Global Poverty Project :

Poverty & Health

By Avramidis Theofilos and Korolis Alexandros

Introduction

The purpose of this project is to categorize countries all around the world based on social-economic and health features in order to provide people from poor countries humanitarian aid.

Table of contents

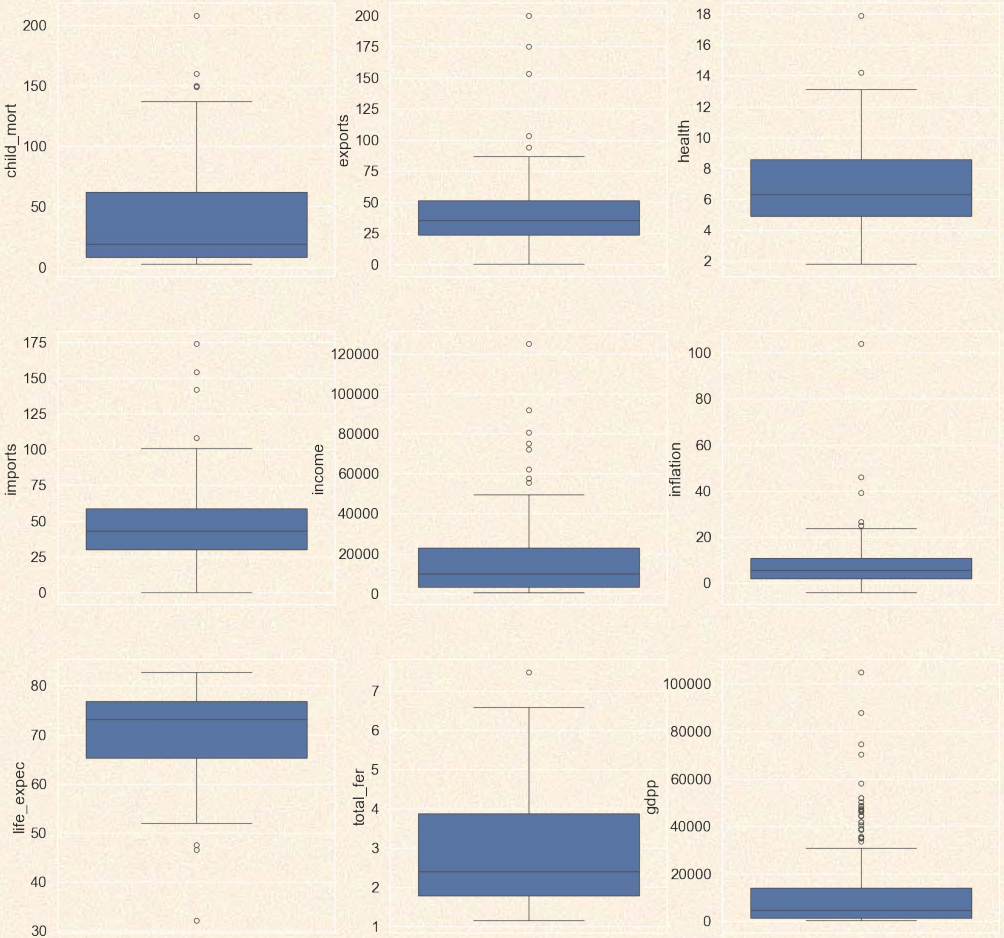
