

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

```
df=pd.read_csv('netflix1.csv')
```

df

	show_id	type	title	director	country	date_added	release_year	rating	duration	listed_in
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	9/25/2021	2020	PG-13	90 min	Documentaries
1	s3	TV Show	Ganglands	Julien Leclercq	France	9/24/2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...
2	s6	TV Show	Midnight Mass	Mike Flanagan	United States	9/24/2021	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mysteries
3	s14	Movie	Confessions of an Invisible Girl	Bruno Garotti	Brazil	9/22/2021	2021	TV-PG	91 min	Children & Family Movies, Comedies
4	s8	Movie	Sankofa	Haile Gerima	United States	9/24/2021	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies
...
8785	s8797	TV Show	Yunus Emre	Not Given	Turkey	1/17/2017	2016	TV-PG	2 Seasons	International TV Shows, TV Dramas
8786	s8798	TV Show	Zak Storm	Not Given	United States	9/13/2018	2016	TV-Y7	3 Seasons	Kids' TV
8787	s8801	TV Show	Zindagi Gulzar Hai	Not Given	Pakistan	12/15/2016	2012	TV-PG	1 Season	International TV Shows, Romantic TV Shows, TV ...
8788	s8784	TV Show	Yoko	Not Given	Pakistan	6/23/2018	2016	TV-Y	1 Season	Kids' TV
8789	s8796	TV	YOM	Not Given	Pakistan	6/7/2018	2016	TV-Y7	1 Season	Kids' TV

Next steps: [Generate code with df](#) [New interactive sheet](#)

```
df.groupby('country')['country'].count().sort_values(ascending=False)
```

country	
country	
United States	3240
India	1057
United Kingdom	638
Pakistan	421
Not Given	287
...	...
Slovenia	1
Puerto Rico	1
Somalia	1
West Germany	1
Zimbabwe	1

86 rows × 1 columns

dtype: int64

```
df['type'].value_counts()
```

```

count
type
Movie    6126
TV Show  2664

dtype: int64

```

```
df.loc[df['release_year'] == 2020]
```

	show_id	type	title	director	country	date_added	release_year	rating	duration	listed_in	
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	9/25/2021	2020	PG-13	90 min	Documentaries	
25	s17	Movie	Europe's Most Dangerous Man: Otto Skorzeny in ...	Pedro de Echave García, Pablo Azorín Williams	Not Given	9/22/2021	2020	TV-MA	67 min	Documentaries, International Movies	
36	s35	TV Show	Tayo and Little Wizards	Not Given	Pakistan	9/17/2021	2020	TV-Y7	1 Season	Kids' TV	
47	s190	TV Show	Bread Barbershop	Not Given	Pakistan	8/28/2021	2020	TV-Y	2 Seasons	Kids' TV, TV Comedies	
58	s48	TV Show	The Smart Money Woman	Bunmi Ajakaiye	South Africa	9/16/2021	2020	TV-MA	1 Season	International TV Shows, Romantic TV Shows, TV ...	
...	
7484	s3288	TV Show	Maradona in Mexico	Not Given	Argentina	11/13/2019	2020	TV-MA	1 Season	Docuseries, Spanish-Language TV Shows	
7517	s3370	TV Show	BoJack Horseman	Not Given	United States	10/25/2019	2020	TV-MA	6 Seasons	TV Comedies	
7537	s3434	TV Show	The Hook Up Plan	Not Given	France	10/11/2019	2020	TV-MA	2 Seasons	International TV Shows, Romantic TV Shows, TV ...	
8685	s8126	TV Show	Super Wings	Not Given	United States	12/1/2020	2020	TV-Y	3 Seasons	Kids' TV, Korean TV Shows	
8687	s8133	TV Show	Surviving R. Kelly Part II:	Not Given	United States	4/13/2020	2020	TV-MA	1 Season	Crime TV Shows, Docuseries	

