Correlation of the Economic Freedom Index with other Indices, and its General Importance Worldwide

- A Data Mining Approach

Andrés José Basile Álvarez

Github: https://github.com/andresbasilea/EconomicFreedom_ADataMiningApproach
Universidad Nacional Autónoma de México
Facultad de Ingeniería
andresbasilealvarez@gmail.com
January, 2023

Abstract

An economically free society is that in which its individuals are free to produce, work, consume and invest in any way they please. There are few published articles that use data mining and artificial intelligence techniques to determine the relationship between economic freedom and other sociological metrics in a country. In this study, we present a statistical and AI analysis of the correlation between different sociological metrics and the economic freedom of a country, with the main purpose of integrating data from different sociological metrics in such a way that it can lead to a better understanding of the implications that economic freedom has in the overall well-being of an individual inside a country. The approach to the analysis consisted on a variable selection process using the Pearson correlation matrix, then analyzing the correlation between different variables and finally building a Random Forest Regressor to predict happiness and crime values using 14 economic variables. The Random Forest Regressor models created presented an r2 score of 85%. These models show the relationship between the economic and societal factors in a country and they could be used to help define economic, security or social policies.

Keywords: Correlation, Economy, Sociological Indices, Data Mining

1 Introduction

Economic Freedom is the fundamental right of every human to control his or her own labor and property. An economically free society is that in which its individuals are free to produce, work, consume and invest in any way they please [Foundation, 2022b].

The Heritage Foundation, creator of the *Index of Economic Freedom*, defines it as the documentation of the positive relationship between economic freedom and a variety of positive social and economic goals, measured based on 12 quantitative and qualitative factors, grouped into four broad categories of economic freedom: **Rule of Law:** property rights, government integrity, judicial effectiveness; **Government Size:** government spending, tax burden, fiscal health; **Regulatory Efficiency:** business freedom, labor freedom, monetary freedom; and **Open Markets:** trade freedom, investment freedom, financial freedom.

The Economic Freedom Index rates each of the twelve economic freedoms on a scale of 0 to 100. Thus, the overall score of a country consists of the non-weighted average (equal importance given to each) of the twelve economic freedoms. The country's scores for the 2022 Economic Freedom Index were calculated based on data which covers the second half of 2020 through the first half of 2021. A more thorough explanation of the elements which constitute this index and the methodology used to construct it can be found at [Foundation, 2022a].

Having mentioned the fundamentals of the primary source of information for this investigation, our main focus throughout this article will be to understand the data which supports the Economic Freedom Index, define the importance of the economic freedoms worldwide and determine the correlation that economic freedom has with other country-based metrics, such as happiness and crime rates.

The previous because, counter intuitively, countries like Costa Rica or Nicaragua, which rank 55 and 122 on the Economic Freedom Index, respectively, rank much better on the happiness index [John F. Helliwell and Wang, 2022], at 23 and 45 (see *fig. 1)*.

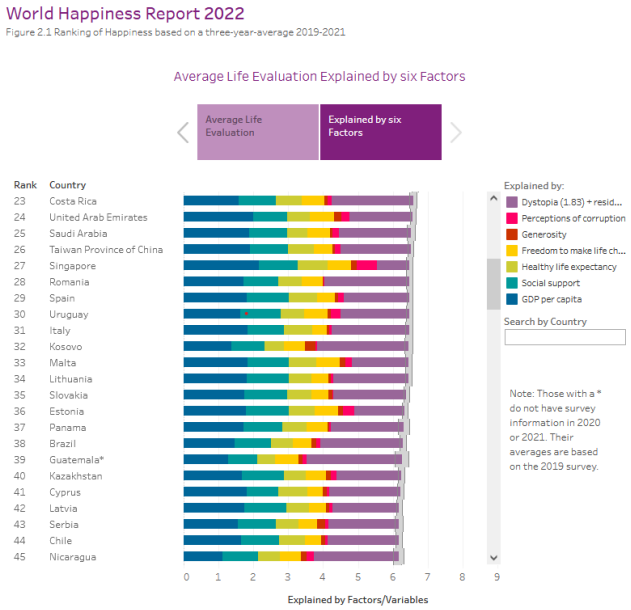


Fig. 1: Ranking of Costa Rica and Nicaragua on the Happiness Index. Extract from [John F. Helliwell and Wang, 2022].

In this study, then, we present a statistical analysis of the correlation between these different metrics and the economic freedom of a country, with the main purpose of integrating data from different sociological metrics in such a way that it can lead to a better understanding of the implications that economic freedom has in the overall well-being of an individual inside a country.

2 Research background and related works

The relationship between a society's well-being and its economic factors has long been studied, as [Granillo, 2019] mentions, the relationship between the human work, creation, consumption, the economy and a nation's welfare was described by Adam Smith (1723-1790) in his book *The wealth of Nations* from 1776.

Nonetheless, today's deeply capitalist societies and their growth and wellness can be understood thanks to the works of [Solow, 1956], as described in the article by [James D. Gwartney, 1999], in which economic freedom is found to be a significant determinant of economic growth.

