Financial Performance Analysis of Netflix 2002 - 2022 and Growth Forecasts

Trang Nguyen

Qi Wang

Introduction

In recent years, streaming services have become increasingly popular, and Netflix is well-known as the most successful streaming platform. Based on the market cap in April 2023, the company's market value is estimated to be about $ 146.67 billion USD that made Netflix gain competitive advantage over other companies in the on-demand media industry in the world[1]. The purpose of this study is to investigate relationships as well as correlations between the numbers of subscribers and financial performance of the company to get valuable insights into the current state of the streaming industry and to gain a better understanding on consumer behavior, market trends, the effectiveness of Netflix's business strategy, company's success, and its future prospects in the next 5 years. To conduct the analysis, categorization of analytical methods is applied in this research, and descriptive analytical methodology, which is known as a scientific research methodology, is utilized in the study to form research questions. The methodology involves formulating hypotheses, refining the research questions, collecting data, testing the hypotheses, and performing analysis to conclude the thorough results of the research.

Data Description

The dataset is obtained from a reliable source Statista. It is available in .xlsx format with a file size of 432 KB. The data collection contains 24 observations and 11 variables, which describe information regarding Netflix's performance from 2002 to 2022 worldwide. The target variables in the data set are qualitative and quantitative variables. Qualitative variables are also known as categorical variables, representing data which cannot measure in the form of number, and it is divided into nominal variables and ordinal variables. In this data source, "World

Region" and "Netflix Subscribers Groups" are targets of nominal variables while "Year" and "Age Group" variables are considered as ordinal variables of the data. By contrast to categorical variables, quantitative variables represent data that is numerical, and it is categorized into discrete variables and continuous variables. Discrete variable in this data file is "Number of Netflix Subscriber". Continuous variables are "Gross Income", "Marketing Expenses", and "Profit". To minimize the impact of data errors on the results, pre-processing steps are considered important steps taken to clean, remove missing, erroneous data, and transform the raw data into proper format for analysis.

Data Exploratory Analysis

Descriptive statistics refer to statistics measures describing characteristics of a dataset. To gain insight into central tendency, variability, and distribution, key descriptive statistics for the four variables of Gross Income, Marketing Expenses, Profits, and Number of Netflix Subscriber is calculated, including measures of mean, standard deviation, median, maximum, and minimum values. Exploratory data analysis is utilized in the analysis as a process of examining the data sets to identify trends, patterns, and other aspects including missing values, outliers, relationships between variables. In this study, box plots and histograms are used in exploratory data analysis to visualize the distribution and variability of the data.

Table 1. Descriptive Statistic of Numbers of Netflix Subscribers and Netflix Financials 2002 to 2022

Descriptive Statistics	Number of Netflix Subscriber (million)	Gross Income (millions USD)	Marketing Expenses (millions USD)	Profit (millions USD)
Mean	68.71	8345.71	908.89	831.61
Std	78.48	10112.78	957.52	1499.35
Min	0.796	150.80	35.78	-21.00
Median	30.36	3609.30	439.21	122.64
Max	230.70	31615.55	2652.46	5116.23

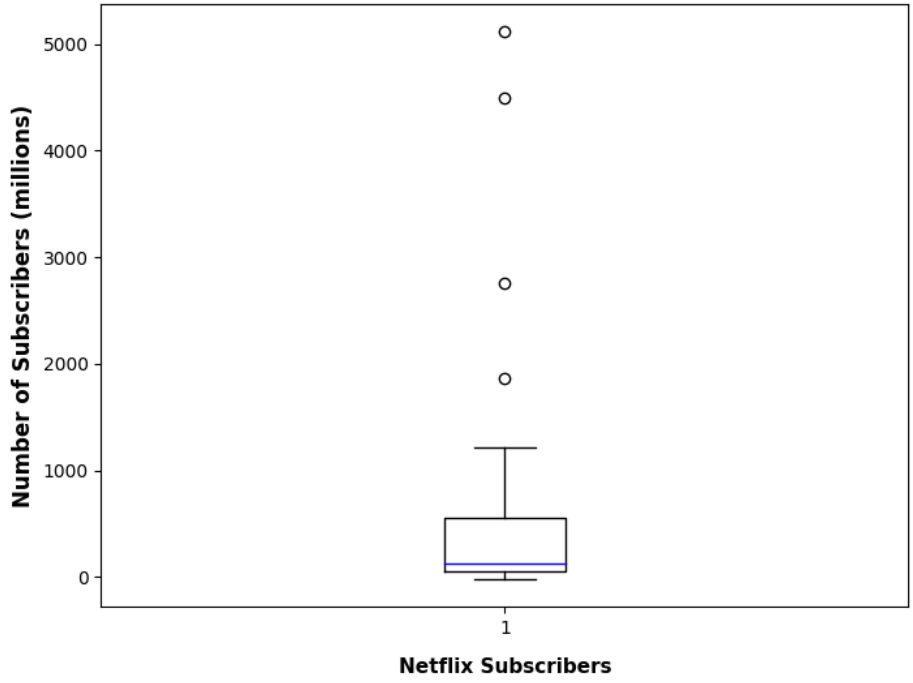

Figure 1. Box Plot of Number of Subscribers

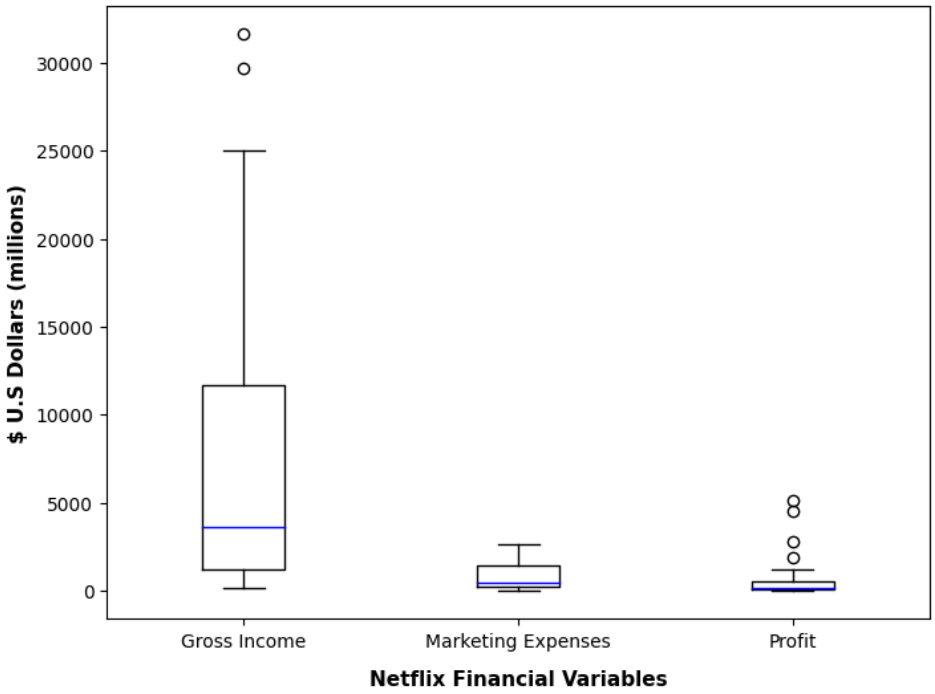

Figure 2. Box Plots of Netflix Financial Variables: Gross Income, Marketing Expenses and Profit

