Q1. Defining Problem Statement and Analysing basic metrics (10 Points)

Ans. Netflix aims to determine which type of shows/movies to produce and how to grow the business in different countries based on the provided dataset. The company aims to leverage its data to determine the types of shows and movies to produce and to identify strategies for growth in different countries. Specifically, the business problem revolves around analyzing the available content to generate insights on content production and business expansion.

Key Objectives:

1. Identify which type of content (movies or TV shows) should be prioritized for production.
2. Determine the best strategies to grow Netflix's subscriber base in various countries.

Analyzing Basic Metrics which include following steps

Load the Data: Import the dataset and understand its structure.

Shape of the Data: Check the number of rows and columns.

Data Types: Examine the data types of all attributes.

Statistical Summary: Generate a statistical summary for numerical attributes.

In [2]:
```
# Load the data
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [4]:
```
from google.colab import drive
drive.mount('/netflix.csv')
```

```
Mounted at /netflix.csv
```

In [27]:
```
df = pd.read_csv('netflix.csv')
```

In [6]:
```
df.head()
```

Out[6]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duratio
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90 mi
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	Seasor
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA	1 Seaso
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	1 Seaso
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	Seasor

In [7]:
```
# Shape of the Data: Check the number of rows and columns.
df.shape
```

Out[7]:
```
(8807, 12)
```

In [8]:
```
# Data Types: Examine the data types of all attributes.
df.dtypes
```

Out[8]:
```
type            object
title           object
director        object
cast            object
country         object
date_added      object
release_year     int64
rating          object
duration        object
listed_in       object
description     object
dtype: object
```

In [15]:
```
# here date in object format this want to convert to day time format
df['date_added'] = pd.to_datetime(df['date_added'], errors='coerce')
```

In [16]:
```
df.head()
```

Out[16]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duratio
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	2021-09-25	2020	PG-13	90 mi
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	2021-09-24	2021	TV-MA	Seasor
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	2021-09-24	2021	TV-MA	1 Seaso
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	2021-09-24	2021	TV-MA	1 Seaso
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	2021-09-24	2021	TV-MA	Seasor

In [12]:
```
# Statistical Summary: Generate a statistical summary for numerical attributes.
numerical_columns = df.select_dtypes(include=['number']).columns
# Generating statical summery
df[numerical_columns].describe()
```

Out[12]:

	release_year
count	8807.000000
mean	2014.180198
std	8.819312
min	1925.000000
25%	2013.000000
50%	2017.000000
75%	2019.000000
max	2021.000000

In [ ]:
```

```

Q2. Observations on the shape of data, data types of all the attributes, conversion of categorical attributes to 'category' (If required), missing value detection, statistical summary (10 Points)

Ans:

In [17]:
```
# Shape of the Data: Check the number of rows and columns.
df.shape
```

Out[17]:
```
(8807, 12)
```

In [18]:
```
## Data Types: Examine the data types of all attributes.
df.dtypes
```

Out[18]:
```
show_id                 object
type                    object
title                   object
director                object
cast                    object
country                 object
date_added      datetime64[ns]
release_year             int64
rating                  object
duration                object
listed_in               object
description             object
dtype: object
```

In [29]:
```
# Converting categorical attribute to category
categorical_columns = ['type', 'title', 'director', 'cast', 'country', 'rating', 'list
```

In [30]:
```
for i in categorical_columns:
  df[i] =df[i].astype('category')
```

In [31]:
```
df.dtypes
```

Out[31]:
```
show_id          object
type           category
title          category
director       category
cast           category
country        category
date_added       object
release_year      int64
rating         category
duration         object
listed_in      category
description    category
dtype: object
```

In [35]:
```
# Treating missing value
missing_values = df.isnull().sum()
```

In [34]:
```
missing_values = missing_values[missing_values > 0]
```

In [37]:
```
missing_values
```

Out[37]:
```
show_id           0
type              0
title             0
director       2634
cast            825
country         831
date_added       10
release_year      0
rating            4
duration          3
listed_in         0
description       0
dtype: int64
```

In [38]:
```
# to find the % of missing value
missing_percentage = (missing_values / len(df)) * 100
missing_percentage
```

Out[38]:
```
show_id          0.000000
type             0.000000
title            0.000000
director        29.908028
cast             9.367549
country          9.435676
date_added       0.113546
release_year     0.000000
rating           0.045418
duration         0.034064
listed_in        0.000000
description      0.000000
dtype: float64
```

In [ ]:
```
# inferance : here the director percengage is '29' so it want to clean very important,
```

In [41]:
```
# cleaning missing data
# we already converted the "director" to category so first we want to add new category
df['director'] = df['director'].cat.add_categories('unknown')
df['director'].fillna('unknown', inplace= True)
```

In [44]:
```
# cleaning missing cast category data with unknown
df['cast'] = df['cast'].cat.add_categories('unknown')
df['cast'].fillna('unknown', inplace= True)
```

In [45]:
```
# cleaning missing country category data with unknown
df['country'] = df['country'].cat.add_categories('unknown')
df['country'].fillna('unknown', inplace= True)
```

In [47]:
```
# droping isnull() value of date added, rating and duration
drop_columns = ['date_added', 'rating', 'duration']
df.dropna(subset=drop_columns, inplace = True)
```

In [48]:
```
df.head()
```

Out[48]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	durati
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	unknown	United States	September 25, 2021	2020	PG-13	90 r
1	s2	TV Show	Blood & Water	unknown	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	Seas
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	unknown	September 24, 2021	2021	TV-MA	1 Seas
3	s4	TV Show	Jailbirds New Orleans	unknown	unknown	unknown	September 24, 2021	2021	TV-MA	1 Seas
4	s5	TV Show	Kota Factory	unknown	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	Seas

In [51]:
```
missing_values = df.isnull().sum()
missing_values = missing_values[missing_values > 0]
missing_values
```

Out[51]:
```
Series([], dtype: int64)
```

In [ ]:
```
# now data ready to analysis
```

Q3. Non-Graphical Analysis: Value counts and unique attributes (10 Points)

In [55]:
```
# here i taken director, cast and country for analysis
# value count

