1. Introduction

This project leverages the skills of Python, SQL and Power BI, using real data of Netflix. Netflix, the global giant in media and video streaming, captivates over 200 million subscribers worldwide with its diverse library of over 8000 movies and TV shows. In this age of digital entertainment, understanding the intricate dynamics of Netflix's content offering is paramount. My "Netflix Power BI Dashboard" delves deep into this ever-expanding universe of entertainment, providing valuable insights and illuminating patterns that can steer content strategies and viewer engagement.

2. Project Objectives

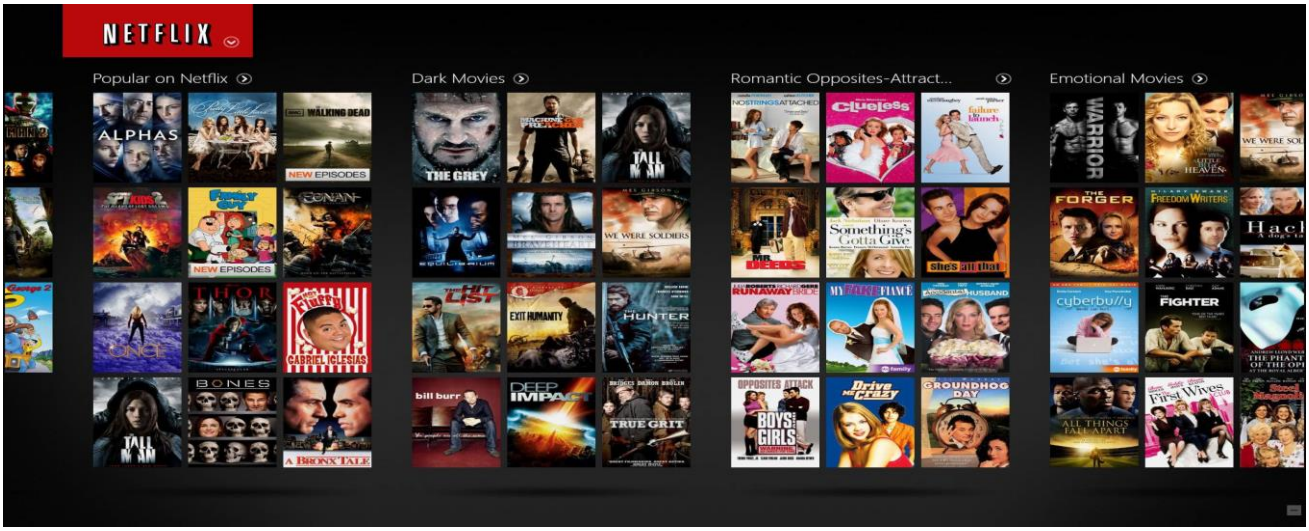
This Power BI project is structured around several key tasks designed to unlock the secrets of Netflix's content ecosystem:

Identifying Genres by Titles
Identifying Ratings by Show ID
Identifying Movies and TV Shows by Release Years
Total Number of Movies and TV Shows
Identifying Top 10 Countries by Movies and TV Shows
Dashboard
Visualization Tools
Tool used: Power BI
Data set used: Netflix Dataset

This Power BI project leverages a rich palette of visualizations, including bar charts, area charts, tree maps, and donut charts. These visuals are carefully crafted to make complex data easily accessible and visually engaging, enhancing the user experience.

3. Conclusion

In a world where data-driven decisions are paramount, our Netflix Power BI Dashboard Project empowers content creators, marketers, and decision-makers in the entertainment industry to navigate the Netflix universe effectively. Explore the nuances of content genres, audience preferences, and international contributions while unveiling the hidden gems within Netflix's content repository. Welcome to the era of data-driven entertainment strategy, brought to life through the power of Power BI.



