Netflix - Interactive Visualisation App

P4

Datavisualisering (DV) (DV3) Exam hand-in

Flemming Steno Kjær

Aalborg University
The technical faculty of Computer Science
Selma Lagerløfts Vej 300
DK-9220 Aalborg Øst

AALBORG
UNIVERSITY

STUDENT REPORT

Faculty of Computer Science
Data science
Selma Lagerløfs Vej 300
9220 Aalborg Øst
http://cs.aau.dk

Theme:

TBA

Project:

DV - Exam

Period:

February 2022 - May 2022

Participants:

Flemming Steno Kjær

Page Number: 11 + front page
Project finished: February 24, 2022

Contents

1 introduction

The interactive visualization program, which will be shown through the project, can be found on [1], and the download for the report will also be included.

1.1 Description of Used Data

For this project, a data set containing over 500 variables has been utilized to describe the use of the well knows Streaming service Netflix[2]. The subscription cost of different services they provide for many countries to the number of accessible shows in each country and their types. These attributes will be utilized through the rapport to get an idea of the pricing to the number of available shows for each country. How the Data set is structured can be seen in table one and elaborated. There are three general themes in the dataset the regions, the accessible library for either tv-shows or movies, and alternative types of Subscriptions for different price ranges.

Country & Country_Code
These categorical attributes explain which country the respected rows information belongs to.

Library
These quantitative attributes explain the amount of either TV shows or movies accessible for each country and the total amount of both.

Subscription
The subscription is spread for three rows, each containing the subscription quantitative price in the respected country.

	Country_code	Country	Total.Library.Size	No..of.TV.Shows	No..of.Movies	Cost.Per.Month...Basic....	Cost.Per.Month...Standard....	Cost.Per.Month...Premium....
1	ar	Argentina	4760	3154	1606	3.74	6.30	9.26
2	au	Australia	6114	4050	2064	7.84	12.12	16.39
3	at	Austria	5640	3779	1861	9.03	14.67	20.32
4	be	Belgium	4990	3374	1616	10.16	15.24	20.32
5	bo	Bolivia	4991	3155	1836	7.99	10.99	13.99
6	br	Brazil	4972	3162	1810	4.61	7.11	9.96
7	bg	Bulgaria	6797	4819	1978	9.03	11.29	13.54
8	ca	Canada	6239	4311	1928	7.91	11.87	15.03

Figure 1

1.2 Description of Tasks

With the available data, it first and foremost leaves the question, if there is any correlation between the prices of the shows and the number of available shows for each country. The added benefits from the increased cost of the subscriptions are worth the money for each country.

Livestreaming services over the last decade have boomed in growth, to the point that it is questionable that we still pay for our TV subscription on the side of the live streaming services such as Netflix. However, because it is such a booming market, Netflix is not the sole distributor of this service. Over the last few years, there has been an ever increasing amount of new streaming services, each demanding a subscription fee and withholding

exclusivity rights over specific shows, so the customers have an exclusive service to provide the desired entertainment.

The broader the range of movies the provider can provide to the customer, the greater the customer will have an interest.

It is a balancing act of being affordable enough for the customer while still being profitable for the provider itself[3].

1.3 Problem Statement

How can interactive visualization be utilized to determine the balancing act of attracting customers while still being profitable?

1. The price for the subscription to the available shows worth - Which country pays the most and least for the amount shown in the library?

2. Are the prices for the different subscriptions unequal distributed in different countries?

2 Analysis

In this section, both the program and the visualizations methods will be analyzed utilizing the knowledge from our Data Visualisation course primarily.

2.1 The Application

The application is built upon R-shiny in R, which were greatly utilized to create plots and at the same time intractable, and not just still images[4].

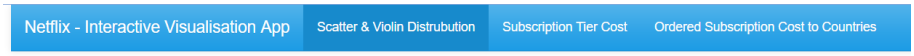

Figure 2

At the top of the application is there a tab menu with the application's title, which you can use to select multiple options for different tabs of illustrations.

