UN Big Data Hackathon 2022

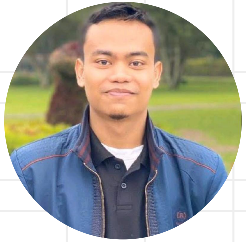
Youth Track

By HRS Tech

The Team

Our Members





Ahmad **Afif** Mauludi
aafifmauludi@gmail.com

Razis Aji Saputro
razisaji1995@gmail.com

Ulan Surjayanto
ultisa222@gmail.com

Bayu Ade Krisna Setiawan
mrbayuade@gmail.com

Outline

01 The Team

02 Introduction & Research Question

03 Dataset

04 Methodology

05 Exploratory Data Analysis

06 Model

07 Dashboard / Visualization

08 Key Finding (s) & Recommendation (s)

09 Conclusions

10 Next Improvement

Introduction

<u>Theme:</u> The rise of food and energy prices affecting vicious cycles of poverty, hunger, and inequalities

"**Rising food** and **energy prices** are putting a strain on families' budgets, making it more difficult to afford necessities such as food, gas, and rent."

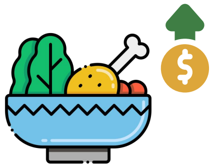
Introduction

Globally, domestic food price inflation remains high. Data from June to September 2022 show that almost all low-income and middle-income countries experienced high inflation.

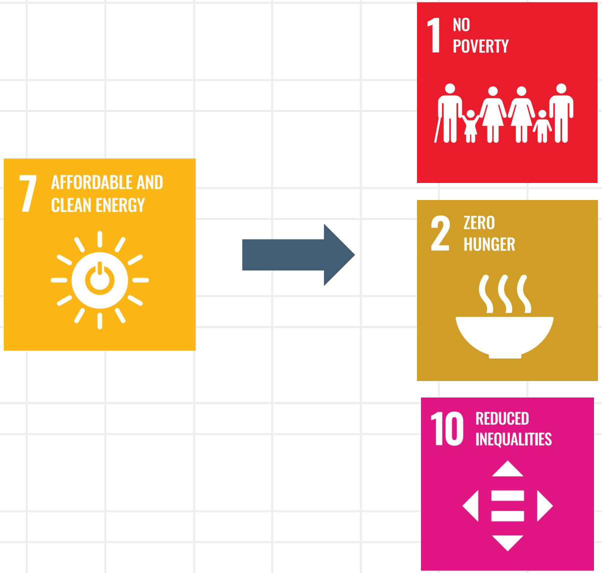
84.2% of low-income countries, 88.9% of lower-middle-income countries, and 93% of upper-middle-income countries experienced inflation levels above 5%, with many experiencing double-digit inflation.

The proportion of high-income countries experiencing high food price inflation has risen to 87.5%.

Source: https://www.worldbank.org/en/topic/agriculture/brief/food-security-update



What we are trying to solve

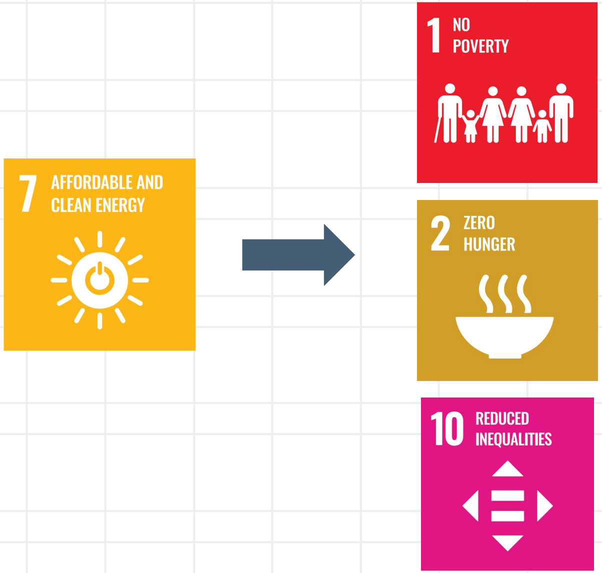


Creating a tool based on data and visualizing the correlation between poverty, hunger, inequalities, and energy consumption & prices to support the decision-makers.

