A Path to Global Economic Prosperity:

Economic Freedom

By Renyu Gu

Abstract

According to the International Monetary Fund, global economic growth was projected to be reduced from 2.3% in 2019 to -4.9% in 2020 due to the negative impact of the Covid-19 pandemic.[1] This downturn affects millions of individuals, as well as governments, as the economic activities have been brought to a near-standstill.[2] Economic freedom has a significant effect on economic growth and prosperity, it also provides a proven formula for economic progress and success. [3] In this report, the analysis of global economic freedom has been studied through data science, referring to an unsupervised machine learning model. The Economic Freedom Index [4] from the Heritage Foundation and the World bank are used as sources for this research. By exploratory analysis and preprocessing, the data has been utilized effectively with the focus of clustering countries and analyzing correlations between each economic factor. Charts and graphs are formed along with the inspection of country clusters and factor correlations, to provide a comprehensive set of principles and facts to global economic prosperity.

Keywords

Economic prosperity, economic freedom, Economic Freedom Index, hierarchical clustering

1. Introduction

According to the World Bank, because Covid-19 took hold at the beginning of 2020, every region is subject to substantial growth downgrades. East Asia and the Pacific will grow by a scant 0.5%. South Asia will contract by 2.7%, Sub-Saharan Africa by 2.8%, Middle East and North Africa by 4.2%, Europe and Central Asia by 4.7%, and Latin America by 7.2%. These downturns are expected to reverse years of progress toward development and tip tens of millions of people back into extreme poverty.[5] Economic prosperity is the key factor that determines the quality of life for individuals as well as the future of a whole nation. The economy must prompt entrepreneurship, innovation, competition, and sustainability of diversity in order to enhance people's living standards.[6] Evidence from the Heritage Foundation and the World Bank has shown that the free-market system that is rooted in the principles of economic freedom— empowerment of the individual, nondiscrimination, and open competition—has fueled unprecedented economic growth.[7] According to figure 1, the global poverty rate has decreased by 67% while the world economy has doubled in size as economic freedom increased, showing that economic freedom brings prosperity. [8] Individuals have economic freedom when (a) property they acquire without the use of force, fraud, or theft is protected from physical invasions by others and (b) they are free to use, exchange, or give away their property if their actions do not violate the rights of others. The benefits of economic freedom are also illustrated in other scientific studies. E.g., Smith [9], Mises [10], and Rothbard [11] promoted the idea that private property is at the core of economic freedom [12]. Gwartney and his group [13] provided statistical support to this thesis using the legal system variable and the protection of property rights in calculating the Economic Freedom Index (IEF). Thus, an index of economic freedom

should measure the extent to which lawfully acquired property is protected and individuals are engaged in voluntary transactions. [14]

In this study, the Economic Freedom Index [4] from the Heritage Foundation is referenced to help readers track over two decades of advancements towards economic freedom, prosperity, and to promote these ideas in their homes, schools, and communities. The Index covers 186 countries, including factors such as Property Rights, Government Integrity, Tax Burden, Fiscal Health, Business Freedom, Monetary Freedom, Trade Freedom, Tariff Rate (%), GDP Growth Rate (%), GDP per Capita (PPP), Unemployment (%), Inflation (%), etc. [7] Country clustering is also found to be critical when examining the global economy. Rather than examining country-level indicators in isolation, clustering offers the opportunity to determine which countries are similar and to explore the relationships between variables driving cluster membership.[16] By using hierarchical clustering methods to classify countries, from the econ-free group to the less free group, both the advantages and disadvantages of economic freedom are identified, providing a clearer picture of the future economic path of each country.