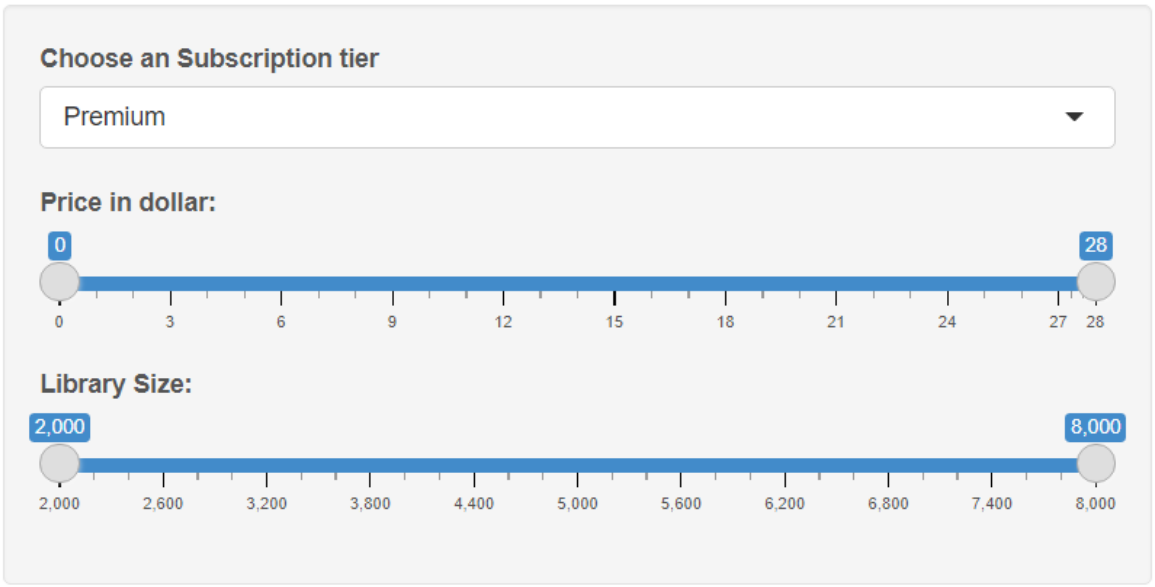

Figure 3

The very first page is the Scatter & Violin distribution. The right side contains multiple options that influence both the scatter plot and the violin plot. The first option is to choose a wished subscription tier from Basic, Standard, and Premium. Each of these subscription tiers benefits the users, but they are not utilized and described in this data set. The next option to manipulate the graphs is the double slider for the price range. It can manipulate the lower bounds and the upper bounds of the price range you wish to look at. This can be utilized if you want to look at a specific price range and are not interested in all the countries' subscriptions prices. Alternatively, the third option allows the user to manipulate the graphs' total library size of TV shows and movies. This one also has a double slider to move the upper and lower bound.