More recent approaches by [Foundation, 2022a] have found interesting relationships between a country's economic freedom and the *Human Development Index Score*, finding a positive correlation of 0.72 in the Middle East / North Africa region, for example (see *fig. 2*). This value indicates the strong bond between economy and general well-being in a country, hypothesis on which we base the analysis done in this study.

The analysis done by [Daniel Gropper, 2011] details the relationship between happiness and economic freedom, underlining the positive relationship between national levels of happiness and economic freedom, as well as the strong positive relationship between GDP per capita and happiness, as shown in *fig. 3*.

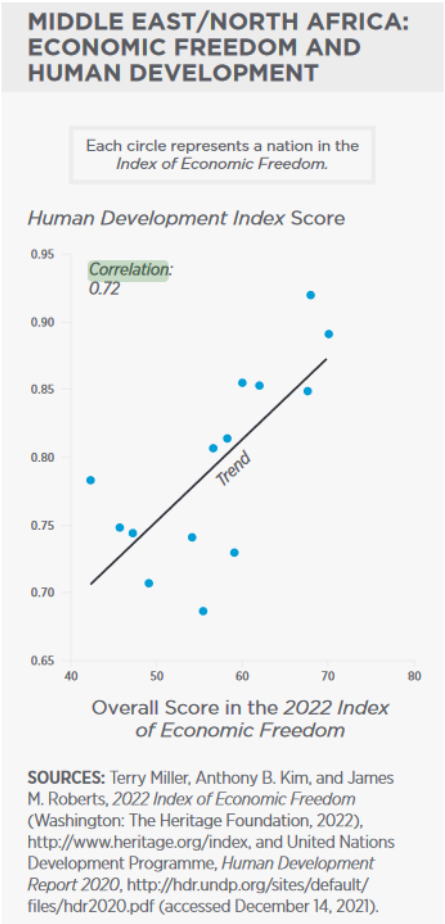


Fig. 2: Economic Freedom and the Human Development Index Score in Middle East / North Africa region. From [Foundation, 2022a].

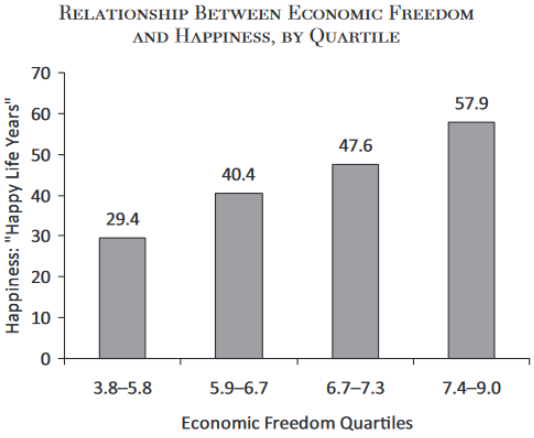


Fig. 3: Economic Freedom and "Happy Life Years". From [Daniel Gropper, 2011]

2.1 Motivation

The enormous development surrounding computers processing and storage capabilities in the last few years has set the appropriate tools needed for a thorough analysis and the leverage of big amounts of data. Economic and finance data has since been one of the key elements of study in data mining and artificial intelligence, with applications in economy government policies or algorithmic trading [Media, 2022]. However, there are few published articles that use data mining and artificial intelligence techniques to determine the relationship between economic freedom and other sociological metrics in a country.

For this reason, we have decided to make an integral, holistic, data mining approach to economic freedom and other sociological metrics, with the main purpose of determining the correlation and impact the economic freedom has in the wellness of a country's individuals.

3 Method

3.1 Data Source

The main data source used, as described in the introduction, is the *Index of Economic Freedom* by the Heritage Foundation. This index, consisting of 12 main factors on a scale of 0 to 100, paired with the *World Happiness* [John F. Helliwell and Wang, 2022], and *Crime Rate* [WPR, 2022] indices by country (2018-2022 period), comprise our source of information. The variables in the data used are described below:

- For Economic Freedom Index: The Economic Freedom Index rates each of the twelve economic freedoms on a scale of 0 to 100. 12 quantitative and qualitative factors, grouped into four broad categories of economic freedom: **Rule of Law:** property rights, government integrity, judicial effectiveness; **Government Size:** government spending, tax burden, fiscal health; **Regulatory Efficiency:** business freedom, labor freedom, monetary freedom; and **Open Markets:** trade freedom, investment freedom, financial freedom. Other variables included in this data source are: Investment Freedom (value from 0 to 100 for each country), Tariff Rate (percentage per country), Income Tax Rate (percentage per country), Corporate Tax Rate (percentage per country), Tax Burden percentage of GDP, Gov't Expenditure Percentage of GDP, Population (millions), GDP (Billions, PPP), GDP Growth Rate (percentage), Unemployment (percentage), inflation (percentage), FDI Inflow (Millions), Public Debt (Percentage of GDP).

- For the World Happiness Index: A value ranking from 0 to 8, reflecting the happiness index of the people in a country, 0 being the unhappiest and 8 being the happiest. This ranking is determined by weight averaging variables such as the degree of social support in a country, the healthy life expectancy at birth, the generosity of a country's population and the perceptions of corruption. More can be found at [John F. Helliwell and Wang, 2022].