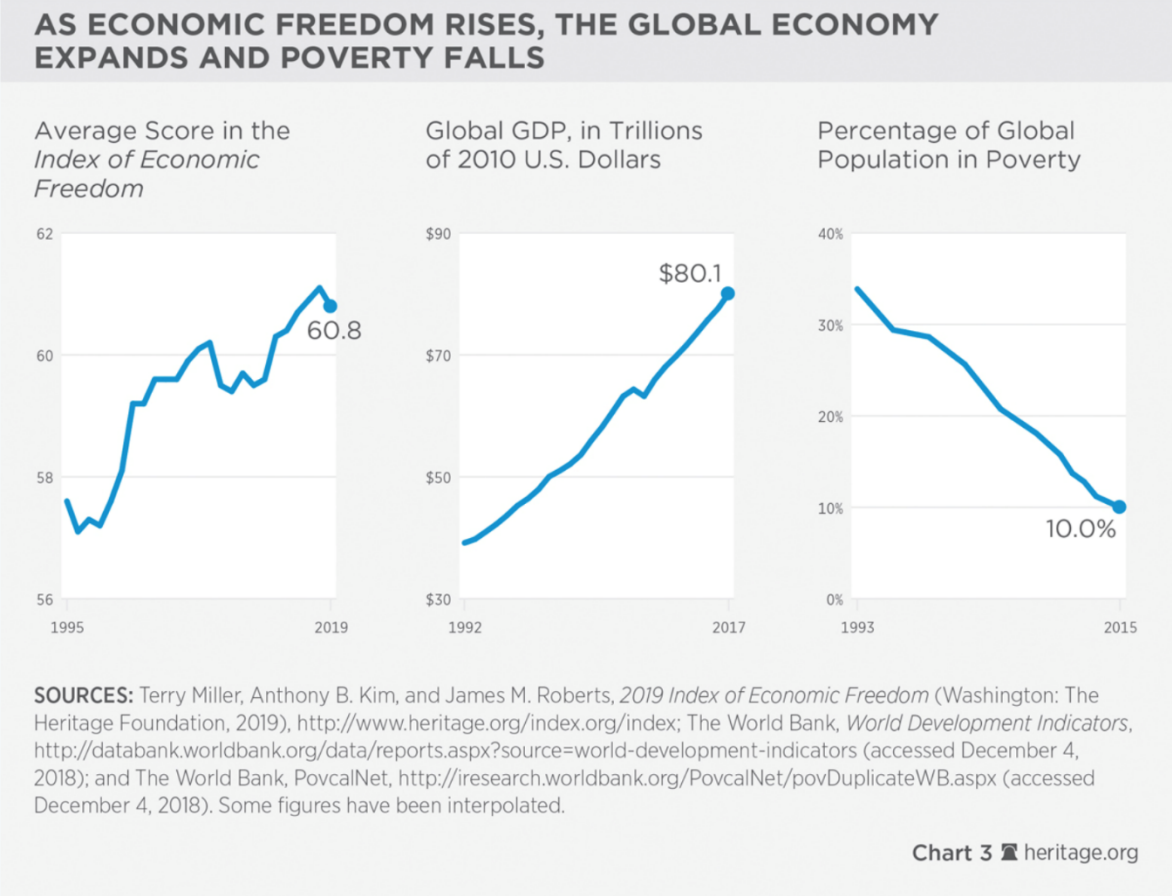


Figure 1: The correlation between global economic freedom and poverty from the 1900s to 2010s (from https://www.heritage.org/index/book/chapter-4?version=164)

2. Method

2.1 Collecting Data and Information

For 26 years, the index of economic freedom published by the Heritage Foundation and the data from the World bank have delivered thoughtful analysis in a clear, friendly, and straight-forward format.[17] The data (last updated 26/02/2019) was presented in CSV format with 186 rows and 34 columns, 12 quantitative and qualitative factors have been categorized into 4 major groups:

- **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)
- **Open Markets** (trade freedom, investment freedom, financial freedom)

For this research, the 2019 economic freedom index has been imported using Python language in the conda environment using the Jupyter Notebook application, where correlations between each factor have been analyzed and the unsupervised learning algorithm (Hierarchical Clustering) has been implemented..

Preprocessing methods have been used to identify missing values and to rearrange effective information from the data. Columns including 'Country Name', 'WEBNAME', 'Region Rank', 'CountryID' are dropped because the data in these columns are repetitive or irrelevant. The missing value is another problem that needs to be addressed for it may weaken the statistical power of the model, and therefore, may lead to invalid conclusions. During preprocessing, missing values have been identified, and the object data type columns have been converted into float data type for better analysis as shown in figure 2. Missing values have been imputed using sklearn K Nearest Neighbors imputer function, which utilizes the k-Nearest Neighbors method to replace the missing values with the mean value from the nearest neighbors found in the dataset. By default, it uses a Euclidean distance metric to impute the missing values.[18]

