Build a Data Pipeline for Processing and Storing Movies Data

THE SITUATION

Created a data pipeline to process and store sales data from a CSV file. The pipeline will transform and load the data with Python into a PostgreSQL and MySQL database for further analysis using Power BI.
Tools required: Python (Pandas), PostgreSQL, MySQL, DBdiagram.io and Power BI

THE STEPS

Used DBdiagram.io:

- To design a dimensional model for TMDB Database.

Used MySQL and PostgreSQL to:

- Create a database title TMDB.
- Define the database schema under the database (TMDB):
  genre,
  director,
  production_company,
  Movies.

Used Python to:

- Use pandas library to load the movies data from the CSV file into pandas dataframe.
- Use pandas to clean and transform the data as needed. For instance, drop some columns, remove duplicates, change date format, create data subset.
- Use the psycopg2 and mysql library to connect to the PostgreSQL and MySQL database.
- Use pandas, psycopg2, and mysql to load the cleaned and transformed data into the database.
- Runed some test queries on the data in the database to ensure the data pipeline is working correctly.

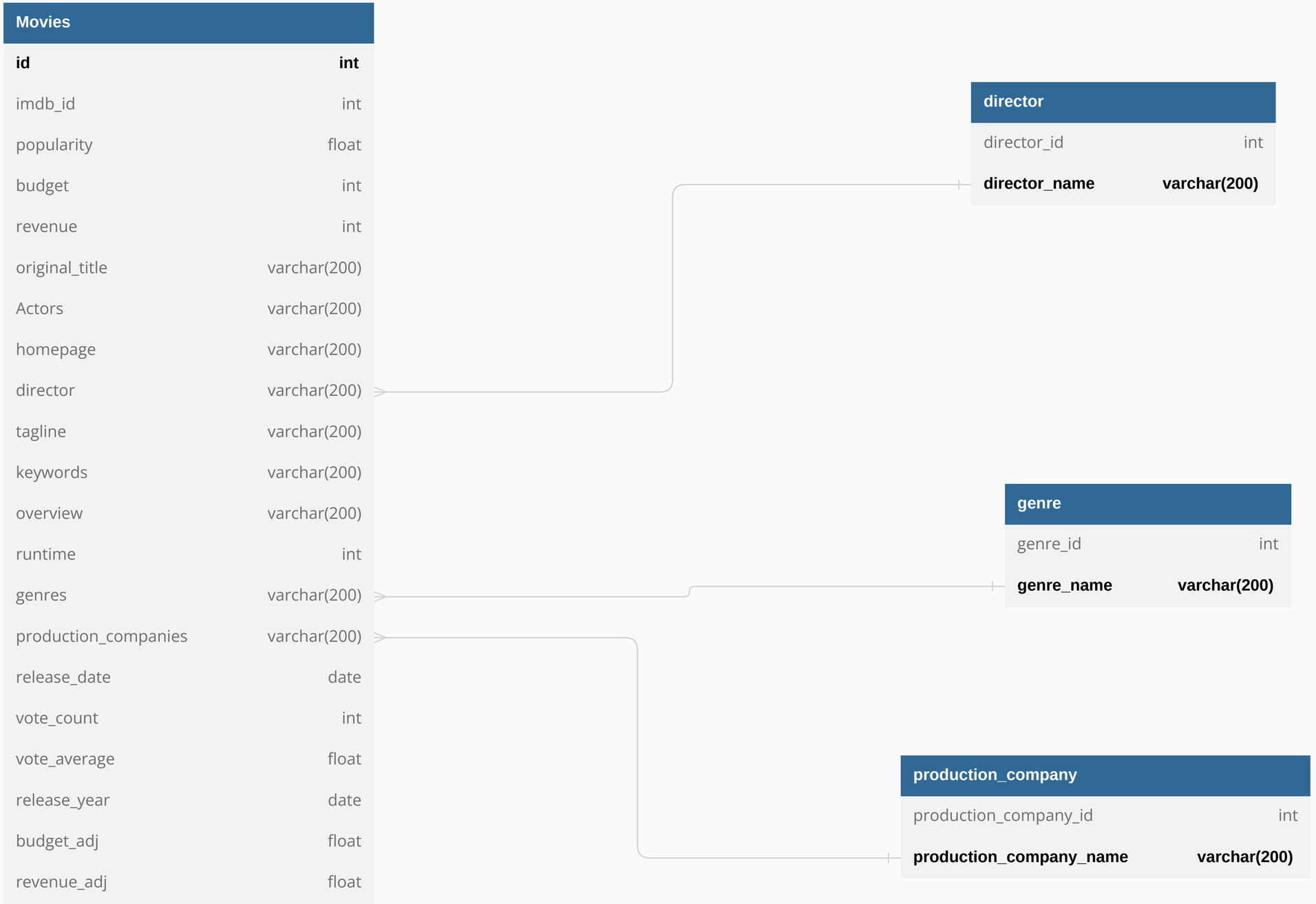
Used PowerBI to:

- To connect the raw data
- Build a relational data model
- Create new calculated columns and DAX measures
- Design an interactive report to analyze and visualize the data

THE SUMMARY

After performing data wrangling, data modelling and data visualization. Here, are the following insights from the dataset.

- Total number of movies produced from 1977 to 2015 was 3,701 with the budget of 142,473,661,668 which generate total revenue of 407,278,108,892 and profit of 264,804,447,224 (all the values are in naira)
- Paramount Pictures is the top spending production house by 12,752,208,041(naira).
- Avatar has the highest profit of 2,544,505,847(naira) which account for 0.94% of the totalrevenue



SQL SCRIPTS FOR TMDB

```
CREATE TABLE genre (
  genre_id INT,
  genre_name VARCHAR(200) PRIMARY KEY
);
CREATE TABLE director (
  director_id INT,
  director_name VARCHAR(200) PRIMARY KEY
);
CREATE TABLE production_company (
  production_company_id INT,
  production_company_name VARCHAR(200) PRIMARY KEY
);
CREATE TABLE Movies (
  id INT PRIMARY KEY,
  imdb_id INT,
  popularity FLOAT,
  budget INT,
  revenue INT,
  original_title VARCHAR(200),
  Actors VARCHAR(200),
  homepage VARCHAR(200),
  director VARCHAR(200),
  tagline VARCHAR(200),
  keywords VARCHAR(200),
  overview VARCHAR(200),
  runtime INT,
  genres VARCHAR(200),
  production_companies VARCHAR(200),
  release_date DATE,
  vote_count INT,
  vote_average FLOAT,
```

```
    release_year DATE,
    budget_adj FLOAT,
    revenue_adj FLOAT,
    FOREIGN KEY (genres)
      REFERENCES genre (genre_name),
    FOREIGN KEY (production_companies)
      REFERENCES production_company (production_company_name),
    FOREIGN KEY (director)
      REFERENCES director (director_name)
);
```

