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

```
data = pd.read_csv('/content/netflix1.csv')
```

```
data
```

	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

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8790 entries, 0 to 8789
Data columns (total 10 columns):
 #   Column        Non-Null Count  Dtype 
---  ------        --------------  ----- 
 0   show_id       8790 non-null   object
 1   type          8790 non-null   object
 2   title         8790 non-null   object
 3   director      8790 non-null   object
 4   country       8790 non-null   object
 5   date_added    8790 non-null   object
 6   release_year  8790 non-null   int64 
 7   rating        8790 non-null   object
 8   duration      8790 non-null   object
 9   listed_in     8790 non-null   object
dtypes: int64(1), object(9)
memory usage: 686.8+ KB
```

```
data.shape
```

```
(8790, 10)
```

```
data.isnull().sum()/len(data)*100
```

	0
show_id	0.0
type	0.0
title	0.0
director	0.0
country	0.0
date_added	0.0
release_year	0.0
rating	0.0
duration	0.0
listed_in	0.0

dtype: float64

```
data.rating.unique()
```

```
array(['PG-13', 'TV-MA', 'TV-PG', 'TV-14', 'TV-Y7', 'TV-Y', 'PG', 'TV-G',
       'R', 'G', 'NC-17', 'NR', 'TV-Y7-FV', 'UR'], dtype=object)
```

```
data.duplicated().sum()
```

```
np.int64(0)
```

```
data.tail()
```

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

```
data['show_id'] = data['show_id'].str.replace('s','')
```

```
data.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

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

type	count
Movie	6126
TV Show	2664

dtype: int64

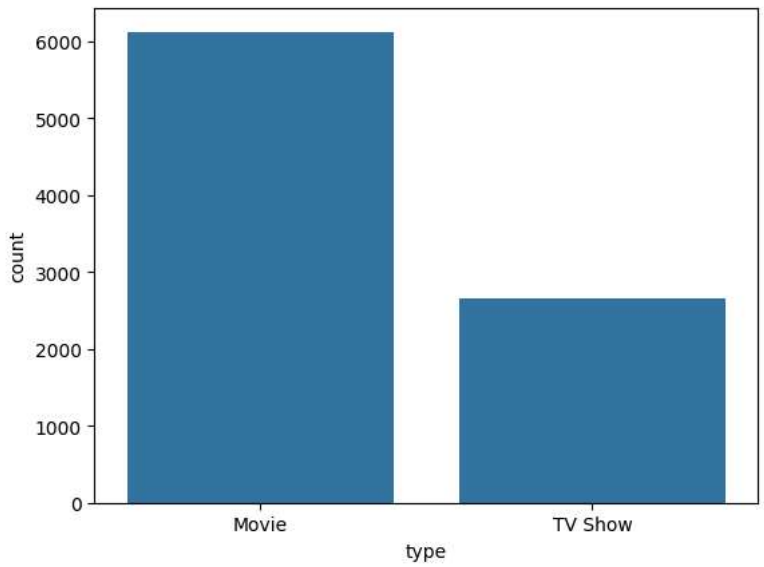
```
data.loc[data['release_year'] == 2000]
```

	show_id	type	title	director	country	date_added	release_year	rating	duration	listed_in

```
data.iloc[100:110]
```

