

Family Wealth - documentation

Authors

Sławomir Gicała

Michał Sokół

Krzysztof Tylka-Suleja

1 Variables description

- 1.1 **earnings**: monthly earnings per month
- 1.2 **cost of living**: monthly cost of products being bought
- 1.3 **real wealth factor**: monthly proportion of cost of living to earnings

2 Data

- 2.1 Location: **dataset**
- 2.2 Modules:
 - 2.2.1 **cost_of_living**
 - 2.2.1.1 name: **cost_of_living.csv**
 - 2.2.1.2 description: monthly cost of living per country per year (in USD)
 - 2.2.2 **cost_of_living_labels**
 - 2.2.2.1 name: **cost_of_living_labels.csv**
 - 2.2.2.2 description: countries clustering in regards to four groups
 - 2.2.3 **earnings_split**
 - 2.2.3.1 name: **earnings_split.csv**
 - 2.2.3.2 description: monthly earnings per country per profession per year (in USD)
 - 2.2.4 **trends_real_wealth_from_2010_to_2022**
 - 2.2.4.1 name: **trends_real_wealth_from_2010_to_2022.csv**
 - 2.2.4.2 description: trend prediction for real wealth factor per country per profession per year
 - 2.2.5 **wealth_factor**
 - 2.2.5.1 name: **wealth_factor.csv**
 - 2.2.5.2 description: real wealth factor per country per profession per year
 - 2.2.6 **wealth_factor_labels**
 - 2.2.6.1 name: **wealth_factor_labels.csv**
 - 2.2.6.2 description: real wealth factor per country per profession clustering in regards to four groups
 - 2.2.7 **wealth_factor_total_labels**
 - 2.2.7.1 name: **wealth_factor_total_labels.csv**
 - 2.2.7.2 description: real wealth factor per country clustering in regards to four groups

3 Analysis

- 3.1 cost of living
 - 3.1.1 location: **cost_of_living_calculation.ipynb**
 - 3.1.2 methods
 - 3.1.2.1 monthly proportion of bought product
 - 3.1.2.2 boxplot per country
- 3.2 real wealth factor
 - 3.2.1 location: **real_wealth_factor_calculation.ipynb**
 - 3.2.2 methods

- 3.2.2.1 boxplot per total mean per country
- 3.2.2.2 boxplot per profession
- 3.2.2.3 boxplot per profession per country
- 3.2.2.4 boxplot per country per profession
- 3.2.2.5 boxplot for each country for each profession

4 Clustering

- 4.1 location: **clustering_calculation.ipynb**
- 4.2 **cost of living**
 - 4.2.1 methods
 - 4.2.1.1 Hopkins test
 - 4.2.1.2 fuzzy c-means clustering
 - 4.2.1.3 partition coefficient and partition entropy coefficient to find best number of clusters
 - 4.2.1.4 boxplot per country per cluster
- 4.3 **real wealth factor**
 - 4.3.1 methods
 - 4.3.1.1 Hopkins test
 - 4.3.1.2 fuzzy c-means clustering
 - 4.3.1.3 partition coefficient and partition entropy coefficient to find best number of clusters
 - 4.3.1.4 boxplot per country per cluster
 - 4.3.1.5 list of overlapping countries between clusters
- 4.4 **cross analysis**
 - 4.4.1 cluster comparison between cost of living and real wealth factor clusters

5 Trend prediction

- 5.1 real wealth factor
 - 5.1.1 location: **real_wealth_trend_calculation.ipynb**
 - 5.1.2 methods
 - 5.1.2.1 linear regression
 - 5.1.2.2 list of professions in regards to country with a tendency to grow
 - 5.1.2.3 tendency to grow heatmap of professions in regards to country