MBA SEMESTER 2

Business Analytics- 4529201

MODULE 1

Business Intelligence:

☐ Definitions and Examples in Business Intelligence

- Business Intelligence is the processes, technologies, and tools that help us change data into information, information into knowledge and knowledge into plans that guide organization

- Technologies for gathering, storing, analyzing and providing access to data to help enterprise users make better business Decisions

- BI Makes companies smarter.

- It is a variety of software applications used to analyze an organization's raw data.

- BI can include data mining, online analytical processing, and business reporting.

- BI is about to delivering relevant and reliable information to the right people at right time to with the goal achieving better decision faster.

- It is a variety of software applications used to analyze an organization's raw data. BI can include data mining, online analytical processing, and business reporting.

- BI helps to make strategic, tactical and operational decisions by providing right information to right person at right place at right time.

What is BI?

What are BI Techniques?

Business intelligence technologies use advanced statistics and predictive analytics to help businesses draw conclusions from data analysis, discover patterns, and forecast future events in business operations. Business intelligence reporting is not a linear practice, rather, it is a continuous, multifaceted cycle of data access, exploration, and information sharing. Common business intelligence functions include:

- Data mining: sorting through large datasets using databases, statistics, and machine learning to identify trends and establish relationships

- Querying: a request for specific data or information from a database

- Data preparation: the process of combining and structuring data in order to prepare it for analysis

- Reporting: sharing operating and financial data analysis with decision-makers so they can draw conclusions and make decisions

- Benchmarking: comparing current business processes and performance metrics to historical data to track performance against industry bests

- Descriptive analytics: the interpretation of historical data to draw comparisons and better understand changes that have occurred in a business

- Statistical analysis: collecting the results from descriptive analytics and applying statistics in order to identify trends

- Data visualization: provides visual representations such as charts and graphs for easy data analysis

What are BI Tools?

Modern business intelligence systems prioritize self-service analysis, empowering businesses to gain insight into their market and improve performance with comprehensive data discovery tools, methods, processes, and platforms. Such business intelligence solutions include:

- Ad hoc analytics: an analysis process designed to answer specific questions on the spot
- Online analytical processing (OLAP): a computing method that enables multi-dimensional analytical queries
- Mobile BI: software that optimizes desktop business intelligence for mobile devices
- Real-time BI: a data analytics approach that delivers real-time information to users by feeding business transactions into a real-time data warehouse
- Operational BI: a data analysis approach that utilizes real-time business analytics to automatically integrate real-time data into operational system for immediate use
- Software-as-a-service BI (SaaS BI): a cloud-hosted, subscription-based delivery model for business intelligence software solutions
- Open source BI (OSBI): business intelligence software solutions that do not require purchasing a software license

- Collaborative BI: the merging of business intelligence software with collaboration tools in order to streamline the sharing process

- Location intelligence (LI): software that is designed to relate geographic contexts to business data

- Data visualization software: facilitates the detection of patterns and correlations by providing visual context

What are BI Platforms?

A business intelligence platform enables businesses to utilize existing data architecture and create custom business intelligence applications that make information available for analysts to query and visualize. Modern business intelligence platforms support self-service analytics, making it easy for end users to create their own dashboards and reports.

Simple user interfaces combined with flexible business intelligence backend software enables users to connect to a range of data sources, including NoSQL databases, Hadoop systems, cloud platforms, and conventional data warehouses, to develop a cohesive view of their diverse data.

As artificial intelligence and machine learning continue to grow, and as businesses strive to be more data-driven and collaborative, so too does business intelligence continue to evolve, enabling users to integrate AI insights and harness the power of data visualizations. Popular business intelligence platform providers include Oracle, Microsoft, IBM, and Salesforce.

What are the benefits of BI?

The importance of business intelligence continues to grow as businesses face an ever-increasing flow of raw data and the challenges of gaining insight from enormous volumes of information (big data). With the employment of business intelligence systems, businesses can gain a comprehensive view of their organization's data and translate it into insights about their business processes, enabling improved and strategic business decisions.

Business intelligence helps organizations analyze data with a historical context, optimize operations, track performance, accelerate and improve decision-making, identify and eliminate business problems and inefficiencies, identify market trends and patterns, drive new revenues and profitability, increase productivity and accelerate growth, analyze customer behavior, compare data with competitors, and ultimately gain a competitive advantage over rival businesses.

Advantages of Business Intelligence

Business Intelligence has a direct impact on an organization's strategic, tactical and operational business decisions. It supports fact-based decision making using historical data rather than assumptions and gut feelings. These 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. Following are some advantages of Business Intelligence:

- A right Business Intelligence software helps to boost productivity in the organization by displaying the gathered data using reports, analytical dashboards, and infographics.

SHREE H.N.SHUKLA COLLEGE OF MANAGEMENT STUDIES, RAJKOT
AFFILIATED TO GUJARAT TECHNOLOGICAL UNIVERSITY

Lalpari Campus, Near Lalpari Lake,	3 – Vaishalinagar
B/H Marketing Yard,	Nr. Amrapali Railway Crossing
Amargadh (Bhichri), Rajkot	Raiya Road, Rajkot

- It provides an organization with a holistic view of the company, segmented into various departments, products, services, etc., making it easier to identify areas that need attention or improvement.

- Complex processes in an organization can be streamlined using advanced automated analytics, enabling a company to reduce time and effort leading to faster and efficient business processes.

