Schema With Example

Dimensional Data Modeling

- **E-R model**
  - **Symmetric**
  - **Divides data into many entities**
  - **Describes entities and relationships**
  - **Seeks to eliminate data redundancy**
  - **Good for high transaction performance**
- **Dimensional model**
  - **Asymmetric**
  - **Divides data into dimensions and facts**
  - **Describes dimensions and measures**
  - **Encourages data redundancy**
  - **Good for high query performance**

Facts/Dimensions

- **Fact**
  - **Central, dominant table**
  - **Multi-part primary key**
  - **Holds millions & billions of records**
  - **Links directly to dimensions**
  - **Stores business measures**
  - **Constantly varying data**

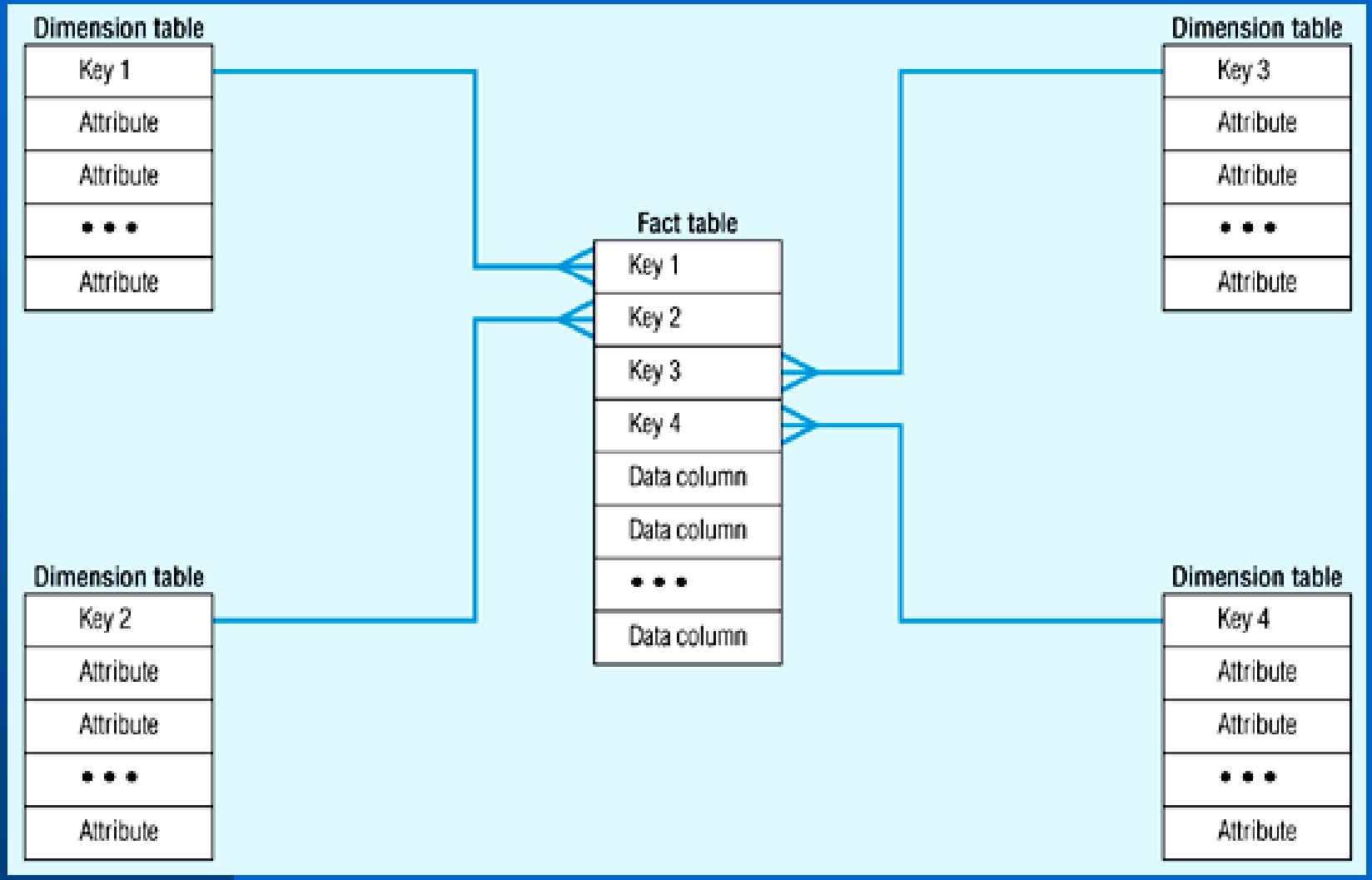
Facts/Dimensions (contd.)

