

Assignment

SC/2017/10291

O.P.M.D.A.AMARASENA

1)

A data warehouse is a process for collecting and managing data from varied sources to provide meaningful business insights. A Data warehouse is typically used to connect and analyze business data from heterogeneous sources.

2)

Data warehouse consists of more information and a data mart consists of less information.

| Data Warehouse | Data Mart |
|---|---|
| Data warehouse is a Centralised system. | decentralised system. |
| In data warehouse, lightly denormalization takes place | While in Data mart, highly denormalization takes place. |
| Data warehouse is top-down model | While it is a bottom-up model. |
| To build a warehouse is difficult | While to build a mart is easy |
| In the data warehouse, a Fact constellation schema is used. | While in this, Star schema and snowflake schema are used. |
| Data Warehouse is flexible | While it is not flexible. |
| Data Warehouse is data-oriented in nature. | While it is the project-oriented in nature |
| Data Warehouse has long life | While data-mart has short life than warehouse |

| | |
|---|-------------------------------------|
| In the Data Warehouse, Data is contained in detail. | summarized form. |
| The Data Warehouse is vast in size. | Data mart is smaller than warehouse |

3)

Star and snowflake schemas are similar at heart: a central fact table surrounded by dimension tables. The difference is in the dimensions themselves. In a star schema each logical dimension is denormalized into one table, while in a snowflake, at least some of the dimensions are normalized.

Star schema advantages

Because a star schema database has a small number of tables and clear join paths, queries run faster than they do against an OLTP system. Small single-table queries, usually of dimension tables, are almost instantaneous.

Star schema disadvantages.

The main disadvantage of the star schema is that it's not as flexible in terms of analytical needs as a normalized data model. Normalized models allow any kind of analytical query to be executed, so long as it follows the business logic defined in the model.

Snowflake schema advantages

Better data quality (data is more structured, so data integrity problems are reduced) Less disk space is used than in a denormalized model.

Snowflake schema disadvantages

The primary disadvantage of the snowflake schema is that the additional levels of attribute normalization adds complexity to source query joins, when compared to the star schema.

Star schemas will only join the fact table with the dimension tables, leading to simpler, faster SQL queries. Snowflake schemas have no redundant data, so they're easier to maintain. Snowflake schemas are good for data warehouses, star schemas are better for data marts with simple relationships.