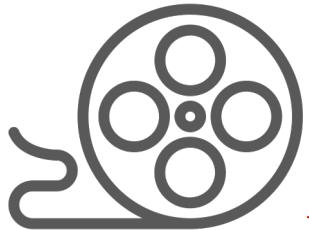


# NETFLIX

## MOVIES AND TV SHOWS Recommendation System

DINA ALSHEHRI



Data Science Bootcamps for SDAIA Academy



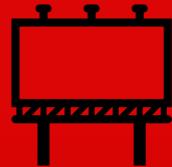


# Outline

- ▶ INTRODUCTION
- ▶ DESIGN AND DATA
- ▶ ALGORITHMS
- ▶ TOOLS
- ▶ COMMUNICATION



# INTRODUCTION



## EDA

Exploratory Data Analysis to know and describe the data in a better way through interactive graphs and visualizations.

## Content-Based Recommendation

A content-based recommendation engine for movies and TV shows on Netflix is presented.

# DESIGN

- The project will be on predicting Netflix Movies and TV Shows based on personal history viewing using recommendation system.
- In order to help understand people preference and how Netflix generate similar catalog of movies and TV shows using genres, title, and release year as features.



# DATA

**DATASET CONTAINS 12 COLUMNS FOR EXPLORATORY ANALYSIS WITH 8807 RECORDS**

**FEATURES:**

- **show\_id**: unique id of each show
- **type**: The category of a show, can be either a Movie or a TV Show
- **title**: Name of the show
- **director**: Name of the director(s) of the show
- **cast**: Name of actors and other cast of the show
- **country**: Name of countries the show is available to watch on Netflix
- **date\_added**: Date when the show was added on Netflix
- **release\_year**: Release year of the show
- **rating**: Show rating on netflix
- **duration**: Time duration of the show
- **listed\_in**: Genre of the show
- **description**: Some text describing the show