- For the Crime Rate Index: Overall crime rate is calculated by dividing the total number of reported crimes of any kind by the total population, then multiplying the result by 100,000. [WPR, 2022]

  All data sources are grouped on a variable called "Country", a string with the name of the country that each element in the data set maps to.

3.2 Exploratory Data Analysis

The exploratory data analysis began with the printing of the dataframes to be used. On a first moment, a dataframe was used for each different source of data (one for World Happiness data for each year of study, one for Economic Freedom data for each year of study, etc.). Later on, the different datasets were analyzed, printing the type of data contained in each, as well as the *null* values, which were later on removed. To facilitate the posterior data analysis performed, the different dataframes created were merged

onto a single one, using the *merge* function from the *pandas* library in Python. The merge was done grouping on "Country", a string variable contained on each dataframe. The merged datasets were analyzed to find atypical values using histograms and boxplots. The variables which showed atypical values were individually inquired into to determine the need to remove them.

3.3 Variable Selection

Once the datasets were merged, the correlation between all variables was calculated using the Pearson method, parameter inside the *corr* function of the *pandas* library. The Pearson's correlation coefficients equation used can be seen below:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad (1)$$

Eq. 1: Equation used to compute the Pearson's correlation coefficients.

The variable selection process was done, firstly, by manually dropping the variables which do not add information to our analysis, like the "Webname" or the region in which a certain country is located. Later on, and for this first approach to the analysis, the calculated variables (those which derive their value from other variables in the set) like Judical Integrity, Government Integrity or Fiscal Health were dropped also. From the different indices to which we are to understand the correlation with the economic freedom variables, we added their values, merging them with the Economic Freedom index variables by country.

A second variable selection process was done after identifying the correlation between the variables on our final dataset. From variables which had a high correlation (either positive or negative), one of them was dropped to continue the analysis and to apply the AI algorithm on a decreased-dimensionality dataset. However, not all of the highly correlated variables were removed from the dataset, as some of them bring important information to our analysis even with the high correlation being present.

3.4 Correlation Analysis and Model Generation

A global correlation analysis between happiness and general economic freedom was done to look into the changing correlations between these two variables. A graph describes the relationship between the variables on the period 2018-2022. Another global analysis was also done to explore the correlation between the general economic freedom and crime, the investment freedom and crime and between happiness and crime on the same period, by mapping the correlation values yearly on a graph. A more specific analysis was done to compare the values of general economic freedom, happiness and crime scores throughout the years in Mexico, to gain a deeper sense of how this variables move with time.

Once the correlation between the variables had been defined, a random forest regression model was implemented, in such a way that the happiness index of a country can be calculated based merely on economic factors, such as inflation, tax rate percentage, GDP and Government Expenditure. This random forest regression was trained with a 80 percent of the total dataset training size, using 5000 estimators, and a maximum depth of 12 for each tree in the random forest. The *RandomForestRegressor* model from the *sklearn* Python library was imported for this purpose; as well as the *sklearn* metrics to determine the *r2 score* or *Coefficient of Determination* of each model. A *Linear Regression* model was also implemented, but the non linearity of the data showed this approach inefficient, as briefly discussed on the results section of this document.

4 Results

4.1 Data Source - Results

The motivation behind this investigation implied the utilization of multiple data sources to create a deeper understanding of the relationship between economic freedom (and economic factors in general) with sociological metrics (happiness and crime indices) for a given period of time (2018-2022). As a consequence, the Exploratory Data Analysis comprised analysis on multiple datasets: five with economic data (one for each year between 2018 and 2022), five for the Happiness Index data and five for the Crime Index data. The variables which form these datasets have been described in 3.1. The final shape of the datasets when loaded into the notebook, are: for Economic Freedom data, the datasets include 34 columns with about 186 entries each; for Happiness Index data, the datasets include 2 columns with about 150 entries each; and for Crime Index data, the datasets include 2 columns with about 130 entries each. The entries on each of the datasets correspond to information from a particular country, and the merging of these datasets will be explained on the following subsection, although it is shown in figure 4.

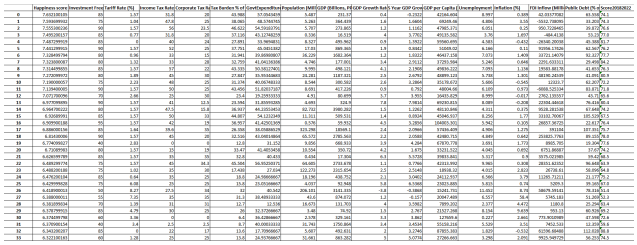


Fig. 4: Concatenated dataset, which includes information from happiness, crime, and economic freedom on 2018-2022. **Note: the text was exported to Microsoft Excel for visualization purposes.**

4.2 Exploratory Data Analysis - Results

The Exploratory Data Analysis done began with the printing of the general information from each dataset used. Included in the information printed was the total sum of *null* entries on each of the datasets. Leaving null valued entries on the datasets would have affected the results obtained by the AI model, reason for which we deleted these entries by using the *dropna* Python included function. It is important to mention that this function was only used for the Economic Freedom datasets, as Happiness and Crime Index datasets had no null valued entries.