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

Out[55]:
```
unknown                                  2621
Rajiv Chilaka                              19
Raúl Campos, Jan Suter                     18
Suhas Kadav                                16
Marcus Raboy                               16
                                         ...
Jose Javier Reyes                           1
Joseduardo Giordano, Sergio Goyri Jr.       1
Khaled Youssef                              1
Louis C.K.                                  0
Alessandro Pepe                             0
Name: count, Length: 4529, dtype: int64
```

In [62]:
```
df['director'].unique()
```

Out[62]:
```
['Kirsten Johnson', 'unknown', 'Julien Leclercq', 'Mike Flanagan', 'Robert Cullen, Jo
sé Luis Ucha', ..., 'Mu Chu', 'Chandra Prakash Dwivedi', 'Majid Al Ansari', 'Peter He
witt', 'Mozez Singh']
Length: 4527
Categories (4529, object): ['A. L. Vijay', 'A. Raajdheep', 'A. Salaam', 'A.R. Murugad
oss', ...,
                            'Óskar Thór Axelsson', 'Ömer Faruk Sorak', 'Şenol Sönme
z', 'unknown']
```

In [56]:
```
df['cast'].value_counts()
```

Out[56]:
```
cast
unknown
825
David Attenborough
19
Vatsal Dubey, Julie Tejwani, Rupa Bhimani, Jigna Bhardwaj, Rajesh Kava, Mousam, Swapn
il                                                  14
Samuel West
10
Jeff Dunham
7

...
Anthony Bourdain
0
Flynn Curry, Olivia Deeble, Madison Lu, Oisín O'Leary, Faith Seci, Joshua Sitch, Heid
i Arena                                              0
Leone Frisa, Paolo Vaccarino, Francesco Migliore, Albrecht Weimer, Giulia Dichiaro, A
lessandra Oriti Niosi, Andreas Segeritz              0
Luke Jurevicius, Craig Behenna, Charlotte Hamlyn, Stavroula Mountzouris, Aletheia Bur
ney                                                  0
Marc Maron, Judd Hirsch, Josh Brener, Nora Zehetner, Andy Kindler
0
Name: count, Length: 7693, dtype: int64
```

In [61]:
```
df['cast'].unique()
```

Out[61]:
```
Gotoas, Samuel Jouy, Nab..., 'Mayur More, Jitendra Kumar, Ranjan Raj, Alam ..., 'Kate
Siegel, Zach Gilford, Hamish Linklater, ..., ..., 'Ali Suliman, Saleh Bakri, Yasa, Al
i Al-Jabri,..., 'Mark Ruffalo, Jake Gyllenhaal, Robert Downey ..., 'Jesse Eisenberg,
Woody Harrelson, Emma Stone,..., 'Tim Allen, Courteney Cox, Chevy Chase, Kate M...,
'Vicky Kaushal, Sarah-Jane Dias, Raaghav Chana...]
Length: 7679
Categories (7693, object): [''Najite Dede, Jude Chukwuka, Taiwo Arimoro, O..., '4Minu
te, B1A4, BtoB, ELSIE, EXID, EXO, Got7, ...,
                            '50 Cent, Ryan Phillippe, Bruce Willis, Rory M..., 'A.J.
LoCascio, Sendhil Ramamurthy, Fred Tatas...,
                            ..., 'Şahin Irmak, İrem Sak, Gonca Vuslateri, Emre ...,
                            'Şükrü Özyıldız, Aslı Enver, Şenay Gürler, Baş..., 'Ṣọpẹ́
Dìrísù, Wunmi Mosaku, Matt Smith, Malai...,
                            'unknown']
```