- Visual infographics and easy to understand reports can be generated using BI software allowing even non-technical individuals to understand the meaning of their metrics.

Business Intelligence by other names

- Reporting
- Decision Support System
- Business Analytics
- Business Insight
- Business Investment
- Competitive Intelligence

☐ BI Components

o Data Warehouse

Data warehousing allows the business owner to go through different data subsets and examine components that could help make the right business decisions. For example, warehousing gives a user an opportunity to monitor certain sales information collected on Mondays for the past 50 weeks.

SHREE H.N.SHUKLA COLLEGE OF MANAGEMENT STUDIES, RAJKOT
AFFILIATED TO GUJARAT TECHNOLOGICAL UNIVERSTY

Lalpari Campus, Near Lalpari Lake,	3 – Vaishalinagar
B/H Marketing Yard,	Nr. Amrapali Railway Crossing
Amargadh (Bhichri), Rajkot	Raiya Road, Rajkot

It helps create important statistics about the business and the industry. Warehousing implies storing formidable amounts of data in numerous special ways, which could be useful for analysis.

Different technologies exist to help the user take advantage of data warehousing quickly and effectively.

Technically, the data warehouse regularly receives data from apps and systems used by the company. The data is formatted and imported to match the data already stored in the warehouse. The processed data becomes readily available to the end-users to help them make business decisions. The organization chooses how often the data should be obtained from apps and systems depending on its particular needs and requirements.

It's worth noting that a data warehouse is different from a simple database. Data warehouses are created to help perform an analysis of a large volume of data. Their goal is to receive and process information in the shortest period of time possible.

o Business Analytics

Business analytics is the process of using quantitative methods to derive meaning from data in order to make informed business decisions.

There are three primary methods of business analysis:

- Descriptive: The interpretation of historical data to identify trends and patterns

- Predictive: The use of statistics to forecast future outcomes

- Prescriptive: The application of testing and other techniques to determine which outcome will yield the best result in a given scenario

Deciding which method to employ is dependent on the business situation at hand.

o Business Performance Management

Corporate Performance Management (CPM) encompasses methods, metrics, activities, and systems, which are used to monitor and manage the business performance of a company. CPM software processes the focused information to turn it into operational plans.

This process and methodology offer business owners an integrated approach to planning, forecasting for finance, sales, marketing, HR, and operations. When this methodology is implemented, it joins company strategies with plans and executions, thus helping a business succeed and improve.

CPM is an important component of business intelligence for companies that are looking for such changes as budget remodeling, cost-cutting, upgrading organization strategy, better KPIs alignment, and improving the process of financial planning.

According to software experts at Tricension, cloud-based CPM software makes it easier to deploy the process, reduce costs, and up the information speed and flexibility. It allows the company to automate numerous manual tasks, and improve the alignment between finance and operations.

o User Interface

The user interface (UI) is the point at which human users interact with a computer, website or application. The goal of effective UI is to make the user's experience easy and intuitive, requiring minimum effort on the user's part to receive maximum desired outcome.

UI is created in layers of interaction that appeal to the human senses (sight, touch, auditory and more). They include both input devices like keyboard, mouse, trackpad, microphone, touch screen, fingerprint scanner, e-pen and camera and output devices like monitors, speakers and printers. Devices that interact with multiple senses are called "multimedia user interfaces". For example, everyday UI uses a combination of tactile input (keyboard and mouse) and a visual and auditory output (monitor and speakers).

Other types of user interfaces can include:

- **Form-based user interface:** Used to enter data into a program or application by offering a limited selection of choices. For example, a settings menu on a device is form-based.
- **Graphical user interface:** A tactile UI input with a visual UI output (keyboard and monitor).
- **Menu-driven user interface:** A UI that uses a list of choices to navigate within a program or website. For example, ATMs use menu-driven UIs and are easy for anyone to use.

- **Touch user interface:** User interface through haptics or touch. Most smartphones, tablets and any device that operates using a touch screen use haptic input.
- **Voice user interface:** Interactions between humans and machines using auditory commands. Examples include virtual assistant devices, talk-to-text, GPS and much more.

SR.NO	QUESTION	ANSWER
1	Which is the processes, technologies and tools that help us change data into information, information into knowledge and knowledge into plans that guide organization?	Business Intelligence
2	Business Intelligence makes companies smarter. True or False	True
3	Business intelligence include data mining, online analytical processing and ………………	Business reporting
4	What decisions helps BI to providing right information to right person at right place at right time?	Strategic, Tactical, Operational
5	Business Intelligence is :Your business, Your management and ……………..	Information technology
6	Business Intelligence terminology are:	Database, Data integration, Data Analysis and Reporting.
7	How many phases are there in business intelligence?	4
8	What is ETL?	Extract , Transform, Load
9	Rational database which contains historical data are stored in?	Data warehouse
10	BI application are divided into how many parts?	2

SHREE H.N.SHUKLA COLLEGE OF MANAGEMENT STUDIES, RAJKOT
AFFILIATED TO GUJARAT TECHNOLOGICAL UNIVERSITY

Lalpari Campus, Near Lalpari Lake,		3 – Vaishalinagar
B/H Marketing Yard,		Nr. Amrapali Railway Crossing
Amargadh (Bhichri), Rajkot		Raiya Road, Rajkot
Ph. No. 9727753360		Rajkot - 360001

11	Business Application can be divided into: one is technology solution and other is………..	Business solutions
12	Full form of CRM	Customer relationship management
13	Who are the users of business intelligence?	IT, power, business, casual and external parties
14	SVOT full form is	Single version of the truth
15	Full form of BI	Business intelligence
16	………… is a set of techniques and processes that can be used to analyse data to improve business performance through fact based decision making.	Business analytics
17	Full form of UI	User interface

Business Analytics:

☐ Introduction to Business Analytics (BA) – Need.