IMPORTING DATA

```
import numpy as np # linear algebra
import pandas as pd # data processing
import matplotlib.pyplot as plt
import seaborn as sns
# Input data files are available in the "../input/" directory.
import os
for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in filenames:
        print(os.path.join(dirname, filename))

# You can write up to 20GB to the current directory (/kaggle/working/) that gets preserved as output when you create a version using "Save & Run All"
# You can also write temporary files to /kaggle/temp/, but they won't be saved outside of the current session. Data: /kaggle/input/netflix-shows/netflix_titles.csv
```

LOADING DATA

```
netflix=pd.read_csv('../input/netflix-shows/netflix_titles.csv')
```

First 5 Values

```
netflix.head()
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, filmm...
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lor...
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	1 Season	Docuseries, Reality TV	Feuds, flirtations and toilet talk go down amo...
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, Romantic TV Shows, TV ...	In a city of coaching centers known to train I...

Last 5 Value

```
netflix.tail()
```

Describe The Data

```
netflix.describe()
```

	release_year
count	8807.000000
mean	2014.180198
std	8.819312
min	1925.000000
25%	2013.000000
50%	2017.000000
75%	2019.000000
max	2021.000000

Information About The Data

```
netflix.info()
```

```
<class 'pandas.core.frame.DataFrame'>

RangeIndex: 8807 entries, 0 to 8806
```

```
Data columns (total 12 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   show_id       8807 non-null   object
 1   type          8807 non-null   object
 2   title         8807 non-null   object
 3   director      6173 non-null   object
 4   cast          7982 non-null   object
 5   country       7976 non-null   object
 6   date_added    8797 non-null   object
 7   release_year  8807 non-null   int64
 8   rating        8803 non-null   object
 9   duration      8804 non-null   object
 10  listed_in     8807 non-null   object
 11  description   8807 non-null   object
dtypes: int64(1), object(11)
memory usage: 825.8+ KB
```

Shape of The Data

```
netflix.shape
```

```
(8807, 12)
```

Name of the Column

```
netflix.columns
```

```
Index(['show_id', 'type', 'title', 'director', 'cast', 'country', 'date_added',
'release_year',          'rating',          'duration',          'listed_in',
'description'],dtype='object')
```

DATA CLEANING

Checking the Null values

```
netflix.isnull().sum()
```

```
show_id            0
type               0
title              0
director        2634
cast             825
country          831
date_added        10
release_year       0
rating             4
duration           3
listed_in          0
description        0
dtype: int64
```

Director has 2634 Null Values ,Cast 825 values , Country 831...

Check Unique Values

```
netflix.nunique()
```

```
show_id         8807
type               2
title           8807
director        4528
cast            7692
country          748
date_added      1767
release_year      74
rating            17
duration         220
listed_in        514
description     8775
dtype: int64
```

Check Duplicated Values

```
netflix.duplicated().sum()
```

```
0
```

Make a copy of Dataset

```
df = netflix.copy()


df.shape
```

```
(8807, 12)
```

Drop null values

```
df=df.dropna()
df.shape
```

```
(5332, 12)
```

Convert Date Time format

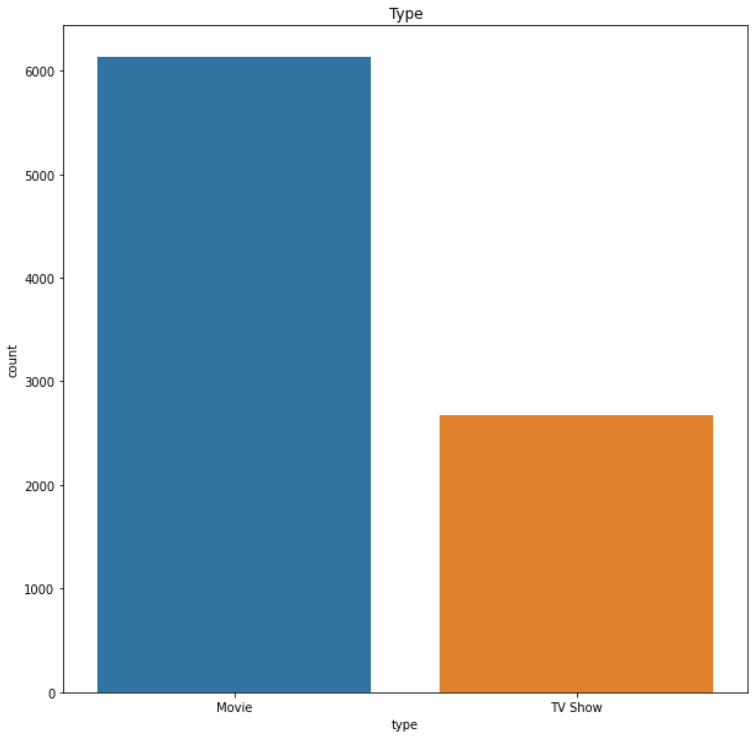
```
df["date_added"] = pd.to_datetime(df['date_added'])
df['day_added'] = df['date_added'].dt.day
df['year_added'] = df['date_added'].dt.year
df['month_added']=df['date_added'].dt.month
df['year_added'].astype(int);
df['day_added'].astype(int);


