

Assignment 1: Visualisation in Applied Data Science 1

Sam Jacob

SRN: 22073281

Msc Data Science (Sandwich)

Repository Link:

https://github.com/Samjacob007/ADS_Assignment/tree/master

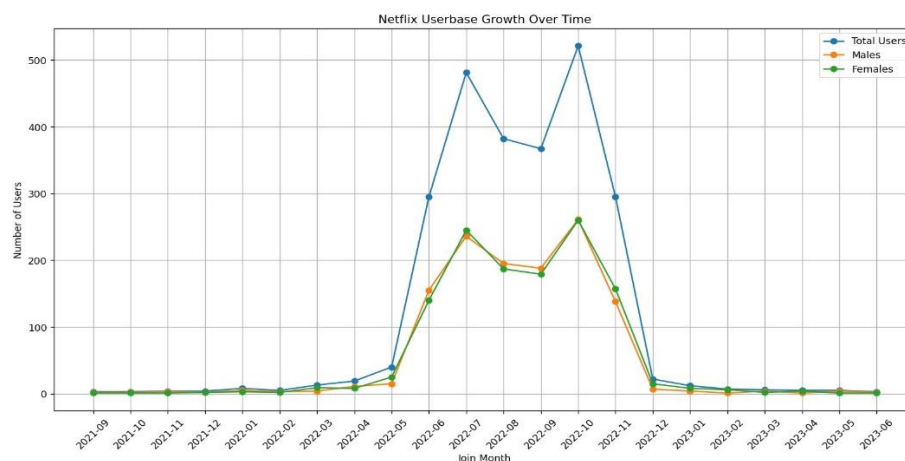
Introduction:

Visualization techniques fall into these three groups. All of the datasets originate from publicly accessible databases. Every data point is explained in great depth in each section. Reading datasets is done with the pandas tool, while plotting graphs is done with the "pyplot" function.

The data was gathered from the "Kaggle" open-source, free database. During my research, I came across an enormous amount of information about a representative sample of Netflix users. The dataset is packed with details like user subscriptions, revenue, account specifics, and user activity. Each row is dedicated to a unique user, identified by their User ID. The dataset includes crucial information such as subscription type (Basic, Standard, or Premium), monthly revenue generated from their subscription, the date they hopped on the Netflix bandwagon (Join Date), the date of their latest payment (Last Payment Date), and even the country they call home. With such rich information, visualizing key aspects of this dataset could provide insightful glimpses into Netflix user trends. The dataset was collected from the website: <https://www.kaggle.com/datasets/arnavsmayan/netflix-userbase-dataset>

LINE PLOT

Line plot demonstrates the Netflix Userbase Growth Over time. (Total users, Males, Females) throughout the years. The definition of the X-axis is the date when subscribers joined in Netflix (Join Month). Y-axis represents the number of users joined in every month on Netflix. The year is taken from 2021-09 to 2023-06.

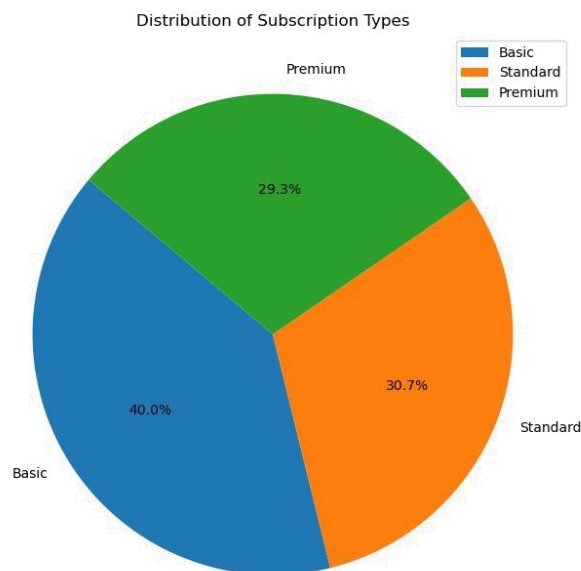


Blue line represents total users, orange line represents male users and green line represents female users. The number of users is constant until 2022 may and afterwards there was a sudden increase in the number. The number of users got fluctuated over some period and there was a sudden decrease by 2022 December. Again, the number became constant till 2023 June.

Line plot is best option to visualise this data. We got a proper visualisation of the users. The graph will help the Netflix to understand their user acquisition. It will help to identify the pattern of user growth. Thereby they can manage their marketing by proper content creation. They will also get a proper overview on the data on males and female perspectives.

PIE CHART

Pie chart demonstrates the distribution of subscription types from 2021 September to 2023 June. The subscription types include Premium, Standard and Basic. Each slice of the pie chart is labelled corresponding their subscription type which will give a proper clarity.

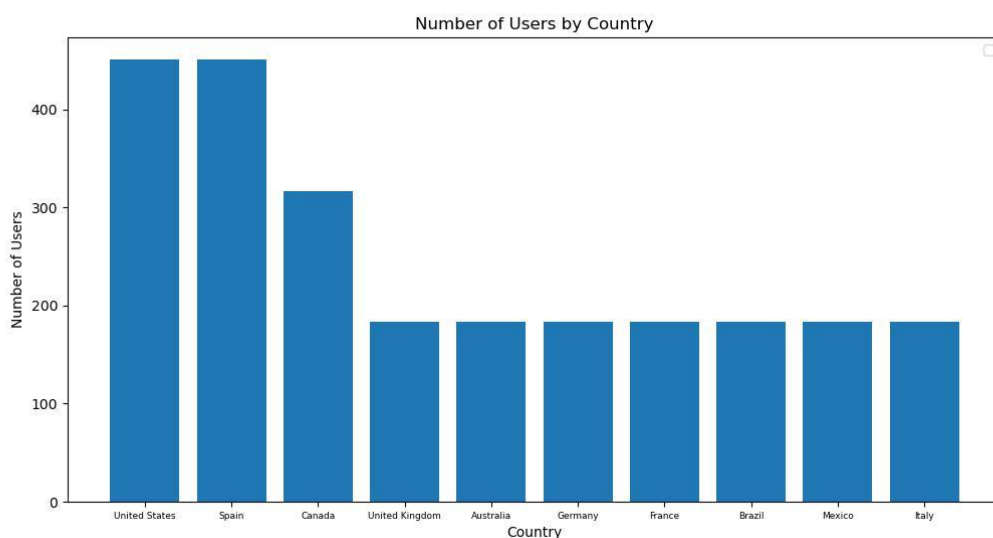


Pie chart will help to know the proper proportion of the subscription types. It will get a clear understanding how users are distributed among these three subscription types. Pie chart will make a quick understanding about the data. It highlights the dominant subscription

category and allows for easy comparison of proportions. Most of the people have basic subscription types i.e., 40% followed by standard with 30.7% and Premium by 29.3%.

BAR CHART

Bar chart demonstrates number of users in country (United States, Spain, Canada, United Kingdom, Australia, Germany, France, Brazil, Mexico, Italy). X axis represent the countries and Y axis represent number of users



Each bar on the graph corresponds to a country, offering a visual comparison of user counts. Most of the users are from United states, Spain and Canada followed by other countries. Other six countries have almost same number of users.

The bar graph visually emphasizes countries with notable Netflix user bases, identifying key markets for the streaming service. Variations in bar heights reveal insights into the popularity of Netflix across different geographical regions. Understanding the geographical distribution can inform decisions related to content localization, regional partnerships, and user engagement initiatives. The bar graph effectively communicates the global reach of Netflix users.