Executive Summary

Background and intent

Background

Netflix, a widely popular streaming service, offering services in 190 countries globally has played a big role in our mental health, especially when the COVID-19 pandemic began shutting down the world. In 2020 and 2021, we generally spent less waking time with people outside our own households. With less time out of the house, Netflix subscriber base increased by multiple folds with the world united with one common pastime - binge watching on Netflix content, a way of self-care to relax and de-stress.

Motivation

In October 2020, Netflix revised its pricing. Further on January 14, 2022, Netflix revised its pricing again for some countries. Netflix Price is different for each country. We wanted to investigate Netflix Pricing related data to see what factors may have helped Netflix's marketing determine pricing for each country. We wanted to see if there is correlation between Pricing with Netflix Library size, and whether there are key country development factors unique to each country which correlate to Netflix Pricing.

We wanted to stretch this analysis further by analyzing if Happiness Score, compiled by Gallup World Poll data (which uses global survey data to report how people evaluate their own lives in more than 150 countries) has any correlation with Netflix Price!

Questions

To address these motivations, our team investigated how key country development factors - Gross Domestic Product ("GDP"), Population Size by Country ("POP"), Consumer Price Index and Inflation Rate ("Inflation"), collectively known as ("Economic Factors"), Happiness Score and Netflix Library Size compare against Netflix Pricing. We hope to discover meaningful insights on association, if any, between Netflix Pricing, Library Size together with Economic Factors and Happiness Score.

> **How will the analysis help us?**
>
> The results will help us understand if there are proxy features which could be considered in building a model to predict Netflix Price for countries where Netflix has no presence in.

> ***" Analyzing and discovering Happiness Score and Economic factors which have a correlation with Netflix Pricing may help build Netflix Price prediction model "***

Objectives

Our main objectives are to:

[1] Provide answers to the following questions:

- How much are the top 10 countries paying for Netflix as compared to the bottom 10 countries? If the subscriber countries are paying more, are they rewarded with a larger Netflix Library Size?
- What is the average price per region, and which region paid the most for Netflix?
- Does GDP, inflation and Population Size affect Netflix Price?
- Which region has the highest Happiness Score? For the region with highest happiness score, does this region have a higher or lower Netflix Price? Further, are there any outliers noted which have potential bias to the analysis performed?

[2] Verify the pre-project assumptions we have:

- Higher GDP will lead to a higher Netflix Price
- Higher Happiness Score will lead to a higher Netflix Price
- Higher POP will lead to higher Netflix Price

Executive Summary

Findings and recommendations

Key findings

- Based on analysis performed, we noted a weak correlation between Netflix Price and Netflix Library Size.
- Prices differ by region and country. Highest Netflix Price noted in Switzerland – US$12.17. Europe has the highest blended average Netflix Price amounting to US$ 8.58.
- Happiness score differ from one country to another. Countries with higher GDP were observed to have higher Happiness Scores and higher Netflix Prices.
- There were a couple of outliers noted for each region in terms of Happiness Score, this will affect the correlation coefficient we calculated between Happiness Score and Netflix Price.
- We looked at the spread and correlation, visualizing the correlation between the various features using an interactive Altair plot hosted on streamlit.
- The chart illustrates the team's thought process of exploring the correlation structure of our dataset using two connected subplots – (i) correlation heatmap and a (i) 2d histogram showing density of points.
- **Some positive trends** were identified, showing that (A) countries with higher GDP per capita have a higher Netflix Price, (B) Higher Happiness Score has a higher Netflix Price. A **weak negative trend** was noted showing that countries with higher populations have lower Netflix Prices, higher inflation has lower Netflix Price.

Limitations

- The source for the Netflix Pricing dataset, is not a complete representation of Netflix pricing in all the countries Netflix offers services. Our dataset only contains 88 countries out of 190 countries Netflix provides streaming services to.
- However, we can use his sample data to make inferences of the population.

Next steps

- Explore outliers discovered in the World Bank Indicators dataset and determine how to address them.
- For the outliers noted in the Happiness dataset, hold discussion with the dataset originator on reasons why some countries score significantly higher / lower than other countries in the same region.
- Due to time constraint and limited Netflix Pricing information available using desktop research, our observational study was performed for data spanned over one year – 2021. The narrow scope may be a limitation on the generalization of the study's findings. Analysis could be done over a longer time period. Stretching across a few years might provide better insights to the correlation between various features.
- On a quarterly / yearly basis, update the Netflix Pricing based on latest price revision and observe if there is any change to the understanding and analysis performed in this report.
- There may be a correlation between popular actors and Netflix Pricing which we did not explore in stage 1 and may want to consider in stage 2 of our project.
- Build Netflix pricing prediction machine learning model using regression and classification.

Methodology

Data Sources

Name	Description	Size	Access
Netflix Pricing for each Country	The dataset has features as follows: 1. Country, 2. Subscription fee, 3. Region, 4. Series library size, 5. Movie library size, 6. Total library size, 7. Subscription fee per title. This dataset enables us to analyze library size vs monthly cost in each part of the world - Bang for buck analysis.	~1.0MB (15,480 rows)	CSV download from comparitech: Which countries pay the most and least for Netflix?
World Happiness Report (from 2015 to 2021)	The dataset has features as follows: 1.GDP per capita, 2. Healthy Life Expectancy, 3.Social support, 4.Freedom to make life choices, 5.Generosity, 6.Corruption Perception. This dataset enables us to understand which features have a higher influence on happiness score.	140.8kB (7 datasets with total 1084 rows)	CSV download from Kaggle: World Happiness Report up to 2022
World Development Indicators	The World Development Indicators (WDI) are the World Bank's first collection of development indicators, compiled from data from officially recognized international sources. The world bank provides the latest and most accurate global development data, including country, regional and global data. We have decided to use it population and inflation data for our analysis.	280.8MB (up to 266 countries data)	CSV download from The World Bank: Data Bank | World Bank Indicators

[3] [4] [5]

We used 10 datasets to create a final dataframe for our analysis

Note [3], [4], [5]: Refer to page 11 of our observational study report for the sources and citation.