df.head(10)
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	day_added	year_added	month_added
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra D...	Faso, United Kin...	2021-09-24	1993	TV-MA	125 min	Movies, International Movies	On a photo shoot in Ghana, an American model s...	24	2021	9
8	s9	TV Show	Baking Show	Andy Devonshire	Mel Giedroyc, Sue Perkins, Mary Berry, Paul Ho...	United Kingdom	2021-09-24	2021	TV-14	9 Seasons	British TV Shows, Reality TV	A talented batch of amateur bakers face off in...	24	2021	9
9	s10	Movie	The Starling	Theodore Melfi	Melissa McCarthy, Chris O'Dowd, Kevin Kline, T...	United States	2021-09-24	2021	PG-13	104 min	Comedies, Dramas	A woman adjusting to life after a loss contend...	24	2021	9
12	s13	Movie	Je Suis Karl	Christian Schwochow	Luna Wedler, Jannis Niewöhner, Milan Peschel, ...	Czech Republic	2021-09-23	2021	TV-MA	127 min	International Movies	After most of her family is murdered in a terr...	23	2021	9
24	s25	Movie	Jeans	S. Shankar	Prashanth, Aishwarya Rai Bachchan, Sri Lakshmi...	India	2021-09-21	1998	TV-14	166 min	International Movies, Romantic Movies	When the father of the man she loves insists t...	21	2021	9
27	s28	Movie	Grown Ups	Dennis Dugan	Adam Sandler, Kevin James, Chris Rock, David S...	United States	2021-09-20	2010	PG-13	103 min	Comedies	Mourning the loss of their beloved junior high...	20	2021	9
28	s29	Movie	Dark Skies	Scott Stewart	Keri Russell, Josh Hamilton, J.K. Simmons, Dak...	United States	2021-09-19	2013	PG-13	97 min	Horror Movies, Sci-Fi & Fantasy	A family's idyllic suburban life shatters when...	19	2021	9
29	s30	Movie	Paranoia	Robert Luketic	Liam Hemsworth, Gary Oldman, Amber Heard, Harr...	States, India, France	2021-09-19	2013	PG-13	106 min	Thrillers	Blackmailed by his company's CEO, a low-level ...	19	2021	9
38	s39	Movie	Birth of the Dragon	George Nolfi	Billy Magnussen, Ron Yuan, Qu Jingjing, Terry ...	United States	2021-09-16	2017	PG-13	96 min	Action & Adventure, Dramas	A young Bruce Lee angers kung fu traditionalis...	16	2021	9
41	s42	Movie	Jaws	Steven Spielberg	Roy Scheider, Robert Shaw, Richard Dreyfuss, L...	United States	2021-09-16	1975	PG	124 min	Classic Movies, Dramas	When an insatiable great white shark terrorize...	16	2021	9

