Laboratory 9

Title of the Laboratory Exercise: Multidimensional data modelling

1. Introduction and Purpose of Experiment

   The multidimensional data model is an integral part of On-Line Analytical Processing (OLAP).Dimensional modelling always uses the concepts of facts, measures, and dimensions. Facts are typically (but not always) numeric values that can be aggregated. Dimensions are groups of hierarchies and descriptors that define the facts. By doing this lab, students will be able to design and implement multi-dimensional data model.

2. Aim and Objectives

   Aim

   - To design and implement concepts of multi-dimensional data modelling

   Objectives

   At the end of this lab, the student will be able to

   - Design star and snowflake schema
   - Implement multi-dimensional data modelling

3. Experimental Procedure

   i. Analyse the problem statement
   ii. Design a data cube which contains a fact table and dimension table
   iii. Document the Results
   iv. Analyse and discuss the outcomes of your experiment

4. Questions

   a. Design the following multi-dimensional data models for the Banking enterprise

      i. Star schema
      ii. Snowflake schema

5. **Calculations/Computations/Algorithms**

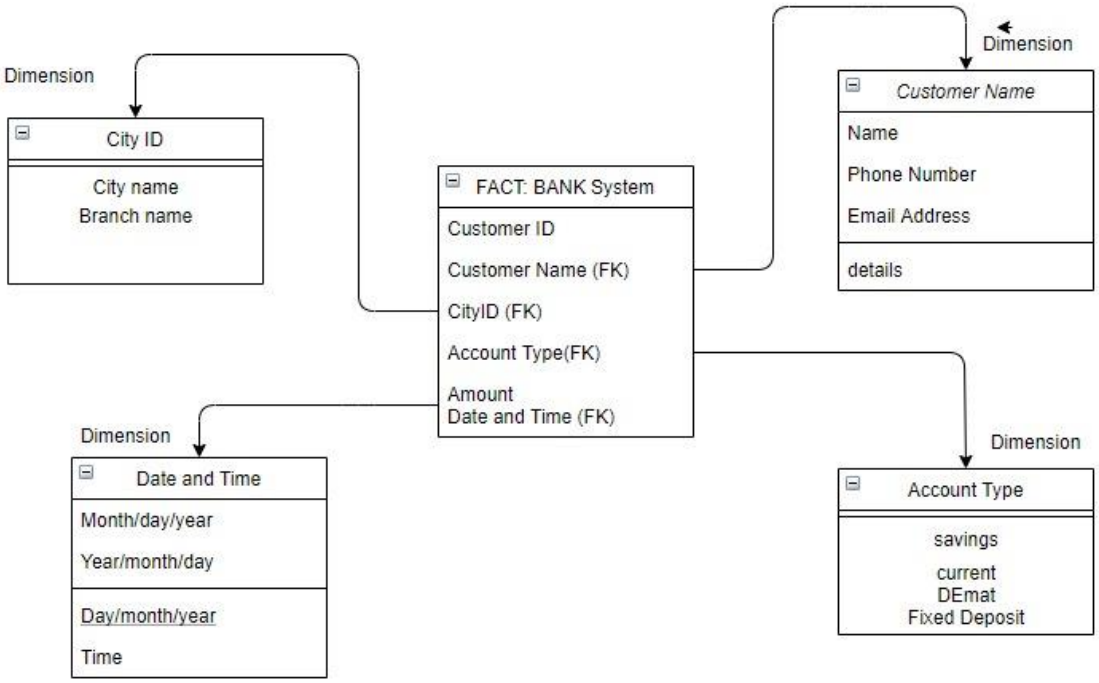

Figure 1 Star Schema

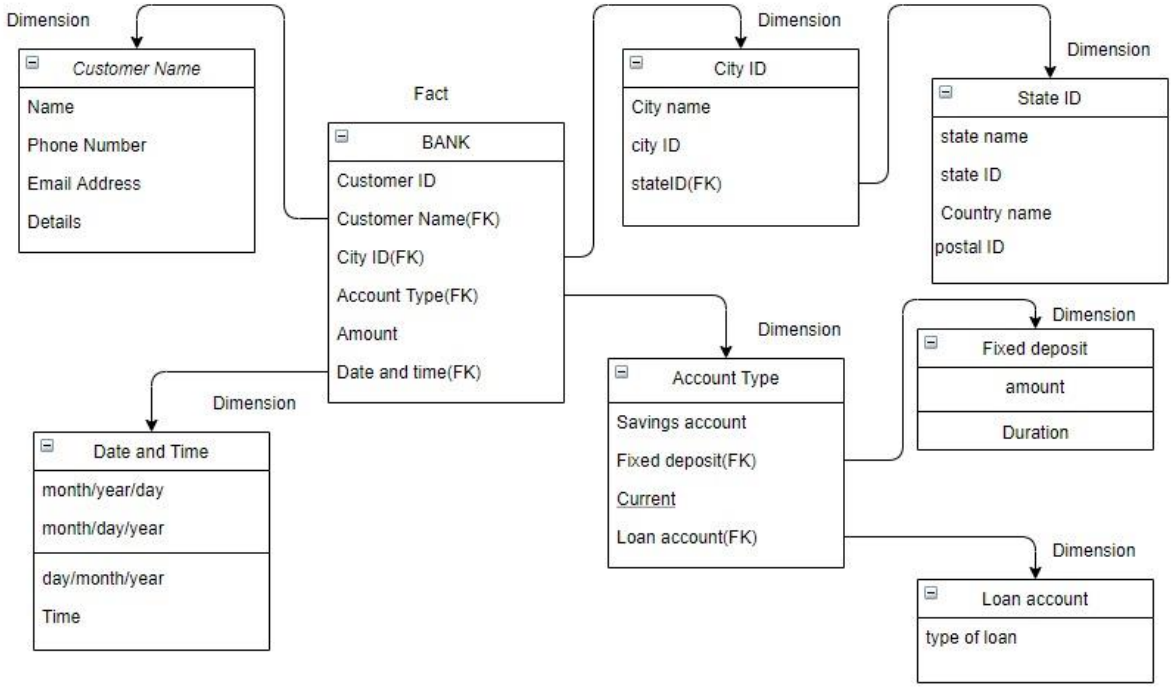

Figure 2 Snowflake schema

Multidimensional data models were implemented for data warehouses and data marts. Star and snow flake models are among the different multidimensional data models. Star schema is the fundamental schema among the data mart schema and it is simplest. This schema is widely used to develop or build a data warehouse and dimensional data marts. It includes one or more fact tables indexing any number of dimensional tables. The star schema is a necessary

case of the snowflake schema. It is also efficient for handling basic queries. The snowflake schema is a variant of the star schema. Here, the centralized fact table is connected to multiple dimensions.

