External Factors That Affect A Country's Economy

Arnav Gupta

Hao Doan

Table of Contents

I. Summary of research questions and results

1. How does the size of a country's government affect the living standards for average citizens?

As a country's government continues to grow, it hinders the growth of a country's economy. Consequently, the living standards for citizens, which include money growth, the ability to start a business, ownership of banks, subsidies, and the overall ranking of the economy, are lowered as well. There is a negative correlation between the size of a country's government and the living standards for average citizens.

2. Which countries allow their citizens to have the most freedom and does that correspond to that given country's economic situation?

The countries that have the highest economic freedom: Australia, Canada, Switzerland, United Kingdom, Hong Kong, Ireland, Mauritius, New Zealand, Singapore, United States.

The countries that have the lowest economic freedom: Angola, Central African Republic, Congo Dem. Rep., Congo Rep. of, Algeria, Egypt, Guinea-Bissau, Iraq, Libya, Sudan, Syria, Chad, Venezuela.

The economic freedom of a country directly correlates to their economic situation, represented by their economic ranking.

3. Can we predict the quartile to which a country's economy belongs based on different factors that affect the economy?

We are able to train a simple machine learning model to classify each country economy to its corresponding quartile based on the different factors of its economy which are Government consumption, Government investment, Labor market regulations, Legal System & Property Rights, Top marginal tax rate, Inflation, Freedom to Trade Internationally, Hiring regulations and minimum wage, Hiring and firing regulations, and Business regulations.
The accuracy of the model we trained is:
90-10 split: 80%
80-20 split: 86%
70-30 split: 73%

II. Motivation and Background

The primary motivation behind this project is to discover a trend in the change of a country's economic status based on various external factors that are presented in the dataset that we are using. By analyzing these factors, we will approximately be able to determine what most influences an economy and assess what a country should prioritize more in order to promote quality living conditions for their citizens. Additionally, this can help citizens with the election voting process as well because understanding what factors most changes an economy will further assist us to make judgement on different policies that presidential candidates put forth.

III. Dataset

The dataset for our project can be found on the following URL:
https://www.fraserinstitute.org/economic-freedom/dataset?geozone=world&page=dataset&min-year=2&max-year=0&filter=0&sort-field=regulation&sort-reversed=0[1]
We are working with 2 datasets. One is the world economic data in 2017 and the other is the world economic data from 1970 to 2017 (master data). Both datasets give rankings of 162 countries in the world based on multiple factors that affect the economy in 2017. The world economic data in 2017 contains 162 rows and 73 columns where each row represents a country and each column is a factor that affects the country's economy. The master data contains 73 columns but 3888 rows since it contains data from 1970-2017. The purpose of using this dataset is to pinpoint a trend in the change of a country's economic status. Analyzing various factors that the dataset provides will help us assess what most influences an economy and what a country should prioritize more. For our project, we are more concerned with the following 25 columns:

- Country name
- Economic Freedom Summary Index
- Overall rank
- Reliability of Police
- Size of Government
- Ownership of Banks
- Impartial Courts
- Starting a business
- Money growth
- Integrity of the legal system
- Transfers and Subsidies

[1] "dataset | Fraser Institute." https://www.fraserinstitute.org/economic-freedom/dataset. Accessed 15 Mar. 2020.

- Freedom to own foreign currency bank accounts
- Freedom of foreigners to visit
- Freedom to Trade Internationally
- Judicial Independence
- Government consumption
- Government investment
- Labor market regulations
- Legal System & Property Rights
- Top marginal tax rate
- Inflation: Most recent year
- Freedom to Trade Internationally
- Hiring regulations and minimum wage
- Hiring and firing regulations
- Business regulations

IV. Methodology

1. How does the size of a country's government affect the living standards for average citizens?

In order to answer this question, we will need to compare the government size to various economic factors such as money growth, ownership of banks, and the overall state of the economy. This will require us to first filter through our data and find the given economic factors that we deem most likely to affect an individual's living conditions. We will then plot the government size vs economic factors graph to analyze the trend. Generally, as the government continues to grow, it can decrease the growth of a country's economy, thusly lowering the living standards for citizens of that specific country.