- ⊡ **(Business) Analytics** is the use of data:
  - ▪ Information Technology
  - ▪ Statistical analysis
  - ▪ Quantitative Methods
  - ▪ Mathematical / Computer-based models

- ⊡ To help managers gain improved insight about their business operations and make better-fact based decisions.

- Business Analytics is a set of techniques and processes that can be used to analyse data to improve business performance through fact-based decision-making.

- Business Analytics is the subset of Business Intelligence, which creates capabilities for companies to compete in the market effectively and is likely to become one of the main functional areas in most companies. Analytics companies develop the ability to support decisions through analytic reasoning.

- Business analytics (BA) is the <u>methodical exploration</u> of an organization's data with emphasis on statistical analysis. Business analytics is used by companies committed to data-driven decision making.

- Business Analytics refers to all methods and techniques that are used by an organization to measure performance.

- Business Analytics are made up of statistical methods that can be applied to specific project ,process or product.

- Business analytics is a mixer of statistical method and technological method for analysis company's data to improve strategic goal.

- It is a data management solution and business intelligence subset.

- It uses methodologies such as data mining, predictive analysis and statistical analysis.

Evolution of Business Analytics

- **Business Intelligence**

- **Information Systems**
- **Statistics**
- **Operations research/management science**

SHREE H.N.SHUKLA COLLEGE OF MANAGEMENT STUDIES, RAJKOT
AFFILIATED TO GUJARAT TECHNOLOGICAL UNIVERSITY

Lalpari Campus, Near Lalpari Lake,   3 – Vaishalinagar

B/H Marketing Yard,   Nr. Amrapali Railway Crossing

Amargadh (Bhichri), Rajkot   Raiya Road, Rajkot

Ph.No–9879775960   Rajkot - 360001

Ph.No–(0281)2471645

⊡ **Decision support system**

Business Intelligence

BI encompasses a wide variety of tools, applications and methodologies that enable organizations to collect data from internal systems and external sources; prepare it for analysis; develop and run queries against that data; and create reports, dashboards and data visualizations to make the analytical results available to corporate decision-makers, as well as operational workers.

Information Systems

An information system (IS) is an organized system for the collection, organization, storage and communication of information.

Statistics

the practice or science of collecting and analysing numerical data in large quantities, especially for the purpose of inferring proportions in a whole from those in a representative sample.

Operations research/management science

It is a discipline that deals with the application of advanced analytical methods to help make better decisions.

SHREE H.N.SHUKLA COLLEGE OF MANAGEMENT STUDIES, RAJKOT
AFFILIATED TO GUJARAT TECHNOLOGICAL UNIVERSITY

Lalpari Campus, Near Lalpari Lake,	3 – Vaishalinagar
B/H Marketing Yard,	Nr. Amrapali Railway Crossing
Amargadh (Bhichri), Rajkot	Raiya Road, Rajkot

Management science (MS), is the broad interdisciplinary study of problem solving and decision making in human organizations, with strong links to management, economics, business, engineering, management consulting, and other sciences.

Decision support system

It is a set of related computer programs and the data required to assist with analysis and decision-making within an organization

Need/Reasons/Importance For BA

One of the fastest growing markets in the 2017 enterprise software landscape.

Today's businesses are growing increasingly digital and are capable of accurately measuring every aspect of their operations, from marketing to human resources, in real-time.

Provides faster and more accurate decisions

Enables businesses to stay on top of the market by revealing sentiments towards the company , as well as its competitors.

Minimizes risk

Offer valuable insights to help businesses make the right choices based on performance, consumer behavior and trends.

⊡ **Organized Work**

SHREE H.N.SHUKLA COLLEGE OF MANAGEMENT STUDIES, RAJKOT
AFFILIATED TO GUJARAT TECHNOLOGICAL UNIVERSITY

Lalpari Campus, Near Lalpari Lake,	3 – Vaishalinagar
B/H Marketing Yard,	Nr. Amrapali Railway Crossing
Amargadh (Bhichri), Rajkot	Raiya Road, Rajkot

Having accurate information, the team is able to work together in an organized manner to come up with a plan that will bring higher chances of success for the company. In this day and age, there are already available business analysis software and applications that enable managers to keep up with advanced consumer shopping trends and also project future trends.

☑ **Assessment of previous business performance**

It is used for predictive analysis, which is typical used to asses' previous business performance. It also clears picture of what is being worked and what is not.

☑ **Pricing decisions**

Business analytics is used to determine pricing of various products in a departmental store based past and present set of information. .i.e. Movie theatre

☑ **Customer Segmentation**

Identifying and targeting key customer groups in different industries.

☑ **Disseminate information to relevant stakeholders through interactive dashboards and reports**

It is used for sharing information to internal and external stakeholders of the company . So many stakeholders can make proper decisions on their basis.