In the snowflake schema, dimensions are present in a normalized from in multiple related tables. The snowflake effect affects only the dimension tables and does not affect the fact tables. For banking enterprise, transaction is taken as the fact table and all the other attributes involved with transaction like customers, bank branches, account details, etc are taken as dimension tables.

Advantages of Star Schema:

i. Simpler Queries:
   Join logic of star schema is quite cinch in compare to other join logic which are needed to fetch data from a transactional schema that is highly normalized.
ii. Simplified Business Reporting Logic:
   In compared to a transactional schema that is highly normalized, the star schema makes simpler common business reporting logic, such as as-of reporting and period-over-period.
iii. Feeding Cubes:
   Star schema is widely used by all OLAP systems to design OLAP cubes efficiently. In fact, major OLAP systems deliver a ROLAP mode of operation which can use a star schema as a source without designing a cube structure.

Advantages of Snowflake Schema:

i. It provides structured data which reduces the problem of data integrity.
ii. It uses small disk space because data are highly structured.

6. Comments

1. Limitations of Experiments: Star model is not flexible in terms if analytical needs as a normalized data model. Snowflaking reduces space consumed by dimension tables, but compared with the entire data warehouse the saving is usually insignificant.

2. Limitations of Results: Star schemas don't reinforce many-to-many relationships within business entities – at least not frequently. Multiple hierarchies can belong to the same dimension has been designed at the lowest possible detail in snow flake model.

3. Learning happened: Learnt multidimensional data models and its uses.