2. Which countries have the highest economic freedom or the lowest economic freedom and does that correspond to that given country's economic situation?

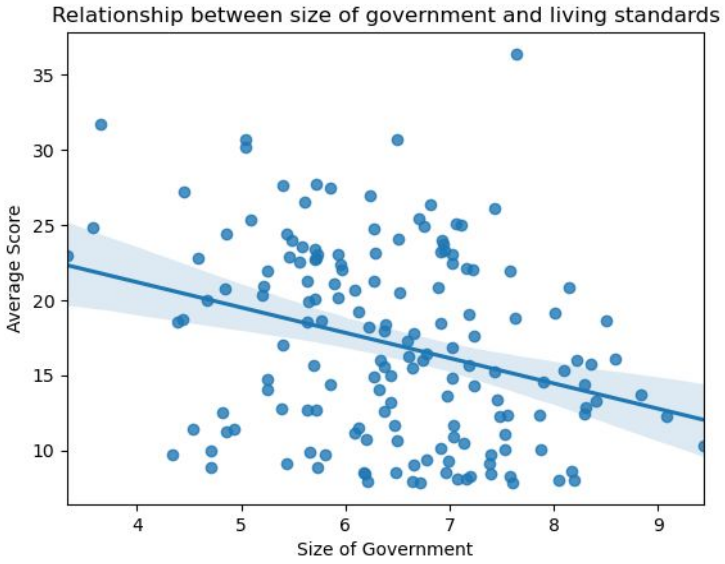
To answer this question, we will find the maximum data points for various columns in the dataset corresponding to a country's freedom such as freedom to trade, freedom of foreigners to visit, freedom of owning bank accounts etc. Then, we will average these values out in order to find the top 3 countries that provide the most freedom for their citizens. After calculating these results, we will plot this value with the country's corresponding economic ranking to assess whether or not a country's freedom has any relationship with the status of the economy.

3. Can we predict the quartile to which a country's economy belongs based on different factors that affect the economy?

To answer this question, we will build a simple machine learning model using scikit-learn library. First, we filter the columns for Government consumption, Government investment, Labor market regulations, Legal System & Property Rights, Top marginal tax rate, Inflation, Freedom to Trade Internationally, Hiring regulations and minimum wage, Hiring and firing regulations, and Business regulations. These columns are chosen based on economic factors listed on wallstreetmojo.com[2] : tax rate, exchange rate, inflation, labor, demand/supply, wages, law and policies, governmental activity and recession. Then, we separate input values and target values. We split the data into two parts: a majority amount of the data is used for training and the left over data is used to test the model. To find the most accurate data split, we will experiment on three kinds of data split: 90-10, 80-20 and 70-30. Next, we train a decision tree classifier model that takes in the filtered columns to predict the quartile to which a country's economy belongs. Finally, we will verify the accuracy of the model by predicting the quartile on the test set.

V. Results

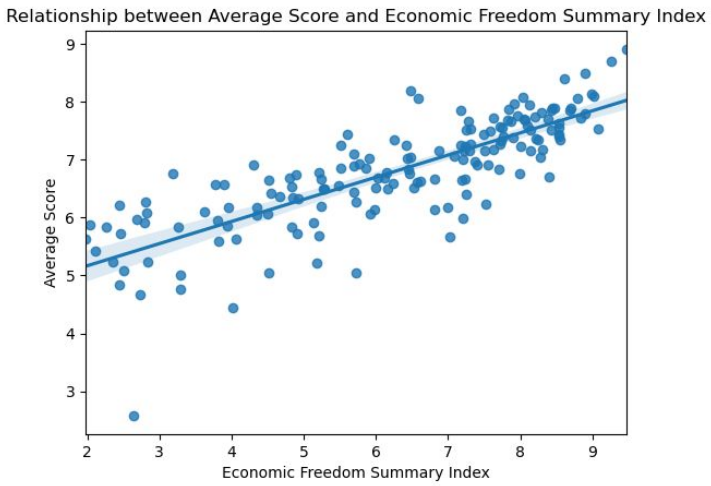
1. How does the size of a country's government affect the living standards for average citizens?



