues
9. Extaction, Transformation and Load (ETL)
10. Social network analysis.

FACULTY OF MANAGEMENT

M.B.A IV - Semester (CBCS) Examination

October - 2020

BUSINESS INTELLIGENCE

Time : 2 Hours] [Max. Marks : 80

PART - A (4 × 5 = 20 Marks)

Note: Answer any four questions.

ANSWERS

1. Business Intelligence **(Unit-I, SQA - 1)**
2. Data Warehousing (DWH) **(Unit-II, SQA - 1)**
3. Business Performance **(Unit-III, SQA - 1)**
4. Data Visualization **(Unit-IV, SQA - 5)**
5. Reality Mining **(Unit-V, SQA - 10)**

PART - B (4 × 15 = 60 Marks)

Note: Answer any four questions.

6. What is Real-Time Business Intelligence? Explain Business Intelligence Architecture. **(Unit-I, Q.No.13, 16)**
7. Write the various Business Intelligence styles with suitable examples. **(Unit-I, Q.No.6)**
8. Explain the Data Warehousing Architectures Models. **(Unit-II, Q.No.7)**
9. What are the Data Mining Functions and Applications in Business? **(Unit-II, Q.No. 20, 22)**
10. What are the steps included in Business Performance Process? **(Unit-III, Q.No.5)**
11. Explain the Applications of Business Performance through Dash Boards. **(Unit-III, Q.No.23)**
12. Explain the Usage and Benefits of Web in Business Analytics. **(Unit-IV, Q.No.8)**
13. What are the techniques of Advanced Business Analytics? **(Unit-IV, Q.No.5)**
14. Explain the factors of Business Intelligence effecting Implementation Stage. **(Unit-V, Q.No.2)**
15. What are the emerging trends in Social Networks and Business Intelligence ? **(Unit-V, Q.No.7)**

FACULTY OF MANAGEMENT

M.B.A IV - Semester (CBCS) Examination

May / June - 2019

BUSINESS INTELLIGENCE

Time : 3 Hours] [Max. Marks : 80

PART - A (5 × 4 = 20 Marks)

[Short Answer type]

Note : Answer all the questions from Part - A and Part - B.

Each question carries 4 marks in Part - A and 12 marks in Part - B.

ANSWERS

1. Business Architecture **(Unit-I, SQA - 6)**
2. Data Mining **(Unit-I, SQA - 5)**
3. BPM Architecture **(Unit-III, SQA - 9)**
4. GPS **(Unit-IV, SQA - 8)**
5. Reality Mining **(Unit-V, SQA - 10)**

PART - B (5 × 12 = 60 Marks)

[Essay Answer type]

6. (a) What is Real Time business intelligence ? Explain its role in business performance. **(Unit-I, Q.No. 13)**

OR

(b) What are the benefits of business intelligence ? **(Unit-I, Q.No. 7)**

7. (a) Explain various types of data warehousing methods. **(Unit-II, Q.No. 3)**

OR

(b) Explain in detail the characteristics and benefits of data mining. **(Unit-II, Q.No. 18, 21)**

8. (a) Compare and contrast between business performance measurement and business intelligence. **(Unit-III, Q.No. 3)**

OR

(b) Explain the architecture of business performance measurement. **(Unit-III, Q.No. 20)**

9. (a) What is business analytics ? Explain various tools and techniques of business analytics. **(Unit-IV, Q.No. 1, 3)**

OR

(b) Explain various methods of data visualization. **(Unit-IV, Q.No. 15)**

10. (a) What are the critical success factors for business intelligence implementation? **(Unit-V, Q.No. 2)**

OR

(b) Explain in detail the emerging trends in business intelligence implementation. **(Unit-V, Q.No. 11, 13, 14, 16)**

FACULTY OF MANAGEMENT

M.B.A IV - Semester (CBCS) Examination

May / June - 2018

BUSINESS INTELLIGENCE

Time : 3 Hours] [Max. Marks : 80

PART - A (5 × 4 = 20 Marks)

[Short Answer type]

ANSWERS

1. Value Chain **(Unit - I, SQA - 5)**
2. Data Integration **(Unit - II, SQA - 3)**
3. Business Performance **(Unit - III, SQA - 1)**
4. GIS **(Unit - IV, SQA - 7)**
5. Business Intelligence **(Unit - V, SQA - 1)**

