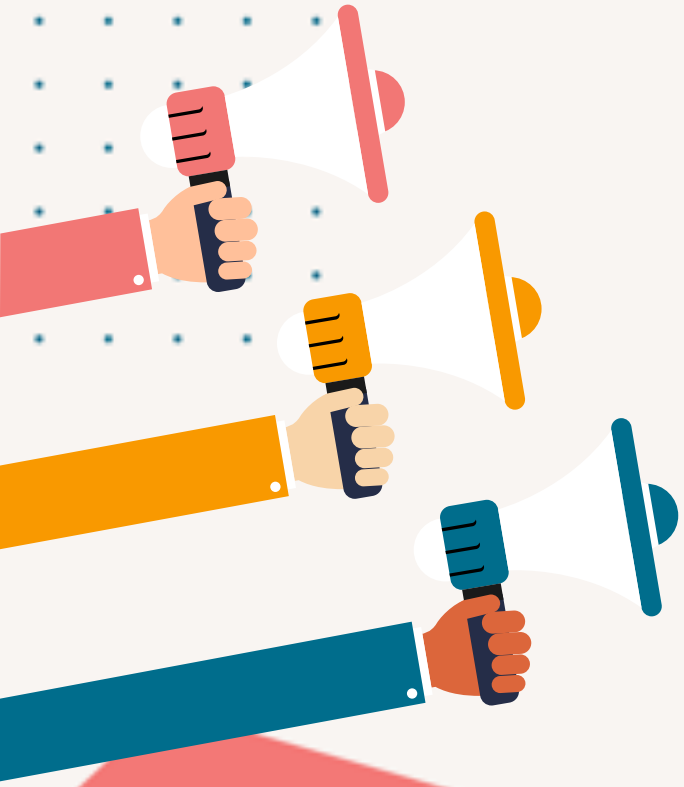
Is Corruption related to Sustainable Development?

Charles - Fabrizio - Ini - Maria - Paul - Thomas - Arthur

SASI Data for a Hot, Crowded, and Unstable World

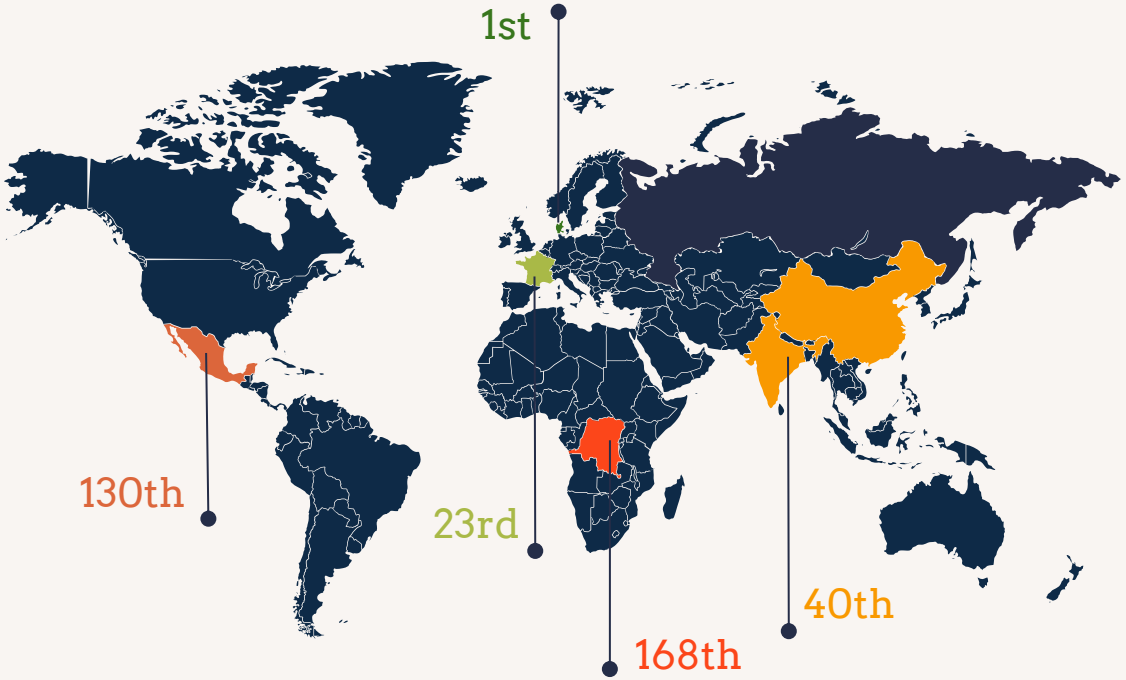
Table of Contents

Context



- Corruption has been around for a very long time
- Divided into 3 different levels
- Empirical evidence suggests that corruption lowers investment and retards economic growth.
- With corruption, public funds are either:
  - Misallocated (partisanship, nepotism)
  - Fewer (tax evasion,)
  - Stolen (embezzlement, bribery)
- Corruption is hard to measure