Read Dataset

In [1]:
```
#First, we need to import pandas as a library for data wrangling
#Then use pd.read_excel function to import the dataset into pandas
# and save the dataset in a variable called Movies
# Therefore, whenever we to read or to import the dataset, just call the variable "Movies" to make the changes.
import pandas as pd
Movies = pd.read_excel(r'C:\Users\Lenovo\Desktop\Personal Documents\DataSet\movies_data.xlsx')
```

In [2]:
```
# Read the dataset with the first top 10 rows.
Movies.head(10)
```

Out[2]:

	id	imdb_id	popularity	budget	revenue	original_title	cast	homepage	director	tagline	...	overview	runtime	genres	product
0	135397	tt0369610	32.985763	150000000	1513528810	Jurassic World	Chris Pratt|Bryce Dallas Howard|Irrfan Khan|V...	http://www.jurassicworld.com/	Colin Trevorrow	The park is open.	...	Twenty-two years after the events of Jurassic ...	124	Action|Adventure|Science Fiction|Thriller	Universa Entertain
1	76341	tt1392190	28.419936	150000000	378436354	Mad Max: Fury Road	Tom Hardy|Charlize Theron|Hugh Keays-Byrne|Nic...	http://www.madmaxmovie.com/	George Miller	What a Lovely Day.	...	An apocalyptic story set in the furthest reach...	120	Action|Adventure|Science Fiction|Thriller	V Picture
2	262500	tt2908446	13.112507	110000000	295238201	Insurgent	Shailene Woodley|Theo James|Kate Winslet|Ansel...	http://www.thedivergentseries.movie/#insurgent	Robert Schwentke	One Choice Can Destroy You	...	Beatrice Prior must confront her inner demons ...	119	Adventure|Science Fiction|Thriller	Entertain Pi
3	140607	tt2488496	11.173104	200000000	2068178225	Star Wars: The Force Awakens	Harrison Ford|Mark Hamill|Carrie Fisher|Adam D...	http://www.starwars.com/films/star-wars-episod...	J.J. Abrams	Every generation has a story.	...	Thirty years after defeating the Galactic Empi...	136	Action|Adventure|Science Fiction|Fantasy	Luc Produc
4	168259	tt2820852	9.335014	190000000	1506249360	Furious 7	Vin Diesel|Paul Walker|Jason Statham|Michelle ...	http://www.furious7.com/	James Wan	Vengeance Hits Home	...	Deckard Shaw seeks revenge against Dominic Tor...	137	Action|Crime|Thriller	Film
5	281957	tt1663202	9.110700	135000000	532950503	The Revenant	Leonardo DiCaprio|Tom Hardy|Will Poulter|Domhn...	http://www.foxmovies.com/movies/the-revenant	Alejandro González Iñárritu	In the 1820s, a frontiersman, Hugh Glass, sets...	...	In the 1820s, a frontiersman, Hugh Glass, sets...	156	Western|Drama|Adventure|Thriller	En Way|Ca
6	87101	tt1340138	8.654359	155000000	440603537	Terminator Genisys	Arnold Schwarzenegger|Jason Clarke|Emilia Clar...	http://www.terminatormovie.com/	Alan Taylor	Reset the future	...	The year is 2029. John Connor, leader of the ...	125	Science Fiction|Action|Thriller|Adventure	Pi
7	286217	tt3659388	7.667400	108000000	595380321	The Martian	Matt Damon|Jessica Chastain|Kristen Wiig|Jeff ...	http://www.foxmovies.com/movies/the-martian	Ridley Scott	Bring Him Home	...	During a manned mission to Mars, Astronaut Mar...	141	Drama|Adventure|Science Fiction	Twent Film C
8	211672	tt2293640	7.404165	74000000	1156730962	Minions	Sandra Bullock|Jon Hamm|Michael Keaton|Allison...	http://www.minionsmovie.com/	Kyle Balda|Pierre Coffin	Before Gru, they had a history of bad bosses	...	Minions Stuart, Kevin and Bob are recruited by...	91	Family|Animation|Adventure|Comedy	Pict
9	150540	tt2096673	6.326804	175000000	853708609	Inside Out	Amy Poehler|Phyllis Smith|Richard Kind|Bill Ha...	http://movies.disney.com/inside-out	Pete Docter	Meet the little voices inside your head.	...	Growing up can be a bumpy road, and it's no ex...	94	Comedy|Animation|Family	Pictures

10 rows × 21 columns

In [3]:
```
# Data summary which includes data types, number of rows
Movies.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10866 entries, 0 to 10865
Data columns (total 21 columns):
 #   Column          Non-Null Count  Dtype
---  ------          --------------  -----
 0   id              10866 non-null  int64
 1   imdb_id         10856 non-null  object
 2   popularity      10866 non-null  float64
 3   budget          10866 non-null  int64
 4   revenue         10866 non-null  int64
 5   original_title  10866 non-null  object
 6   cast            10790 non-null  object
 7   homepage        2936 non-null   object
 8   director        10822 non-null  object
 9   tagline         8042 non-null   object
 10  keywords        9373 non-null   object
 11  overview        10862 non-null  object
 12  runtime         10866 non-null  int64
 13  genres          10843 non-null  object
 14  production_companies  9836 non-null  object
 15  release_date    10866 non-null  object
 16  vote_count      10866 non-null  int64
 17  vote_average    10866 non-null  float64
 18  release_year    10866 non-null  int64
 19  budget_adj      10866 non-null  float64
 20  revenue_adj     10866 non-null  float64
dtypes: float64(4), int64(6), object(11)
memory usage: 1.7+ MB
Movies.isna().sum()
```

In [4]:
```
# From the data summary, we observed that columns like imdb_id, cast, homepage,director,tagline,keywords,overview, genres, and production_companies
  # contain null values because the count rows is less than the total entries.(10866)
  # To verify we will use isna to confirm the total number of null values
  Movies.isna().sum()
```

Out[4]:
```
id                         0
imdb_id                   10
popularity                 0
budget                     0
revenue                    0
original_title             0
cast                     76
homepage               7930
director                 44
tagline                2824
keywords               1493
overview                  4
runtime                   0
genres                   23
production_companies   1030
release_date              0
vote_count                0
vote_average              0
release_year              0
budget_adj                0
revenue_adj               0
dtype: int64
```

Data Wrangling

In [5]:
```
# Remove the delimiter from the generes columns
Movies['genres'] = Movies['genres'].str.split('|').str.get(0)
Movies['production_companies'] = Movies['production_companies'].str.split('|').str.get(0)
Movies
```