- **Dimensions**
  - **Single join to the fact table (single primary key)**
  - **Stores business attributes**
  - **Attributes are textual in nature**
  - **Organized into hierarchies**
  - **More or less constant data**
  - **E.g. Time, Product, Customer, Store, etc.**

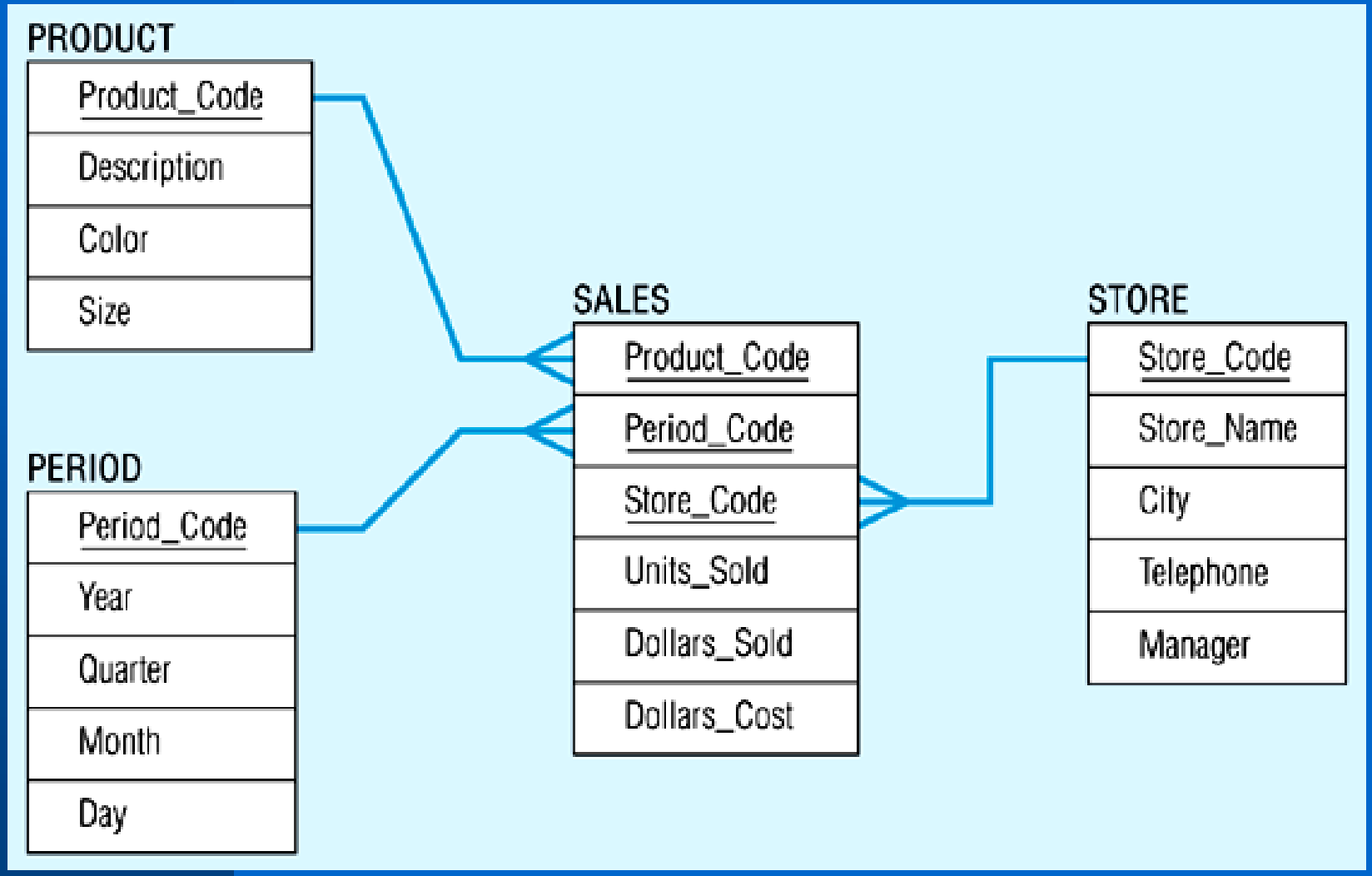
Star/Snowflake schema

- **Star schema**
  - **Fact surrounded by 4-15 dimensions**
  - **Dimensions are de-normalized**
- **Snowflake schema**
  - **Star schema with secondary dimensions**
  - **Don’t snowflake for saving space**
  - **Snowflake if secondary dimensions have many attributes**