Corruption Perception Index 2019



Denmark

Score: 87/100

France

Score: 69/100

China/India

Score: 41/100

Mexico

Score: 29/100

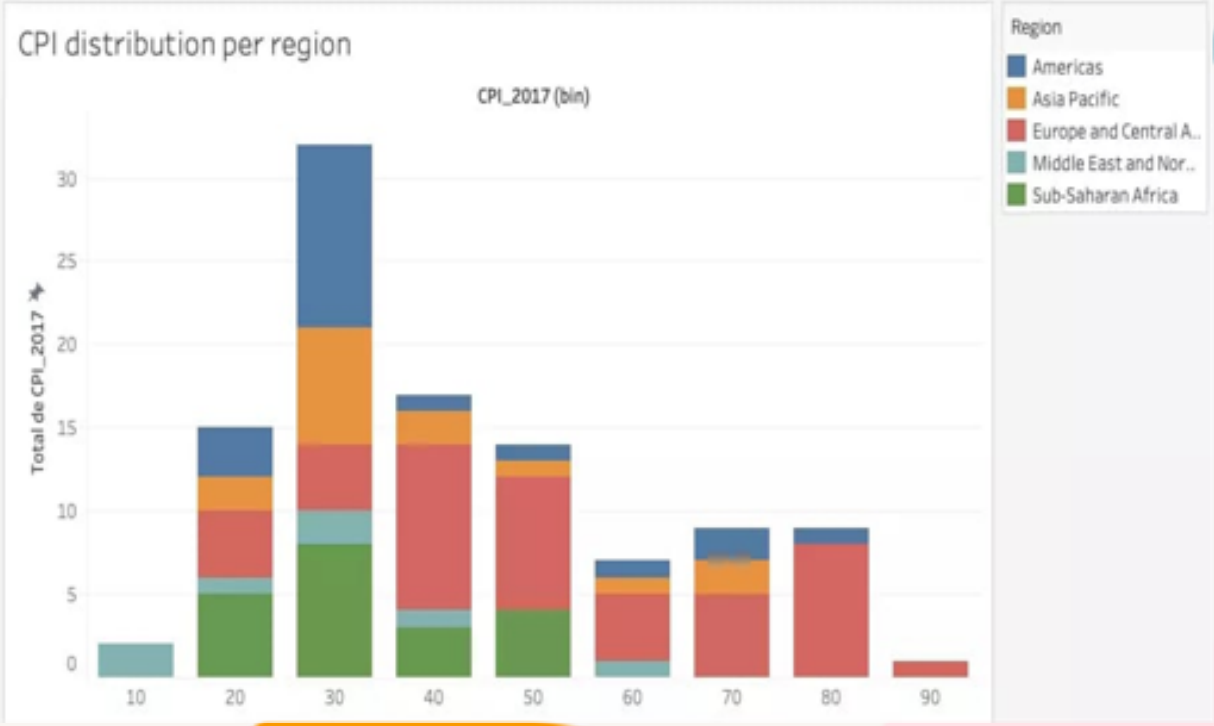
DRC

Score: 18/100

Source: Transparency International 2020

Regional distribution of CPI

- Most countries have a CPI score between 20 and 60
- Sub-Saharan African and Middle East < 60
- Americas and Asia Pacific are low, with few exceptions
- Europe has the **highest** CPI scores
- The only non-European country over 80 score is **Canada**



World Development Indicators

Comparing CPI to:

GDP per capita

Market value of all the final goods and services produced / population (1676)

Human Development Index

People, opportunities, choice (1990)

Sustainable Development Index

Ecological efficiency of human development (2015)

Overview

01 Linear regression

02 Tree Regression

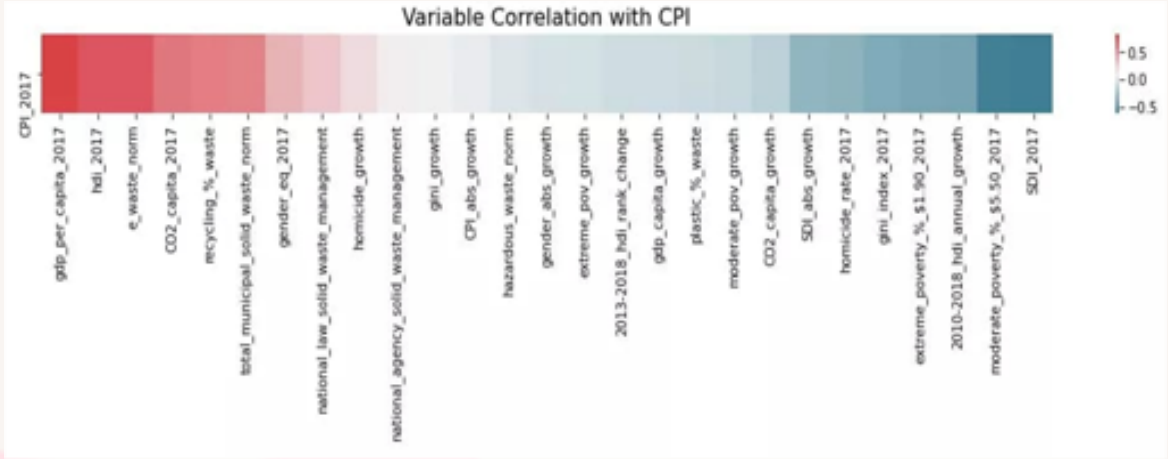
03 Selection of variables to study

04 World Overview

05 Conclusions

Linear Regression

The heatmap below presents **the correlation of the variables with CPI**. In red are the variables positively correlated with CPI and in blue the ones negatively correlated. Only 27 variables were kept, removing those that were too much correlated with each other, i.e. redundant.

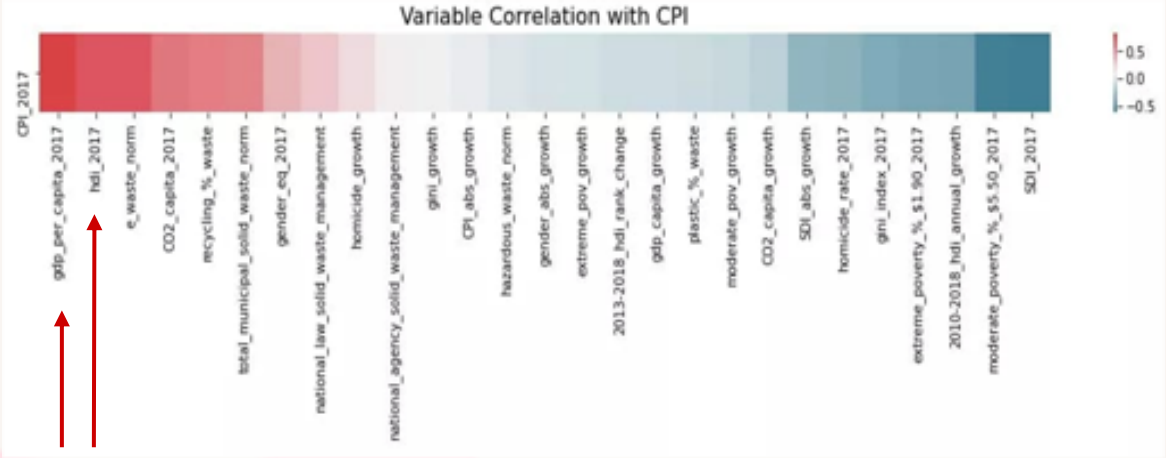


Linear Regression

The heatmap below presents **the correlation of the variables with CPI**. In red are the variables positively correlated with CPI and in blue the ones negatively correlated. Only 27 variables were kept, removing those that were too much correlated with each other, i.e. redundant.