01
EDA

Exploratory Data Analysis ,
introduction to the dataset

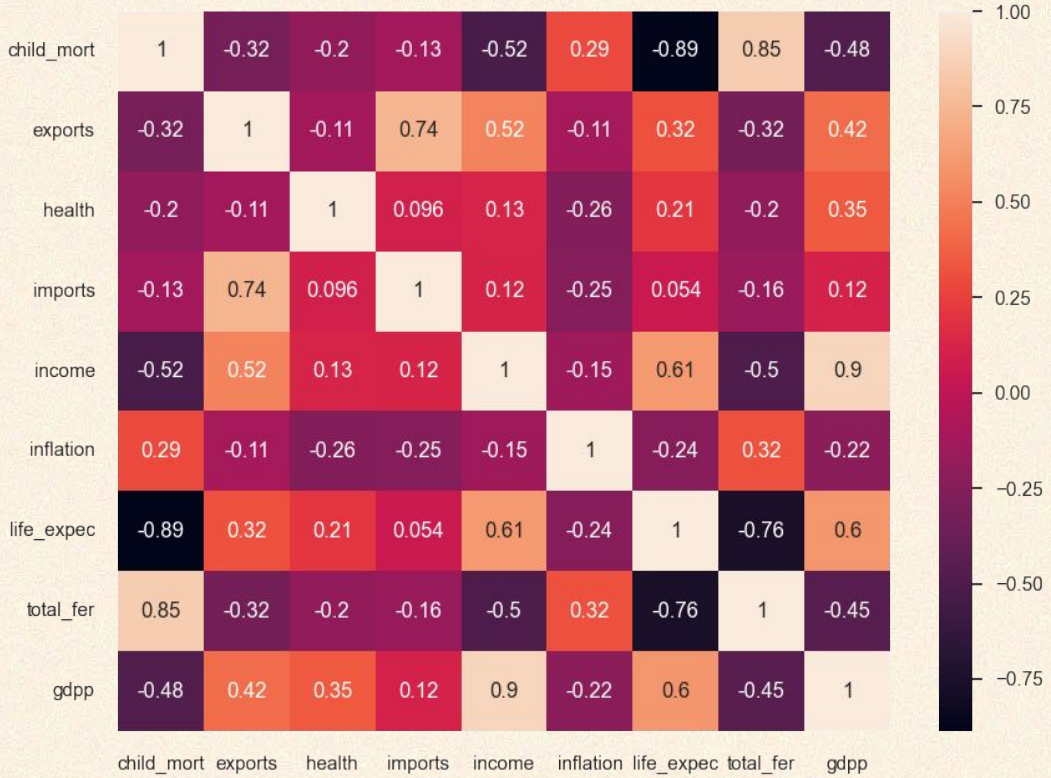
EDA

- **Observations** from **167 countries** around the world .
- Based on **10 features (**related to socio-economic and health factors**) .**
- **No missing** or **duplicate values** .
- Existence of **extreme values** in the dataset .

