



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

## MBA SEMESTER 2

### Business Analytics- 4529201

#### MODULE 2

##### Types of Digital Data:

□ Definition, Sources, Storage and Characteristics of Structured, Unstructured and Semi Structured Data

- ▶ Digit refers to number
- ▶ Represents in binary form
- ▶ Binary digit (bit) = 0 or 1
- ▶ Example :
  - Data                      Binary representation
  - Letter A                0100 0001
  - Number of distinct bit combinations that can be produced is given by the formula  $2^n$ .
- ▶ Adding 1 to the power doubles the number of distinct data items that can be encoded
- ▶  $2^1$  - 2 items
- ▶  $2^2$  - 4 items
- ▶  $2^3$  - 8 items
- ▶ It is the language used in most present day by computers.
- ▶ Data received from electronic devices or network is known as 'Digital Data'.
- ▶ Digital data is any information we represent in binary (base 2) form.



Lalpari Campus, Near Lalpari Lake,

3 – Vaishalinagar

B/H Marketing Yard,

Nr. Amrapali Railway Crossing

Amargadh (Bhichri), Rajkot

Raiya Road, Rajkot

▶ A string of ones and zeros

Rajkot - 360001

▶ In this system, there represents an increasing power of 2, 2 to the power of 0, 2 to the power of 1, and so on.

Phone-202912471645

▶ For example, a person might say, 'Hi' to a friend, it will be converted in binary form and then will be reached to the receiver.

▶ A computer would represent that as '0100100001101001' in base 2. Or a person might hand you a picture to look at. A computer would break the picture down into a 2-dimensional series of dots.

▶ 0,1 - Binary Number

▶ 8 bits – 1 byte

▶ 1,2,3,.....,9,10 = base 0 & 1

▶ Base-2 = 0,1,10,11,100,101,110,111

## Sources of Data

### Primary Data

- ▶ Primary data means original data that has been collected specially for the purpose in mind. It means someone collected the data from the original source first hand.
- ▶ Data collected this way is called primary data.
- ▶ Primary data has not been published yet and is more reliable, authentic and objective. Primary data has not been changed or altered by human beings; therefore its validity is greater than secondary data.

### Sources of data

- ▶ Survey
- ▶ Questionnaire



Lalpari Campus, Near Lalpari Lake,

B/H Marketing Yard,

Amargadh (Bhichri), Rajkot

▶ Interview Ph. No. 9727753360

▶ Observations

3 – Vaishalinagar

Nr. Amrapali Railway Crossing

Raiya Road, Rajkot

Rajkot - 360001

Ph.No-(0281)2471645

## ADVANTAGES

- Data interpretation is better.
- Targeted Issues are addressed.
- Correct data

## DISADVANTAGES

- High Cost
- Time Consuming
- Inaccurate Feed-backs
- More number of resources is required

## Secondary Data

- ▶ Secondary data is the data that has been already collected by and readily available from other sources. When we use Statistical Method with Primary Data from another purpose for our purpose we refer to it as Secondary Data.
- ▶ It means that one purposes Primary Data is another purposes Secondary Data. So that secondary data is data that is being reused. Such data are more quickly obtainable than the primary data.
- ▶ These secondary data may be obtained from many sources, including literature, industry surveys, compilations from computerized databases and information systems, and computerized or mathematical models of environmental processes.



**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

**Sources of data**

Ph. No. 9727753360

Rajkot - 360001

Ph.No-(0281)2471645

- ▶ Published Printed Sources
- ▶ Books
- ▶ Journals/periodic
- ▶ Magazines/Newspapers
- ▶ Published Printed Sources

**ADVANTAGES**

- ▶ It is economical. It saves efforts and expenses.
- ▶ Time saving.
- ▶ It helps to make primary data collection more specific since with the help of secondary data, we are able to make out what are the gaps and deficiencies and what additional information needs to be collected.
- ▶ It provides a basis for comparison for the data that is collected by the researcher.

**DISADVANTAGES**

- ▶ Accuracy of secondary data is not known.
- ▶ Data may be outdated.

**These classification are also the part of Big data**

**Big Data:** It is the huge volume of data which can not be stored and processed using traditional system within given time.

**1. Structured data**

The data that does have proper format to associate to it can be referred to as Structured Data.



Lalpari Campus, Near Lalpari Lake,

3 – Vaishalinagar

B/H Marketing Yard,

Nr. Amrapali Railway Crossing

Amargadh (Bhichri), Rajkot

Raiya Road, Rajkot

Rajkot - 360001

These data are stored in databases.

Structured data is highly organized information that uploads neatly into a relational database

It concerns all data which can be stored in database SQL

- ▶ For Example,
  - An employee table in a database
  - A Data Mart / Data Warehouse
  - Data in Excel format etc
- ▶ Structured data can be handled through traditional system (RDBMS System) or Hadoop system.
- ▶ Hadoop is not a database, it is basically a distributed file system which is used to process and store large data.
- ▶ The simplest way to manage information.
- ▶ Structured data is relatively simple to enter, store, query, and analyze, but it must be strictly defined in terms of field name and type.

Characteristics

- ▶ Organized data
- ▶ Easy to analyze and interpret
- ▶ These data are stored in databases.
- ▶ Mostly used by Lower level employees
- ▶ E.g. Excel spreadsheets, files.

Sources of Structured Data

1. Machine Generated :
  - Sensor data



Lalpari Campus, Near Lalpari Lake,

3 – Vaishalinagar

B/H Marketing Yard,

Nr. Amrapali Railway Crossing

Amargadh (Bhichri), Rajkot

Raiya Road, Rajkot

- Web log data
- Financial data (i.e. stock trading data)
- Point of sale data (i.e. swipe machine)

Rajkot - 360001

Ph.No–(0281)2471645

## 2. Human Generated :

- Input data (human enters data such as name, age etc.)
- Click stream data ( i.e. by clicking link)
- Gaming related data (i.e. steps taken by human in games)

## Storage of Structured Data



## Basis of Management Terminology

### ► The Data Hierarchy

- 8 bits => 1 byte => 1 character
- Field – (attributes)
  - a logical grouping of characters into a word, a small group of words, or a complete number (Fields describe attributes)
- Record (entity)



Lalpari Campus, Near Lalpari Lake,

3 – Vaishalinagar

B/H Marketing Yard,

Nr. Amrapali Railway Crossing

Amargadh (Bhichri), Rajkot

Raiya Road, Rajkot

- a logical grouping of related fields (Records describe entities)
- File –
  - a logical grouping of related records
- Database –
  - a logical grouping of related files

Ph.No–(0281)2471645

## 2.Semi-structured

- The data which does not have any proper format associate to it is referred to as Semi structured data.
- For example, Data which is presented in e-mail messages, word processing documents etc.
- It can only be managed through Hadoop System
- The data that can be converted into structured form after some pre-processing is called semi-structured data.
- The structured data which does not conform with formal structure of data models in context of relationships is semi-structured data.
- This again represents another 5% of the total available data.

### Characteristics

- ▶ Does not have any proper format
- ▶ It can only be managed through Hadoop System
- ▶ 5% of the total available data.



**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

Rajkot - 360001

**Text (E-mails,  
Computer languages)**

Ph.No-0281)2471645

**Web Server  
(Accessing websites)**

**Sensor Data (Wireless  
technology, GPS )**

### 3. Unstructured data

- The data which does not have any format associate to it is referred to as Unstructured data.
- Unstructured data represent around 80% of data.
- All the remaining data having no structure at all, falls into this category.
- It often includes text and multimedia content.
- Examples include videos, photos, audio files, presentations, web pages and many other kinds of business documents.
- Those data which can not be stored inside the RDBMS is called unstructured data.
- Unstructured data is everywhere.
- In fact, most individuals and organizations conduct their lives around unstructured data.





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

- Just as with structured data, unstructured data is either machine generated or human generated
- Unstructured data may have its own internal structure, but does not conform neatly into a spreadsheet or database.
- Most business interactions, in fact, are unstructured in nature.
- Today more than 80% of the data generated is unstructured.
- The fundamental challenge of unstructured data sources is that they are difficult for nontechnical business users and data analysts alike to understand, and prepare for analytic use.