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 183 entries, 0 to 185
Data columns (total 5 columns):
 #   Column                  Non-Null Count  Dtype
---  ------                  --------------  -----
 0   Population (Millions)   183 non-null    float64
 1   GDP (Billions, PPP)     183 non-null    float64
 2   GDP per Capita (PPP)    183 non-null    float64
 3   Unemployment (%)        178 non-null    float64
 4   FDI Inflow (Millions)   180 non-null    float64
dtypes: float64(5)
memory usage: 8.6 KB
```

Figure 2: the result of conversion of feature columns from object type to float type

2.2 Exploratory Data Analysis

After preprocessing the data, the process of Exploratory Data Analysis (EDA) has been facilitated. Seaborn distribution plot function combines the matplotlib histogram function with the seaborn kernel density estimate (KDE) plot and rugplot (marginal distributions) functions.[19]

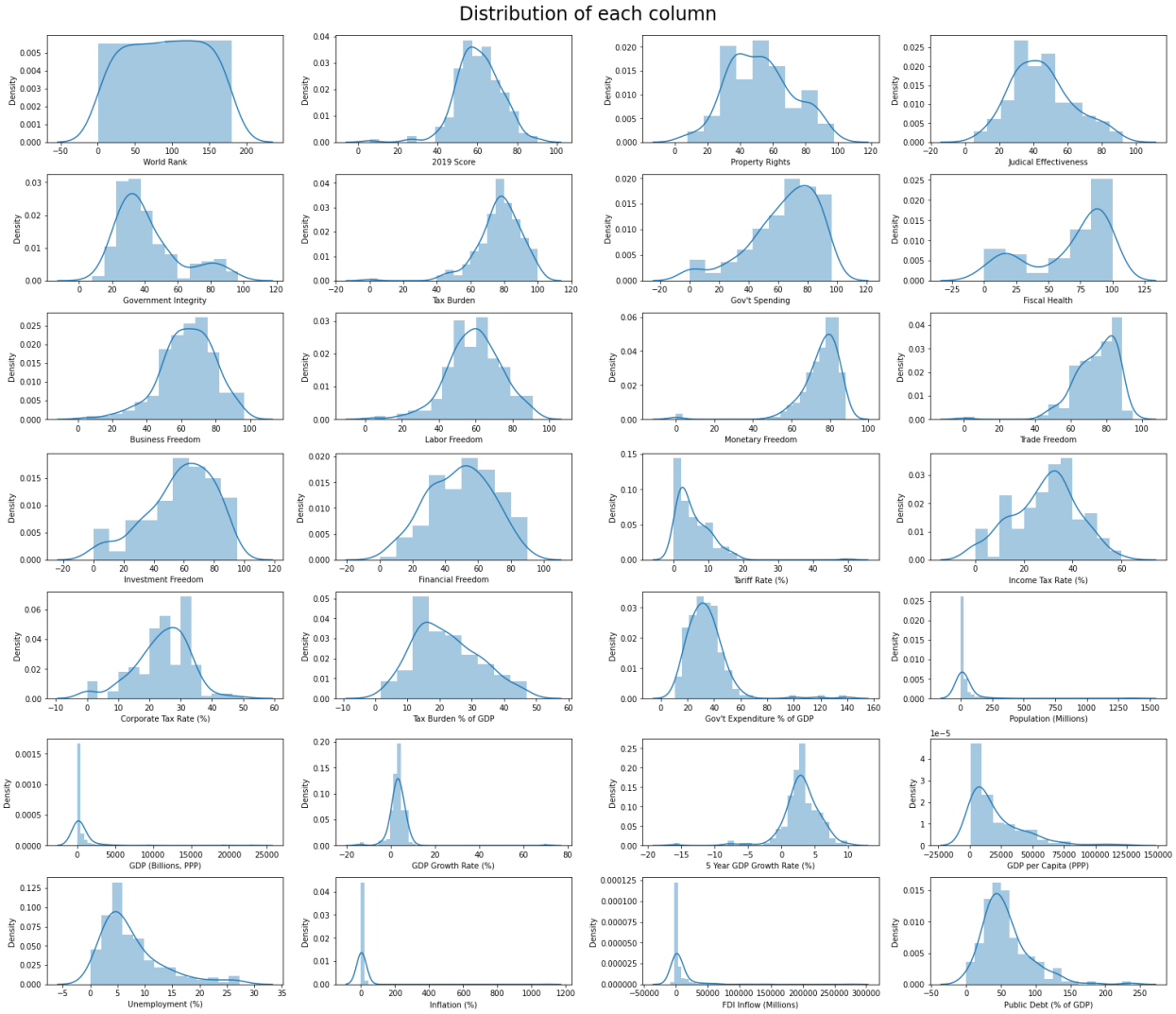
In figure 3, there are 28 sections, each presents one column/feature. The distribution plot shows a univariate distribution of observations. The x-axis shows bins (ranges) of the variable and the y-axis is the probability density function for the kernel density estimation. The World Rank column has been dropped because it is the result of other columns.

Government integrity is derived primarily from Transparency International's Corruption Perceptions Index (CPI) for 2011, which measures the level of corruption in 183 countries. The CPI is based on a 10-point scale in which a score of 10 indicates very little corruption and a score of 0 indicates a very corrupt government. In scoring freedom from corruption, the Index converts the raw CPI data to a scale of 0 to 100 by multiplying the CPI score by 10. The higher

the level of corruption, the lower the level of overall economic freedom and the lower a country's score.[20] According to the government integrity section, there is a peak around 30, demonstrating that the majority group of countries has government integrity around the 20-40 range, having a probability density around 0.03. There is a small peak around 70-90, which indicates that there is around 0.5% group of countries that are less corrupted and have more economic freedom than countries that earned the government integrity score of 20-50.

Per capita Gross Domestic Product (GDP) is a metric that breaks down a country's economic output per person and is calculated by dividing the GDP of a country by its population.[21] Throughout the history of the Index, the per capita economic growth rates of countries that have grown in economic freedom the most are on average at least 50 percent higher than those of countries where freedom has stagnated or slowed [16], presenting that countries that have higher GDP per capita (most are on an average at least 50% higher than the unfree ones) tend to be more econ-free than those who have lower GDP per capita. According to figure 3, the graph of GDP per capita (PPP) demonstrates that the number of countries that are not econ free is significantly higher than econ-free countries, as the density of countries who have around 0-10000 GDP per capita appear to be the highest. As the GDP per capita (PPP) rate goes higher, there is a lower density of countries presented in the graph. As a result, the least number of countries have the highest rate of GDP per capita (PPP) and are the most economically free.

Figure 3: Distribution of each column