In [59]:
```
df['country'].value_counts()
```

Out[59]:
```
country
United States
2809
India
972
unknown
829
United Kingdom
418
Japan
243

...
Ireland, Canada, Luxembourg, United States, United Kingdom, Philippines, India
1
Ireland, Canada, United Kingdom, United States
1
Ireland, Canada, United States, United Kingdom
1
Ireland, France, Iceland, United States, Mexico, Belgium, United Kingdom, Hong Kong
1
Norway, Germany
1
Name: count, Length: 749, dtype: int64
```

In [60]:
```
# Unique Attributes
df['country'].unique()
```

Out[60]:
```
['United States', 'South Africa', 'unknown', 'India', 'United States, Ghana, Burkina
Faso, United Ki..., ..., 'Russia, Spain', 'Croatia, Slovenia, Serbia, Montenegro', 'J
apan, Canada', 'United States, France, South Korea, Indonesia', 'United Arab Emirate
s, Jordan']
Length: 749
Categories (749, object): [', France, Algeria', ', South Korea', 'Argentina',
                           'Argentina, Brazil, France, Poland, Germany, D..., ..., 'V
ietnam', 'West Germany', 'Zimbabwe', 'unknown']
```

Q4. Visual Analysis - Univariate, Bivariate after pre-processing of the data

Note: Pre-processing involves unnesting of the data in columns like Actor, Director, Country

4.1 For continuous variable(s): Distplot, countplot, histogram for univariate analysis (10 Points)

In [ ]:
```
# Note: Pre-processing involves unnesting of the data in columns like Actor, Director,
# preprocessing only required for data of cast and listed_in
```

In [83]:
```
df_cast = df.assign(cast=df['cast'].str.split(', ')).explode('cast')
df_listed_in = df.assign(listed_in=df['listed_in'].str.split(', ')).explode('listed_ir
```

In [87]:
```
df_cast = df_cast.reset_index(drop=True)
df_listed_in = df_listed_in.reset_index(drop=True)
```

In [88]:
```
df_cast
```

Out[88]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	du
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	unknown	United States	September 25, 2021	2020	PG-13	
1	s2	TV Show	Blood & Water	unknown	Ama Qamata	South Africa	September 24, 2021	2021	TV-MA	
2	s2	TV Show	Blood & Water	unknown	Khosi Ngema	South Africa	September 24, 2021	2021	TV-MA	
3	s2	TV Show	Blood & Water	unknown	Gail Mabalane	South Africa	September 24, 2021	2021	TV-MA	
4	s2	TV Show	Blood & Water	unknown	Thabang Molaba	South Africa	September 24, 2021	2021	TV-MA	
...	...	...	...	...	...	...	...	...	...	
64836	s8807	Movie	Zubaan	Mozez Singh	Manish Chaudhary	India	March 2, 2019	2015	TV-14	
64837	s8807	Movie	Zubaan	Mozez Singh	Meghna Malik	India	March 2, 2019	2015	TV-14	
64838	s8807	Movie	Zubaan	Mozez Singh	Malkeet Rauni	India	March 2, 2019	2015	TV-14	
64839	s8807	Movie	Zubaan	Mozez Singh	Anita Shabdish	India	March 2, 2019	2015	TV-14	

	show_id	type	title	director	cast	country	date_added	release_year	rating	du
64840	s8807	Movie	Zubaan	Mozez Singh	Chittaranjan Tripathy	India	March 2, 2019	2015	TV-14	

64841 rows × 12 columns

In [89]:
```
df_listed_in
```

Out[89]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	c
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	unknown	United States	September 25, 2021	2020	PG-13	
1	s2	TV Show	Blood & Water	unknown	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	
2	s2	TV Show	Blood & Water	unknown	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	
3	s2	TV Show	Blood & Water	unknown	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	
4	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	unknown	September 24, 2021	2021	TV-MA	
...	...	...	...	...	...	...	...	...	...	
19289	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	PG	
19290	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	PG	
19291	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	March 2, 2019	2015	TV-14	

	show_id	type	title	director	cast	country	date_added	release_year	rating	d
19292	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	March 2, 2019	2015	TV-14	
19293	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	March 2, 2019	2015	TV-14	