Based on Figure 3, it is discovered that the distribution of the four variables appears to be positively skewed, with a longer tail to the right side of the distribution. This also indicates that Netflix's data is expanding at a fairly consistent rate, which is reflected in the relatively high values observed, and the positively skewed feature can also be indicated by the mean being greater than the median (Table 1). Therefore, there is a possibility of extreme values, as shown in the boxplots. In addition, it can be seen from the standard deviation that the variability in gross income and profits is relatively high, while the variability in marketing expenses and the number of subscribers is relatively low. From the histograms (Figure 3), there is a trend of an increasing number of data points in the higher value range, indicating a potential positive trend in the future.

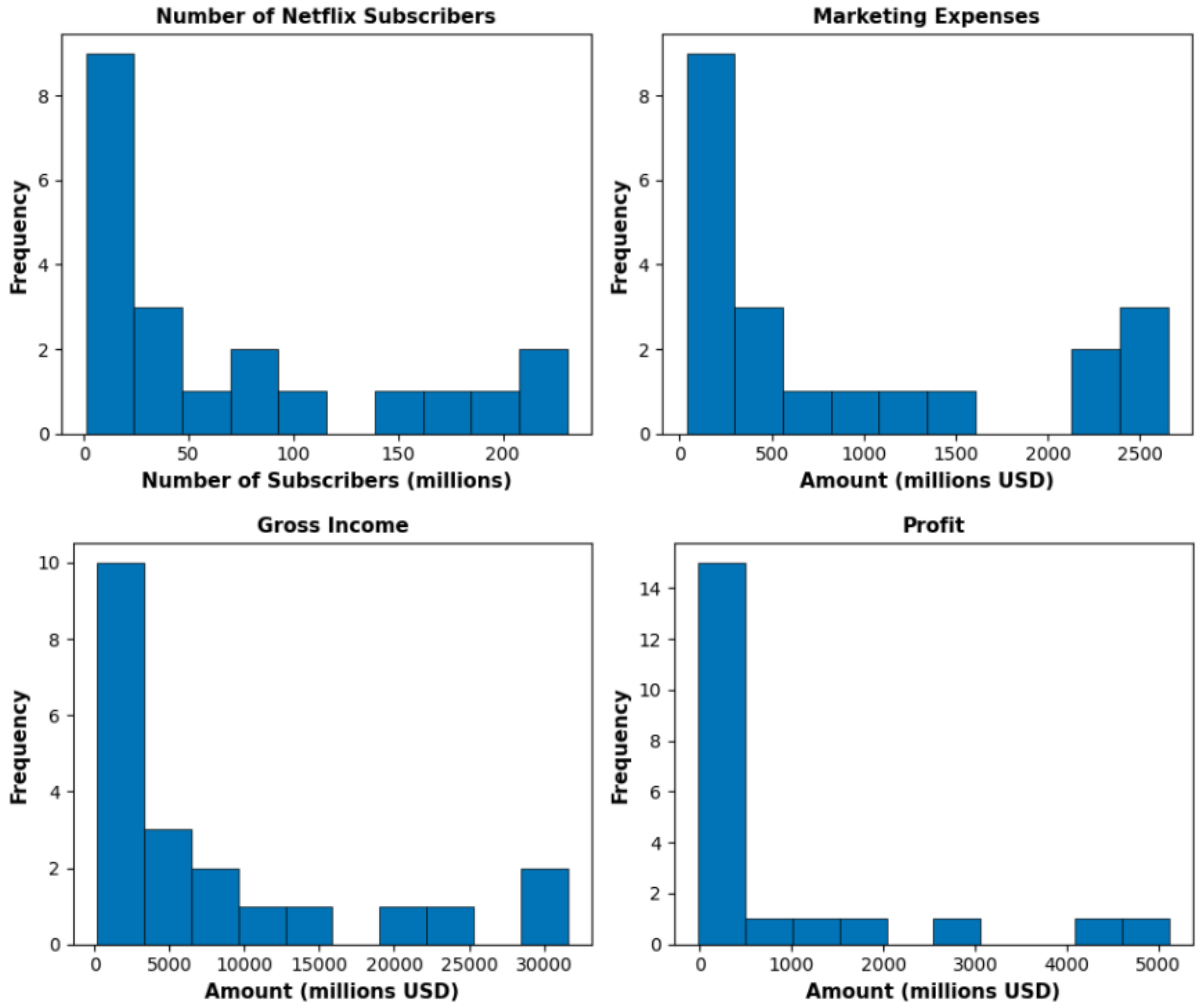

Figure 3. Histograms of Number of Netflix Subscribers, Marketing Expenses, Gross Income, Marketing Expenses

Data Visualization Design

Types of data visualizations are used in this project based on the purpose of the study and types of data. Appropriate color schemes are selected when creating data visualizations that involve quantification or categorization. To convey information clearly, it is essential to be cautious and thoughtful that using too many colors or using colors that are too bright or too dark can detract from the clarity of the visualization. To emphasize key findings, besides using color, other pre-attentive attributes such as size, length, position, etc. are also used to draw attention to important points. Below are specific examples of different data visualizations involved in this project.

To analyze different aspects of Netflix's performance and Netflix's user demographics, we use the map to show the geographic distribution of Netflix in different regions, and a bar graph to display top countries with the most Netflix subscribers. Categorical palette is used to distinguish different variables in the bar graph. Scatter plots are also utilized in this study because they are the best data visualization type for continuous variable information, and they are very helpful in discovering inferences and viewing potential data relationships between financial performance of the company and subscriber growth. Patterns including the overall trend, the strength of the relationship between variables, clusters of points and outliers can be easily identified based on the distribution of points in the scatterplot. Furthermore, to analyze the financial performance of Netflix from 2002 to 2022 and predict future performance of the company's growth, the line chart is considered as the most appropriate data visualization tool to identify the trends over a period of time and allow for potential forecasting. After discovering the potential correlation between these variables, a table listing the values of various correlation coefficients is utilized to quantify the strength of the relationship. By using the table, it can be easy to make comparisons across categories. In addition, to display the strength of relationship between multiple variables, a heatmap with a sequential palette is also used for visualization,

which allows us to visualize the correlation coefficients and highlight the weakness as well as the strongest relationships in the data.

Results

Data Visualizations of Netflix Distribution

The global distribution of Netflix subscribers is displayed in Figure 5. As of 2022, Netflix has estimated that the streaming service is available in 7 regions including North America, Central America, South America, Europe, Middle East, Africa and Asia, and the streaming platform has been expanding in over 190 countries[2].