UN Big Data Hackathon

HRS TECH | SUSTAINABLE DEVELOPMENT GOALS

Research Question



#1 How is the correlation between **energy prices** & **energy consumption***?

#2 How is the impact between renewable & non-renewable **energy prices** on **poverty, hunger and inequality**?

#3 How is the impact between renewable & non-renewable **energy consumption*** on **poverty, hunger and inequality**

*consumption per capita

Dataset

Poverty

- Poverty Headcount Ratio National Poverty Lines
- Multidimensional poverty index
- Multidimensional poverty headcount ratio, household (% of total households)

etc

Source: World Bank

Hunger (Food Insecurity)

- Prevalence of severe food insecurity in the population (%)
- Prevalence of moderate or severe food insecurity in the population (%)
- Prevalence of undernourishment (% of population)

Source: World Bank

Inequalities

- Gini Index

Source: World Bank

Energy

- Energy Consumption
- Energy Price

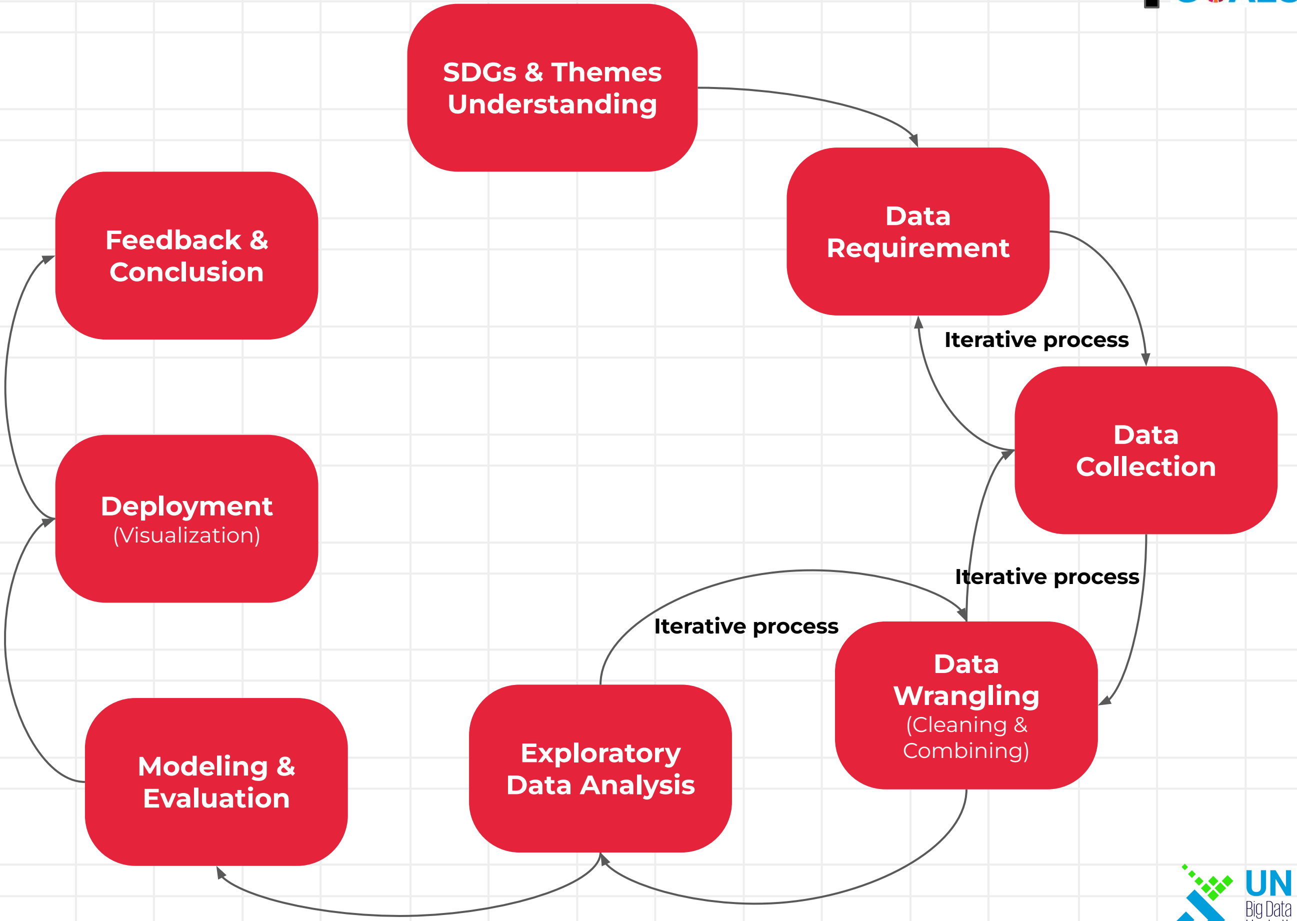
etc

Source: Our World in Data & World Bank

Data Documentation   **Final Combined Data Set**

Methodology

-



Exploratory Data Analysis (EDA)