	show_id	type	title	director	country	date_added	release_year	rating	duration	listed_in
494	351	Movie	Space Cowboys	Clint Eastwood	United States	8/1/2021	2000	PG-13	130 min	Action & Adventure, Dramas, Sci-Fi &
100	370	TV Show	Myth & Mogul: John DeLorean	Not Given	Pakistan	7/30/2021	2021	TV-14	1 Season	British TV Shows, Crime TV Shows, Docuseries
101	377	TV Show	Transformers: War for Cybertron Kingdom	Not Given	Pakistan	7/29/2021	2021	TV-Y7	1 Season	Action & Adventure, Anime Series
717	567	Movie	Charlie's Angels	McG	United States	7/1/2021	2000	PG-13	98 min	Comedies
102	380	TV Show	Tattoo Redo	Not Given	Pakistan	7/28/2021	2021	TV-MA	1 Season	Reality TV
103	382	TV Show	The Snitch Cartel: Battlefield Earth	Not Given	Pakistan	7/28/2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, Spanis...
859	780	Movie	Origins	Roger Christian	United States	6/2/2021	2000	PG-13	118 min	Action & Adventure, Cult Movies, Sci-Fi & Fantasy
104	398	TV Show	Feels Like Ishq	Not Given	Pakistan	7/23/2021	2021	TV-MA	1 Season	International TV Shows, Romantic TV Shows, TV ...
2644	3473	Movie	Rugrats in Paris: The Movie	Stig Bergqvist, Paul Demeyer	Germany	10/1/2019	2000	G	79 min	Children & Family Movies, Comedies
105	483	TV Show	How to Become a Tyrant	Not Given	Pakistan	7/9/2021	2021	TV-MA	1 Season	Docuseries
106	82	Movie	Kate	Cedric Nicolas-Troyan	United States	9/10/2021	2021	R	106 min	Action & Adventure
3420	4724	Movie	Fiza	JJC Skillz, Mohamed Hunke	India	8/2/2018	2000	TV-14	163 min	Dramas, International Movies, Music & Musicals
107	85	Movie	Omo Ghetto: the Saga	Khalid Akindele	Nigeria	9/10/2021	2020	TV-MA	147 min	Action & Adventure, Comedies, Dramas
3585	4957	Movie	Phir Bhi Dil Hai	Aziz Mirza	India	4/1/2018	2000	TV-14	159 min	Comedies, Dramas,

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



			Dreams		States					Faith & Spir...
							2000	PG-13	104 min	Comedies, Dramas
							2000	R	102 min	Comedies, Cult Movies, Dramas
							2000	G	79 min	Children & Family Movies
							2000	R	110 min	Dramas, Independent Movies, International Movies
							2000	PG-13	116 min	Dramas
							2000	TV-14	132 min	Comedies, International Movies, Romantic Movies
							2000	PG-13	120 min	Action & Adventure, International Movies, Roma...
			Dragonheart: A New		United					Action & Adventure,

```
top_10 = data['country'].value_counts().nlargest(10)
top_10
```

					States					Movies, Thrillers
5018	6913	Movie	Hamara Dil Aapke Paas Hai	Satish Kaushik	India	3/1/2018	2000	TV-14	158 min	Dramas, International Movies, Music & Musicals
5097	7014	Movie	How the Grinch Stole Christmas	Ron Howard	United States	6/1/2017	2000	PG	105 min	Children & Family Movies, Comedies
5283	7248	Movie	Kya Kehna	Kundan Shah	India	4/1/2018	2000	TV-PG	149 min	Dramas, International Movies, Romantic Movies
5339	7317	Movie	Little Nicky	Steve Brill	United States	12/1/2020	2000	PG-13	90 min	Comedies
5345	7328	Movie	Ramaiya Vastavaiya	Bhanwar	India	12/2/2017	2000	TV-PG	137 min	Comedies, Dramas,

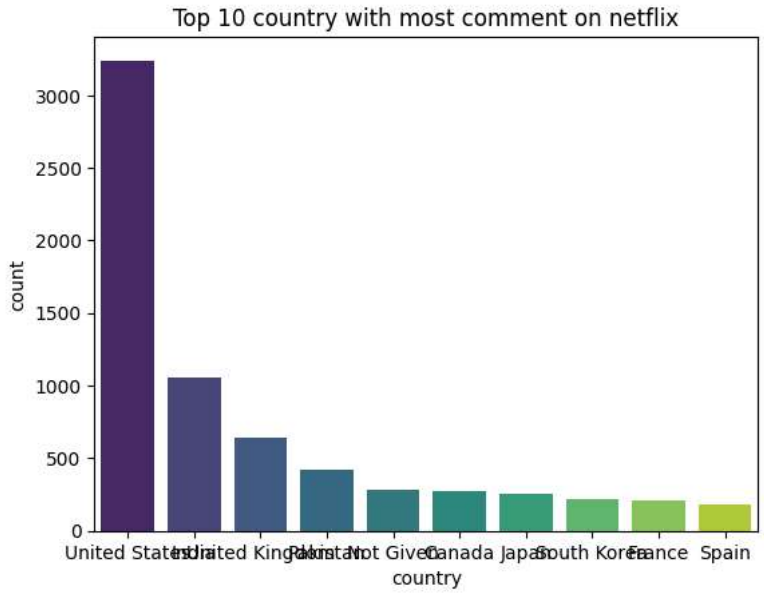
3645 7705 Movie Papa the Great Bhagyaraj India 12/8/2017 2000 TV-PG 137 min International Movies