Methodology

Data Manipulation

Data Consistency

- Renamed the columns in a more readable manner and ensured that the columns have a consistent name in different datasets.
- Some countries and regions in each dataset had different names. For example, Countries - Hong Kong, Taiwan in the Netflix Pricing were documented as Hong Kong S.A.R. of China and Taiwan Province of China in the Happiness dataset. A total of ten changes to country names made. We queried each country via a regex function accompanied by an .isin() function with checks using .query() and renamed the differences noted to ensure we do not lose any essential data.

Data Cleaning

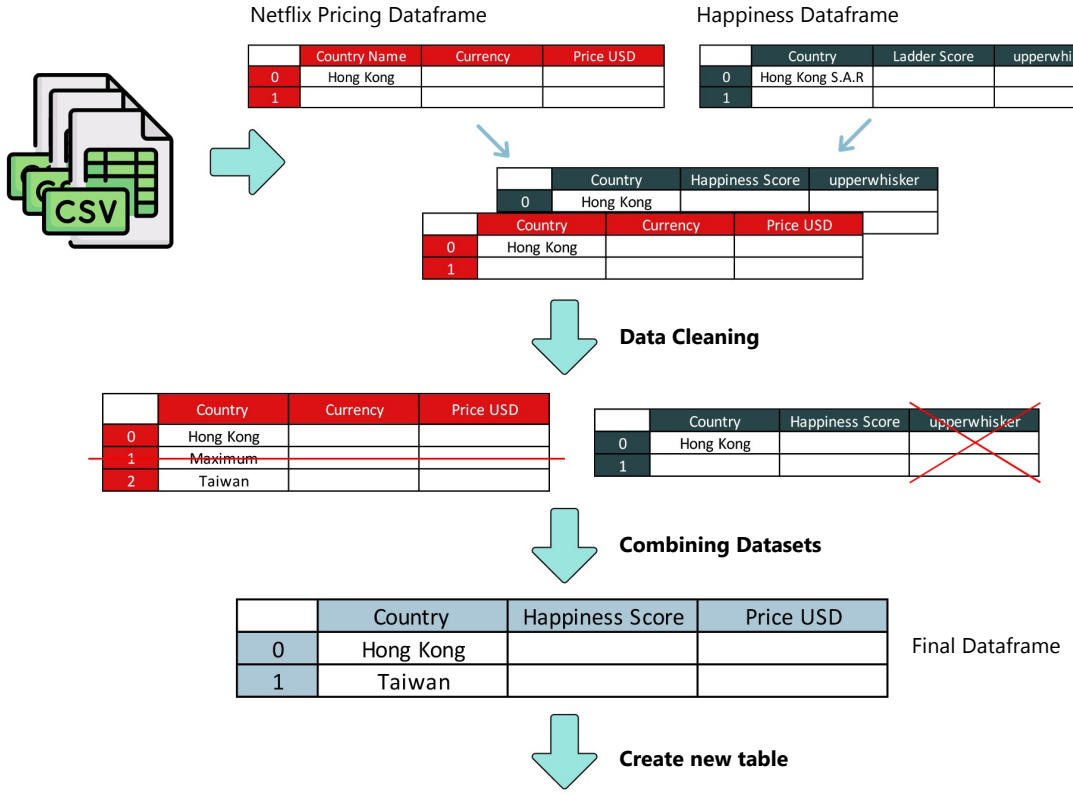
- Removed inaccurate and noisy values in the dataset, such as min, average and max from the country column.
- After we had a good understanding of the datasets, we removed the unnecessary columns, namely - 'Dystopia + residual', 'lowerwhisker', 'upperwhisker'.

Combining Datasets

- After cleaning each dataset, we performed a merge with all the required columns, creating a new usable dataset used for our analysis.
- Populated the lost data values in the 'Region' column after performing a merge between the various datasets.

Data Calculation

- Create new table with maximum, mean, and minimum values over the requested axis by region to gain insights.



Region	No. of countries	Mean price	Minimum price	Median Price	Maximum price
Western Europe	23	8.58	7.57	7.76	12.17
North America and ANZ	5	8.18	7.15	7.43	9.99
Middle East and North Africa	19	7.59	2.48	7.99	9.33

Analysis and Visualizations

Netflix Price and Library by Top 10, Bottom 10 Countries

Analysis:

One of the questions we had was to determine the amount top 10 countries paying for Netflix Price as compared to the bottom 10 countries. And, for the countries that were paying more, are they getting a bigger Netflix Library Size?

We present in the charts on the right the top and bottom 10 countries by price and looked for possible association between price and library size.

Findings:

