DECOMPOSING AND CLUSTERING WORLD ECONOMIC FREEDOM CLOUD

ABSTRACT

Economic freedom is the fundamental right of every human to control his or her own labor and property. In an economically free society, individuals are free to work, produce, consume and invest in any way they please. For much of the human history, most individuals have lacked economic freedom and opportunity, condemning them to poverty and deprivation. Therefore, understanding the way these problems emerge is of high importance. To do so, we will use the Economic Freedom index provided by heritage.com covering 12 freedoms-from property rights to financial freedom- in 184 countries.

The index of Economic Freedom documents the positive relationship between economic freedom and a variety of positive social and economic goals. The ideals of economic freedom are strongly associated with healthier societies, cleaner environments, greater per capita wealth, human development, democracy and poverty elimination.

This analysis tries to the partition of 184 countries given by the heritage.org based on their economic freedoms and tries to figure out possible explanations why certains countries are good at it and why some others are not. We make use of the unsupervised statistical technique of Principal Component Analysis in order to decompose the high-dimensional indicator space and evaluate the economic-freedom position of each country after which a Cluster Analysis is performed to demostrate how countries are grouped in the cloud of economic freedom.

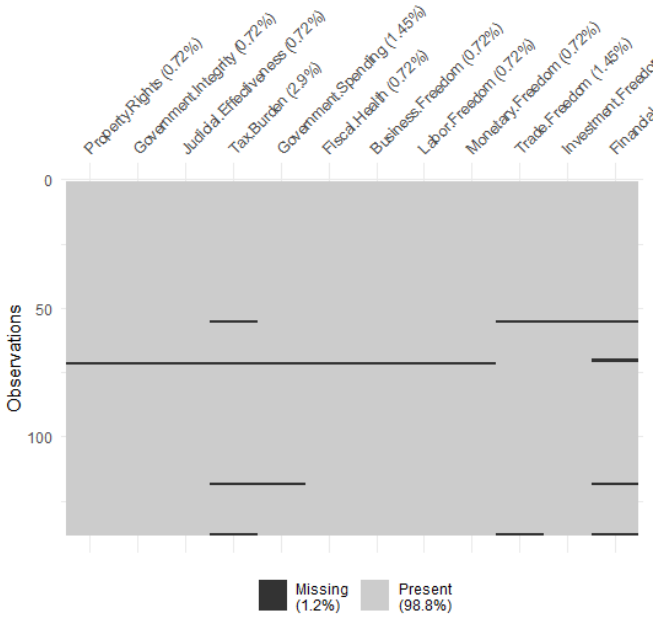
ECONOMIC FREEDOM INDEX STATISTICS

Heritage provides official and realiable data source for measuring Economic Freedom. They measure the economic freedom based on 12 quantitative and qualitative factors, grouped into four broad categories of economic freedom:

1. Rule of Law(property rights, government integrity, judicial effectiveness)
2. Government size(government spending, tax burden, fiscal health)
3. Regulatory Efficiency(business freedom, labor freedom, monetary freedom)
4. Open Markets (trade freedom, investment freedom, financial freedom)

