3.3 SQL for Data Analysts

1. **SELECT**

```sql
SELECT category_id, name FROM category;
```

category_id [PK] integer	name character varying (25)
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

2. **INSERT and CREATE**

```sql
INSERT INTO category (name) VALUES ('Thriller'), ('Crime'),
('Mystery'), ('Romance'), ('War');
```

```sql
-- new table "category" gets created
CREATE TABLE category
(
    -- columns of the table are getting defined
    -- "category_id" has to be an integer, mustn't be zero
    category_id integer NOT NULL DEFAULT
nextval('category_category_id_seq'::regclass),
    -- "name" has to be text, mustn't be zero
    name text COLLATE pg_catalog."default" NOT NULL,
    -- "last_updated" has to be in a timestamp format, mustn't be zero
(automatically generated)
    last_update timestamp with time zone NOT NULL DEFAULT now(),
    -- sets "category_id" as the primary key of this table
    -- "category_pkey" is the name of this constraint
```

```
        CONSTRAINT category_pkey PRIMARY KEY (category_id)
);
```

3. UPDATE

```
SELECT film_id FROM film WHERE title = 'African Egg';
```

film_id [PK] integer
5

(row: 1)

```
UPDATE film_category SET category_id = 17 WHERE film_id = 5;
```

film_id [PK] smallint	category_id [PK] smallint	last_update timestamp without time zone
5	17	2022-07-15 16:37:06.165666

(row: 1)

4. DELETE

```
DELETE FROM category WHERE name = 'Mystery';
```

5. SQL vs Excel

Often, I feel like Excel would be easier – especially when there is not so much data – because I can directly see the data itself. This is exactly why I really like (or almost need) an ERD, as I use it as an overview to understand the architecture behind it. SQL commands feel clearer and more straight forward to utilize than excel functions. For example: When I use the DELETE command of SQL, I know that it works, whereas with Excel, I am often unsure whether what I wanted was really deleted.

6. Find typos:

```
CREATE TABLE employee
(
    employee_id INT NOT NULL AUTO_INCREMENT,
    name VARCHAR(50) NOT NULL,
    contact_number VARCHAR(30),
    designation_id INT,
    last_update TIMESTAMP NOT NULL DEFAULT now(),
    CONSTRAINT employee_pkey PRIMARY KEY (employee_id)
);
```

Corrected some typos and added stuff that I thought would make sense :)