Most positively correlated variables:

- GDP per capita
- Human Development Index



Linear Regression

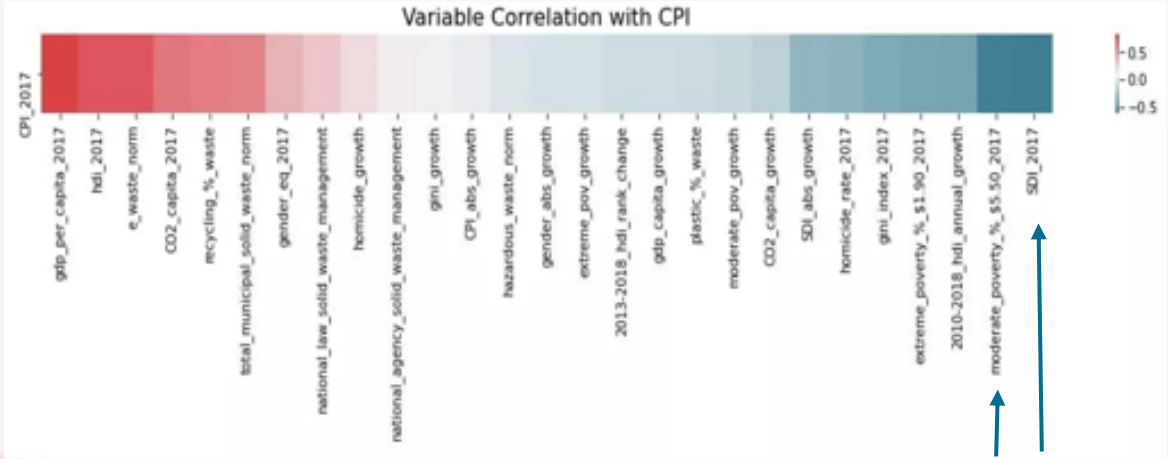
The heatmap below presents **the correlation of the variables with CPI**. In red are the variables positively correlated with CPI and in blue the ones negatively correlated. Only 27 variables were kept, removing those that were too much correlated with each other, i.e. redundant.

Most positively correlated variables:

- GDP per capita
- Human Development Index

Most negatively correlated variables:

- Poverty
- Sustainable Development Index (not relevant!)



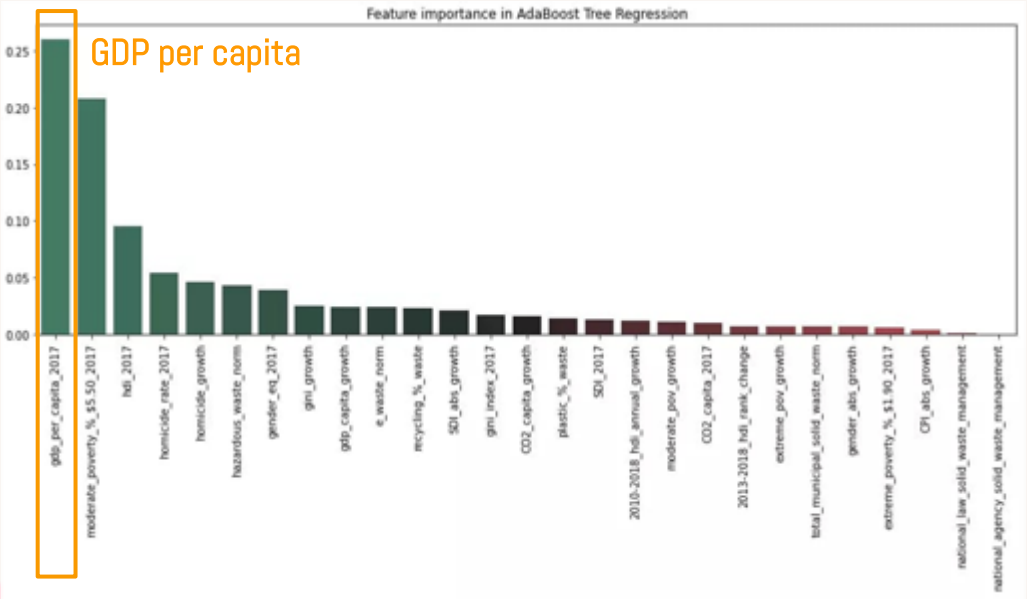
Tree Regression

This model is based on probability trees, which **can capture non-linear relationships** between the different variables and CPI.

This model gave the best results at predicting CPI.

Results:

- GPD per capita is the most relevant variable to predict CPI, explaining more than 25% of CPI variance.
- Meaning that the highest the GDP per capita is, the less corrupted the country is perceived.



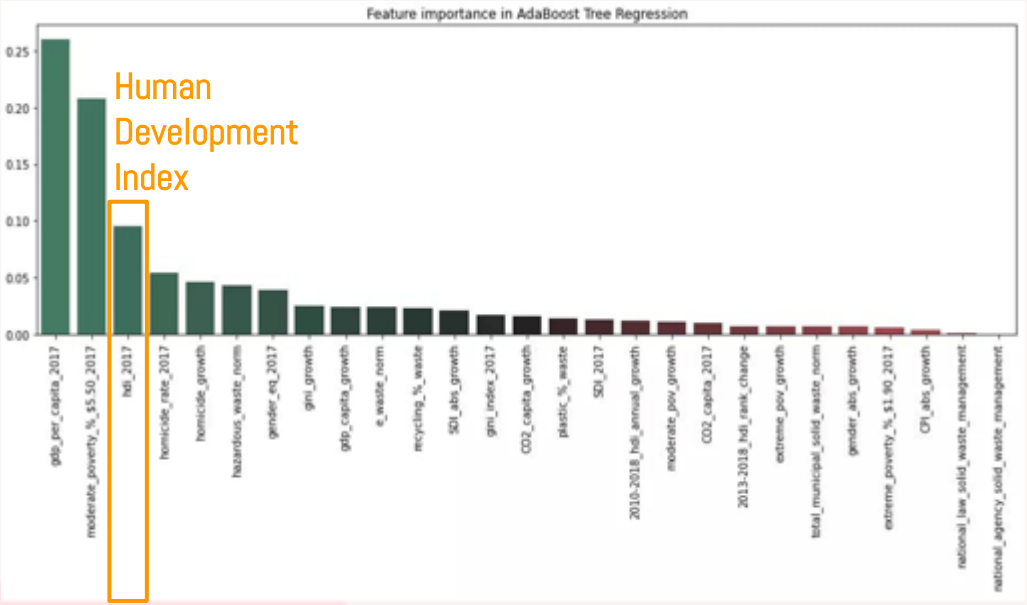
Tree Regression

This model is based on probability trees, which **can capture non-linear relationships** between the different variables and CPI.

This model gave the best results at predicting CPI.