Out[5]:

	id	imdb_id	popularity	budget	revenue	original_title	cast	homepage	director	tagline	...	overview	runtime	genres	production_companies	re
0	135397	tt0369610	32.985763	150000000	1513528810	Jurassic World	Chris Pratt|Bryce Dallas Howard|Irrfan Khan|V...	http://www.jurassicworld.com/	Colin Trevorrow	The park is open.	...	Twenty-two years after the events of Jurassic ...	124	Action	Universal Studios	20
1	76341	tt1392190	28.419936	150000000	378436354	Mad Max: Fury Road	Tom Hardy|Charlize Theron|Hugh Keays-Byrne|Nic...	http://www.madmaxmovie.com/	George Miller	What a Lovely Day.	...	An apocalyptic story set in the furthest reach...	120	Action	Village Roadshow Pictures	20
2	262500	tt2908446	13.112507	110000000	295238201	Insurgent	Shailene Woodley|Theo James|Kate Winslet|Ansel...	http://www.thedivergentseries.movie/#insurgent	Robert Schwentke	One Choice Can Destroy You	...	Beatrice Prior must confront her inner demons ...	119	Adventure	Summit Entertainment	20
3	140607	tt2488496	11.173104	200000000	2068178225	Star Wars: The Force Awakens	Harrison Ford|Mark Hamill|Carrie Fisher|Adam D...	http://www.starwars.com/films/star-wars-episod...	J.J. Abrams	Every generation has a story.	...	Thirty years after defeating the Galactic Empi...	136	Action	Lucasfilm	20
4	168259	tt2820852	9.335014	190000000	1506249360	Furious 7	Vin Diesel|Paul Walker|Jason Statham|Michelle ...	http://www.furious7.com/	James Wan	Vengeance Hits Home	...	Deckard Shaw seeks revenge against Dominic Tor...	137	Action	Universal Pictures	20
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
10861	21	tt0060371	0.080598	0	0	The Endless Summer	Michael Hynson|Robert August|Lord Tally Ho! B...	NaN	Bruce Brown	NaN	...	The Endless Summer, by Bruce Brown, is one of ...	95	Documentary	Bruce Brown Films	
10862	20379	tt0060472	0.065543	0	0	Grand Prix	James Garner|Eva Marie Saint|Yves Montand|Bri...	NaN	John Frankenheimer	Cinerama sweeps YOU into a drivers seat and sp...	...	Grand Prix stars James Garner as Pete Aron is fired by his te...	176	Action	Cherokee Productions	
10863	39768	tt0060161	0.065141	0	0	Beregis Avtomobil	Innokentiy Smoktunovskiy|Oleg Efremov|Georgi Z...	NaN	Eldar Ryazanov	NaN	...	An insurance agent who moonlights as a car cra...	94	Mystery	Mosfilm	
10864	21449	tt0061177	0.064317	0	0	What's Up, Tiger Lily?	Tatsuya Mihashi|Akiko Wakabayashi|Mie Hama|Joh...	NaN	Woody Allen	WOODY ALLEN STRIKES BACK!	...	In comic Woody Allen's film debut, he took the ...	80	Action	Benedict Pictures Corp.	
10865	22293	tt0060666	0.035919	19000	0	Manos: The Hands of Fate	Harold P. Warren|Tom Neyman|John Reynolds|Dan...	NaN	Harold P. Warren	It's Shocking! It's Beyond Your Imagination!	...	A family gets lost on the road and stumbles up...	74	Horror	Norm-Iris	

10866 rows × 21 columns

In [6]:
```
# 1. Change release date format since the data is in a string format.
  Movies['release_date'] = pd.to_datetime(Movies['release_date'])
  Movies['release_date'].head(10)
```

Out[6]:
```
0   2015-06-09
1   2015-05-13
2   2015-03-18
3   2015-12-15
4   2015-04-01
5   2015-12-25
6   2015-06-23
7   2015-09-30
8   2015-06-17
9   2015-09-09
Name: release_date, dtype: datetime64[ns]
```

In [7]:
```
# 2. Remove duplicates.
  Movies.drop_duplicates(inplace = True)
  Movies.shape
```

Out[7]:
```
(10865, 21)
```

In [8]:
```
# 3. Remove data value that are zero in budget and revenue column.
  Movies = Movies[Movies['budget'] != 0]
  Movies = Movies[Movies['revenue'] != 0]
```

In [9]:
```
Movies.rename({"cast" : "Actors"}, axis = 1, inplace = True)
  Movies
  Movies = Movies
  movies
```

Out[9]:

	id	imdb_id	popularity	budget	revenue	original_title	Actors	homepage	director	tagline	...	overview	runtime	genres	production_companies	relea
0	135397	tt0369610	32.985763	150000000	1513528810	Jurassic World	Chris Pratt|Bryce Dallas Howard|Irrfan Khan|V...	http://www.jurassicworld.com/	Colin Trevorrow	The park is open.	...	Twenty-two years after the events of Jurassic ...	124	Action	Universal Studios	20
1	76341	tt1392190	28.419936	150000000	378436354	Mad Max: Fury Road	Tom Hardy|Charlize Theron|Hugh Keays-Byrne|Nic...	http://www.madmaxmovie.com/	George Miller	What a Lovely Day.	...	An apocalyptic story set in the furthest reach...	120	Action	Village Roadshow Pictures	20
2	262500	tt2908446	13.112507	110000000	295238201	Insurgent	Shailene Woodley|Theo James|Kate Winslet|Ansel...	http://www.thedivergentseries.movie/#insurgent	Robert Schwentke	One Choice Can Destroy You	...	Beatrice Prior must confront her inner demons ...	119	Adventure	Summit Entertainment	20
3	140607	tt2488496	11.173104	200000000	2068178225	Star Wars: The Force Awakens	Harrison Ford|Mark Hamill|Carrie Fisher|Adam D...	http://www.starwars.com/films/star-wars-episod...	J.J. Abrams	Every generation has a story.	...	Thirty years after defeating the Galactic Empi...	136	Action	Lucasfilm	20
4	168259	tt2820852	9.335014	190000000	1506249360	Furious 7	Vin Diesel|Paul Walker|Jason Statham|Michelle ...	http://www.furious7.com/	James Wan	Vengeance Hits Home	...	Deckard Shaw seeks revenge against Dominic Tor...	137	Action	Universal Pictures	20
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
10822	396	tt0061184	0.670274	7500000	33736689	Who's Afraid of Virginia Woolf?	Elizabeth Taylor|Richard Burton|George Segal|S...	NaN	Mike Nichols	You are cordially invited to ... George and Martha, ...	...	Mike Nichols's film version of Edward Albee's play ...	131	Drama	Chimault Productions	20
10828	5760	tt0061107	0.402730	3000000	13000000	Tom Curtain	Paul Newman|Julie Andrews|Ja Kedrova|Hans|S...	NaN	Alfred Hitchcock	It tears you apart with suspense!	...	An American scientist is persuaded publicly to defect to ...	128	Mystery	Universal Pictures	20
10829	6644	tt0061619	0.395668	4653000	6000000	El Dorado	John Wayne|Robert Mitchum|James Caan|Charlene ...	NaN	Howard Hawks	It's the Big One with the Big Two	...	Cole Thornton is a gunfighter to hire, joins fo...	120	Action	Paramount Pictures	20
10835	5923	tt0060934	0.299911	12000000	20000000	The Sand Pebbles	Steve McQueen|Richard Attenborough|Richard Cher...	NaN	Robert Wise	This is the heroic story of the men ... on the U.S.S. ...	...	Engineer Jake Holman arrives aboard the gunboa...	182	Action	Twentieth Century Fox Film Corporation	20
10848	2161	tt0060397	0.207257	5115000	12000000	Fantastic Voyage	Stephen Boyd|Raquel Welch|Edmond O'Brien|Donal...	NaN	Richard Fleischer	A Fantastic and Spectacular Voyage ... Through ...	...	The science of miniaturization has been unlock...	100	Adventure	Twentieth Century Fox Film Corporation	20