EDA

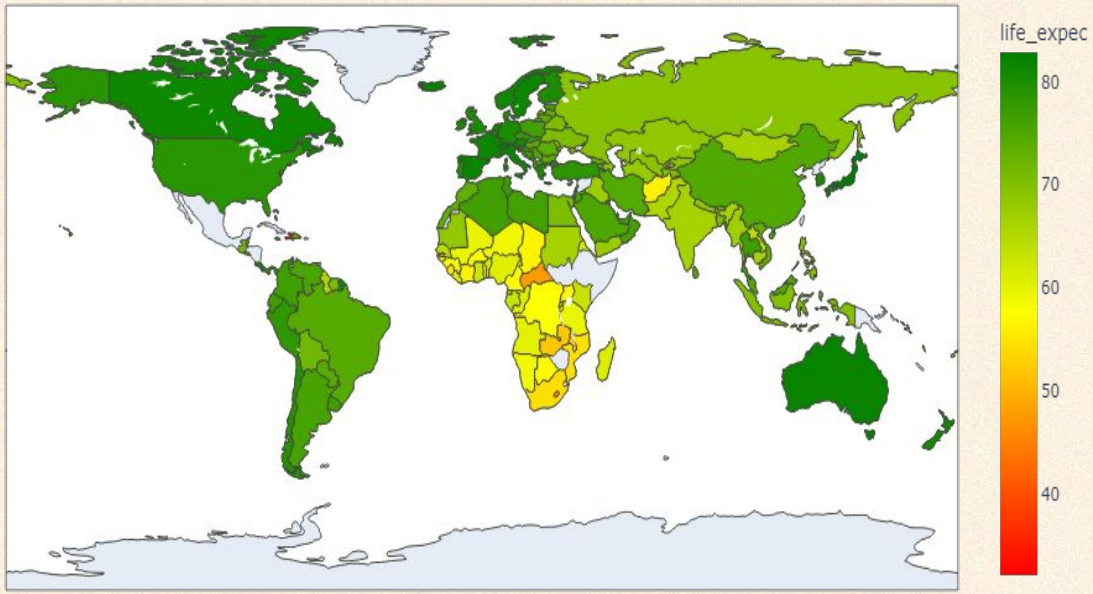

child_mort
exports
health
imports
income
inflation
life_expec
total_fer
gdpp

EDA



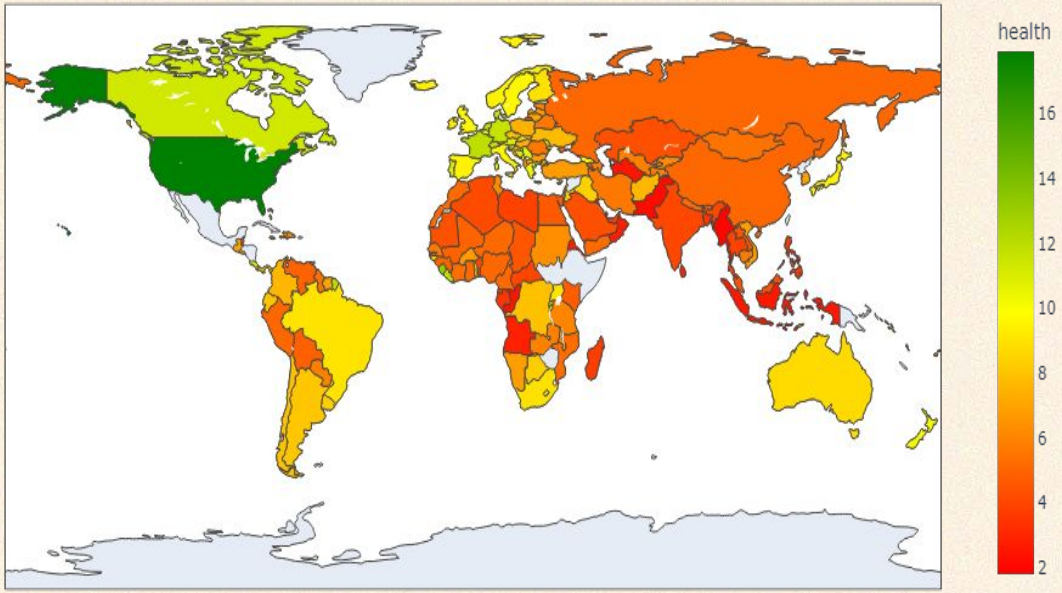
- Life expectancy ~ gdpp
- Income ~ child mortality
- Income ~ life expectancy
- Income ~ health

EDA



- **Lesotho** has the l**owest average life expectancy** with 46.5 years
- **Japan** has the **highest average life expectancy** with 82 years

EDA



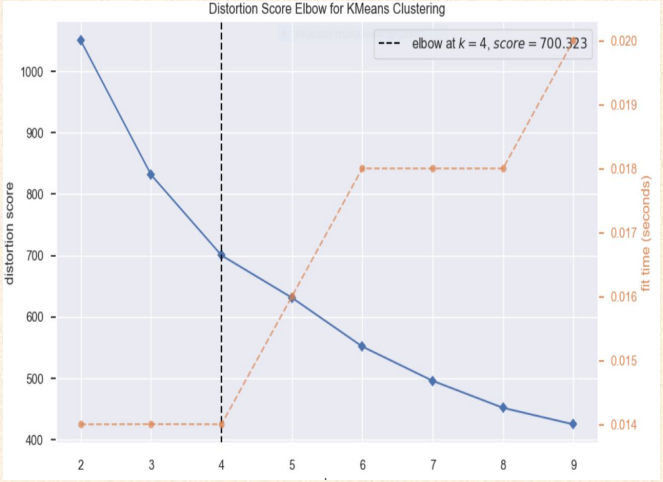
- **Qatar** has the **lowest total health spending** per capita. (1.81%)
- **United States** has the **highest total health** spending per capita. (17.9%)

02
Modeling

Problem modeling and categorization

Modeling

- **Scale** the data, in order to compare and decide the category.
- Applied **k-means** (optimum distance between each data point and a centroid).
- Evaluation based on **elbow visualization** and **silhouette**.