Results:

- HDI is the second most relevant variable to predict CPI, explaining around **10% of CPI variance.**
- Therefore, **the highest the HDI, the less corrupted** the country will be perceived.



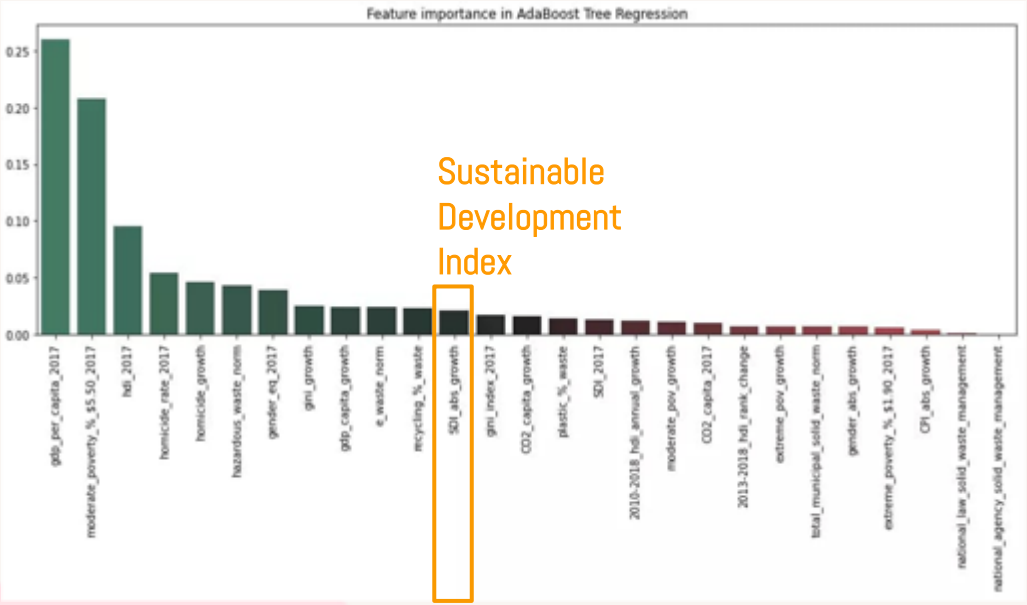
Tree Regression

This model is based on probability trees, which **can capture non-linear relationships** between the different variables and CPI.

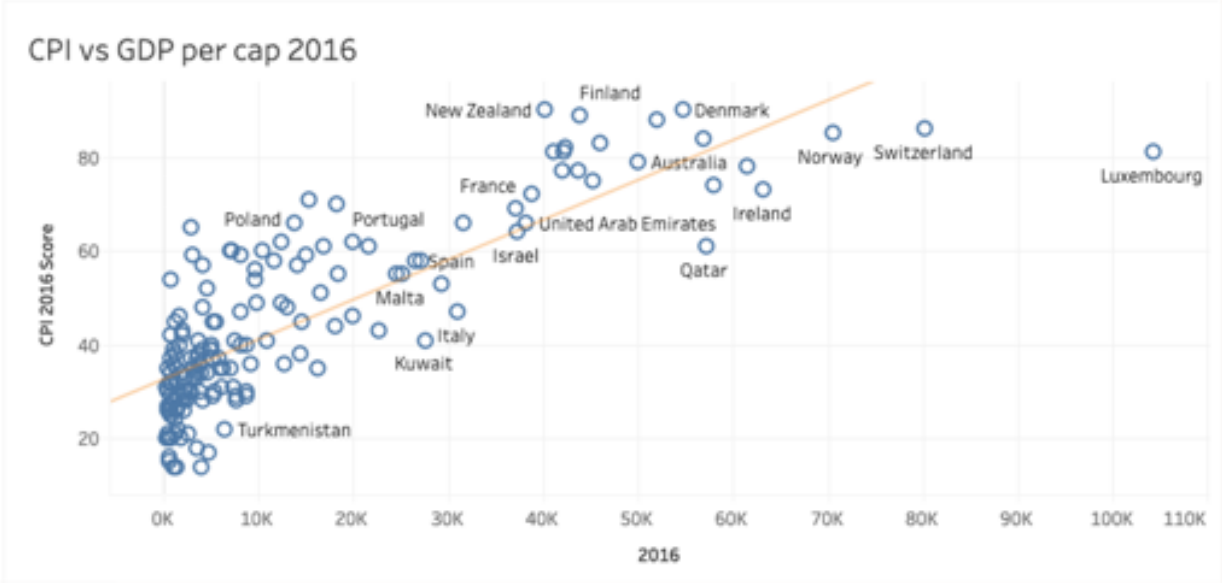
This model gave the best results at predicting CPI.

Results:

- SDI cannot be considered a relevant variable to predict CPI, since it explains less than 5% of CPI variance.
- Therefore, not much can't be said about its impact on CPI



Correlation between CPI and GDP per capita



The overall correlation coefficient between CPI and HDI is given by **r squared = 0.66** which is considered a moderate linear relationship.

Looking more into detail, per region the situation is as follows:

Americas: r sq = 0.52
Asia Pacific: r sq = 0.80
Europe r sq = 0.70
Middle East r sq = 0.61
Sub-Saharan Africa: r sq = 0.28

Linear relationship.

