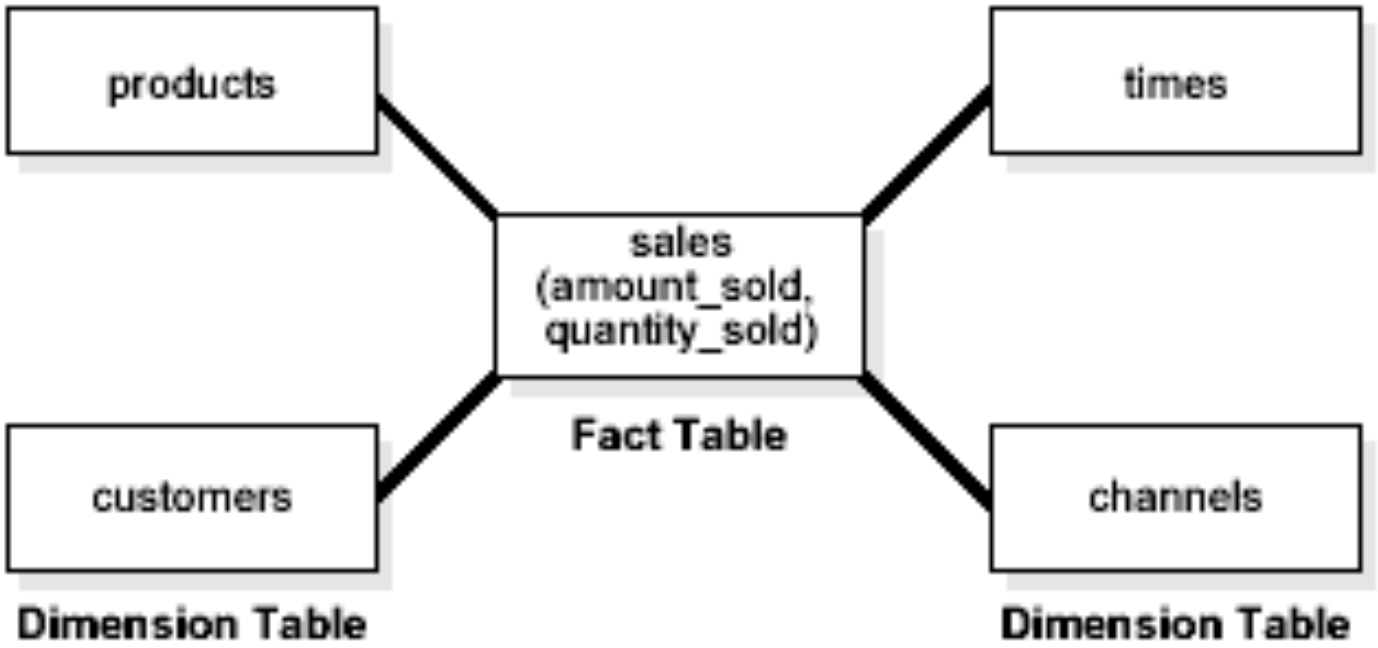
Star Schema

Star Schemas

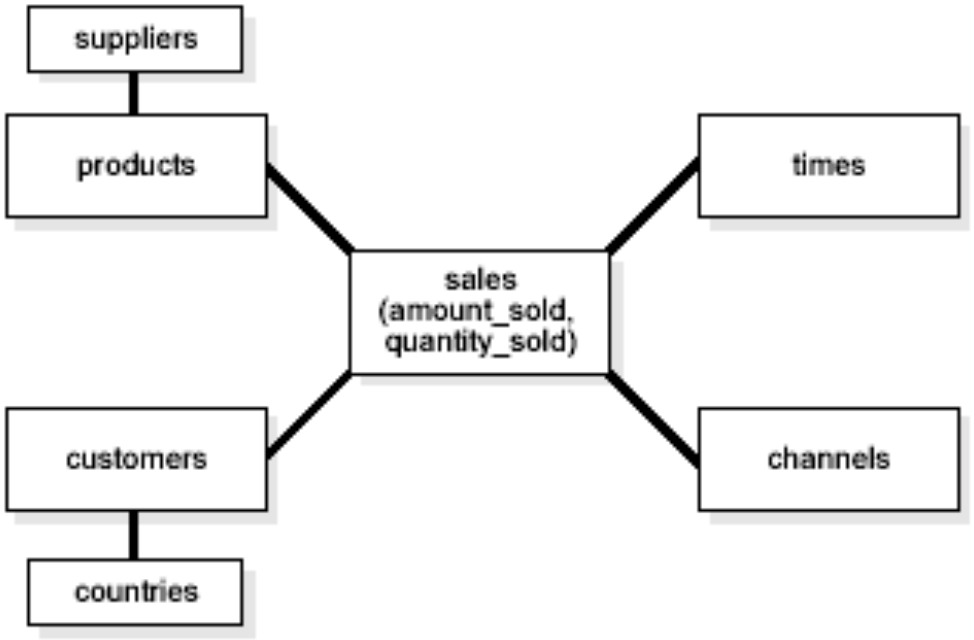
- The **star schema** is the simplest data warehouse schema. It is called a star schema because the diagram resembles a star, with points radiating from a center.
- The center of the star consists of one or more fact tables and the points of the star are the dimension tables, as shown below:



Star Schemas

- A star schema is characterized by one or more very large **fact tables** that contain the primary information in the data warehouse and a number of much smaller **dimension tables** (or lookup tables), each of which contains information about the entries for a particular attribute in the fact table.
- In fact tables primary keys are formed from the concatenation of all the columns that are foreign keys referencing related dimension tables.
- A **star query** is a join between a fact table and a number of dimension tables.
- Each dimension table is joined to the fact table using a primary key to foreign key join, but the dimension tables are not joined to each other.
- A **star join** is a primary key to foreign key join of the dimension tables to a fact table.
- The main advantages of star schemas are that they:
  - Provide a direct and intuitive mapping between the business entities being analyzed by end users and the schema design.
  - Provide highly optimized performance for typical data warehouse queries.

Snowflake Schemas



- The snowflake schema is a more complex data warehouse model than a star schema, and is a type of star schema.
- It is called a snowflake schema because the diagram of the schema resembles a snowflake.
- Snowflake schemas normalize dimensions to eliminate redundancy. That is, the dimension data has been grouped into multiple tables instead of one large table.