2.3 Correlation between each feature column

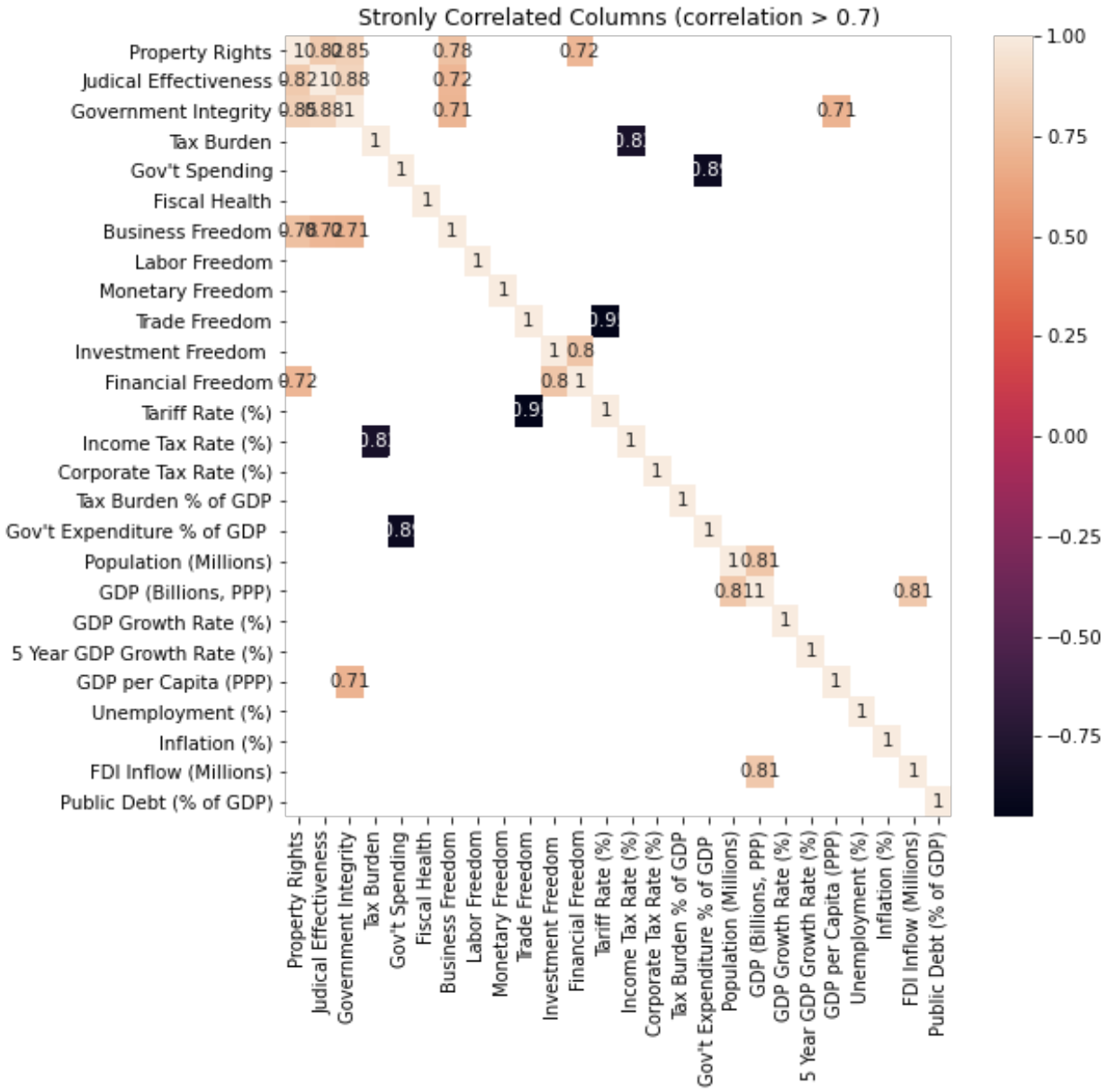


Figure 4: Correlation (>|0.7|) heat map between features

To better analyze the relationship between each column/feature, a heatmap (figure 4) is generated to demonstrate strong correlations between certain factors. Both x-axis and y-axis are labeled with each distinct column name. The vertical color stripes on the right illustrate the

extent of the correlation along with the matching color. The color varies from light to dark, from top to bottom. The light color represents a strong positive correlation between the two factors, and the dark color square shows that the two factors are negatively correlated. In figure 4, the mark of the color square only shows when the correlation coefficient is greater than |0.7| to highlight the strong correlation.

According to figure 4, government integrity appears to have a 0.85 correlation coefficient with property rights, suggesting that they are positively correlated. Furthermore, government integrity is shown to be positively correlated with judicial effectiveness with a coefficient of 0.88, and with business freedom with a coefficient of 0.71. There is also a positive correlation between government integrity and GDP per capita (PPP), suggesting that as government integrity enhances, judicial effectiveness, business freedom, and GDP per capita (PPP) would also increase correspondingly**.**

2.4 Hierarchy clustering

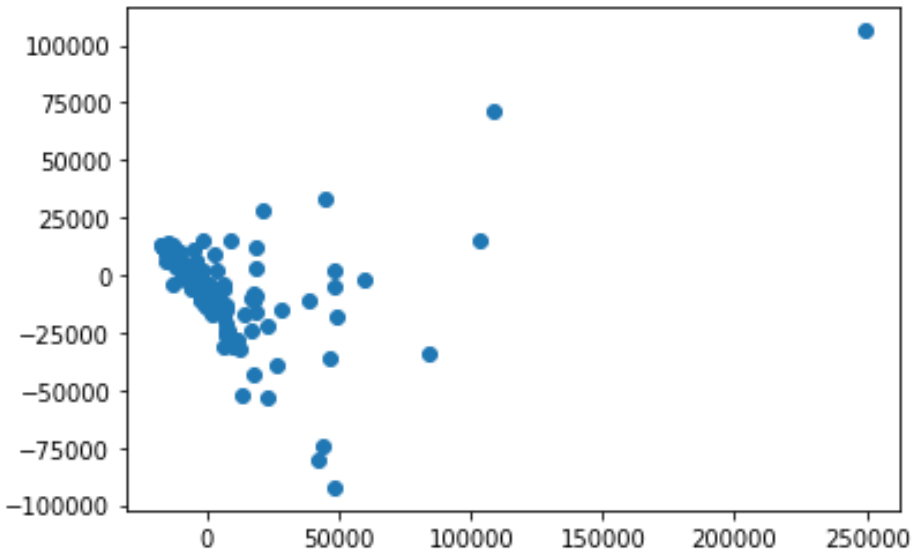


Figure 5: 2-D graph for all countries using PCA

A large dataset is commonly difficult to interpret. Feature extraction was performed using principal component analysis (PCA) as the preparatory step for hierarchy clustering. PCA is a technique for reducing the dimensionality of such datasets, increasing interpretability, but at the same time minimizing information loss.[22] This method allows the dataset to reduce dimensionality while preserving as much of its variability as possible. Through feature extraction, variable inputs are combined in a specific way, and the "least important" variables are dropped while the most valuable parts of the variables are retained. In addition, the assumptions of a linear model require each of the "new variables" to be independent of one another during the process, providing an additional benefit to PCA.[23] In figure 5, the x and y axes represent principal components 1 and 2.

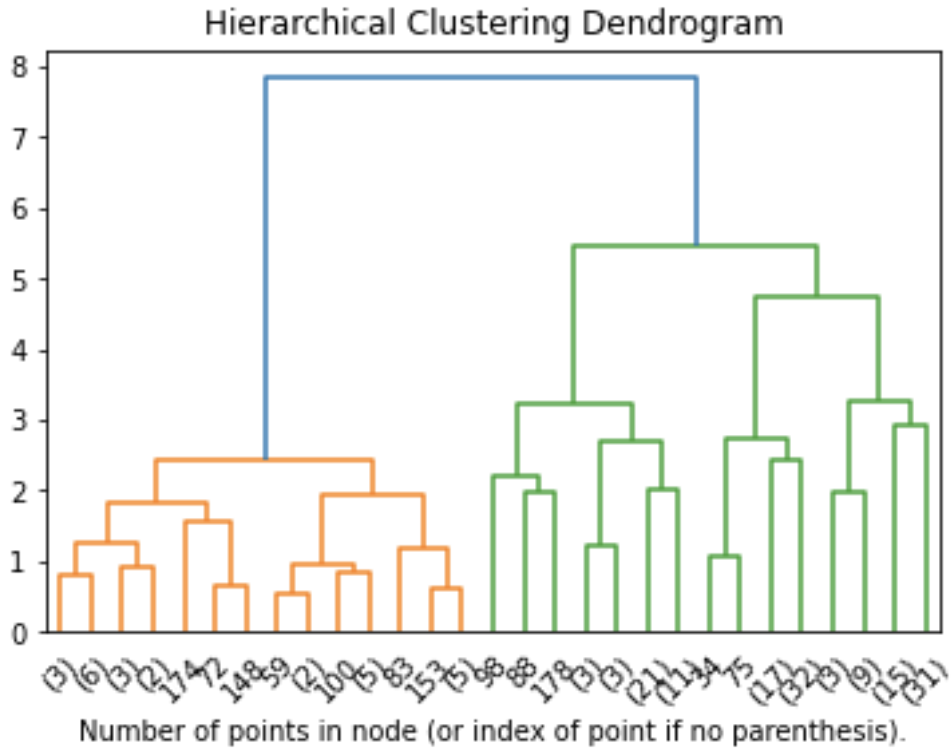


Figure 5: Hierarchical Clustering Dendrogram

Agglomerative Clustering is a member of the Hierarchical Clustering family which works by merging every single cluster in the process that is repeated until all the data have becomes one cluster.[24] During the clustering process, each data point is initially assigned a single cluster. Then the distance between each cluster is measured and the distance matrix is calculated using Euclidean distance measurement:

$$d = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

Euclidean distance calculation

After the measurement, linkage criteria are used to merge the clusters. Using this approach, a given cluster is merged into another when a member of that cluster is close to at least one member of the other cluster. Hence, this requires that only a "single" connection be found between cases in the two clusters before merging them together.[25] Sequentially, updating the distance matrix, and finally repeating the whole process until every data point becomes one cluster.[24]