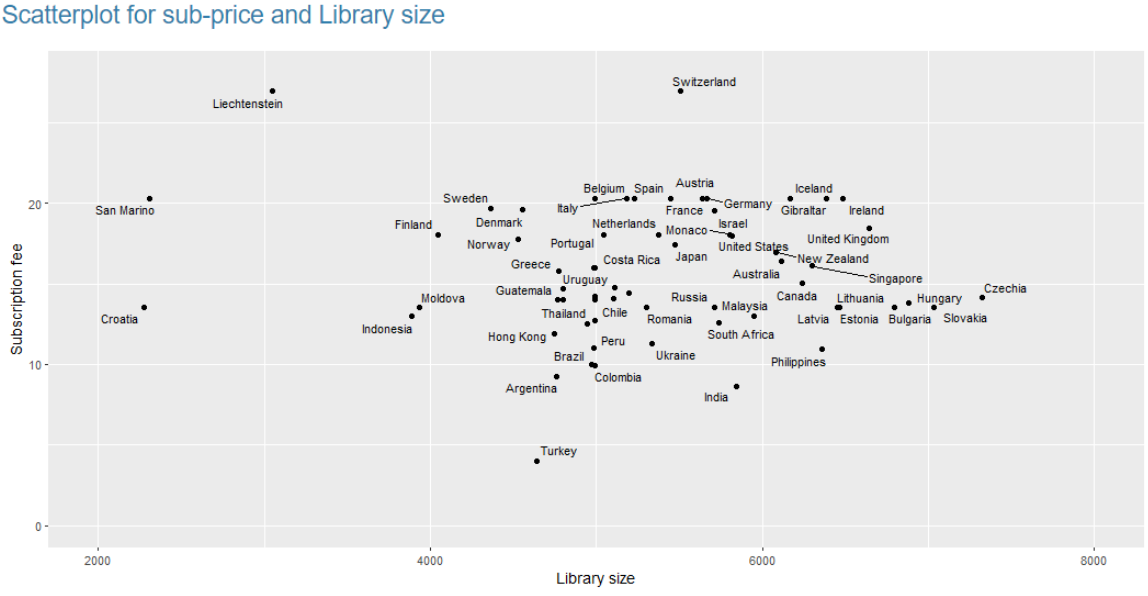

Figure 4

This page contains two graphs. The first is a scatter plot representing how the countries are distributed in the fee they have to pay for Netflix and the library size they end up getting access to. On the X-axis is the library size, and on the y-axis is the subscription fee. The dots in the scatter points in the graphs are the countries noted by their name near the dot they represent. A detail worth noticing is that the names do not overlap each other and become a blur. The scatter plot is to first and foremost get a first impression of the distribution. If the first graph looked into was a t-distribution, it would be hard to interpret any patterns. And some methods cannot be utilized since the two selected attributes are quantitative values.

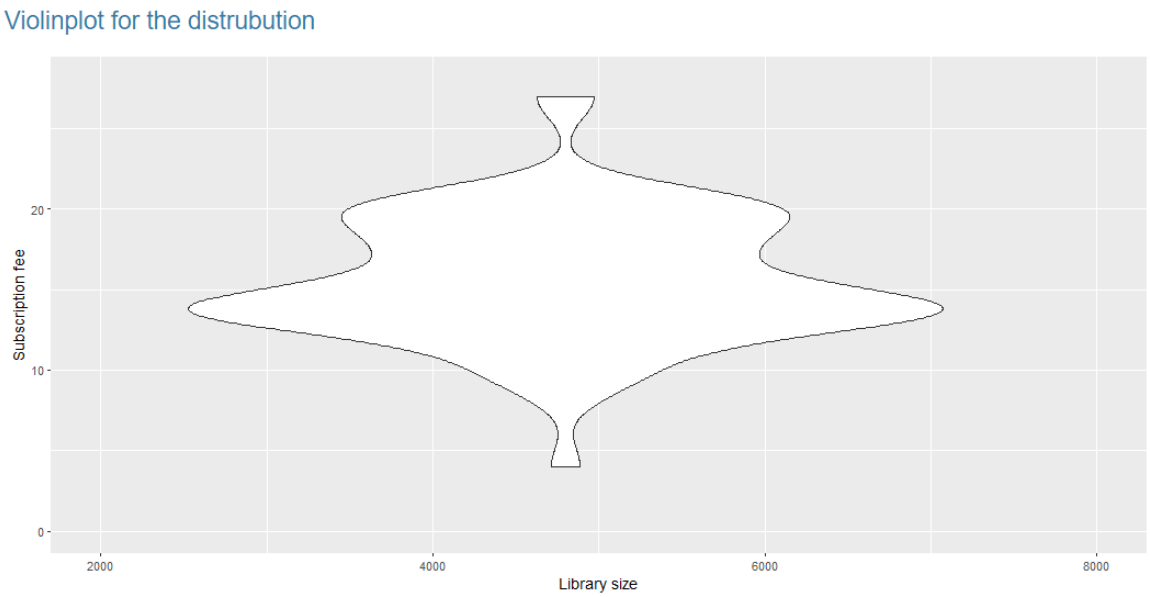

Figure 5

The Violin plot represents the distribution by becoming broader where multiple countries are clustered. The catch is that the violin only looks at the number of countries in a price range, not the library size. So having them both simulate the same values can contribute to a better understanding of how they are spread for each tier.

