

ASSIGNMENT NO. 01

Q1) What is BI? Explain its components of Architecture?

Ans

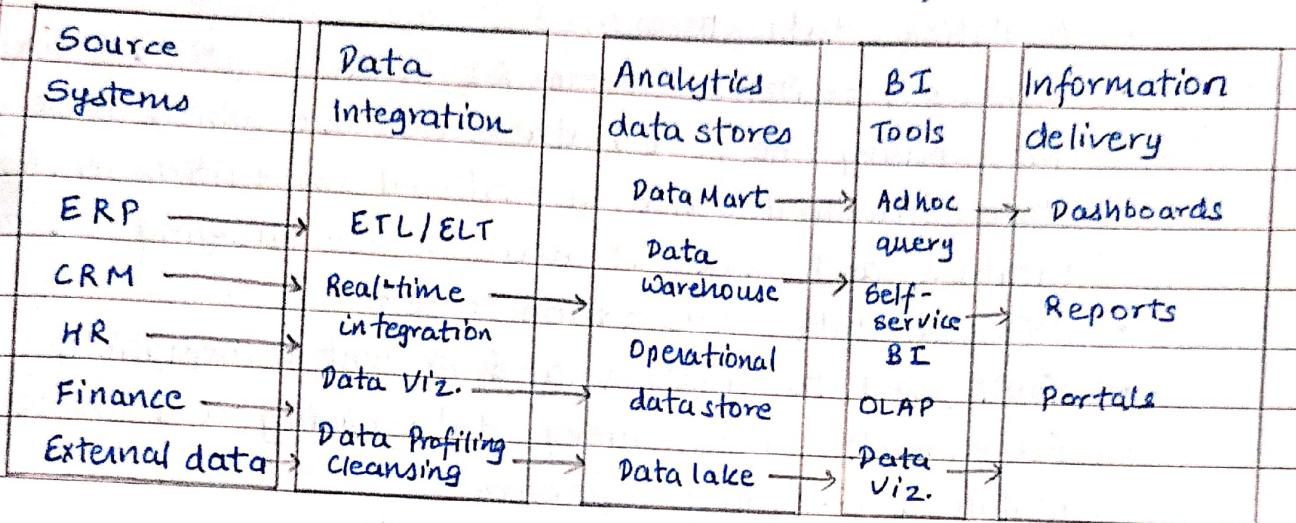


Fig: Architecture of BI

I) Business Intelligence [BI] :-

1. BI comprises the strategies and technologies used by enterprises for the data analysis and management of info.
2. Common function of business technologies include reporting, online analytical processing, analytics, dashboard development, data mining, process mining, predictive analytics and so on.
3. BI can handle large amount of structured as well as unstructured data to identify, develop and otherwise create new strategic business opportunities.

II) Components of BI Architecture :-

- **Source Systems:-**
These are all of the systems that capture and hold the transactional and operational data identified as essential for enterprise BI program.
- They also include secondary sources such as market data and customer databases from outside information providers.
- **Data integration and cleansing tools:-**
To effectively analyse the data collected for a BI program, an organization must integrate and consolidate different

data sets to create unified views of them.

- Most widely used data integration process is Extract, Transform and Load [ETL] software.
- Analytics data stores:-
- Various repositories where BI data is stored & managed.
- The primary one is a data warehouse which usually stores structural data in relational or multidimensional database and makes it available for querying and analysis.
- BI and data visualization tools:-
- Tools used to analyze and present information.
- For example ad hoc query, data mining and online analytical processing software.
- Dashboards, portals and reports:-
- Give business users visibility into the results of BI and analytics applications.
- Reports tend to present data in a more static form.

q2) Explain role of end user assumption and setting up data for BI?

Ans I] End-user assumptions:-

1. End-user assumption refers to the assumptions made by BI designers and developers about the end users of the BI solution.
2. It is important to involve end users in the design and development process to ensure assumptions made about their needs and requirements are accurate.
3. By involving end-users developers get a better understanding of users perspective and experiences, which can get more user friendly & effective solutions.
4. This can simply be done by taking continuous feedback.