$$D(X, Y) = \min_{x \in X, y \in Y} d_{x,y}$$

Mathematical formula for linkage criteria

According to figure 5, it appears that there are two stable ways to cluster the data: dividing data into 2 clusters and into 4 clusters. To secure a meaningful categorization of various types of countries, countries clusters are arranged into 4 separate clusters. Figure 6 depicts the PCA colored by hierarchy cluster labels.

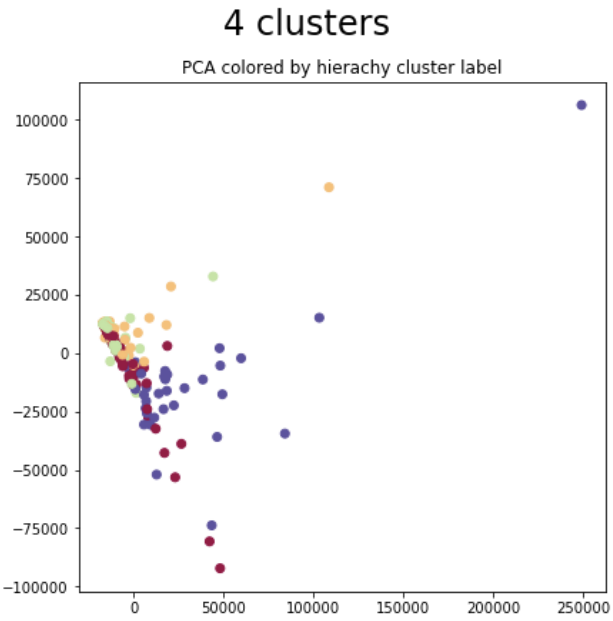


Figure 6: 4 country clusters in PCA analysis

4 cluster	2019 Score
0	65.086538
1	47.785088
2	57.625000
3	76.068966

Figure 7: each cluster's average economic freedom score

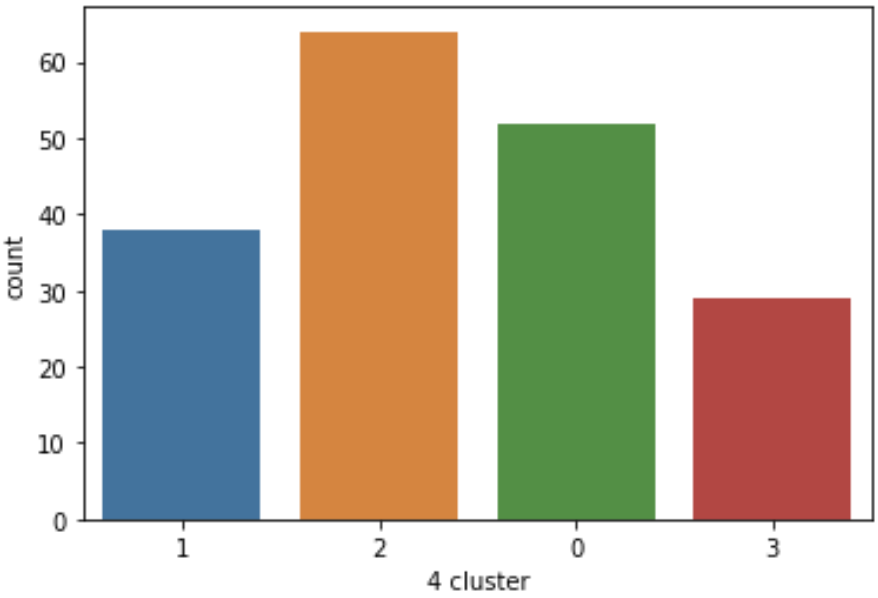


Figure 8: Country Count Plot in Four Clusters

In figure 8, the x-axis represents the 4 different clusters, and the y-axis represents the number of countries. From each cluster's mean economic freedom score (shown in figure 7), cluster 1 represents the least econ-free country; cluster 2 represents the mostly unfree country; cluster 0 represents the moderately free country; cluster 3 represents the freest country.

3. Result

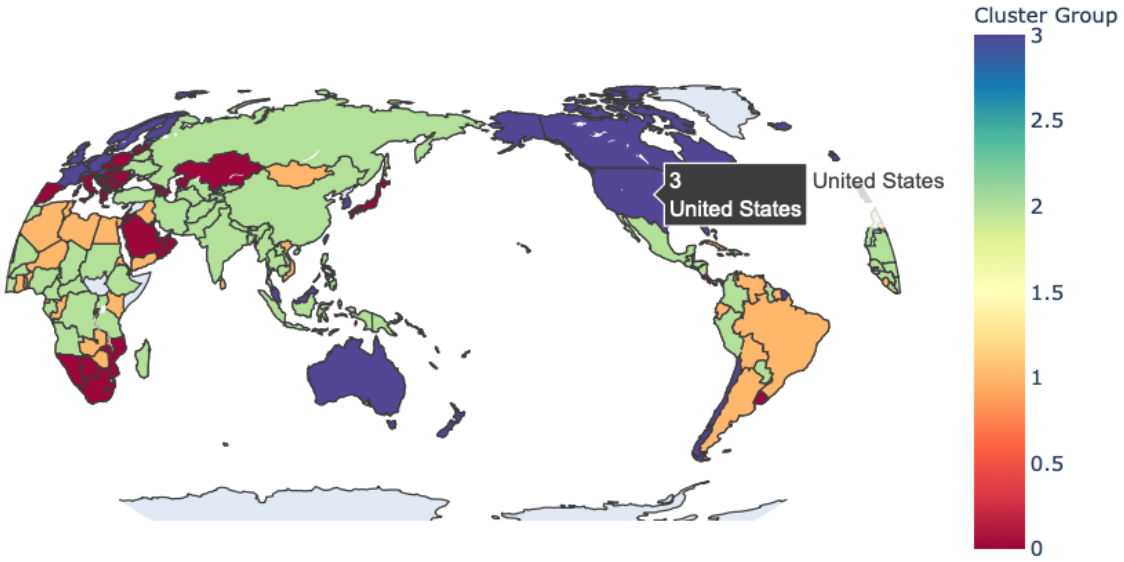


Figure 9: The interactive four cluster world map.

From economically unfree to free, 1 (yellow) - 2 (green) - 0 (red) -3 (purple)

For better visualization, country clusters are converted into an interactive world map (figure 9) using iplot from Plotly (Python Open Source Graphing Library). This interactive map in figure 9 was exported to the html file: file:///Users/Renyu.Gu/Downloads/map4.html. In figure 9, the color strip on the right demonstrates the corresponding color to the four cluster groups. According to figure 8 7, the economic freedom score of each cluster is in the sequence: 1<2<0<3 (from most unfree to most free). Matching this sequence with the color strip in figure 9, yellow represents the most unfree cluster, green represents the moderately unfree cluster, red represents the moderately free country cluster, and the purple section at the 3 represents the freest cluster. According to the map, most unfree countries are concentrated on the South America continent and North Africa, countries on other continents such as Mongolia, Vietnam, Iraq, and Yemen. Most moderately unfree countries are concentrated in Central Africa, central America, Asia, and

eastern Europe, including China, Russia, Mexico, etc. Most moderately free countries are in South Africa and East Europe; Kazakhstan, Saudi Arabia and South Africa are also in the moderately free clusters. Most free countries are concentrated on the North American continent, Europe, and Australia including the United States, United Kingdom, France, Australia, Canada, etc.

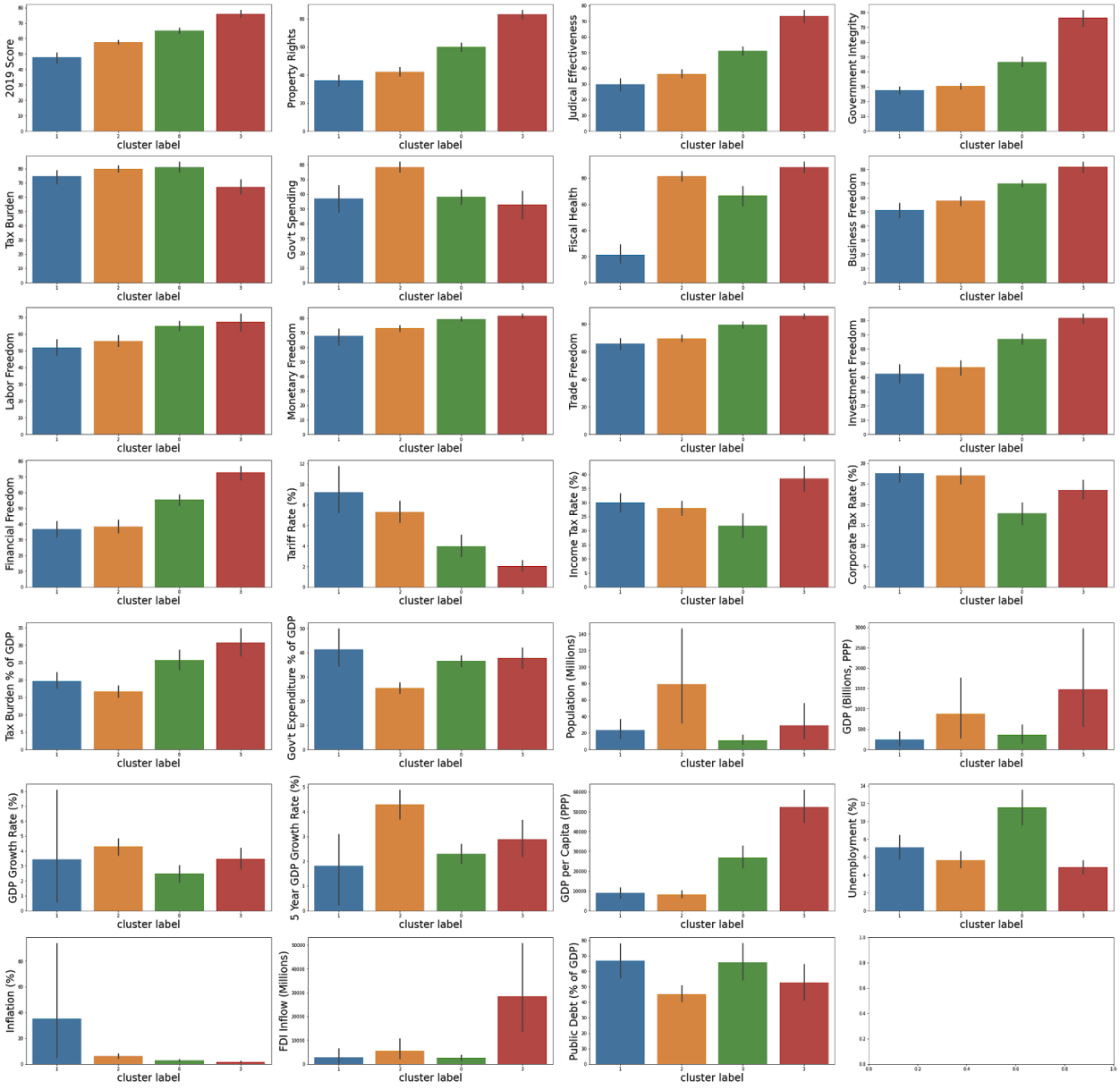


Figure 10: Barplot of each Feature Colored by Four Cluster Labels

Details of cluster features are shown in figure 10. Each feature column is represented by one of the 27 subplots. In the x-axis, the clusters remain in a sequence: 1<2<0<3 (from the most unfree to the freest), this is presented in the first subplot, in which the y-axis is the 2019 score. As the cluster is freer, the score is higher. Many features appear to be positively correlated with the 2019 score, including government integrity, judicial effectiveness, property rights, business freedom, trade freedom, financial freedom, PPP, etc. In this feature group, the most unfree clusters have the lowest value, and the freest