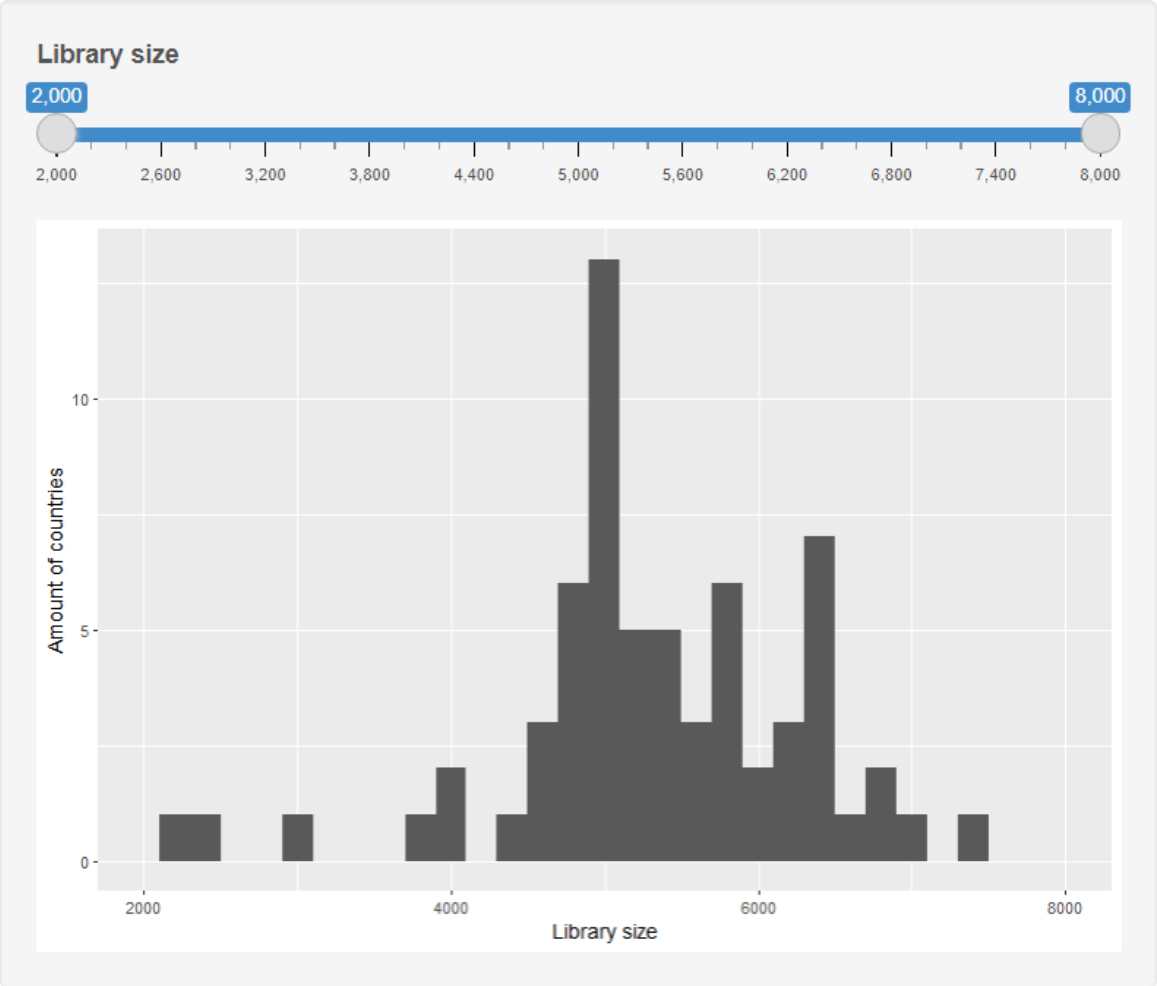

Figure 6

The second page, 'Subscription Tier Cost, holds three graphs. The graph to the left is a histogram. It does not hold much value in this case, further elaborated in the usability test. The histogram can only be used to describe a single parameter, which is the Library size. It has a pre-selected bin-with of 200. If we look at the two other graphs, they represent a bar chart for the subscription tier. Moreover, the Bar chart is not confused with a histogram since the bar chart utilizes both Quantitative and categorical data. The color chosen is the representative, for they are not the same value in each bar. they each are their own category we compare to each other, which are the subscription tiers and their prices. These can also be manipulated by choosing some pre-selected countries of interest from the tab. Denmark is chosen since it is what we have to pay for it. And then there is Turkey. As we can see in the last scatter plot, Turkey had the cheapest access to Netflix in the whole dataset. Likewise, Liechtenstein was the most expensive. Moreover, South Korea was a random pick to get an idea of not just the borders of the price range but also another random country to see if there are any similarities.