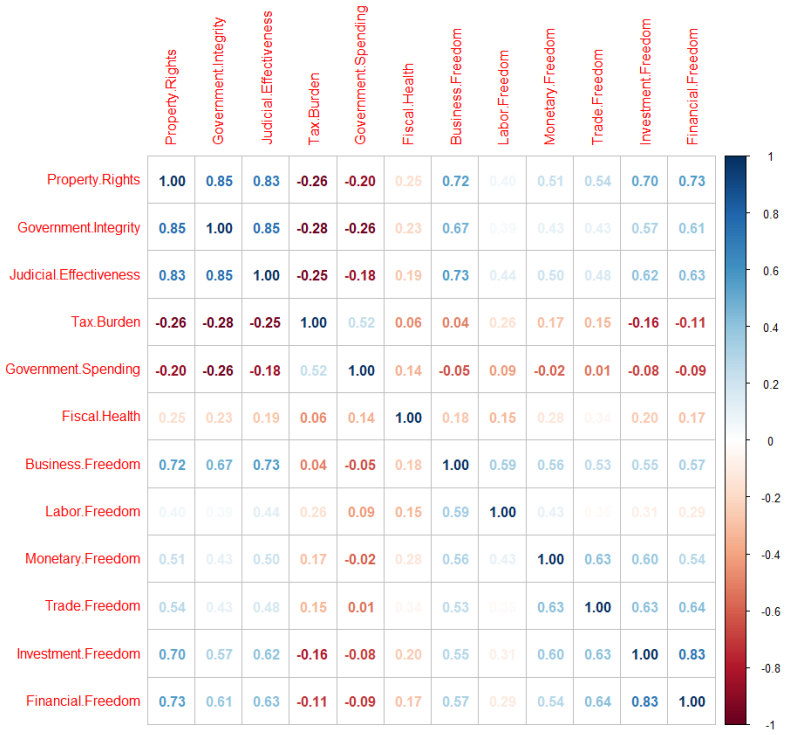
Each of the twelve economic freedoms within these categories is graded on a scale of 0 to 100. A country's overall score is derived by averaging these twelve economic freedoms, with equal weight being given to each.

NAs handling



This plot shows the locations of the missing values in a given row and column. 1.2% of the whole dataset is missing which amounts to only five countries which are Liechtenstein, Yemen, Iraq, Libya and Syria. They are missing because probably it is substantially difficult in these countries to collect the data due to interal conflict that has been going on for years or because they are too small to provide such data(i.e. Liechtenstein). We therefore decided to drop them because it is a negligible part of the whole dataset.

Multicollinearity

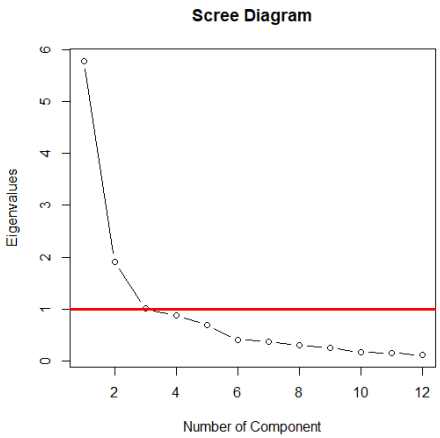


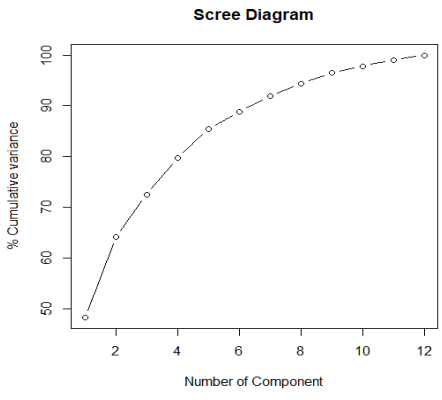
	Property.Rights	Government.Integrity	Judicial.Effectiveness	Tax.Burden	Government.Spending	Fiscal.Health	Business.Freedom	Labor.Freedom	Monetary.Freedom	Trade.Freedom	Investment.Freedom	Financial.Freedom
Property.Rights	1.00	0.85	0.83	-0.26	-0.20	0.25	0.72	0.40	0.51	0.54	0.70	0.73
Government.Integrity	0.85	1.00	0.85	-0.28	-0.26	0.23	0.67	0.38	0.43	0.43	0.57	0.61
Judicial.Effectiveness	0.83	0.85	1.00	-0.25	-0.18	0.19	0.73	0.44	0.50	0.48	0.62	0.63
Tax.Burden	-0.26	-0.28	-0.25	1.00	0.52	0.06	0.04	0.26	0.17	0.15	-0.16	-0.11
Government.Spending	-0.20	-0.26	-0.18	0.52	1.00	0.14	-0.05	0.09	-0.02	0.01	-0.08	-0.09
Fiscal.Health	0.25	0.23	0.19	0.06	0.14	1.00	0.18	0.15	0.28	0.34	0.20	0.17
Business.Freedom	0.72	0.67	0.73	0.04	-0.05	0.18	1.00	0.59	0.56	0.53	0.55	0.57
Labor.Freedom	0.40	0.38	0.44	0.26	0.09	0.15	0.59	1.00	0.43	0.36	0.31	0.29
Monetary.Freedom	0.51	0.43	0.50	0.17	-0.02	0.28	0.56	0.43	1.00	0.63	0.60	0.54
Trade.Freedom	0.54	0.43	0.48	0.15	0.01	0.34	0.53	0.36	0.63	1.00	0.63	0.64
Investment.Freedom	0.70	0.57	0.62	-0.16	-0.08	0.20	0.55	0.31	0.60	0.63	1.00	0.83
Financial.Freedom	0.73	0.61	0.63	-0.11	-0.09	0.17	0.57	0.29	0.54	0.64	0.83	1.00