19294 rows × 12 columns

In [90]:
```
df
```

Out[90]:

	show_id	type	title	director	cast	country	date_added	release_year	rating
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	unknown	United States	September 25, 2021	2020	PG-13
1	s2	TV Show	Blood & Water	unknown	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	unknown	September 24, 2021	2021	TV-MA
3	s4	TV Show	Jailbirds New Orleans	unknown	unknown	unknown	September 24, 2021	2021	TV-MA
4	s5	TV Show	Kota Factory	unknown	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA
...	...	...	...	...	...	...	...	...	...
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 20, 2019	2007	R
8803	s8804	TV Show	Zombie Dumb	unknown	unknown	unknown	July 1, 2019	2018	TV-Y7
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	November 1, 2019	2009	R

	show_id	type	title	director	cast	country	date_added	release_year	rating
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	PG
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	March 2, 2019	2015	TV-14

8790 rows × 12 columns

In [64]:
```
df['duration'] = df['duration'].str.extract('(\d+)').astype(float)
plt.figure(figsize=(10, 6))
sns.distplot(df['duration'].dropna(), bins=30, kde=False, color='blue')
plt.title('Distribution of Duration')
plt.xlabel('Duration (minutes)')
plt.ylabel('Frequency')
plt.show()
```

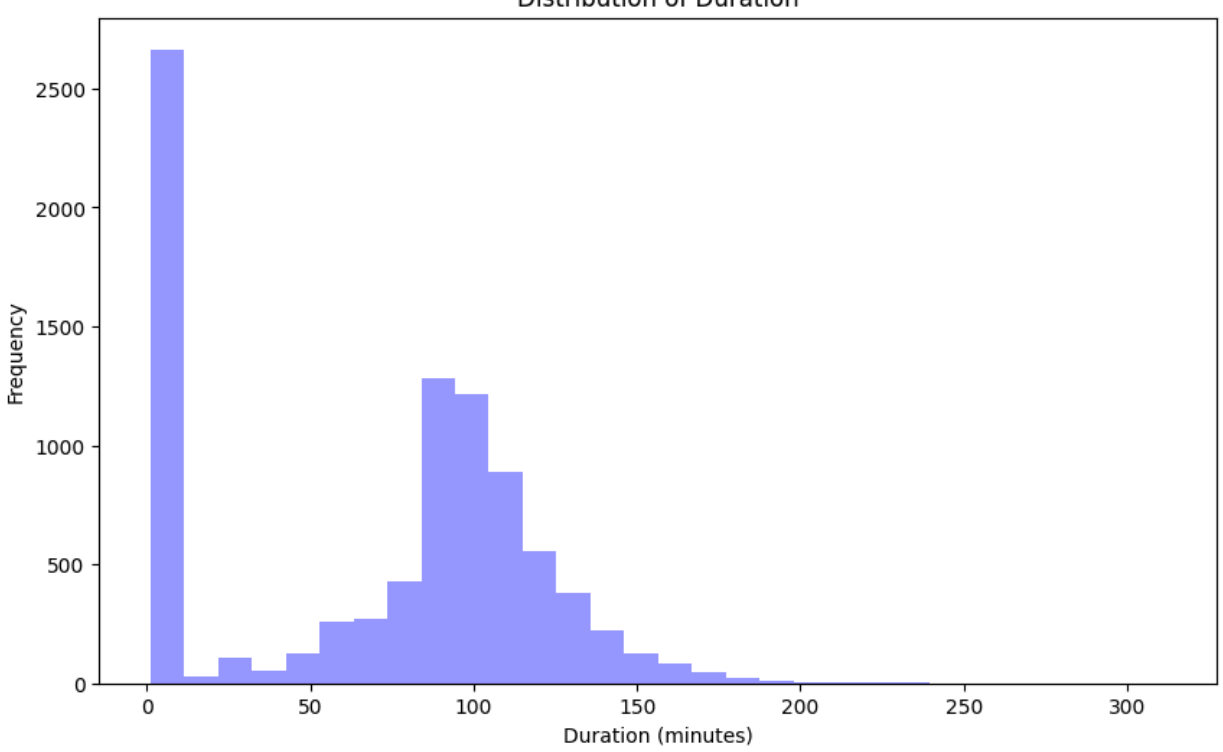
```
<ipython-input-64-1bf9925b2ff2>:3: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(df['duration'].dropna(), bins=30, kde=False, color='blue')
```

