

Concept Hierarchies

“A concept hierarchy defines a sequence of mappings from a set of low-level concepts to higher-level, more general concepts”

Ex: Location

It is defined with Country

It is defined with state, and country

It is defined with city, state, and country

It is defined with street, city, state, and country

Concept Hierarchies for Location

street < city < province or state < country

Concept Hierarchies

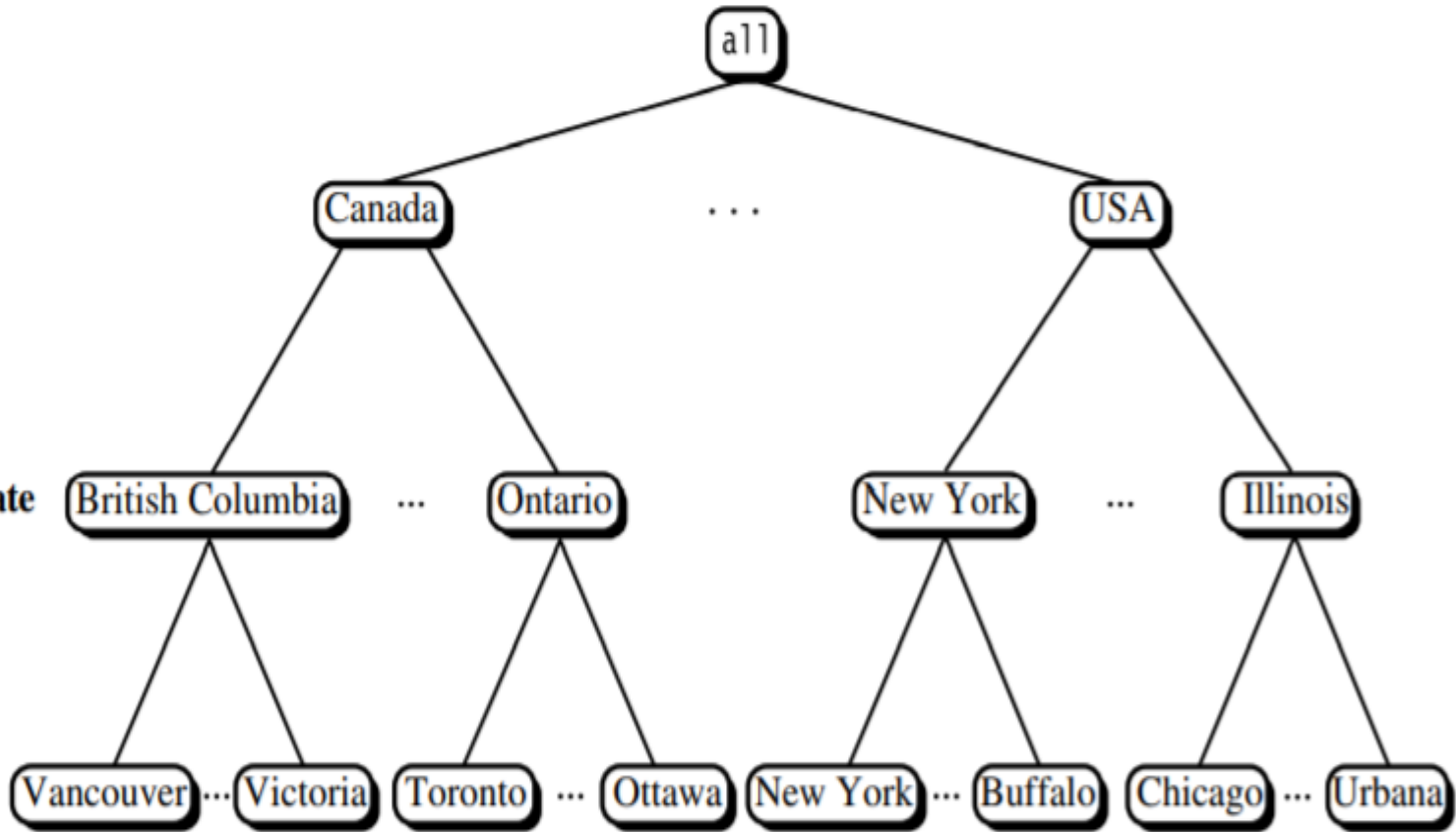
location

all

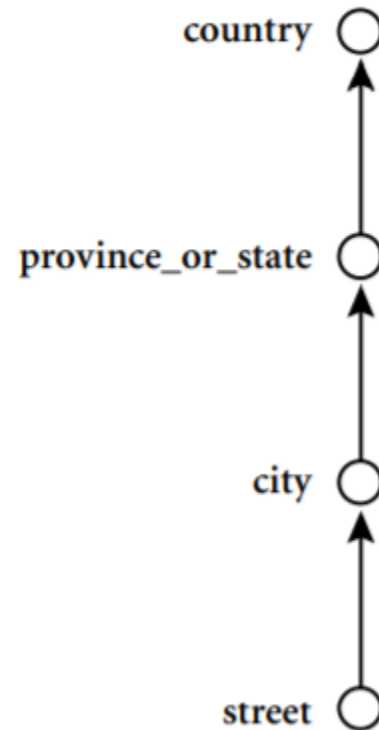
country

province_or_state

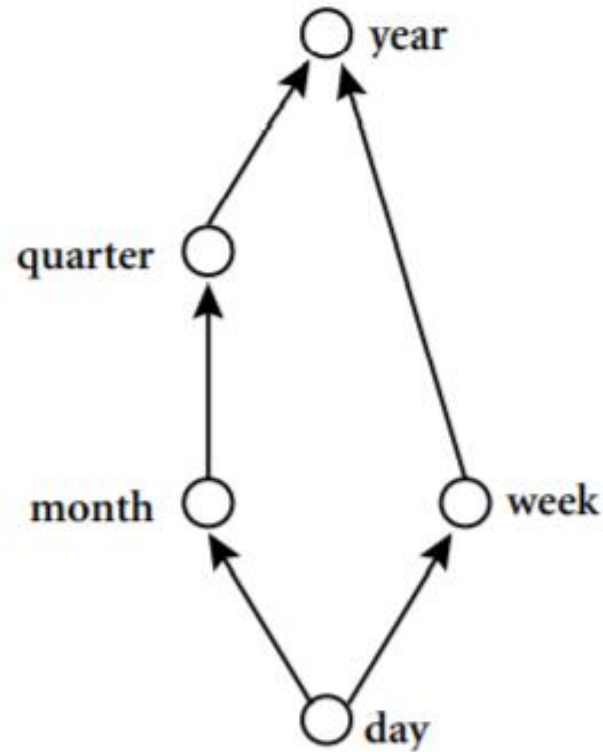
city



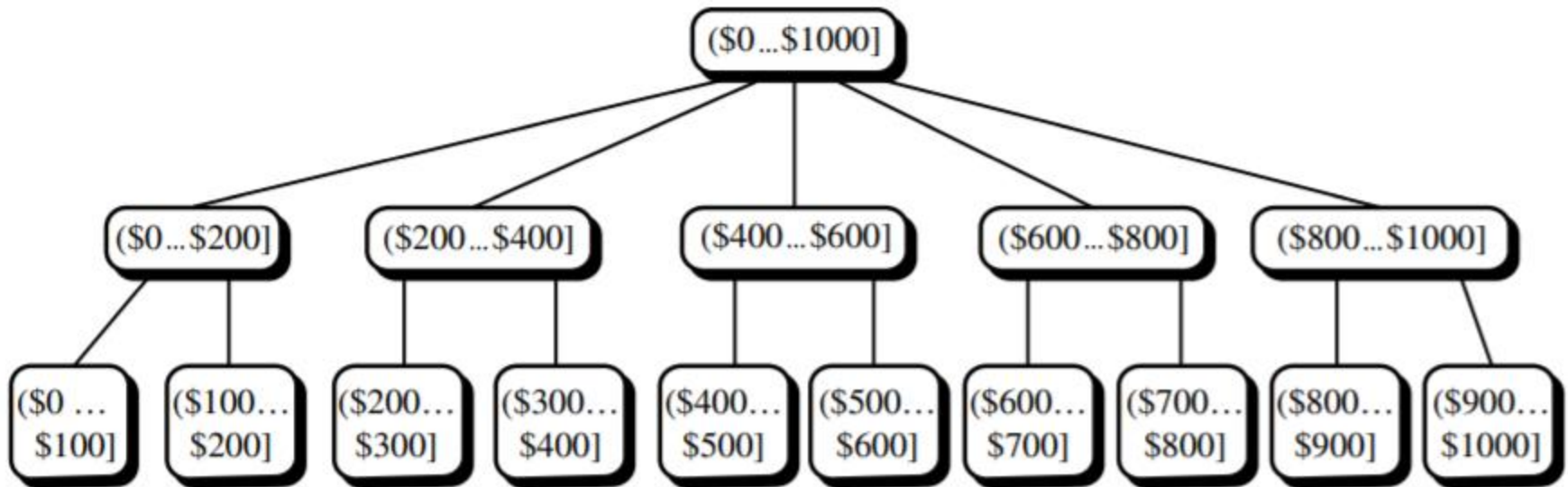
Concept Hierarchies



Concept Hierarchies



Concept Hierarchies



OLAP Operations in the Multidimensional Data Model



How are concept hierarchies useful in OLAP?