☑ **Improved customer service**

Improved customer service keep track of frequent customer queries which prevents businesses from repeating mistakes and improving customer satisfaction.

☑ **Merchandising**

Determining what to sell and buy

- **Social Media**

- Understand trends and customer perception which will help managers and product designers.

Ph.No–(0281)2471645

☐ Components (Business Context, Technology, Data Science)

(1) Business Context

▣   Any business analytics project starts with a business context and continue with asking the right questions.

▣   To help with business decisions, one should ask questions, he/she wants to gain insights into before starting the data collection process.

Based on the company's strategy, goals, budget and target customers, one should prepare a set of questions that will help him/her through the data analysis.

(2) Technology

▣ Technology is also necessary to analyze the data.

▣ IT can be used for acquisition, storage, preparation, analysis and dissemination of data.

▣ Companies use many softwares for analysis.

▣ Technology is important to implement solution. e.g., in the case of targeted advertising, technology can be used to personalize advertisements to be sent to individual customers.

(3) Data Science

▣ Most significant component

▣ Comprises of statistical techniques, deep learning, machine learning etc.

▣ Aim of this component is to identify the best fit technique in current context.

SHREE H.N.SHUKLA COLLEGE OF MANAGEMENT STUDIES, RAJKOT
AFFILIATED TO GUJARAT TECHNOLOGICAL UNIVERSITY

Lalpari Campus, Near Lalpari Lake,	3 – Vaishalinagar
B/H Marketing Yard,	Nr. Amrapali Railway Crossing
Amargadh (Bhichri), Rajkot	Raiya Road, Rajkot

▣ Multiple models are developed for solving the problems using available techniques and some of them are selected for deployment of the business analytics solution.

Types (Descriptive, Predictive and Prescriptive).

1. Descriptive Analytics: The use of data to <u>understand past and current</u> business performance and make informed decisions.

2. Predictive Analytics: <u>Predict the future</u> by examining historical data, detecting patterns or relationships in these data and then connect them with future.

3. Prescriptive Analytics: Identify the best alternatives to minimize or maximize some objectives.

1. Descriptive Analytics

- 90% of organizations today use descriptive analytics which is the most basic form of analytics.

- Easiest and quickest part

- It helps in answering this question:"What has happened?".

- The simplest class of analytics, one that allows you to condense big data into smaller, more useful information."

- The main objective of descriptive analytics is <u>to find out the reasons behind precious success or failure in the past</u>. The 'Past' here, refers to any particular time in which an event had occurred and this could be a month ago or even just a minute ago.

- Most of the social analytics are descriptive analytics. They summarize certain groupings based on simple counts of some events. The number of followers, likes, posts, fans are mere event counters. These metrics are used for social

analytics like average response time, average number of replies per post, number of page views, etc. that are the outcome of basic arithmetic operations.

- <u>The vast majority of big data analytics used by organizations falls into the category of descriptive analytics.</u>

Characteristics

1. Mostly used by organizations
2. Based on historical data
3. Most of the social analytics are this kind of analytics

2. **Predictive Analytics**

- The subsequent step in data reduction is predictive analytics.
- It is used by businesses to study the data to find answers to the question "What could happen in the future based on previous trends and patterns?"
- "The purpose of predictive analytics is NOT to tell you what will happen in the future. It cannot do that. In fact, no analytics can do that.
- Predictive analytics can only forecast what might happen in the future, because all predictive analytics are probabilistic in nature."
- Provides better recommendations and more future looking answers to questions that cannot be answered by BI.
- Helps predict the likelihood of a future outcome by using various statistical and machine learning algorithms but the accuracy of predictions is not 100%, as it is based on probabilities.

SHREE H.N.SHUKLA COLLEGE OF MANAGEMENT STUDIES, RAJKOT
AFFILIATED TO GUJARAT TECHNOLOGICAL UNIVERSITY

Lalpari Campus, Near Lalpari Lake,	3 – Vaishalinagar
B/H Marketing Yard,	Nr. Amrapali Railway Crossing
Amargadh (Bhichri), Rajkot	Raiya Road, Rajkot

Organizations should capitalise on hiring a group of data scientists who can develop statistical and machine learning algorithms to leverage predictive analytics and design an effective business strategy.

Characteristics

- ▣ Next step of descriptive analysis
- ▣ It answers what will be happened in future on the basis of past actions.
- ▣ It also helps to find out reasons of past occurrence.

3. **Prescriptive Analytics**

- ■ Prescriptive analytics is the next step of predictive analytics that adds the spice of manipulating the future.
- ■ Prescriptive analytics advises on possible outcomes and results in actions that are likely to maximise key business metrics. It basically uses simulation and optimization to ask "What should a business do?"
- ■ Prescriptive analytics are comparatively complex in nature and many companies are not yet using them in day-to-day business activities, as it becomes difficult to manage.
- ■ Prescriptive analytics if implemented properly can have a major impact on business growth.
- ■ Large scale organizations use prescriptive analytics for scheduling the inventory in the supply chain, optimizing production, etc. to optimize customer experience.