Some examples of unstructured data:

- ▶ Satellite images
- ▶ Photographs and video files and audio files.
- ▶ Text internal to your company (Word doc,PDF file etc.)
- ▶ Social media data
- ▶ Mobile data
- ▶ website content

Characteristics

- ▶ Majority of data in any organization is in unstructured form.
- ▶ Text and multimedia content are of this category.
- ▶ This is the data which can not be stored inside the RDBMS.
- ▶ It is either machine generated or human generated.

Sources of Unstructured data

- Videos
- Emails



**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

- Chats
- Reports
- Webpage

**One word question answer**

Sr. no	Questions	Answer
1	Which of the following is false with respect to digital data transmission?	Only restricted to communication between computers
2	Why is digital data not easily affected by noise?	Cannot easily change binary 1 to 0
3	What is the process of digital communication where a threshold value is set at the receiver to reduce noise?	Signal regeneration
4	Data received from electronic devices or network is known as	Digital data



**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

5	How many sources of digital data is	Rajkot - 360001 Ph.No-(0281)2471645
6	Which data is the easiest and trickiest to secure?	Internal transactional data
7	Which data is easiest to control?	Syndicated data
8	Which data is least secure source of data ?	Harvest data
9	How many types of digital data are there?	3
10	The data that can be converted into structured form after pre processing is called.....	Semi structured
11	An employee table in a database is	Structured data
12	The data which does not have any format associate to it's referred to as	Unstructured data
13	Unstructured data either .....	Machine or human generated
14	Unstructured data often includes	Text and multimedia
15	Semi structured data managed through	Hadoop system

**Data Warehouse:**

Definition, characteristics, framework

- ▶ DW is actually a set of new concept.
- ▶ It is an important tool which is involved into a current technology.
- ▶ With the help of DW, it becomes easier for any organization to count all the types of problem and it will provide the key information of concerned application.



Lalpari Campus, Near Lalpari Lake,

3 – Vaishalinagar

B/H Marketing Yard,

Nr. Amrapali Railway Crossing

Amargadh (Bhichri), Rajkot

Raiya Road, Rajkot

- ▶ Gathering this information is all well and good, but many firms are struggling with their attempts to put this collected knowledge to any meaningful use.

Due to its volume, increased data becomes more problematic for effective analysis.

Data warehouses are key to solving this paradox

History of Data warehouse

Here are some key events in evolution of Data Warehouse-

- ▶ 1960- Dartmouth and General Mills in a joint research project, develop the terms dimensions and facts.
- ▶ 1970- A Nielsen and IRI introduces dimensional data marts for retail sales.
- ▶ 1983- Tera Data Corporation introduces a database management system which is specifically designed for decision support
- ▶ Data warehousing started in the late 1980s when IBM worker Paul Murphy and Barry Devlin developed the Business Data Warehouse.
- ▶ However, the real concept was given by Inmon Bill. He was considered as a father of data warehouse. He had written about a variety of topics for building, usage, and maintenance of the warehouse & the Corporate Information Factory.

## Data Warehouse

Definition :

A data warehouse is a powerful database model that significantly enhances the user's ability to quickly analyze large, multidimensional data sets. It cleanses and organizes data to allow users to make business decisions based on facts.

Date warehousing is an aspect to gather data from multiple sources into central repository, called Data warehouse.



Lalpari Campus, Near Lalpari Lake,

3 – Vaishalinagar

B/H Marketing Yard,

Nr. Amrapali Railway Crossing

Amargadh (Bhichri), Rajkot

Raiya Road, Rajkot

According to William H. Inmon, a leading architect in the construction of data warehouse systems, "A data warehouse is a subject – oriented, integrated, time variant and non- volatile collection of data in support of management's decision making process."

“A data warehouse is simply a single complete, and consistent store of data obtained from a variety of sources and made available to end users in a way they can understand and use it in a business context.”

The data warehouse's greatest strength is getting relevant insight and information into the hands of decision-makers in a timely manner. This enables businesses to keep up with the pace of change, high-competition and digital transformation.

### Characteristics of Data Warehouse

There are four prominent data warehouse characteristics:

1. Subject-Oriented : A data warehouse can be used to analyze a particular subject area. For example, “sales” can be a particular subject. It will provide a simple and conceptual view of any subject.
2. 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.
3. Time-variant: Data is organized via time-periods (weekly, monthly, annually, etc.).
4. Non-volatile: Once data is in the data warehouse, it will not change. So, historical data in a data warehouse should never be altered. A data warehouse is



Lalpari Campus, Near Lalpari Lake,

3 – Vaishalinagar

B/H Marketing Yard,

Nr. Amrapali Railway Crossing

Amargadh (Bhichri), Rajkot

Raiya Road, Rajkot

not updated in real-time. It is periodically updated via the uploading of data, protecting it from the influence of momentary change. Ph.No-(0281)2471645

## Benefits of DW

### 1. The Enablement of Better Decision-Making

- ▶ As companies are now able to get closer to their consumers than ever before, the corporate decision-makers no longer have to make important business decisions based on partial or limited data. They're now backed up by facts and statistics housed within data warehouses that can be recalled ad hoc.

### 2. Quick and Easy Data Access

- ▶ If there's one thing the application economy has taught us, it's that speed is everything. Users can access an array of information, stored across multiple sources, almost instantly. It means you won't be wasting time attempting to manually pull information from various sources, or seeking help from your IT department.

### 3. Consistent Quality Data

- ▶ Data warehouses gather information from countless sources, but they convert it into a unified format to be used throughout your organization. You can have confidence that each of your departments will be producing results which are in line and consistent with each other, which in turn ensures company-wide accuracy.



Lalpari Campus, Near Lalpari Lake,

3 – Vaishalinagar

B/H Marketing Yard,

Nr. Amrapali Railway Crossing

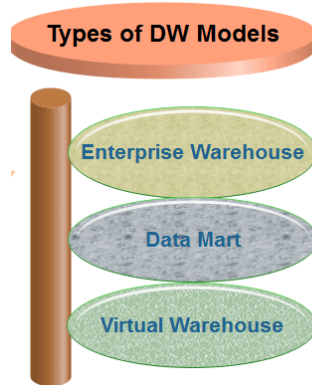
Amargadh (Bhichri), Rajkot

Raiya Road, Rajkot

9727753360

Rajkot - 360001

Ph.No-(0281)2471645



## Enterprise Warehouse

- ▶ In computing, a data warehouse (DW or DWH), also known as an enterprise data warehouse (EDW), is a system used for reporting and data analysis, and is considered a core component of business intelligence.
- ▶ An enterprise data warehouse is a unified database that holds all the business information an organization and makes it accessible all across the company.

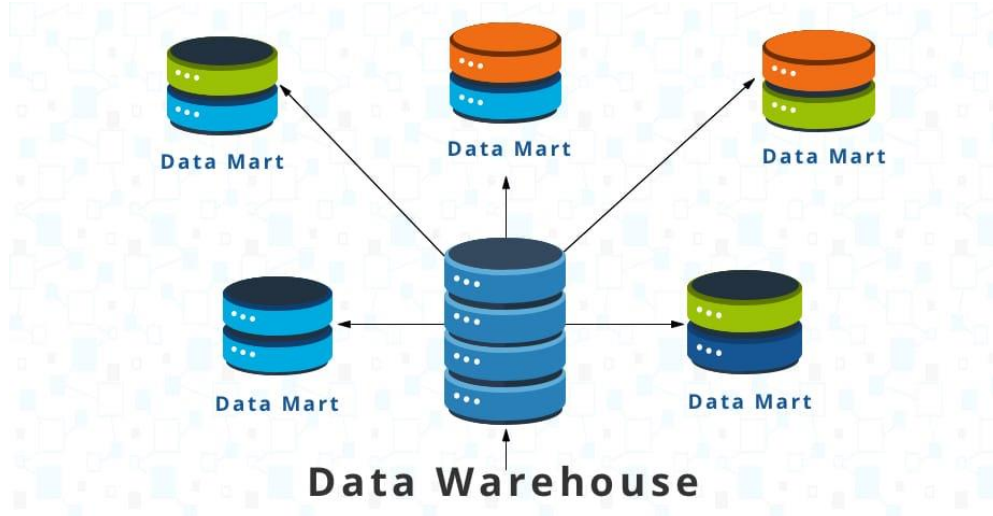
An Enterprise Data Warehouse (EDW) is a form of corporate repository that stores and manages all the historical business data of an enterprise.