[2] "Economic Factors - WallStreetMojo." https://www.wallstreetmojo.com/economic-factors/. Accessed 15 Mar. 2020.

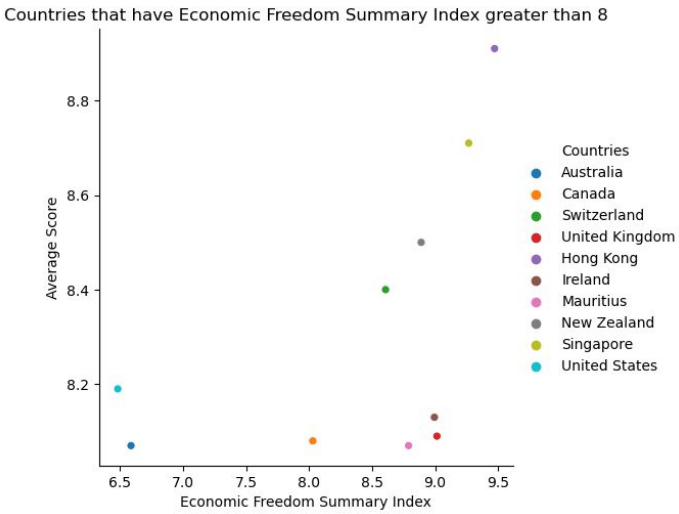
As we can see by the scatterplot and the line of best fit, it shows a clear negative correlation between the size of government and living standards. As the size of the government increases, the average living condition score decreases. Additionally, some of points along the scatterplot seem to be dispersed across the entirety of the graph instead of following the pattern of the line of best fit. This is because the size of government is not the only factor that affects an individual's living conditions. Taking money growth as an example, there could be many reasons why a country may experience a surplus or shortage of money such as natural disasters, breakthroughs in technology, increase in labor etc. The size of a government is just one of the many factors that affect living conditions, which is why the data in the plot does not exhibit a pattern as obvious as our other plots.

2. Which countries have the highest economic freedom or the lowest economic freedom and does that correspond to that given country's economic situation?



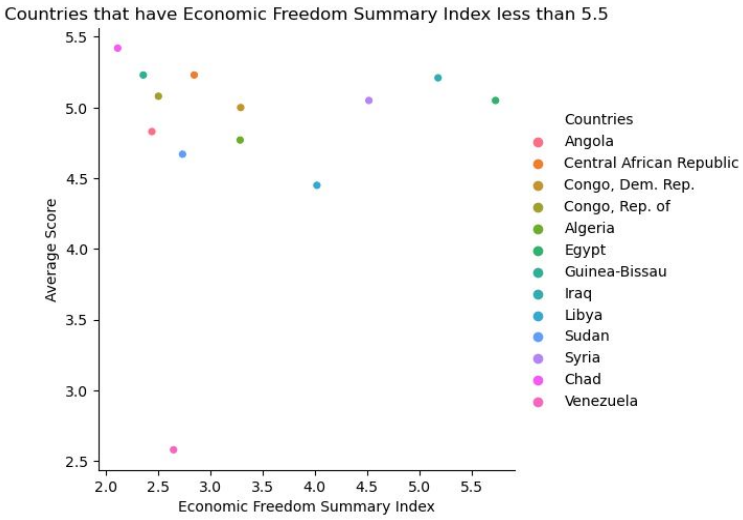
Based on the data in 2017, the result shows that there is a linear relationship between the average score of a country's freedom versus Economic Freedom Summary Index. This relationship is represented by a straight line in the graph above. Thus, we may conclude that there is a linear relationship between the freedom scores and a country's economic situation. The more freedom a country has, the better its economy.