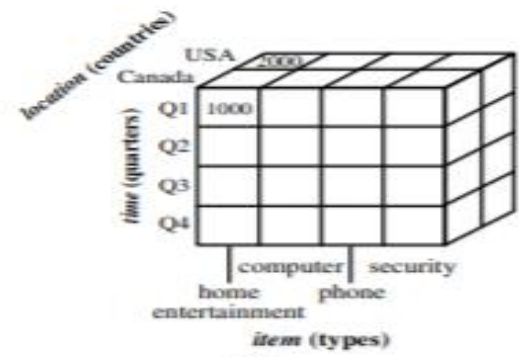
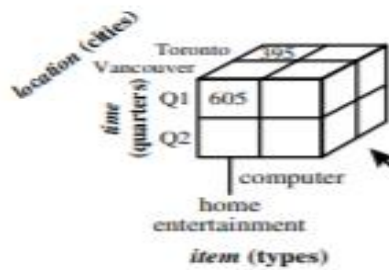
“” In the multidimensional model, data are organized into multiple dimensions, and each dimension contains multiple levels of abstraction defined by concept hierarchies”

OLAP Operations in the Multidimensional Data Model



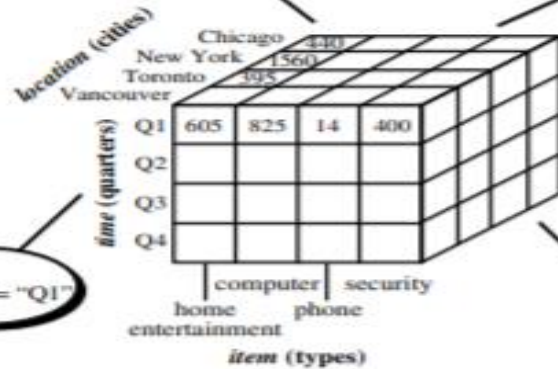
OLAP Operations

1. Roll-Up
2. Drill-down
3. Slice and Dice
4. Pivot



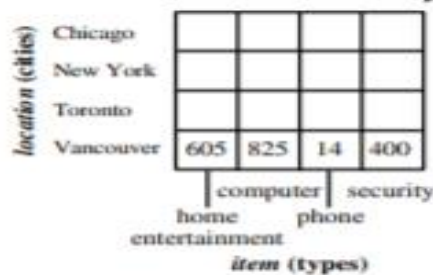
dice for
 (location = "Toronto" or "Vancouver")
 and (time = "Q1" or "Q2") and
 (item = "home entertainment" or "computer")

roll-up
 on location
 (from cities to countries)

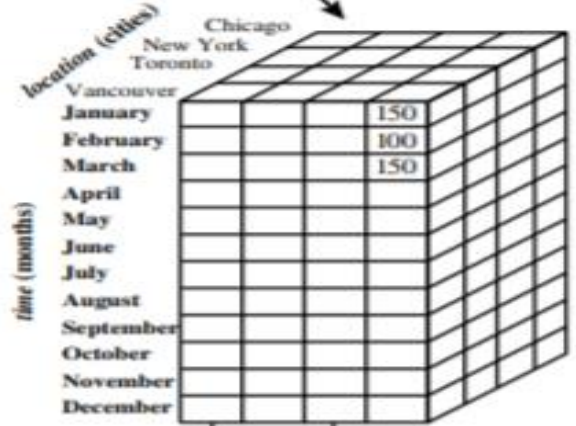
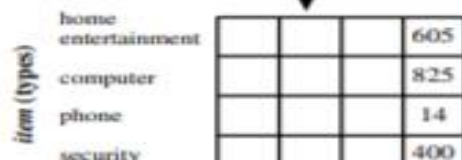


slice
 for time = "Q1"

drill-down
 on time
 (from quarters to months)