## Data Mart



- ▶ The data mart is a subset of the data warehouse and is usually oriented to a specific business line or team. Whereas data warehouses have an enterprise-wide depth, the information in data marts pertains to a single department.
- ▶ In some deployments, each department or business unit is considered the owner of its data mart including all the hardware, software and data.
- ▶ This enables each department to isolate the use, manipulation and development of their data.



### Virtual Data Warehouse

- ▶ A virtual warehouse is another term for a data warehouse.
- ▶ The data found in a virtual warehouse is usually copied from multiple sources throughout a production system. This is done so related data can be searched quickly and without accessing the entire system.
- ▶ Essentially, these are multiple databases connected virtually, so they can be queried as a single system





Lalpari Campus, Near Lalpari Lake,

3 – Vaishalinagar

B/H Marketing Yard,

Nr. Amrapali Railway Crossing

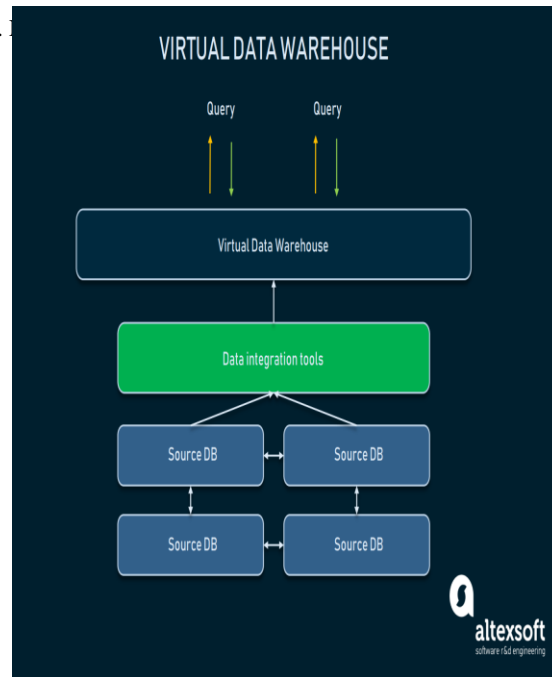
Amargadh (Bhichri), Rajkot

Raiya Road, Rajkot

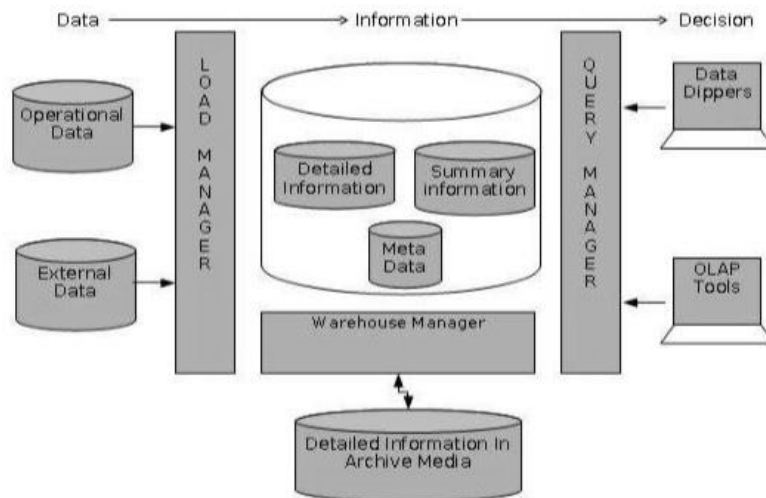
Ph. No.

Rajkot - 360001

Phone - (0281)2471645



## FRAMEWORK/ARCHITECTURE OF DATA WAREHOUSE



- ▶ **Operational Source:** The source of data for the DW are collected from-
  - The data from the main frame OS in traditional network
  - Data can be come from external DB.



Lalpari Campus, Near Lalpari Lake,

3 – Vaishalinagar

B/H Marketing Yard,

Nr. Amrapali Railway Crossing

Amargadh (Bhichri), Rajkot

Raiya Road, Rajkot

- Load Manager: Ph No - 9727753360

Rajkot - 360001

It performs all the operations associated with extraction policy and load the data into DW.

- Warehouse Manager:
  - The warehouse manager performs all the operations associated with the management of data in DW.
  - It will analyze the data consistency, generation of aggregation, backing up and archive the data.
- Query Manager:
  - Detailed Data: It stores all the data from DB schema.
  - Highly summarized & lightly summarized data : It will store the pre-defined lightly & highly summarized data which are generated by warehouse manager.
  - Archive & Back-up data : The detailed & summarized data will be stored for the purpose of archive & Back-up data
- Metadata of Warehouse: Data Warehouse will store all the metadata which are used by the all process in the warehouse.
- End user/Client: It will provide the information to the business manager for strategic decision making.

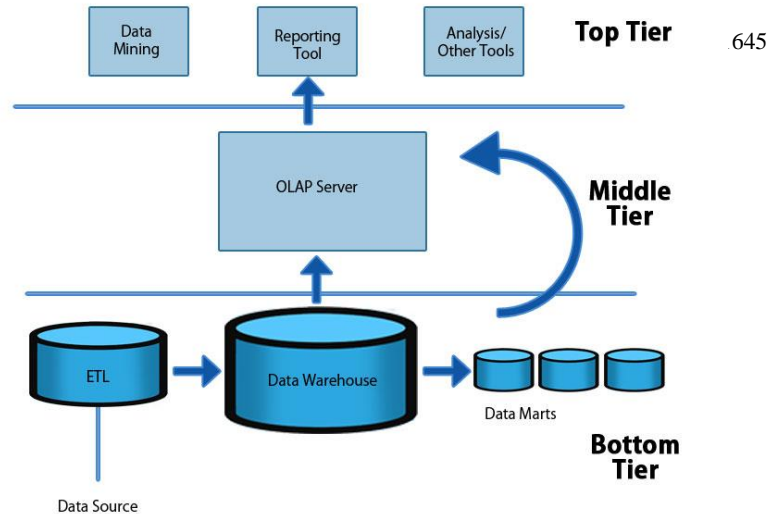
3 Tier Dataware house architecture



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

Lalpari Campus, Near Lalpari Lake,  
B/H Marketing Yard,  
Amargadh (Bhichri), Rajkot

3 – Vaishalinagar  
Nr. Amrapali Railway Crossing  
Raiya Road, Rajkot



- ▶ **Bottom Tier** – The bottom tier of the architecture is the data warehouse database server. It is the relational database system. We use the back end tools and utilities to feed data into the bottom tier. These back end tools and utilities perform the Extract, Clean, Load, and refresh functions.
- ▶ **Middle Tier** – In the middle tier, we have the OLAP Server that can be implemented in either of the following ways.
  - By Relational OLAP (ROLAP), which is an extended relational database management system..
  - By Multidimensional OLAP (MOLAP) model, which directly implements the multidimensional data and operations.
- ▶ **Top-Tier** – This tier is the front-end client layer. This layer holds the query tools and reporting tools, analysis tools and data mining tools.



**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

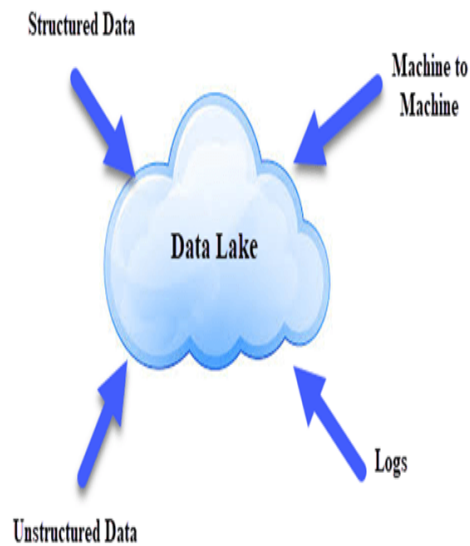
Raiya Road, Rajkot

Sr. No.	Ph. No. 9727753360 Questions	Rajkot - 360001 Answer
1.	1. _____ is a subject-oriented, integrated, time-variant, non-volatile collection of data in support	Data warehousing
2	The data Warehouse is_____.	Read only
3	Expansion for DSS in DW is_____.	Decision support system
4	The important aspect of the data warehouse environment is that data found within the data warehouse is	Subject oriented integrated& time variant
5	The time horizon in Data warehouse is usually _____.	5-10 years
6	The data is stored, retrieved & updated in _____.	OLTP
7	_____describes the data contained in the data warehouse.	Meta data
8	_____predicts future trends & behaviours, allowing business managers to make proactive, knowledge-driven decisions.	Data mining
9	_____ is the heart of the warehouse.	Data warehouse database servers
10	_____ is the specialized data warehouse database.	Redbrick
11	_____ is held in the catalogue of the warehouse database system.	Algorithmic level metadata
12	_____maps the core warehouse metadata to business concepts, familiar and useful to end users.	Application level meta data
13	The source of all data warehouse data is the_____.	Operational environment
14	Data warehouse contains _____data that is never found in the operational environment.	Summary

Data lake



- ▶ A DATA LAKE is a storage area that holds a vast amount of raw data in its native format, including structured, semi-structured and unstructured data.
- ▶ The term was given by James Dixon, CTO at Pentaho.
- ▶ Amazon, Google, Microsoft, Tera data, Zeloni are the major companies in this area.