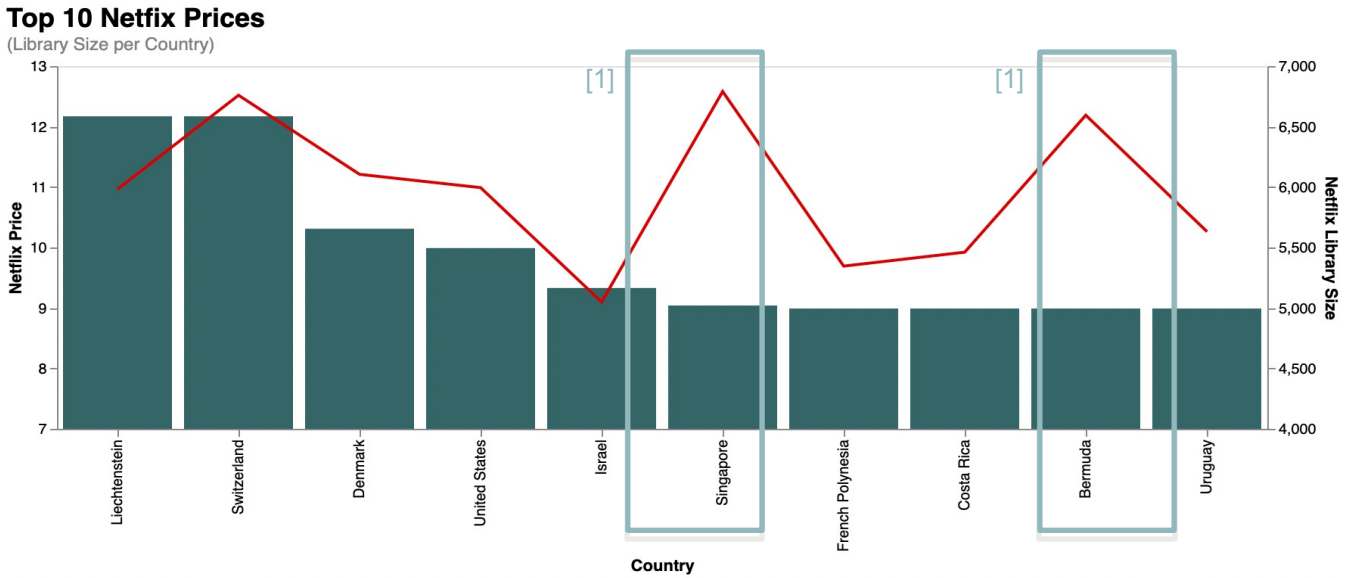
Countries like Singapore and Bermuda in the Top 10 chart, are paying a lower Netflix Price as compared to the other countries in Top 10 but receiving a larger library size.

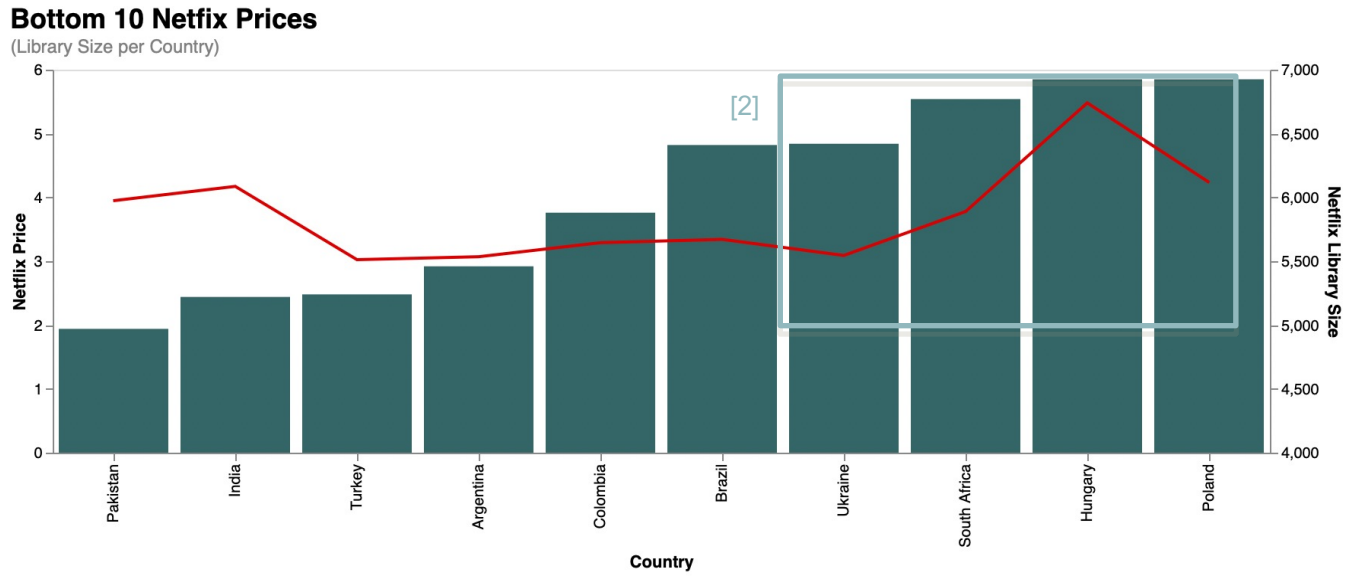
Singapore is the biggest winner, with a Netflix price of US$8.52, it is paying less than Switzerland but getting a larger Netflix Library Size as compared to Switzerland. [1]

With regards to the bottom 10 countries in terms of Netflix Price, we noted a gradual increase in content as the price increases. [2]

Insight:

As per the above analysis, we do not observe a clear relationship between Netflix Price and Netflix Library Size.





Analysis and Visualizations

Netflix Price By Region Analysis

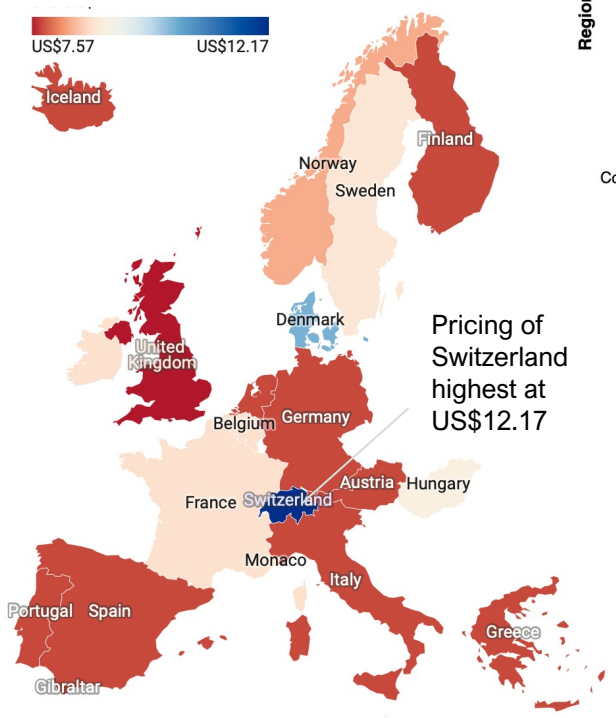
Analysis

We were not convinced after our observational study performed on Top 10 and Bottom Library Size that there was an association, if any, between Netflix Pricing and Library Size. We think there are other variables which have a correlation with Netflix Pricing.