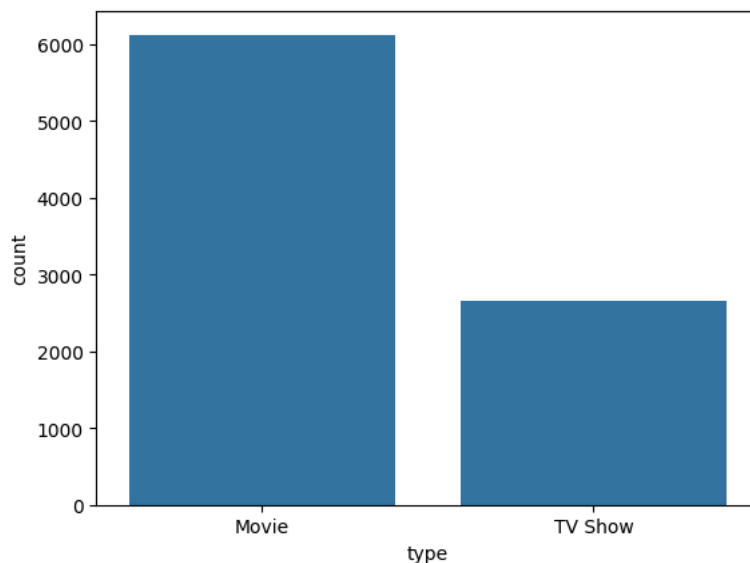
```
df.iloc[100:100]
```

```
show_id type title director country date_added release_year rating duration listed_in
```

```
df.query('release_year == 2020')
```

	show_id	type	title	director	country	date_added	release_year	rating	duration	listed_in
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	9/25/2021	2020	PG-13	90 min	Documentaries
25	s17	Movie	Europe's Most Dangerous Man: Otto Skorzeny in ...	Pedro de Alcn, Pablo Azorn Williams	Not Given	9/22/2021	2020	TV-MA	67 min	Documentaries, International Movies
36	s35	TV Show	Tayo and Little Wizards	Not Given	Pakistan	9/17/2021	2020	TV-Y7	1 Season	Kids' TV
47	s190	TV Show	Bread Barbershop	Not Given	Pakistan	8/28/2021	2020	TV-Y	2 Seasons	Kids' TV, TV Comedies
58	s48	TV Show	The Smart Money Woman	Bunmi Ajakaiye	South Africa	9/16/2021	2020	TV-MA	1 Season	International TV Shows, Romantic TV Shows, TV ...
...
7484	s3288	TV Show	Maradona in Mexico	Not Given	Argentina	11/13/2019	2020	TV-MA	1 Season	Docuseries, Spanish-Language TV Shows
7517	s3370	TV Show	BoJack Horseman	Not Given	United States	10/25/2019	2020	TV-MA	6 Seasons	TV Comedies
7537	s3434	TV Show	The Hook Up Plan	Not Given	France	10/11/2019	2020	TV-MA	2 Seasons	International TV Shows, Romantic TV Shows, TV ...
8685	s8126	TV Show	Super Wings	Not Given	United States	12/1/2020	2020	TV-Y	3 Seasons	Kids' TV, Korean TV Shows
8687	s8133	TV Show	Surviving R. Kelly Part II:	Not Given	United States	4/13/2020	2020	TV-MA	1 Season	Crime TV Shows, Docuseries

```
sns.countplot(x='type',data=df)
plt.show()
```



```
top10 = df['country'].value_counts().nlargest(10)
display(top10)
```

country	count
United States	3240
India	1057
United Kingdom	638
Pakistan	421
Not Given	287
Canada	271
Japan	259
South Korea	214
France	213
Spain	182

