ALY6030 Final Project Assignment

Yash Patel

Northeastern University

Table of Contents

[ALY6030 Final Project Assignment 3](#_Toc22983999)

[Part 1 3](#_Toc22984000)

[1.1 3](#_Toc22984001)

[1.2 3](#_Toc22984002)

[Part 2 4](#_Toc22984003)

[2.1 4](#_Toc22984004)

[2.2 4](#_Toc22984005)

[2.3 4](#_Toc22984006)

[Part 3 6](#_Toc22984007)

[Part 4 7](#_Toc22984008)

[4.1 7](#_Toc22984009)

[4.1.1 7](#_Toc22984010)

[4.2 7](#_Toc22984011)

[4.2.1 7](#_Toc22984012)

[4.2.2 8](#_Toc22984013)

[4.3 8](#_Toc22984014)

[4.3.1 8](#_Toc22984015)

[4.3.2 8](#_Toc22984016)

ALY6030 Final Project Assignment

# Part 1

Normalization

## 1.1

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

There are 3 fact variables in my fact table, namely – fill\_date, copay and insurancepaid. The fill\_date variable is a non-additive fact type. And the copay and insurancepaid variables are semi-additive fact type.

-2 fill date is a dimension it is non-numeric and can’t be added generally be speaking it represents a point in time.

-2 copay and insurance are fully additive

## 1.2

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

The amount of money paid for a drug when a member filed an insurance claim on a day. correct

# Part 2

Primary and Foreign Key Setup in MySQL

## 2.1

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

The member\_id is the natural primary key for the Member table. The member\_gender\_id is the natural primary key for the Gender table. The drug\_ndc is the natural primary key for the Drug\_Name table. The drug\_form\_code is the natural primary key for the Drug\_Form table. And the drug\_brand\_generic\_code is the natural primary key for the Drug\_Brand table.

## 2.2

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?

For the Member table, I designated the member\_gender\_id as foreign key, I referenced it from the Gender table. For the Drug\_Name table, I designated drug\_form\_code and drug\_brand\_generic\_code as the foreign keys, I referenced it from the Drug\_Form and Drug\_Brand tables respectively. And for the Claim table, I designated member\_id and drug\_ndc as the foreign keys, I referenced it from the Member and Drug\_Name tables respectively.

## 2.3

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?

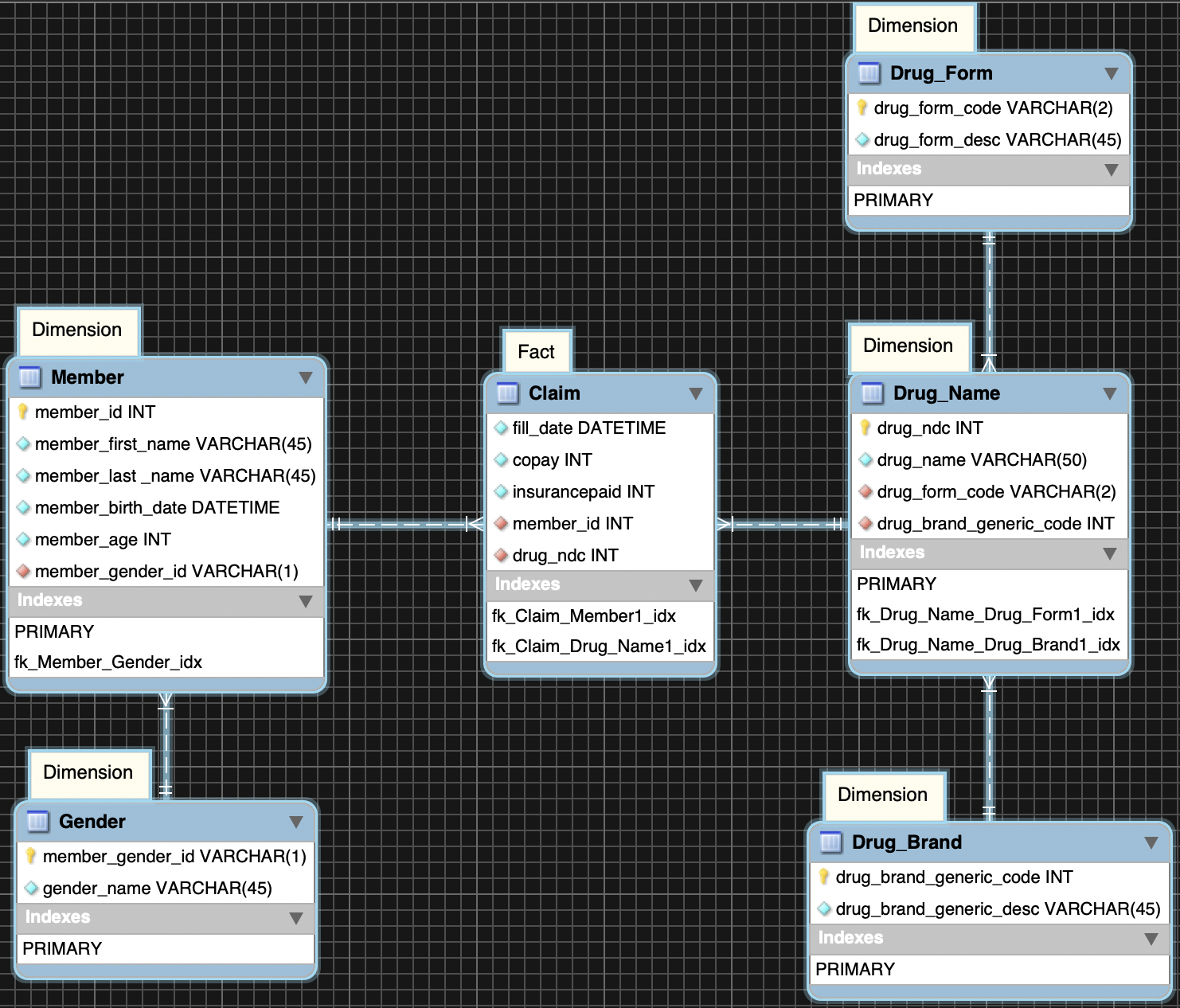
For member\_id foreign key in claim table, I used cascade option on deletion as well as update because I want every data related to the member to reflect the current data or if we delete the member, I want every data related to that member deleted from the database.