We have a few variables showing multicollinearity caused by the very structure of the dataset. For example, it is highly intiutive the positive relationship between government integrity,property rights and judicial effectiveness since government integrity is a sort of guaarantee for such institutions. They are highly related with economic variables as well since such institutions are essetial for having economic freedom related to business, investment and finance. We therefore expect to see such variables' loading vectors closely projected to each other. In particular, government integrity, property rights and judicial effectiveness should set out themselves relatively apart from others in the biplot.

PRINCIPAL COMPONENT ANALYSIS

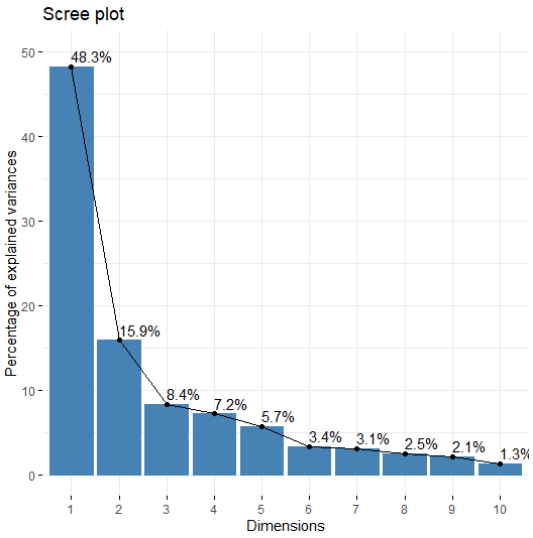
Here we want to figure out amount of information carried by each country namely its score, collecting all the correlated information contained in whole dataset and shrink it. This is done through the PCA analysis that respons to such aim by reducing the dataset dimension to a few components that are keeping most of the information contained in it. In particular, two is a desirable number of dimensions for graphical applications.





Any principal component with $\lambda > 1$ is able to explain more variance than just one variable. We have three eigenvalues satisfying this conditions however the third one is right at the border.

With only two eigenvalues, we are able to explain around 64% of the variablity contained in the whole dataset that is good enough for our exploratory purpose.



The first eigenvalue explains over 48% of the variance while the second around 16%.

Dimensions	eigenvalue	variance.percent	cumulative.variance.percent
Dim.1	5.79069043194095	48.2557535995079	48.2557535995079
Dim.2	1.90993354319429	15.9161128599524	64.1718664594603
Dim.3	1.00202912529292	8.35024271077434	72.5221091702347
Dim.4	0.866775199506875	7.22312666255729	79.745235832792
Dim.5	0.685181820492614	5.70984850410512	85.4550843368971
Dim.6	0.404623269607264	3.37186058006053	88.8269449169576
Dim.7	0.370328251812346	3.08606876510289	91.9130136820605
Dim.8	0.297950604407282	2.48292170339402	94.3959353854545
Dim.9	0.254180566923283	2.11817139102736	96.5141067764819
Dim.10	0.159636472374626	1.33030393645522	97.8444107129371
Dim.11	0.148883175148435	1.24069312623696	99.0851038391741
Dim.12	0.109787539299111	0.914896160825928	100

