

Business Intelligence & its Tools

Assignment - 01

Sehpaathi
Date 08/10/21
Page No.

Section - A

1. What is Business Intelligence (BI)?
Business Intelligence is a set of processes, architectures, and technologies that convert raw data into meaningful information that drives profitable business actions.
2. Define Business Intelligence. Give some examples.
Business intelligence is the process by which enterprises use strategies and technologies for analyzing current and historical data. Ex:- data warehouses, dashboards, cloud data services etc.
3. Briefly discuss history of Business Intelligent.
Business intelligence ~~as it~~ is said to have evolved from the decision support systems (DSS) that began in the 1960s and developed throughout the mid-1980s.
4. Explain the concept of the data warehouse.
A data warehouse is a type of data management system that is designed to enable and support business intelligence activities & analytics.
5. what is offline Extract, Transform, & Load,

The process of extracting data from source systems and bringing it into the data warehouse is commonly called ETL, which stands for extraction, transformation, and loading.

6. What do you understand by data-mining engines?

Data mining engines helps to extract information from huge sets of data. It is the procedure of mining knowledge from data.

7. "Implementing a data mart can be less expensive than implementing a data warehouse". Elucidate.

Data warehouse is a large repository of data collected from different sources whereas Data mart is only subtype of a data warehouse, hence is less expensive.

8. What is the purpose of business intelligence systems?

The purpose of business intelligence systems is to help companies make better decisions by showing present and historical data within their business context.

9. What are the key advantages of business intelligence system?

- Fast and more accurate reporting.
- analysis as planning.
- Improved employee satisfaction and improved data quality.
- Benefits achieved include reducing costs.

10. Write the brief description of Business Intelligence Applications?

Business intelligence systems combine data gathering, data storage, and knowledge management with data analysis to evaluate and transform complex data into meaningful, actionable information.

Section - B

1) What is the multi-dimensional analysis?
Multidimensional analysis refers to the process commonly used in data warehousing applications of examining data using various combinations of dimensions.

Dimensions are categories used to classify data such as time, geography, a company's departments, product lines.

2) Briefly explain the dimension attribute properties.

A dimension attribute is a column in a dimension table. Each attribute describes a level of summary within a dimension hierarchy. The dimension elements define the hierarchical

relationship within a dimensional table. The dimension attributes describe the dimension elements in terms that are familiar to users.

- 3) Discuss about the key attribute of dimension attributes.

Each dimension contains a key attribute. The key attribute is the attribute in a dimension that identifies the columns in the dimension main table that are used in foreign key relationships to the fact table.

- 4.) What is the dimension hierarchy? Explain with example.

A dimension hierarchy denotes how data is organized at various levels of aggregation. An analyst uses a dimensional hierarchy to identify various trends at one level. For example, one possible hierarchy in the data dimension is Year > Quarter > Month > Day.

5. Discuss the various type of hierarchy. Two types of hierarchies are the subsumptive containment hierarchy and the compositional containment hierarchy. A subsumptive subsumes its children, and a compositional

hierarchy is composed of its children.

- 6) "Parent-child hierarchy is actually value-based hierarchy". Comment.

Parent-child hierarchies are often used to represent charts of accounts, stores, salespersons and such. Parent-child hierarchies have a peculiar way of storing the hierarchy in the sense that they have a variable depth. In this pattern we should have to use parent-child hierarchies to store budget, actual and forecast values in a report.

- 7) What is ragged hierarchy? Give the suitable example.

A ragged hierarchy is a user-defined hierarchy that has an uneven number of levels. A ragged hierarchy is different in that the logical parent of at least one member is not in the level immediately above the member.

Ex.: Consider a salesperson hierarchy that has two top-level aggregates, foreign and domestic.

