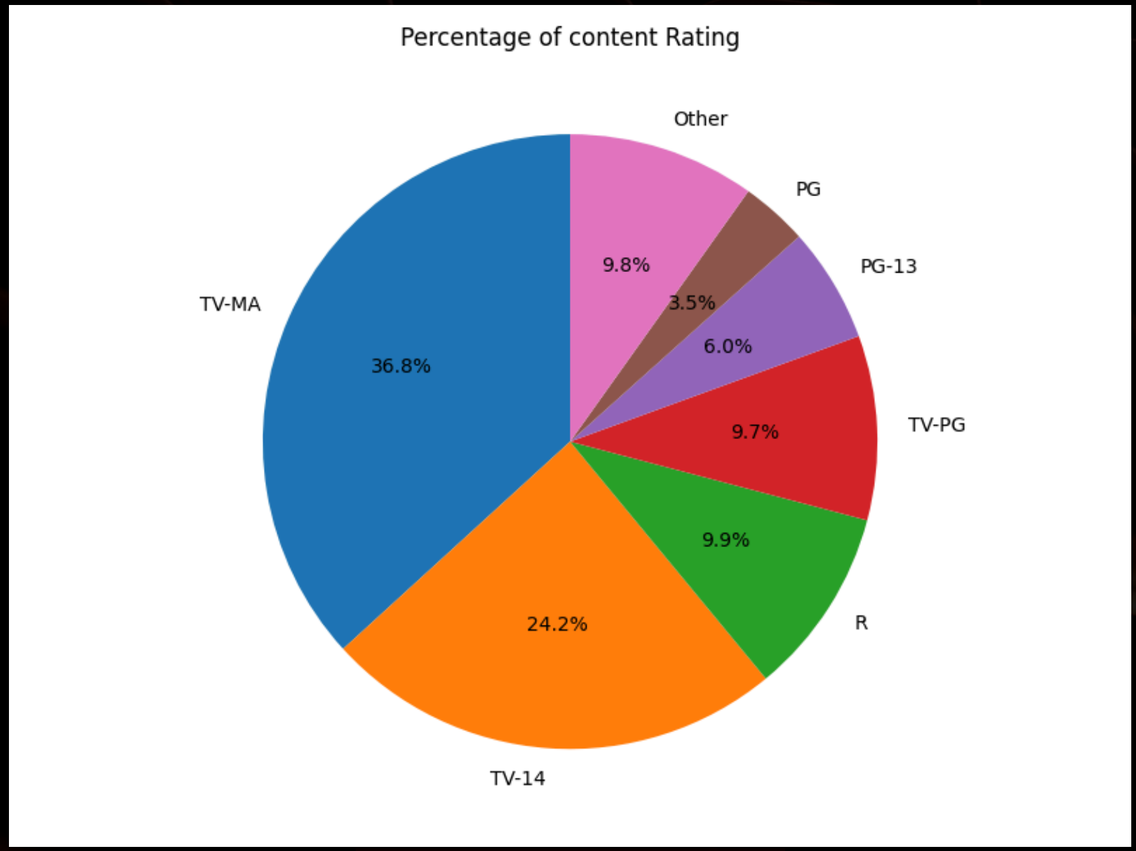
Netflix Data Visualization

```python
import pandas as pd
import matplotlib.pyplot as plt
df=pd.read_csv('netflix_titles.csv')
df.columns
df=df.dropna(subset=['type','release_year','rating','country','description','duration'])

# Count ratings
rating_counts = df['rating'].value_counts()

# Show only top 6 ratings, rest as "Other"
top_ratings = rating_counts[:6]
other_sum = rating_counts[6:].sum()
rating_labels = list(top_ratings.index) + ['Other']
rating_values = list(top_ratings.values) + [other_sum]

# PIE_CHART:- Content_Rating
plt.figure(figsize=(8,6))
plt.pie(rating_values, labels=rating_labels,autopct='%1.1f%%',startangle=90)
plt.title('Percentage of content Rating')
plt.tight_layout()
plt.savefig('content_rating_PieChart.png')
plt.show()
```