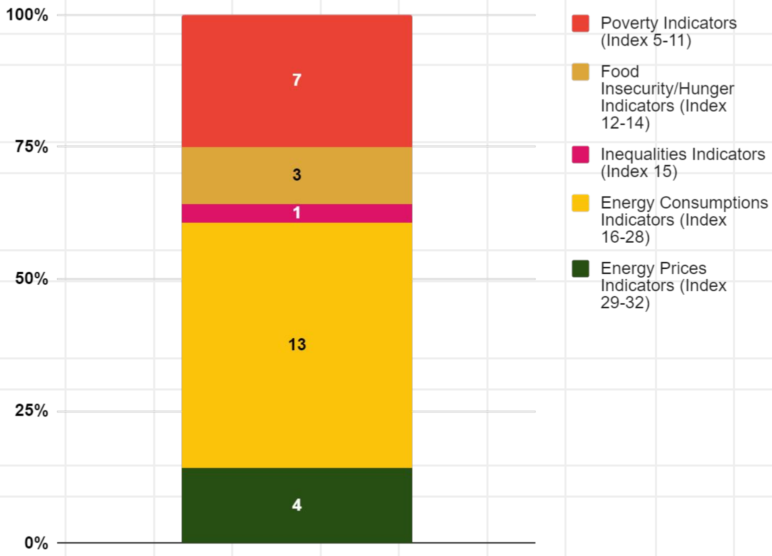
Data Info

33 Columns & 16476 Rows

2 Categorical, 30 Numerical, and 1 Date Feature (s)

our_df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 16476 entries, 0 to 16475
Data columns (total 33 columns):
 #   Column                              Non-Null Count  Dtype
---  ------                              --------------  -----
 0   country                             16476 non-null  object
 1   region                              16251 non-null  object
 2   year                                16476 non-null  datetime64[ns]
 3   population                          16371 non-null  float64
 4   gdp                                 11004 non-null  float64
 5   SI.POV.NAHC                         960 non-null    float64
 6   SI.POV.GAPS                         1668 non-null   float64
 7   SI.POV.LMIC.GP                      1668 non-null   float64
 8   SI.POV.UMIC                         1667 non-null   float64
 9   SI.POV.MDIM                         431 non-null    float64
 10  SI.POV.MDIM.MA                      357 non-null    float64
 11  SI.POV.MDIM.FE                      357 non-null    float64
 12  SN.ITK.SVFI.ZS                      665 non-null    float64
 13  SN.ITK.MSFI.ZS                      671 non-null    float64
 14  SN.ITK.DEFC.ZS                      2816 non-null   float64
 15  SI.POV.GINI                         1663 non-null   float64