Strong Moderate Weak

Correlation between CPI and Human Development

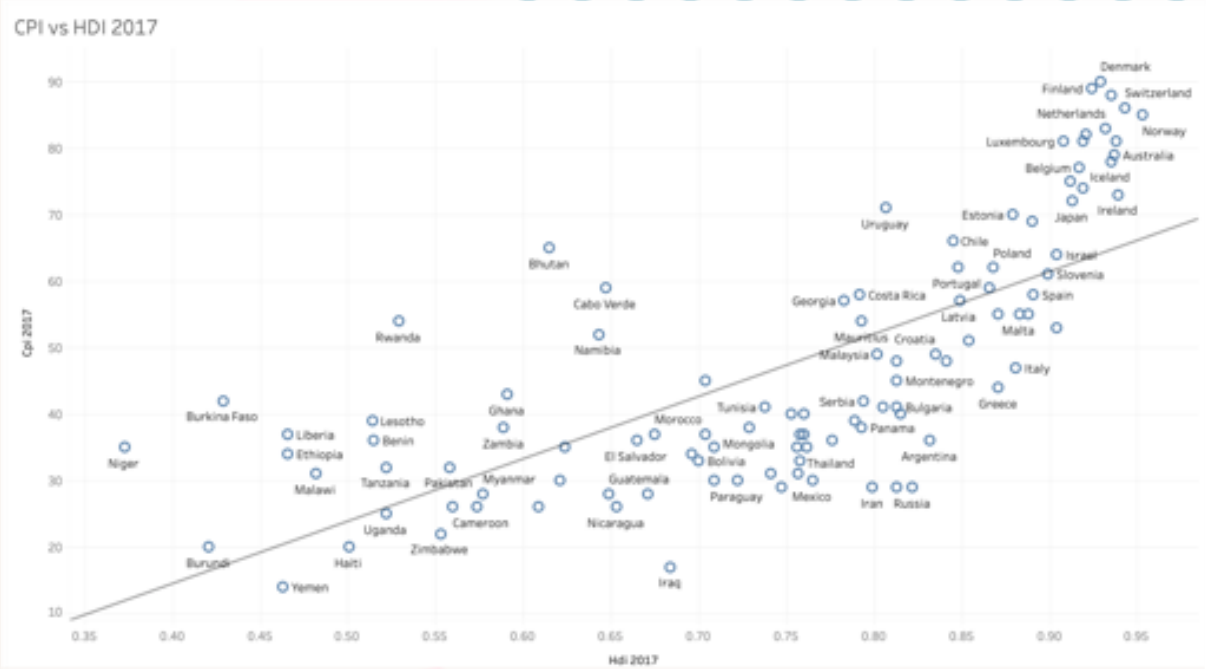
The overall correlation coefficient between CPI and HDI is given by

r squared = 0.51

which is considered a moderate linear relationship.

Looking more into detail, per region the situation is as follows:

- Americas: r sq = 0.63
- Asia Pacific: r sq = 0.48
- Europe r sq = 0.78
- Middle East r sq = 0.64
- Sub-Saharan Africa: r sq = 0.29

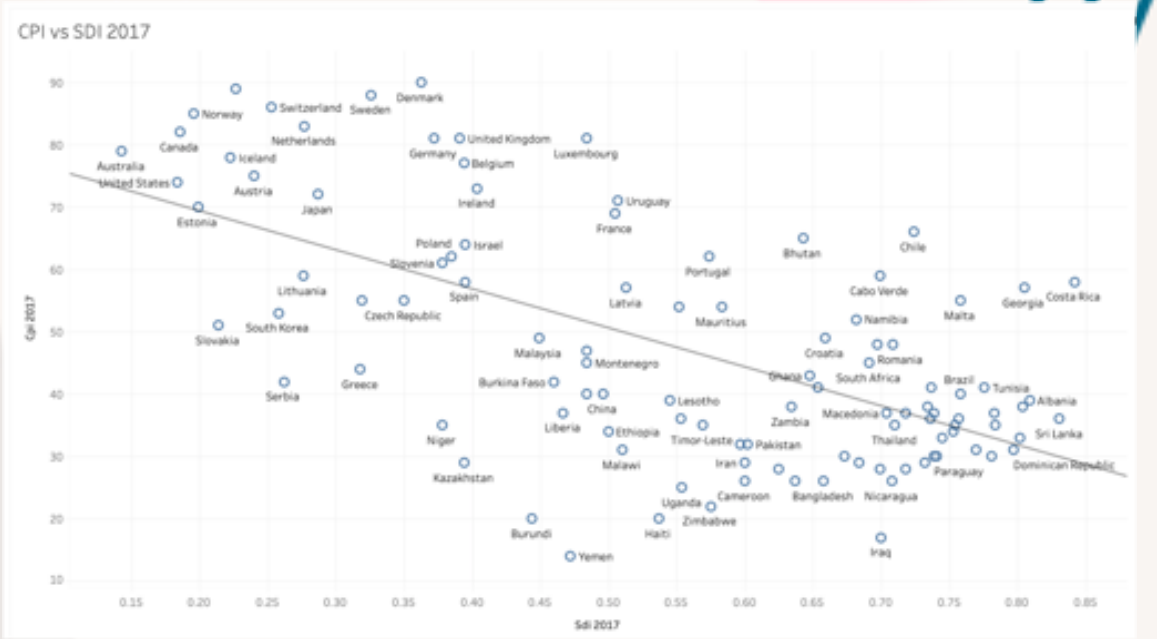


Linear relationship.

Strong Moderate Weak

Correlation between CPI and Sustainable Development

OUT



The overall correlation coefficient between CPI and SDI is given by **r squared = 0.38** which is considered a weak linear relationship.

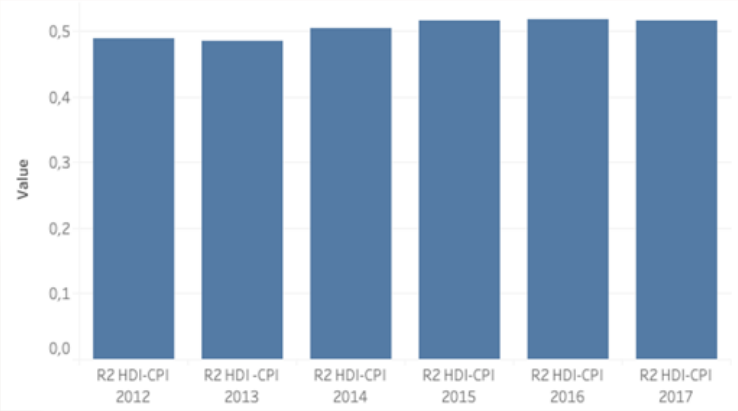
While, per region, it looks like this:

- Americas: r sq = 0.43
- Asia Pacific: r sq = 0.56
- Europe r sq = 0.42
- Middle East r sq = 0.06
- Sub-Saharan Africa: r sq = 0.16

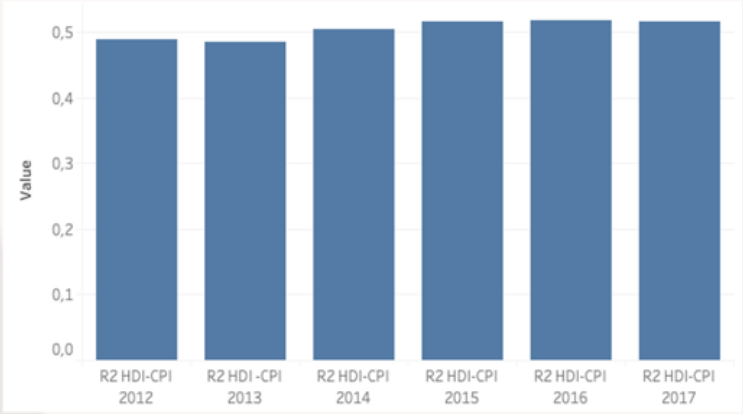
Linear relationship.