3854 rows × 21 columns

Create Dataset for Genres, Production Companies, and Director

In [10]:
```
# Create a dataframe Genres
  Genre = Movies.drop_duplicates(subset = ['genres'])
  Genre_table = Genre.drop(["budget_adj","revenue_adj","budget","overview","keywords","tagline","homepage","imdb_id","id","popularity","budget","revenue","original_title","Actors","d
  Genre_table.sort_index(ignore_index = True, inplace = True)
  Genre_table.head(5)
```

Out[10]:

	genres
0	Action
1	Adventure
2	Western
3	Science Fiction
4	Drama

In [11]:
```
Director = Movies.drop_duplicates(subset = ["director"])
  Director_table = Genre.drop(["budget_adj","revenue_adj","budget","overview","keywords","tagline","homepage","imdb_id","id","popularity","budget","revenue","original_title","Actors"
  Director_table.sort_index(ignore_index = True, inplace = True)
  Director_table.head(5)
```

Out[11]:

	director
0	Colin Trevorrow
1	Robert Schwentke
2	Alejandro González Iñárritu
3	Alan Taylor
4	Ridley Scott

In [12]:
```
Production_Company = Movies.drop_duplicates(subset = ["production_companies"])
  Production_Company_table = Genre.drop(["budget_adj","revenue_adj","budget","overview","keywords","tagline","homepage","imdb_id","id","popularity","budget","revenue","original_title
  Production_Company_table.sort_index(ignore_index = True, inplace = True)
  Production_Company_table.head(5)
```

Out[12]:

	production_companies
0	Universal Studios
1	Summit Entertainment
2	Regency Enterprises
3	Paramount Pictures
4	Twentieth Century Fox Film Corporation

Load Data in MySQL

In [13]:
```
pip install sqlalchemy

Requirement already satisfied: sqlalchemy in c:\users\lenovo\anaconda3\lib\site-packages (1.4.32)
Requirement already satisfied: greenlet!=0.4.17 in c:\users\lenovo\anaconda3\lib\site-packages (from sqlalchemy) (1.1.1)
Note: you may need to restart the kernel to use updated packages.
```

In [14]:
```
import sqlalchemy
```

In [15]:
```
pip install mysql

Requirement already satisfied: mysql in c:\users\lenovo\anaconda3\lib\site-packages (0.0.3)
Requirement already satisfied: mysqlclient in c:\users\lenovo\anaconda3\lib\site-packages (from mysql) (2.1.1)
Note: you may need to restart the kernel to use updated packages.
```

In [16]:
```
engine1 = sqlalchemy.create_engine("mysql://sammy:password@localhost/tmdb")
```

In [17]:
```
movies_to_sql(name = 'movies', con = engine1, index = False, if_exists = 'replace')
```

Out[17]:
```
3854
```

In [18]:
```
Genre_table_to_sql(name = 'genre', con = engine1, index = False, if_exists = 'replace')
```

Out[18]:
```
19
```

In [19]:
```
Director_table_to_sql(name = 'director', con = engine1, index = False, if_exists = 'replace')
```

Out[19]:
```
19
```

In [20]:
```
Production_Company_table_to_sql(name = 'production_company', con = engine1, index = False, if_exists = 'replace')
```

Out[20]:
```
19
```

Load Data in PostgreSQL

In [21]:
```
engine = sqlalchemy.create_engine("postgresql+psycopg2://Uche:diamond@localhost/TMDB")
```

In [22]:
```
movies_to_sql(name = 'movies', con = engine, index = False, if_exists = 'replace')
```

Out[22]:
```
854
```

In [23]:
```
Director_table_to_sql(name = 'director', con = engine, index = False, if_exists = 'replace')
```

Out[23]:
```
19
```

In [24]:
```
Production_Company_table_to_sql(name = 'production_company', con = engine, index = False, if_exists = 'replace')
```