 16  energy_per_capita                   9521 non-null   float64
 17  biofuel_cons_per_capita             838 non-null    float64
 18  coal_cons_per_capita                4196 non-null   float64
 19  fossil_energy_per_capita            4193 non-null   float64
 20  gas_energy_per_capita               4219 non-null   float64
 21  hydro_energy_per_capita             4215 non-null   float64
 22  low_carbon_energy_per_capita        4215 non-null   float64
 23  nuclear_energy_per_capita           4058 non-null   float64
 24  oil_energy_per_capita               4222 non-null   float64
 25  other_renewables_energy_per_capita  4172 non-null   float64
 26  renewables_energy_per_capita        4215 non-null   float64
 27  solar_energy_per_capita             4172 non-null   float64
 28  wind_energy_per_capita              4172 non-null   float64
 29  EP.PMP.DESL.CD                      1582 non-null   float64
 30  NY.COAL.RT.ZS                       6910 non-null   float64
 31  NY.PETR.RT.ZS                       7054 non-null   float64
 32  NY.TOTL.RT.ZS                       7145 non-null   float64
dtypes: datetime64[ns](1), float64(30), object(2)
memory usage: 4.1+ MB
```



The Indicators

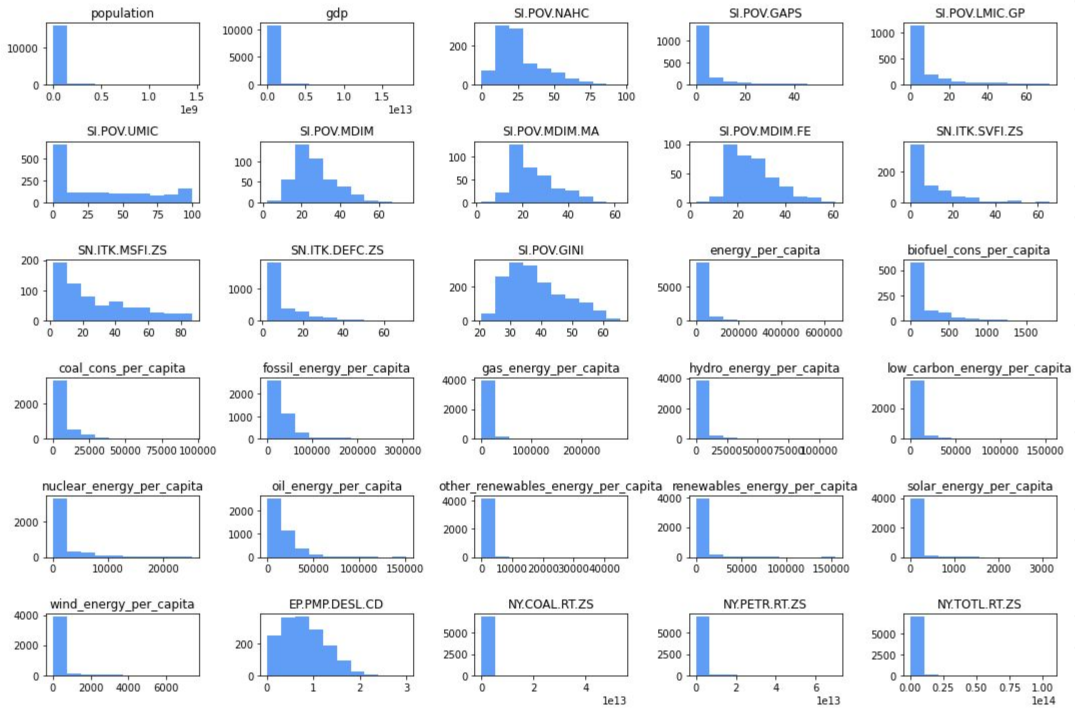
Data Documentation

UN Big Data Hackathon | EDA

HRS TECH | SUSTAINABLE DEVELOPMENT GOALS

Data Distribution

numerical



categorical

	count	unique	top	freq
country	16476	221	Zimbabwe	122
region	16251	6	Asia	4401