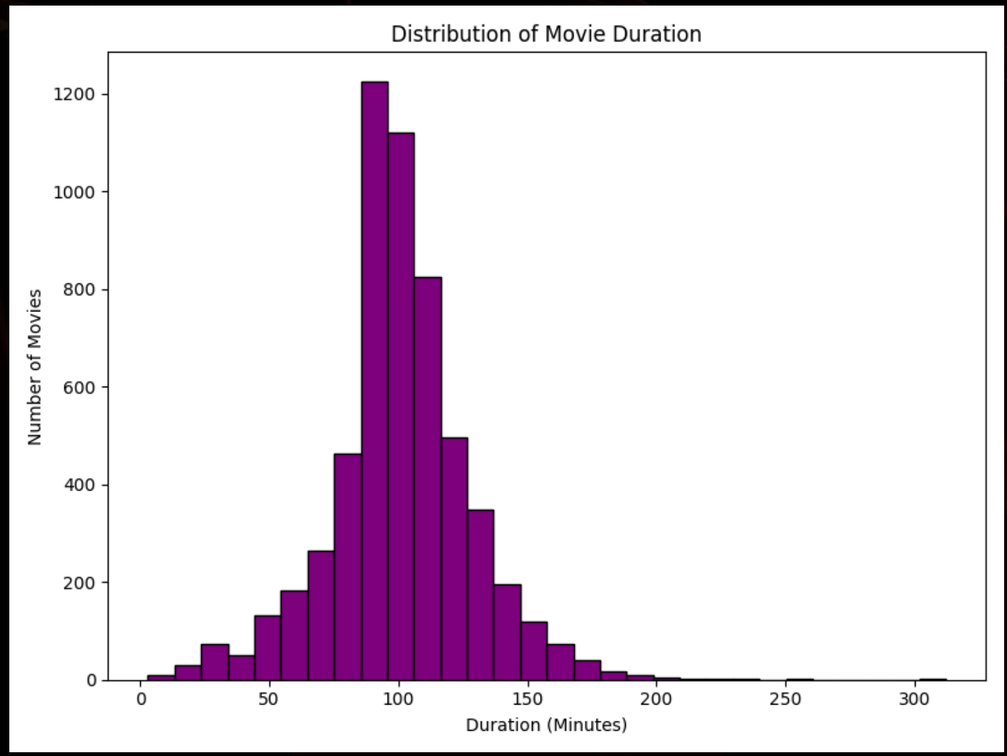


PIE CHART

Netflix Data Visualization

```
import pandas as pd
import matplotlib.pyplot as plt
df=pd.read_csv('netflix_titles.csv')
df.columns
df=df.dropna(subset=['type','release_year','rating','country','description','duration'])

# HISTOGRAM:- Movie_Duration_Histogram
movie_df=df[df['type']=='Movie'].copy()
movie_df['duration_int'] = movie_df['duration'].str.replace('min','').astype(int)
plt.figure(figsize=(8,6))
plt.hist(movie_df['duration_int'],bins=30,color='purple',edgecolor='black')
plt.title('Distribution of Movie Duration')
plt.xlabel('Duration (Minutes)')
plt.ylabel('Number of Movies')
plt.tight_layout()
plt.savefig('movie_duration_histogram.png')
plt.show()
```