Characteristics

- ▣ Next step to predictive analytics

SHREE H.N.SHUKLA COLLEGE OF MANAGEMENT STUDIES, RAJKOT
AFFILIATED TO GUJARAT TECHNOLOGICAL UNIVERSITY

Lalpari Campus, Near Lalpari Lake,   3 – Vaishalinagar

B/H Marketing Yard,   Nr. Amrapali Railway Crossing

Amargadh (Bhichri), Rajkot   Raiya Road, Rajkot

Ph.No–9727753100   Ph.No–(0281)2471645

- ▣ Provides guidelines regarding how to make business more efficient
- ▣ Identifies data uncertainties to make future decisions
- ▣ Mostly used by financial institutions

▢ Business Intelligence versus Business Analytics.

Point of difference	Business Intelligence (BI)	Business Analytics (BA)
Meaning	BI is a software application which can include data mining, online analytical processing and business reporting.	Business Analytics is the methodical exploration of the data with emphasis on statistical analysis.
Objective	To collect and represent data in an under stable manner	To derive insights and understanding
Example	Dashboard, Report Generation, ETL etc	Data Mining, Predictive Modeling , Forecasting etc.
Focus	Present	Present and Future
Data Type	Structured Data	Structured and unstructured Data
Users	It includes business users	It includes data scientists, Business Analyst, Business users.
Answers the questions:	• What happened? • When did it happen? • Who is accountable for what happened?	• Why did it happen? • Will it happen again? • What will happen if we change x ?

SHREE H.N.SHUKLA COLLEGE OF MANAGEMENT STUDIES, RAJKOT
AFFILIATED TO GUJARAT TECHNOLOGICAL UNIVERSITY

Lalpari Campus, Near Lalpari Lake,	3 – Vaishalinagar
B/H Marketing Yard,	Nr. Amrapali Railway Crossing
Amargadh (Bhichri), Rajkot	Raiya Road, Rajkot

	• How many? • How often? • Where did it happen?	• What else does the data tell us that we never thought to ask? • What is the best that can happen?
Scope	It is a **broad** term	It is a subset of BI

o **OLTP v/s OLAP**

Point of Difference	OLAP	OLTP
Application based	It is subject-oriented	It is transaction-oriented
The number of users	The number of users are few in OLAP	The number of users are multiple in OLTP
Source of data	OLAP data comes from the various OLTP Databases	OLTPs are the original source of the data.
Purpose of data	To help with planning, problem solving, and decision support	To control and run fundamental business tasks
What the data	Multi-dimensional views of various kinds of business activities	Reveals a snapshot of ongoing business processes
Inserts and Updates	No inserting and updating data	Short and fast inserts and updates initiated by end users

Queries	Often complex queries involving aggregations	Relatively standardized and simple queries Returning relatively few records
Processing Speed	Depends on the amount of data involved; batch data refreshes and complex queries may take many hours	Typically very fast
Space Requirements	Larger due to the existence of aggregation structures and history data; requires more indexes than OLTP	Can be relatively small if historical data is archived
Data storage	Terabytes	Gigabytes
DB Size	100 GB-TB	100 MB-GB
Engine	OLAP is the Reporting engine OLAP is the reporting engine on top of the DW/BI that collates a question and provides the answer the user is looking for.	OLTP is the business process engine A sales order is recorded in an OLTP engine, as is the quotations, the materials, prices, shipping and other business processes. relevant data is then

	Ph. No. 9727753360	extracted into a data warehouse. Ph.No. (0281)2471645
Example	Knowledge workers	IT Professionals
Data	OLAP stores historical information	OLTP stores real-time ,up to date information
Function	Decision support	Day to day operations

o OLAP Operations

o Data models for OLTP (ER model) and OLAP (Star & Snowflake Schema)

- ▣ OLAP uses Data warehouse.

- ▣ A data warehouse requires to maintain a schema.

- ▣ Schema is a logical description of the entire database.

- ▣ It includes the name and description records of all record types including all associated data-items.

- ▣ It is important to understand two basic terms before understanding schemas, viz. fact and dimensions form.

What are Facts and Facts tables?

- ▣ Facts relate to **events.**

- ▣ Fact tables contain the data relating to a particular business.

- ▣ Each row represents a single event associated with a process.

  For example,

A retail company might have fact tables related to customer purchases, customer service, telephone calls and product returns.

SHREE H.N.SHUKLA COLLEGE OF MANAGEMENT STUDIES, RAJKOT
AFFILIATED TO GUJARAT TECHNOLOGICAL UNIVERSITY

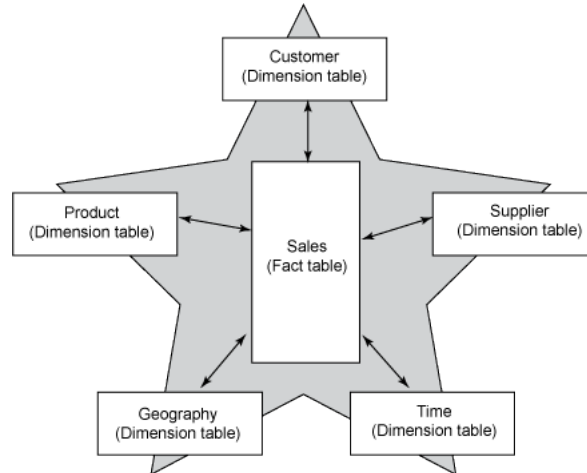| Lalpari Campus, Near Lalpari Lake,	3 – Vaishalinagar
B/H Marketing Yard,	Nr. Amrapali Railway Crossing
Amargadh (Bhichri), Rajkot	Raiya Road, Rajkot

- ▣ The information contained within a fact table is typically numeric data and it is often data that can be manipulated.