As described on 3.3, a first drop of variables which do not add information to our analysis was done manually. From the Economic Freedom datasets, the complete list of columns dropped is: World Rank (we only care for the Economic Freedom Score), CountryID, WEBNAME (does not add useful information for our analysis), Region and Region Rank(we used a global approach, not a regional approach), Country (this column appeared twice in the datasets), Property Rights (calculated variable), Judical Effectiveness (calculated variable), Government Integrity (calculated variable), Tax Burden (calculated variable), Gov't Spending (calculated variable), Fiscal Health (calculated variable), Business Freedom (calculated variable), Labor Freedom (calculated variable), Monetary Freedom (calculated variable), Trade Freedom (calculated variable), Financial Freedom (calculated variable).

The merging of the datasets was firstly done by merging the Economic Freedom, Happiness Index and Crime Index datasets by year (i.e. Economic Freedom 2018 dataset with the Happiness Index 2018 and Crime Index 2018 datasets). After the merging into 5 distinct datasets was done, the first analysis was the search for atypical values inside the datasets. For this purpose, a histogram plot was done for each variable inside the datasets, as shown in the figure 4.

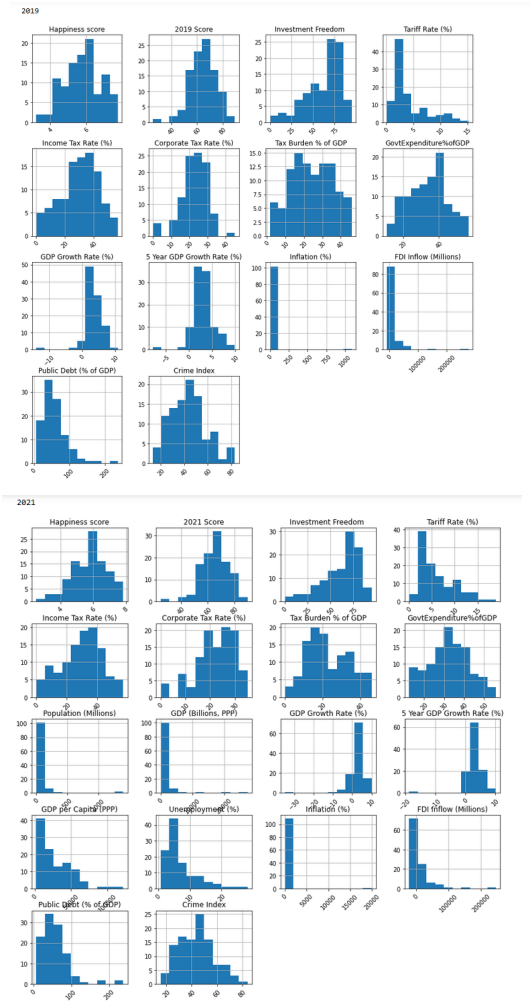
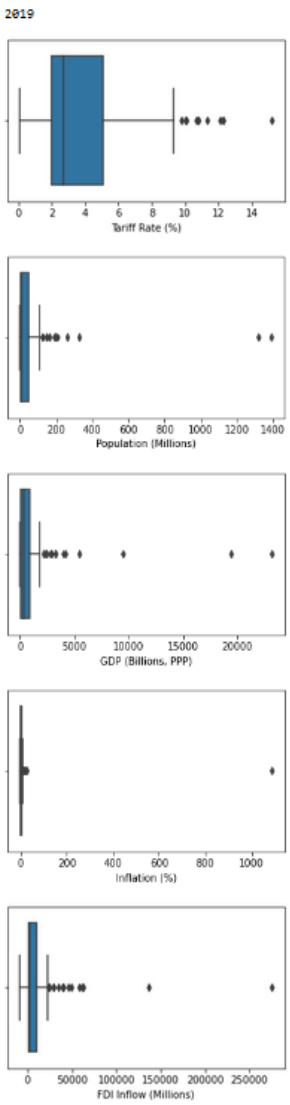


Fig. 5: Histogram plot for each variable inside the datasets, in order to find possible atypical values. As an example, only the histograms for 2019 and 2021 are included. From looking at these histograms, it is evident that variables like Tariff Rate, Population, GDP, Inflation and FDI Inflow present atypical values, which need to be analyzed. **Note: full period (2018-2022) results can be found at the public Github repository created for this paper, on *https://github.com/andresbasilea/EconomicFreedom_ADataMiningApproach*.**

Once the histograms were created and analyzed, it is obvious that variables such as Tariff Rate, Population, GDP, Inflation and FDI Inflow present atypical values, which need to be analyzed. In order to do this task, we created boxplots specifically for these variables in each of the datasets. The boxplot analysis helped us identify on a clearer fashion the existing atypical values and the range until which we can consider as "normal" the values of a variable. Those variables with values outside of the norm, were analyzed individually.