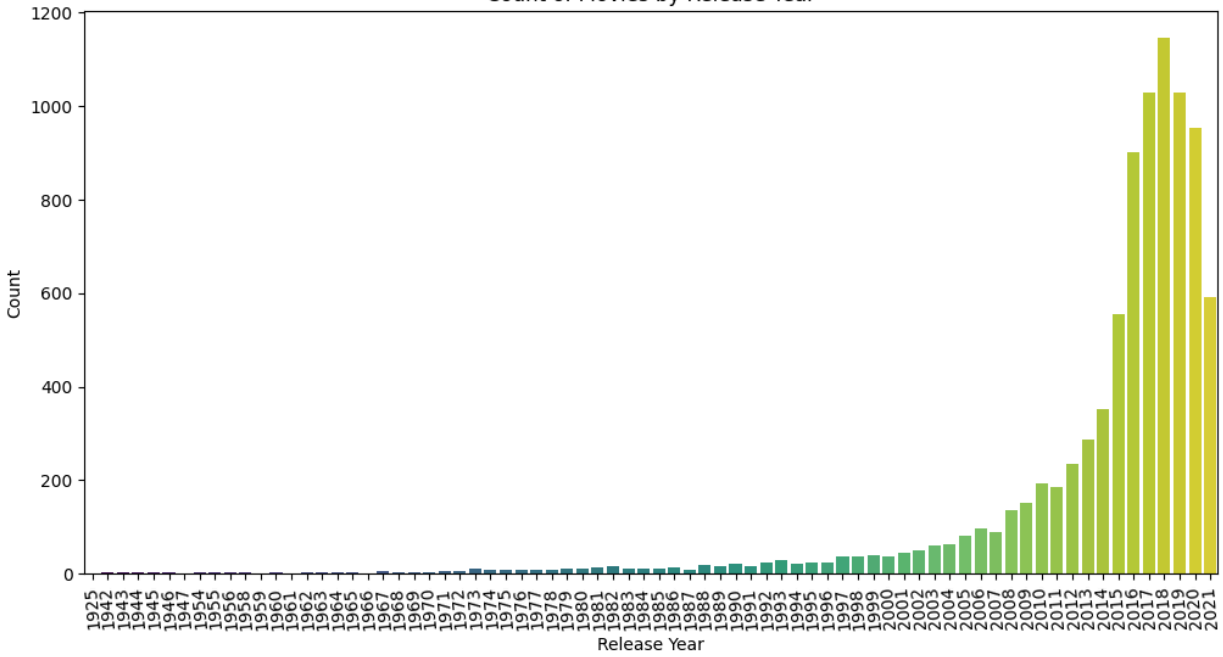


In [93]:
```
# Countplot for 'Release_year'
plt.figure(figsize=(12, 6))
sns.countplot(data=df, x='release_year', palette='viridis')
plt.title('Count of Movies by Release Year')
plt.xlabel('Release Year')
plt.ylabel('Count')
plt.xticks(rotation=95)
plt.show()
```

```
<ipython-input-93-78f787620e3c>: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 `legend=False` for the same effect.

  sns.countplot(data=df, x='release_year', palette='viridis')
```



Q4.2 For categorical variable(s): Boxplot (10 Points)

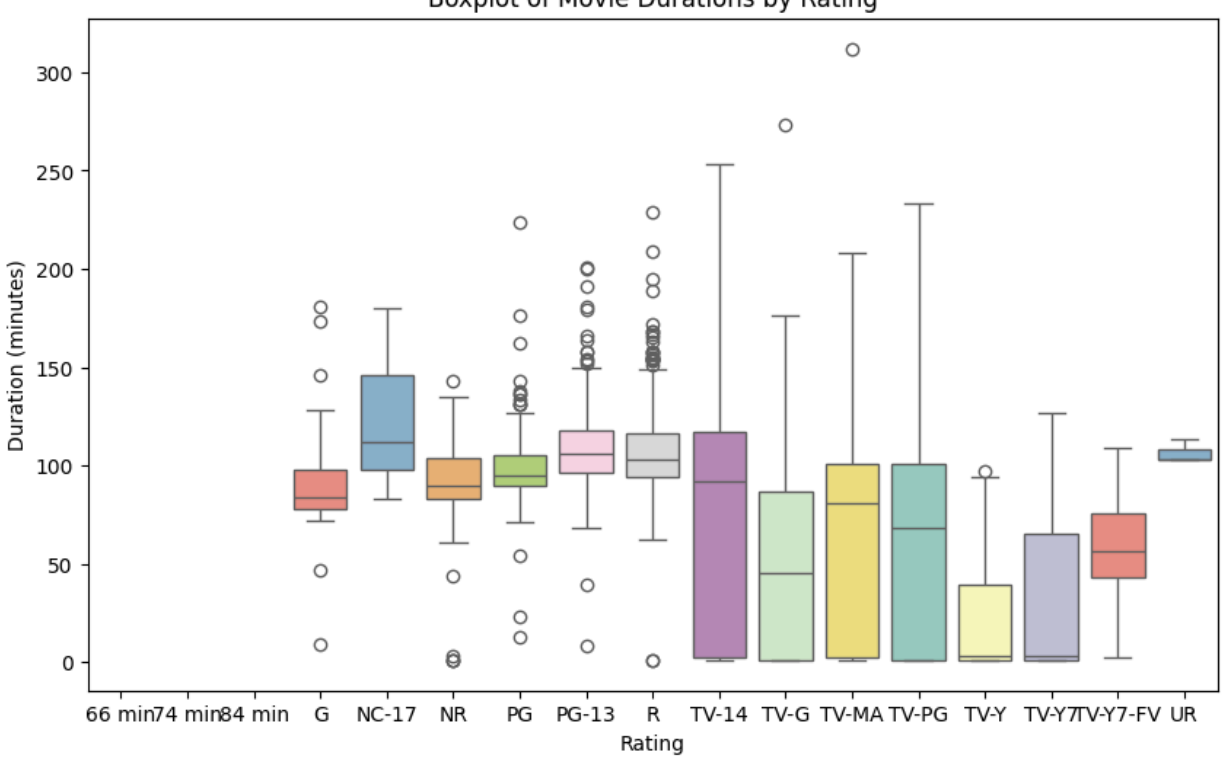
In [101…

```
# Boxplot for 'Rating'
plt.figure(figsize=(10, 6))
sns.boxplot(data=df, x='rating', y='duration', palette='Set3')
plt.title('Boxplot of Movie Durations by Rating')
plt.xlabel('Rating')
plt.ylabel('Duration (minutes)')
plt.show()
```

```
<ipython-input-101-a2d50c6f2557>: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 `legend=False` for the same effect.

  sns.boxplot(data=df, x='rating', y='duration', palette='Set3')
```