pivot



Questions

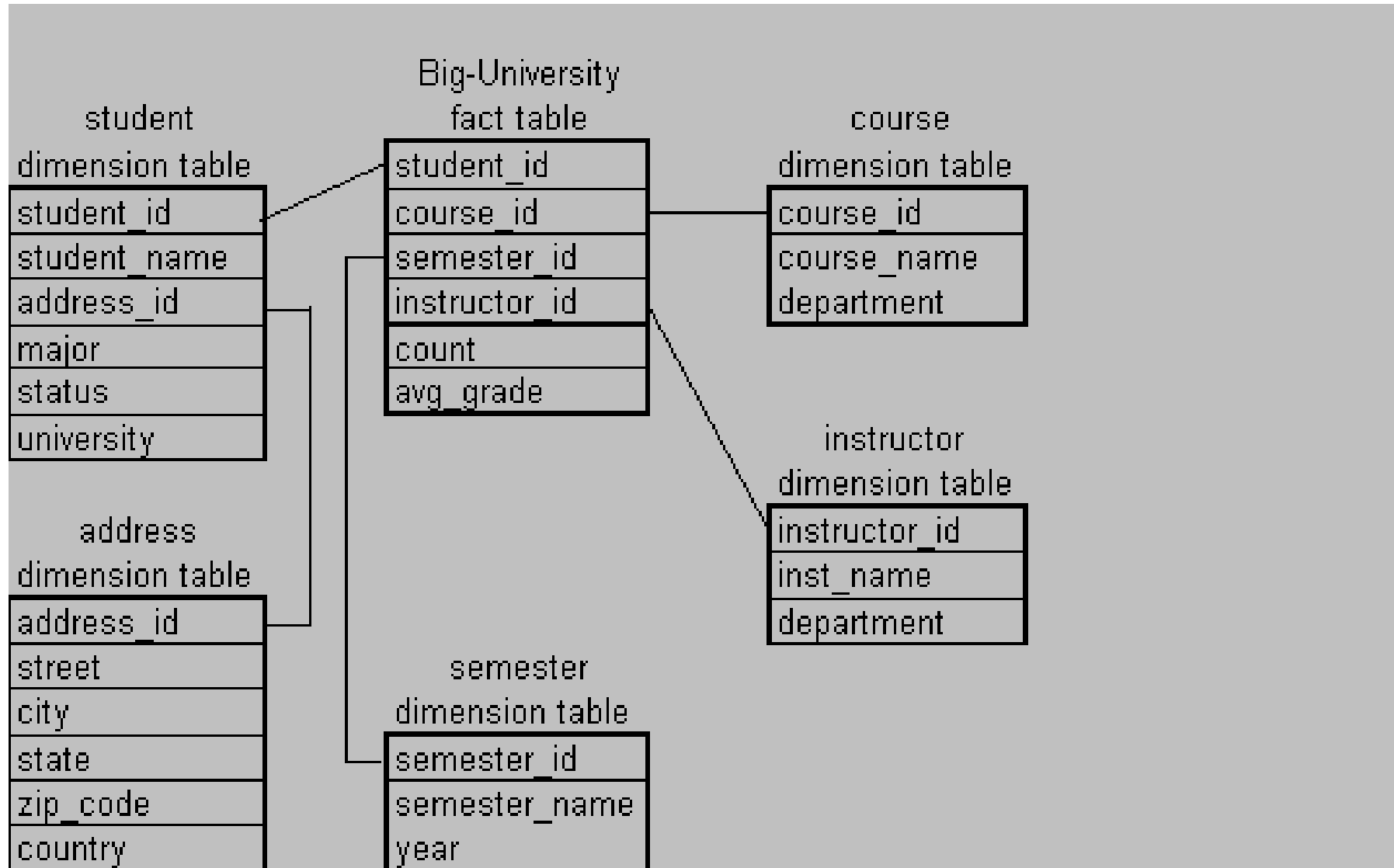


Suppose that a data warehouse for Big University consists of the following four dimensions: student, course, semester, and instructor, and two measures count and avg grade. When at the lowest conceptual level (e.g., for a given student, course, semester, and instructor combination), the avg grade measure stores the actual course grade of the student. At higher conceptual levels, avg grade stores the average grade for the given combination

Draw a snowflake schema diagram for the data warehouse.

Starting with the base cuboid [student, course, semester, instructor], what specific OLAP operations (e.g., roll-up from semester to year) should one perform in order to list the average grade of CS courses for each Big University student.