Star schema


Dimension table
Key 1
Attribute
Attribute
• • •
Attribute
Dimension table
Key 3
Attribute
Attribute
• • •
Attribute
Fact table
Key 1
Key 2
Key 3
Key 4
Data column
Data column
• • •
Data column
Dimension table
Key 2
Attribute
Attribute
• • •
Attribute
Dimension table
Key 4
Attribute
Attribute
• • •
Attribute

Star schema example

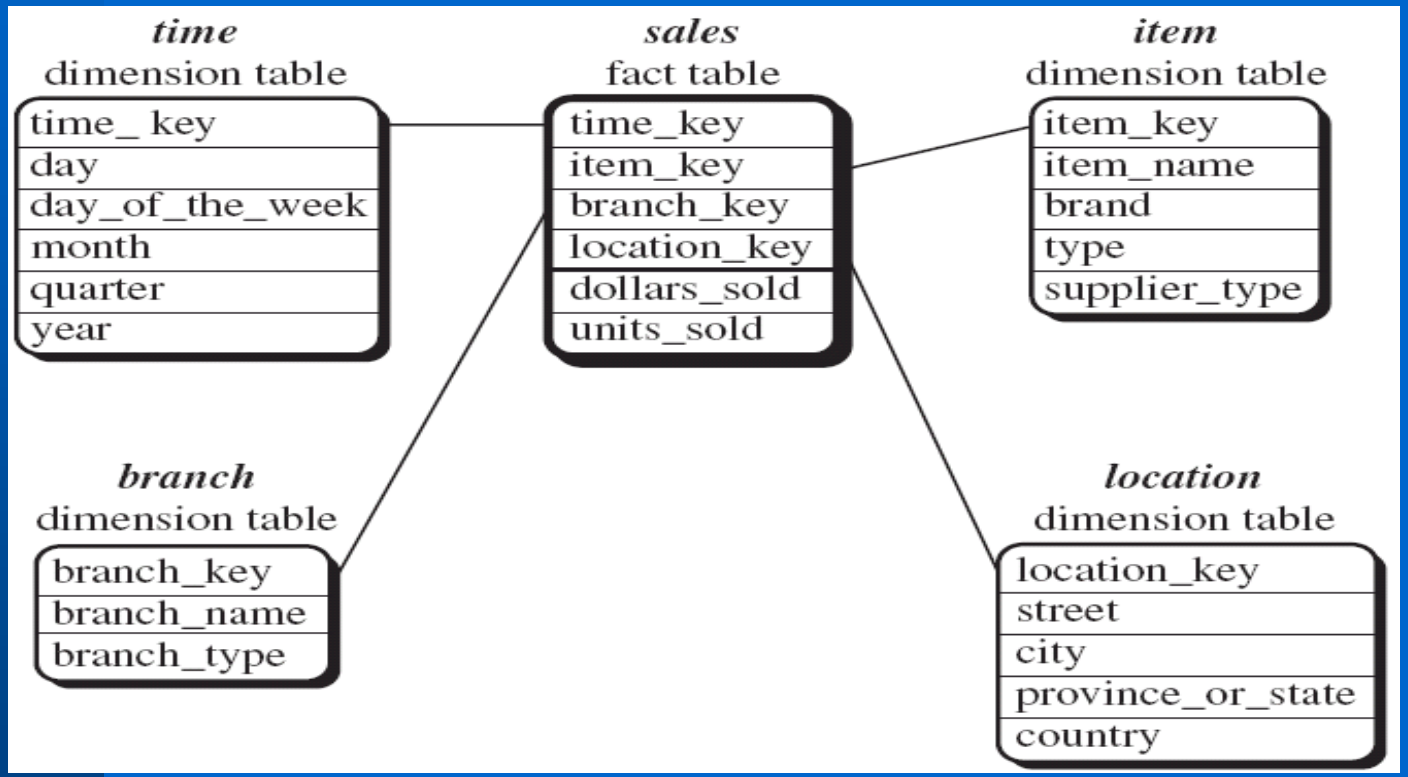


Conceptual Modeling of Data Warehouses

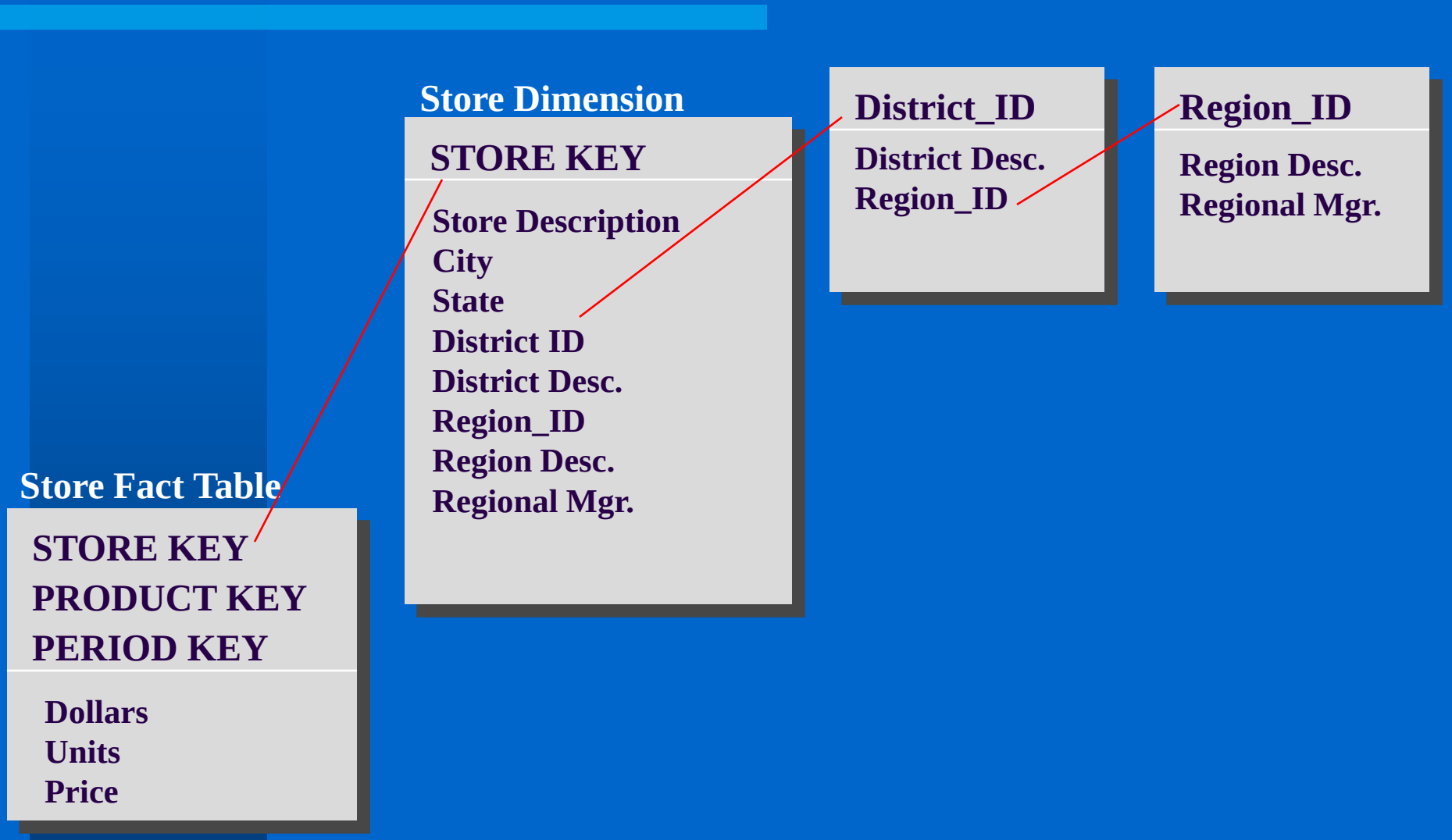
- **Star schema: A fact table in the middle connected to a set of dimension tables**