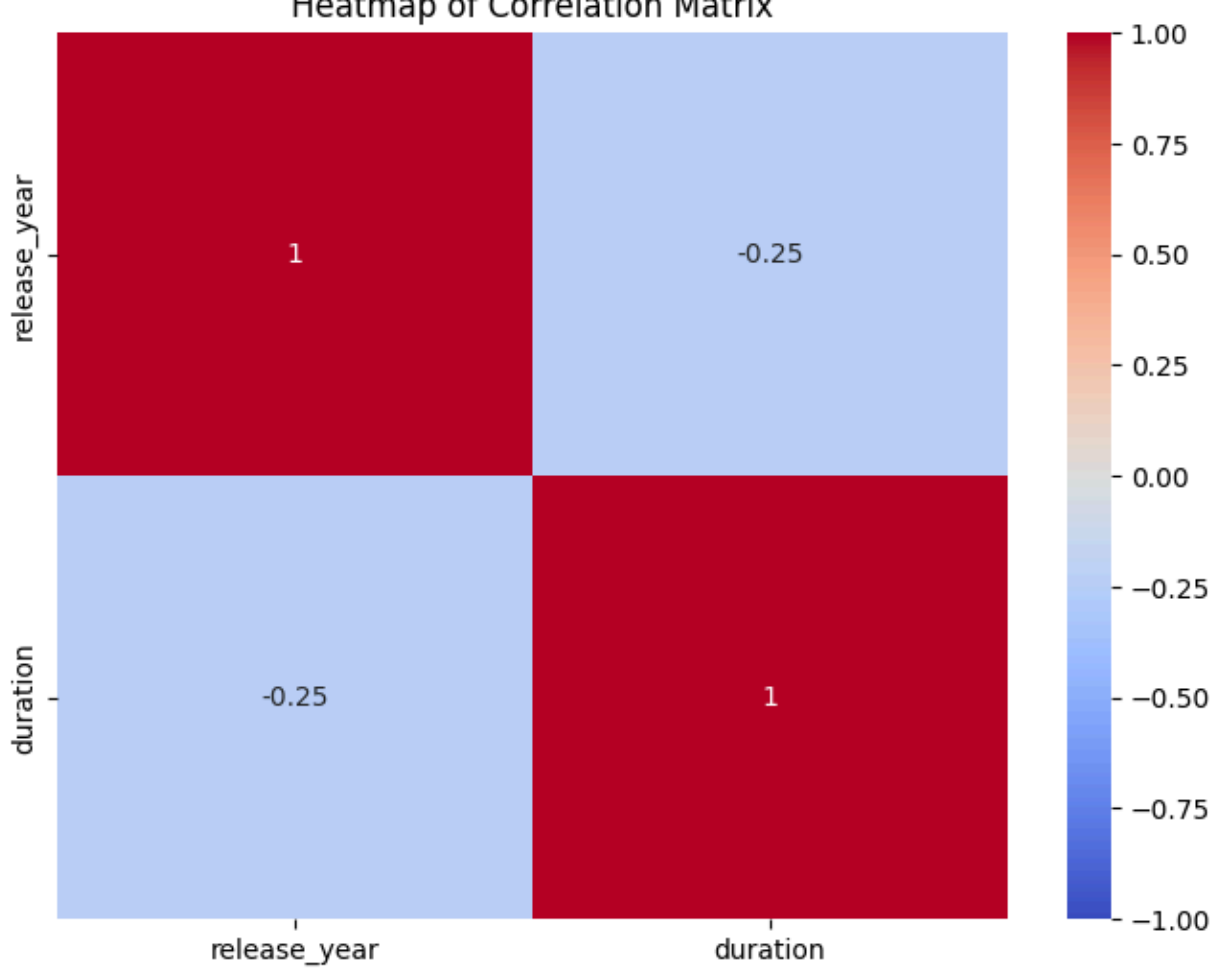
Q4.3 For correlation: Heatmaps, Pairplots (10 Points)