dtype: int64

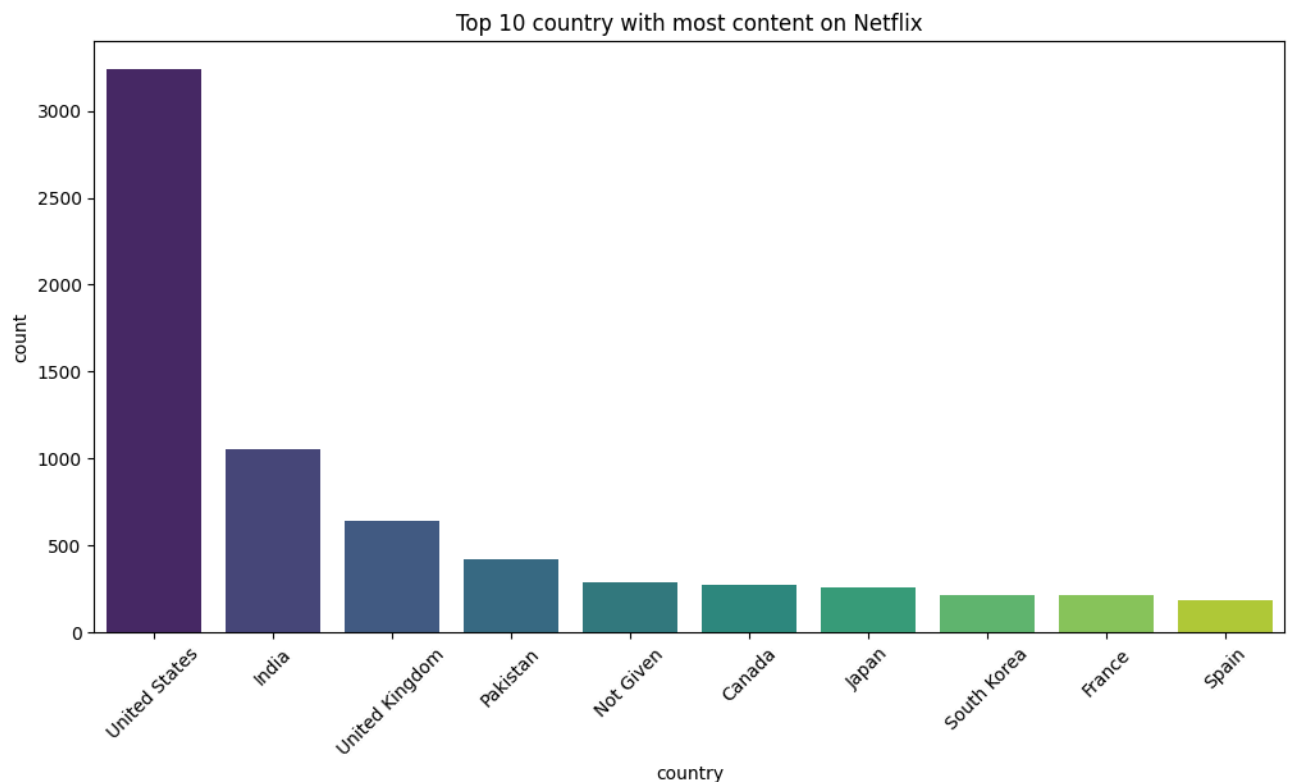
```
plt.figure(figsize=(12,6))
sns.barplot(x=top10.index, y=top10.values, palette="viridis")

plt.xlabel('country')
plt.ylabel("count")
plt.title("Top 10 country with most content on Netflix")
plt.xticks(rotation=45)
plt.show()
```

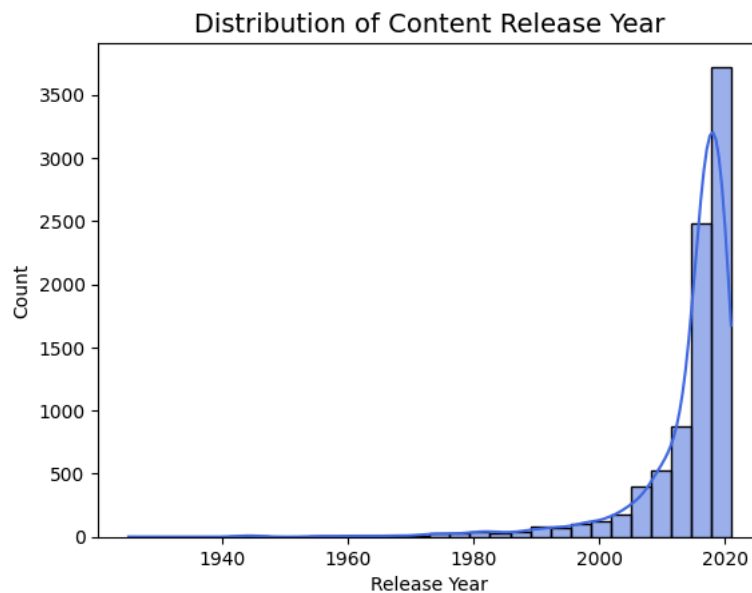
/tmp/ipython-input-2361458008.py:2: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `order` to the original order of the categories.

```
sns.barplot(x=top10.index, y=top10.values, palette="viridis")
```



```
sns.histplot(df["release_year"], bins=30, kde=True, color="royalblue")
plt.title("Distribution of Content Release Year", fontsize=14)
plt.xlabel("Release Year")
plt.ylabel("Count")
plt.show()
```