Answers



- **Roll-up on course from course_id to department.**
- **Roll-up on student from student_id to university.**
- **Dice on course, student with department = "CS" and university = "big-university"**
- **Drill-down on student from university to student_name.**

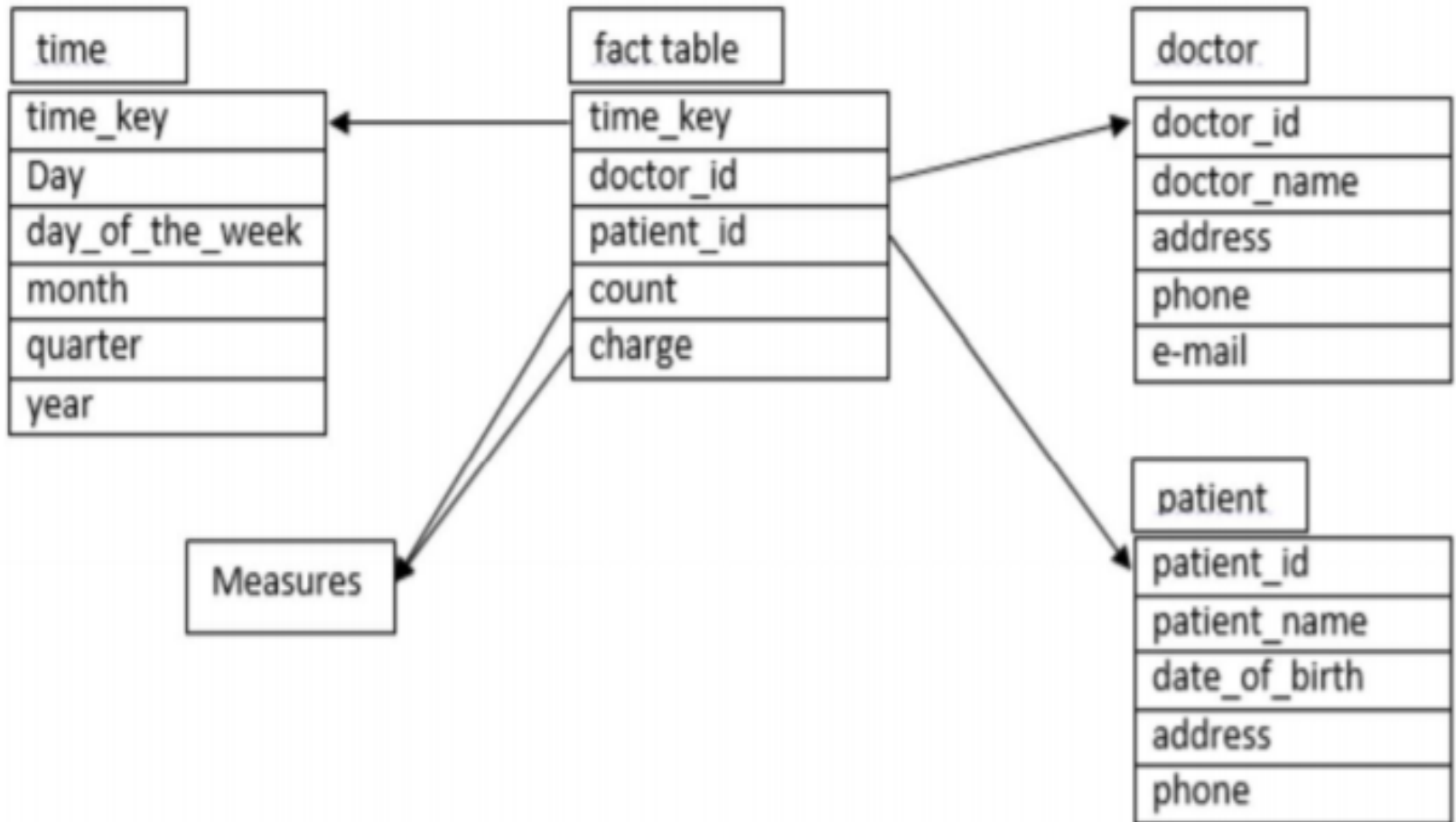
Question



Suppose that a data warehouse consists of the three dimensions time, doctor, and patient, and the two measures count and charge, where charge is the fee that a doctor charges a patient for a visit. a.

Draw a schema diagram for the above data warehouse using one of the schemas. [star, snowflake]

Answer



Data Warehouse Architecture (Three- Tier)



- Top-Down Approach
- Bottom-Up Approach
- Or Combination of Both

- Data Warehouse Design Process:
 - Choose a business process to model
 - Choose the grain of the business process
 - Choose the dimensions
 - Choose the measures

Data Warehouse Architecture (Three-Tier)

