A Comprehensive Study of Global Economic Freedom and Happiness

Gopalan Iyengar
19D170009
Undergraduate Student
Dept. of Mechanical Engineering
Indian Institute of Technology Bombay

Shreyas Nadkarni
19D170029
Undergraduate Student
Dept. of Electrical Engineering
Indian Institute of Technology Bombay

***Abstract*—In the modern world, the overall happiness of the human race has empirically been dependent on multiple factors, only one of which will be scrutinized in this report, that is economic well-being. A comprehensive analysis of the factors affecting economic freedom and worldwide happiness standards has been made. A more detailed analysis of one facet was opted for over an inch-deep mile-wide analysis. This project report comprises a statistical study of data collected from various sources such as the Economic Freedom Index Report and World Happiness Report. An attempt has been made to explore the correlation of human happiness and the macro-economic scenario, and to suggest economic fields in which which countries could possibly improve their policies, so that the overall welfare(hereafter termed as 'Happiness Index/Happiness') of the citizens rises. After performing extensive exploratory and descriptive data analyses on data from multiple sources, it was found that happiness index does indeed have a high positive correlation with the economic freedom index. The major factors affecting the economic freedom index of a country are: property rights, judicial effectiveness, government integrity, financial freedom and business freedom. Region-wise analysis was also performed to examine dependency of happiness and economic freedom on the region.**

I. Introduction

For this project, the main objective was to relate how the worldwide happiness index has been affected by economic factors, economic freedom specifically. On viewing some existing analyses, it was found that the analyses were either completely oriented towards using models to estimate the data, used some specific country's data, or were majorly qualitative in nature. This project aims to provide the reader with a simplistic qualitative and quantitative analysis of data-sets consisting of data from most countries, in an ordered tabular and graphical manner. The data used in this project consists of happiness index data, and economic and overall freedom index data from multiple reliable sources. We will be exploring the distribution, correlation and cohesion of the data accumulated from different sources.

Objectives:

- To analyse the Economic Freedom Index Data for 2019 of several countries and explore its dependence on factors such as property rights, judicial effectiveness, individual freedoms (financial, monetary, labour,etc), per-capita GDP, population, interest rates and tax rates.
- To find how much the Economic Freedom Index affects the overall happiness of citizens.
- To compare statistics of various regions with each other and analyse whether economic freedom and well-being of a country's citizens depend on the region or continent the country is in.
- To construct a regression model to accurately predict the happiness score of a country, given the year, region, and economic freedom index.

Related Work:

- 'Is Economic Freedom the Secret to Happiness?'[1]- Work which was freely available seemed to be purely qualitative, and based on data related only to the USA.
- 'Economic Freedom and Happiness'[2]
- 'Happiness and economic freedom: Are they related?' [3]

II. Datasets

Data from multiple sources has been collected and utilized in this project. The data which is essential for the above mentioned analyses are the economic freedom index data, and the world happiness index data.

A. Economic Freedom Index for 2019

This dataset is owned by The Heritage Foundation. It has several attributes which contribute in the calculation of Economic Freedom Scores. The data is for the year 2019, for a total of 186 countries across the globe. Most of the data is numerical, though we have one categorical variable called "Region" which takes 5 possible values: "Asia-Pacific", "Europe", "Americas", "Middle East and North Africa" and "Sub-Saharan Africa". There are continuous variables which contain several attributes which might be important determinants of economic freedom. Our target variable is "2019 Score" which is a measure of the overall economic freedom for each country in the year 2019. This extensive data is used in the analysis to understand the dependence of the economic freedom on each of the factors and also the correlation between each pair of the factors thereby moving towards understanding the key elements for global economic prosperity. It is available as a file named economic_freedom_index2019_data.csv.

B. World Happiness Report

The World Happiness Report is an annual publication of the United Nations Sustainable Development Solutions Network. It contains data on the overall happiness of people in various countries based on a survey. Citizens' happiness is one of the most important factor which governments take into account during policy-making which is the reason why this report continues to get global recognition. The columns like "GDP per Capita", "Family", etc quantify the extent to which these factors influence happiness and the "Happiness Score" is a sum of all these factors. We have the data for years 2015 to 2019 in the form of separate files, namely 2015.csv to 2019.csv. We have directly used the value of final happiness scores for years 2015, 2016, 2017 and 2019 since we have the data on economic freedom in these years and our objective is to compare the happiness levels with citizens' economic freedom, without going into the depths of each factor which influences happiness.