```
value counts dari column country
Zimbabwe       122
Morocco        122
Ecuador        122
Myanmar        122
Netherlands    122
              ...
Eritrea         27
Palestine       24
Timor           18
Montenegro      17
South Sudan     10
Name: country, Length: 221, dtype: int64
-----------------------------

value counts dari column region
Asia        4401
Africa      4139
Americas    3378
Europe      3355
Oceania      938
0             40
Name: region, dtype: int64
-----------------------------
```

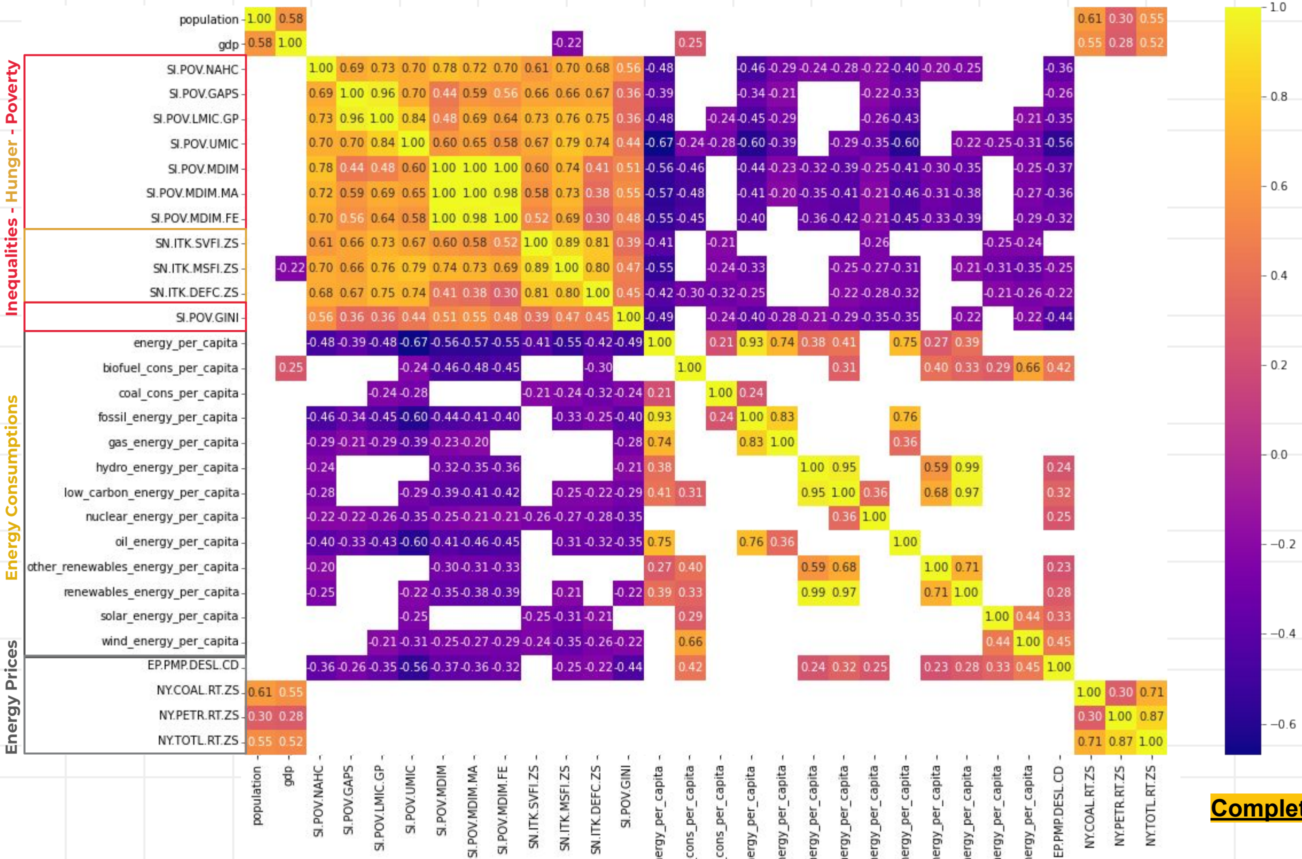
date