### Objective of Data Lake

- ▶ The idea of data lake is to have a single store of all data in the enterprise ranging from raw data to transformed data which is used for various tasks including reporting, Visualization, analytics and machine learning.
- ▶ With the increase in data volume, data quality, and metadata, the quality of analyses also increases.
- ▶ It offers competitive advantage to the organization.
- ▶ It offers business ability to move quickly for any decision.

### Tools of Data Lake



**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. 97277533

Rajkot - 360001

Ph.No-(0281)2471645



- Data Quality : Checking the correctness
- Metadata Management : Data above the data
- Security : Use protective tools, E-Governance
- Admin : Includes procedures , planning etc
- Audit/ Logging : Check out the effectiveness
- Data Life cycle management : Life cycle of data ( Share, Archive and Destroy)
- Data Scheduling : Scheduling the process (hourly , weekly, etc)
- Lineage : Know the stages of data ( Initially how was the data, then how has data been transformed)

### Importance of Data Lake

- ▶ The ability to collect more data, from more sources, in less time, and empowering users to collaborate and analyze data in different ways leads to better, faster decision making.
- ▶ Improved customer interactions



Lalpari Campus, Near Lalpari Lake,

B/H Marketing Yard,

Amargadh (Bhichri), Rajkot

3 – Vaishalinagar

Nr. Amrapali Railway Crossing

Raiya Road, Rajkot

Rajkot - 360001

Ph.No–(0281)2471645

- ▶ Improve R&D Innovation choices
- ▶ Increase operational efficiencies

## Various Perspectives of managing Data

### 1. Data Lifecycle Perspective

- ▶ Data should be acquired and maintained only to meet a specific need.

Various phases of DLC are as follows.

- ▶ Plan: A documented sequence of intended actions to identify and secure resources & gather, maintain and utilize data.
- ▶ Acquire: Acquisition involves collecting data.
- ▶ Process: It includes actions performed on data to verify, organize and extract data in an appropriate output.
- ▶ Analyze: It involves actions performed on data that help describe facts, detect patterns and test hypothesis.
- ▶ Preserve: It involves actions to keep data for some period of time.
- ▶ Publish/share: The ability to prepare and disseminate data to public and to other agencies.
- ▶ Describe (documentation, Metadata): Updated and complete metadata and documentation to maintain quality of data.
- ▶ Manage quality: Methods must be employed to ensure that data are properly collected, handled, used at all stages of DLC. This is commonly known as “QA/QC”.
- ▶ Back up & security: It is a process to protect data from accidental data loss. It includes making additional copies of data files or database.



Lalpari Campus, Near Lalpari Lake,

B/H Marketing Yard,

Amargadh (Bhichri), Rajkot

3 – Vaishalinagar

Nr. Amrapali Railway Crossing

Raiya Road, Rajkot

Rajkot - 360001

## 2. Data Storage Perspective

The data storage is the process of storing the data on computer's hard drive, or any other storage medium.

Examples of storage area- CPU, RAM, Hard drive, optical drivers etc.

## 3. Data Processing and Analysis Perspective

- Data processing is the prime requirement of the businesses for generating valuable insights from data.
- Computers have been used since decades for processing of data.
- Since the era of mainframe PCs, mini PCs, Desktop PCs, laptops and now smart phones.
- For analysis, company can use spreadsheets, ETL Tools, Data Mining Tools, Big Data Analytics Tools etc.

## 4. Data from Decision Support Perspective

- Measuring the value of data is a boundless process with endless options – whether it is structured or unstructured.
- Following are the criteria for analyzing the value of information.
  - Effectiveness
  - Efficiency
  - Context

## 5. Data from Quality Management Perspective

- The organization need to create some checks to control data quality.





Lalpari Campus, Near Lalpari Lake,

3 – Vaishalinagar

B/H Marketing Yard,

Nr. Amrapali Railway Crossing

Amargadh (Bhichri), Rajkot

Raiya Road, Rajkot

- The quality management process shall take into consideration following criteria.

- Accuracy
- Validity
- Reliability
- Timeliness
- Completeness

6. Data from Technological Influence Perspective

- It includes impact of data technologies impacting the business value. For example,
- Internet on Things (IoT) : It is the network of physical devices, vehicles, home appliances and other items embedded with electronics, software, sensors, actuators, and connectivity which enables these objects to connect and exchange data
- Cloud storage: The ability of storing data on cloud.
- Machine Learning: The ability of machines to make the decisions without human interventions.

Sr. No	Questions	Answer
1	What is a storage area that holds a vast amount of raw data in its native format, including structured, semi-structured and unstructured data?	Data lake
2	A Data lake is a storage area that holds a vast amount of raw data in its native format, including which kind of data?	structured, semi-structured and unstructured data
3	Data lake, the term was given by Whom?	James Dixon



**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

4	James Dixon given which term?	Data lake
5	Which are the major companies in the area of Data lake?	Amazon, Google, Microsoft, Tera data, Zeloni
6	With the increase in data volume, data quality, and metadata, the _____ of analyses also increases in data lake.	Quality
7	With which kind of data quality of data lake increases?	increase in data volume, data quality, and metadata
8	How many tools a data lake have?	Eight
9	What is identifies by data quality?	Correctness of data
10	Correctness of data is identified by what?	Data quality
11	Admin tool in data lake can be used for?	Planning and procedure
12	Which tool of data lake for planning and procedure can be used for?	Admin tool
13	The data storage is the process of storing the data on what?	Computer's hard drive
14	What is the use of computer hard drive?	Data storage
15	For analysis perspective which tools are used?	ETL Tools, Data Mining Tools, Big Data Analytics Tools etc
16	Which are the criteria for analyzing the value of information?	Effectiveness, efficiency and context
17	In which perspective data found like; accuracy, validity, reliability, timeliness and completeness?	Data from quality management perspective
18	The data which is used for technological purpose, what it is called?	Data from technological influence perspective
19	The ability of storing data on cloud is called?	Cloud storage
20	The ability of machines to make the decisions without human interventions is what?	Machine learning



Lalpari Campus, Near Lalpari Lake,

B/H Marketing Yard,

Amargadh (Bhichri), Rajkot

Ph. No. 9727753360

3 – Vaishalinagar

Nr. Amrapali Railway Crossing

Raiya Road, Rajkot

Rajkot - 360001

Ph.No-(0281)2471645

## Business Reporting, Visual Analytics:

### □ Definition, concepts

- Definition

Business reporting is at the heart of strong and sustainable organizations, financial markets, and economies.

It's the process of converting data into information.

Reporting is the structuring of information in such a way that it can be used to measure and monitor business performance.

- Business reporting or enterprise reporting refers to both "the public reporting of operating and financial data by a business enterprise, "and "the regular provision of information to decision-makers within an organization to support them in their work.”
- It is a fundamental part of the larger movement towards improved business intelligence and knowledge management.
- Implementation often involves extract, transform, and load (ETL) procedures in coordination with a data warehouse and then using one or more reporting tools.
- Reports can be distributed in print form, via email or accessed via a corporate intranet.
- Reporting can also be used for verification and cross-checks. Audit teams like FINRA and SEC adhere to reports for all business firms.