ones have the highest values. Among 27 sections, tariff rates and inflation sections that are negatively correlated with economic freedom score. The tariff rate and inflation in the most unfree clusters are the highest, and it descends as the clusters become freer. Other sections appear to be less ordered in sequence. For instance, the fiscal health section shows that the more unfree country has significantly lower values compared to the other three clusters. The population section also appears that the moderately unfree cluster gets the highest number, and the other three clusters have significantly less population.

4. Conclusion

Utilizing the Python language, this study processed and analyzed the data from the Heritage Foundation. Principal component analysis and hierarchical clustering were employed. The results derived from the interactive world map show that the most unfree countries are generally located in Africa and South America; moderately unfree countries are geographically concentrated in central Africa, eastern Europe, Asia, and Central America. Moderately free countries are mostly located in South Africa, eastern Europe, and Saudi Arabia. Most econ-free countries are in North America, Europe, and Australia. The barplot of each feature colored by four cluster labels is created to pinpoint the correlation between each feature and the extent of economic freedom. From figure 10, factors such as government integrity, GDP per capita (PPP), judicial effectiveness, property rights, business freedom, investment freedom, among others, are shown to be positively correlated to the economic freedom score, in which the freest countries have the highest value. This demonstrates that in order to become econ-free and enjoy greater economic prosperity, a country should improve these positively correlated factors primarily by developing government regulations to promote domestic entrepreneurship and eliminating government corruption. There are also factors that are negatively correlated with the economic freedom score: tariff rate and inflation. The most unfree clusters have the highest value in these two factors and have the opposite trend in the economic freedom score. This shows that to have economic prosperity and freedom, a country should promote free trade by lowering its tariff rate, and manage inflation.

Other features appear to have a relatively weak correlation to economic freedom. The unemployment rate, for instance, has the highest value on cluster 0 (moderately free). Despite cluster 0, the rest of the 3 clusters are negatively correlated to the economic freedom score.

Analysis showed that the major reason why moderately free clusters have the highest employment was the presence of many southern African countries in this cluster, which are at the top of the world rank of the unemployment rate.[26] South Africa, for example, has a significantly higher unemployment rate than other countries due to a high search cost for workers, compounded by high turnover in the labor market, and high reservation wages due to familial support.[27] The GDP plot also appears to have a weak correlation because despite the other three clusters being positively correlated to the economic freedom score, cluster 2 (moderately unfree) disrupted the order and appears to have higher GDP than cluster 0 (moderately free). This could be explained by the GDP world rank.