For member\_gender\_id foreign key in member table, drug\_form\_code and drug\_brand\_generic\_code foreign keys in drugname table and drug\_ndc foreign key in claim table – I used restrict on deletion as well as update because I wanted every field to reflect some data rather than showing N/A.

correct

# Part 3

Entity Relationship Diagram



-4 PK not identified in the fact table

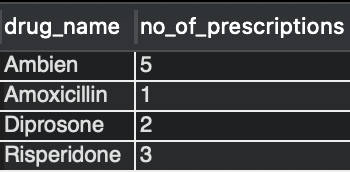
Image 4.01

# Part 4

Analytics and Reporting

## 4.1

Write a SQL query that identifies the number of prescriptions grouped by drug name.



### 4.1.1

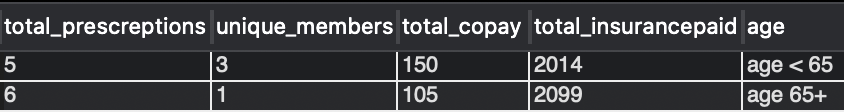
How many prescriptions were filled for the drug Ambien?

There were 5 prescriptions filled for drug Ambien.

correct

## 4.2

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’.



### 4.2.1

How many unique members are over 65 years of age?

There is only 1 unique member who is over 65 years of age. correct

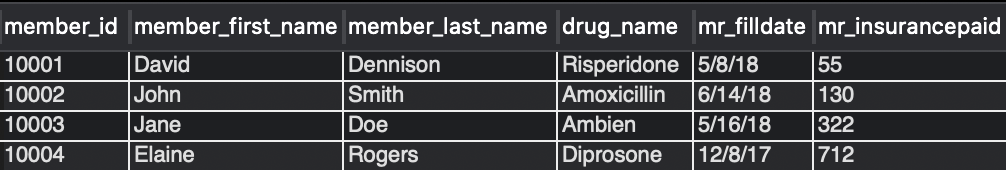
### 4.2.2

How many prescriptions did they fill?

Members over 65 years of age filled 6 prescriptions and members under 65 years of age filled 5 prescriptions.

## 4.3

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.



### 4.3.1

For member ID 10003, what was the drug name listed on their most recent fill date?

Ambien is listed as the drug name on the most recent fill date for member ID 10003.

### 4.3.2

How much did their insurance pay for that medication?

The insurance paid $322 for that medication.

correct

Grade 92