- **It contains:**
  - **A large central table (fact table)**
  - **A set of smaller attendant tables (dimension table), one for each dimension**

Star schema


time
dimension table
time_ key
day
day_of_the_week
month
quarter
year
sales
fact table
time_key
item_key
branch_key
location_key
dollars_sold
units_sold
item
dimension table
item_key
item_name
brand
type
supplier_type
branch
dimension table
branch_key
branch_name
branch_type
location
dimension table
location_key
street
city
province_or_state
country

Snowflake schema example

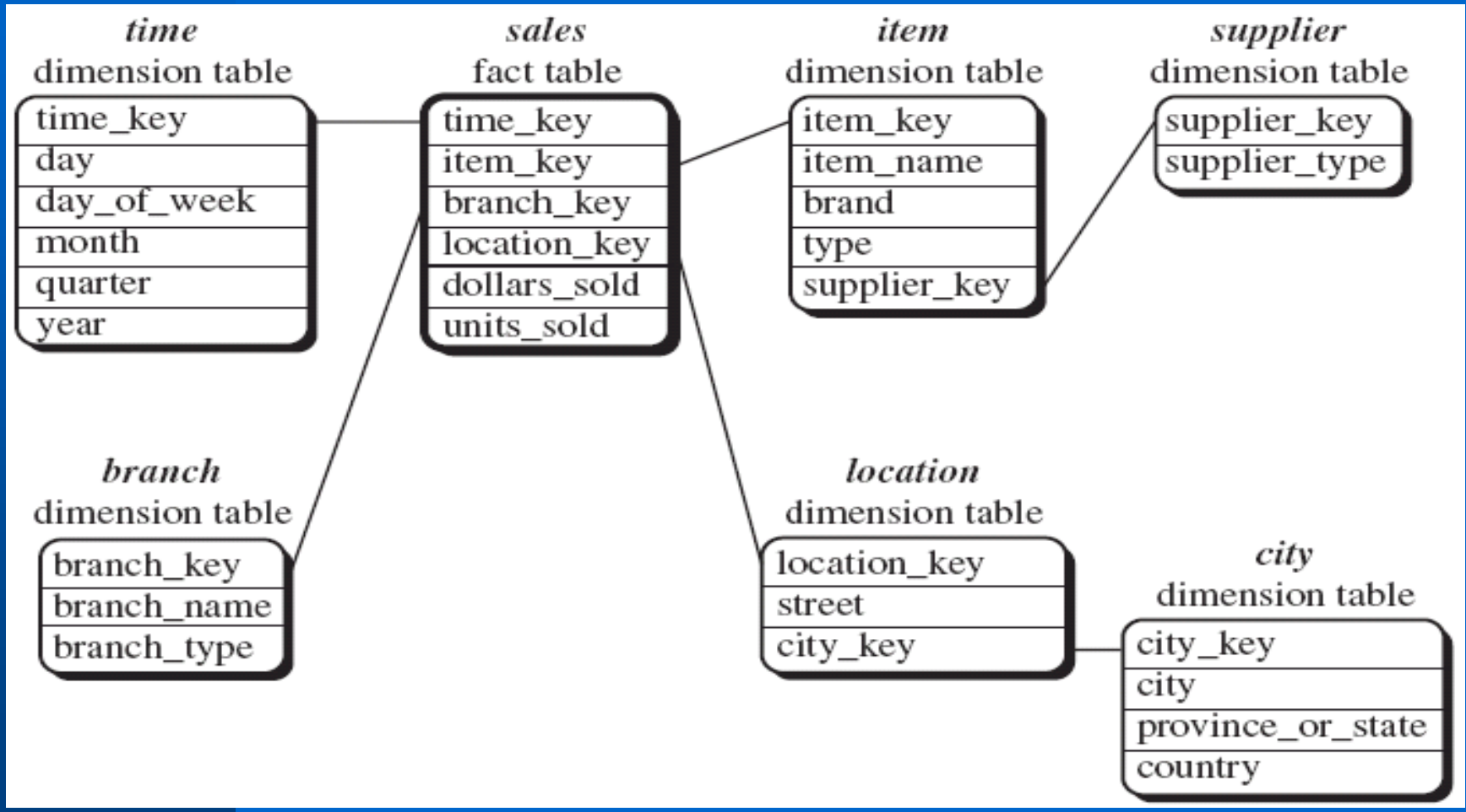


Conceptual Modeling of Data Warehouses

- **Snowflake schema: A refinement of star schema where some dimensional hierarchy is further splitting (normalized) into a set of smaller dimension tables, forming a shape similar to snowflake**