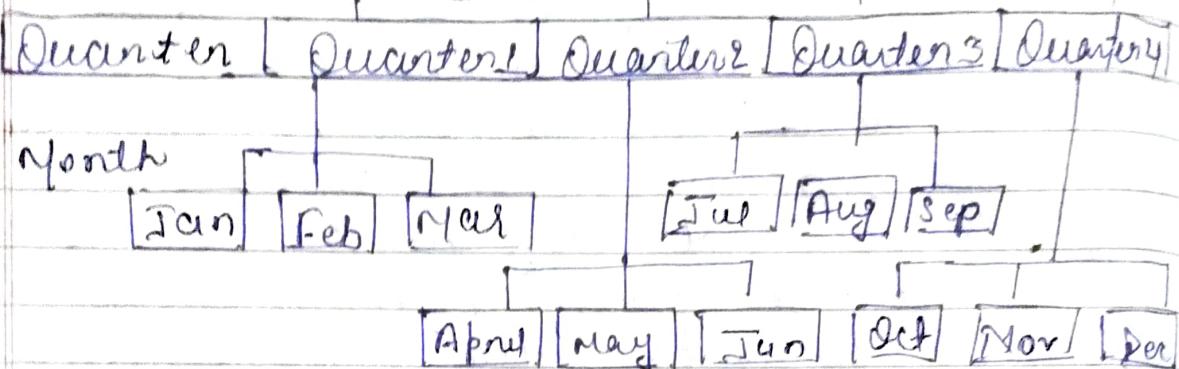
- 8) Explain the relationship between the levels and members of the

Calender dimension

The relationship b/w the levels & members of the calender dimensions

Levels Members

Year 2004



9) What is dimension?

A dimension is a structure that categorizes facts and measures in order to enable users to answer business questions.

Commonly used dimensions are people, products, place and time.

10) Explain the dimensional model.

A dimensional model is a database structure that is optimized for online queries and data warehousing tools. It is composed of fact and "dimension" tables.

Section - C

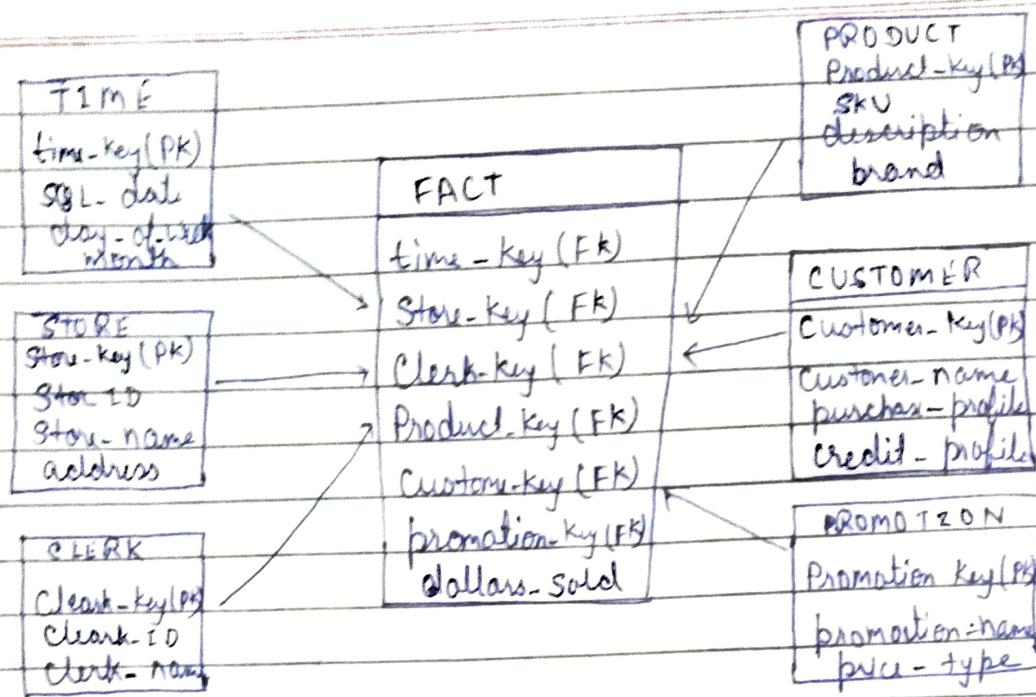
- 1) "Dimensions define the axis of enquiry of a fact." Elucidate.

Dimensions model comprises of a fact table and numerous dimensional tables and is used for assessing summarized data. Since Business Intelligence reports are used for assessing the facts (aggregates) across various dimensions, dimensional data modelling prefer the modelling technique in a BI environment.

