National University of Computer and Emerging Sciences

Lahore Campus

Data Warehousing and Business Intelligence (DS3003)

Sessional-I Exam

Total Time (Hrs.): 1

Total Marks: 25

Total Questions: 3

Date: September 23rd 2024

Course Instructor(s)

M. Ishaq Raza

_____ _____ _____

Roll No Section Student Signature

Solution

Do not write below this line.

Note: Please ensure that you attempt all questions and their respective parts in the given order.

Consider the following case study for the next two questions:

Bill Date Dim: Bill Date, Bill Day Desc, Bill Month ID, Bill Month Desc, Bill Year ID, Bill Year Desc

Customer Dim: Customer Code, Customer Desc, City ID, City Desc, Country ID, Country Desc

Sales Rep Dim: Sales Rep No, Sales Rep Desc, Channel ID, Channel Desc

Rate Plan Dim: Rate Plan Id, Rate Plan Desc, Rate Plan Type Code, Rate Plan Type Desc

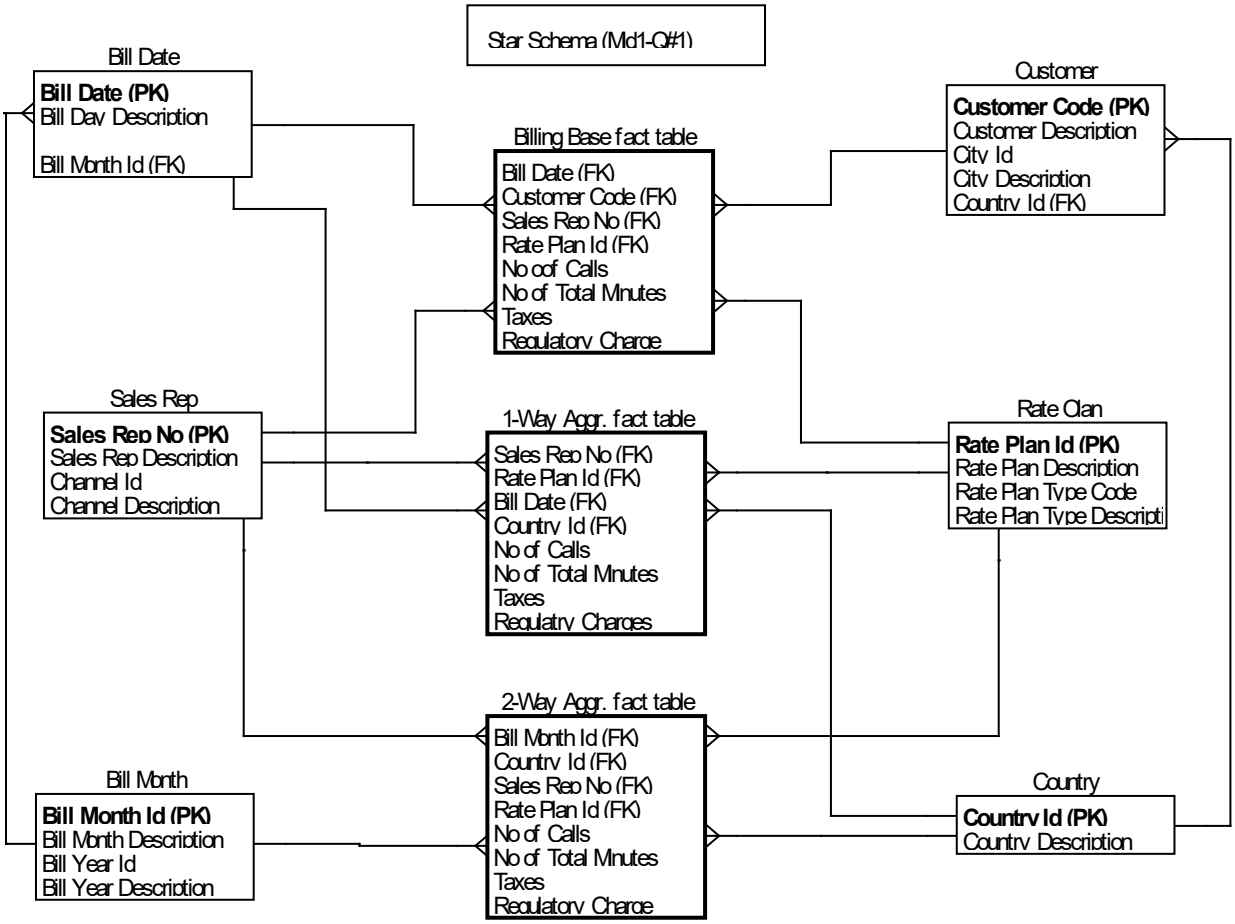
Billing Fact: Bill Date, Customer Code, Sales Rep No, Rate Plan Id, No of Calls, No of Total Minutes, Taxes, Regulatory Charge

Assume: 10,000 customers, 200 cities, 5 countries, 40 sales rep, 4 channels, 30 rate plans, 3 rate plan types, and 3 years billing history.

CLO # 2: Demonstrate an understanding of the fundamental concepts of the Star and the Snowflake Schema; learn how to design the schema of a DW based on these two models.

Q. No 1: Draw the appropriate star schema that includes a base fact table, a one-way aggregate fact table (along Customer Country), and a two-way aggregate fact table (along Bill Month and Customer Country). Show the primary keys, foreign keys and all the relationships between the dimensions and fact tables. Note: Draw only one diagram that includes base fact table as well as aggregate fact tables. [10]

Ans:



CLO # 2: Demonstrate an understanding of the fundamental concepts of the Star and the Snowflake Schema; learn how to design the schema of a DW based on these two models.

Q. No 2: Estimate the size (in number of rows) of the above customer dimension table, sales rep dimension table, billing base fact table, and both the aggregate fact tables. [5]

Ans:

Customer Dimension: 10000+5= 10005 rows

Sales Rep Dimension: 40 rows

Base Fact Table: (3x365) x 10000 x 40 x 30 = 13,140,000,000 rows

Aggregate Fact Table1 (along country:

3x365 (day) x 5 (country) x 40 (sales rep) x 30 (rate plan) = 6,570,000 rows

Aggregate Fact Table2 (along month & country):

3x12 (month) x 5 (country) x 40 (sales rep) x 30 (rate plan) = 216,000 rows

CLO # 2: Demonstrate an understanding of the fundamental concepts of the Star and the Snowflake Schema; learn how to design the schema of a DW based on these two models.

Q. No 3: Briefly answer the following questions. [10]

a. Pick any one architecture for building a data warehouse and list the advantages and disadvantages of that architecture.

b. What are the different types of OLAP? Which type of OLAP can handle large amounts of data? Justify your answer.

c. How does a snowflake schema differ from a star schema? Name two advantages of the snowflake schema.

d. Differentiate between pre-join denormalization and column-replication denormalization techniques. Explain with an example.

e. When would you use partitioned cubes in multidimensional online analytical processing (MOLAP)?

Ans: See Textbook/Notes.