[28] China and India are placed at the top of the GDP world rank, and these two countries are both in cluster 2, which lifts the average GDP of cluster 2 to a higher level, causing it to be higher than cluster 3. Furthermore, the FDI Inflow plot also shows a weak correlation trend. This could be explained by the FDI net flow data from the World Bank, which indicates that the United States has the highest FDI net flow, and it is 2 times as much as the net flow of the second country, China.[29] This shows that the most economically free country receives significantly more foreign investments compared to the rest of the world.

Thus, to have economic freedom that leads to economic prosperity, a country is required to develop government regulations to promote entrepreneurship, to encourage foreign trade, and to control inflation. Still, many questions remain. For example, other variables that are not included in this data may be significant to economic freedom. Furthermore, the proportion of the "2019 score" factor's influence is unknown, compared to the rest of the features in the process of clustering. A different clustering result could have been reached if completed without the "2019

score" variable. Also, it could be beneficial to compare 2019's data to 2020's data for further exploration of this issue. Future studies are necessary to address these unsolved problems.

Renyu Gu is a junior at Saint Andrew's School, where she enjoys learning economics, computer science, and biology. In 2021, Renyu won First Place and the Special Award of the International Science and Engineering Fair in Delaware. She also enjoys volunteer activities. Renyu volunteers with Operation Smile China and is a co-funder of Art Bond, a non-profit social enterprise that raises funds on an online platform through the sale of paintings created by people with autism to benefit local autism organizations.

5. References

[1] "World Economic Outlook Update, June 2020: A Crisis Like No Other, An Uncertain Recovery." International Monetary Fund, *IMF*, 1 June 2020, www.imf.org/en/Publications/WEO/Issues/2020/06/24/WEOUpdateJune2020.

[2] "The Global Economic Outlook During the COVID-19 Pandemic: A Changed World." *World Bank*, 8 June 2020, www.worldbank.org/en/news/feature/2020/06/08/the-global-economic-outlook-during-the-covid-19-pandemic-a-changed-world.

[3] "Economic Freedom: Policies for Lasting Progress and Prosperity." *Economic Freedom: Policies for Lasting Progress and Prosperity | 2020 Index of Economic Freedom Book*, 2020, The heritage foundation, www.heritage.org/index/book/chapter-2.

[4] Economic Freedom Index, "About The Index." *2020 Index of Economic Freedom | The Heritage Foundation*, 2020, www.heritage.org/index/about?version=1104.

[5] "The Global Economic Outlook During the COVID-19 Pandemic: A Changed World."*World Bank*, 2020, www.worldbank.org/en/news/feature/2020/06/08/the-global-economic-outlook-during-the-covid-19-pandemic-a-changed-world.

[6] *Recover Better Economic and Social Challenges and Opportunities* . United Nations, Department of Economic and Social Affairs, New York, 2020.

[7] "The Power of Economic Freedom." *The Power of Economic Freedom | 2020 Index of Economic Freedom Book*, 2020, www.heritage.org/index/book/chapter-4?version=164.

[8] "About The Index." *2020 Index of Economic Freedom | The Heritage Foundation*, www.heritage.org/index/about.

[9] Smith, Adam. *An Inquiry into the Nature and Causes of the Wealth of Nations*. Liberty Press, 1981.