Facts are normally calculated data like dollars' worth or sales or income. They correspond to the aim of a condensed support analysis
"Dimensions define the axis of enquiry of a fact".

For ex:- Product, Region and Time are the axis of enquiry of the sales detail.

- 2) Give examples of dimensional model.



3) what do you understand by fact table?

Fact table generally represent a process or reporting environment that is of value to the organization. It is important to determine the identity of the fact table and specify exactly what it represents. A fact table typically corresponds to an associative entity in the E-R model.

They must be listed in a logical fact table. Each measure has its own aggregation rules such as ADD, AVG, MIN or MAX.

Aggregation rules define the way by which business would like to contrast standards of a measure value.

4) Explain the types of measure and fact table.

Types of measure:-

Various types of measure in a fact table are:

- Additive - Measures that can be added across any dimensions are additive measures.
- Semi additive - Measures that can be added across only some dimensions are semi additive.
- Non Additive - Measures that cannot be added across any dimension.

Types of fact tables:

- Transactional: - A transactional table is the most basic and fundamental type of fact table.
- Periodic snapshots: It takes a picture of the moment, where the moment could be anything like performance summary of a salesman over the previous 3 months.
- Accumulating Snapshots: → In this type of fact table the activity of a process is shown.

5) "Dimension tables consist of attributes that describe fact records in the fact table." Discuss.

Dimension tables consist of attributes that describe fact records in the fact table. Some of these attributes provide descriptive information; others are used to specify how fact table data should be summarized to provide useful information to the person who is analysing the information. Every dimension has a set of descriptive attributes. Dimension tables contain attributes that describe business entities.

Example:- The client dimension can contain attributes like C-No., Area, State, Country etc.

Section - 2

- 1) Define the concept of Surrogate key. Also write down the advantages and disadvantages.

A Surrogate key in a database is a unique identifier for either an entity in the modelled world or an object in the database. The surrogate key is not derived from application data. Surrogate keys are keys that are maintained within the data warehouse instead of keys taken from source data systems.

Advantages of Surrogate keys.

- Immutability:- Surrogate keys do not change while the row exists.

- Change in Requirements : Attributes that uniquely recognize an entity might change over the time.
- Performance :- Surrogate keys tend to be a compact data type, such as as a four-type integer.
- Uniformity : → when every table has a uniform surrogate key.
- Validation : → It is possible to design key-values that are in co-ordination with a well-known pattern which can be automatically verified.

Disadvantages of Surrogate Keys :-

- Referential Integrity : → Referential integrity must be maintained between all dimension tables.
- Shared Dimensions : → To maintain consistency dimension tables that are shared are created.

2) what do you understand by alternative tables used in data warehousing?

Auxiliary Table

This table is created with the SQL statements CREATE AUXILIARY TABLE and is used to hold the data for a column that is defined in a base table.

Base Table

The most common type of table is base table. You can create a base table with the SQL CREATE TABLE Statement.

Clone Table

A table that is structurally identical to a base table is known as clone table.

Empty Table

A table with zero rows is an empty table.

History Table

A history table is used by Database to store historical versions.

Materialized Query Table

Materialized query tables are useful for complex queries that run on large amounts of data.

Result Table

A table that contains a set of rows that a database selects or generates, directly or indirectly, from one or more base tables in response to an SQL Statement is known as result table.

Temporal Table

A temporal table is a table that records the period of time when a row is valid.

XML Table

It is a special table that holds only XML data.

3) Briefly explain about multidimensional OLAP.

In MOLAP, data is stored in a multi-dimensional cube. It fulfills the requirements for an analytic application, where you require to access only summarized level of data. The storage is not in the relational database, but in proprietary formats.

Advantages

- MOLAP cubes are built for fast data retrieval and are thus optimal for slicing operations.
- MOLAP can perform complex calculations quickly.

Disadvantages:-

- MOLAP is limited in the amount of data it can handle because all the calculations are performed when the cube is built.

- Cube technology generally do not already exist in the organization, therefore, to adopt MOLAP technology, chances are additional investments in the form of human and capital is needed.