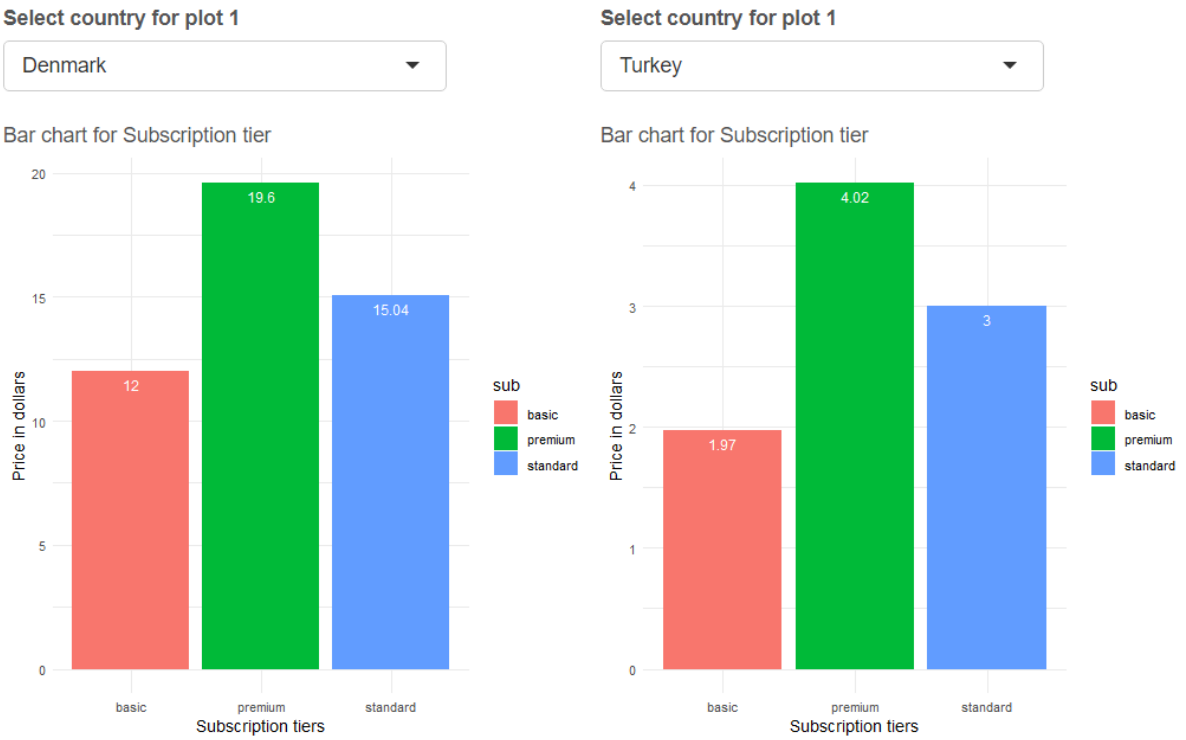

Figure 7

These bar charts can not only be utilized to compare singular countries' subscription tiers but because they are by the side of another bar chart containing the same information, they can also be used to compare countries' tiers to each other. Furthermore, it is even clearer which tier is what by the color.

The bar chart and the histogram hold an advantage for comparison since they utilize easily understandable length, which is the tallest or lowest of the bars/bins.