DATA VISUALIZATION

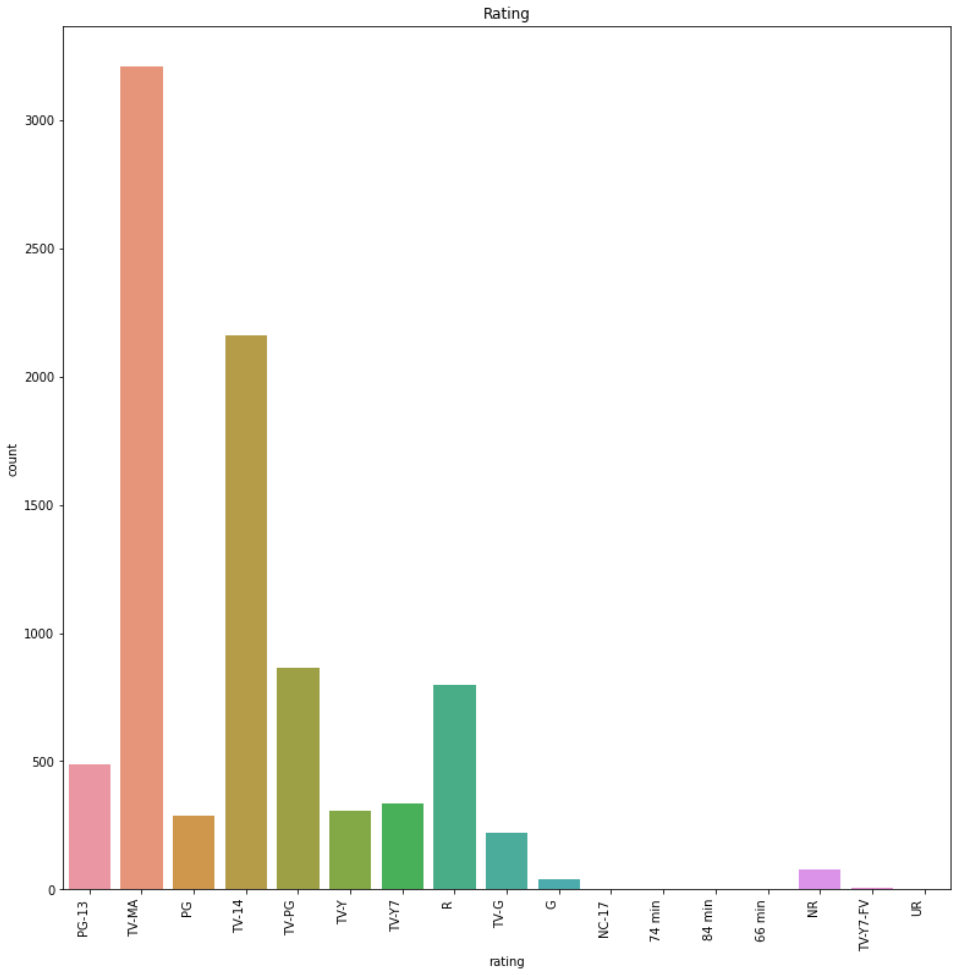
Type: Movie and TV Shows

```
sns.countplot(netflix['type'])
fig = plt.gcf()
fig.set_size_inches(10,10)
plt.title('Type')
```