- **However, the snowflake structure can reduce the effectiveness of browsing, since more joins will be needed**

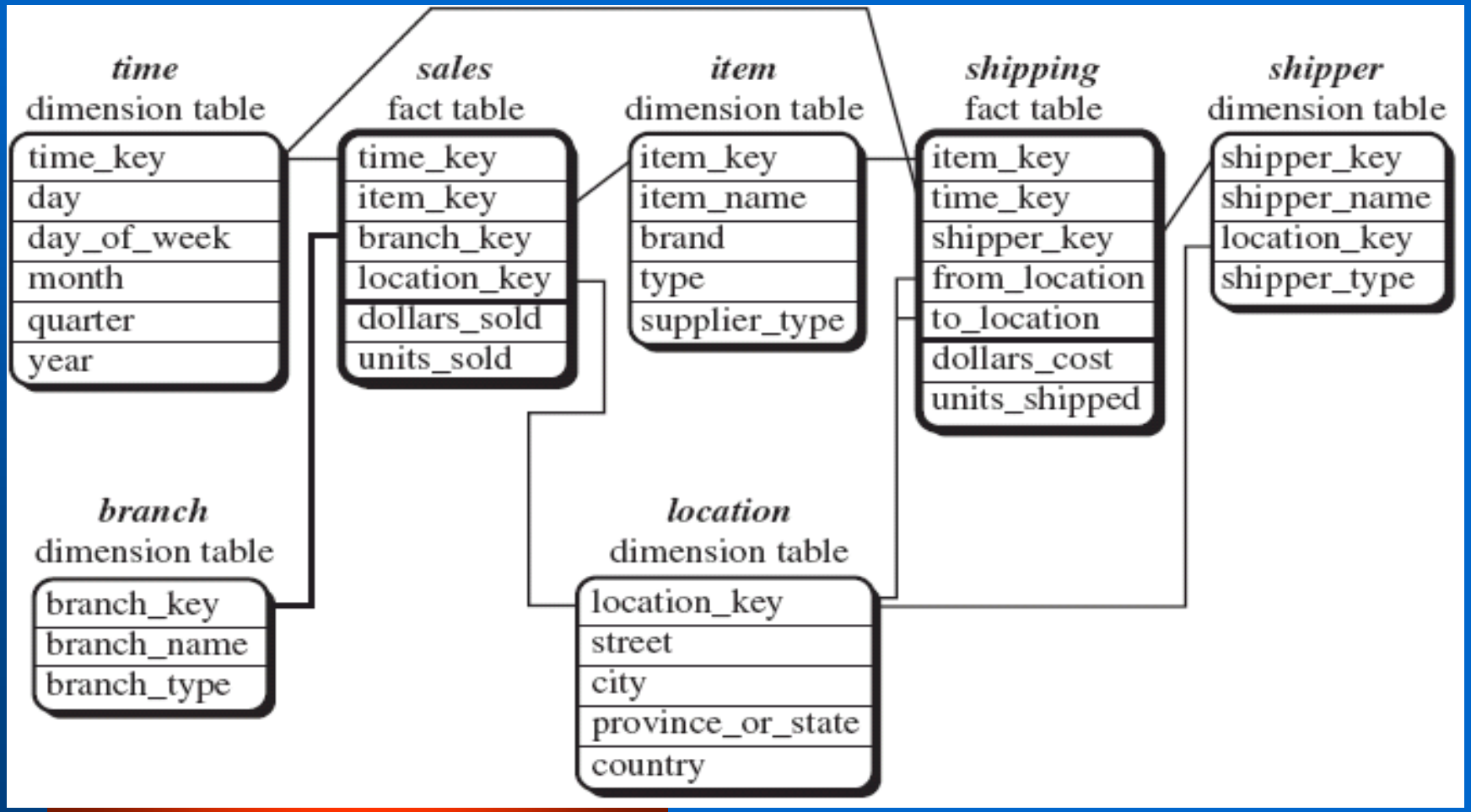
Snowflake schema



Conceptual Modeling of Data Warehouses

- **Fact constellations: Multiple fact tables share dimension tables, viewed as a collection of stars, therefore called galaxy schema or fact constellation**