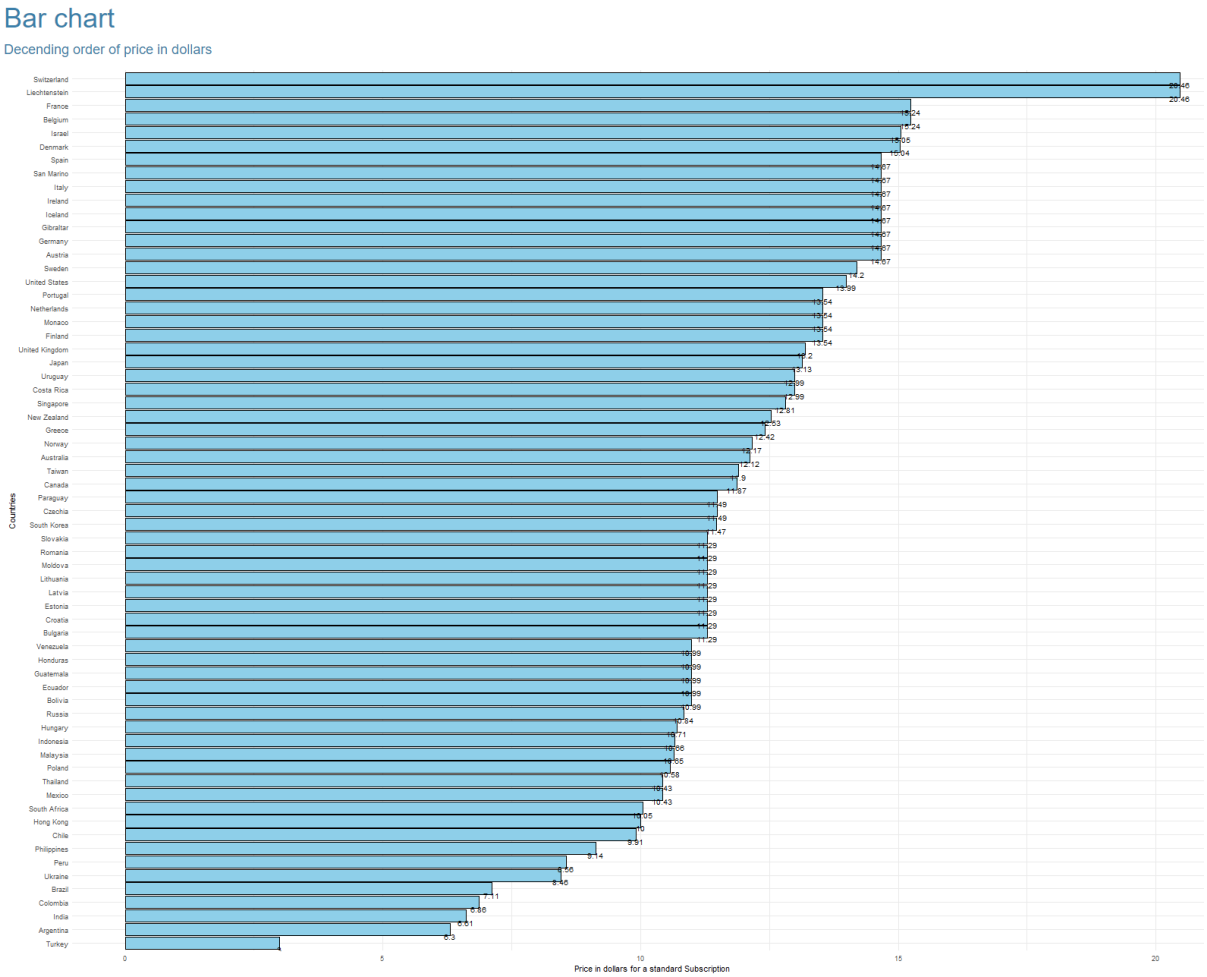

Figure 8

The third page, 'Ordered Subscription Cost to Countries', contains a singular list of ordered countries after the standard subscription fee in dollars. It is rather large to contain all the countries the data set had. The bar chart has been tilted horizontally for better readability for countries' names. The typical user has it easier scrolling up and down if the person does not contain a large enough screen. This does, though, also promote some problems which will be discussed in the usability test. The method of order could have been done to an even more extreme by adding a hue that gradually changed the hue. It was chosen not to include this and add numbers as an additional label on the bars of their value since it would become over stimulating for such a large graph.

3 Ethical Concerns

This section will discuss several ethical concerns about the project and the visualizations.

3.1 Social Relevance

The vast majority of the users are unaware of how much Netflix chooses to vary the amount of content they deliver to several countries and the money they fee for their service. Especially nowadays, where more and more choose to circumvent these circumstances by

utilizing VPN tools to bypass regions and buy these services across borders. This project opens for the discussion of why they choose to have such high prices in several countries across the world with even less content to enjoy. Alternatively, it is because of increased competition in specific regions where they try to moderate their pricing to be favored compared to their competition[5].