Ph.No–(0281)2471645

What are Dimensions and Dimension Tables?

- ▣ Dimensions describe the objects involved in a business intelligence effort.

- ▣ Dimensions relate to people, items or other objects.

For Example,

- ■ In the retail company, purchases returns are facts, while customers, employees, items are dimensions and should be contained in dimension tables.

- ▣ Dimension tables contain details about each instance of an object.

For example,

- ■ The items dimension table would contain a record for each item sold in the store. It includes information such as the cost of the item, color, sizes etc.

- ▣ Dimension tables are generally small in size then fact table.

- ▣ Typical fact tables store data about sales while dimension tables data about geographic region, clients, products etc.

Types of Schemas

1. Star Schema

- It is the simplest data warehouse schema.

- It is called as 'Star Schema' because the diagram looks like a star, with points radiating from a center.

- The center of the star consists of fact table and the points of the star are the dimensions table.

SHREE H.N.SHUKLA COLLEGE OF MANAGEMENT STUDIES, RAJKOT
AFFILIATED TO GUJARAT TECHNOLOGICAL UNIVERSITY

Lalpari Campus, Near Lalpari Lake,   3 – Vaishalinagar

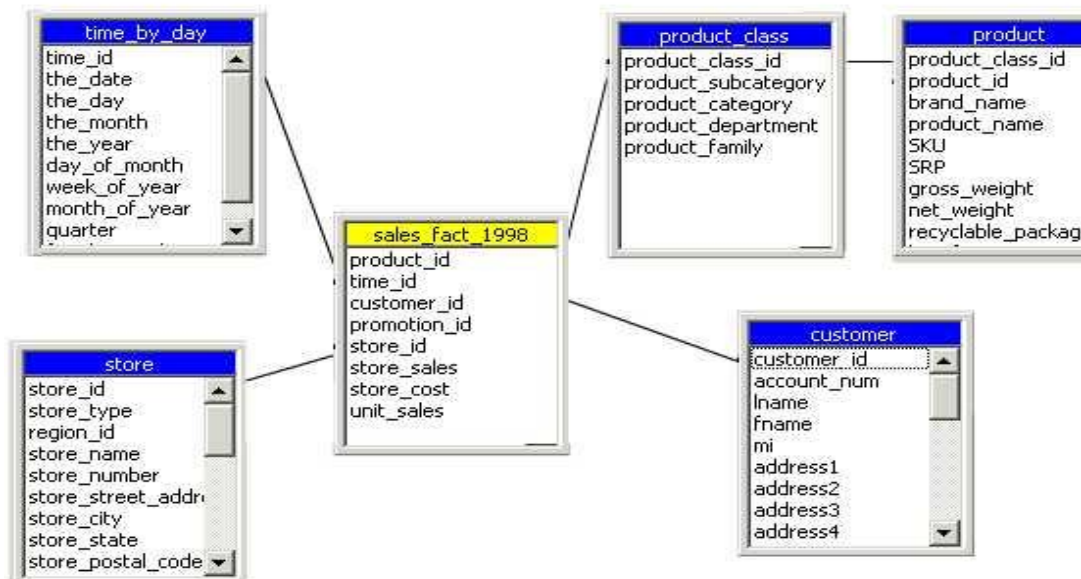B/H Marketing Yard,   Nr. Amrapali Railway Crossing

Amargadh (Bhichri), Rajkot   Raiya Road, Rajkot

Ph.No 9727753360   Rajkot–360001

Ph.No–(0281)2471645

- Each dimension in a star schema is represented with only one-dimension table.

- This dimension table contains the set of attributes.

This diagram shows the sales data of a company respect to the 4 dimensions, namely time, Item, branch and location.

SHREE H.N.SHUKLA COLLEGE OF MANAGEMENT STUDIES, RAJKOT
AFFILIATED TO GUJARAT TECHNOLOGICAL UNIVERSITY

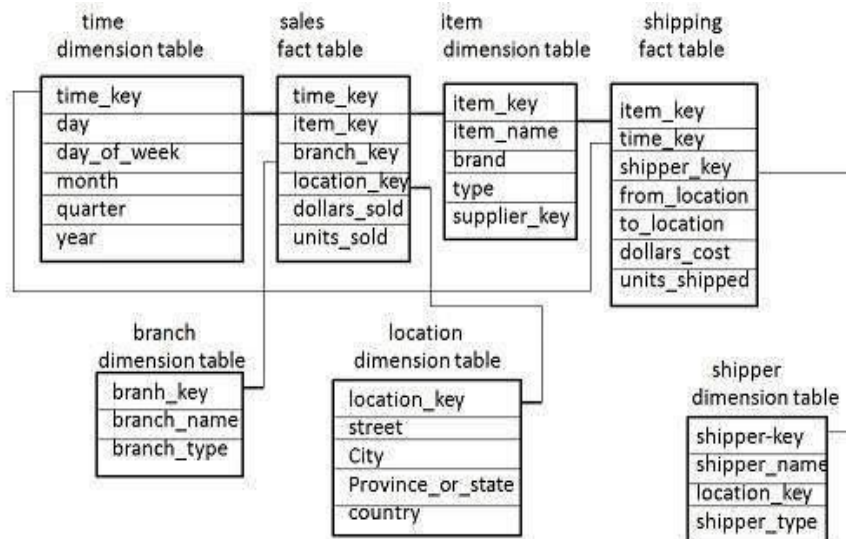|	
Lalpari Campus, Near Lalpari Lake,	3 – Vaishalinagar
B/H Marketing Yard,	Nr. Amrapali Railway Crossing
Amargadh (Bhichri), Rajkot	Raiya Road, Rajkot
Ph. No. 9727753360	Rajkot - 360001
	Ph.No–(0281)2471645

2. Snowflake Schema

- It is a logical arrangement of tables in a multidimensional databases such that the entity relationship diagram looks like a snowflake shape.