Strong Moderate Weak

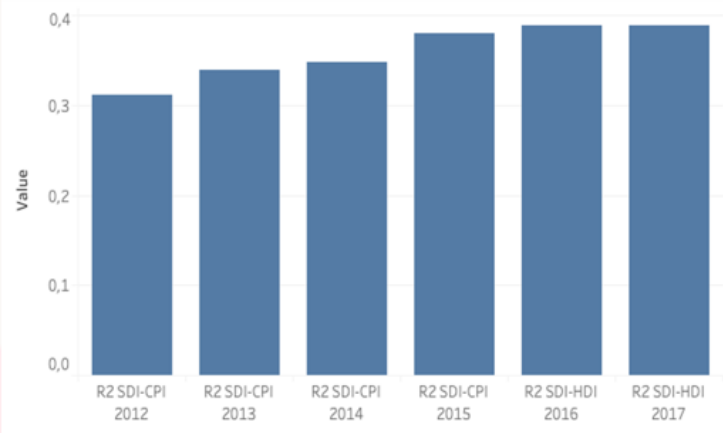
CPI & HDI correlation in terms of R2 (2012-2017)



CPI & HDI correlation in terms of R2 (2012-2017)

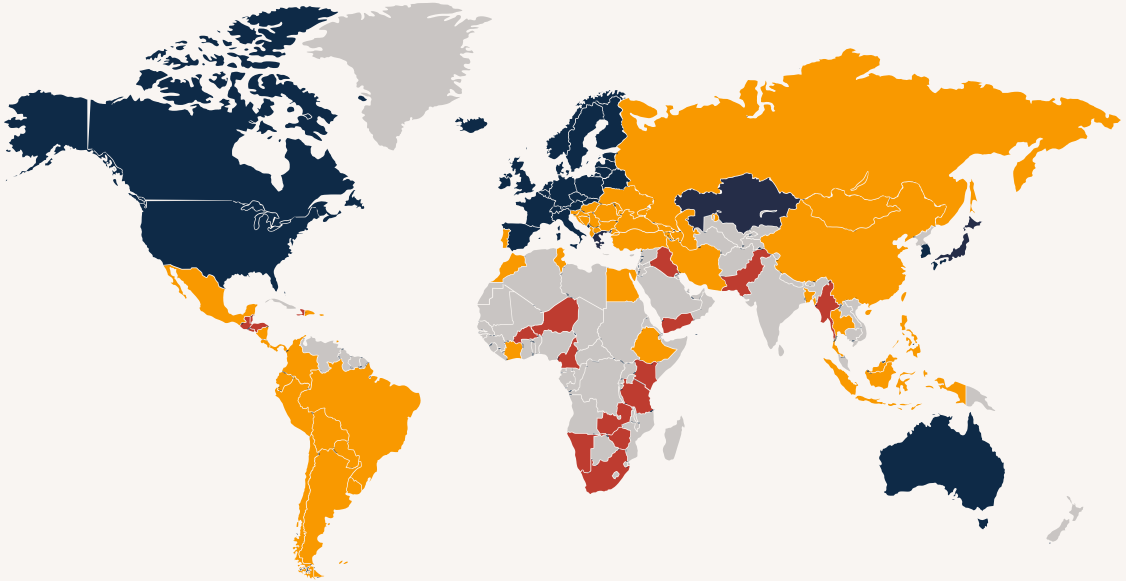


CPI & SDI correlation in terms of R2 (2012-2017)



How does corruption work?