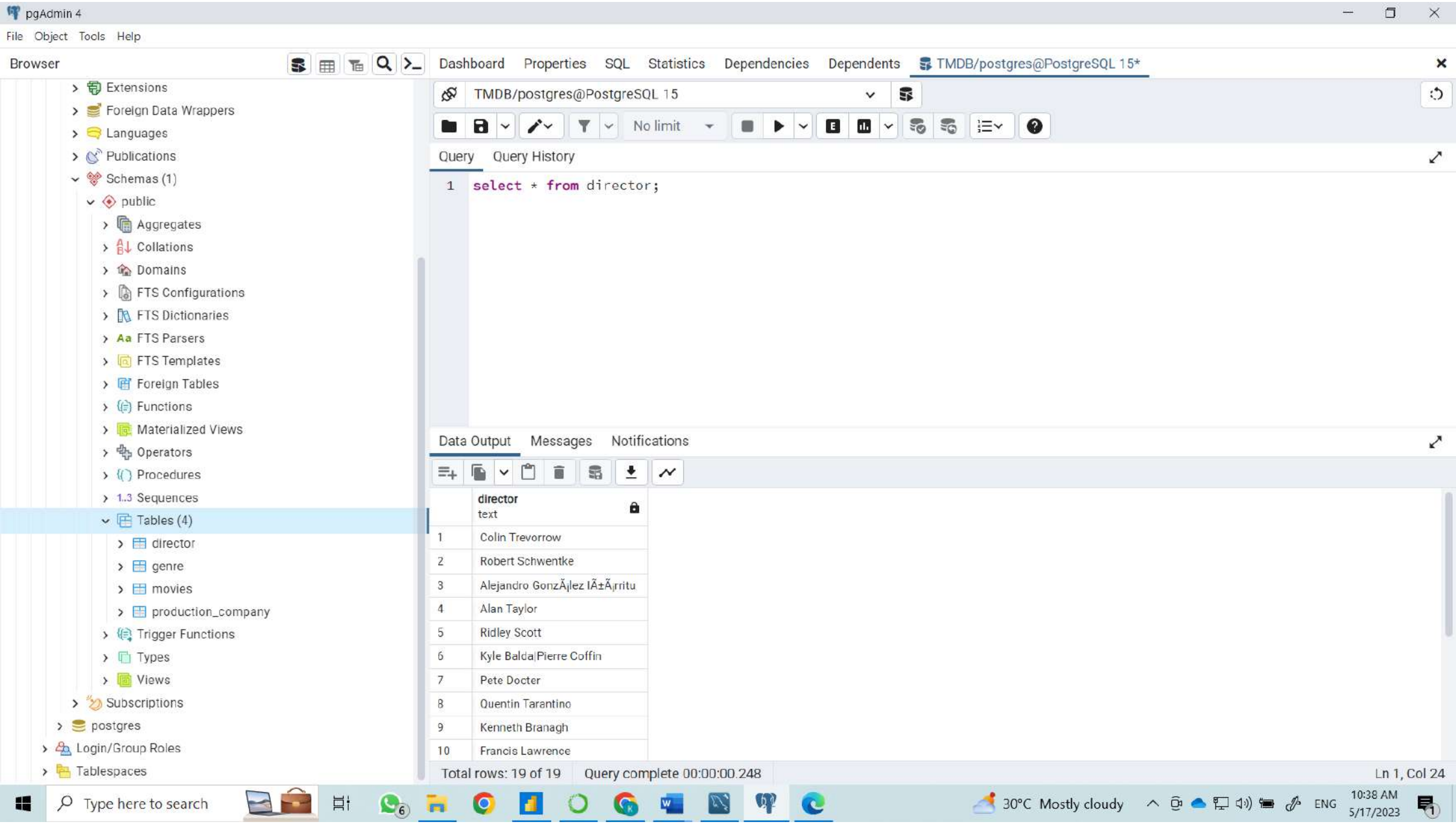
Out[24]:
```
19
```

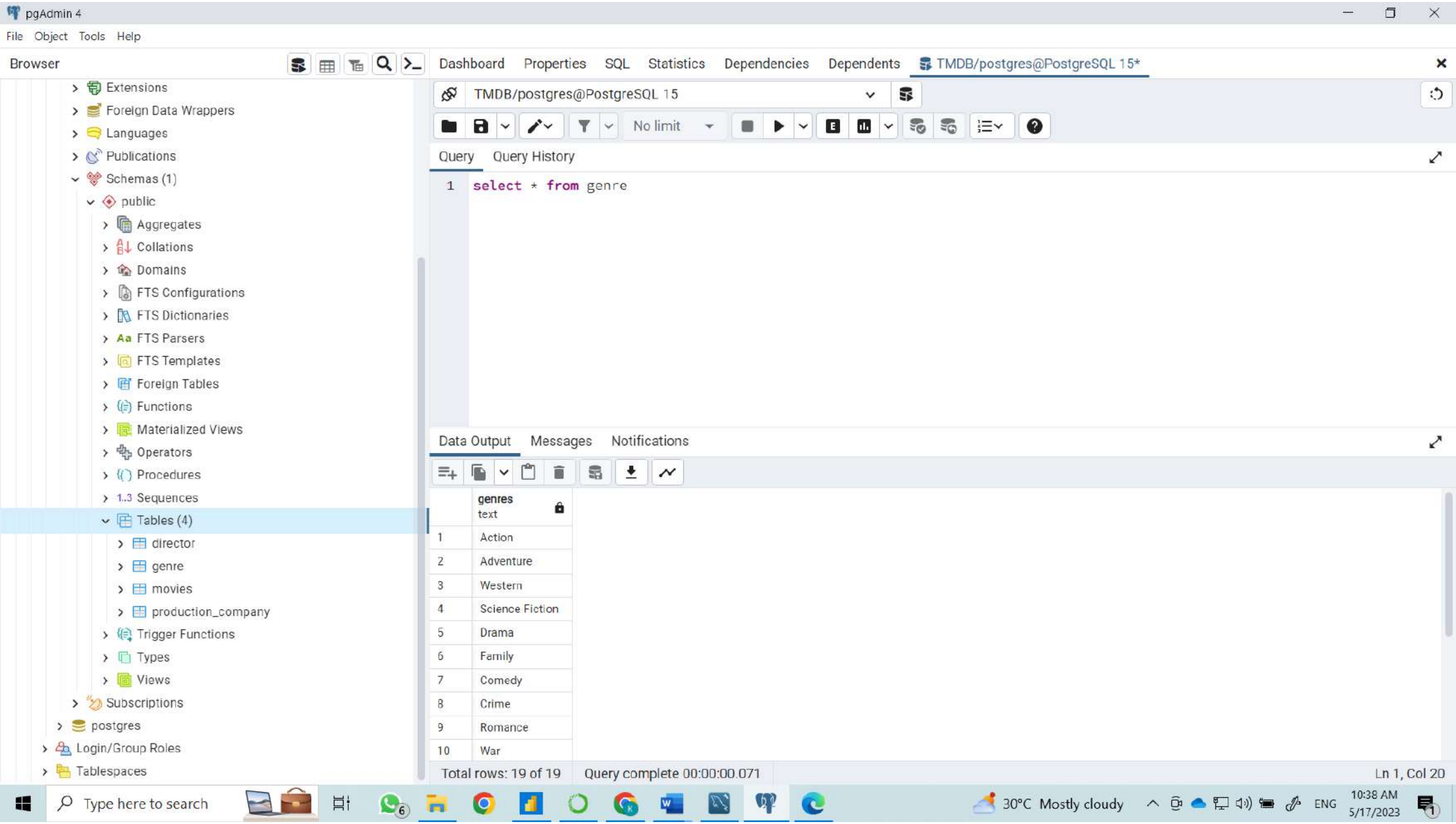
In [25]:
```
Genre_table_to_sql(name = 'genre', con = engine, index = False, if_exists = 'replace')
```