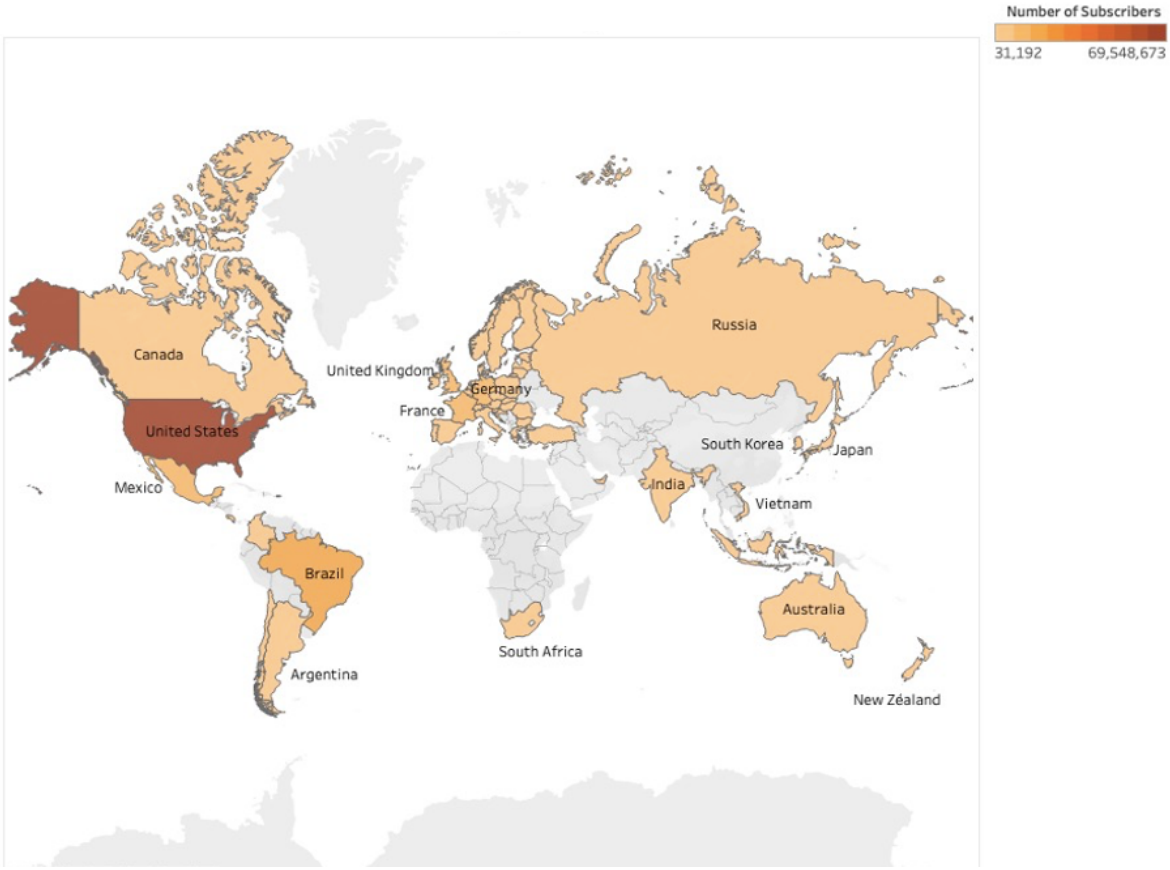

Figure 5. Netflix Subscribers by Country in 2022

According to Table 2, there were 76.73 million subscribers in Europe, Middle East, and Africa in 2022. 74.30 million subscribers were in the United States and Canada. Latin America and Asia Pacific had 41.70 million and 38.02 million subscribers, respectively. However, in comparison to other countries in Europe, Latin America and Asia regions, the United States

was considered to be the top countries with the most Netflix subscribers with 69.55 million subscribers in 2022, followed by Brazil with 18.84 million subscribers and the United Kingdom with 13.18 million subscribers, as shown in Figure 6.

Table 2. Netflix Subscribers in Different Regions in 2022

World Regions	Subscriber Numbers
Europe, Middle East, and..	76.73 M
U.S.A. and Canada	74.30 M
Latin America	41.70 M
Asia Pacific	38.02 M