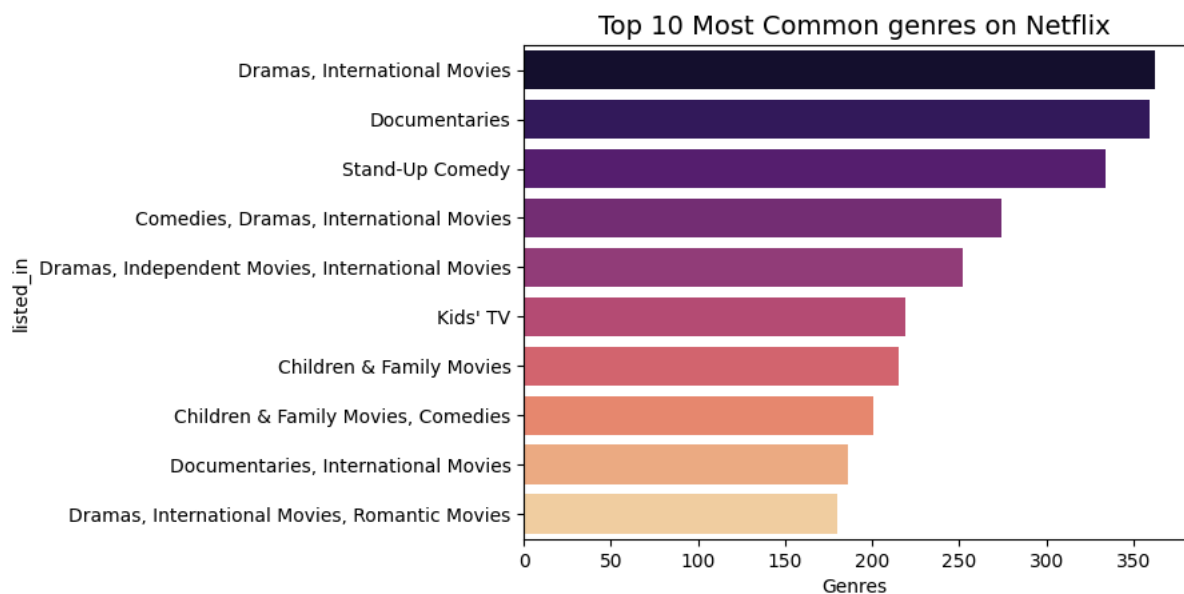


```
top_genres = df["listed_in"].value_counts().head(10)
sns.barplot(x=top_genres.values, y=top_genres.index, palette="magma")
plt.title("Top 10 Most Common genres on Netflix", fontsize=14)
plt.xlabel("Genres")
plt.show()
```

/tmp/ipython-input-1605797719.py:2: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `y` variable to `hue` and set

```
sns.barplot(x=top_genres.values, y=top_genres.index, palette="magma")
```



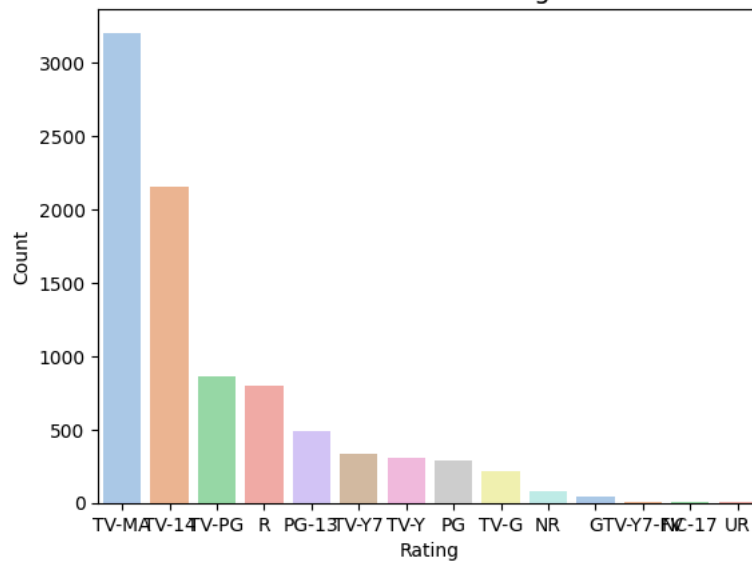
```
rating_counts = df["rating"].value_counts()
sns.barplot(x=rating_counts.index, y=rating_counts.values, palette="pastel")
plt.title("Distribution of Content Ratings on Netflix", fontsize=14)
plt.xlabel("Rating")
plt.ylabel("Count")
plt.show()
```

```
/tmp/ipython-input-2033981661.py:2: FutureWarning:
```

```
Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set
```

```
sns.barplot(x=rating_counts.index, y=rating_counts.values, palette="pastel")
```