K-means clustering k=3



Cluster 1

Score: 87/100

Cluster 2

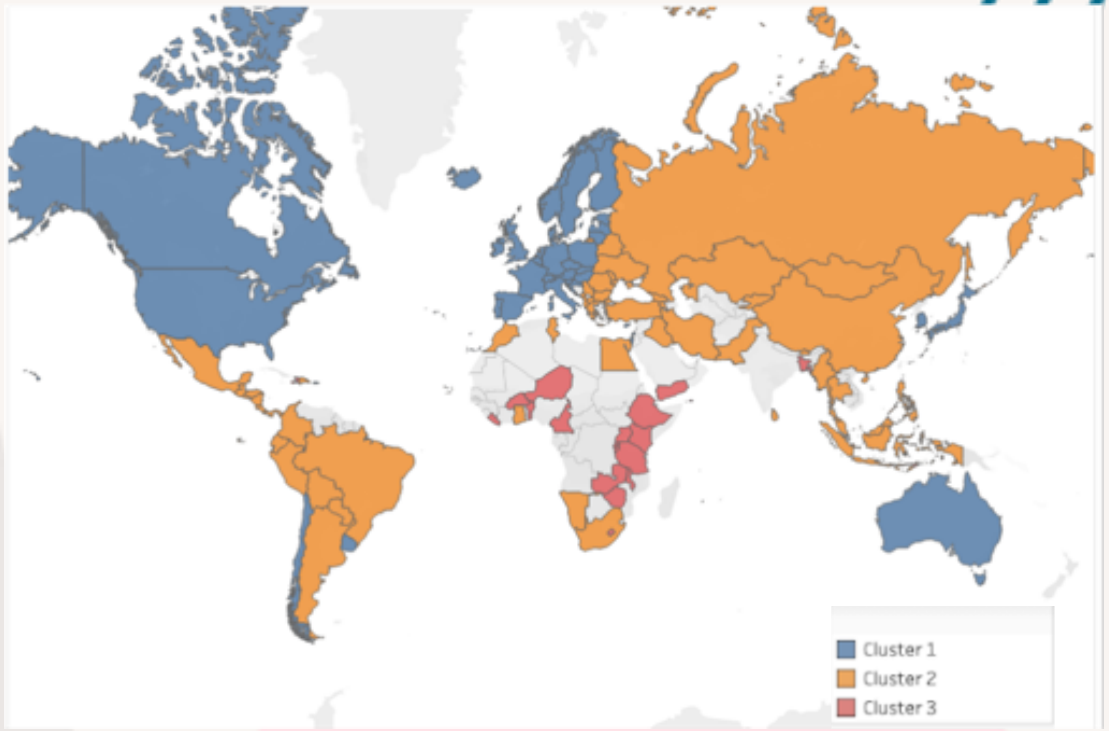
Score: 69/100

Cluster 3

Score: 41/100

No Data

A world overview showing Economic variables cluster



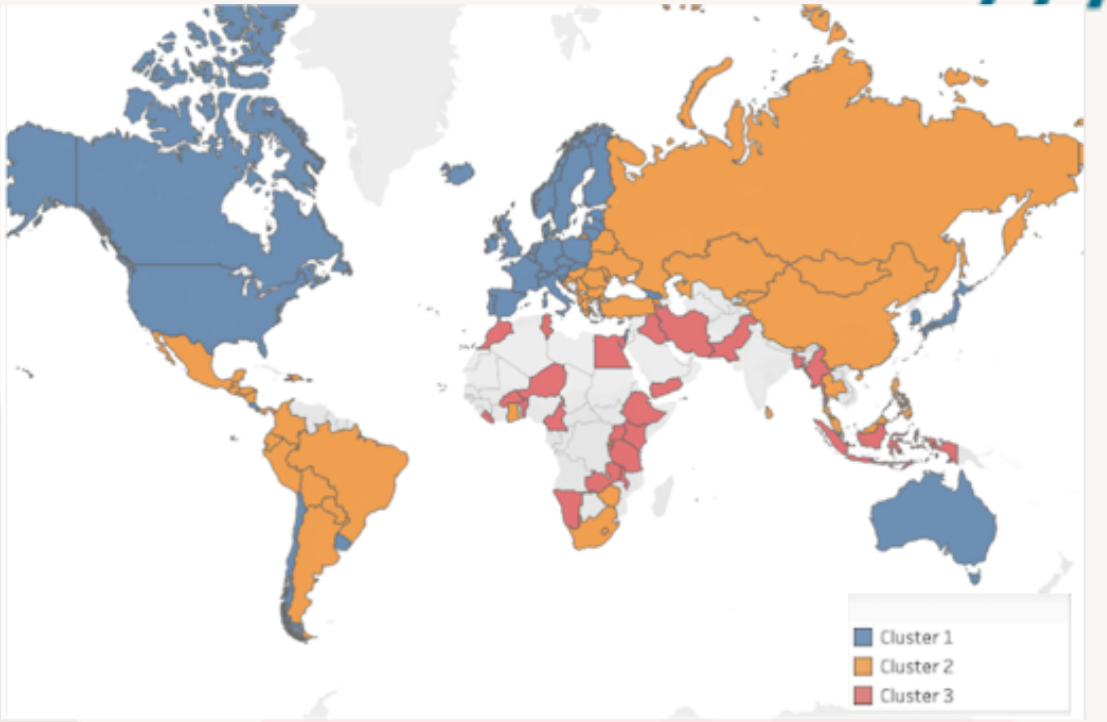
Details of clusters

Cluster 1: mainly "Northern countries" composed by North America, Northern and Western Europe, and Australia and Japan. Additionally we find Chile and Uruguay.

Cluster 2: a huge cluster composed by Eastern Europe, Russia, Latin America, Asia, Maghreb , Middle East plus South Africa and Namibia.

Cluster 3: mainly sub saharan countries (with the exceptions of Namibia and South Africa).

A world overview showing Human Development variables cluster



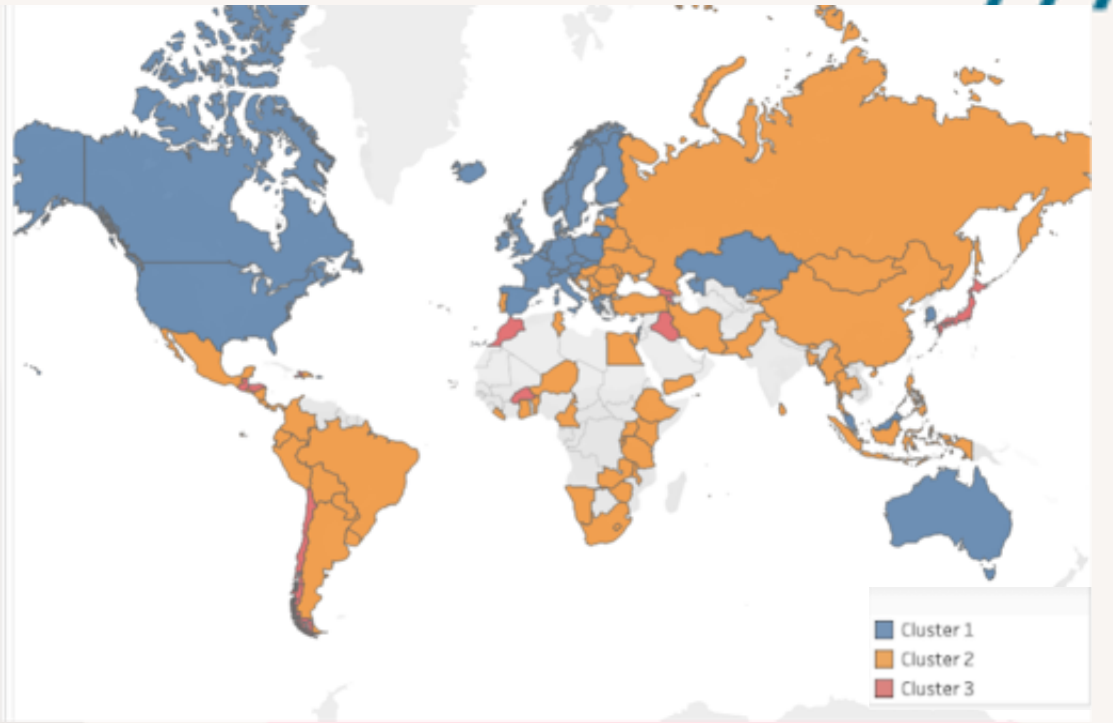
Details of clusters

Cluster 1: mainly "Northern countries" composed by North America, Northern and Western Europe (without Portugal), and Australia and Japan. Additionally we find Kazakhstan.

Cluster 2: a huge cluster composed by Eastern Europe (plus Portugal), Russia, Latin America, Asia (without Kazakhstan), Middle East (without Iraq) and African countries (without Morocco).

Cluster 3: only Morocco, Iraq and Burkina Faso.

A world overview showing Environmental variables cluster



Details of clusters

Cluster 1: mainly "Northern countries" composed by North America, Northern and Western Europe, and Australia and Japan. Additionally we find Chile and Uruguay.

Cluster 2: a huge cluster composed by Eastern Europe, Russia, Latin America, Eastern and Central Asia, plus South Africa and Zimbabwe.

Cluster 3: mainly African countries (with the exceptions of Zimbabwe and South Africa), Middle East, and Southern Asia.