- It is represented by centralized fact tables which are connected to multiple dimensions.

- Some dimension tables in the Snowflake schema are normalized.

- The normalization splits up the data into additional tables.

- The snowflake schema is similar to the star schema.

- Here, tables are split into multiple tables.
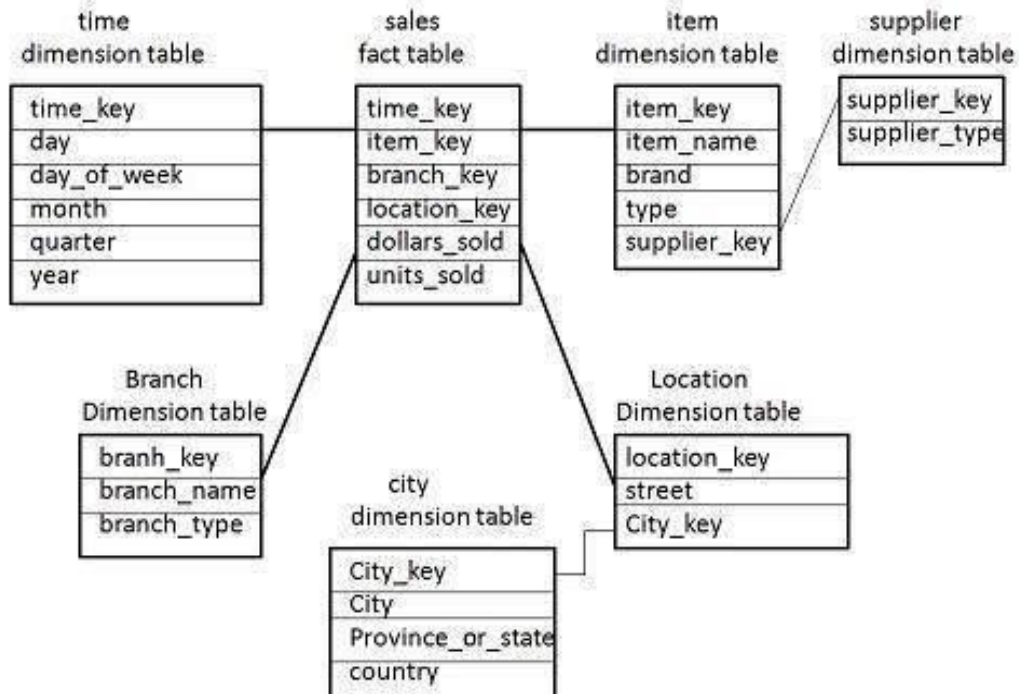
3. Fact Constellation Schema

- A fact constellation has multiple fact tables. It is also known as galaxy schema.

- The sales fact table is same as that in the star schema.

- **E.g. :** The shipping fact table has the five dimensions, namely item_key, time_key, shipper_key, from_location, to_location . The shipping fact table also contains two measures, namely dollars sold and units sold.

- It is also possible to share dimension tables between fact tables. For example, time, item, and location dimension tables are shared between the sales and shipping fact table.

The following diagram shows two fact tables, namely sales and shipping.

Sr. no	Questions	Answer
1.	OLAP stands for?	Online Analytical Processing Server

2	OLAP based on?	Multidimensional data model
3.	How many types of OLAP servers?	4
4	_____ uses array-based multidimensional storage engines for multidimensional views of data.	MOLAP
5	ROLAP servers are placed between relational back-end server and client front-end tools. True and false	True
6	Which of the following is false about OLAP?	Number or users is in millions
7	The pivot operation is also known as?	Rotation
8	The generalization of cross-tab which is represented visually is _____ which is also called as data cube.	Two dimensional cube
9	The process of viewing the cross-tab (Single dimensional) with a fixed value of one attribute is	Slicing
10	Roll-up performs aggregation on a data cube	Yes
11	In SQL the cross-tabs are created using	Pivot
12	How many grouping is possible in this rollup?	4
13	What do data warehouses support?	OLAP
14	In which of the following, data are stored, retrieved and updated?	OLTP
15	Full for of OLTP	Online transaction processing
16	OLTP deals with large numbers of transactions by a ………….	Large number of users
17	Purpose of OLTP is	To control and run fundamental business tasks