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

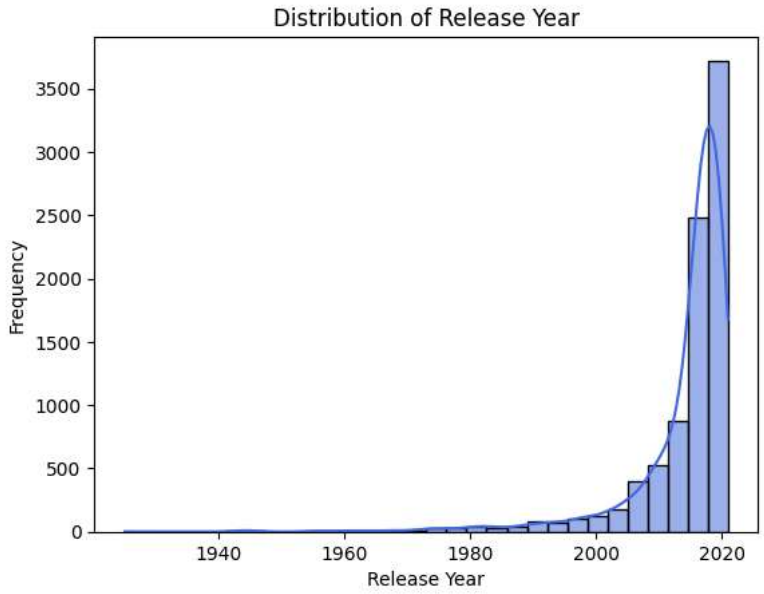
dtype: int64

```
sns.barplot(x = top_10.index,y = top_10.values,palette ="viridis")
plt.show

plt.xlabel('country')
plt.ylabel('count')
plt.title("Top 10 country with most comment on netflix")
plt.show()
```

```
/tmp/ipython-input-919513308.py:1: FutureWarning:

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

  sns.barplot(x = top_10.index,y = top_10.values,palette ="viridis")
```

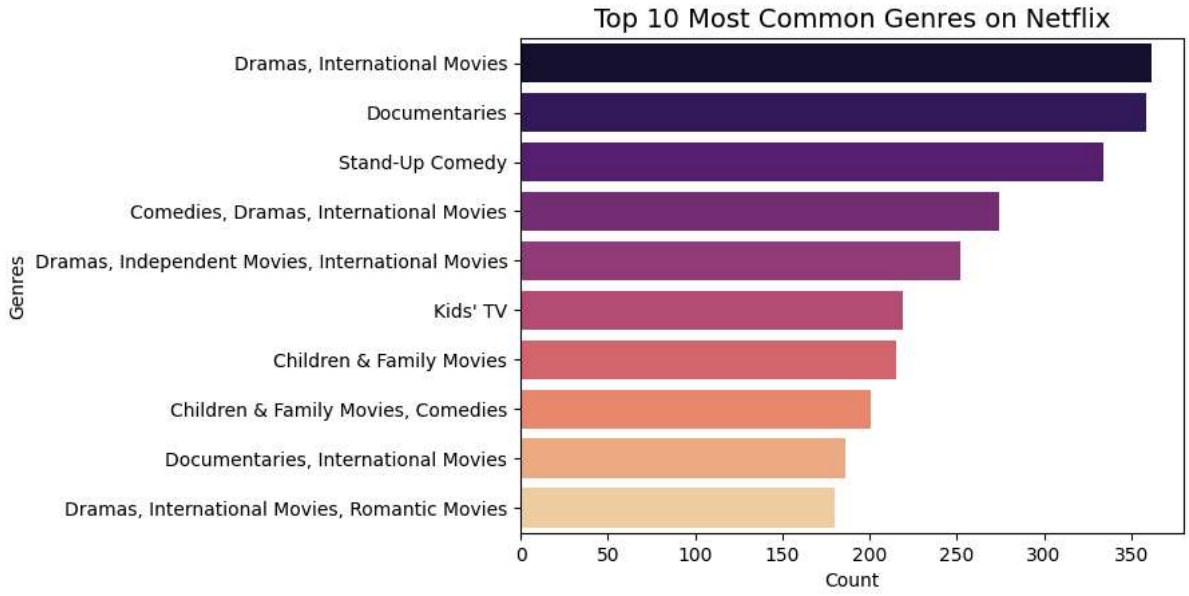



```
sns.histplot(data["release_year"],bins=30,kde=True,color="royalblue")
plt.title("Distribution of Release Year")
plt.xlabel("Release Year")
plt.ylabel("Frequency")
plt.show()
```