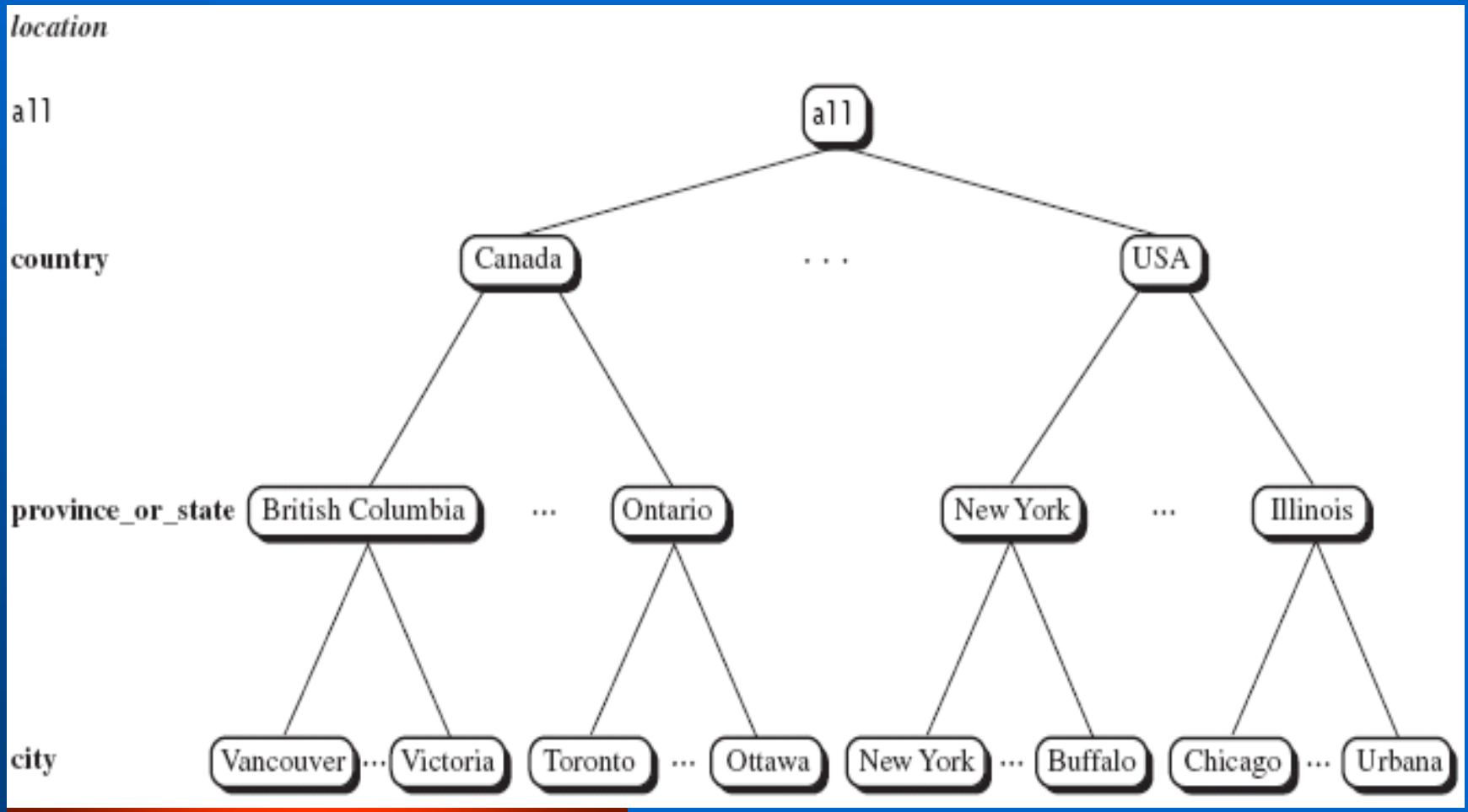
Fact constellations



Concept Hierarchies

- **A Concept Hierarchy defines a sequence of mappings from a set of low-level concepts to high-level**
- **Consider a concept hierarchy for the dimension “Location”**

Concept Hierarchies



Concept Hierarchies

- **Concept hierarchies may also be defined by grouping values for a given dimension or attribute, resulting in a set-grouping hierarchy**