PART - B (5 × 12 = 60 Marks)

[Essay Answer type]

6. (a) Briefly explain the history and evolution of Business Intelligence. **(Unit - I, Q.No. 4)**

OR

(b) Explain the role of business intelligence in organizational development. **(Unit - I, Q.No. 12)**

7. (a) What is data warehousing ? Briefly explain the frame work for data warehousing. **(Unit - II, Q.No. 1, 6)**

OR

(b) Describe different tools and techniques of data mining. **(Unit - II, Q.No. 23)**

8. (a) Discuss the process of business performance measurement. **(Unit - III, Q.No. 4)**

OR

(b) Explain the various applications of business performance measurement. **(Unit - III, Q.No. 21)**

9. (a) Explain the benefits of Business Analytics. **(Unit - IV, Q.No. 8)**

OR

(b) Compare and contrast between GIS and GPS. **(Unit - IV, Q.No. 18)**

10. (a) Briefly explain the managerial issues related to business intelligence implementation. **(Unit - V, Q.No. 3)**

OR

(b) Explain the role of social networks in business intelligence. **(Unit - V, Q.No. 11)**

FACULTY OF MANAGEMENT

M.B.A IV - Semester (CBCS) Examination

Model Paper

BUSINESS INTELLIGENCE

Time : 3 Hours] [Max. Marks : 80

PART - A (5 × 4 = 20 Marks)

[Short Answer type]

ANSWERS

1. Various stages of business intelligence. **(Unit - I, SQA - 3)**
2. Define Data Warehousing. **(Unit - II, SQA - 1)**
3. Define business performance measurement. **(Unit - III, SQA - 2)**
4. Geographical Information System. **(Unit - IV, SQA - 7)**
5. Advantages of RFID **(Unit - V, SQA - 9)**

PART - B (5 × 12 = 60 Marks)

[Essay Answer type]

6. a) Briefly explain the various benefits of Business Intelligence. **(Unit - I, Q.No. 7)**

OR

b) Elucidate the concept of Real-time Business Intelligence. **(Unit - I, Q.No. 13)**

7. a) What are the characteristics of data ware housing? **(Unit - II, Q.No. 2)**

OR

b) Define (i) Cluster Analysis (ii) Association Rule Mining **(Unit - II, Q.No. 24)**

8. a) Outline the process of business performance Measurement. **(Unit - III, Q.No. 4)**

OR

b) Elucidate the Architecture of Business Performance Measurement. **(Unit - III, Q.No. 20)**

9. a) Explain briefly about Vendor Support for Web Analytics. **(Unit - IV, Q.No.7)**

OR

b) What is Data visualization? Explain the importance of data visualization? **(Unit - IV, Q.No.18)**

10. a) Explain the Managerial Issues Related to Business Intelligence Implementation ? **(Unit - V, Q.No.3)**

OR

b) What do you understand by reality mining ? **(Unit - V, Q.No.16)**

FACULTY OF MANAGEMENT

M.B.A IV - Semester (CBCS) Examination

July / August - 2021

BUSINESS INTELLIGENCE

Time : 2 Hours] [Max. Marks : 80

PART - A (4 × 5 = 20 Marks)

Note: Answer any four questions.

1. Define Business Intelligence **(Unit-I, SQA-1)**
2. Data Warehousing integration
3. Performance dash boards
4. GIS vs. GPS
5. Reality Mining

PART - B (4 × 15 = 60 Marks)

Note: Answer any four questions.

6. Explain benefits of Business Intelligence in detail.
7. Explain in detail about Business Intelligence Value Chain.
8. What are the sources and types of data for a Data Warehouse?
9. Describe characteristics and benefits of Data Mining.
10. Explain in detail about BPM Architecture.
11. Explain in detail about Business Performance Measurement (BPM) process.
12. Define Business Analytics (BA). Discuss various tools and techniques of BA.
13. Define Data Visualization. Write about new direction in Data Visualization.
14. Explain about Managerial Issues related to BI Implementation.
15. Discuss impact and application of BI on Social networks, RFID and collaborative decision making.