18	In OLTP data storage is in	100 Giga bytes
19	Which data warehouse is the simple?	Star schema
20	Each dimensions of the star schema is represented with what?	One dimension table
21	How many schemas used by data warehouse?	3
22	Which schema has multiple dimension used by data warehouse?	Fast Constellation Schema
23	A fact constellation has multiple fact tables. True/False	True
24	What is also known as galaxy schema?	Fact Constellation Schema
25	___ is a random error or variance in measured variables.	Noise
26	State true or false I. BI applications can also help managers to be better informed about actions that a company's competitors are taking II. BI can help companies share selected strategic information with business partners. III. BI 2.0″ is used to describe the acquisition, provision and analysis of "real-time" data	i-T, ii-T, iii-T.
27	. ___ routines attempt to fill in missing values, smooth out noise while identifying outlines, and correct inconsistencies in the data.	Data cleaning
28	___ is used to refer to systems and technologies that provide the business with the means for decision-makers to extract personalized meaningful information about their business and industry	Business Intelligence
29	In ___ each value in a bin is replaced by the mean value of the bin.	Smoothing by bin means

30	___ regression involves finding the "best" line to fit two variables so that one variable can be used to predict the other.	Linear
31	works to remove the noise from the data that includes techniques like binning, clustering, and regression.	Smoothing
32	Redundancies can be detected by correlation analysis. (True/False)	True
33	___ Technique uses encoding mechanisms to reduce the data set size.	Data compression
34	In which Strategy of data reduction redundant attributes are detected.	Dimension reduction
35	___ Hierarchies can be used to reduce the data by collecting and replacing low-level concepts by higher-level concepts.	Concept
36	The ___ rule can be used to segment numeric data into relatively uniform, "natural" intervals.	3-4-5
37	Oracles, SQL/Server, DB2 are examples for ___.	DBMS
38	Data Base Management System (DBMS) supports query languages. (True/False)	True
39	The ___ item sets find all sets of items (items sets) whose support is greater than the user-specified minimum support, σ.	Frequent set
40	A frequent set is a ___ if it is a frequent set and no superset of this is a frequent set.	Maximal frequent set
41	techniques are used to detect relationships or associations between specific values of categorical variables in large data sets	Association rule mining
42	A Decision Tree is a ___ model	Predictive ;model
43	Using a decision tree, only categorical variables would be modeled. (True/False).	False

44	Clustering is an unsupervised learning method (True/false).	False
45	Neural networks are made up of many ___.	Artificial neurons
46	For a given transaction database T, a ___ is an expression of the form X => Y, where X and Y are subsets of A and X => Y holds with confidence Ƭ, if Ƭ% of transactions in D support X also support Y.	Association rule
47	The ___ rule describes associations between quantitative items or attributes.	Quantitative association
48	The ___ step eliminates the extensions of (k-1) – item sets, which are not found to be frequent, from being considered for counting support.	Pruning
49	In the first phase of the Partition algorithm, the algorithm logically divides the database into a number of ___.	non – overlapping partitions
50	The a priori algorithm operates in a ___ and ___.	Bottom-up, breadth-first search method.
51	___ algorithm works like a train running over the data, with stops at intervals M between transactions. When the train reaches the end of the transaction file it completes one path.	DIC Algorithm
52	FP–Tree Growth Algorithm can be implemented in ___ Phases	Two
53	FP – tree stands for ___.	Frequent pattern tree
54	Data mining systems should provide capabilities to mine association rules at multiple levels of abstraction and traverse easily among different abstraction spaces (True/False).	True
55	Which one of the following is alternative search strategies for mining multiple-level associations with reduced support?	Level – by level independent, Level – cross-

	Ph. No. 9727753360 Rajkot 160001 Ph.No– (0281)2471645	filtering by a single item, Level – cross-filtering by k – itemset
56	Which of the following is NOT a common binning strategy?	Equilength binning
57	Association rules that involve two or more dimension or predicates can be referred to as ___.	Multidimensional association rules
58	An algorithm that performs a series of "walks" through itemset space is called a ___.	Random walk algorithm
59	What are knowledge type constraints?	They specify the type of knowledge to be mined.
60	A standard measure of within-cluster similarity is ___.	variance
61	The process of grouping a set of physical or abstract objects into classes of similar objects is called ___.	Cluster
62	Clustering may also be considered as ___.	Segmentation
63	Clustering is also called:	Segmentation, Compression & Partitions with similar objects
64	Clustering is used only in data mining (True/False).	True
65	Clustering is a form of learning by observation rather than ___.	By example
66	Weight and height of an individual fall into ___ kind of variables.	Continuous
67	. In the K-means algorithm for partitioning, each cluster is represented by the ___ of objects in the cluster.	Means

68	K-means clustering requires prior knowledge about number clusters required as its input.(True/False).	True
69	. One form of unsupervised learning is ___.	Clustering
70	___ software provides a set of partitioned clustering algorithms that treat the clustering problem as an optimization process.	CLUTO
71	Data classification is a ___ step process.	Two
72	___ can be viewed as the construction and use of a model to assess the class of an unlabeled sample, or to assess the value or value ranges of an attribute that a given sample is likely to have.	Prediction
73	___ of data removes or reduces noise (by applying smoothing techniques) and the treatment of missing values.	Pre-processing
74	___ method refers to the ability to construct the model efficiently given a large amount of data.	Scalability
75	What is a decision tree?	This is a flow – chart – like a tree structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and leaf nodes represent classes or class distribution
76	The basic algorithm for decision tree induction is a ___ algorithm.	greedy
77	The ___ measure is used to select the test attribute at each node in the tree.	information gain

78	A user session is a ___ record spanning the entire Web.	Clickstream record
79	___ is simple text files that are automatically generated every time someone accesses one Website.	Log File
80	Which files are frequently used in sequential mining.	Web log files