We performed a regional analysis on Netflix Price and noted that on average, countries in Western Europe are paying more for Netflix than the rest of the regions.

From the Region Analysis of Netflix Price vs Happiness score, we noted that Western European Countries pay a higher price and have higher happiness scores when compared to other regions.

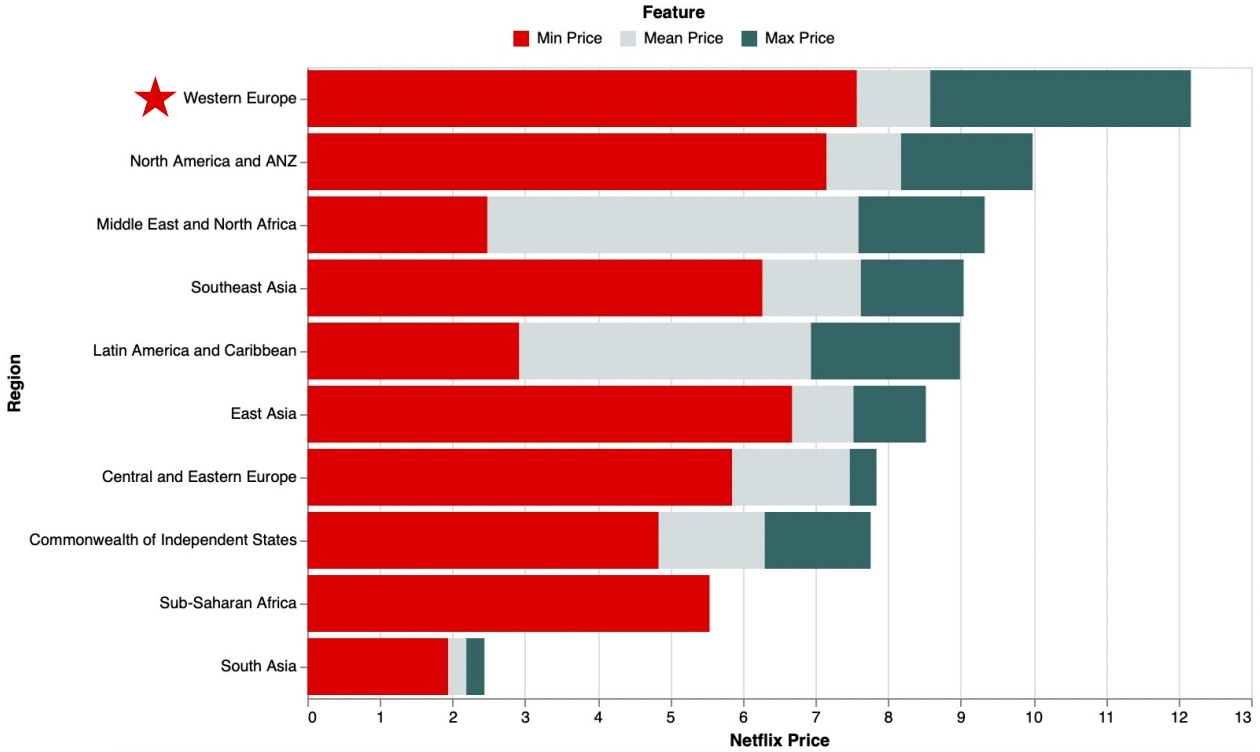
Average Netflix Price in Western European countries was notably higher when compared to other regions. We were unable to decode any strong correlation between inflation, population, and Netflix Pricing, which we will discuss further in following pages of the report.



Pricing of Switzerland highest at US$12.17

Visualized with datawrapper | Data source: Netflix dataset raw data processed | Created on 2022-10-13

Netflix Pricing Across Regions
(Minimum, Mean, Maximum Price)



Region	No. of countries	Mean price	Minimum price	Median Price	Maximum price
Western Europe	23	8.58	7.57	7.76	12.17
North America and ANZ	5	8.18	7.15	7.43	9.99
Middle East and North Africa	19	7.59	2.48	7.99	9.33
Southeast Asia	5	7.62	6.27	7.56	9.04
Latin America and Caribbean	14	6.94	2.92	7.99	8.99
Central and Eastern Europe	13	7.47	5.85	7.76	7.84
Commonwealth of Independent States	2	6.30	4.84	6.30	7.76
East Asia	4	7.52	6.68	7.45	8.52
Sub-Saharan Africa	1	5.54	5.54	5.54	5.54
South Asia	2	2.19	1.94	2.19	2.44
Global	**88**	**7.59**	**1.94**	**7.76**	**12.17**

Analysis and Visualizations

Netflix Price compared to GDP and Happiness Score

Analysis:

One of the questions we had was to address which countries/regions are paying the most for Netflix. We present in the charts on the right the top region by price and happiness score, Western Europe. In the chart in the upper right, we observe Switzerland as the country with the highest Happiness Score (7.571) and Netflix Price ($USD 12.17).

