

Ex no 6
blalzs

Exploratory Data Analysis with Python.

Aim:

To analyze the distribution of netflix genre, and to identify trends in content addition over time below

Program code:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

df = pd.read_csv("netflix_titles.csv")
print(df.info())
print(df.head())
print(df.describe(include="all"))

print("number of unique directors:",
      df["director"].nunique())
print(df["type"].value_counts())
print(df["release_year"].value_counts(),
      head())
print(df.groupby(['country', 'type']).size().sort_values(),
      ascending=False)
```


Shot on OnePlus

[df['date_added'] = pd.to_datetime(df[

Output:

[`df`] < class: Dataframe
[`df`] < class: DataFrame

(`df`) < 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	web_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	data-added	8797	non-null	object
7	release-year	8807	non-null	object
8	rating	8803	non-null	object
9	duration	8804	non-null	object

No. of unique countries: 748

Shot on OnePlus

No. of unique directors: 4528

```
df.dropna(subset=['date-added'], inplace=True)
df.set_index('date-added', inplace=True)
monthly_content = df.groupby('M').size()
plt.figure(figsize=(12, 6))
monthly_content.plot()
plt.title('netflix. Content added over time')
plt.xlabel('Date')
plt.grid(True)
plt.show()

plt.figure(figsize=(10, 5))
sns.histplot(df['release_year'], bins=30)
plt.title("Distribution of release years")
plt.xlabel("Release year")
plt.ylabel("Number of titles")
plt.show()

top_genres = genres.value_counts().head(10)
top_genres.plot(kind='barh', color='red')
plt.title("Top 10 Genre on netflix")
plt.xlabel("Count")
plt.show()
```



Name : cont, d type : int64

Country type

United States movie: 2058

India movie 893

United States TV Show 760

United Kingdom TV Show 760

Spain Movie 97

Egypt Movie 92

97 & 1990s begin

Netflix Content added over time

(Over time Netflix added 1990s, 1980s, 1970s)

(Stop off begin)

(Stop off begin)

((1990s begin))

stop off begin

"backward vision forward" begin

(Japan, Brazil off begin)

(Japan) - Japan begin

backward vision forward begin

(Brazil) - Brazil begin

((Brazil)) vision 3rd world begin

autumn 1st world begin



Shot on OnePlus

Result:

The analysis reveals trends in Netflix's content library, highlighting the growth of content over time, top countries, popular genres, and the balance between movies and TV shows.



Shot on OnePlus