Modeling

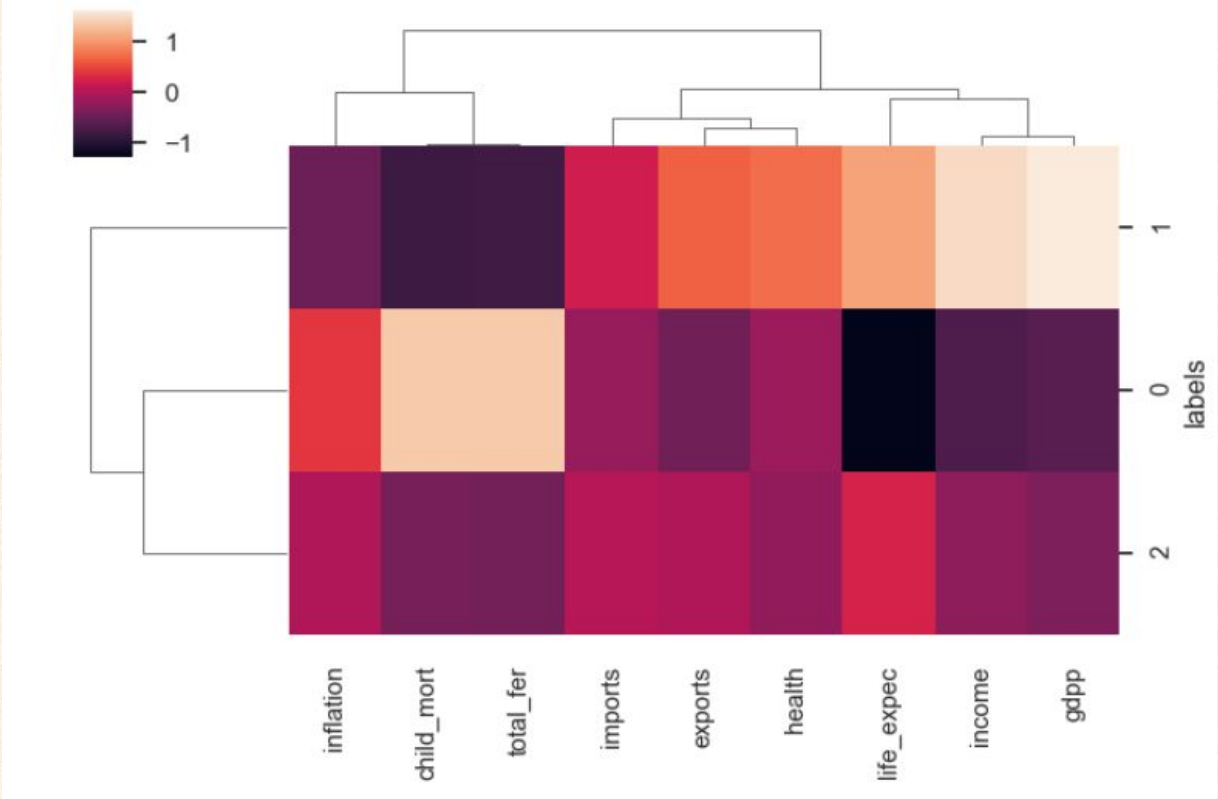
- Distortion : the distance between the data points in the same category. (How dense a category is.)
- How far the categories are. (The further the better)