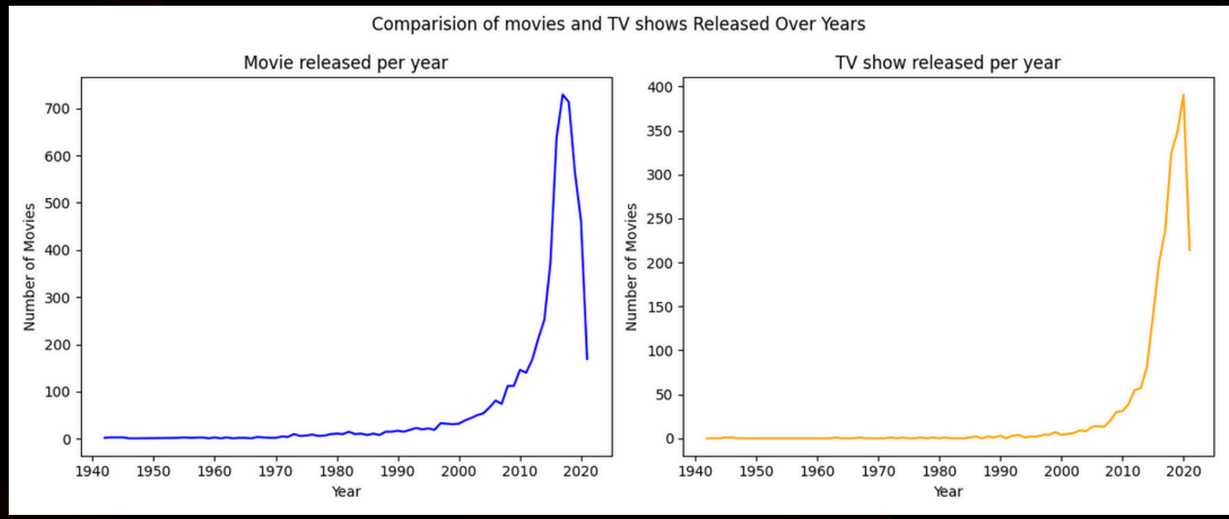


HISTOGRAM

Netflix Data Visualization

```python
import pandas as pd
import matplotlib.pyplot as plt
df=pd.read_csv('netflix_titles.csv')
df.columns
df=df.dropna(subset=['type','release_year','rating','country','description','duration'])

# SUBPLOT:- Movies_Tv_Show_Comparision
content_by_year=df.groupby(['release_year','type']).size().unstack().fillna(0)
fig,ax=plt.subplots(1,2,figsize=(12,5))
#first subplot:movies
ax[0].plot(content_by_year.index,content_by_year['Movie'],color='blue')
ax[0].set_title('Movie released per year')
ax[0].set_xlabel('Year')
ax[0].set_ylabel('Number of Movies')
# second subplot:TV show
ax[1].plot(content_by_year.index,content_by_year['TV Show'],color='orange')
ax[1].set_title('TV show released per year')
ax[1].set_xlabel('Year')
ax[1].set_ylabel('Number of Movies')
fig.suptitle('Comparision of movies and TV shows Released Over Years')
plt.tight_layout()
plt.savefig('movies_tv_show_comparision.png')
plt.show()
```

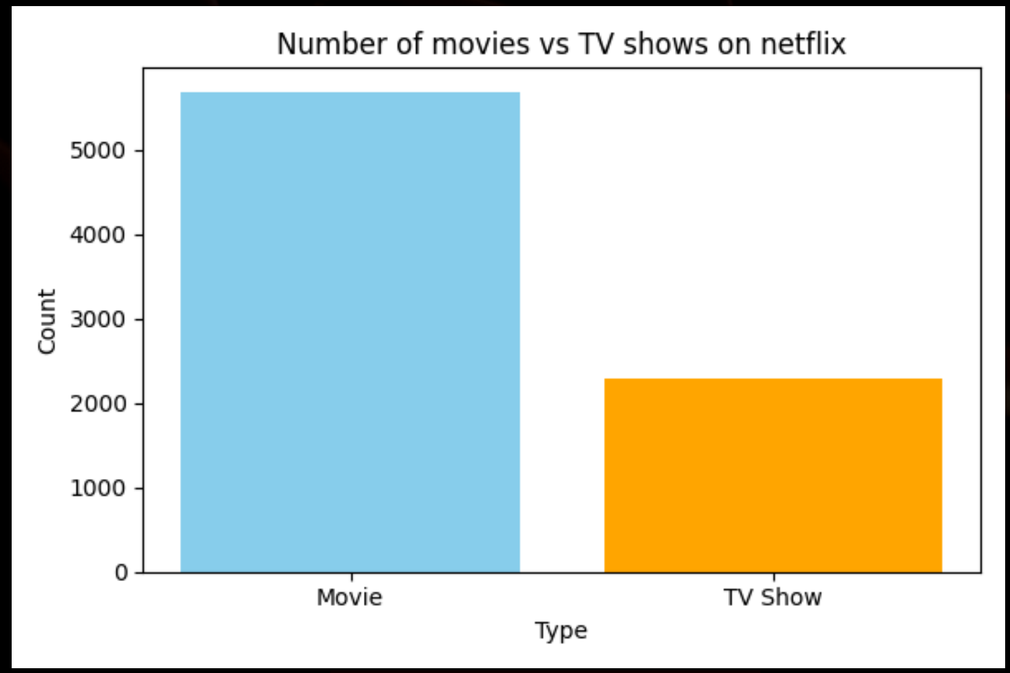


SUBPLOT

Netflix Data Visualization

```python
import pandas as pd
import matplotlib.pyplot as plt
df=pd.read_csv('netflix_titles.csv')
df.columns
df=df.dropna(subset=['type','release_year','rating','country','description','duration'])

# BAR_CHART:- Movies_VS_Tvshows
type_count=df['type'].value_counts()
plt.figure(figsize=(6,4))
plt.bar(type_count.index,type_count.values,color=['skyblue','orange'])
plt.title('Number of movies vs TV shows on netflix')
plt.xlabel('Type')
plt.ylabel('Count')
plt.tight_layout()
plt.savefig('movies_vs_tvshows.png')
plt.show()
```