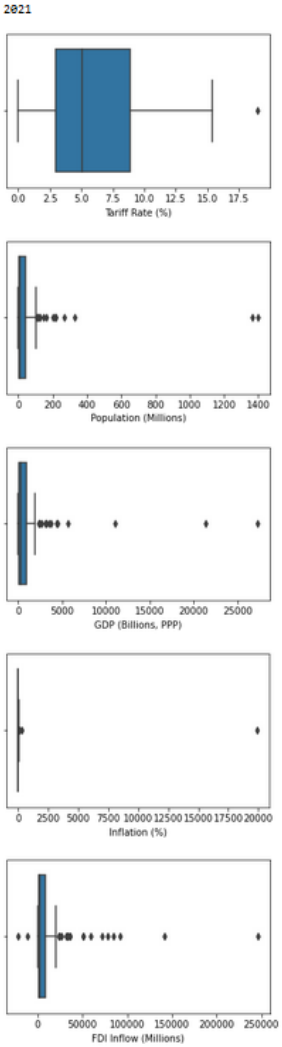


Fig. 6: Boxplots to analyze the atypical values inside the variables previously identified on the Histogram plots. As an example, the boxplots for 2019 and 2021 datasets are included.

Once the boxplots were printed, the use of the *max* Python function allowed us to see the maximum values for each column of each year's dataset, and from the information gathered before, a simple for loop was created to obtain the complete row of entries that satisfy the condition of "outside the norm" defined by the boxplot for each variable. From this information, we concluded the following:

- Population maximums correspond to India's or China's population, which is much higher compared with other countries. We will leave this information as it is.
- 'Tariff Rate (%)' atypical value in Zimbabwe for 2022 might correspond to the economic crisis the country is facing. Tariff rates might had been raised in order to encourage national production and consumption. As the data is correct, we will leave it as it is.
- 'GDP' maximums correspond to highly industrialized and economically advanced countries (as United States and China). Even though this values are much larger compared to other countries' GDP's, we will leave this data as it is.
- Venezuela's hyperinflation throws data which is extremely out of range (compared to other countries' inflation). We should delete Venezuela's registries on all dataframes. However, as this information is correct, we will leave it as it is.
- FDI Inflow of the countries with the biggest economies are out of the normal range when compared to other countries. We will leave this data as it is.

```
 2019
Country                          Zimbabwe
Happiness score                   7.7689
2019 Score                          89.4
Investment Freedom                    95
Tariff Rate (%)                     15.2
Income Tax Rate (%)                   57
Corporate Tax Rate (%)                45
Tax Burden % of GDP                 45.9
GovtExpenditure%ofGDP               56.6
Population (Millions)             1390.1
GDP (Billions, PPP)              23159.1
GDP Growth Rate (%)                 10.9
5 Year GDP Growth Rate (%)           9.9
GDP per Capita (PPP)              124529
Unemployment (%)                    27.3
Inflation (%)                     1087.5
FDI Inflow (Millions)             275381
Public Debt (% of GDP)              236.4
Crime Index                          83.2
dtype: object
```

Fig. 7: Results of using Python's *max* function on each of the variables of the datasets. As an example, the figure shows the maximum values for the 2019 dataset.

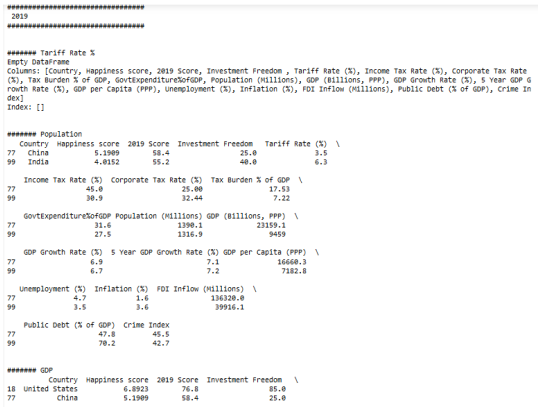


Fig. 8: A simple *for* loop was created to obtain the complete row of entries that satisfy the condition of "outside the norm" defined by the boxplot for each variable. As an example, the figure shows the rows which contain maximum values for some of the variables for the 2019 dataset.

4.3 Variable Selection - Results

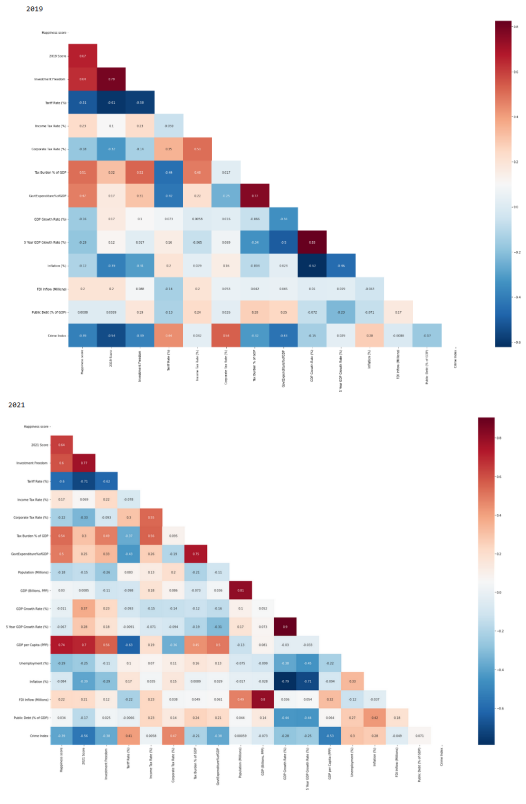


Fig. 9: Correlation plots created using the Pearson method (computed with 1). As an example, only the 2019 and 2021 correlation plots are shown. The difference in the amount of columns for 2021 is due to the increase in columns from the original [Foundation, 2022b] 2022 Economic Freedom dataset.