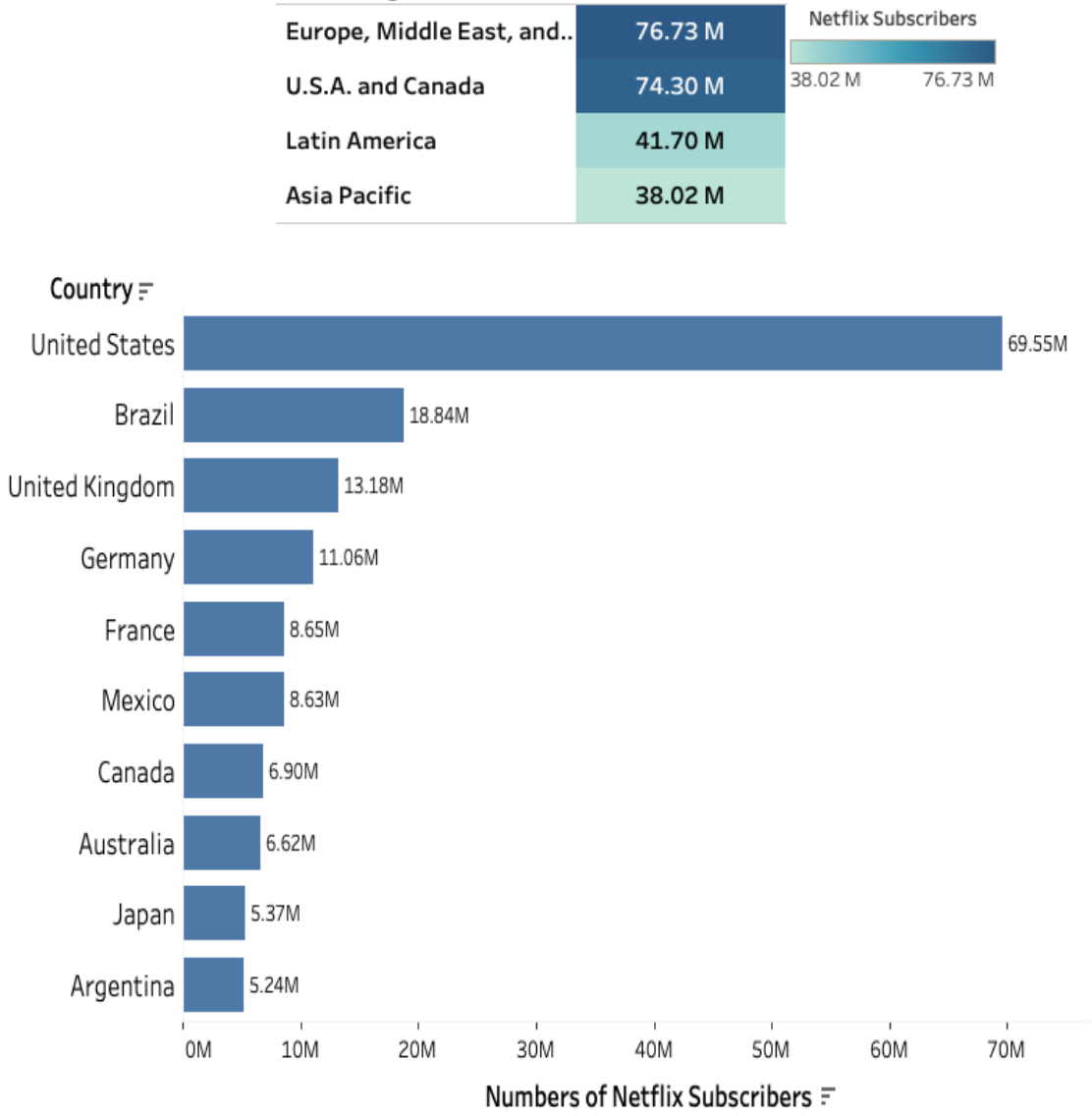

Figure 6. Top 10 Countries with the Most Netflix Subscribers in 2022

Based on the study's findings, there were 230.7 million Netflix paid subscribers worldwide in 2022 (Figure 8). Gen Z and Millennial age groups comprise 70% and 65 % of the Netflix customer base; by contrast, Gen Z and Baby Boomers make up only 54% and 39%,

respectively among Netflix consumers (Figure 7). It is concluded that Netflix's popularity is more attractive to younger customers rather than older generations.