```
year range is between 1900-01-01 00:00:00 and 2021-01-01 00:00:00
```

Completed Code .ipynb   Completed Code .html

EDA

Heat Map Correlation



Poverty, Hunger, and **Inequality** have **Negatives Correlation** with **Energy Consumption** (both renewable & non-renewable)

The vicious cycle (Poverty, Hunger, and Inequality) **positively correlates.**

The Pump price for diesel fuel (EP.PMP.DESL.CD) is **negatively correlated with The Vicious Cycles** and **positively correlated** with some **renewable energy consumption**

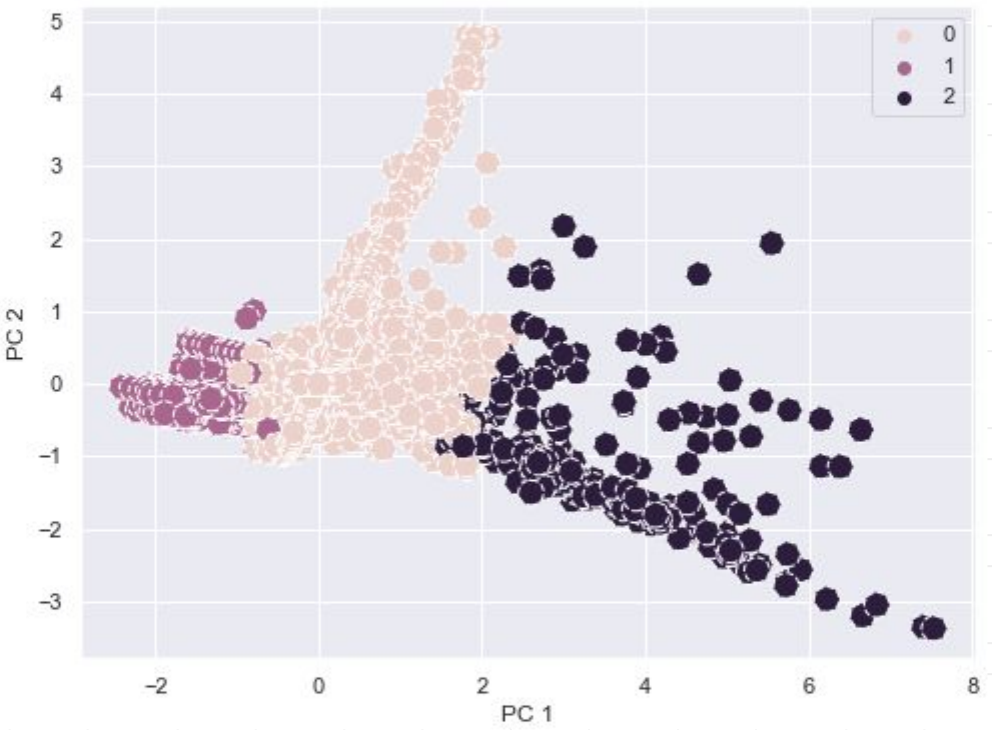
Completed Code .ipynb **Completed Code .html**

Modeling

We use clustering to creating a cluster based on poverty, hunger and inequalities index into three clusters.

- cluster 0 has medium poverty, medium food insecurity and high gini index (**fair wellbeing cluster**)
- cluster 1 has low poverty, low food insecurity and low gini index (**good wellbeing cluster**)
- **cluster 2** has high poverty, high food insecurity and medium gini index (**poor wellbeing cluster**)

From this clustering method we found that **for the latest year records**, all countries (n=221) are segmented in cluster 0 (**fair wellbeing**)



*) Wellbeing based on poverty, hunger and inequality

Notes on Clustering

Our objectives on this clustering are to provide the world countries segmentation based on poverty, hunger, inequalities, and also rising energy price indicators and deploy it on the dashboard.

We standardized, imputed for missing data, and selected the features for clustering. In this clustering, we selected Poverty Headcount Ratio National Poverty Lines, the poverty gap at $2.15 a day, the poverty gap at $3.65 a day, the Poverty headcount ratio at $6.85 a day, all three food insecurity indicators, Gini index, and Diesel Pump Price.

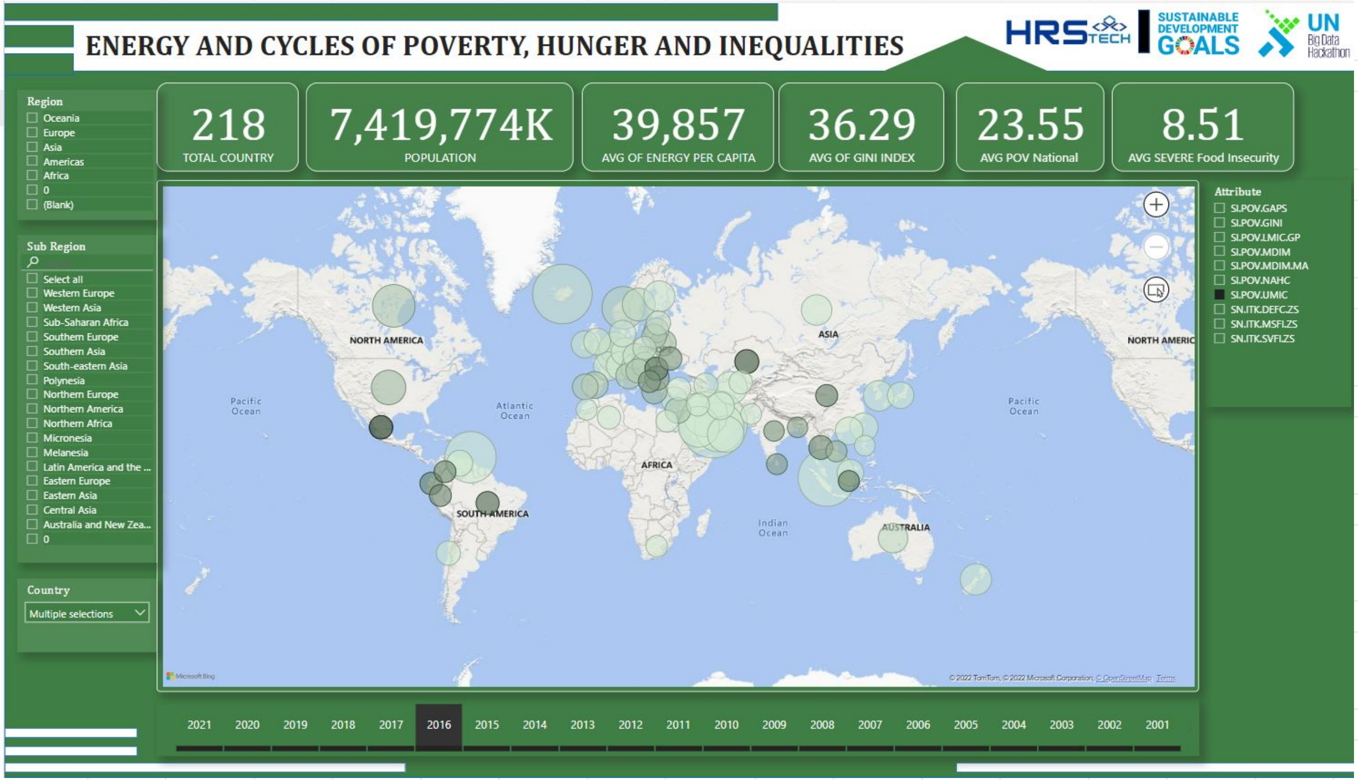