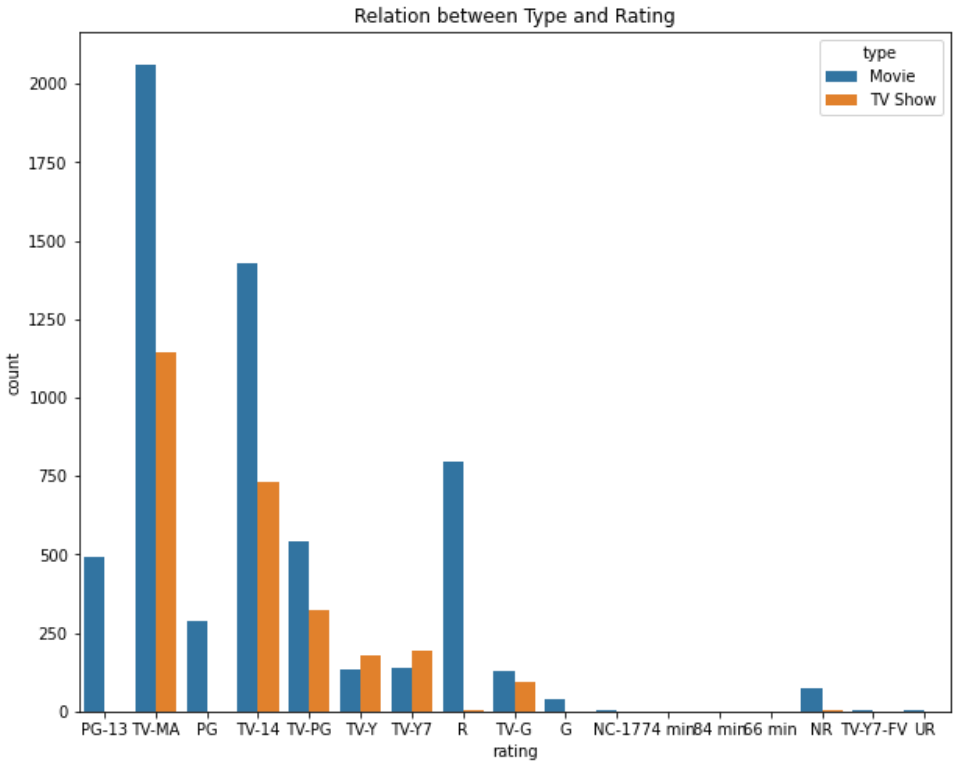
Rating of shows and movies

```
sns.countplot(netflix['rating'])
sns.countplot(netflix['rating']).set_xticklabels(sns.countplot(netflix['rati
ng']).get_xticklabels(), rotation=90, ha="right")
fig = plt.gcf()
fig.set_size_inches(13,13)
plt.title('Rating')
```



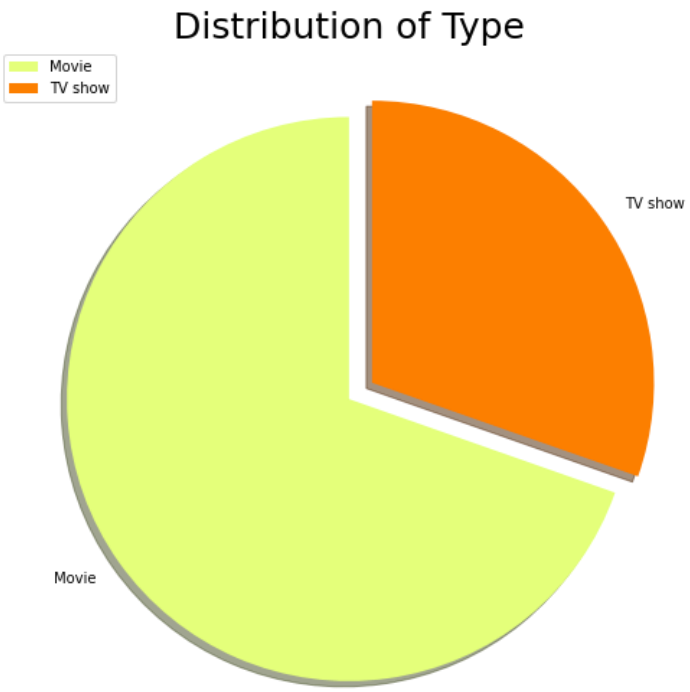
Relation Between Type and Rating

```
plt.figure(figsize=(10,8))
sns.countplot(x='rating',hue='type',data=netflix)
plt.title('Relation between Type and Rating')
plt.show()
```