C. Economic Freedom Of The World

This data is published by Fraser Institute, Vancouver, Canada and it gives an insight into the extent to which economic freedom depends on countries' policies, in five major areas:

- Size of the government
- Legal System and Property Rights
- Sound Money and Inflation
- Freedom to trade internationally
- Regulation

There are multiple sub-attributes within each of the above areas whose detailed data is available in the file efw_cc.csv. However to keep our analysis simple, we have taken just the final values of these five attributes as features in determining the economic freedom score. We have data from 1970 to 2016 for each country and using statistical techniques, we have used it to find whether the variables follow a normal distribution. It is also used to find the correlation between each of these areas in order to compare the results with the ones obtained from the Economic Freedom Index data and ensure that both of the analyses give us similar results to draw a conclusion.

D. Human Freedom Index

The Human Freedom Index depicts a broad picture of the overall freedom that individuals enjoy on economic, civil and personal fronts. It is a huge dataset spanning 243 columns and containing 10 years' data of 162 countries. The dataset which we have used is hfi_cc_2019.csv which contains data from 2008 to 2017. For simplicity, we only make use of the columns which give a final estimate of economic freedom, personal freedom and overall human freedom, dropping the sub-attributes in each of them. The dataset is primarily used for the values of economic freedom for years 2015 to 2017 for analysing the correlation with the happiness scores of countries for these years from the World Happiness Index data, and also to find the dependence of happiness on personal and overall freedom, in addition to economic freedom

III. Analysis Pipeline

Summary: The data was assimilated from multiple sources. The economic freedom index data was taken from the following files:

- economic_freedom_index2019_data.csv
- efw_cc.csv
- hfi_cc_2019.csv

The happiness index data was used from the following files:

- 2015.csv, 2016.csv, 2017.csv, 2019.csv

s The files 2018.csv and hfi_cc_2018.csv were not used due to lack of economic freedom data for 2018, and outdated data respectively.
The data was cleaned and ordered. Extensive exploratory and descriptive data analyses were performed on these datasets individually, especially the Economic Freedom Index data for 2019. Then, merging the appropriate year-wise economic freedom index and happiness index data, a joint analysis was performed, and this merged data was analysed for the correlation between economic freedom and the happiness. Subsequently, all the years' data was combined and a few regression models were built and tested, which could serve to accurately predict the dataset parameters.

A. Economic Freedom Index data for 2019

- The data was cleaned, that is null entries and redundant columns were removed, data-types were analysed and established, and the data was sorted.
- The mean, standard deviation, etc. of the Economic Freedom Index (referred to as '2019 Score'), were studied, an overall plot of score vs. country was made, and region wise analysis was done.
- The Pearson Correlation of score with the different parameters (columns) was analysed, and a heatmap was created to easily visualize this alongwith correlation between every two columns. Some bubble/regression plots were created for important parameters, and some which had a high correlation.
- Normality of parameters and score was tested using log likelihood, chi-squared goodness-of-fit test for a normal distribution, and QQ-plots.
- A list of the top 20 countries by 2019 score was created for future comparison.

B. Happiness Index 2015-2019

- Separate dataframes for the year wise happiness index data were created and cleaned, for the purpose of using the happiness index in further analyses.
- The mean, standard deviation, of the dataframes etc. were analysed.

C. Year Wise Economic Freedom Index

- The data was loaded and cleaned, and five major relevant columns are chosen for analysis.
- These five major fields consist of various sub-fields which eventually weighed in to their super-field. These fields

are- 'Size of Government', 'Property Rights', 'Sound Money', 'Trade', and 'Regulation'.
- It was observed that here, the economic freedom index is the mean of the ratings in these five major fields.
- The Pearson Correlation of score with the different major fields was found, and a correlation heatmap was created to visualize this easily.
- The data was sorted by the average value of score for each country, and this was subjected to further analyses.
- Normality of the data of these 5 major fields and score was tested using log-likelihood, chi-squared goodness-of-fit test for a normal distribution, and QQ-plots.

D. Effect of Economic Freedoms on Happiness of Citizens

This is the essence of the project.