We continued the analysis with the generation of correlation matrices and correlation plots. The correlation matrices were generated for each of the five datasets, using the Pearson method. As seen in figure 9, certain variables are strongly correlated (Happiness score with 2019 Economic Freedom Score, for example). Some of the variables strongly correlated were kept in the dataset, as the information they provide is valuable. However, in the case of highly correlated variables like GDP Growth Rate and 5 Year GDP Growth Rate, the 5 year GDP growth rate variable was eliminated from the dataset, as the information they can provide to the model is practically the same. Afterwards, the five datasets were concatenated into a single one by placing one below the other, giving us a dataframe shaped with 15 distinct columns and 522 rows. We will use this dataframe to analyze the correlation between the variables more deeply, as well as to generate a Random Forest Regression model to predict the Happiness and Crime rates for a given country by using its economic data. The final dataframe variables used for the prediction model are: Happiness score / Crime score (variable to be predicted), Investment Freedom, Tariff Rate, Income Tax Rate, Corporate Tax Rate, Tax Burden % of GDP, Gov't Expenditure % of GDP, Population, GDP, GDP Growth Rate, Unemployment, Inflation, FDI Inflow, Public Debt % of GDP, and the Economic Freedom Score.

4.4 Correlation Analysis and Model Generation - Results

As mentioned in 3.4, a global correlation analysis between happiness and general economic freedom was done to look into the changing correlations between these two variables. Figure 10 shows the graph created to exhibit this relationship. On it, we see that the minimum correlation value between happiness and the general economic freedom appears in 2018, with a 0.65 (out of 1.0, which is the maximum positive correlation possible) positive correlation. The maximum value of correlation between these variables was achieved on 2022, with a value of almost 0.75 positive correlation (a 15% increase from the 2018 correlation value).
The cause of this increase is out of the scope of this paper, however, the worldwide increase in the correlation between these variables might be due to the Covid-19 global pandemic which started on 2020, were a big percentage of the population was forced to remain inside their homes. The pandemic worsened the economic conditions, as well as pre-existing mental health conditions, and the intolerance for economic and health uncertainty and decreased quality of social interactions (all of them being online) [Lyubomirsky, 2022] might explain the higher correlation between happiness and economic freedom, as they both decreased importantly, though this is subject to a deeper examination.

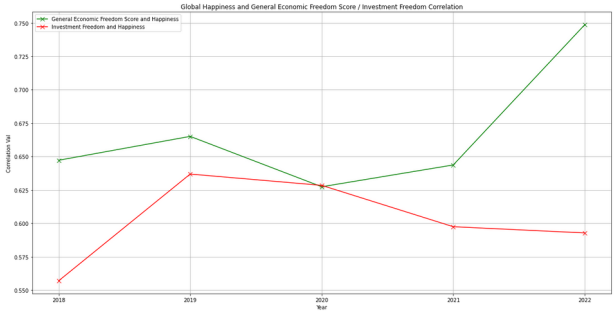


Fig. 10: Graph that shows the correlation between Happiness and the General Economic Freedom Score, as well as Happiness and Investment Freedom throughout the years studied.

Nevertheless, the correlation between investment freedom and happiness (shown in red in figure 10) peaked on 2019 with a value of 0.637, and decreased ever since.

Similarly, a correlation analysis between crime and general economic freedom was done to look into the correlation between these two variables. The analysis also includes the correlation between crime and investment freedom, as well as between crime and happiness, as seen on figure 11. On it, we see that the maximum correlation value between crime and the general economic freedom appears in 2021, with a -0.56 (out of -1.0, which is the maximum negative correlation possible) negative correlation. The minimum value of correlation between these variables was achieved on 2022, with a value of -0.525 negative correlation. Thus, it has remained practically unchanged in the past few years, with a high negative correlation denoting the intrinsic relationship between the economic conditions of a country and the level of crime that can be expected; the better the economic conditions, the lesser the crime rate.

The analysis between crime and investment freedom (red on figure 11) and crime and happiness shows lesser negative correlation between the variables, but still an almost unchanged value throughout the years. As with economic freedom and crime, the graph shows that countries with more investment freedom show a lower crime rate; and as the happiness of a country's population increases, the crime rate decreases.