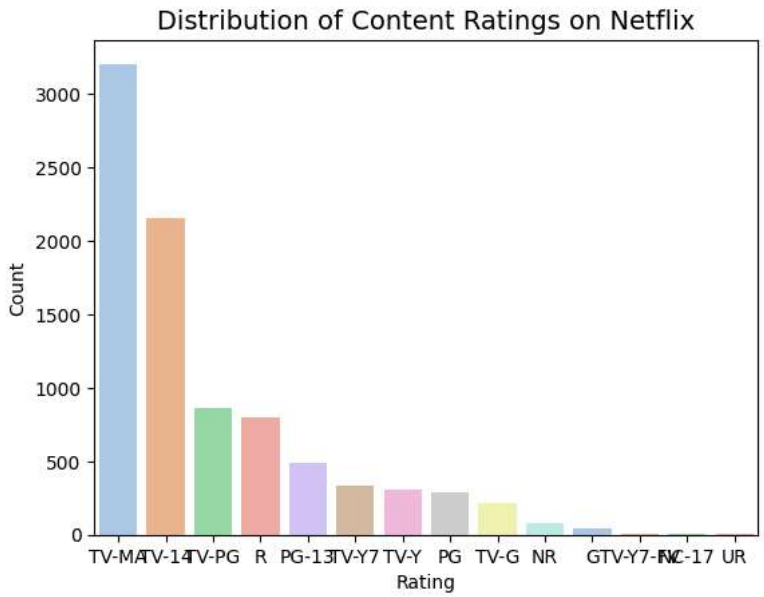

```
top_genres = data[ "listed_in"]. value_counts().head(10)
sns.barplot(x=top_genres.values, y=top_genres. index, palette="magma")
plt.title("Top 10 Most Common Genres on Netflix"
, fontsize=14)
plt.xlabel( "Count")
plt.ylabel ("Genres")
plt.show()
```

```
/tmp/ipython-input-474789517.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 `l

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

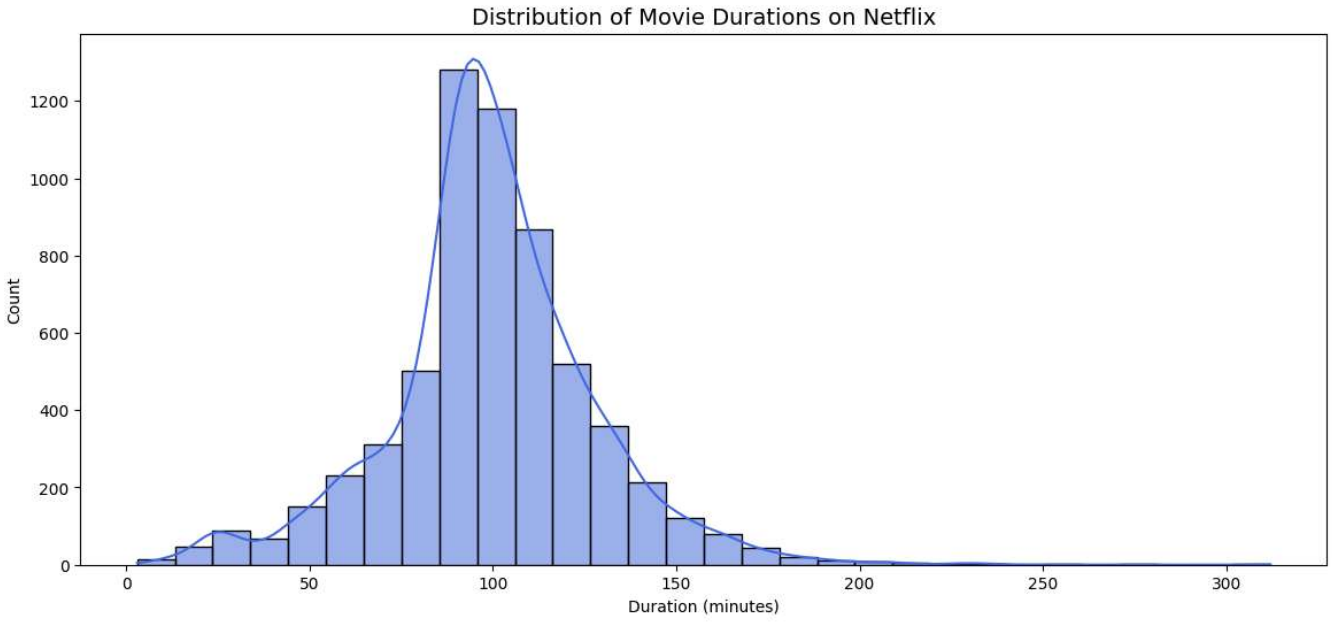