1) 2019 Data:
   - Only those fields which had common countries from the happiness index dataset (2019.csv) and the economic freedom dataset (economic_freedom_index2019.csv) were chosen.
   - Normality of economic freedom index and happiness index for 2019 was tested using log-likelihood, chi-squared goodness-of-fit test for a normal distribution, and QQ-plots.
   - The hypothesis that the two indices come from the same distribution was tested, i.e. the hypothesis that the two distributions have the same mean was tested by a Welch T-Test, and that the two have the same median was tested by a Wilcoxon Ranked Sign test.
   - The Pearson Correlation of economic freedom index and happiness index was computed, and a regression/bubble plot was made to visualize this correlation, along with region and population.
   - A region-wise analysis was done, i.e. average economic freedom index and happiness scores vs. region.
2) 2017 Data:
   - A similar analysis as 2019 was made for the data for 2017. The essential differences are as given below.
   - The economic freedom index data was taken from the 'Human freedom Index' dataset, i.e. hfi_cc_2019.csv, which contains freedom index data for 2015, 2016, and 2017, apart from others.
   - This dataset contain three columns equivalent to freedom indexes-
     a) Economic Freedom
     b) Personal Freedom
     c) Overall Human Freedom (mean of previous 2)
   - The dataset was cleaned and sorted, and values for 2017 were extracted.
   - The regions were a bit more detailed in this dataset, so they were accordingly grouped to their respective regions while merging data of happiness index (df2017).
   - Population data was not used.
   - Correlation and respective graphs were made with respect to all three equivalents of a freedom index.
3) 2015 & 2016 Data:
   - Exactly the same analysis was done for 2015 and 2016 as was done for 2017.

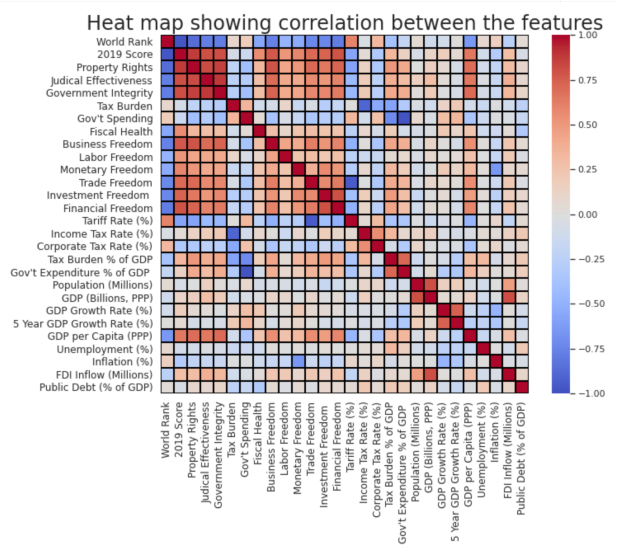
E. Application of Regression Models to consolidated data

- The goal is to build a regression model capable of predicting the happiness score of an entity(country-which is the index of the dataframes) given the region, year and the economic freedom index.
- The happiness index vs. economic freedom index dataframes which were separated by year, are now concatenated to form a significantly sized dataset for the regression model.
- Upon one-hot encoding the variables 'year' and 'region', a test split of 5% was utilized.
- Three regression models are used for testing:
  1) Multi-Layered Perceptron Regression (MLPRegressor)
  2) Lasso Regression
  3) Ridge Regression
- A grid-search was performed in a reasonable space of hyper-parameters for each model, in which cross-validation was used to verify the score (negative rms error) for hyper-parameter estimation.
- Upon comparison of the models' losses given their best hyper-parameters, a suitable best performance model was chosen.

IV. Results and Discussion

A. Economic Freedom Index data for 2019

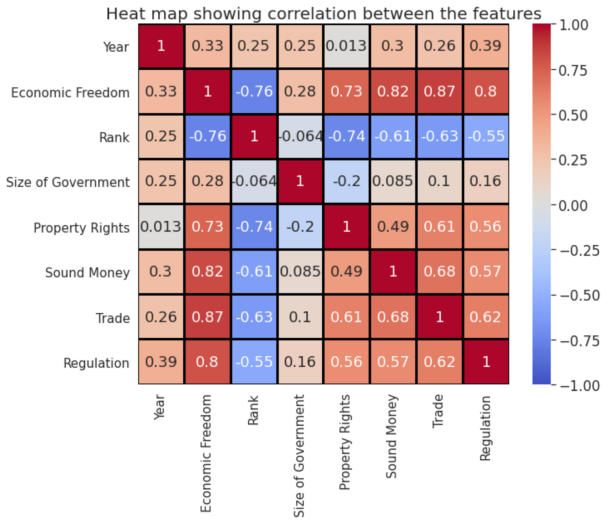
- The average score in 2019 was 61.35 and the standard deviation was 10.29 (out of 100).
- a lot of European countries had high scores as compared to other regions and the average, while the Sub Saharan African Countries mostly had low scores.
- The attributes which are significantly impacting the economic freedom score (correlation higher than 0.75) are Property Rights, Judicial Effectiveness, Government Integrity, Financial Freedom and Business Freedom.



