**Part 1) Normalization**

1. For each fact variable in your fact table, what type of fact is it? Additive, semi-additive, or non-additive?

The ‘copay’, ‘insurancepaid’ facts are Additive facts that can be summed up through all dimensions be it the ‘member\_id’, ‘fill\_date’, and ‘drug\_ndc’.

1. In your fact table, describe the *grain* in one sentence. What does each fact row represent?

The granularity of this dimensional model containing, drug\_details, member\_id is insurancepaid and copay to get the drug by day.

**Part 2) Primary and Foreign Key Setup in MySQL**

1. What are the primary keys you designated for each of your tables? For each PK, is it a natural key or a surrogate key?

|  |  |  |
| --- | --- | --- |
| **TABLE NAME** | **PRIMARY KEY** | **KEY TYPE ()** |
| dim\_memberdetails | member\_id | Natural |
| dim\_drugform | drug\_form\_code | Natural |
| dim\_drug\_bg\_code | drug\_brand\_generic\_code | Natural |
| dim\_drugdetails | drug\_ndc | Natural |
| fact\_insurance\_info | No Primary Key defined for the fact table as we are never going to query for single record in the fact table. | |

1. What are the foreign keys you designated for each of your tables? For each FK, which table did you reference where that FK is listed as the PK?

|  |  |  |
| --- | --- | --- |
| **TABLE NAME** | **FOREIGN KEY** | **TABLE NAME WHERE FK is referenced as PK** |
| dim\_drugdetails | drug\_form\_code | dim\_drugform |
| dim\_drugdetails | drug\_brand\_generic\_code | dim\_drug\_bg\_code |
| fact\_insurance\_info | member\_id | dim\_memberdetails |
| fact\_insurance\_info | drug\_ndc | dim\_drugdetails |

1. For each FK, what did you tell MySQL to in case of deletion or update (CASCADE, SET NULL, or RESTRICT)? Why did you select the option that you did for each FK?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **TABLE NAME** | **FOREIGN KEY** | **constraint on FK** | **Reason for the constraint** | **TABLE NAME WHERE FK is referenced as PK** |
| dim\_drugdetails | drug\_form\_code | ON DELETE RESTRICT  ON UPDATE RESTRICT | A patient has used drug in mentioned form, there are different forms of drugs of the same name. This can cause ambiguity if the drug form gets changed or deleted and the patient can receive incorrect drug. Hence, there can be no change in the form. | dim\_drugform |
| dim\_drugdetails | drug\_brand\_generic\_code | ON DELETE RESTRICT ON UPDATE CASCADE | Records for this should not be deleted but if updated in the parent table, the change should be reflected in the child table. | dim\_drug\_bg\_code |
| fact\_insurance\_info | member\_id | ON UPDATE CASCADE ON DELETE RESTRICT | There can be changes to member names and certain updates in the age of the member or updating the incorrect birthdates hence on update cascade. But the record should not be deleted from the parent table. | dim\_memberdetails |
| fact\_insurance\_info | drug\_ndc | ON DELETE RESTRICT | Drugs are already taken by the patients and those records should be maintained for future need. Drug detaild should not be deleted from the parent table. | dim\_drugdetails |

Write a SQL query that identifies the number of prescriptions **grouped by drug name**. Paste your output to this query in the space below here; your code should be included in your .sql file

**A picture containing table

Description automatically generated**

How many prescriptions were filled for the drug Ambien?

Answer: 5

Write a SQL query that counts total prescriptions, counts unique (i.e. *distinct*) members, sums copay $$, and sums insurance paid $$, for members grouped as either ‘age 65+’ or’ < 65’. **Use case statement logic** to develop this query like lecture 3.

Table

Description automatically generated

How many unique members are over 65 years of age?

Calendar

Description automatically generated

How many prescriptions did they fill?

Answer: 6

Write a SQL query that identifies the **amount paid by the insurance** for the **most recent prescription fill date**. Use the format that we learned with SQL Window functions. Your output should be a table with **member\_id, member\_first\_name, member\_last\_name, drug\_name, fill\_date (*most recent*), and most recent insurance paid.**

Table

Description automatically generated with medium confidence

Also answer these questions: **For member ID 10003**, what was the drug name listed on their most recent fill date?

Answer: Ambien

How much did their insurance pay for that medication?

Answer: 322