As noted in our pre-project assumptions, we confirm here that higher happiness scores are associated with higher Netflix prices. However, we are unable to observe a clear relationship between Logged GDP per capita, Netflix Price and Happiness Scores. We will examine this relationship with a correlation analysis on page 10 of this report.

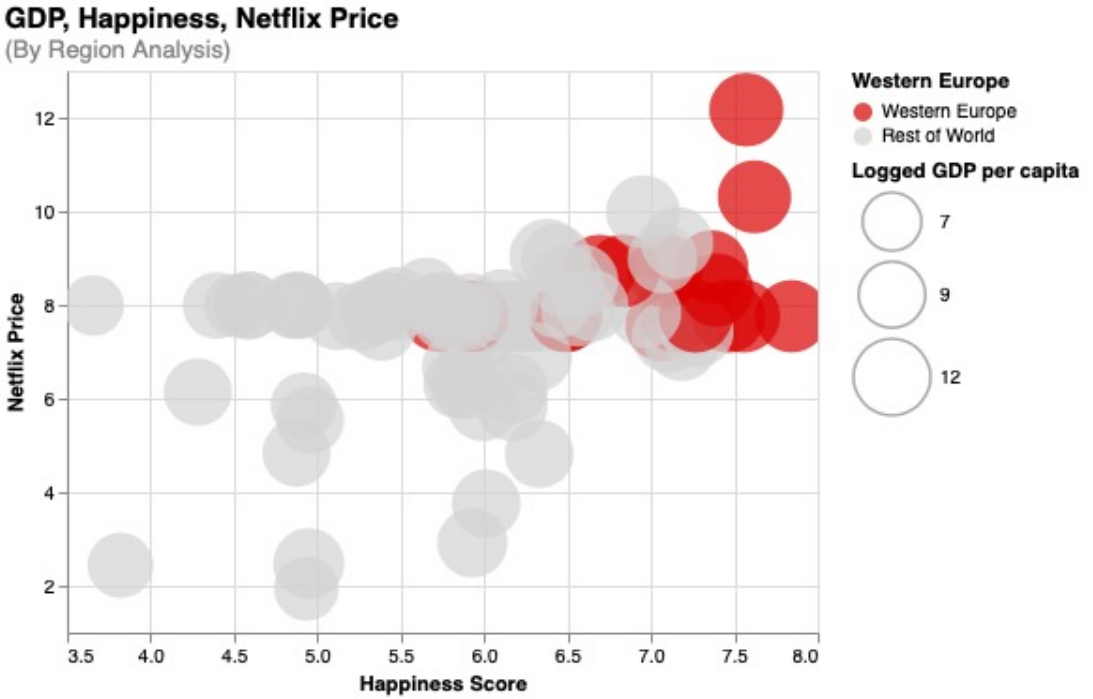
Findings:

Higher GDP and higher Happiness Score are both associated with higher Netflix prices. We also found a positive correlation between Netflix Price and Happiness score. In other words, countries with higher GDP have higher happiness scores.

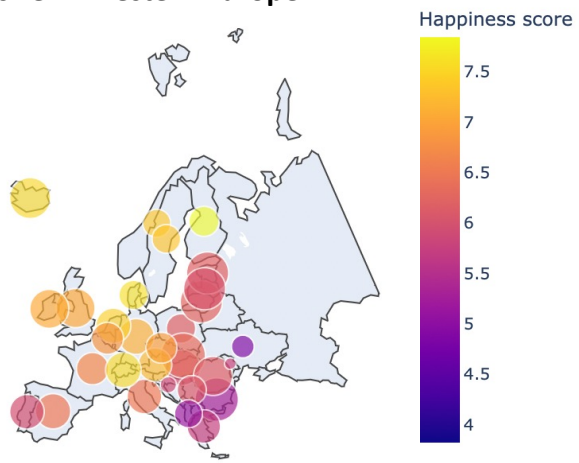
In the chart in the lower right, we mapped happiness scores with a focus on Western Europe, to show the distribution within the happiest region.

Insight:

As per the above analysis, we observe higher Netflix prices with higher happiness scores, with Western Europe scoring highly in happiness and paying the most for Netflix.



Happiness and library size in Western Europe



Analysis and Visualizations

Inflation and Population compared to Netflix price

Analysis and Findings

Population vs Netflix Price by Country

We wanted to find out if Population size plays a role in Netflix Price. We noted from the chart on the right that as Population size increases, Netflix Price decreases. There is a negative correlation between Population and Netflix Price. [1]

Outlier- We noted that India, highlighted with a callout box, has a significantly high population. If we are to remove India from the dataset, there will be a stronger negative correlation. [2]