- The data which could be considered near normal accounting for a 95% confidence interval are 2019 Score, Business Freedom, Labor Freedom, Financial Freedom, Income Tax Rate(%), and Gov't Expenditure(% of GDP).

B. Year-Wise Economic Freedom Index

- It was found that the size of government was the only factor out of five which did not affect economic freedom greatly.



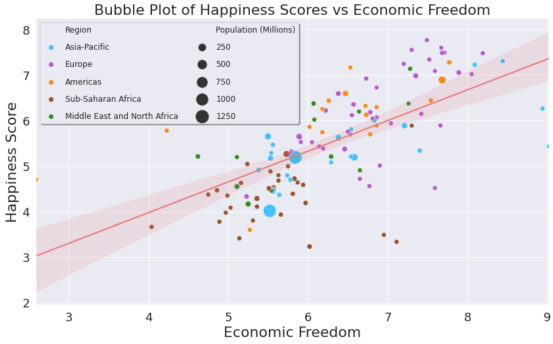
- The data was sorted by the average value of score for each country over all the years, and this was subjected to normality tests. The data which could be considered near normal accounting for a 95% confidence interval are Economic Freedom, Size of Government, and Regulation.

C. Effect of Economic Freedoms on Happiness of Citizens

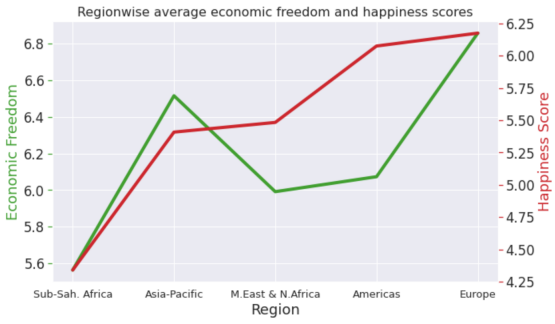
This is the essence of the project.

1) 2019 Data:

- Economic Freedom and Happiness Index for 2019, distributed by country, may be considered a normal distribution accounting for a 95% confidence interval.
- The null hypotheses for both the Welch T-Test and Wilcoxon Ranked Sign Test were rejected, implying that the Economic Freedom and Happiness Index were **not** from the same distribution.
- The Pearson Correlation Coefficient of economic freedom index and happiness index was computed to be approximately **0.63**, implying a moderately strong positive correlation.



- The region-wise distribution is as follows:



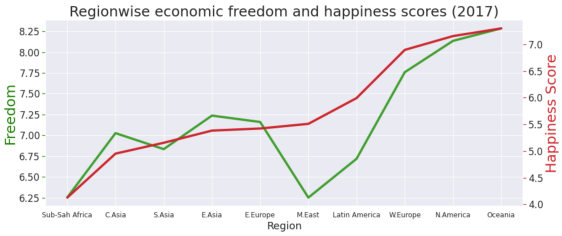
2) 2017 Data:

- Economic Freedom and Happiness Index for 2017, distributed by country, may **not** be considered a normal distribution accounting for a 95% confidence interval, whereas Personal Freedom and Overall Freedom may be considered normally distributed.
- The null hypotheses for both the Welch T-Test and Wilcoxon Ranked Sign Test between Economic Freedom and Happiness Index were rejected, implying that these were **not** from the same distribution.
- The Pearson Correlation Coefficient of Happiness Score and:

  a) Economic Freedom: 0.55
  b) Personal Freedom: 0.59
  c) Overall Human Freedom: 0.61

  This is consistent with data from 2019, and implies a moderately high correlation between happiness score and freedom.