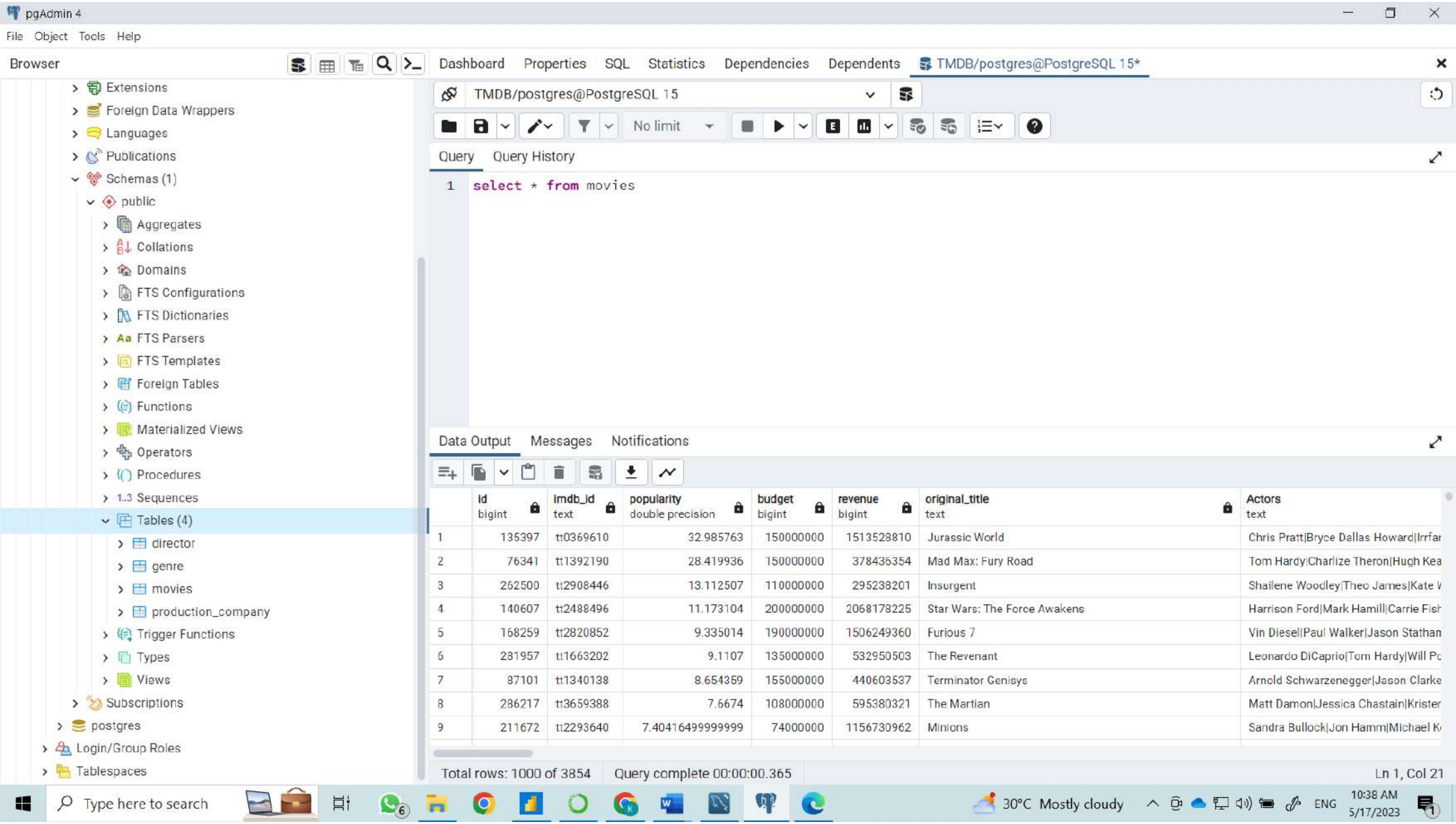
Out[25]:
```
19
```

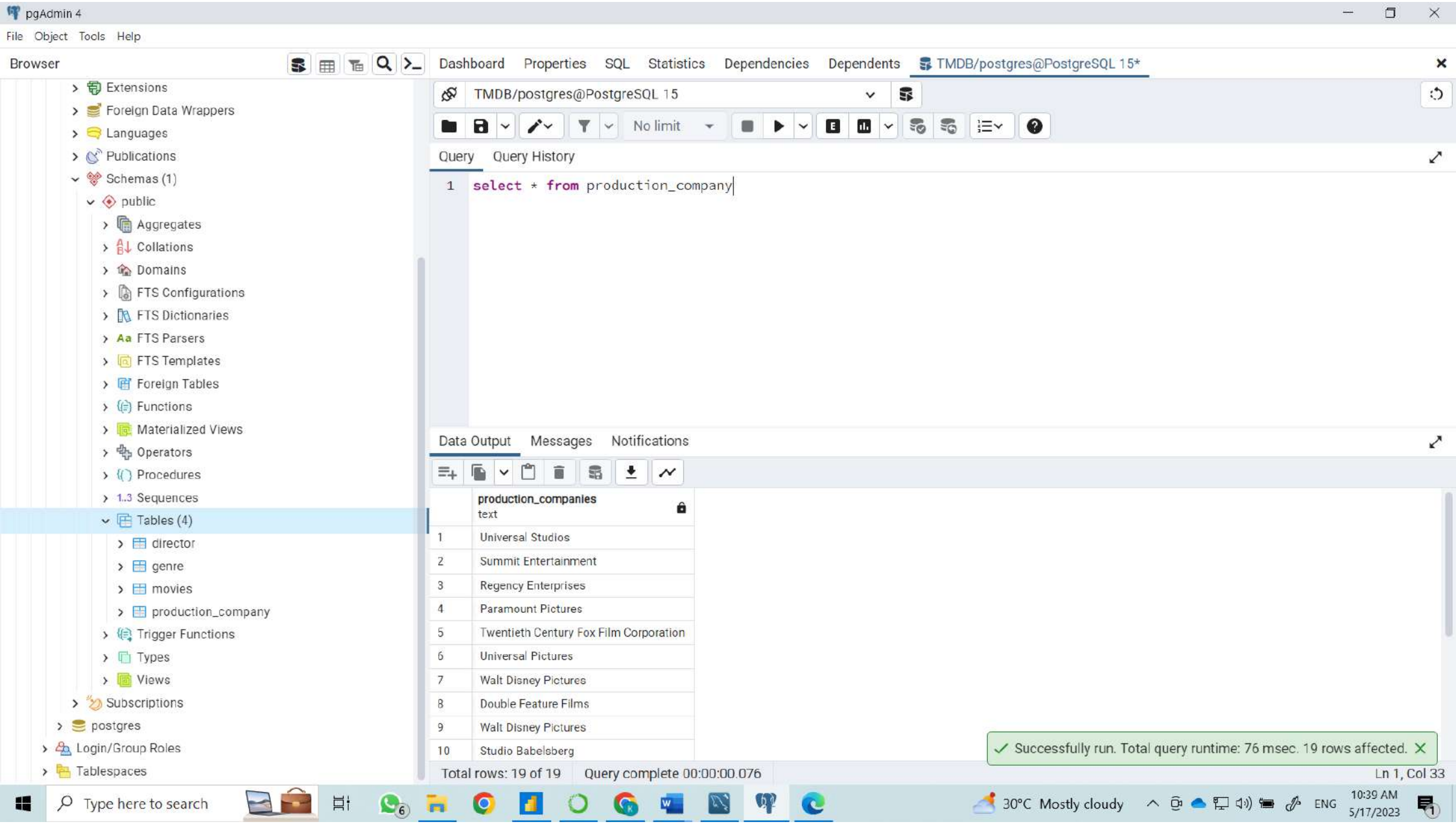
In [ ]:
```

```


pgAdmin 4
File Object Tools Help
Browser
Dashboard Properties SQL Statistics Dependencies Dependents TMDB/postgres@PostgreSQL 15*
Extensions
Foreign Data Wrappers
Languages
Publications
Schemas (1)
public
Aggregates
Collations
Domains
FTS Configurations
FTS Dictionaries
FTS Parsers
FTS Templates
Foreign Tables
Functions
Materialized Views
Operators
Procedures
Sequences
Tables (4)
director
genre
movies
production_company
Trigger Functions
Types
Views
Subscriptions
postgres
Login/Group Roles
Tablespaces
TMDB/postgres@PostgreSQL 15
No limit
Query Query History
1 select * from director;
Data Output Messages Notifications
director
text
1 Colin Trevorrow
2 Robert Schwentke
3 Alejandro GonzÃ¡lez IÃ±Ã¡rritu
4 Alan Taylor
5 Ridley Scott
6 Kyle Balda|Pierre Coffin
7 Pete Docter
8 Quentin Tarantino
9 Kenneth Branagh
10 Francis Lawrence
Total rows: 19 of 19
Query complete 00:00:00.248
Ln 1, Col 24
Type here to search
30°C Mostly cloudy
ENG
10:38 AM
5/17/2023