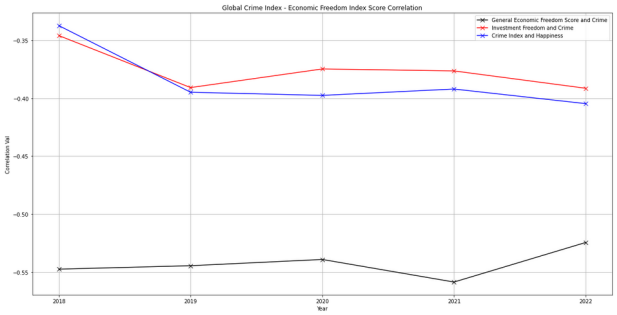


Fig. 11: Graph that shows the correlation between the Crime Index and the General Economic Freedom Score, as well as the Crime Index and Investment Freedom throughout the years studied. Finally, the correlation between Happiness and the Crime Index throughout the years is presented in blue.

Understanding this relationship on a deeper scale requires an examination which is outside of the scope of this paper, however, a clear picture of the behaviour of these variables can be depicted by looking at the graphs presented.
To obtain deeper insight from the datasets, the study data was reduced to a single country, Mexico. This was done to see the variation of the values of happiness, crime and economic freedom, not only the correlation between them. To be able to compare this variables, they were normalized using the min-max normalization formula (see 2) for Mexico in the period 2018-2022.

$$Normalized_x = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{2}$$

The results on figure 12 show that the maximum happiness score in Mexico was achieved on 2019, and it has been decreasing every year until 2022, with the lowest score. Similarly, the general economic freedom score increased from 2019 to 2020, but then decreased on 2021 and reached a minimum on 2022. The crime rate in Mexico, on the other hand, increased from 2018 to 2021 and had a big decrease from 2021 to 2022. Multiple factors could explain this behaviour; however, the decrease in the happiness score in the last 3 years might be due to the toll the pandemic has had on Mexican families, with consumer prices increasing, general economic freedom decreasing and crime rates reaching 5 year maximums, variables deeply correlated as seen on figures 10 and 11 on a worldwide scale.

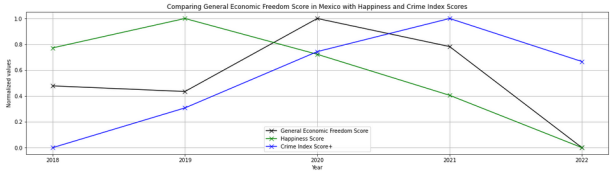


Fig. 12: Comparison that shows the variation of the Happiness Score, Crime Index and Economic Freedom Score in Mexico in the period 2018-2022. The values of the variables have been normalized so that they can be compared directly.

Once the correlation between the variables had been defined, a random forest regression model was implemented, in such a way that the happiness index of a country can be calculated based merely on economic factors, such as inflation, tax rate percentage, GDP and Government Expenditure. Two Random Forest Regressor objects were created, one with the variable "Happiness score" as the value to predict and one with the variable "Crime Index" as the variable to predict. These Random Forest Regressor objects were trained with a 80 percent of the total dataset size, using 5000 estimators, and a maximum depth of 12 for each tree in the random forest. The *RandomForestRegressor* model from the *sklearn* Python library was imported for this purpose; as well as the *sklearn* metrics to determine the *r2 score* or *Coefficient of Determination* of each model. A *Linear Regression* model was also implemented, but the non linearity of the data showed this approach inefficient, with a score of the Coefficient of Determination of less than 60%, reason for which the Random Forest Regression was the best approach.
With our model, the happiness prediction based on 14 economic variables, showed the following results:

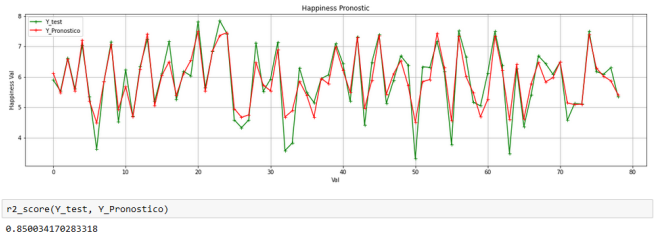


Fig. 13: Happiness Prediction using the Random Forest Regressor. The graph shows the prediction values for each entry in the test set in red, and the original test value for each entry in green. It can be seen that the Random Forest Regressor values are very similar to the real data.

A coefficient of determination of 85% indicates that the model prediction fits the observed data with a value of 0.85 out of 1, which means that 85% of the time, the model will fit the observed data perfectly. Furthermore, a prediction object was created so that predictions of happiness in a country can be achieved only by entering manually the 14 class variables values (this values are often published by the government on a quarterly, bi-annually or annually basis, making the predictor object useful to have an approximate of how happy the inhabitants of a country are or will be). The happiness score of a country, as mentioned before, has an important effect on the way people in that country behave and work, so knowing

how the economic factors are affecting the happiness of the people can be beneficial in order to take economic measures, or create economic and social policies.
In a similar manner, the crime prediction based on 14 economic variables, showed the following results:

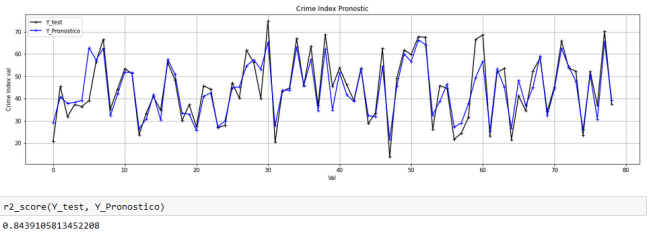


Fig. 14: Crime prediction using the Random Forest Regressor. The graph shows the prediction values for each entry in the test set in blue, and the original test value for each entry in black. It can be seen that the Random Forest Regressor values are very similar to the real data.