Conclusions

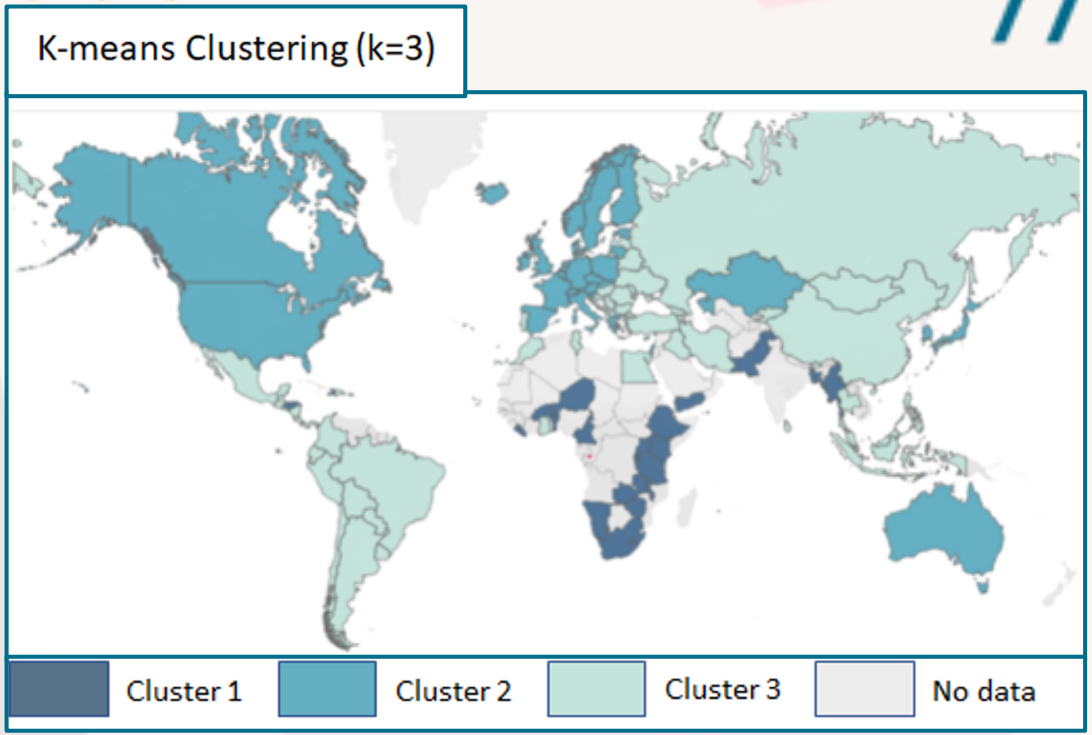
- The Corruption Perception Index is correlated to some of the selected world development indicators.
- Traditional measures (GDP per capita & Poverty) are still more correlated to CPI than Sustainability measures (HDI & SDI).
- Social variables (= HDI) still have a higher correlation than environmental variables.
- Environmental variables are growing in importance.

THANK YOU!

Charles - Fabrizio - Ini - Maria Jesus - Paul

SASI Data for a Hot, Crowded, and Unstable World

A rough world overview

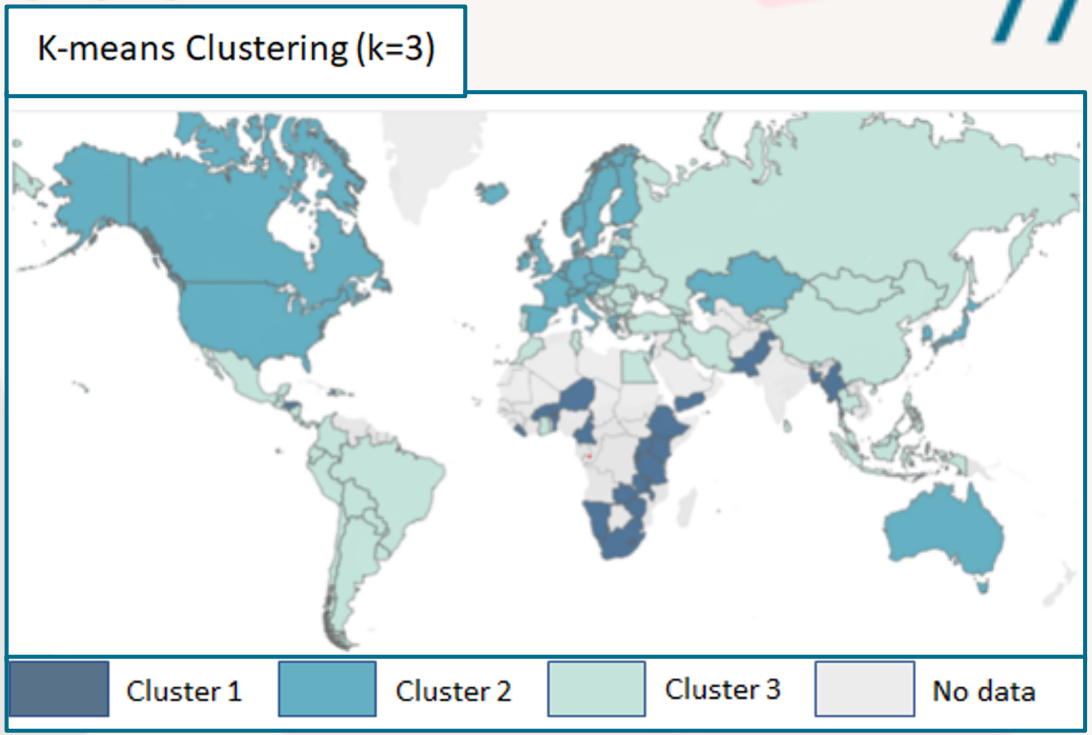


Countries that behave similarly

This map was obtained thanks to an algorithms that has treated the data gathered from the 27 selected variables used for our study.

The algorithm has created three main clusters of countries that encounter some similarities in their behaviours.

A rough world overview



Details of clusters

Cluster 1: mainly composed by African countries plus Pakistan and Myanmar.

Cluster 2: huge clusters composed by countries from different areas, and with very distinct CPI rankings.

Cluster 3: is a group of countries with high GDP per capita, high HDI, and low level of corruption.

A rough world overview

Countries that behave similarly

This map was obtained thanks to an algorithms that has treated the data gathered from the 27 selected variables used for our study.

The algorithm has created three main clusters of countries that encounter some similarities in their behaviours.

Correlation between CPI and Sustainable Development

Not correlated