II] Setting up data for BI:-

1. It refers to the process of collecting, organizing and preparing data for use in BI solutions.
2. This process involves identifying sources of data, cleaning and validating the data, and transformation of data into a format that can be used in BI solutions.
3. Quality and accuracy of data used in BI solutions can have significant impact on effectiveness of solution.

Q3) Explain DSS ? Definition and Development of DSS ?

- Ans 1. A decision support system [DSS] is a computer program application used to improve company's decision making process.
2. It Analyzes large amount of data and presents an organization with best possible options available.
 3. It brings together data and knowledge from different areas and source to provide users with information beyond usual reports and summaries.
- Development of DSS:-
 - A DSS consists of three parts:-
 - a) Knowledge database . b) Software c) User Interface
- a) Knowledge database:-
- An integral part of DSS database.
 - Contains information from both internal and external sources.
- b) Software System:-
- It is composed of management systems for models.
 - Organizations use model to predict how outcomes will change with different adjustments to system.
- c) User Interface:-
- It enables easy system navigation.
 - Primary goal of DSS's UI is to make it easier for the user to manipulate data stored in it.

Q4) Explain four stages of Simon's design making process?

Ans Simon's four stages of decision making process are :-

(i) Intelligence. (ii) Design. (iii) Choice. (iv) Implementation

i) Intelligence:-

- It consists of gathering information by examining reality, then identifying and defining the problem.
- In this phase problem ownership should also be established.

ii) Design:-

- Consists of determining alternatives and evaluating them.
- If evaluation requires construction of model it is completed in this phase as well.

iii) Choice:-

- This phase consists of selecting a tentative solution and testing its validity.

iv) Implementation:-

- It consists of putting the selected solution into effect.

Q5) Explain common strategies and approach for decision makers?

Ans Top three common strategies and approach for decision makers are as follows :-

I) ~~Intuitive-Emotional~~: Analytical decision making:-

- It uses logic, data and facts to make a rational decision.
- This strategy is an excellent choice if we have access to all the information you need to assess a situation accurately.

- It usually follows well ordered set / sequence of steps that help break complex problem into a simple one.

II) Command decision making:-

- Leaders who use this strategy make a choice without listening to input from other people.

- This approach is useful in urgent or time sensitive situations.
- It provides team members with a clear sense of direction in fast-paced work environments.

III) Collaborative decision-making :-

- It involves taking input from other people.
- Using this strategy one can help ensure that the opinions of team members are valued.
- Many leaders choose to incorporate feedback from clients, vendors and industry experts in their collaborative decision making process.

ASSIGNMENT 2

Q1) Explain data warehouse? Explain its type and architecture?

Ans 1. A data warehouse is a heterogeneous collection of different data sources organised under a unified schema.

2. With the evolution of technology and demands of the data-driven economy, multi-cloud architecture allows for the portability to relocate data and workloads as the business expands, both geographically and among the major cloud vendors.

- Types of data warehouse architecture:-

I] Single-tier architecture:-

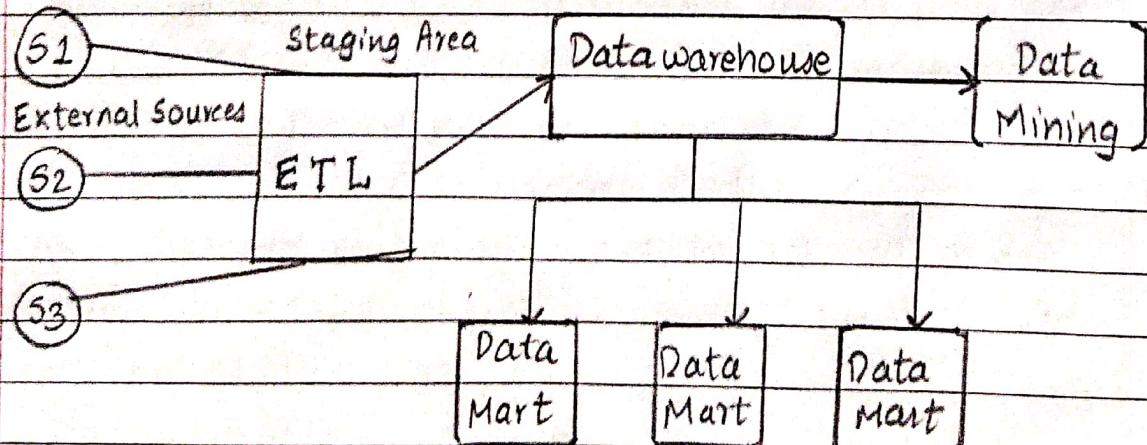
- It aims to deduplicate data to minimize the amount of stored data.