# ALGORITHMS

## CLEANING THE DATASET

### **dropna()**

to drop any  
row with  
null value

### **Imputation**

fillna  
function

# ALGORITHMS

## Content-based Filtering

**Recommender Systems**

**Uses similarities  
in features to  
make decisions**

**Nearest Neighbour Algorithm**

**To Calculate  
The “Distance”  
Between The  
Target Tv Show**

**MinMaxScaler**

**Transform features  
by scaling each  
feature to a given  
range**

# TOOLS

- ✓ NUMPY
- ✓ PANDAS
- ✓ MATPLOTLIB
- ✓ SEABORN
- ✓ SKLEARN.LINEAR\_MODEL

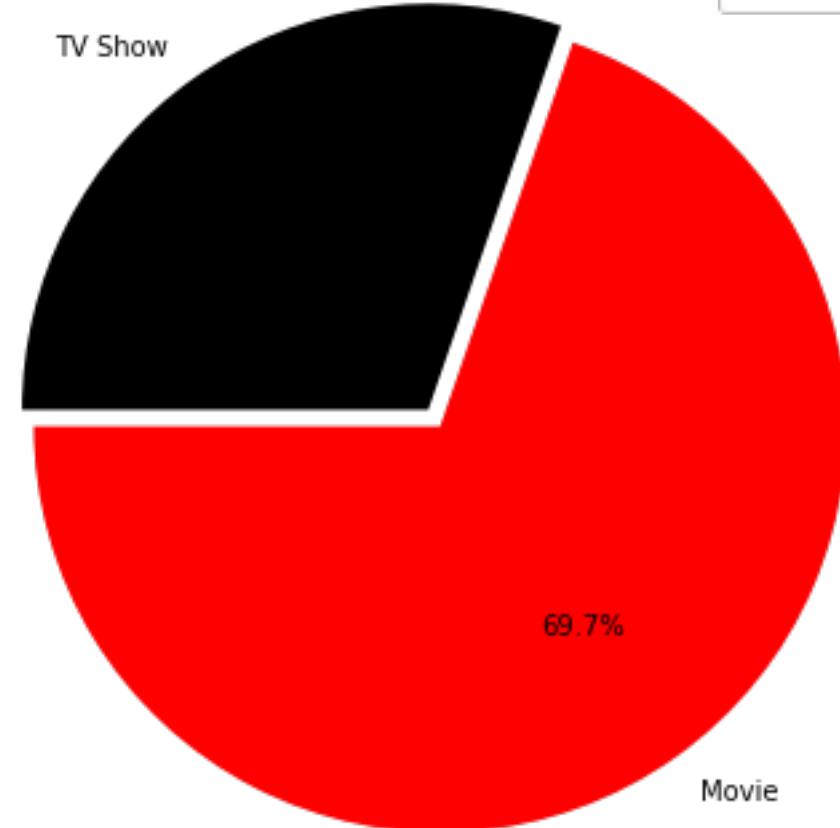


# CONCLUSIONS

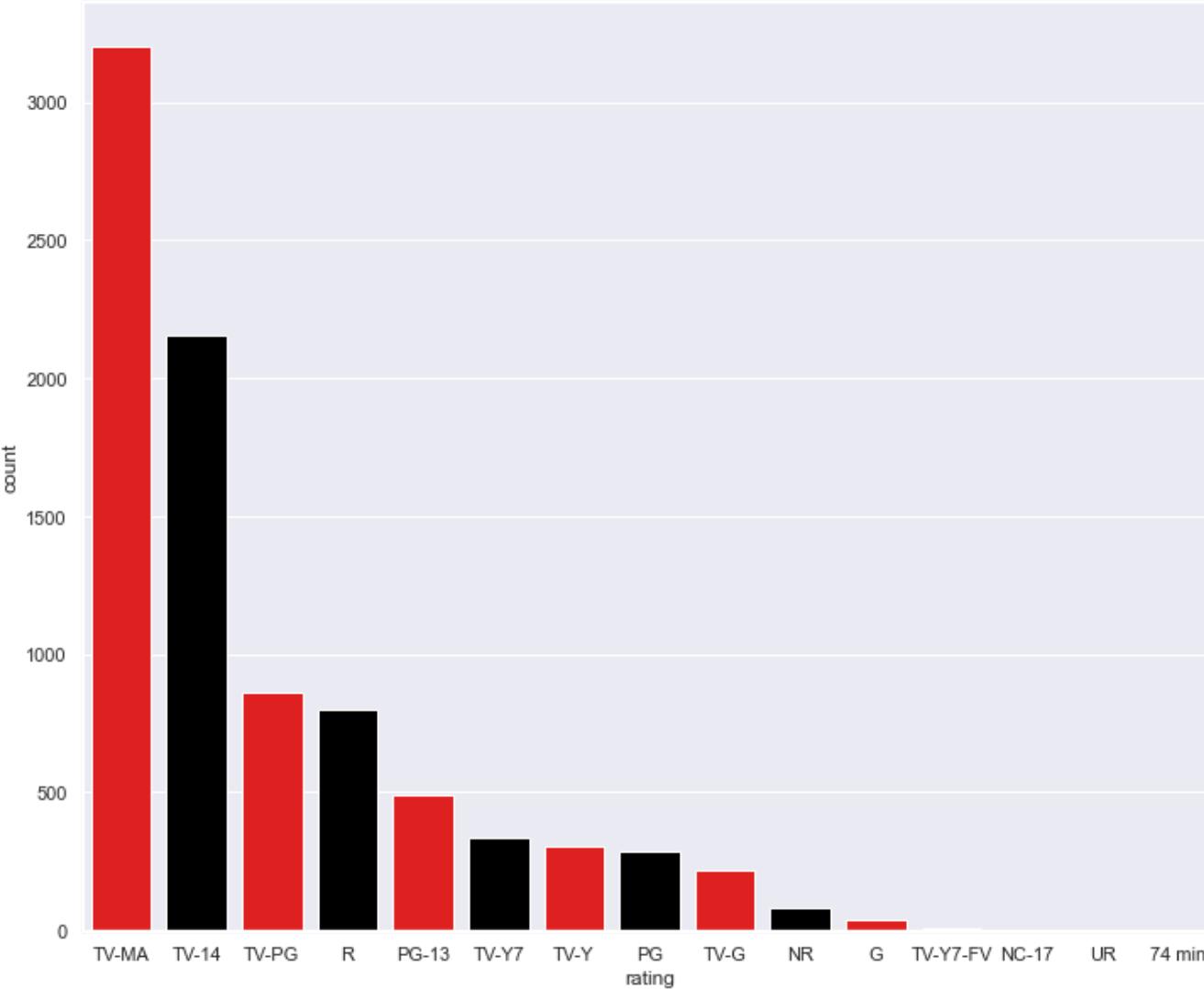


The most content type on Netflix is movies

Percentage of Netflix Titles that are either Movies or TV Shows



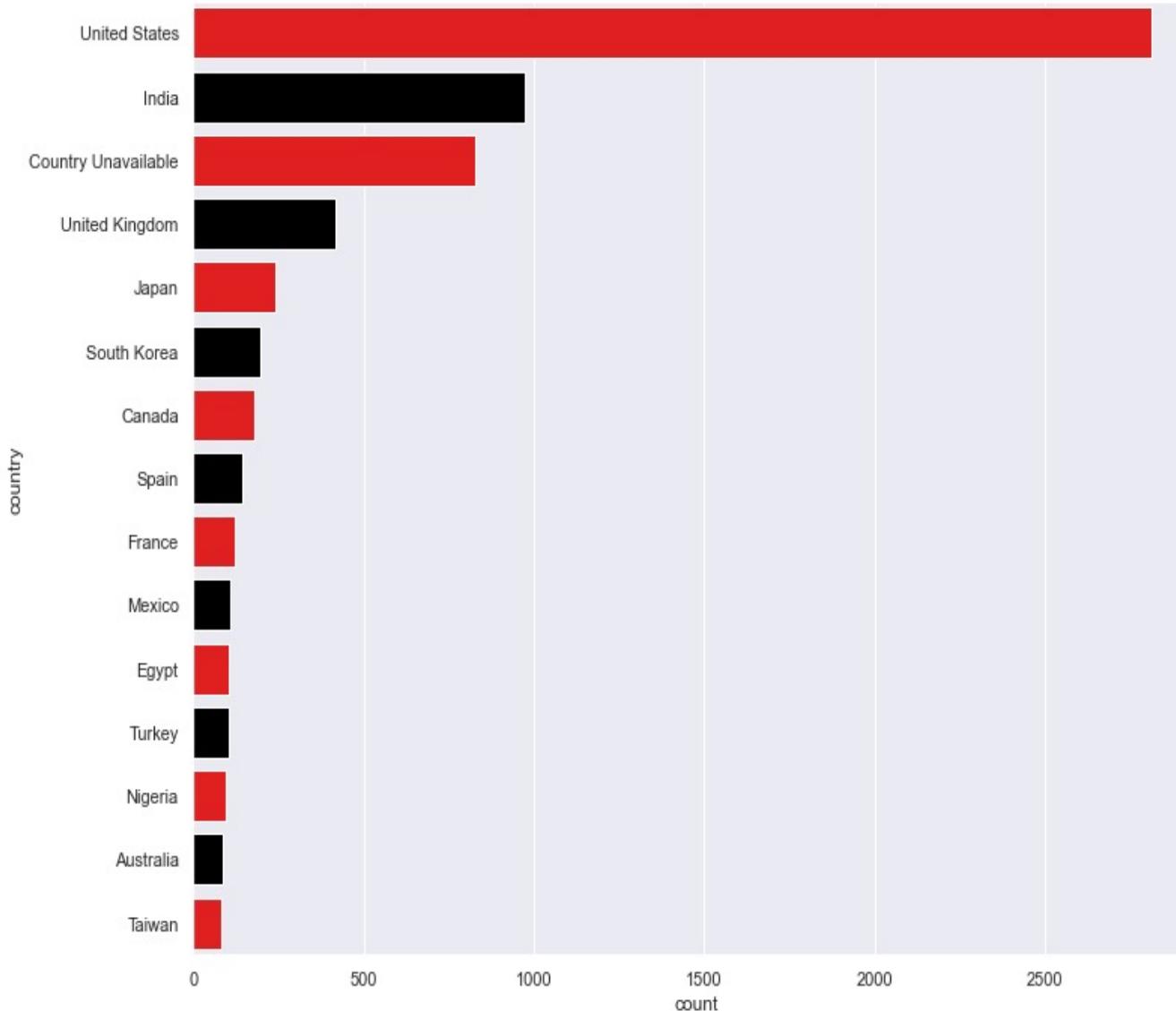
# CONCLUSIONS



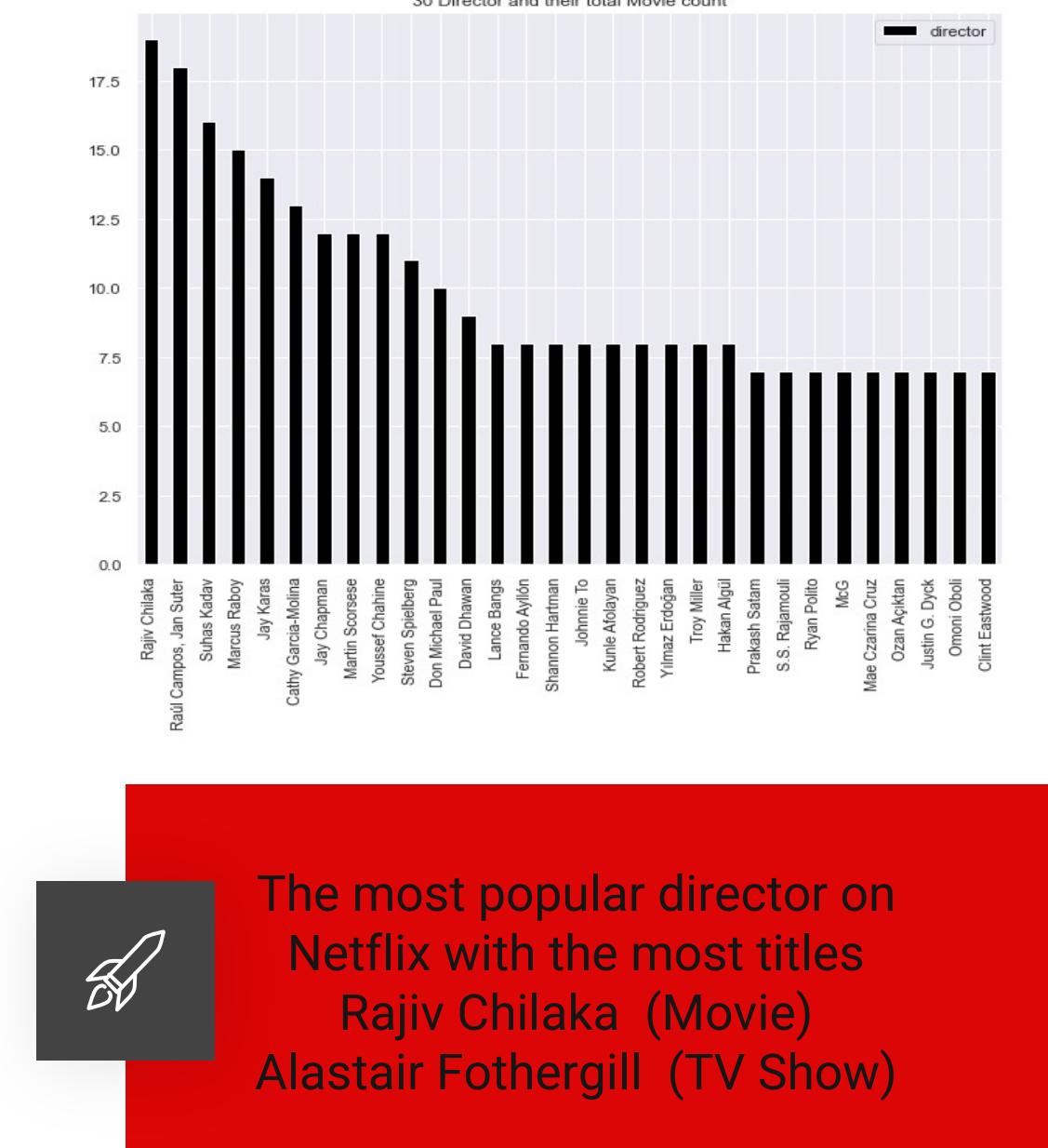
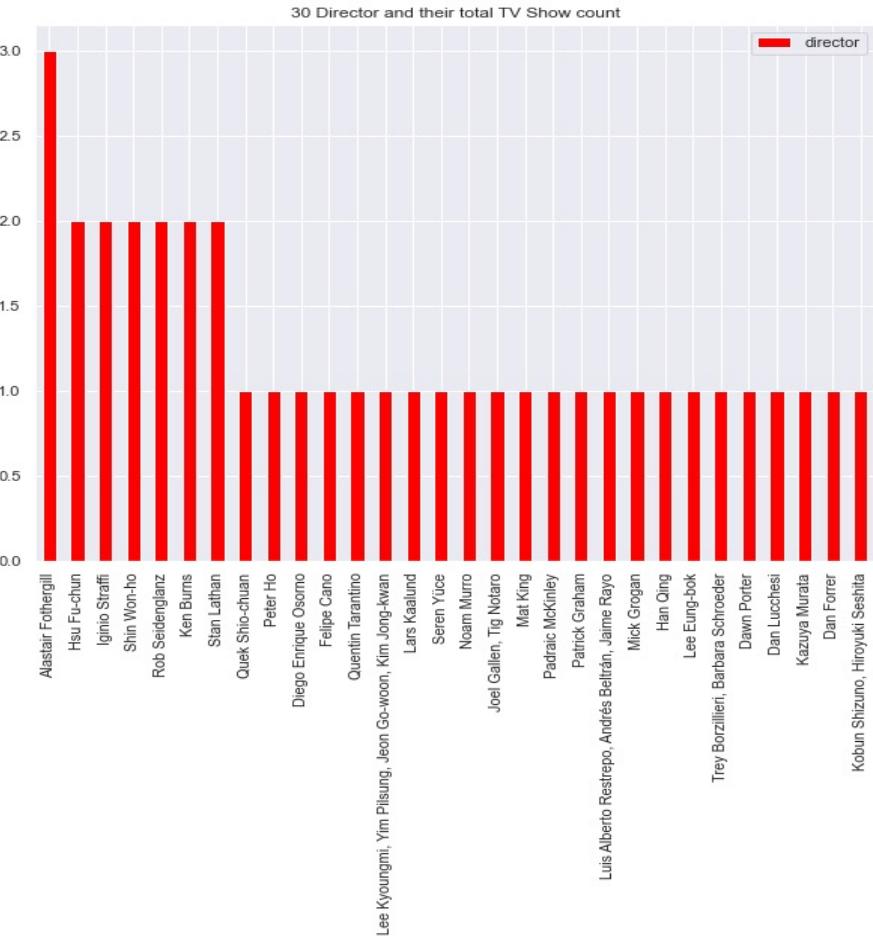
 The largest count of Movies and TV shows is made with a "TV-MA" rating

# CONCLUSIONS

 United States have highest rated content on the Netflix.



# CONCLUSIONS



# CONCLUSIONS

```
recommending_per_watch_history("Emily in Paris")
```

```
[ 'DASH & LILY',
  'The Baker and the Beauty',
  'Emily in Paris',
  'On My Block',
  'Ginny & Georgia']
```

```
recommending_per_watch_history("Workin' Moms")
```

```
[ 'The Ingenuity of the Househusband',
  "Workin' Moms",
  "Creator's File: GOLD",
  'Why Are You Like This',
  'Masameer County']
```

```
recommending_per_watch_history("Lupin")
```

```
[ 'Lupin', 'Ganglands', 'Bangkok Breaking', 'Dealer', 'Fauda']
```

```
recommending_per_watch_history("Anne with an E")
```

```
[ 'The Prince of Tennis ~ Match! Tennis Juniors ~',
  'Candy Online',
  'Anne with an E',
  'Baby',
  'Love 101']
```

# REFERENCES

Kaggle.com. 2021. *Netflix Movies and TV Shows*. [online] Available at: <<https://www.kaggle.com/shivamb/netflix-shows>> [Accessed 25 November 2021].

# Thank You

NETFLIX