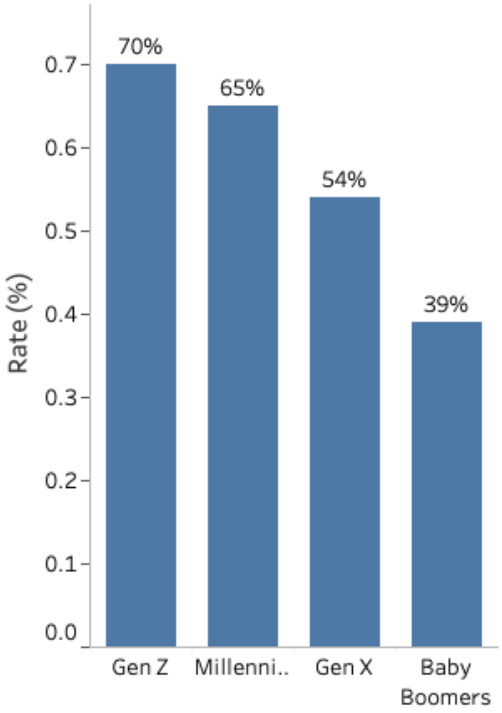

Figure 7. Netflix's Age Demographics 2022

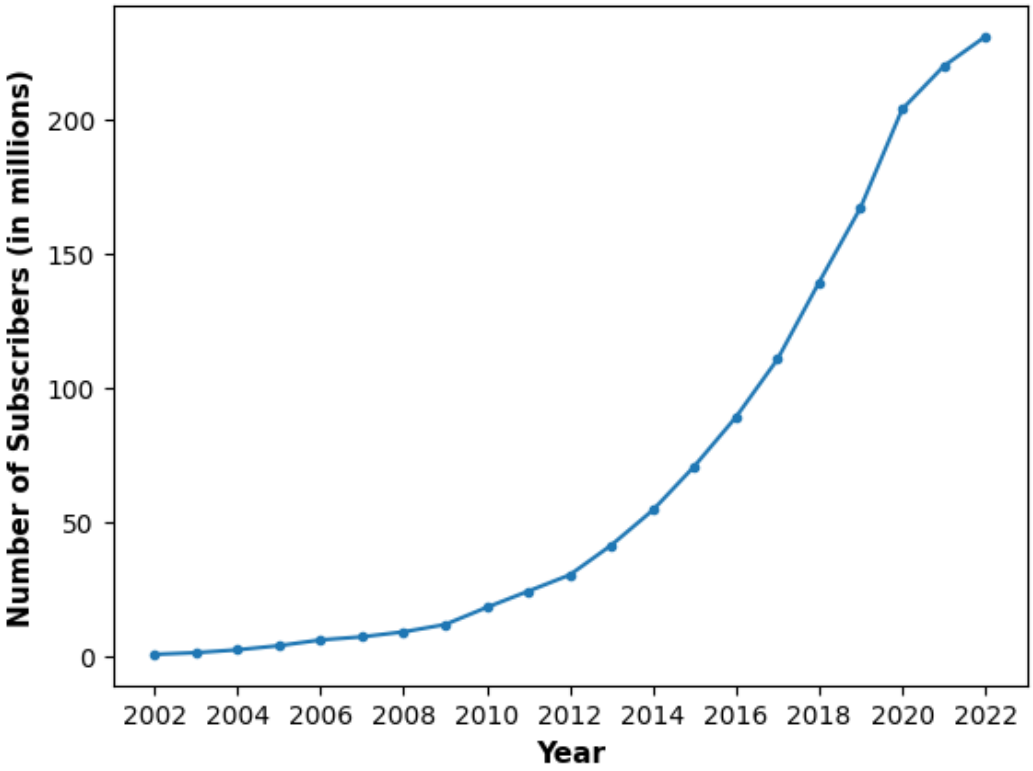

Figure 8. Trends in the Numbers of Netflix Subscribers from 2002 to 2022

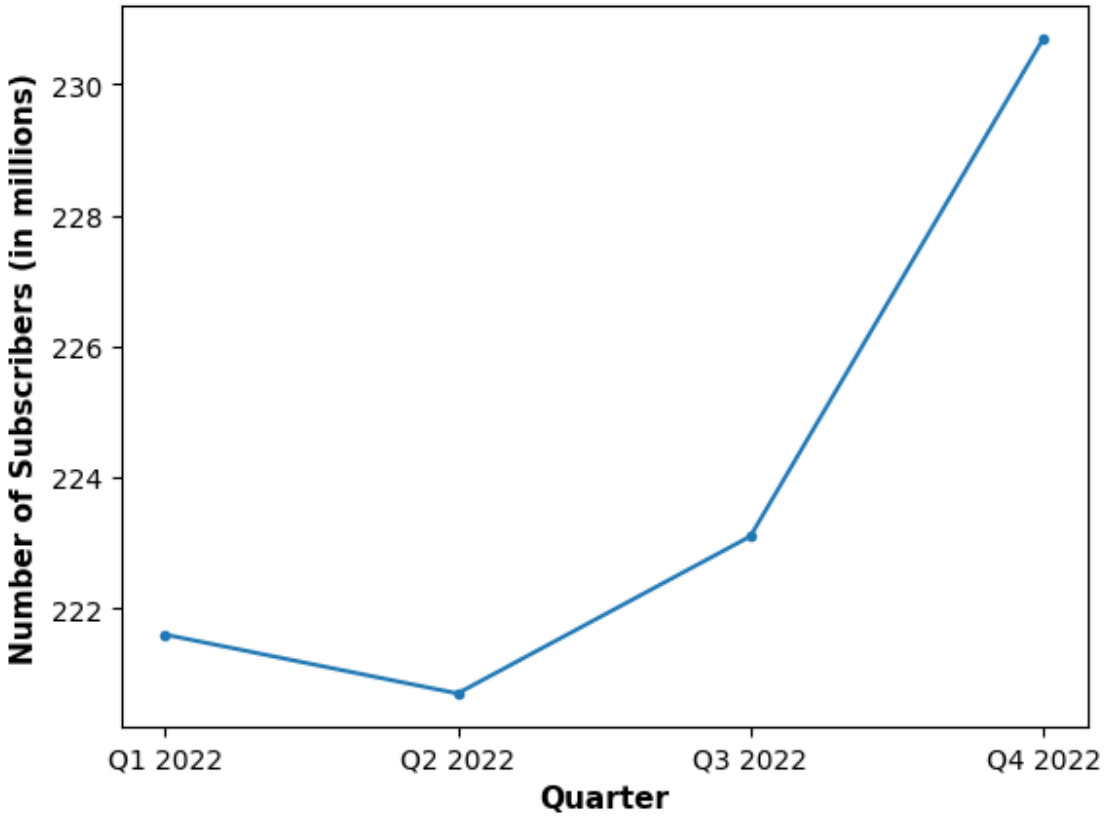

Figure 9. Trends in Numbers of Netflix Subscribers from Q1 2022 to Q4 2022

Overall, there has been a significant growth in the number of Netflix subscribers over the years. According to Figure 8, Netflix's streaming business has expanded from 796 thousand subscribers in 2002 and reaching 230.7 million subscribers by 2022. However, as seen in Figure 9, there is a trend highlighted between the first and second quarter of the year 2022. At the beginning of the year 2022, due to the war in Ukraine, the company stopped Netflix's streaming services in Russia. Additionally, the company increased subscription cost in the United States and Canada along with too many competitors available in streaming service; as a consequence, there were a lot of subscribers who canceled Netflix accounts, leading to a million subscribers dropping between April and July of 2022. Nevertheless, Netflix's subscriber growth is surging again in the third quarter of the year.

Financial Performance Analysis and Predictions

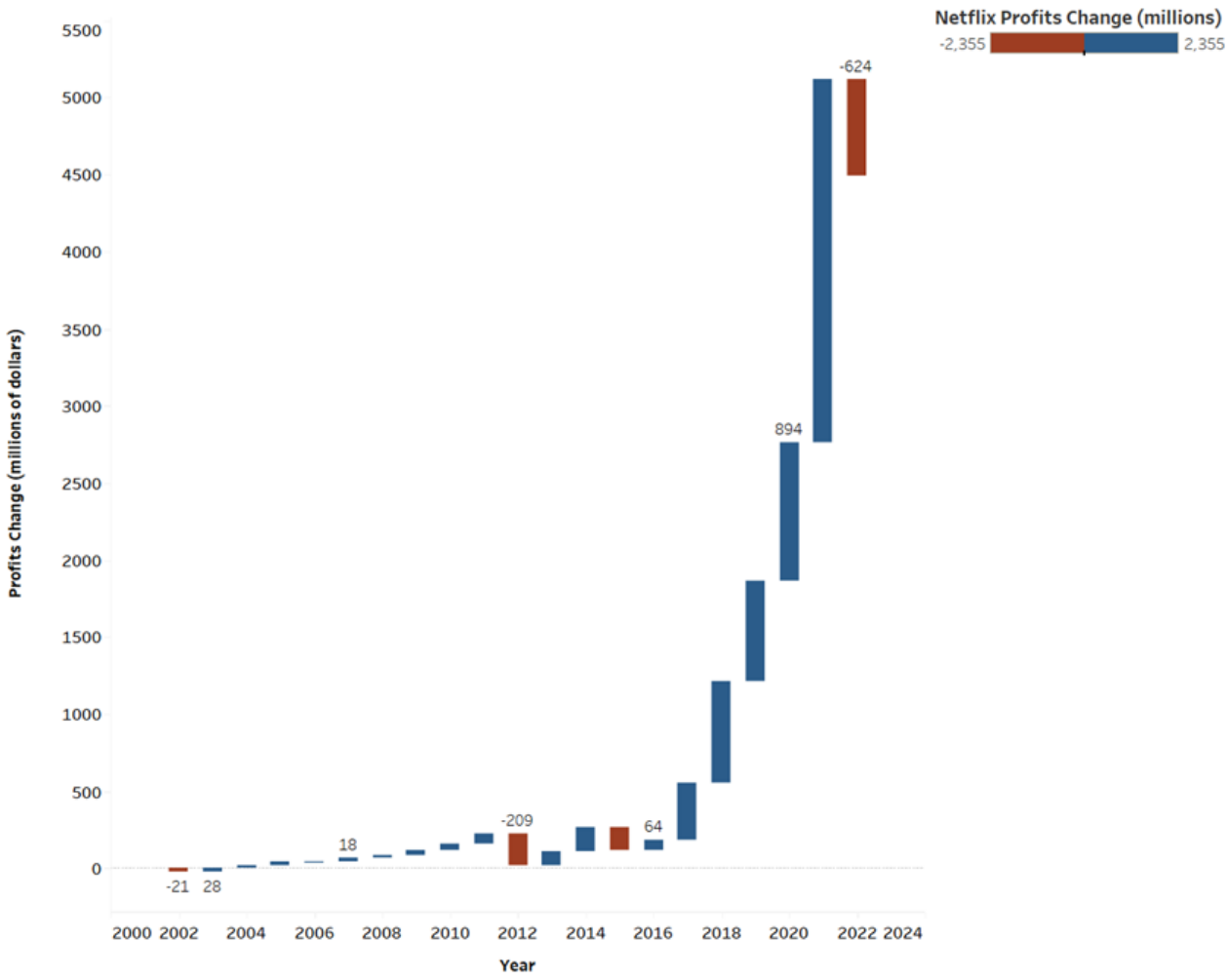

Figure 10. Changes in Netflix Profit from 2002 to 2022

Over the years, Netflix is still considered to be the fastest growing streaming service and a dominant company in the on-demand media industry. In its early stages of 2022, the company was primarily focused on growing the DVD rental business. As demonstrated in Figure 10, the company's net income lost around $21 million USD because the company's expenses including marketing, operations, and content costs, exceeded its revenue from DVD rentals. However, in 2021, the company's profit was achieved to approximately $5,116.23 millions. In 2022, the company's profit was around $4491.92 millions, a $624 millions decline from 2021 due to the war in Ukraine and an increase of subscription cost.