II] Two-tier architecture:-

- It separates physical data sources from the data warehouse making it incapable of expansion or supporting many end users.

III] Three-tier architecture:-

- The bottom tier, database of data warehouse servers.
- Middle tier, an online analytical processing (OLAP) server provides abstracted view of database for end users.
- The top tier, a front-end client layer consisting of the tools and API used to extract data.



Q6) What are the cubes? Explain Drill down, roll up, slice and Dice and pivot?

Ans :-

- Cubes :-
 - Grouping of data in a multidimensional matrices are called as Data cubes.
 - The general idea of this approach is to materialize certain expensive computations that are frequently required.
- Roll-up :-
 - This operation aggregates certain similar data attributes having the same dimensions together.
 - For example, if the data cube displays the daily income of a customer, we can use a roll-up operation to find the monthly income of his salary.
- Drill-down :-
 - This operation works opposite to the roll-up operation.
 - It allows us to take particular information and then subdivide it further for coarser granularity analysis.
- Slicing :-
 - This operation filters the unnecessary portions.
 - Suppose in a particular dimensions, the user does not need everything for analysis, rather a particular attribute.
- Dicing :-
 - This operation does a multidimensional cutting, that not only cuts only one dimension but also can go to another dimensions and cut a certain range of it.
 - As a result, it looks more like a subcube.
- Pivot :-
 - It transforms the data cube in terms of view.
 - It does not change the data present in the data cubes, it is very important from a viewing point of view.

Q4) Differentiate between OLAP and OLTP?

Ans Category

OLAP

OLTP

1. Definition

Online Analytical
Processing

Online Transaction
Processing

2. Known as

Online database query
management system.

Operational current
data.

3. Method
used

Data Warehouse.

Standard DBMS

4. Normalization

Not Normalized.

Normalized [3NF]

5. Volume of
data.

Relatively large [in
TB, PB]

Relatively small [in
GB, MB].

6. Speed of
queries

Relatively slow due to
large data.

Relatively fast.

Q5) Explain the following term DSS, MIS, EIS, Dashboard.

Ans

- DSS [Decision Support System] :-

- Computer program application used to improve a company's decision making process.

- It analyzes large amount of data and presents an organization with the best decision possible.

- MIS [Management Information System] :-

- A department that oversees hardware and software solutions dedicated in making critical business decision.

- These information systems are designed to gather and analyze data and produce business reports that help with crucial decision making.

- EIS [Executive Information System] :-
- It is a decision support system (DSS) used to assist senior executive in decision making process.
- It does this by providing easy access to important data needed to achieve strategic goals in an organization.
- Dashboard :-
- A dashboard is a progress report and considered a form of data visualization.
- It often provides at-a-glance views of key performance indicators relevant to particular objectives.

q7) Explain ROLAP and MOLAP.

Ans • ROLAP [Relational Online Analytical Processing] :-

1. In ROLAP data is stored in the form of tables, columns & rows.
2. It can handle large volume of data using relational database model.
3. ROLAP lies between the front end and back end of the system and stores each warehouse data effectively.
4. ROLAP can handle large amount of data using relational database model.