pgAdmin 4
File Object Tools Help
Browser
Dashboard Properties SQL Statistics Dependencies Dependents TMDB/postgres@PostgreSQL 15*
Extensions
Foreign Data Wrappers
Languages
Publications
Schemas (1)
public
Aggregates
Collations
Domains
FTS Configurations
FTS Dictionaries
FTS Parsers
FTS Templates
Foreign Tables
Functions
Materialized Views
Operators
Procedures
Sequences
Tables (4)
director
genre
movies
production_company
Trigger Functions
Types
Views
Subscriptions
postgres
Login/Group Roles
Tablespaces
TMDB/postgres@PostgreSQL 15
No limit
Query Query History
1 select * from genre
Data Output Messages Notifications
genres text
1 Action
2 Adventure
3 Western
4 Science Fiction
5 Drama
6 Family
7 Comedy
8 Crime
9 Romance
10 War
Total rows: 19 of 19
Query complete 00:00:00.071
Ln 1, Col 20
Type here to search
30°C Mostly cloudy
ENG
10:38 AM
5/17/2023


pgAdmin 4
File Object Tools Help
Browser
Dashboard Properties SQL Statistics Dependencies Dependents TMDB/postgres@PostgreSQL 15*
Extensions
Foreign Data Wrappers
Languages
Publications
Schemas (1)
public
Aggregates
Collations
Domains
FTS Configurations
FTS Dictionaries
FTS Parsers
FTS Templates
Foreign Tables
Functions
Materialized Views
Operators
Procedures
Sequences
Tables (4)
director
genre
movies
production_company
Trigger Functions
Types
Views
Subscriptions
postgres
Login/Group Roles
Tablespaces
TMDB/postgres@PostgreSQL 15
No limit
Query Query History
1 select * from movies
Data Output Messages Notifications
id bigint
imdb_id text
popularity double precision
budget bigint
revenue bigint
original_title text
Actors text
1 135397 tt0369610 32.985763 150000000 1513528810 Jurassic World Chris Pratt|Bryce Dallas Howard|Irrfar
2 76341 tt1392190 28.419936 150000000 378436354 Mad Max: Fury Road Tom Hardy|Charlize Theron|Hugh Kea
3 262500 tt2908446 13.112507 110000000 295238201 Insurgent Shailene Woodley|Theo James|Kate
4 140607 tt2488496 11.173104 200000000 2068178225 Star Wars: The Force Awakens Harrison Ford|Mark Hamill|Carrie Fish
5 168259 tt2820852 9.335014 190000000 1506249360 Furious 7 Vin Diesel|Paul Walker|Jason Stathan
6 281957 tt1663202 9.1107 135000000 532950503 The Revenant Leonardo DiCaprio|Tom Hardy|Will Pc
7 87101 tt1340138 8.654359 155000000 440603537 Terminator Genisys Arnold Schwarzenegger|Jason Clarke
8 286217 tt3659388 7.6674 108000000 595380321 The Martian Matt Damon|Jessica Chastain|Krister
9 211672 tt2293640 7.40416499999999 74000000 1156730962 Minions Sandra Bullock|Jon Hamm|Michael K
Total rows: 1000 of 3854
Query complete 00:00:00.365
Ln 1, Col 21
Type here to search
30°C Mostly cloudy
ENG
10:38 AM
5/17/2023


pgAdmin 4
File Object Tools Help
Browser
Dashboard Properties SQL Statistics Dependencies Dependents TMDB/postgres@PostgreSQL 15*
Extensions
Foreign Data Wrappers
Languages
Publications
Schemas (1)
public
Aggregates
Collations
Domains
FTS Configurations
FTS Dictionaries
FTS Parsers
FTS Templates
Foreign Tables
Functions
Materialized Views
Operators
Procedures
Sequences
Tables (4)
director
genre
movies
production_company
Trigger Functions
Types
Views
Subscriptions
postgres
Login/Group Roles
Tablespaces
TMDB/postgres@PostgreSQL 15
No limit
Query Query History
1 select * from production_company
Data Output Messages Notifications
production_companies
text
1 Universal Studios
2 Summit Entertainment
3 Regency Enterprises
4 Paramount Pictures
5 Twentieth Century Fox Film Corporation
6 Universal Pictures
7 Walt Disney Pictures
8 Double Feature Films
9 Walt Disney Pictures
10 Studio Babelsberg
Successfully run. Total query runtime: 76 msec. 19 rows affected.
Total rows: 19 of 19
Query complete 00:00:00.076
Ln 1, Col 33
Type here to search
30°C Mostly cloudy
ENG
10:39 AM
5/17/2023