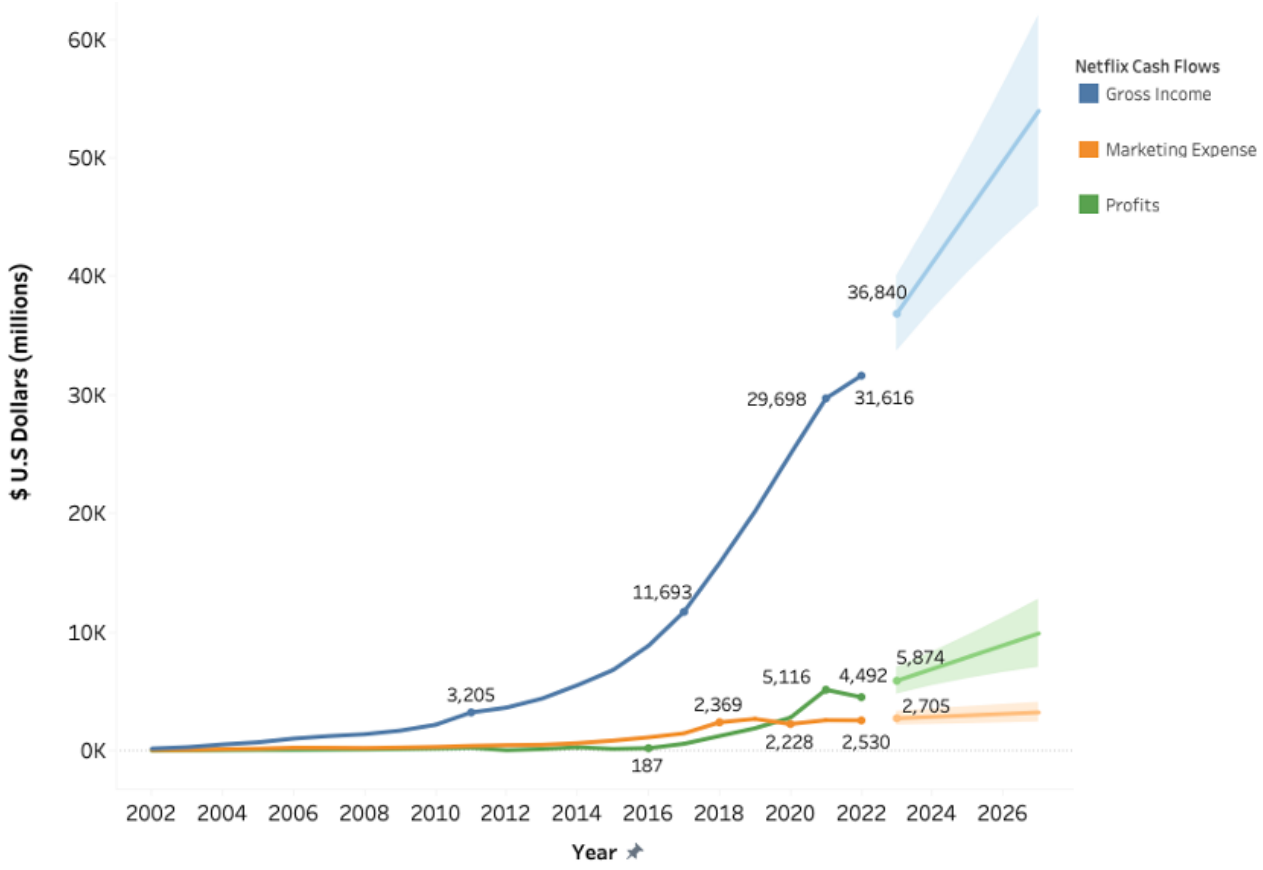

Figure 11. Netflix's Marketing Expense, Gross Income, Profits 2002 to 2022, and Growth Forecasts

Gross income, Marketing expenses and Profits have been analyzed to better study the financial performance of Netflix over the past 21 years. From Figure 11, we can see that there was rapid growth in the middle years. However, in recent years, the growth of all three variables has slowed down, and there have even been fluctuations, which is similar to the changes in the numbers of subscribers we analyzed earlier. These similarities in trend lines seem to reflect an association among these variables. Overall, these three variables have shown a relatively stable growth since 2002. Based on the overall trend, by using Tableau's forecasting function, we can get positive trend lines of these three variables, as indicated by the shaded area in the graph and predicted data for the next five years as shown in Table 3. These values can provide a useful reference for predicting future growth.

Table 3. Growth Forecast of Netflix Financial Performance 2023 to 2027

Year	Gross Income (millions of dollars)	Profits (millions of dollars)	Marketing Expenses (millions of dollars)
2023	36840	5874	2705
2024	41118	6870	2830
2025	45396	7866	2955
2026	49673	8862	3079
2027	53951	9858	3204