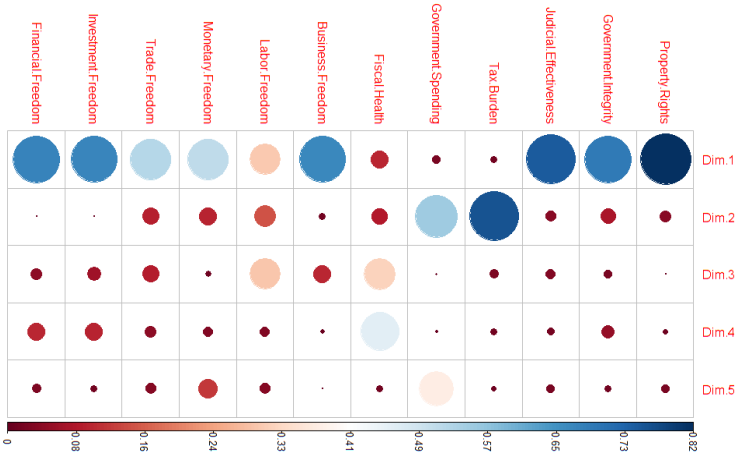
As shown in the table, eigenvalues drop to under 1 after the third component which would add only 8% of variance if included in the analysis.

The sum of all the eigenvalues give a total variance of 12. The proportion of variation explained by each eigenvalue is given in the second column. For example, 5.79 divided by 12 equals to .4825 or about 48.25% of the variation is explained by this first eigenvalue. The cumulative percentage explained is obtained by adding the successive proportions of variation explained to obtain the running total.

RESULTS

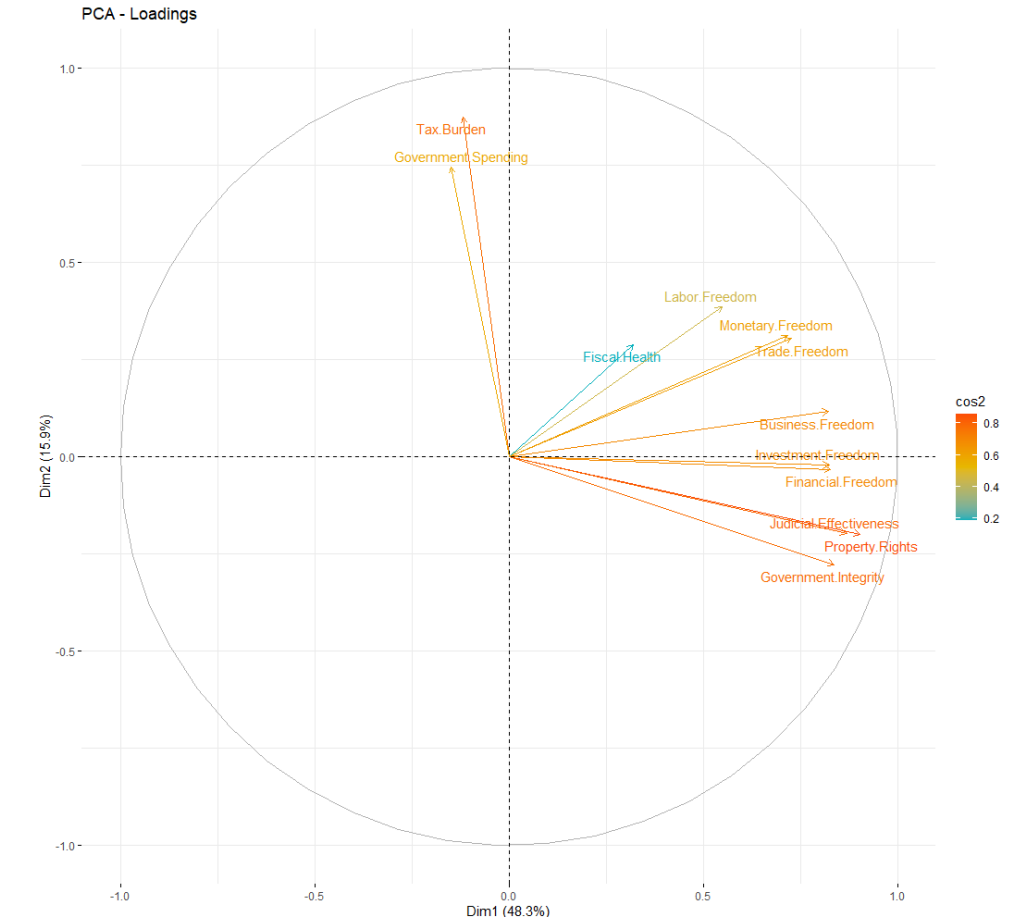
We first present the result of correlation circle between the principal components and the features. The correlation between a variable and a principal component (PC) is used as the coordinates of the variable on the PC. The representation of variables differs from the plot of the observations: The observations are represented by their projections but the variables are represented by their correlations.