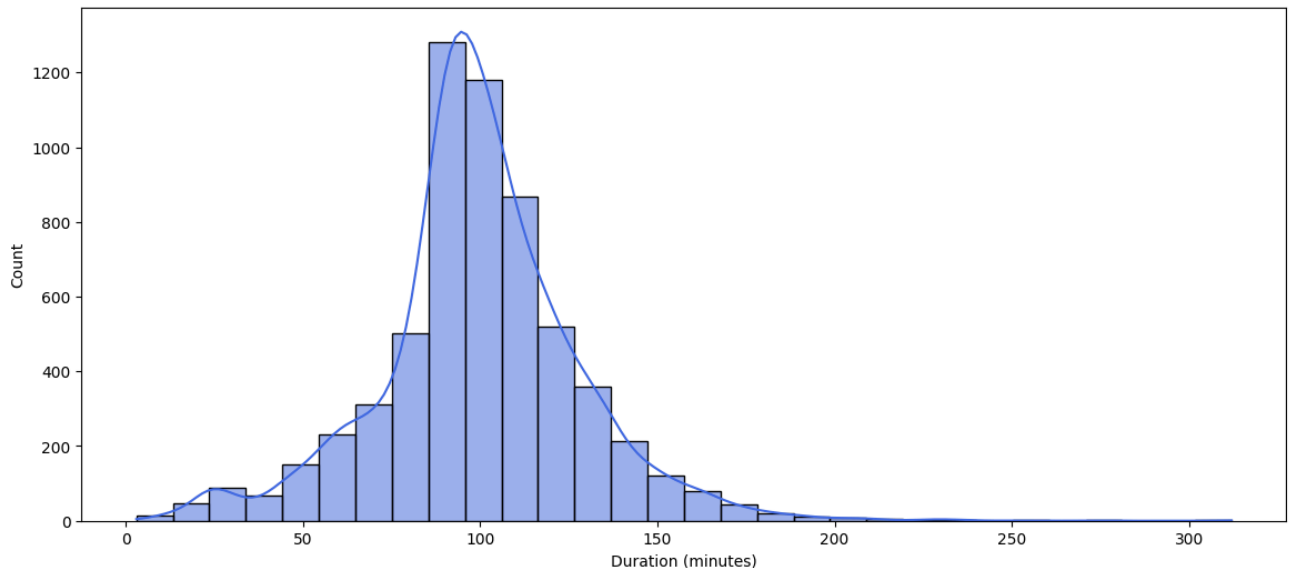
Distribution of Content Ratings on Netflix



```
movies_df = df[df["type"] == "Movie"].copy()
movies_df["duration"] = movies_df["duration"].str.replace(" min", "").astype(float)
```

```
plt.figure(figsize=(14, 6))
sns.histplot(movies_df["duration"], bins=30, kde=True, color="royalblue")
plt.title("Distribution of Movie Durations on Netflix", fontsize=14)
plt.xlabel("Duration (minutes)")
plt.ylabel("Count")
plt.show()
```

Distribution of Movie Durations on Netflix



```
plt.figure(figsize=(14, 6))
sns.boxplot(data=movies_df, x="rating", y="duration", palette="coolwarm")
plt.title("Distribution of Movie Durations Across Different Ratings", fontsize=14)
plt.xlabel("Rating")
plt.ylabel("Duration (minutes)")
plt.xticks(rotation=45)
plt.show()
```

```
/tmp/ipython-input-2404038288.py:2: FutureWarning:
```

```
Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set
```

```
sns.boxplot(data=movies_df, x="rating", y="duration", palette="coolwarm")
```