03

Labelling

Analyze and labeling different categories

Labelling

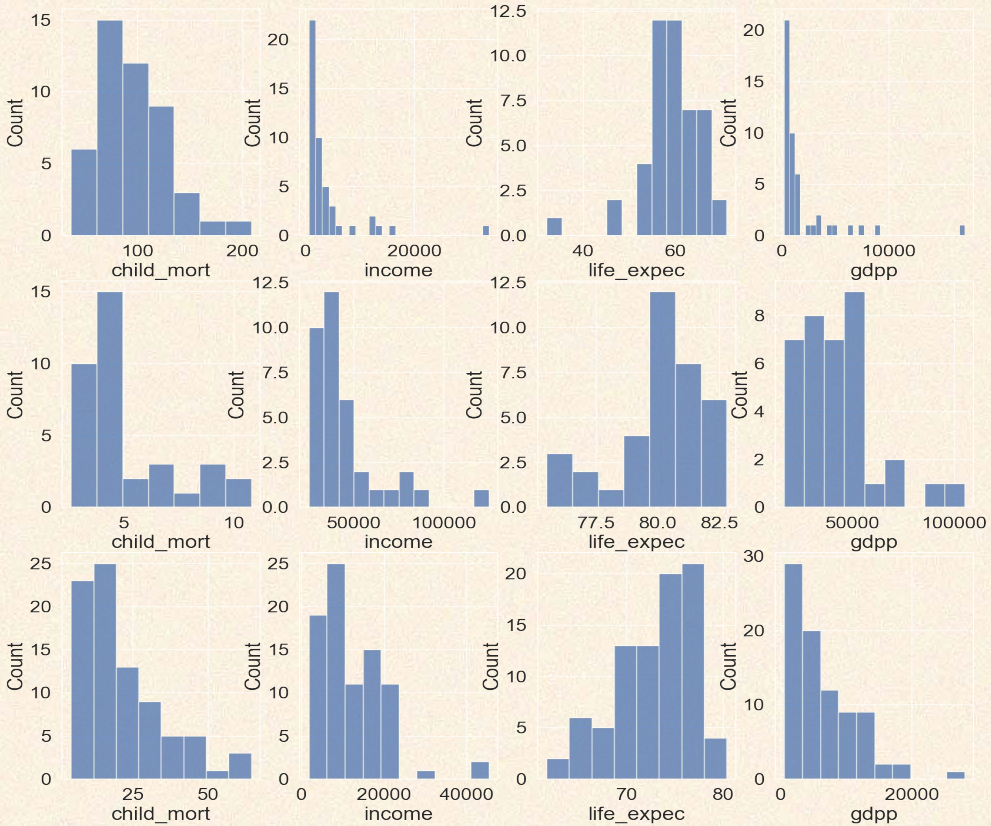

1
0
−1
inflation
child_mort
total_fer
imports
exports
health
life_expec
income
gdpp
1
0
2
labels

Labelling

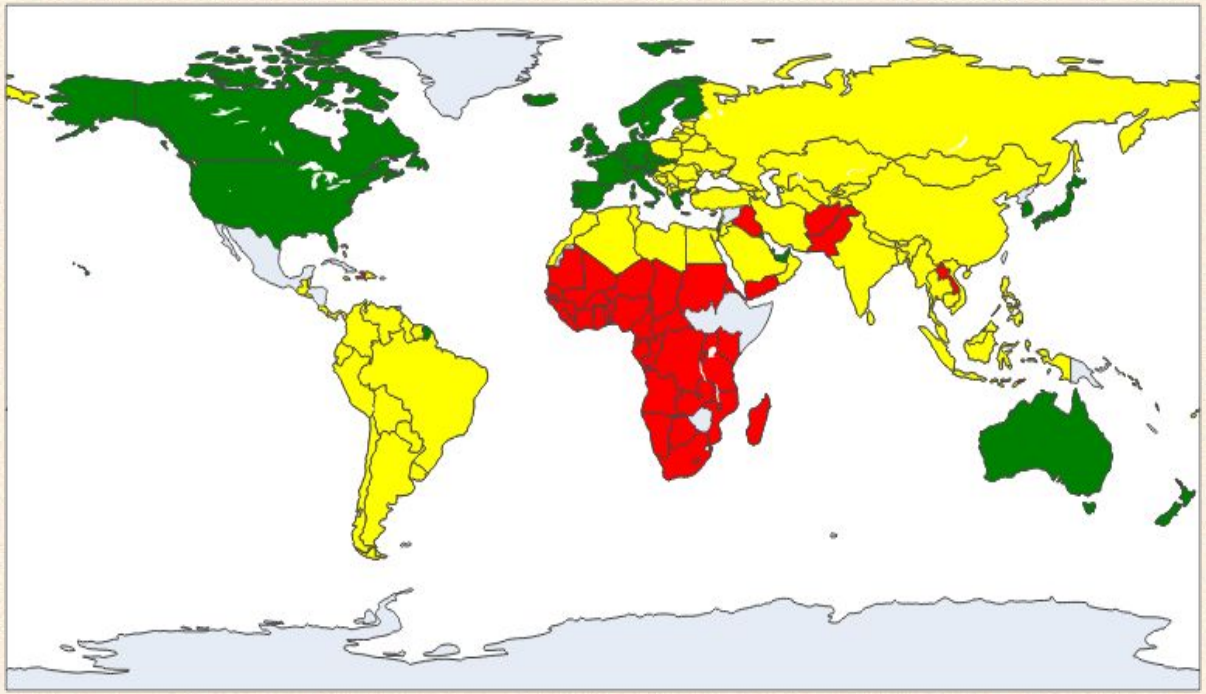
Cluster 0 :

Cluster 1:

Cluster 2:



Labelling



- Green : Developed Countries
- Yellow : Least Developed Countries
- Red : Developing countries

04
Conclusion

Inferences and final review

Conclusion

- We can **categorize countries** based on the features of our dataset.
- Characteristics of **Developed Countries** are : higher GDPP, Income and Life expectancy .
- Characteristics of **Developing Countries** are : higher Child Mortality, higher total Fertility .
- **Least Developed Countries** fall in between the Rich and Poor Countries .

Thank You !!!