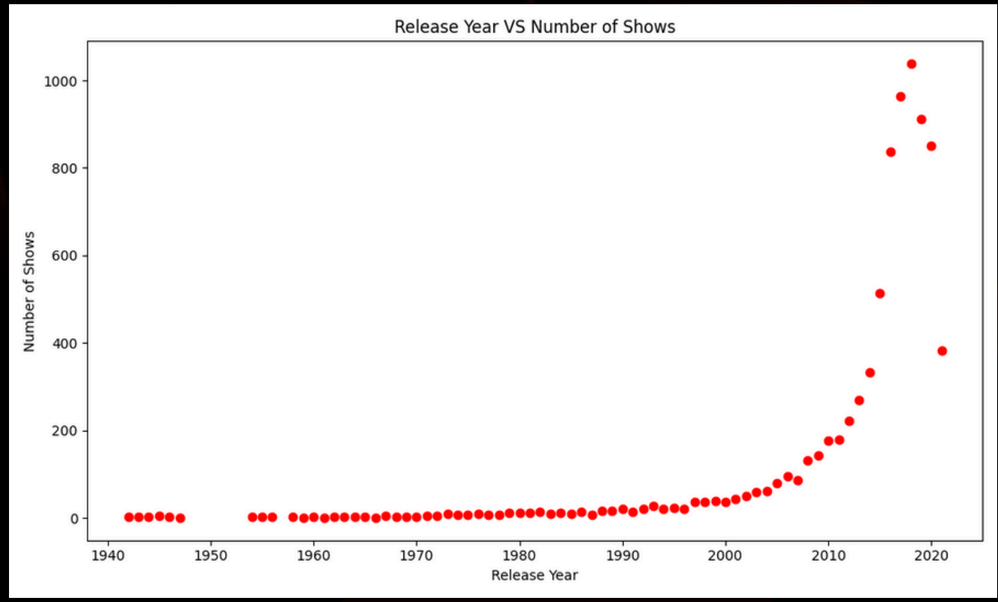


BAR CHART

Netflix Data Visualization

```python
import pandas as pd
import matplotlib.pyplot as plt
df=pd.read_csv('netflix_titles.csv')
df.columns
df=df.dropna(subset=['type','release_year','rating','country','description','duration'])

# SCATTER_PLOT:- Release_Year_Scatter
release_counts = df['release_year'].value_counts().sort_index()
plt.figure(figsize=(10,6))
plt.scatter(release_counts.index,release_counts.values,color='red')
plt.title('Release Year VS Number of Shows')
plt.xlabel('Release Year')
plt.ylabel('Number of Shows')
plt.tight_layout()
plt.savefig('release_year_scatter.png')
plt.show()
```

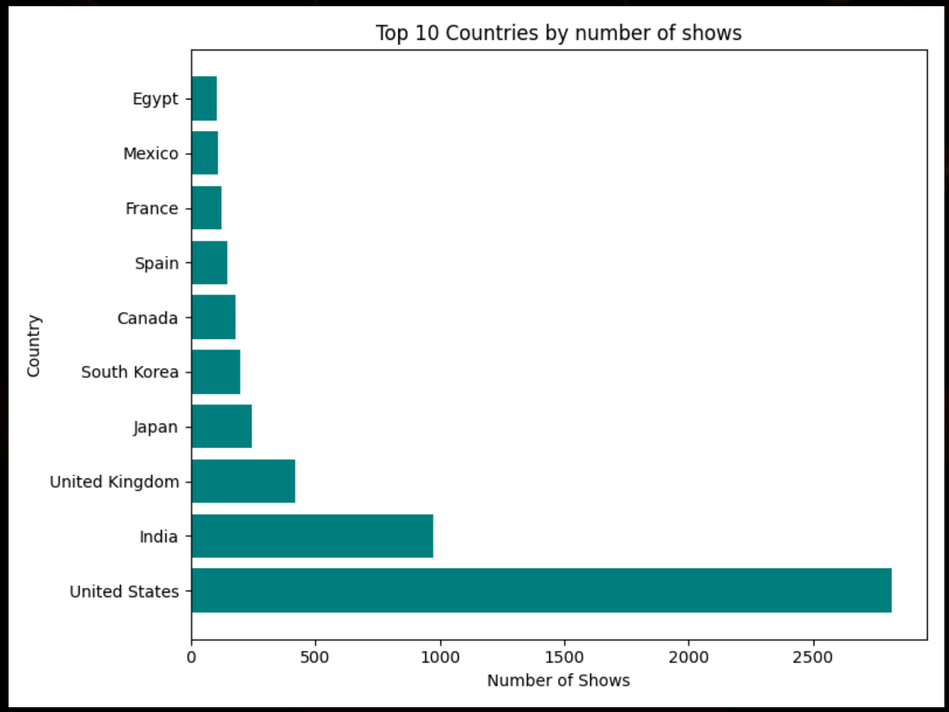


SCATTER PLOT

Netflix Data Visualization

```python
import pandas as pd
import matplotlib.pyplot as plt
df=pd.read_csv('netflix_titles.csv')
df.columns
df=df.dropna(subset=['type','release_year','rating','country','description','duration'])

# HORIZONTAL_BAR_CHART:- Top10_Countries
country_counts=df['country'].value_counts().head(10)
plt.figure(figsize=(8,6))
plt.barh(country_counts.index,country_counts.values,color='teal')
plt.title('Top 10 Countries by number of shows')
plt.xlabel('Number of Shows')
plt.ylabel('Country')
plt.tight_layout()
plt.savefig('top10_countries.png')
plt.show()
```



HORIZONTAL BAR CHART

Thank *You*

Gmail:- shivamk66297@gmail.com