```
rating_counts = data["rating"].value_counts()

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-3591801135.py:3: FutureWarning:

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

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

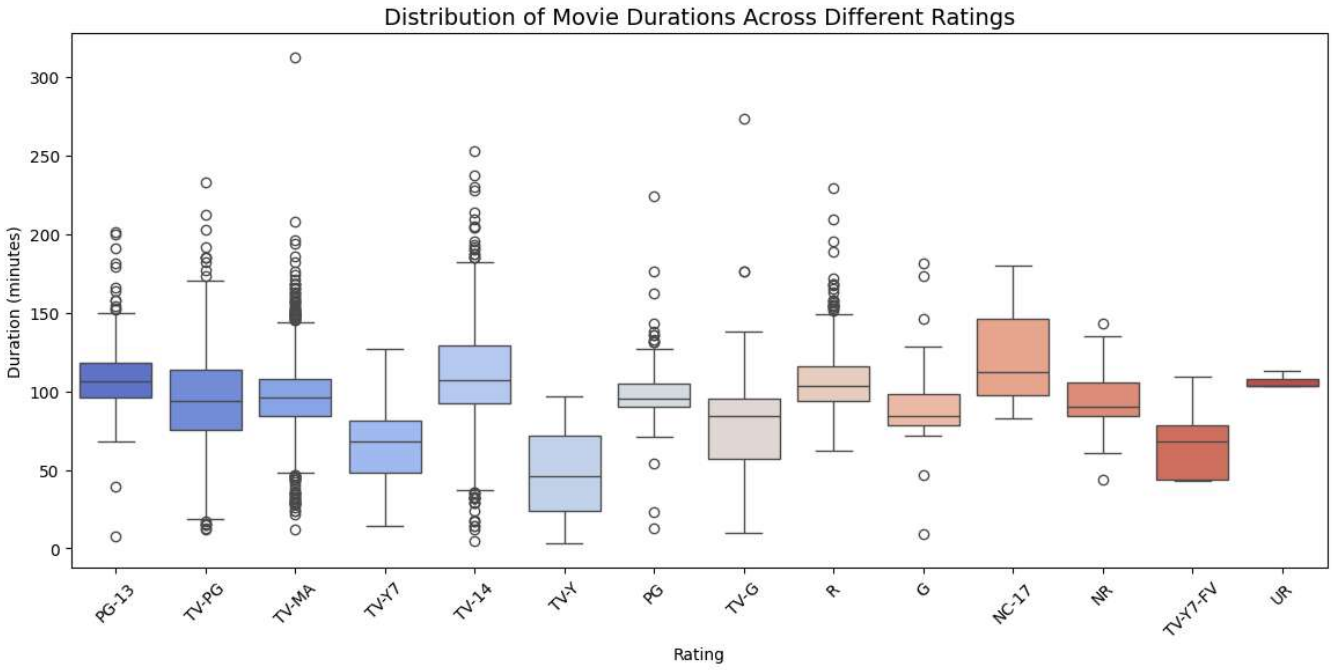



```
movies_df = data[data["type"] == "Movie"].copy()

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()
```



```
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-1700444524.py:3: FutureWarning:

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

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




```
data.columns
```

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

```
from sklearn.preprocessing import LabelEncoder
le=LabelEncoder()
a=['show_id', 'type', 'director', 'title', 'country', 'date_added', 'rating', 'duration', 'listed_in']
for i in a:
 data [i]=le.fit_transform(data[i])
```

```
data.head()
```