Inflation, Netflix Price, GDP by Country

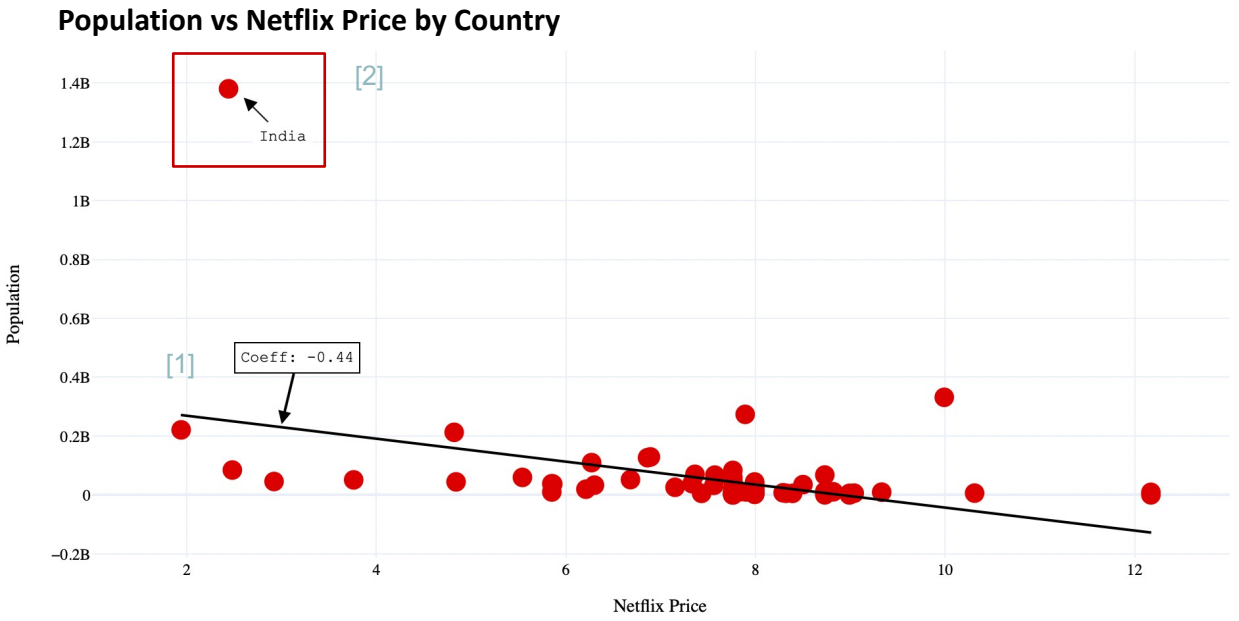
There was a country that is extremely anomalous – Lebanon with an inflation rate of 84.8 was significantly higher when compared to the next country with the second highest inflation rate of 12.3. Therefore, we removed Lebanon from the dataset.

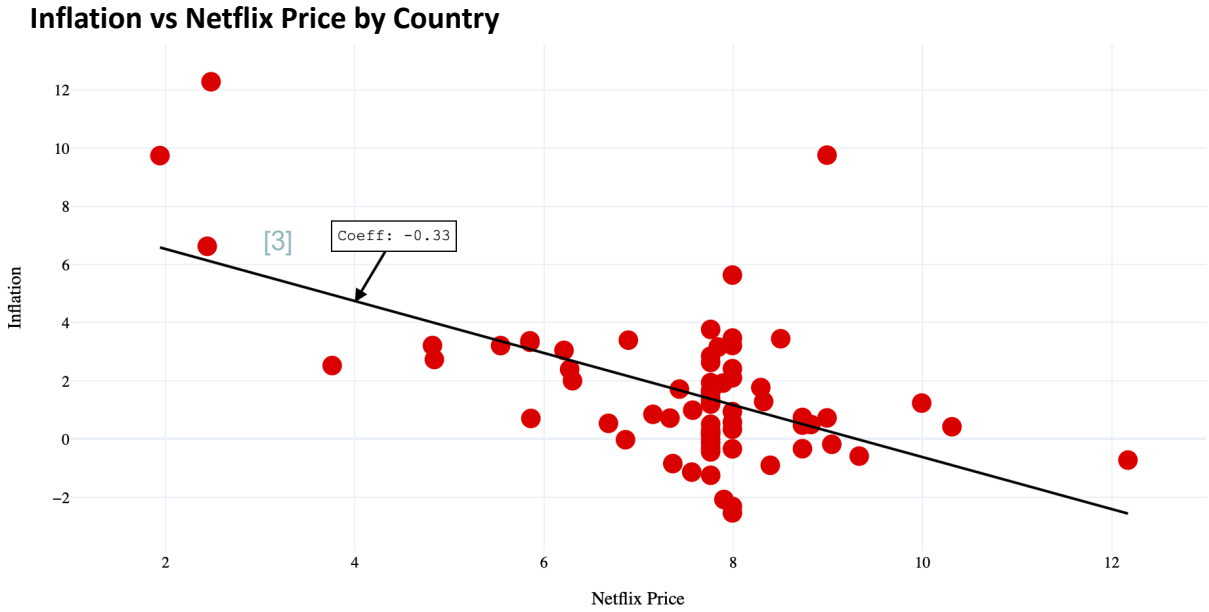
We noted that as Inflation decreases, Netflix Price decreases. We observed a negative correlation between inflation and Netflix Price". [3]

Insight:

The results show a **negative trend** that a country with

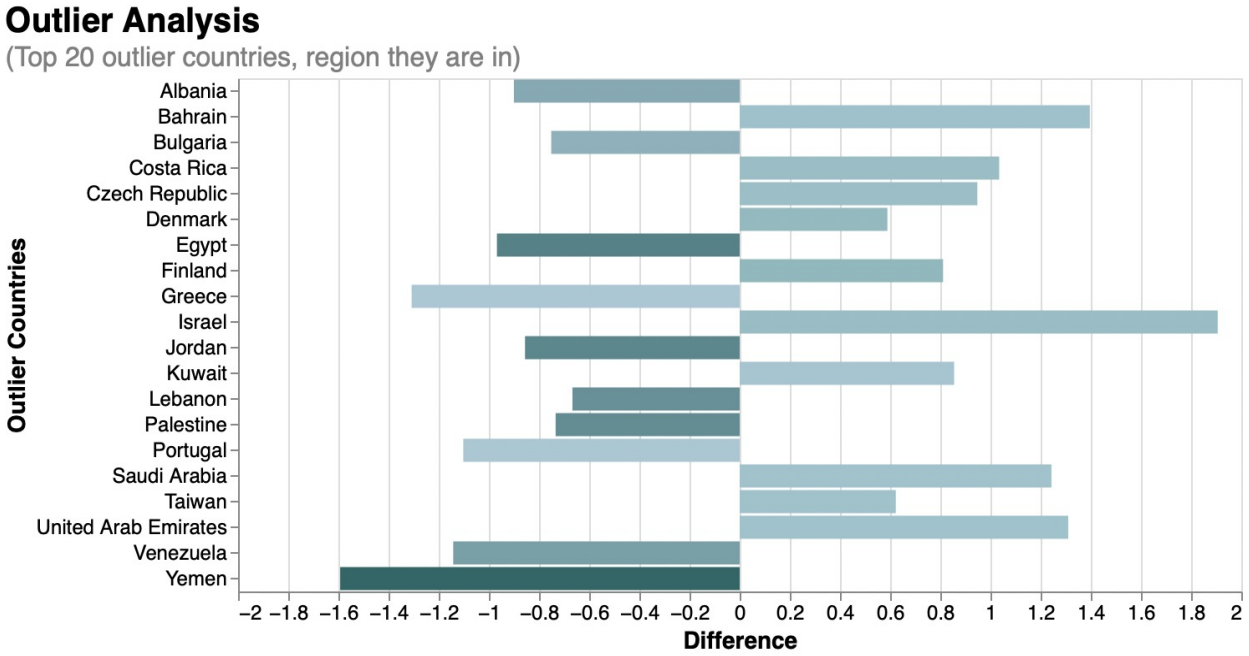
- higher Population has a lower Netflix Price [1],
- higher Inflation has a lower Netflix Price [3] .