Relationship Analysis

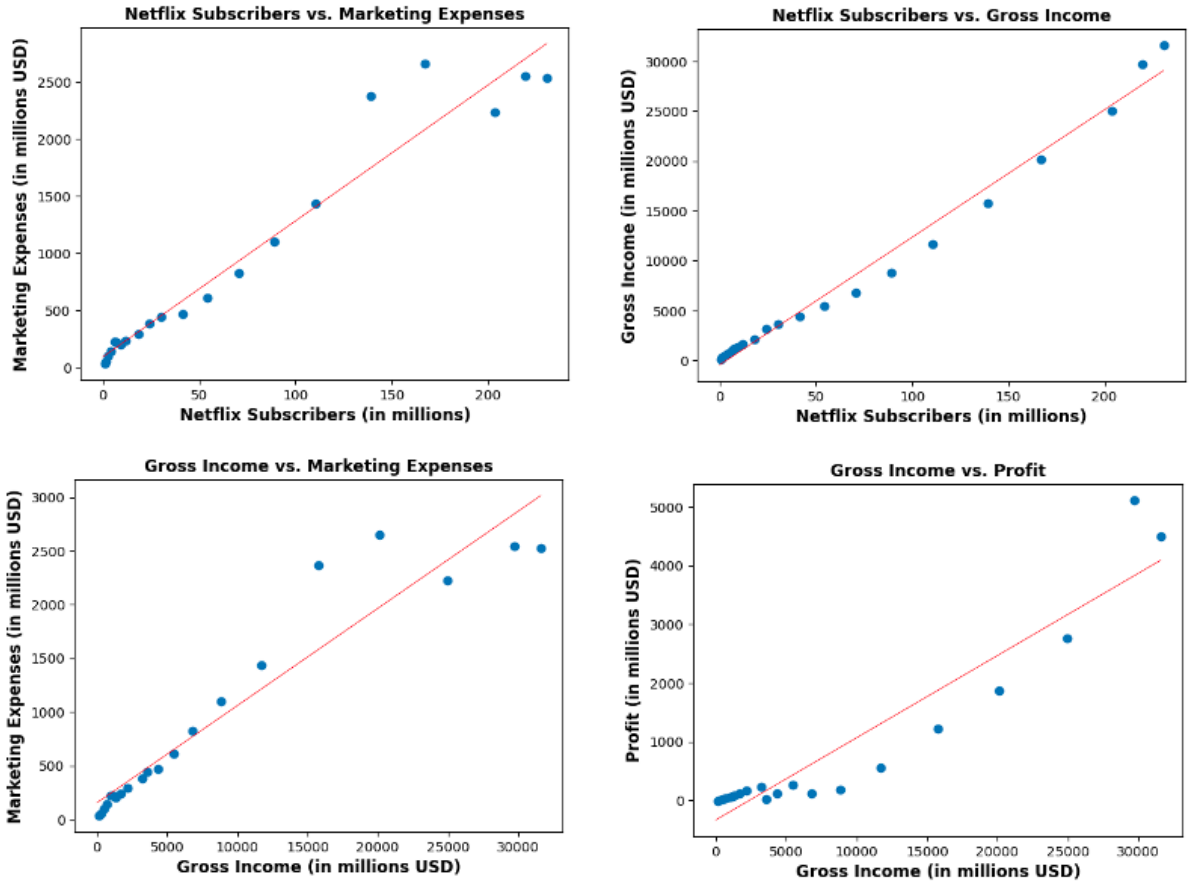

Figure 12. Relationships between Netflix Subscribers, Gross Income, Marketing Expenses, and Profit

To discover the relationship and interpret the mutual influence between variables, we create scatterplots for visualization. From the observation of Figure 12, the points tend to move upwards, an initial inference can be made that there is a positive correlation between these variables. Additionally, the points in Figure 12 also tend to form an upwards straight line. It

seems that there is a linear relationship among these four variables. However, it is worth noting that in Figure 12, there are some points that are far away from the straight line. These outliers may affect the strength of the correlation.

To further investigate the relationship between variables, we calculate the correlation coefficients between these variables to quantify the strength of relationship. We then used a heatmap to display the strength of the relationship (Figure 13). The correlation coefficient between gross income and the numbers of subscribers is the largest, with the darkest color in the chart, indicating the strongest linear relationship between these two variables. Even the lightest-colored area on the heatmap, represented by a value of 0.84. It can be concluded that there is indeed a positive correlation between these variables.

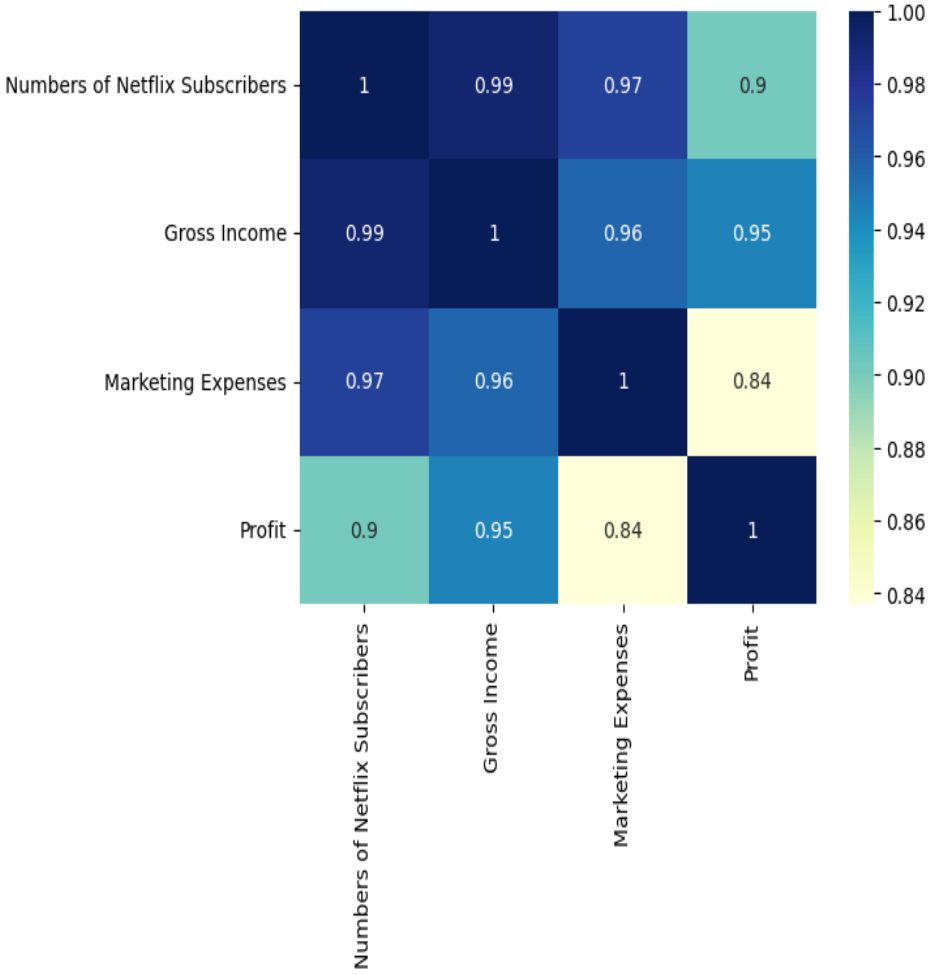

Figure 13. Heatmap Correlation between Netflix Subscribers, Gross Income, Marketing Expenses, and Profit

Discussion and conclusion

For findings about the financial performance, in recent years, competition in streaming services, and an economic downturn have led to a slowdown in subscriber growth and, therefore, a slowdown in gross income growth. Moreover, there have also been some fluctuations in marketing expenses and profits in recent years. Overall, however, in the past 21 years, Netflix's data has shown a relatively steady growth with a positive overall trend line. In addition, based on the current streaming market and our data analysis, it is expected that Netflix's profits will probably continue to steadily increase in the next five years. Based on the previous relationship analysis and the analysis of the geographical distribution of subscribers, Netflix will continue to invest heavily in marketing expenses in order to maintain steady financial growth and user stickiness. Meanwhile, to stimulate the growth of gross income, it will also actively expand its global market, optimize pricing to attract more users, and the Asia-Pacific region is likely to have the largest growth potential in the future.

There are several limitations in our project that need to be addressed. Firstly, some of our findings and conclusions are based on our personal interpretation of the data, which may have some biases. Secondly, the analysis of the relationship between variables is only based on the correlation coefficients, which cannot establish causation. It is possible that other variables or factors may also affect the relationship between the variables.

Additionally, to improve our visualization, interactive features or animations can be incorporated into visualizations, which can enhance the user experience. And based on our project, there are some suggestions for future research:

1. Future studies could consider including more variables to get a more complete picture of Netflix's financial performance.

2. Analyze the reasons for the slowdown in growth in recent years in order to get more effective and accurate predictions of future development trends.