Indicators	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5	communality
Property.Rights	0.903673894655057	-0.199297654700422	-0.0329394664186658	0.096284916231259	0.152340677986112	0.889909538761003
Government.Integrity	0.837017031409686	-0.277938625241318	-0.143384123220817	0.225879109500373	0.119637429999668	0.863740883828326
Judicial.Effectiveness	0.868275777080463	-0.197643892687018	-0.179129223442665	0.139802393884063	0.14594475637612	0.865897793321731
Tax.Burden	-0.117325209980524	0.874189471484714	-0.162671435627019	-0.121345979332084	-0.0918252268176782	0.827591151900834
Government.Spending	-0.149057771159788	0.744714306336731	0.044284498432632	0.0555510931820532	0.607753722868522	0.951229245621421
Fiscal.Health	0.31989104722888	0.287628984077428	0.564586597897522	0.680213772807437	-0.120511262093742	0.981032482112455
Business.Freedom	0.822653528791828	0.116043470168614	-0.315687673295671	0.0874370785588845	0.00354046790166395	0.797541400093131
Labor.Freedom	0.547475102316944	0.384972279432201	-0.544431529911246	0.176505875259614	-0.189079621474776	0.81124376160788
Monetary.Freedom	0.717581335306479	0.310988104648286	0.10711089485726	-0.175749799573599	-0.34614018805885	0.773810339649662
Trade.Freedom	0.726463014039094	0.304106829186474	0.296024394985415	-0.202707988684592	-0.18790343968905	0.784258148074626
Investment.Freedom	0.823413930011764	-0.0217110153624277	0.251031430951204	-0.308154755240311	0.121141262923406	0.851133206410788
Financial.Freedom	0.827091338494753	-0.032066521955178	0.207201179255254	-0.317331600988781	0.168766558688518	0.85722216904578



Here we see how our features are correlated with the given dimensions. The first and most important dimension is highly correlated with all the variables except Fiscal Health, Tax Burden and Government Spending. This dimension seems like an indexed shrunk to only one number showing how a country is developed with respect to its economic policies and democratic institutions. Therefore, we can say being projected on the negative side of the X axis representing the first dimension will show us the poor and economically underdeveloped countries. We expect the EU countries to cluster around this happy cone. On the other hand, the second PC is highly correlated with Fiscal Health, Tax Burden and Government Spending. So we recover the information that is not contained in

the first PC about these three features of our dataset. This dimension seems to measure the overall democratic and economic problems in a given country. We expect underdeveloped countries to be correlated with this dimension.



As we expected from our multicollinearity, government integrity, judicial effectiveness and propertiy rights show and lie over almost exact the same direction having the highest correlation with the first PC. On the other hand, Tax Burden and Government Spending becoming best friends on the second quadrant well set apart from other features being the infamous features of this index.

Here is the biplot of our PCA result. As expected, many developed EU and non EU countries fall in the happy fourth quadrant having high correlations with the first PC. On the other hand, worst peforming countries with respect to Economic Freedom is on the second quadrant having high Government Spending and Tax Burden. This is really intiutive to understand. Since, in these countries it is substantially difficult to invest and enjoy financial and economic opportunites equally, all the economy seems to be run by government with government spending coming through high taxes. As an example, we can count Afghanistan and Turkmenistan.