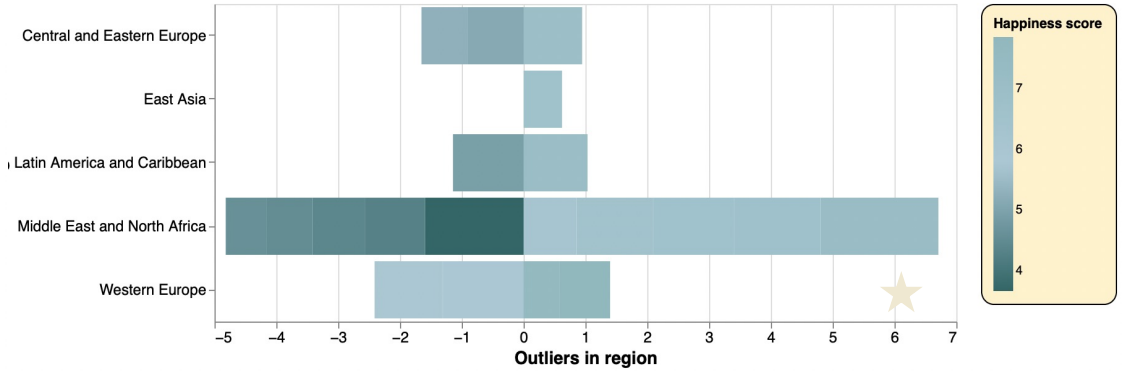


Analysis and Visualizations

Outlier Analysis – Outliers Noted



Rank	Country	Region	Happiness score	Average Happiness Score in region	Difference
Higher than region mean - Top 5 Outliers					
1	Israel	Middle East and North Africa	7.16	5.51	1.64
2	Bahrain	Middle East and North Africa	6.65	5.51	1.13
3	United Arab Emirates	Middle East and North Africa	6.56	5.51	1.05
4	Saudi Arabia	Middle East and North Africa	6.49	5.51	0.98
5	Costa Rica	Latin America and Caribbean	7.07	6.14	0.93
Lower than region mean - Top 5 Outliers					
1	Greece	Western Europe	5.72	7.03	(1.31)
2	Jordan	Middle East and North Africa	4.40	5.51	(1.12)
3	Portugal	Western Europe	5.93	7.03	(1.10)
4	Lebanon	Middle East and North Africa	4.58	5.51	(0.93)
5	Albania	Central and Eastern Europe	5.12	6.03	(0.92)



Analysis – Outliers noted

We present in the chart on the left with table as follows, the (1) Top countries whereby their happiness scores are higher than the mean of the region and the (2) bottom countries whereby their happiness score is lower than the mean of the region the country is in.

From this analysis, we can conclude that there are several countries where the happiness score differs from the mean of each region, this will impact the Correlation Coefficient calculation.

Findings

★ Out of the 10 regions under analysis, Middle East and North Africa presented the greatest number of outlier countries. These outlier countries will affect the calculation of the Correlation Coefficient.

To consider if we are to remove these outliers from the dataset, we should reach out to team that prepared this data and find out the basis for collection of data from these countries. These outliers will affect the correlation coefficient calculation for each region.

Analysis and Visualizations

Correlation analysis

Spearman's rank correlation coefficient was used to explore the correlations since none of the attributes were normally distributed, and we assumed linear relationships. None of the outliers were removed for purpose of our analysis.

For the chart on the right, we used Altair as the visualization tool to decode which features of Happiness Score, Economic Factors correlate to Netflix Price. We performed the analysis using (1) World – All 88 Countries, and (2) Europe region – 23 countries. We chose Europe instead of other regions as Europe commands the highest Netflix Price as discussed in page 5 of the report.