Before the cluster also we determined the number of clusters considering the result of the elbow method and silhouette score and we decided to make three clusters as segmentation.

We trained data for clustering and use Principal Component Analysis for dimensionality reduction and help us visualize the clusters.

Unfortunately, because uncompleted data in big numbers, we cannot deploy the clusters yet on our dashboard.

Dashboard / Visualization

Dashboard / Visualization



Dashboard Link

Key Finding (s) & Recommendation (s)

#1 When fossil energy prices are volatile, investors tend to put money into the renewable energy industry. An additional set of actions to save energy, diversify supplies, and replace fossil fuels by hastening the deployment of renewable energy. Increasing energy efficiency and savings, as well as expanding renewables, are expected to lower energy prices while hastening the green transition.

#2 The rise of diesel fuel prices surprisingly has a negative correlation with “the vicious cycle”. It means the higher the price, the poverty, hunger, and inequality indicators would be lower!

#3 Consumption of renewable and nonrenewable energies per capita correlates negatively with poverty, hunger, and inequality. Energy bills consume a significant portion of consumers' income, limiting their ability to cover other expenses, and resulting in energy poverty.

The Conclusions

–

#1 The rise of the diesel fuel price is positively correlated with renewable energy consumption.

#2 The rise of diesel fuel prices has a negative correlation with “the vicious cycle” .

#3 The higher the vicious cycle, lower the energy consumption.

Next Improvement

1. Check whether the diversity of energy source and utility are impacting the vicious cycle (poverty, hunger, and inequalities)