To see which countries allow their citizens to have the most freedom we produced the graph of countries that have the Economic Freedom Summary Index greater than 8.



Based on this graph, the countries having the most economic freedom are Australia, Canada, Switzerland, United Kingdom, Hong Kong, Ireland, Mauritius, New Zealand, Singapore, United States. Hong Kong has the highest Economic Freedom Summary Index. To better see what it is, we filter the Economic Freedom Index of Hong Kong and obtain the result: ('Hong Kong', 8.91).

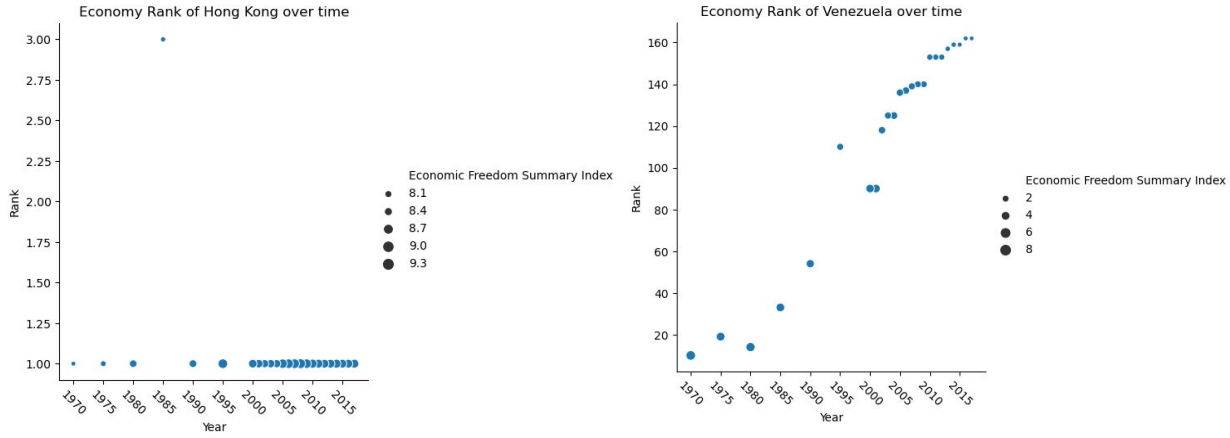
On the other side of the spectrum, we also look at the countries that have an Economic Freedom Summary Index less than 5.5.



Based on this graph, the countries having the least economic freedom are Angola, Central African Republic, Congo Dem. Rep., Congo Rep. of, Algeria, Egypt,

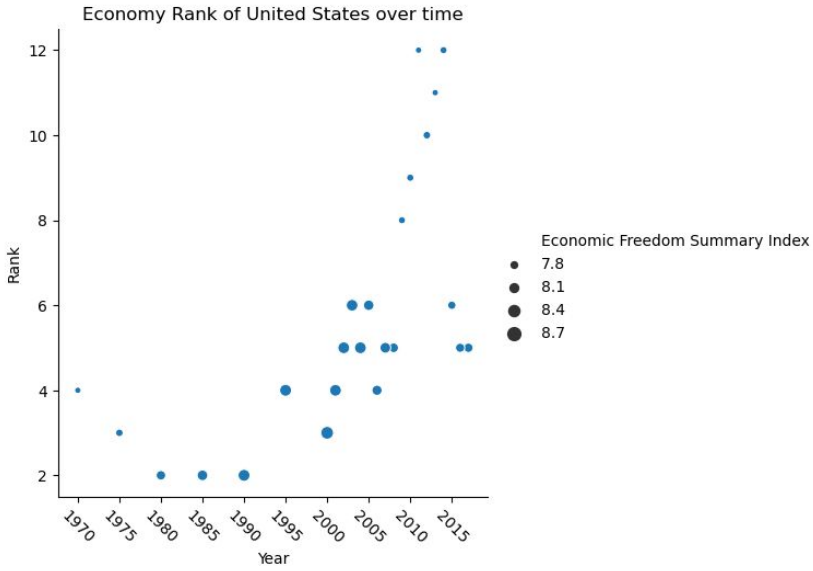
Guinea-Bissau, Iraq, Libya, Sudan, Syria, Chad, Venezuela. Venezuela has the highest Economic Freedom Summary Index. To better see what it is, we filter the Economic Freedom Index of Venezuela and obtain the result: ('Venezuela', 2.58).