A coefficient of determination of 84.39% indicates that the model prediction fits the observed data with a value of 0.8439 out of 1, which means that 84.39% of the time, the model will fit the observed data perfectly. On the same way as with the happiness prediction object, a crime prediction object can be created that predictions of crime in a country can be achieved only by entering manually the 14 class variables values. The crime rate of a country, as mentioned before, has an important effect on the way people in that country behave and the general well-being one can expect living in that country, so knowing how the economic factors are affecting the crime rate in a country can be beneficial in order to take economic measures, or create security, social and economic policies.

5 Concluding remarks

This article has been proof of the existing relationship between the Economic Freedom Index and other sociological metrics such as happiness and crime rate. The data mining approach used during this study allowed a more integral and holistic vision of economic and societal factors and its intrinsic relationship.
Obtaining the data for this paper was relatively straightforward, as it is all publicly available and free to download. A data dictionary was written to understand better the definition of each variable and the values in which they vary.
An exploratory data analysis was performed, where a first look into the variables was done, followed by the cleaning of data as well as the removal of unwanted values. Later, the search for atypical values inside the datasets was executed to find those values which could be entering noise into our data. Lastly, a correlation analysis was carried out following the Pearson coefficients methodology.

After the exploratory data analysis was performed, the variables were selected and the dimensionality decreased by identifying those variables with high coefficients of correlation and removing those which provided similar information to other variables in the dataset.

A correlation analysis was performed to understand the variation of correlation between the economic freedom index and other indices as happiness and crime throughout the years 2018-2022. It was found that happiness and the general economic freedom appears in 2018, with a 0.65 (out of 1.0, which is the maximum positive correlation possible) positive correlation. The maximum value of correlation between these variables was achieved on 2022, with a value of almost 0.75 positive correlation (a 15% increase from the 2018 correlation value). The pandemic worsened the economic conditions, as well as pre-existing mental health conditions, and the intolerance for economic and health uncertainty and decreased quality of social interactions (all of them being online) [Lyubomirsky, 2022] might explain the higher correlation between happiness and economic freedom, as they both decreased importantly, though this is subject to a deeper examination out of the scope of this paper. We also saw that the maximum correlation value between crime and the general economic freedom appears in 2021, with a -0.56 (out of -1.0, which is the maximum negative correlation possible) negative correlation. The minimum value of correlation between these variables was achieved on 2022, with a value of -0.525 negative correlation. Thus, it has remained practically unchanged in the past few years, with a high negative correlation denoting the intrinsic relationship between the economic conditions of a country and the level of crime that can be expected; the better the economic conditions, the lesser the crime rate. It was also found that as the happiness of a country's population increases, the crime rate decreases. An analysis was also performed to understand the changing values of happiness, crime and economic freedom specifically for Mexico. Finally, a Random Forest Regressor model was created to predict the values of happiness and crime in a country based on 14 economic factors such as inflation and GDP, obtaining a 85% coefficient of determination for the happiness model and an 84.39% coefficient of determination for the crime rate model.

Further work should include a deeper analysis on why the correlation between this indices is so high, trying to find a causality relationship which could be beneficial to create

new economic, security, and social policies. An analysis by nation, instead of a worldwide analysis could be more useful if the model created here was pretended for use when creating governmental policies.

References

[Daniel Gropper, 2011] Daniel Gropper, Robert Lawson, J. T. (2011). Economic freedom and happiness. *Cato Journal 31: 2*.

[Foundation, 2022a] Foundation, H. (2022a). *2022 Index of Economic Freedom*.

[Foundation, 2022b] Foundation, H. (2022b). About the index. *2022 Index of Economic Freedom*.

[Granillo, 2019] Granillo, M. G. (2019). *Breve Historia de las Doctrinas Económicas*.

[James D. Gwartney, 1999] James D. Gwartney, Robert A. Lawson, R. G. H. (1999). Economic freedom and the environment for economic growth. *Journal of Institutional and Theoretical Economics (JITE) 155: 643-663*.

[John F. Helliwell and Wang, 2022] John F. Helliwell, Richard Layard, J. D. S. J.-E. D. N. L. B. A. and Wang, S. (2022). World happiness report. *World Happiness Report*.

[Lyubomirsky, 2022] Lyubomirsky, K. O.-M. . S. (2022). Social connection and well-being during covid-19. *World Happiness Report*.

[Media, 2022] Media, M. (2022). 'self-driving' ai advances in buy-side trading. *Traders Magazine*.

[Solow, 1956] Solow, R. (1956). *A Contribution to the Theory of Economic Growth*.

[WPR, 2022] WPR (2022). Crime rate by country. *World Population Review*.