Lalpari Campus, Near Lalpari Lake,

3 – Vaishalinagar

B/H Marketing Yard,

Nr. Amrapali Railway Crossing

Amargadh (Bhichri), Rajkot

Raiya Road, Rajkot

- Standard Business Reporting is a group of international programs instigated by a number of governments with the end of make business the centre when it comes to managing business-to-government reporting obligations
- Reports are all about data. Without accurate data, there's no way to produce accurate reports.
- But times are changing. Almost any business application you need is now available in the cloud. That means it's online, so you can access it from anywhere and at any time – using a tablet, laptop or smartphone.
- So a lot of your business data can now be stored online. This makes reporting much easier.

#### FIND THE BEST REPORTING SOFTWARE

- Ease of use

Reporting software should be simple and straightforward, giving you the information you need quickly.

- Compatibility

Reporting software needs to connect to as many of your business applications as possible. That means it should be compatible with your accounting, inventory management and other soft wares. The more data you can 'feed' it, the more accurate the reports will be.

- Flexibility

Good reporting software will be flexible enough to give you the reports you want.

#### WAYS TO USE BUSINESS REPORTS

- Budgeting



Lalpari Campus, Near Lalpari Lake,

3 – Vaishalinagar

B/H Marketing Yard,

Nr. Amrapali Railway Crossing

Amargadh (Bhichri), Rajkot

Raiya Road, Rajkot

- With accurate costs you can create an accurate budget. That means you won't over spend or under spend in important areas.

Ph.No-(0281)2471645

- Staffing

Identify your best workers, improve training plans and set some realistic goals for the future.

- Client engagement

Generate reports of jobs completed for a client, or number of hours worked.

Use these reports at client meetings to enhance your business relationship.

- Managing stock

Once you know how product sales vary on a seasonal basis, you can improve inventory management so you never buy too little – or too much.

- Applying for funding

Accurate financial reports are more likely to convince a bank or other lender to invest in your business.

- Seeing the big picture

You can find out which jobs are in progress, and at what stage, at any time.

This enables you to stay on track. It lets you see what's really happening in your business.

### Tools for Business Reporting

- Zoho Analytics
- Agency Analytics
- Wrike
- Tap Reports
- Altair Smart Sight



Lalpari Campus, Near Lalpari Lake,

B/H Marketing Yard,

Amargadh (Bhichri), Rajkot

3 – Vaishalinagar

Nr. Amrapali Railway Crossing

Raiya Road, Rajkot

Rajkot - 360001

Ph.No–(0281)2471645

- **Spotlight Reporting** 9727753360

## □ Different types of charts and graphs

- Line Chart
- Bar Chart
- Pie Chart
- Scatter Plot
- Bubble Chart
- Tree maps
- Dual axis bar chart
- Dual axis line chart
- Line Chart
- A line chart shows the relationship of one or more measures over some interval, such as time or a series of ranges. You can measure a single measure or you can show the relationships among multiple measures (multivariate analysis), such as the leading or lagging relationship between advertising and sales over time.



**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

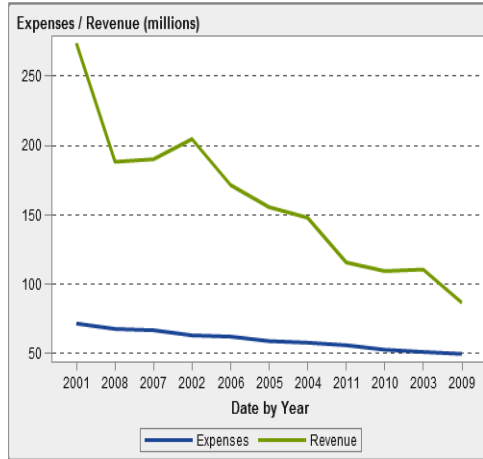
Amargadh (Bhichri), Rajkot

Raiya Road, Rajkot

Ph. N *A Line Chart*

Rajkot - 360001

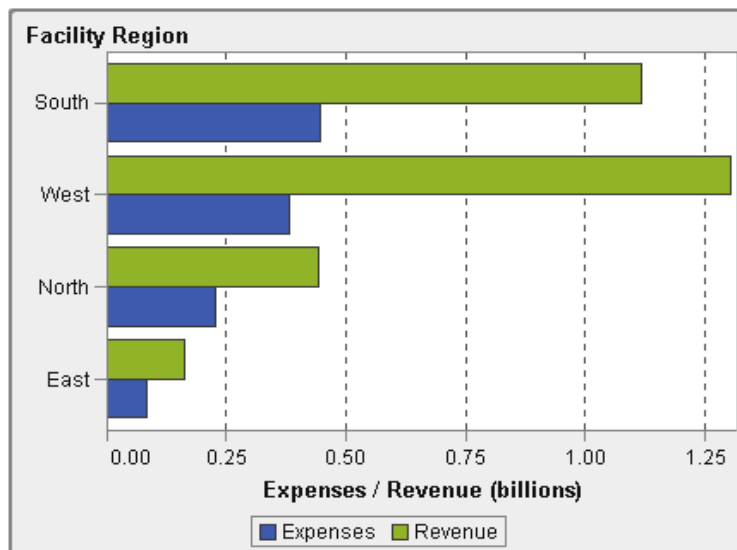
No-(0281)2471645



- Bar Chart
- A bar chart consists of vertical or horizontal bars that represent quantitative data.

Use bar charts to compare data that is aggregated by the distinct values of a category.

*A Bar Chart*





Lalpari Campus, Near Lalpari Lake,

B/H Marketing Yard,  
Amargadh (Bhichri), Rajkot

Ph. No. 9727753360

3 – Vaishalinagar

Nr. Amrapali Railway Crossing

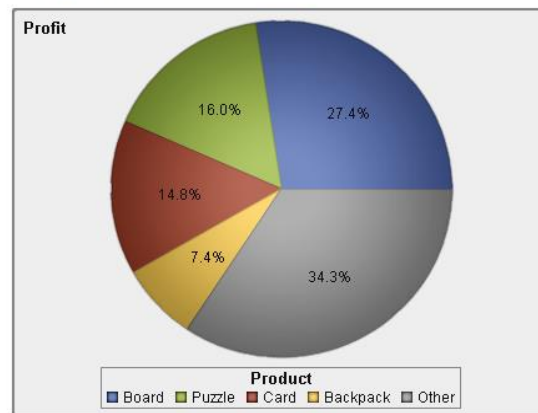
Raiya Road, Rajkot

Rajkot - 360001

Ph.No-(0281)2471645

- Pie Chart
- A pie chart displays a part-to-whole relationship in a circle divided into multiple slices for each value of a category data item based on a single measure data item.
- Each slice represents the relative contribution of each part to the whole.
- Effective pie charts limit the number of slices to 5 or 6.

A Pie Chart



- Scatter Plot
- A scatter plot is a two-dimensional plot
- It graphically shows the relationship between two variables in a process.
- Each marker (represented by a symbol such as a dot, a square, or a plus sign) represents an observation.
- Identifies a pattern that may cause a quality problem.
- Use a scatter plot to examine the relationship between numeric data items.





Lalpari Campus, Near Lalpari Lake,

3 – Vaishalinagar

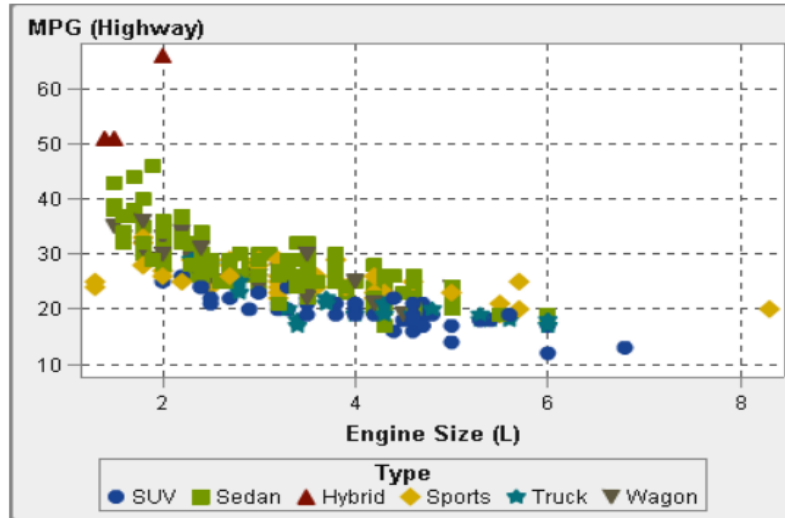
B/H Marketing Yard,

Nr. Amrapali Railway Crossing

Amargadh (Bhichri), Rajkot

Raiya Road, Rajkot

### A Scatter Plot



- Bubble Chart
- A bubble plot is a variation of a scatter plot in which the markers are replaced with bubbles.
- A bubble plot displays the relationships among at least three measures.
- Two measures are represented by the plot axes, and the third measure is represented by the size of the bubbles.
- Each bubble represents an observation.
- A bubble plot is useful for data sets with dozens to hundreds of values.