To expand our scope of research, we also look at how the economy of Hong Kong and Venezuela developed over the period from 1970-2017. We chose Hong Kong and Venezuela as objects of study because Hong Kong has the highest economic rank in 2017 and Venezuela has the lowest economic rank in 2017. The graphs of the economy of Hong Kong and Venezuela from 1970 to 2017 are as follows:



We can see that the economy of Hong Kong remains strong and stable throughout the period 1970-2017. Hong Kong's Economic Rank always stays at #1, its Economic Freedom Index is always above 8.0. However, in 1985, Hong Kong dropped to #3 in Rank with an Economic Freedom Index of 8.1. On the contrary, the economy of Venezuela varies wildly throughout the period 1970-2017. It's Economic Rank drops from #12 to #160 throughout the years. Venezuela had a strong Economic Freedom Index in the 1970s and 1980s with values around 6.0-8.0 but these values dropped significantly over time.

We are also particularly interested with the situation of the United States. Below is the graph of the development of the economy of the United States over the period 1970-2017.



The United States' economy is mostly in the top 10 of the world from 1970 to 2017. It is the most developed in the 1980s and 1990s with the Economic Freedom Index of 8.1 - 8.4. We noticed a significant drop in Economy Rank during the late 2000s and early 2010s. This period corresponds with the Great recession of the United States' economy due to the housing bubble.

3. Can we predict the quartile to which a country's economy belongs based on different factors that affect the economy?

After performing experiment with splitting data by a 90-10 split, 80-20 split, 70-30 split. We obtain the following accuracy:
90-10 split: 0.8 ~ 80%
80-20 split: 0.8666666666666667 ~ 86%
70-30 split: 0.7333333333333333 ~ 73%
We may conclude that the 80-20 split gives us the best accuracy. Either of the three splits give us pretty high accuracy. Although the accuracy may change each time we run the model, under our observation, the accuracy of the model is above 60% almost 95% the time. This signifies that our model works quite efficiently in predicting the quartile to which a country's economy belongs.
We may also conclude that the list of factors that affect the economy are:

- Government consumption
- Government investment
- Labor market regulations
- Legal System & Property Rights

- Top marginal tax rate
- Inflation: Most recent year
- Freedom to Trade Internationally
- Hiring regulations and minimum wage
- Hiring and firing regulations
- Business regulations

VI. Reproducing results

To reproduce the results, perform the following steps:

1. Download submission from Gradescope.
2. Unzip the file downloaded that contains the code files and the data files.
3. Open Ed -> Click on Workspace tab.
4. Create a new workspace and choose a name of your choice.
5. Upload all the files unzipped to the workspace.
6. Open terminal.
7. Use the command

    $ python project.py

in the terminal. All the graphs will be recreated in the directory and the output will be displayed in the terminal.

8. To test the functions that output numeric values, use this command in the terminal:

    $ python project_test.py

If no error occurs, it means that the code passed all test cases.

VII. Work Plan Evaluation

To work on the project, we share the same workspace on Ed and collaborate on the coding part. The report is written collaboratively on Google Doc. We divide our project into 3 parts in order to answer the 3 research questions. Each part is divided into 3 subparts: coding, testing, and discussion of the result. The details of the work plan are listed below.