- The region-wise distribution is as follows:



3) 2015 & 2016 Data:

- Economic Freedom and Happiness Index for 2015 **and** 2016, distributed by country, may **not** be considered a normal distribution accounting for a 95% confidence interval, whereas Personal Freedom and Overall Freedom may be considered normally distributed.
- The null hypotheses for both the Welch T-Test and Wilcoxon Ranked Sign Test between Economic Freedom and Happiness Index were rejected for both 2015 **and** 2016, implying that these were **not** from the same distribution.
- The Pearson Correlation Coefficient for 2015 of Happiness Score and:

  a) Economic Freedom: 0.46
  b) Personal Freedom: 0.54
  c) Overall Human Freedom: 0.55

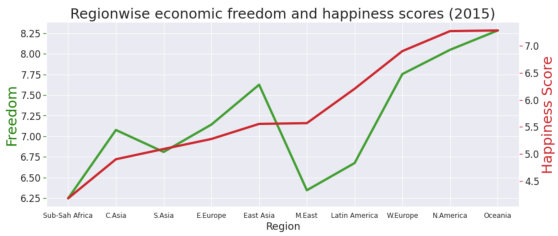
  This is consistent with data from 2019 and 2017, and implies a moderately high correlation between happiness score and freedom.

- The Pearson Correlation Coefficient for 2016 of Happiness Score and:

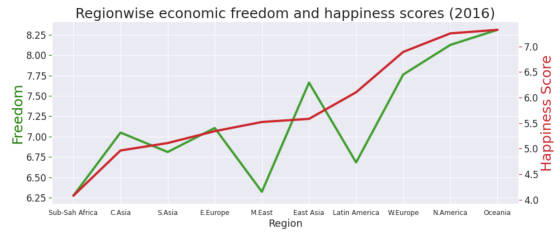
  a) Economic Freedom: 0.50
  b) Personal Freedom: 0.57
  c) Overall Human Freedom: 0.58

  This is consistent with data from 2019 and 2017, and implies a moderately high correlation between happiness score and freedom.

- The region-wise distribution for 2015 is as follows:



- The region-wise distribution for 2016 is as follows:



D. Application of Regression Models to consolidated data

- 5 test runs for each model were run, and the model which had the least average rms error was chosen. The results for a specific split of randomized test-train split are:

	RMSE Value 1	RMSE Value 2	RMSE Value 3	RMSE Value 4	RMSE Value 5	Average RMSE Value
MLPRegressor	0.678646	0.671305	0.672063	0.698173	0.667948	0.677627
Lasso	0.720634	0.720634	0.720634	0.720634	0.720634	0.720634
Ridge	0.721278	0.721278	0.721278	0.721278	0.721278	0.721278

- It is observed that minimum average RMSE is: 0.677 for 'MLPRegressor' Regression Model. Thus, we can conclude that this model is appropriate for use on this dataset.

E. Conclusions

- We can conclude that happiness index has a good positive amount of correlation with the economic freedom index. This is a pretty strong conclusion for data collected from various sources.
- We can also say that personal freedom and overall freedom also have a good correlation with happiness, from our analysis of data for 2015-17. This is also a strong conclusion.
- Since no economic freedom index data was known during the lifetime of the project, the absence of analyses for 2018 is a huge gap, which may have lead to skewed results in the prediction model. Thus the prediction model might not be as accurate as one may think.
- This has been a simplistic view of the correlation between economic freedom and happiness, and can be used as a stepping stone for those who might want to attempt a more in-depth analysis of this topic in a specific field/for a specific country, region or time period.

ACKNOWLEDGMENTS

This project was done as a part of the course DS203: Programming For Data Science, under C-MInDS, IIT Bombay. We would sincerely like to thank our course instructors Prof. Amit Sethi, Prof. Manjesh Hanawal, Prof. Sunita Sarawagi and Prof. S. Sudarshan for the valuable knowledge they have given us through this course alongwith the specific guidance for this project. We would also like to express our gratitude to our Teaching Assistant, Mr. Parikshit Bansal, who helped us sail through this course.

REFERENCES

[1] Jackson, J. Free to Be Happy: Economic Freedom and Happiness in US States. J Happiness Stud 18, 1207–1229 (2017). https://doi.org/10.1007/s10902-016-9770-9

[2] Daniel M. Gropper, Robert A. Lawson, and Jere T. Thorne Jr., (2011), Economic Freedom and Happiness, Cato Journal, 31, issue 2, p. 237-255, https://EconPapers.repec.org/RePEc:cto:journl:v:31:y:2011:i:2:p:237-255.

[3] Ilkay Yilmaz, and Mehmet Nasih, "Happiness and economic freedom: Are they related?", SHS 28 Web of Conferences,01109 (2016) RPTSS 2015

APPENDIX

Notebook used-

- COLAB: https://colab.research.google.com/drive/13WLa6xK6AEG7JyrLAX4NNllYChiHyHjB?usp=sharing
- GitHub: https://github.com/gopalaniyengar/DS203-Project/blob/main/19D170009_19D170029.ipynb