• MOLAP [Multidimensional Online Analytical Processing] :-

1. MOLAP creates data cubes to retrieve data from multiple dimensions.
2. To handle data scattered matrix technology is used.
3. Speed of response to a query in MOLAP is very high and it also allows faster index to previously summarized data.
4. However storage utilization of MOLAP is low if the data sets are scattered.

NAME:- AFNAN ATTAR PRN:- F19112003 CLASS:- BE
 COMP. II SUBJECT:- LP VT [BI] ASSIGNMENT :- 01

Q1) What do you understand by BI?

Ans 1. BI stands for business intelligence.

2. It is a technology driven method that helps users analyze and visualize business data from thousands of data sources and share insights across their organizations.
3. It also provides other information that may help corporate executives, business managers and other users make business decision.
4. Common functions of BI include but are not limited to data mining, process mining, reporting, online analytical processing, prescriptive analysis.

Q2) What is Power BI?

- Ans 1. Power BI is a Business Analytics solution or a Business Intelligence and Data Visualization tool.
2. It is developed by Microsoft in the year 2014.
 3. It enables users to convert data from thousands of data sources into interactive dashboards and BI reports.
 4. It is a cloud-based self service BI tool that collects application, multiple software connectors and services used to organize raw data into important content.

Q3) What are the essential applications that use Power BI?

Ans Microsoft's Power BI is mostly used by the following applications.

Page No.	
Date	

- 1) Business & Data Analyst
- 2) Project Management Office (PMO)
- 3) Developer & Database Admin.
- 4) Data Scientists
- 5) IT Team / Professionals
- 6) Consumers for final report.

Q4) In how many formats Power BI is available in the market?

Ans It is available in 3 formats:-

- 1) Power BI desktop:- For desktop computers.
- 2) Power BI service:- Online SaaS.
- 3) Mobile Power BI apps:- For Android/iOS devices.

Q5) What do you mean by the term Power BI Desktop?

- Ans:
- 1. Power BI Desktop is a free business intelligence app.
 - 2. It is used to install on the computer system to work with Power BI services.
 - 3. It offers shaping, modelling, data exploration, and report creation using high interactive visualization.
 - 4. It allows us to save all our work to a file to publish reports and data to the Power BI site so that we can share it to others too.

NAME:- AFNAN ATTAR PRN:- F19112003 CLASS:- BE COMP II
 SUBJECT:- LPVI [BI] ASSIGNMENT NO. :- 02

Q1) Explain how ETL works?

Ans Let us see what happens in each step of the process:-

I) Extraction:-

1. During data extraction, raw data is copied or exported from source location to a staging area.
2. Data management teams can extract data from various data sources that can be structured or unstructured.
3. Sources include but are not limited to:-
 - SQL or NoSQL servers.
 - CRM and ERP systems
 - Flat Files, Emails, Web Pages.

II) Transform:-

1. In this process data is transformed and consolidated for its intended analytical use case.
2. It may involve:-
 - Filtering, cleaning, de-duplicating, validating and authenticating data.
 - Performing calculations, translations or summarization based on raw data.
 - Conducting audits to ensure data quality & compliance.
 - Removing, encrypting, or protecting data governed by industrial or governmental regulators.
 - Formatting the data into tables or joined tables to match the schema of the target warehouse.

III) Load:-

Page No.	
Data	

1. In this step the transformed data is moved from staging area into a target data warehouse.
2. It involves initial loading of all data followed by periodic loading of incremental data changes, and less often, full refreshes to erase and replace data in the warehouse.

Q2) What are the benefits and challenges of ETL?

- Ans :-
1. ETL solutions improve quality by performing data cleansing prior to loading the data to a different repository.
 2. A time-consuming batch operation, ETL is recommended for creating smaller target repository more often than not that require less frequent updating while other methods such as:-
- i) Extract, Load, Transform [ETL]
 - ii) Change Data Capture [CDC]
 - iii) Data Visualization
- are used to integrate increasingly larger volumes of data that changes or real-time data streams.