Lalpari Campus, Near Lalpari Lake,

3 – Vaishalinagar

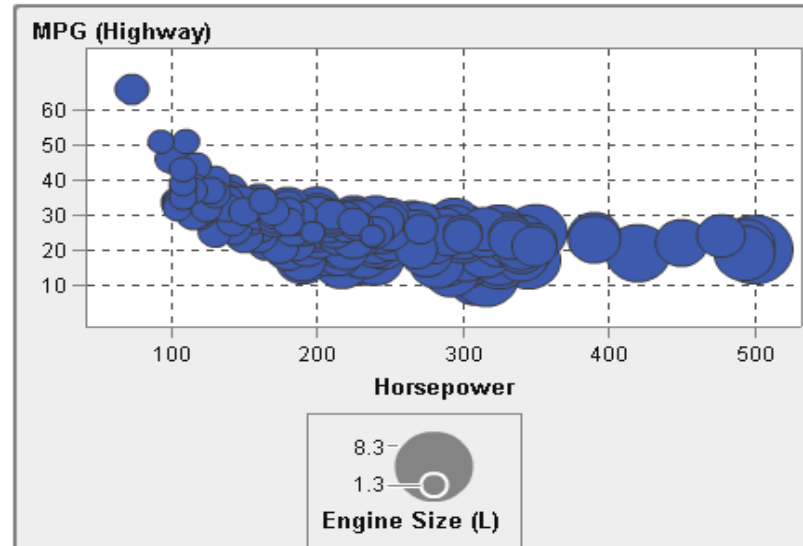
B/H Marketing Yard,

Nr. Amrapali Railway Crossing

Amargadh (Bhichri), Rajkot

Raiya Road, Rajkot

### A Bubble Plot



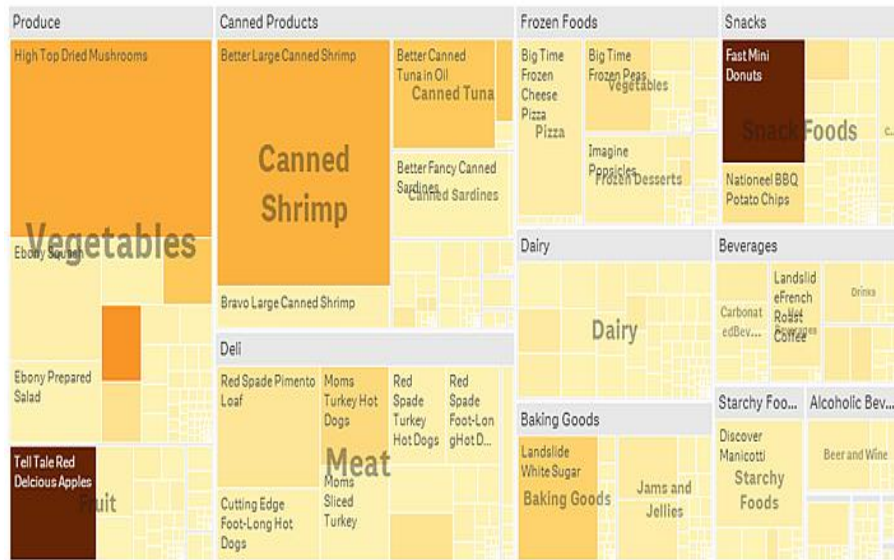
### Tree Maps

- A treemap displays your data as a set of rectangles (called tiles). Each tile represents a category node or a hierarchy node.
- The color of each tile represents the value of the first measure. The size of each tile represents the value of the second measure. (There are two data roles for measures in a treemap—Size and Color.)
- Use a treemap when space is constrained and you have a large amount of hierarchical data that you need to get an overview of. Treemaps should primarily be used with values that can be aggregated.



Lalpari Campus, Near Lalpari Lake,  
 B/H Marketing Yard,  
 Amargadh (Bhichri), Rajkot

3 – Vaishalinagar  
 Nr. Amrapali Railway Crossing  
 Raiya Road, Rajkot



- In this image you have several product groups, such as Produce, Canned Products, and Frozen Foods. Each product group consists of a large rectangle.
- You can regard the product groups as branches of the tree. When you select a product group, you drill down to the next level, the product type, for example, Vegetables, Meat, and Dairy. You can regard the product types as sub-branches of the tree. The branches have leaves.
- A leaf node's rectangle has an area proportional to a specific dimension of data
- In this example, the items Ebony Squash, Bravo Large Canned Shrimp, Red Spade Pimento Loaf, and so on, are the leaves. The leaf nodes are colored to show a separate dimension of the data.

□ **Emergence of data visualization and visual analytics**

- Visual analytics works towards representing data in an easily understandable format.



Lalpari Campus, Near Lalpari Lake,

3 – Vaishalinagar

B/H Marketing Yard,

Nr. Amrapali Railway Crossing

Amargadh (Bhichri), Rajkot

Raiya Road, Rajkot

- This helps in the easier understanding of complex data and facilitates reasoning and decision-making based on large and complex data sets.
- Visual analytics does not work with raw and unstructured data. It employs data mining algorithms to cleanse the data and then decides how to display the data.
- For visual analytics, the data is first evaluated using software tools and evaluation models, methods, and theories and involves both the users and tasks along with the data.
- Visual analytics is both data-driven and user-driven.
- Since visual analytics assists in applying human judgments to reach evidence-based conclusions.
- Data visualization and visual analytics definitely are not the same thing. At the same time, they are two parts of the same coin that aim to make data more understandable and more effective and hence more usable and make good use of the sea of data at our disposal.

### Emergence of data visualization and visual analytics

- Pre-17th Century: Early maps and diagrams
- 1600-1699: Measurement and theory
- 1700-1799: New graphic forms
- 1800-1850: Beginnings of modern graphics
- 1850–1900: The Golden Age of statistical graphics



Lalpari Campus, Near Lalpari Lake,

B/H Marketing Yard,

Amargadh (Bhichri), Rajkot

3 – Vaishalinagar

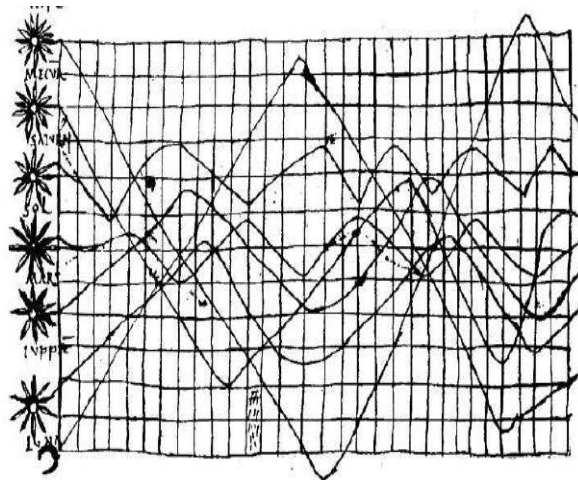
Nr. Amrapali Railway Crossing

Raiya Road, Rajkot

Rajkot - 360001

Ph.No-(0281)2471645

- 1900-1950: The modern dark ages
- 1950-1975: Re-birth of data visualization
- 1975-present: High-D, interactive and dynamic data visualization
- Pre-17th Century: Early maps and diagrams
- The earliest seeds of visualization arose in geometric diagrams
- In the 14th century, the idea of a plotting a theoretical function (as a proto bar graph), and the logical relation between tabulating values and plotting them appeared in a work followed somewhat later by the idea of a theoretical graph of distance vs. speed.
- By the 16th century, techniques and instruments for precise observation and measurement of physical quantities, and geographic position were well-developed (for example, a “wall quadrant” constructed by Tycho Brahe [1546-1601])