3. Further analyze the relationship between variables and obtain the relationship function via using a regression method.

4. Increase the granularity of the data, such as using quarterly data to gain more insights into trends and patterns of the financial performance of Netflix.

References

Data source: Statista
Software: Tableau, Python
Other sources:

1. Market capitalization of Netflix (NFLX (Company's Market Cap, 2023)
2. Countries Where Netflix is Available (Netflix, 2023)

Appendix

Code snippets:

```
# Check current working directory
import os
print(os.getcwd())
from google.colab import drive
drive.mount('/content/drive') # mount the drive
cwd = '/content/drive/MyDrive/Data_Viz_Lab'  # set the current working directory
os.chdir(cwd)   # change the working directory
print(os.getcwd())

import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from IPython.core.interactiveshell import InteractiveShell
InteractiveShell.ast_node_interactivity = "all"

# Load data set from my drive
from google.colab import drive
drive.mount('/content/drive/')

path_in_drive = "/content/drive/My Drive/Data_Viz_Lab/GroupDataProject.xlsx"
df = pd.read_excel(path_in_drive)
df.head(5)
# check the dimension
df.shape

# Descriptive statistics of the variables
stat= df.describe().loc[:,['Gross Income','Marketing Expenses','Profits','# of Subscribers']]
stat.to_csv('descriptive statistics.csv', index=True, header=True)
```

```
# Create Box Plot of Number of Netflix Subscribers
import matplotlib.pyplot as plt
num_netflix_subscribers = [0.796, 1.41, 2.48, 4.02, 6.15, 7.32, 9.16, 11.89, 18.26, 24.3,
30.36, 41.43, 54.47, 70.83, 89.09, 110.64, 139.25, 167.09, 203.66, 219.7, 230.7]
fig, ax = plt.subplots(figsize=(8, 6))
boxes = ax.boxplot(num_netflix_subscribers, vert=True)

for line in boxes['medians']:
  line.set_color('blue')
# Set the labels for the x-axis and y-axis
ax.set_xlabel('Netflix Subscribers', weight = 'bold')
ax.set_ylabel('Number of Subscribers (in millions)', weight = 'bold')
plt.show()

# Create Box Plots for Gross Income, Marketing Expenses and Profit
gross_income = [150.80, 270.40, 500.60, 682.20, 996.70, 1205.30, 1364.70, 1670.30,
2162.60, 3204.60, 3609.30, 4374.56, 5504.66, 6779.51, 8830.67, 11692.71, 15794.34,
20156.45, 24996.06, 29697.84, 31615.55]
marketing_expenses = [35.78, 49.95, 100.53, 144.56, 225.44, 218.21, 199.71, 237.74, 293.84,
381.27, 439.21, 469.94, 607.19, 824.09, 1097.52, 1436.28, 2369.47, 2652.46, 2228.36,
2545.15, 2530.00]
profit = [-21.00, 7.00, 22.00, 42.00, 49.00, 66.61, 83.03, 115.86, 160.85, 226.13, 17.15,
112.40, 266.80, 122.64, 186.68, 558.93, 1211.24, 1866.92, 2761.40, 5116.23, 4491.92]
# Combine the data into a list of lists
data = [gross_income, marketing_expenses, profit]
# Set the labels for the x-axis
labels = ['Gross Income', 'Marketing Expenses', 'Profit']
# Create the box plot
fig, ax = plt.subplots(figsize=(8, 6))
boxes = ax.boxplot(data, vert=True)
ax.set_xticklabels(labels, rotation=45)
ax.set_xlabel('Netflix Financial Variables')
ax.set_ylabel('$ U.S Dollars (millions)')
for line in boxes['medians']:
  line.set_color('blue')
# Set the font weight of the x and y axis labels
ax.set_xticklabels(labels, rotation=0)
```

```
ax.set_xlabel('Netflix Financial Variables', weight='bold', labelpad=10)
ax.set_ylabel('$ U.S Dollars (millions)', weight='bold')
plt.show()

# Calculate the correlation coefficients
corr = df[['Gross Income','Marketing Expenses','Profits','# of Subscribers']].corr()
corr
corr.to_csv('values.csv', index=True, header=True)

# Draw a heatmap
correlation_ = sns.heatmap(df[['Gross Income','Marketing Expenses','Profits','# of Subscribers
']].corr(),cmap='YlGnBu')
correlation_.set(title="Correlation coefficients between four variables")

# Draw Scatterplots
# Netflix subscribers vs. Gross income
plt.scatter(num_netflix_subscribers, gross_income)
plt.title("Netflix Subscribers vs. Gross Income", weight='bold', fontsize=12)
plt.xlabel("Netflix Subscribers (in millions)", weight='bold', fontsize=12)
plt.ylabel("Gross Income (in millions USD)", weight='bold', fontsize=12)
x = np.linspace(0, max(num_netflix_subscribers), 100)
poly1d_fn = np.poly1d(np.polyfit(num_netflix_subscribers, gross_income, 1))
plt.plot(x, poly1d_fn(x), color='red', linewidth=0.5)
plt.show()

# Netflix subscribers vs. Marketing expenses
plt.scatter(num_netflix_subscribers, marketing_expenses)
plt.title("Netflix Subscribers vs. Marketing Expenses", weight='bold', fontsize=12)
plt.xlabel("Netflix Subscribers (in millions)", weight='bold', fontsize=12)
plt.ylabel("Marketing Expenses (in millions USD)", weight='bold', fontsize=12)
x = np.linspace(0, max(num_netflix_subscribers), 100)
poly1d_fn = np.poly1d(np.polyfit(num_netflix_subscribers, marketing_expenses, 1))
plt.plot(x, poly1d_fn(x), color='red', linewidth=0.5)
plt.show()

# Gross income vs. Marketing expenses
plt.scatter(gross_income, marketing_expenses)
plt.title("Gross Income vs. Marketing Expenses", weight='bold', fontsize=12)
plt.xlabel("Gross Income (in millions USD)", weight='bold', fontsize=12)
plt.ylabel("Marketing Expenses (in millions USD)", weight='bold', fontsize=12)
```

```
x = np.linspace(0, max(gross_income), 100)
poly1d_fn = np.poly1d(np.polyfit(gross_income, marketing_expenses, 1))
plt.plot(x, poly1d_fn(x), color='red', linewidth=0.5)
plt.show()

# Gross income vs. Profit
plt.scatter(gross_income, profit)
plt.title("Gross Income vs. Profit", weight='bold', fontsize=12)
plt.xlabel("Gross Income (in millions USD)", weight='bold', fontsize=12)
plt.ylabel("Profit (in millions USD)", weight='bold', fontsize=12)
x = np.linspace(0, max(gross_income), 100)
poly1d_fn = np.poly1d(np.polyfit(gross_income, profit, 1))
plt.plot(x, poly1d_fn(x), color='red', linewidth=0.5)
plt.show()
```