What we observed:

From the global perspective, for the 88 countries under review, we noticed:

- Moderate positive correlation between Netflix Price and GDP per capita which amounts to Global: 0.33, Europe: 0.51.
- Moderate positive correlation between Netflix Price and Happiness Score which amounts to Global: 0.34, Europe: 0.33
- Moderate negative correlation between Netflix Price and Population which amounts to Global: -0.44, Europe: -0.33
- Weak negative correlation between Netflix Price and Inflation which amounts to Global: -0.33, Europe: -0.24

Conclusion:

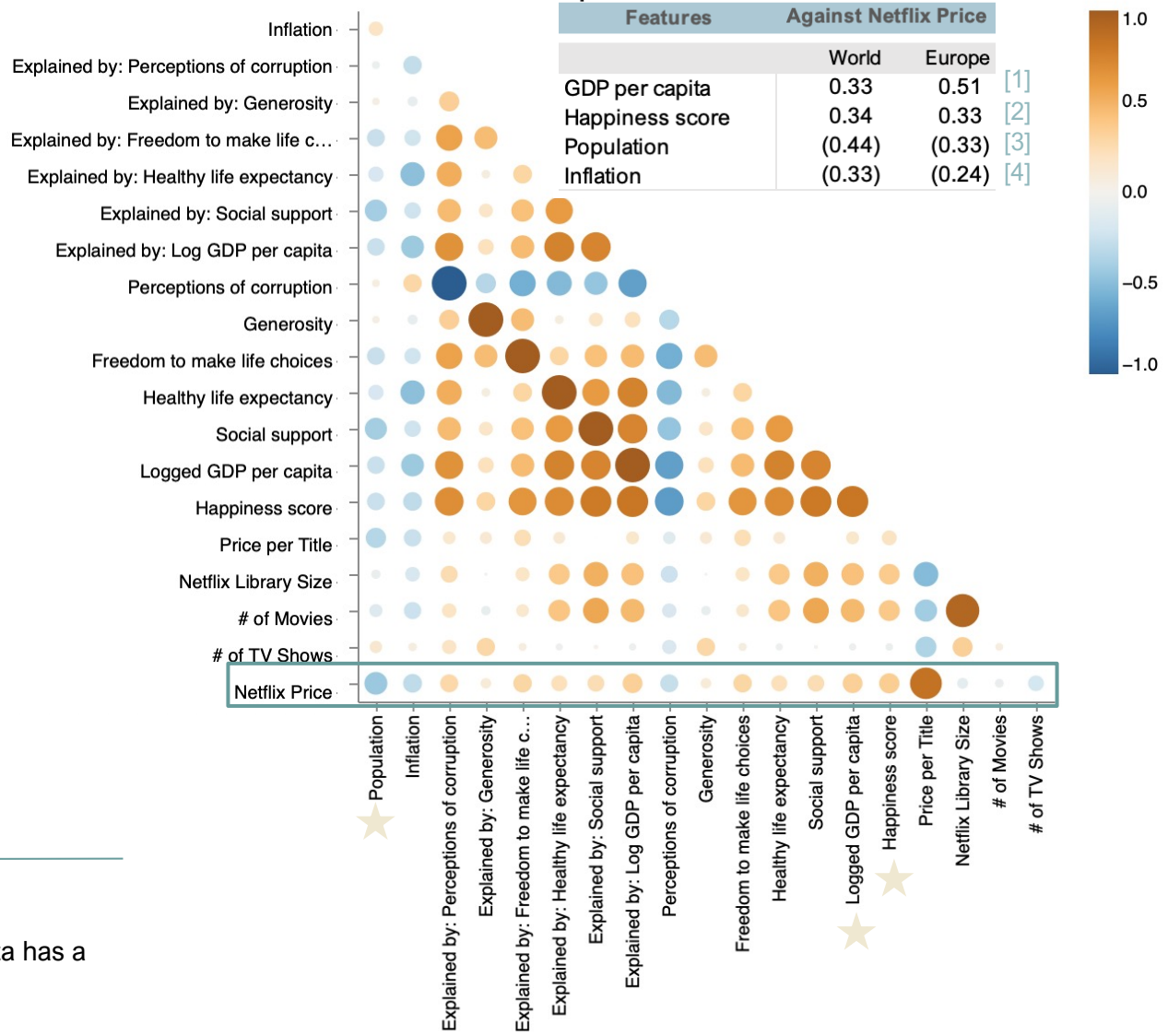
The results show a **weak positive trend** that countries with [1] higher GDP per capita has a higher Netflix Price, [2] higher Happiness Score has a higher Netflix Price.

A **weak negative trend** that countries with [3] higher population has lower Netflix Price, [4] higher inflation has lower Netflix Price.

Spearman Correlation

Spearman's rank correlation coefficients

Features	Against Netflix Price		
	World	Europe	
GDP per capita	0.33	0.51	[1]
Happiness score	0.34	0.33	[2]
Population	(0.44)	(0.33)	[3]
Inflation	(0.33)	(0.24)	[4]

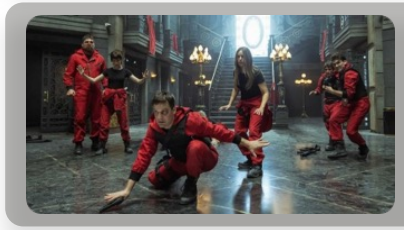


Note: Click on this streamlit link for interactive chart. We visualized the correlation between the various features using an interactive Altair plot. Works well on google chrome browser.

Statement of Work

All Members 1344 hours

Report writing, Ensure reproducibility, reviews



Steven Ong 1344 hours

Report design, manipulation, visualization, create modules

John Kaspers 1344 hours

Debugging, Manipulation, visualization, editing report content

Chi Huen Fong 1344 hours

Manipulation, visualization, code maintenance

Sources and Citation

[1] Netflix plans and pricing. Available at <https://help.netflix.com/en/node/24926>

[2] External data visualization tool. Available at <http://www.datawrapper.de/

[3] “THE NEW NETFLIX LOGO 2022.” Google, eDigital, https://images.app.goo.gl/oGUdhqkVHdrg3Sro6. Accessed 14 Oct. 2022.

[4] “The World Happiness Report.” Google, Worldhappiness.report, https://images.app.goo.gl/QyhjeQizrtkqUYJo9. Accessed 14 Oct. 2022.

[5] “The World Bank.” Google, Kindpng, https://images.app.goo.gl/zkmiQr35oeRQzgdR8. Accessed 14 Oct. 2022.