Andrew Smith

Career Foundry Data Immersion

January 13, 2022

3.3: SQL for Data Analysts (Answers 3.3)

**Step 1:**

Query: SELECT category\_id,name FROM category

Output:

Cateogry\_id/Name

1 "Action"

2 "Animation"

3 "Children"

4 "Classics"

5 "Comedy"

6 "Documentary"

7 "Drama"

8 "Family"

9 "Foreign"

10 "Games"

11 "Horror"

12 "Music"

13 "New"

14 "Sci-Fi"

15 "Sports"

16 "Travel"

**Step 2:**

a)

INSERT INTO category(category\_id,name) Values(17,'Thriller')

INSERT INTO category(category\_id,name) Values(18,'Crime')

INSERT INTO category(category\_id,name) Values(19,'Mystery')

INSERT INTO category(category\_id,name) Values(20,'Romance')

INSERT INTO category(category\_id,name) Values(21,'War')

b)

In the create statement, the first constraint applied was “NOT NULL” for the category\_id which makes sure each category\_id value isn’t empty. This is important because if any category\_id value was missing, then the constraint would make you input the proper value so that there are not any empty rows, and the table would function properly. The “NOT NULL” is also applied to the name and last\_update categories to make sure any category name and timestamp is not missing. The last constraint is the “PRIMARY KEY” for the category\_id which gives all values in the category a unique ID. This ensures that the “PRIMARY KEY” values are unique and can be identifiable in a large dataset. In addition, the “PRIMARY KEY” values cannot contain any null or duplicate values which forces the table to be fully functional.

**Step 3:**

a)

SELECT \*

FROM film

WHERE title='African Egg'

-film\_id= 5

SELECT \*

FROM film\_category

WHERE film\_id=5

-category\_id=8

b)

UPDATE film\_category

SET category\_id=17

WHERE film\_id=5

**Step 4:**

DELETE FROM category

WHERE category\_id=19

**Step 5:**

In my opinion, I think some of these steps in Excel would be much more difficult than SQL since finding and updating values in the dataset would take much longer. Commands in SQL allow for faster searching and updates to the data rather than tediously having to change each individual row or column in Excel. On the other hand, simple tasks like in Step 1 would likely be faster in Excel since you could add a filter to the columns and look at the genre list rather than having to type a command. All in all, SQL will be more helpful for analyzing larger datasets.