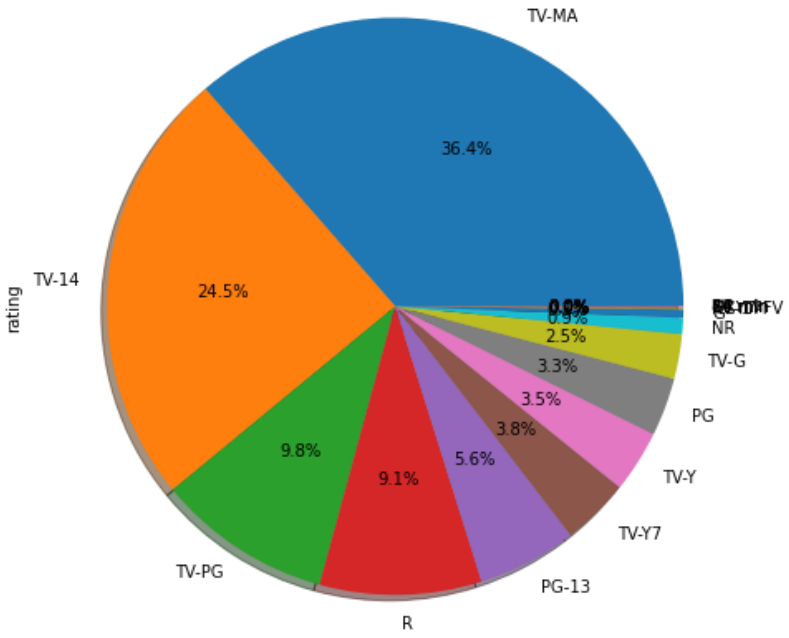
Pie-chart for the Type: Movie and TV Shows

```
labels = ['Movie', 'TV show']
size = netflix['type'].value_counts()
colors = plt.cm.Wistia(np.linspace(0, 1, 2))
explode = [0, 0.1]
plt.rcParams['figure.figsize'] = (9, 9)
plt.pie(size,labels=labels, colors = colors, explode = explode, shadow = True,
startangle = 90)
plt.title('Distribution of Type', fontsize = 25)
plt.legend()
plt.show()
```



Pie-chart for Rating

```
netflix['rating'].value_counts().plot.pie(autopct='%1.1f%%',shadow=True,figsize=(10,8))
plt.show()
```

