Output :

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

RangeIndex = 8807 entries, 0 to 8806

Data columns (Total 12 columns)

#	Column	Non-Null Count	Dtype
0	type	8807	Non-Null
1	title	8807	Non-Null
2	director	6173	Non-Null
3	cast	7982	Non-Null
4	country	7976	Non-null
5	date_added	8797	Non-null
6	release_year	8807	Non-null
7	rating	8803	Non-null
8	duration	8804	Non-Null

Show_id	Type	Title	Director
0	S1	Johnson	Kristen Johnson
1	S2	Blood & Water	NaN
2	S3	Cranglands	Julien Leclercq
3	S4	Jailbirds	NaN

EXP. NO : 6

9/9/25

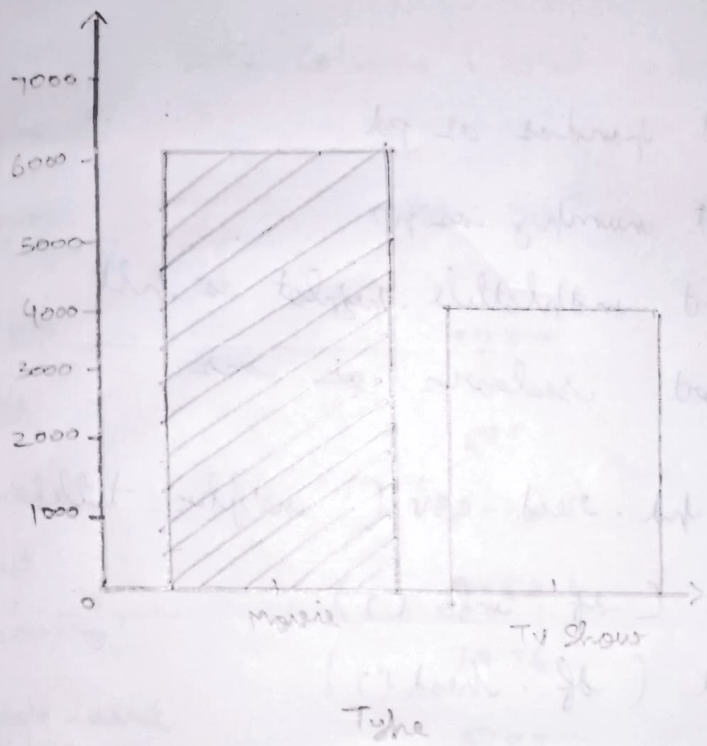
Exploratory Data Analysis with Python

Aim : By using EDA & neccessary libraries to perform the neccessary modules.

Program :

```
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(f"{df.describe(include='all')}")
print("Number of unique countries :", df['country'].nunique())

print("Number of unique directors :" df['director'].nunique())

print(df['type'].value_count())

df['date_added'] = pd.to_datetime(df['date_added'],
                                  format='mixed', errors='coerce')

df.set_index('date_added', inplace=True)
monthly_content = df.resample('M').size()
plt.figure(figsize=(0.6))
```

Count of Movie & TV Show:



```
monthly_content.plot()
plt.xlabel("Date")
plt.ylabel("Number of titles added")
plt.grid(True)
plt.show()

# content type count
sns.countplot(data=df, x='type', palette='set')
plt.title("Count of Movies & TV Shows")
plt.show()

# Top 10 countries with most count
top_countries = df['country'].value_count().head(10)
top_countries.plot(kind='bar', color='skyblue')
plt.title("Top 10 countries by titles")

# Genere Frequency:
genres = df['listed_in'].str.split(',', expand=True)
top_genres = genres.value_count().head(10)
top_genres = plot(kind="bauh", color='coral')
plt.title = ("Top 10 Genres of Netflix")
plt.show()
```

Result:

Thus, the required programming for EDA has been executed successfully.