- 1600-1699: Measurement and theory



Lalpari Campus, Near Lalpari Lake,

3 – Vaishalinagar

B/H Marketing Yard,

Nr. Amrapali Railway Crossing

Amargadh (Bhichri), Rajkot

Raiya Road, Rajkot

- Among the most important problems of the 17th century were those concerned with physical measurement— of time, distance, and space for astronomy, surveying, map making, navigation and territorial expansion.
- This century also saw great new growth in theory and the dawn of practical application— the rise of analytic geometry and coordinate systems, the birth of probability theory (Pascal and Fermat), and the beginnings of demographic statistics (John Graunt) and “political arithmetic” (William Petty)— the study of population, land, taxes, value of goods, etc. for the purpose of understanding the wealth of the state.
- By the end of this century, the necessary elements for the development of graphical methods were at hand— some real data of significant interest, some theory to make sense of them, and a few ideas for their visual representation. Perhaps more importantly, one can see this century as giving rise to the beginnings of visual thinking.
- 1700-1799: New graphic forms
- Johann Lambert [1728–1777] introduced the ideas of curve fitting from empirical data points.
- He used various sorts of line graphs and graphical tables to show periodic variation, for example, in air and soil temperature.
- William Playfair [1759–1823] is widely considered the inventor of most of the graphical forms widely used today— first the line graph and bar chart later the pie chart and circle graph.



Lalpari Campus, Near Lalpari Lake,  
 B/H Marketing Yard,  
 Amargadh (Bhichri), Rajkot

3 – Vaishalinagar  
 Nr. Amrapali Railway Crossing  
 Raiya Road, Rajkot

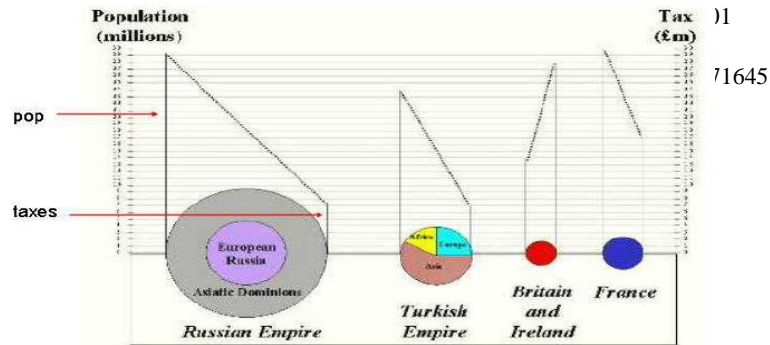


Figure 6: Re-drawn version of a portion of Playfair's 1801 pie-circle-line chart, comparing population and taxes in several nations.

- 1800-1850: Beginnings of modern graphics
- With the fertilization provided by the previous innovations of design and technique, the first half of the 19th century witnessed explosive growth in statistical graphics and thematic mapping.
- In statistical graphics, all of the modern forms of data display were invented: bar and pie charts, histograms, line graphs and time-series plots, contour plots, scatterplots, and so forth.
- In 1801, the first geological maps were introduced in England by William Smith [1769–1839].



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

Lalpari Campus, Near Lalpari Lake,

3 – Vaishalinagar

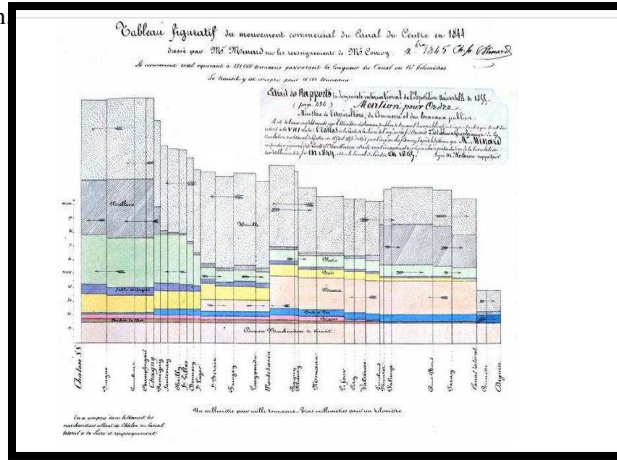
B/H Marketing Yard,

Nr. Amrapali Railway Crossing

Amargadh (Bhichri), Rajkot

Raiya Road, Rajkot

Ph



360001

81)2471645

- 1850–1900: The Golden Age of statistical graphics
- By the mid-1800s, all the conditions for the rapid growth of visualization had been established.
- In Golden Age, there occurred unparalleled beauty and many innovations in graphics and thematic cartography.
- The French Albums de Statistique Graphique were discontinued in 1897 due to the high cost of production
- 1900-1950: The modern dark ages
- If the late 1800s were the “golden age” of statistical graphics and thematic cartography, the early 1900s can be called the “modern dark ages” of visualization.
- Graphic innovation was also awaiting new ideas and technology:





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

Lalpari Campus, Near Lalpari Lake,

3 – Vaishalinagar

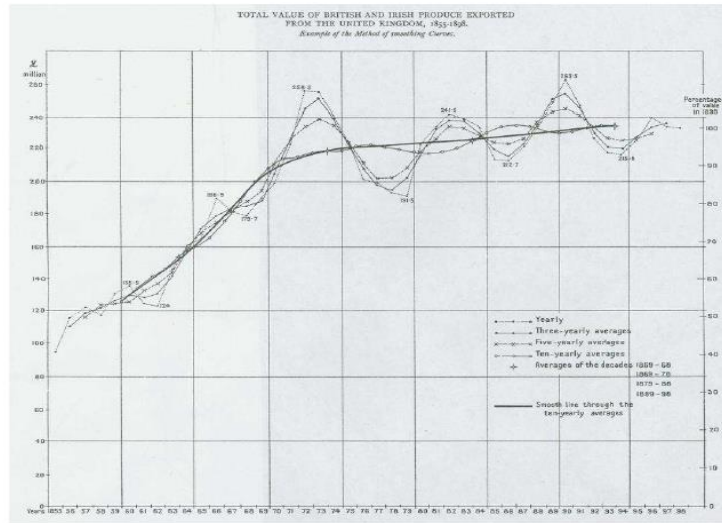
B/H Marketing Yard,

Nr. Amrapali Railway Crossing

Amargadh (Bhichri), Rajkot

Raiya Road, Rajkot

- The development of the machinery of modern statistical methodology, and the advent of the computational power and display devices which would support the next wave of developments in data visualization.



- 1950–1975: Re-birth of data visualization
- Data visualization began to rise from dormancy in the mid 1960s.
- Shortly, there occurred the invention of a wide variety of new, simple, and effective graphic displays, under the rubric of “Exploratory Data Analysis” (EDA)— stem-leaf plots, boxplots, hanging rootograms, two-way table displays, and so forth, many of which entered the statistical vocabulary and software implementation.
- By the end of this period significant intersections and collaborations would begin: (a) computer science research (software tools, C language, UNIX, etc.) at Bell Laboratories (Becker,1994) (b) developments in data analysis.



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

- By the close of this period, the first discovery of modern GIS and interactive systems for 2D and 3D statistical graphics would appear. These would set goals for future development and extension.
- 1975–present: High-D, interactive and dynamic data visualization
- During the last quarter of the 20th century, data visualization has blossomed into a mature.
- Various software tools for a wide range of visualization methods and data types are available for every desktop computer were invented.
- With this disclaimer, a few major themes stand out:
  - New methods for visualizing high-dimensional data
  - The development of highly interactive statistical computing systems
  - Extensions of classical linear statistical modelling to ever wider domains

### Short Question Answer

Sr. No.	Question	Answer
1	Data can be visualized using?	Graphs, charts and maps
2	Data visualization is also an element of the broader with what?	Data presentation architecture
3	Which method shows hierarchical data in a nested format?	Tree maps



**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

4	Which is used to find inference for 1 proportion using normal approx?	prop.test() Ph.No-(0281)2471645
5	Which is used to find the factor congruence coefficients?	factor.congruence
6	Data visualization decrease the insights and take slower decisions. Statement is true or false	False
7	Common use cases for data visualization include?	Politics, Sales, marketing & Healthcare
8	Which of the following plots are often used for checking randomness in time series?	Autocorrelation
9	Data visualization is related with what?	Pictorial representations
10	Which of the following are Use of data visualization?	See context of data, Clear data understanding and finding pattern in data
11	Which of the following value is provided by kind keyword for barplot?	Bar
12	You can create a scatter plot matrix using the __ method in pandas.tools.plotting.	scatter_matrix
13	Plots may also be adorned with error bars or tables.	True