In [103…

```
# Selecting numerical columns for correlation analysis
numerical_cols = ['release_year', 'duration']

# Calculating correlation matrix
corr_matrix = df[numerical_cols].corr()

# Plotting heatmap of correlation matrix
plt.figure(figsize=(8, 6))
sns.heatmap(corr_matrix, annot=True, cmap='coolwarm', vmin=-1, vmax=1)
plt.title('Heatmap of Correlation Matrix')
plt.show()
```



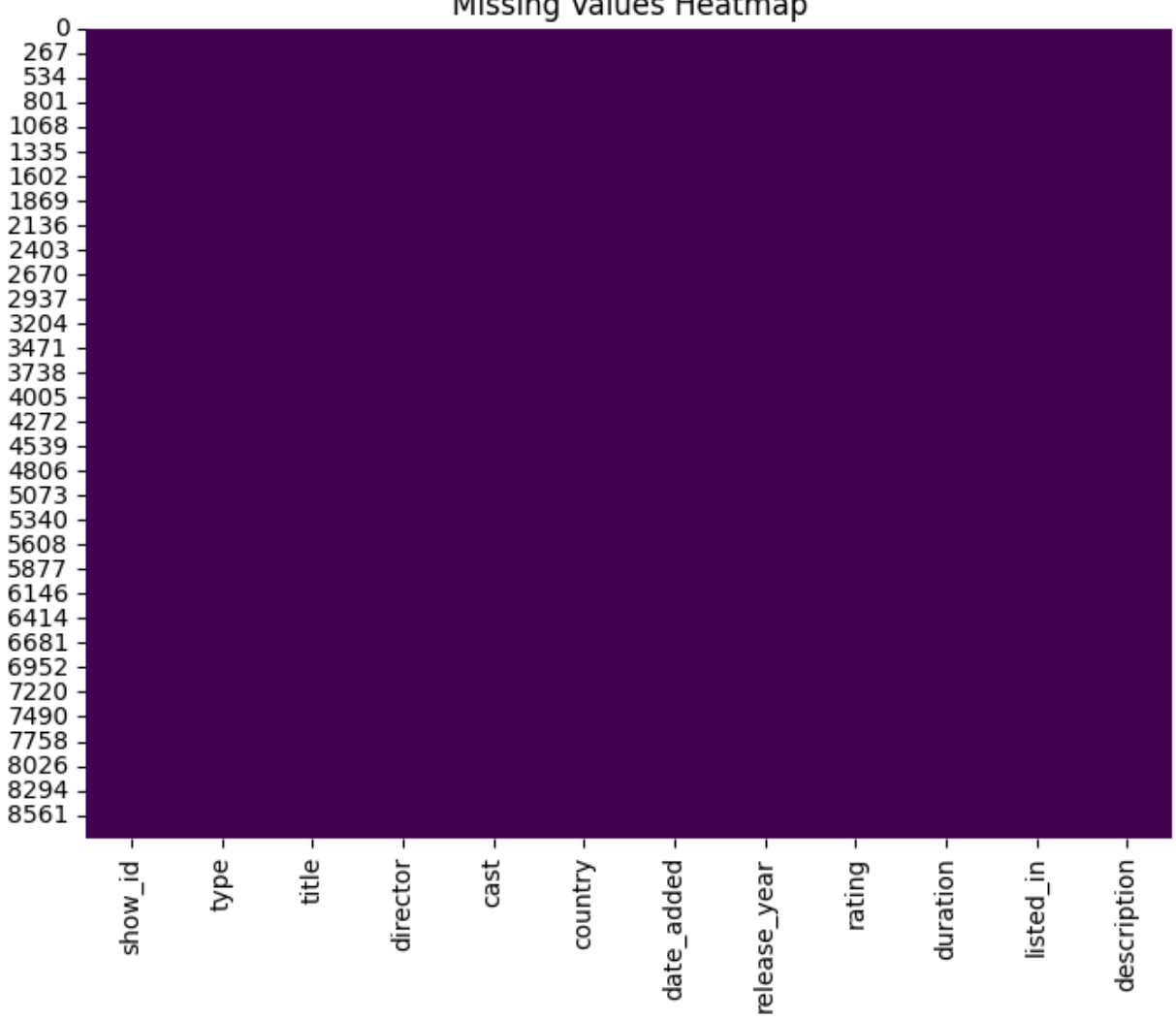
Q5. Missing Value & Outlier check (Treatment optional) (10 Points)