Part 1: Analysis of the effect of a country's size of government on the living standards of citizens (eta 5 hours) (Arnav)

1) Code:

    Write code to:

    - Filter/clean up the dataset
    - Find the overall economy score of the country that has the maximum size of government
    - Find the overall economy score of the country that has the minimum size of government

- Filter the dataset based on the columns that best represent a country's living conditions
- Graph the economy score of each country with respect to its size of government and government investment

2) Test:
   Test the code by utilizing smaller datasets to see if the graph appropriately matches the real data that we are utilizing

3) Discussion:
   Draw conclusions based on the results

Cleaning up the data was a fairly simple process because we figured out beforehand which columns we wanted to include and which columns we wanted to omit for our project. Part 1 did not take relatively long to code, however what did take a long time was mapping out which columns to include for the graph. Because there are so many factors that go into a country's living conditions, it took time to find the right columns that displayed a distinct relationship between the size of the government and living standards.

Part 2: Analysis of economic freedom (eta 5 hours) (Hao)

1) Code:
   Write code to:
   - Sort the countries based on economic freedom summary index
   - Find the average score among freedom to own foreign currency bank, freedom of foreigners to visit, freedom to trade internationally of each country
   - Sort the countries based on the average score calculated
   - Find the number of country that has the economic freedom summary index greater than 8
   - Graph the average score among freedom to own foreign currency bank, freedom of foreigners to visit, freedom to trade internationally of each country, judicial independence with respect to the economic freedom summary index

2) Test:
   Test if the code produces correct graphs that match the real data that we are utilizing
   Test if the code outputs the correct country that has the highest economy score and lowest economy score by utilizing smaller datasets
3) Discussion:

Draw conclusions based on the results

For this part, we add one more column to the average on which is the judicial independence which we think plays a role in determining the economic freedom summary index. We also find the countries that have the economic freedom summary index less than 5.5. In addition, we expanded our scope of research and looked into the economic development of Hong Kong, Venezuela, and the United States over the period 1970-2017.

Part 3: Machine Learning model (eta 5 hours) (Arnav and Hao)

1) Code:
   Write code to:
   - Filter the columns for overall economy score, government investment, legal system and property rights, business regulations and freedom to trade internationally then separate input values and target values
   - Divide data: 90% percent of the data for training and 10% of the data to test the model
   - Train a decision tree regressor model that takes in government investment, legal system and property rights, business regulations and freedom to trade internationally scores
   - Predict the overall economy score of a country
   - Verify the accuracy of the model

2) Test:
   Check if the accuracy is high enough and acceptable.

3) Discussion:
   Draw conclusions based on the results

For this part, we decided to change our machine learning model to a quartile classifier instead of a economy score regressor. This is because we think that the factors we chose might not be perfectly exact and therefore predicting the economy score of a country is going to be inaccurate. However, predicting which quartile the economy of a country belongs to is more feasible and therefore the model will give higher accuracy**.** We decided to experiment on different data splits to find the most accurate one. Also, based on research, we chose a different set of factors that affect the economy for the model to learn.

Overall, our work plan was accurate. We extended our scope of research on part 2 and performed more experiments for part 3. For this, the work time of each part is more that what we anticipated. It took us around 7 hours for each part. However, we had the chance to explore further on the topic we are interested in.

VIII. Testing

We tested the code with numeric outputs by using the assert equals function in python. We came up with test cases by writing down the expected output for each test case and checked whether it is equal to the output of our code. For the functions that plotted graphs, since there was no numeric output, we did not utilize any assert equals function for our tests. Instead, we created smaller data sets to ensure that the results of our graphs were accurate. Since our code passed all of our test cases, our results are accurate and trustworthy. Additionally, our code can be trusted because our dataset comes from a respected source (Fraser Institute), which means that the data plotted in our graph is accurate as well.

IX. Collaboration

This project was created through the collaboration of Hao Doan and Arnav Gupta.