**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

14	Which of the following plots are often used for checking randomness in time series?	Autocorrelation
15	Where plots are used to visually assess the uncertainty of a statistic	Bootstrap
16	Which of the following is not a challenge in Big Data Visualization?	Version
17	Which of the following is not a problem in Big Data Visualization?	Scaled Data
18	Which of the following is a problem in Big Data Visualization?	Visual Noise
19	Which of the candidate is suitable for interactive visualization?	Type of Visual, Cardinality & Size of data
20	Visual Mapping is important for what?	Remapping
21	Data visualization techniques are:	Scatter Plot, Line Chart & Pie Chart
22	Column graph is another name for what?	Bar Chart
23	Graphic innovation was also awaiting new _____	ideas and technology
24	The French Albums de Statistique Graphique were discontinued in _____ due to the high cost of production.	1897
25	EDA stands for _____	Exploratory Data Analysis



**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

26	_____ is data that represents other forms of data using specific machine language systems that can be interpreted by various technologies.	Digital data
27	The most fundamental of these systems is a _____	binary system
28	The goal of _____ is to create a trove of historical data that can be retrieved and analyzed to provide useful insight into the organization's operations.	data warehousing
29	in a series of binary characters, traditionally _____	ones and zeros, or "on" and "off" values.
30	_____ is the secure electronic storage of information by a business or other organization.	Data warehousing
31	A data warehouse never put emphasis only _____	current operations.
32	A data warehouse is always a _____	subject oriented
33	Data warehousing is a vital component of _____	business intelligence.
34	A good _____ system makes it easier for different departments within a company to access each other's data.	data warehousing
35	A data warehouse is built by _____ from various sources of data such that a mainframe and a relational database.	integrating data



**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

36	Data is _____ at particular intervals.	Rajkot - 360006 Ph.No-(0281)2471645	read-only and refreshed
37	How many types of data operations done in the data warehouse?		Two types
38	The term <i>data lake</i> is often associated with _____ object storage.		Hadoop-oriented
39	_____ can encompass hundreds of terabytes or even petabytes, storing replicated data from operational sources, including databases and SaaS platforms.		Data lakes
40	Data stored in a _____ can be anything, from completely unstructured data like text documents or images, to semistructured data such as hierarchical web content, to the rigidly structured rows and columns of relational databases.		Data lake
41	Two types of data operations done in the data warehouse are: _____		Data Loading Data Access
42	A _____ is a storage repository that holds a vast amount of raw data in its native format until it is needed.		data lake
43	data lake provides a single place to save and access _____		valuable enterprise data.
44	From the following option which option is true for feature of data lake?		Sophisticated access control mechanisms.



**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

45	_____ must be able to set permissions for keeping data secure and private when and where it needs to be.	Data owners Ph.No: (0281)2471645
46	A _____ is an impersonal presentation of facts.	report
47	A _____ may be defined as an organized statement of facts or events or any situation relating to business or commercial interests prepared after an investigation and presented to the interested persons with or without recommendations.	business report
48	A business report is usually a type of _____	upward communication
49	_____ as an objective-based and orderly communication of the information having facts for serving a specific business purpose.	business report
50	From the following option which option is true for business report?	Correct and unbiased facts
51	_____ reports are based upon a specific purpose. These are of two types, i.e. ordinary or daily reports and special-purpose reports.	Importance-based or frequency-based reports
52	_____ types of reports are general reports that are prepared and forwarded to management or higher authority as a routine activity or at specific intervals.	Ordinary or daily routine reports



**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

53	_____ business report is written and forwarded to the senior officials or management for a special purpose or on their special request	Special purpose report
54	The report that is prepared in a specific or pre-determined format and as per well-established process is known as a _____ in business.	formal report
55	Which report may be submitted either daily or weekly, quarterly, monthly or annually, etc?	Ordinary or daily routine reports
56	_____ type of report includes a user-specific format or structure that is according to the convenience of the writer and submitted directly to the desired authority as and when required.	Informal business report
57	_____ is based on the maximum available information related to any matter or situation.	Informative business report
58	Which types of reports include the views or recommendations of the report writer based on facts?	Special purpose report
59	_____ includes different related facts like the cause of any issue occurred, opinions, required actions, and recommendations of the reporter.	Interpretation-based report
60	_____ type of report is a complete record of concerns or matters discussed in the meeting.	Discussion-based Reports
61	_____ is the summary of details discussed during any meeting to send information for the press	Summary reports





**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

	release of shares with shareholders of the organizations.	Rajkot - 360001 Ph.No-(0281)2471645
62	_____contain analytical information and facts.	Analytical reports
63	Which report include data and information, the presentation of the data in tabulated form, explanation based on analytics, and findings or interpretations?	Analytical reports
64	Frequency curve is _____	Asymptotic to x-axis
65	A Histogram containing a set of _____	Adjacent Rectangles
66	A frequency curve touches x-axis _____	No
67	A histogram is _____	A frequency graph
68	Which of the following is an example of compressed data?	Frequency distribution
69	For geographically base data, the bars are used:	Horizontal
70	A(n) _____ is a graphical representation in which the sample space is represented by a rectangle and events are represented as circles	Venn diagram
71	In a Pie chart, usually, the arrangement of the sectors is:	Anti-clockwise
72	A graphical device used for enumerating sample points in a multiple-step experiment is a	<ul style="list-style-type: none"> <li>• Bar chart</li> <li>• Pie chart</li> </ul>



**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)247164	• Histogram Ans: None of these
73	If the frequency curve has a longer tail to left, the distribution is		Negatively skewed
74	A circle in which sectors represents various quantities is called		Pie chart
75	Decumulative frequency is presented by		Less than ogive
76	Component bar charts are used when data is divided into:		Groups
77	A graphical method of representing the sample points of a multiple-step experiment is		A tree diagram
78	_____ represent diagrams of a mathematical or statistical function, while a chart is a graphical representation of the data		Graphs
79	The _____ representations are common methods to get visual inspection about data.		graphical & charting
80	The most commonly used graphical summaries of the data are _____		bar charts, histograms, pie charts, line graphs.



**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

81	A population value that is of interest to us and that we would like to estimate is called:	Ph.No-(0281)2471645
82	A summary measure that is computed from a sample to describe a characteristic of the population is called	A statistic
83	If you have data on the yearly average temperature at cape town international airport from 1900 to 2000, and if you are particularly interested in change over time, what is the most effective graphical display?	Line chart
84	If you have data on house prices and the distance of each of those houses from the city centre, and you are curious whether there is an association between distance from the city centre and the price of the house, with which of the following graphical techniques could you most easily see whether there is indeed such a relationship?	A scatter diagram
85	The graph which shows the changes over a specific time period is called	line graph
86	The vertical axis of bar graph is also known as	y-axis
87	Considering the line graph, the x-axis represents	time period
88	Considering the line graph, the y-axis represents	subject of measurement
89	The horizontal axis of bar graph is also known as	x-axis
90	In order to develop a relative frequency distribution, each frequency count must be divided by:	the total number of data values.



**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

91	The most effective technique to display either continuous variables or discrete variables that have many possible outcomes is	a grouped data frequency distribution.
92	All of the following are criteria for constructing classes in a grouped frequency distribution except	All of the above are criteria for constructing classes in a grouped frequency distribution.
93	Classes in a frequency distribution that do not overlap so that a data value can be placed in only one class are said to be	mutually exclusive classes.
94	The data for an ogive is found in which distribution?	A cumulative relative frequency distribution
95	Which of the following is not a characteristic of bar charts?	Multiple variables must be graphed on separate graphs.
96	Which of the following is not a characteristic of stem & leaf diagrams.	Stem values are single digit numbers.



**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

97	Which of the following is true about line charts? <small>Ph.No. 071753310</small>	<small>Rajkot - 360008</small> <small>Ph.No-(0281)2471645</small> Straight lines connect consecutive points.
98	On a scatter diagram, what values are placed on the horizontal axis?	The independent variable
99	A line graph which is a whole unbroken line is called a:	Linear graph
100	Which point lies only on y-axis?	(0,-2)
101	The point (-2,-2) is:	Equidistant from x-axis and y-axis.