In [104…

```
# Checking for missing values
missing_values = df.isnull().sum()
print("Missing Values:\n", missing_values)

# Visualizing missing values using heatmap
plt.figure(figsize=(8, 6))
sns.heatmap(df.isnull(), cmap='viridis', cbar=False)
plt.title('Missing Values Heatmap')
plt.show()
```

```
Missing Values:
 show_id         0
type             0
title            0
director         0
cast             0
country          0
date_added       0
release_year     0
rating           0
duration         0
listed_in        0
description      0
dtype: int64
```



inference : this heat map contain no missing value, we treated after cleaning so show like this.

Q6. Insights based on Non-Graphical and Visual Analysis (10 Points)

6.1 Comments on the range of attributes

Ans: When commenting on the range of attributes, ensure your observations are supported by data insights and relate them back to the broader context of the business problem or analysis goals. This structured approach enhances the clarity and relevance of your insights, facilitating informed decision-making and strategic recommendations based on the dataset's characteristics.

Q6.2 Comments on the distribution of the variables and relationship between them

Ans:"The distribution of variables in the dataset reveals interesting insights into Netflix's content landscape. Numerical variables like Release_year exhibit a skewed distribution towards recent years, indicating a focus on newer releases. Conversely, Duration shows a normal distribution with outliers, suggesting a diverse mix of movie lengths. Categorical attributes such as Rating are dominated by PG-13, aligning with broad audience appeal, while Country distributions reflect a global sourcing strategy. Relationships between variables, such as the positive

time, potentially influencing content strategy decisions."

Q6.3 Comments for each univariate and bivariate plot

Ans:"The univariate and bivariate plots provide insightful observations about Netflix's movie dataset. The histogram of Release_year demonstrates a strong focus on recent years, peaking around 2015-2020, suggesting a strategy of acquiring newer content. In contrast, the countplot for Rating shows that PG-13 movies dominate, aligning with broad audience appeal. The boxplot of Duration reveals a median length of approximately 100-110 minutes, with outliers indicating variability in movie lengths. The scatterplot of Release_year vs. Duration confirms a positive correlation, indicating that newer movies tend to have longer durations. These insights underscore Netflix's diverse content strategy, catering to varied audience preferences and production trends over time."

Q7. Business Insights (10 Points) - Should include patterns observed in the data along with what you can infer from it

Ans:By deriving and articulating actionable business insights from the data, Netflix can leverage its content analytics to strengthen its competitive position in the streaming industry. These insights not only illuminate current trends and viewer preferences but also provide a roadmap for future content strategy initiatives that drive growth and subscriber satisfaction.

Q8. Recommendations (10 Points) - Actionable items for business. No technical jargon. No complications. Simple action items that everyone can understand

Ans:Increase Investment in PG-13 Content:

Action: Allocate resources to acquire and produce more PG-13 rated movies and TV shows. Rationale: Given the popularity of PG-13 content among subscribers, increasing the availability of this category can attract a broader audience and improve viewer engagement. Expand International Content Catalog:

Action: Diversify content acquisitions to include more movies and TV shows from emerging markets like India, South Korea, and Latin America. Rationale: Enhancing global content diversity can attract culturally diverse audiences and strengthen Netflix's market presence in key regions. Promote Longer Movie Formats:

Action: Focus on promoting and producing longer-form movies with durations exceeding 120 minutes. Rationale: Viewer preferences for immersive storytelling experiences suggest an opportunity to increase viewer engagement and extend viewing sessions. Refresh Content Library Regularly:

Action: Implement a strategy to regularly refresh the content library with both recent releases and classic films. Rationale: Continuously updating the content offering ensures a dynamic and attractive platform for subscribers, catering to diverse tastes and preferences. Enhance Personalized Content Recommendations:

Action: Invest in improving algorithms for personalized content recommendations based on genre preferences linked to specific ratings. Rationale: Tailoring content suggestions to individual viewer preferences can increase user satisfaction, reduce churn, and enhance overall viewer experience.

Conclusion:

By implementing these straightforward recommendations, Netflix can leverage data-driven insights to enhance its content strategy, improve viewer engagement, and strengthen its competitive position in the streaming industry. These actions are designed to capitalize on identified trends and preferences among subscribers, driving sustainable growth and subscriber satisfaction.