Service Tiers

Their service tier gives few bonuses, but it is a conclusive factor for some. Else they will not be interested. In Denmark, there are only three things that differentiate: the video quality and bit-rate household member count. The Basic subscription gives full access to their catalog of movies and TV shows, but you can only watch in 480p and with what they define as "good" bit-rate, which is the phone quality screen a decade ago and give access to a singular user. The Standard subscription gives 1080p, which is acceptable for laptop use and phones for the most part with "better" bit-rate and access to two users in a household. However, for those who want to look at the shows and movies on the TV of modern standards, they would likely need to pay for the Premium subscription for getting access to 4k video quality with the "best" bit-rate" for four users in a household[6].

3.2 Visualisations

Bar Chart for Subscription Tiers

They can be hard to grasp since their peak all go to the top of the graph whenever you change them and compare them with a side-by-side comparison. If you had, for example, Liechtenstein next to the Turkish graph and you did not look at the y-axis numbers nor the labeled price cost on the bar itself, they would look like they are equal. Though if the graph did not scale, it would be challenging for the Turkish to see the difference between its subscription tiers. For example, the danish Basic is closer to Premium in price than the Turkish Basic to their Premium. So it is essentially not good at comparing their values, only their distribution.

Violin Plot

It was mentioned that it was not a suitable parameter alone to look at the distribution of the countries unless you wanted to look at the countries' fee distribution alone without looking at the library size. And then still, even if they are both together, they are hard to determine how the countries are distributed at a quick glance. The violin plot, together with the scatter plot, would not be an excellent method to present the distribution to the general public because it is harder to understand when it is not just the size but shape of the violin plot. It has to be used with the scatter plot, which is essentially a cluster of dots for the untrained eye.

4 Usability Test

in order to check how the program performs, it will be put to a User test where the user will be presented with some objective they have to figure out by utilizing the program. Afterwords it will be concluded how well it lives up to the objectives.

4.1 User Objectives

1. Which two countries have the least amount of content in their Netflix library?
2. What is the mean for library size?
3. How is Denmark Premium subscription pricing compared to other countries?
4. What is the mean pricing for a Standard subscription ?
5. What is the value and country which pays the most?

4.2 Results

1. Slightly confusing, was not written in the tab page names, nor headline for any graph.24
2. Straightforward and clear but boring image.6
3. Quite clear and easy to find, but double-checked on multiple graphs.7
4. Very clear but not totally certain.5
5. extremely easy and straightforward.8

4.3 Conclusion

To conclude, The project gives what it was meant to do, but it was not entertaining to play around with for the most part. This could be changed with Plotly and utilizing linked brushing, which could make it more fun to play around with, in general, more interactive and not just tabs and sliders.

The Histogram6 was meant to interact with the bar charts to highlight the bin with the countries' library size. While rest will be grayed out, and the bin would also have a label of the library size count. An alternative that would be possible with Plotly to make it interactive with any country if there was a linked plot to it.

Not the whole data set was utilized. I did not find the specific amount of movies and TV shows to have any exciting influence on the other attributes alone.

5 Bibliography

[1] Flemming Steno Kjær. *GitHub for Project.*
URL: `https://github.com/flemmingsteno/Netflix---Interactive-Visualisation-App` (visited on 2022).

[2] CC0: Public Domain. *Netflix subscription fee in different countries.*
URL: `https://www.kaggle.com/prasertk/netflix-subscription-price-in-different-countries` (visited on 2022).

[3] Patrick Seitz. *Is Netflix Stock A Buy After Video Streamer's Q4 Earnings Report?*
URL: `https://www.investors.com/research/netflix-stock-buy-now/` (visited on 2022).

[4] RStudio Inc. *R Shiny.*
URL: `https://shiny.rstudio.com/gallery/` (visited on 2022).

[5] REBECCA MOODY. *Which countries pay the most and least for Netflix?* URL: `https://www.comparitech.com/blog/vpn-privacy/countries-netflix-cost/` (visited on 2022).

[6] Netflix. *Netflix subscription.*
URL: `https://www.netflix.com/signup/planform` (visited on 2022).