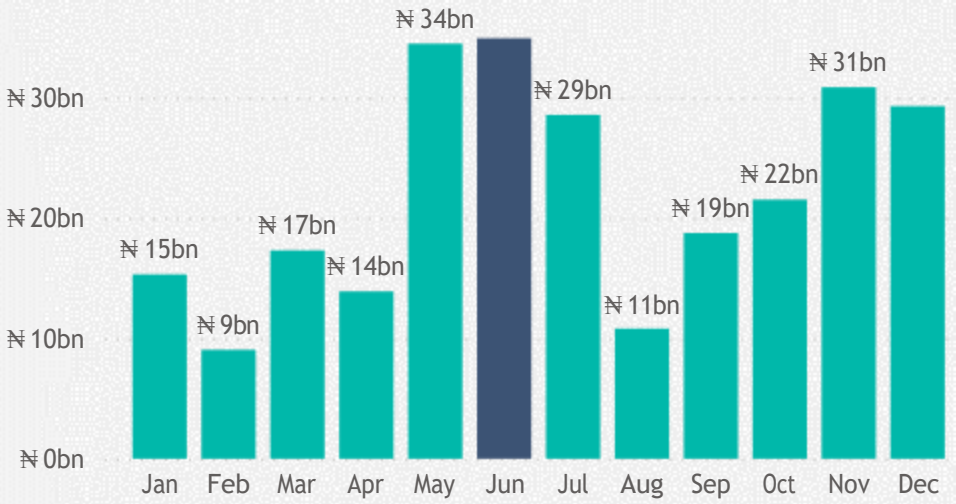
TMDB REPORT

Select Year to Filter: All

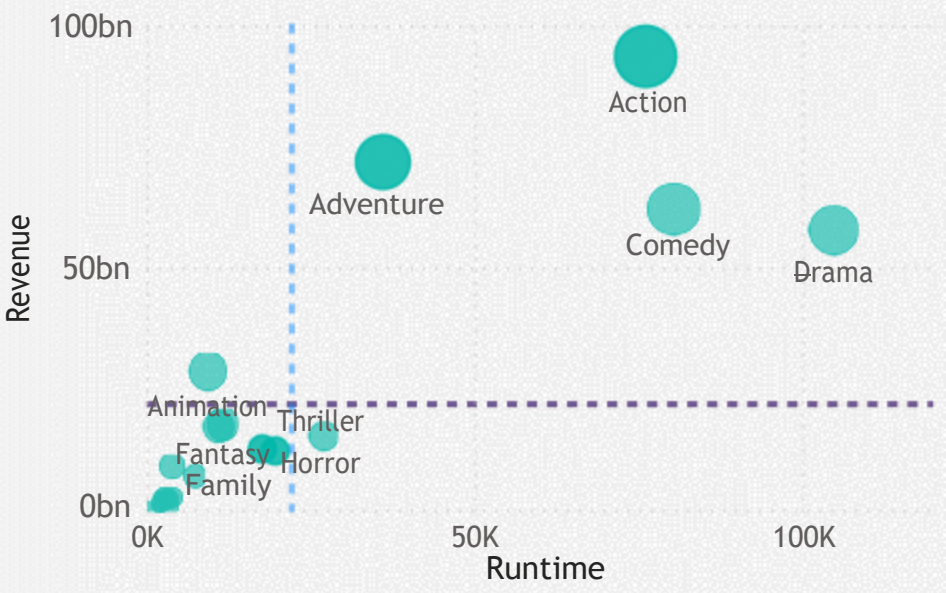
Summary (Comparison between current year selected vs the previous year)

Total Movies	Budget	Revenue	Profit
3701	₦ 142,474M	₦ 407,278M	₦ 264,804M
▲ +4.49% | 3542	▲ +5.29% | ₦ 135,311M	▲ +6.87% | ₦ 381,079M	▲ +7.75% | ₦ 245,768M

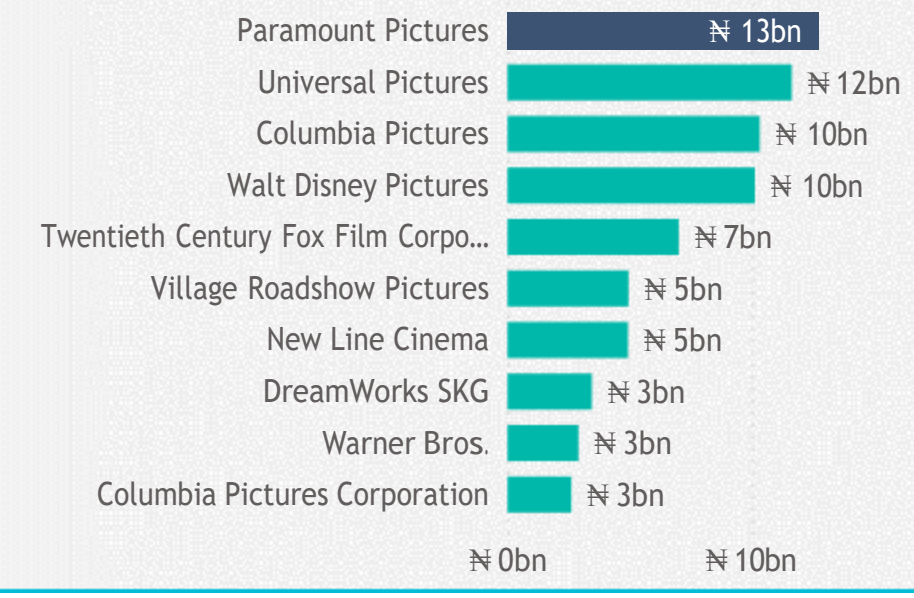
Which month had the highest and lowest profit?



Revenue vs Runtime by Genre.



Top 10 Spending Production Houses



Quick Analysis (Movies by Budget, Revenue and Profit)

Movies	Budget	Revenue	Profit
Avatar	₦ 237,000,000	₦ 2,781,505,847	₦ 2,544,505
Star Wars: The Force Awakens	₦ 200,000,000	₦ 2,068,178,225	₦ 1,868,178,
Titanic	₦ 200,000,000	₦ 1,845,034,188	₦ 1,645,034
The Avengers	₦ 280,000,000	₦ 1,568,080,742	₦ 1,288,080
Jurassic World	₦ 150,000,000	₦ 1,513,528,810	₦ 1,363,528
Furious 7	₦ 190,000,000	₦ 1,506,249,360	₦ 1,316,249
Avengers: Age of Ultron	₦ 280,000,000	₦ 1,405,035,767	₦ 1,125,035
Harry Potter and the Deathly Hallows: Part 2	₦ 125,000,000	₦ 1,327,817,822	₦ 1,202,817
Frozen	₦ 150,000,000	₦ 1,274,219,009	₦ 1,124,219
Iron Man 3	₦ 200,000,000	₦ 1,215,439,994	₦ 1,015,439
Minions	₦ 74,000,000	₦ 1,156,730,962	₦ 1,082,730
Transformers: Dark of the Moon	₦ 195,000,000	₦ 1,123,746,996	₦ 928,746
The Lord of the Rings: The Return of the King	₦ 94,000,000	₦ 1,118,888,979	₦ 1,024,888
Skyfall	₦ 200,000,000	₦ 1,108,561,013	₦ 908,561
The Net	₦ 22,000,000	₦ 1,106,279,658	₦ 1,084,279,
The Dark Knight Rises	₦ 250,000,000	₦ 1,081,041,287	₦ 831,041
Pirates of the Caribbean: Dead Man's Chest	₦ 200,000,000	₦ 1,065,659,812	₦ 865,659
Toy Story 3	₦ 200,000,000	₦ 1,063,171,911	₦ 863,171
Alice in Wonderland	₦ 200,000,000	₦ 1,025,467,110	₦ 825,467
Pirates of the Caribbean: On Stranger Tides	₦ 380,000,000	₦ 1,021,683,000	₦ 641,683
The Hobbit: An Unexpected Journey	₦ 250,000,000	₦ 1,017,003,568	₦ 767,003,
The Dark Knight	₦ 185,000,000	₦ 1,001,921,825	₦ 816,921,
Harry Potter and the Philosopher's Stone	₦ 125,000,000	₦ 976,475,550	₦ 851,475,
Despicable Me 2	₦ 76,000,000	₦ 970,761,885	₦ 894,761,
Pirates of the Caribbean: At World's End	₦ 300,000,000	₦ 961,000,000	₦ 661,000
The Hobbit: The Desolation of Smaug	₦ 250,000,000	₦ 958,400,000	₦ 708,400
The Hobbit: The Battle of the Five Armies	₦ 250,000,000	₦ 955,119,788	₦ 705,119
Harry Potter and the Deathly Hallows: Part 1	₦ 250,000,000	₦ 954,305,868	₦ 704,305
Harry Potter and the Order of the Phoenix	₦ 150,000,000	₦ 938,212,738	₦ 788,212
Harry Potter and the Half-Blood Prince	₦ 250,000,000	₦ 933,959,197	₦ 683,959
The Lord of the Rings: The Two Towers	₦ 79,000,000	₦ 926,287,400	₦ 847,287
Star Wars: Episode I - The Phantom Menace	₦ 115,000,000	₦ 924,317,558	₦ 809,317,
Total	₦ 143,383,048,064	₦ 415,024,218,442	₦ 271,641,170