[10] *Human Action: a Treatise on Economics*, by Ludwig Von Mises, Ludwig Von Mises Institute, 1998, pp. 31–33.

[11] Rothbard, Murray N. *What Has the Government Done to Our Money?* LvMI, 2010.

[12] G. Žurga, *"Contemporary values in public administration and indication of their further development,"* Administratie si Management Public, vol. 29, pp. 108-127, 2017.

[13] J. Gwartney, R. Lawson and E. Gartzke, *Economic Freedom of the World: 2008 Annual Report*, Fraser Institute Publishing House, Canada, pp. 54-59, 2008.

[14] Stansel, Dean, *An Economic Freedom Index for U.S. Metropolitan Areas,* Journal of Regional Analysis and Policy, Forthcoming, August 2012
Available at SSRN: https://ssrn.com/abstract=2131179.

[15] *CHAPTER 4 THE POWER OF ECONOMIC FREEDOM*. 2019, www.heritage.org/index/pdf/2019/book/chapter4.pdf.

[16] A. Grien, S. Sethi, L. Tatum, *"A dynamic analysis of country clusters, the role of corruption, and implications for global firms"* Journal of Economics and Business, vol. XIII, No. 2 pp. 33-60, 2010.

[17] "About The Index." *2020 Index of Economic Freedom | The Heritage Foundation*, www.heritage.org/index/about?version=961.

[18] Htoon, Kyaw Saw. "A Guide To KNN Imputation." *Medium*, Medium, 3 July 2020, medium.com/@kyawsawhtoon/a-guide-to-knn-imputation-95e2dc496e.

[19] "Seaborn.distplot¶." *Seaborn.distplot - Seaborn 0.11.0 Documentation*, seaborn.pydata.org/generated/seaborn.distplot.html.

[20] "Government Integrity." *Government Integrity: Index of How Corruption Affects Economic Freedom*, www.heritage.org/index/freedom-from-corruption.

[21] Staff, Investopedia. "Per Capita GDP Definition." *Investopedia*, Investopedia, 26 Oct. 2020, www.investopedia.com/terms/p/per-capita-gdp.asp.

[22] I. Jolliffe, J. Cadima, *"Principal component analysis: a review and recent developments"*, The Royal Society Publishing, Vol. 374, Issue 2065, 13 Apr. 2016, https://royalsocietypublishing.org/doi/10.1098/rsta.2015.0202

[23] Brems, Matt. "A One-Stop Shop for Principal Component Analysis." *Medium*, Towards Data Science, 10 June 2019, towardsdatascience.com/a-one-stop-shop-for-principal-component-analysis-5582fb7e0a9c.

[24] Wijaya, Cornellius, "Breaking down the agglomerative clustering process" *Medium*, towardsdatascience.com/breaking-down-the-agglomerative-clustering-process-1c367f74c7c.

[25] Cornwell, B. Linkage Criteria for Agglomerative Hierarchical Clustering. In *Social Sequence Analysis: Methods and Applications* (Structural Analysis in the Social Sciences, pp. 270-274), 2015, https://www.cambridge.org/core/books/social-sequence-analysis/linkage-criteria-for-agglomerative-hierarchical-clustering/9DA28CBBBA94983FF4FE7F568E742760.

[26] "Unemployment Rate." *Unemployment Rate - Countries - List*, tradingeconomics.com/country-list/unemployment-rate.

[27] Banerjee, Abhijit, et al. "Why Has Unemployment Risen in the New South Africa.", *National Bureau of Economics Research*, 2007, doi:10.3386/w13167, https://www.nber.org/system/files/working_papers/w13167/w13167.pdf.

[28] "COUNTRY COMPARISON :: GDP (PURCHASING POWER PARITY)." *Central Intelligence Agency*, Central Intelligence Agency, www.cia.gov/library/publications/the-world-factbook/rankorder/2001rank.html.

[29] "Foreign Direct Investment, Net Inflows (BoP, Current US$)." *The World Bank Data, International Monetary Fund, Balance of Payments database, supplemented by data from the United Nations*, data.worldbank.org/indicator/BX.KLT.DINV.CD.WD?most_recent_value_desc=true.