2. Completing the *null* data on dataset for future analysis.
3. Collect and add **Renewable Energy Potential Data** for more completed analysis and decision support tools on how to utilize and prioritize the renewable energy accelerating manufacturer based on poverty, hunger, and inequalities indicators.
4. **Deployment** of “the model” on visualization and predicting the missing and future indicators.
5. Creating tools based on the model to customized and predict certain indicators.
6. Add more feature on dashboard such as advanced filter for energy consumptions based on the source of energy (renewable/non-renewable, type of energy, etc)

THANK YOU

Appendix 1 : Color Palette Inspiration

SDGs #1

SDGs #2

SDGs #7

SDGs #10

SDGs #11

SDGs #13

2022 Theme

Using Big Data and Data Science to develop ideas and solutions to address Global Challenges and help achieve Sustainable Development Goals; notably to support policies caused by:

- The disruption to Global Value Chains and Economic Globalization due to disasters, conflicts, restrictions, blockages
- The impact of Climate Change on society as part of monitoring SDG 13
- The rise of food and energy prices affecting vicious cycles of poverty, hunger, and inequalities

#E5243B #DDA63A #3F7E44 #ffffffff #000000ff #808284 #DD1367

SDGs Brand Guidelines

Appendix 2 : The Assets

Logo Usage

[SDGs Brand Guidelines]

Font Usage

Montserrat - Normal
Montserrat - Bold
Montserrat - Black

Icons

source source source source source

UN Big Data Hackathon

HRS TECH | SUSTAINABLE DEVELOPMENT GOALS

Appendix 3 : SDGs Logo

1 NO POVERTY

2 ZERO HUNGER

7 AFFORDABLE AND CLEAN ENERGY

10 REDUCED INEQUALITIES

11 SUSTAINABLE CITIES AND COMMUNITIES

13 CLIMATE ACTION

1 NO POVERTY

2 ZERO HUNGER

7 AFFORDABLE AND CLEAN ENERGY

10 REDUCED INEQUALITIES

11 SUSTAINABLE CITIES AND COMMUNITIES

13 CLIMATE ACTION