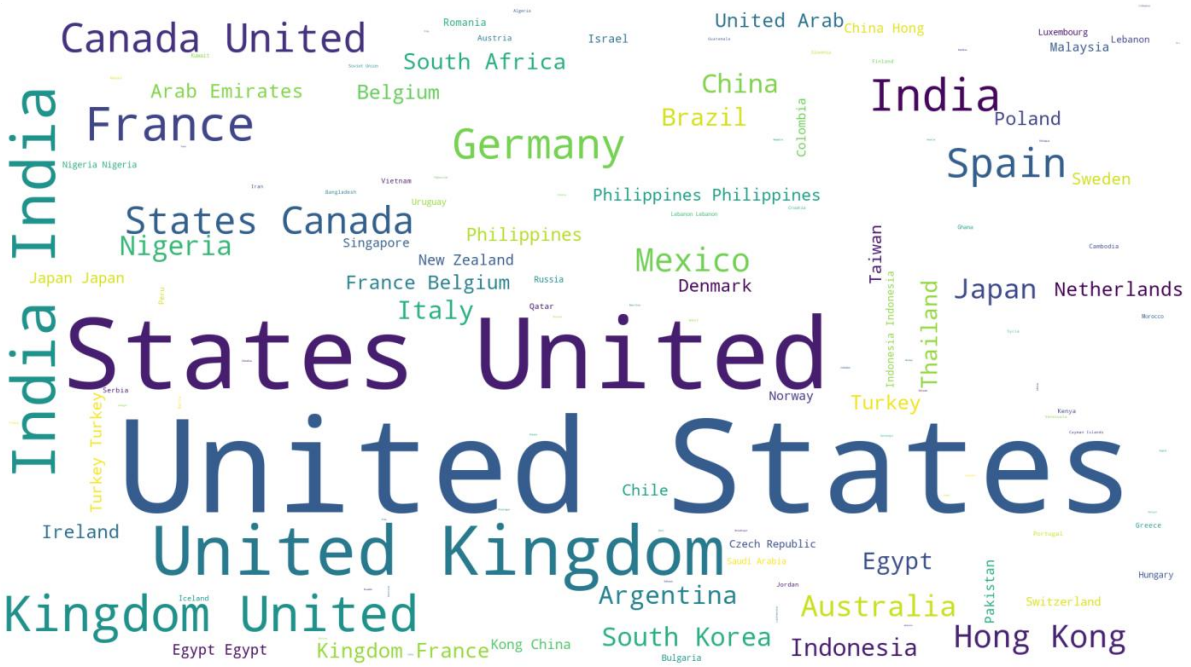


WORDCLOUD

```
from wordcloud import WordCloud
```

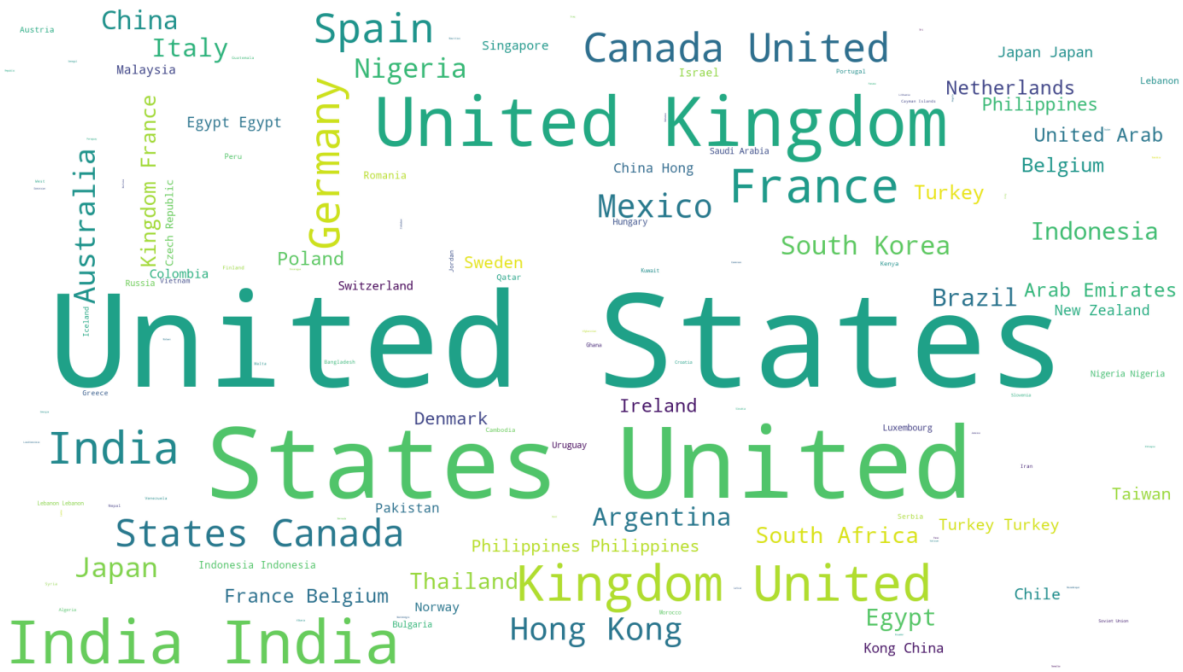
Country

```
plt.subplots(figsize=(25,15))
wordcloud = WordCloud(
                          background_color='white',
                          width=1920,
                          height=1080
                         ).generate(" ".join(df.country))
plt.imshow(wordcloud)
plt.axis('off')
plt.savefig('country.png')
plt.show()
```



Cast in the Show

```
plt.subplots(figsize=(25,15))
wordcloud = WordCloud(
                          background_color='white',
                          width=1920,
                          height=1080
                         ).generate(" ".join(df.country))
plt.imshow(wordcloud)
plt.axis('off')
plt.savefig('country.png')
plt.show()
```



Directors

```
plt.subplots(figsize=(25,15))
wordcloud = WordCloud(
                          background_color='white',
                          width=1920,
                          height=1080
                         ).generate(" ".join(df.director))
plt.imshow(wordcloud)
plt.axis('off')
plt.savefig('director.png')
plt.show()
```

Categories

```
plt.subplots(figsize=(25,15))
wordcloud = WordCloud(
                          background_color='white',
                          width=1920,
                          height=1080
                         ).generate(" ".join(df.listed_in))
plt.imshow(wordcloud)
plt.axis('off')
plt.savefig('category.png')
plt.show()
```