Collected Data

Marian Petruk

May 18 2017

Contents

1 Indicators **2**

2 Example **2**

3 Amount of the data **3**

1 Indicators

The data constist of eleven indicators. Seven of them are collected from the section "economy and growth" and four other are from different sections. From these indicators, it is possible to find a characteristic for a particular year. We are interested to analyse the trends in the period from 1990 to 2016. The indcators are presented below.

```
    #                       Indicators                          #
    # ----------------------------------------------------------#
    #                   Economy and growth                      #
    # ----------------------------------------------------------#

    gdp = "NY.GDP.MKTP.CD"  # GDP (current US dollars)
    gdp_per_capita = "NY.GDP.PCAP.CD"  # GDP per capita (current US dollars)
    gross_savings = "NY.GNS.ICTR.ZS"  # Gross savings (% of GDP)
    inflation_gdp = "NY.GDP.DEFL.KD.ZG"  # Inflation, GDP deflator (annual %)
    imports = "NE.IMP.GNFS.ZS"  # Imports of goods and services (% of GDP)
    inflation_consumer_prices = "FP.CPI.TOTL.ZG"  # Inflation, consumer prices (annual %)
    gni = "NY.GNP.MKTP.PP.CD"  # GNI, PPP (current international dollars)

    # ----------------------------------------------------------#
    #                  Some other indicators                    #
    # ----------------------------------------------------------#

    total_population = "SP.POP.TOTL" # total population indicator
    life_expectancy = "SP.DYN.LE00.IN"  # Life expectancy at birth, total (years)
    high_tech_exports = "TX.VAL.TECH.CD"  # High-technology exports (current US£)
    science_tech_articles = "IP.JRN.ARTC.SC"  # Scientific and technical journal articles
```

2 Example

The table below shows the part of the code and result of the run. The data structure (2D Array) presents the "total population" indicator of Ukraine and Poland from 1990 to 2016:

```
>>> dataset = api.get_dataset(total_population, iso_country_codes, date="1990:2017")
>>> gdp_2d_array = Array2D(27, 2)
>>> gdp_2d_array[(0, 0)] = "UA"
>>> gdp_2d_array[(0, 1)] = "PL"
>>> for i in range(1, 27):
>>> ...  gdp_2d_array[(i, 0)] = dataset.as_dict()['UA'][str(i-1+1990)]
>>> ...  gdp_2d_array[(i, 1)] = dataset.as_dict()['PL'][str(i-1+1990)]
>>> print(gdp_2d_array)
UA PL
51892000.0 38110782.0
52000470.0 38246193.0
52150266.0 38363667.0
52179210.0 38461408.0
```

```
51921041.0 38542652.0
51512299.0 38594998.0
51057189.0 38624370.0
50594105.0 38649660.0
50143939.0 38663481.0
49673350.0 38660271.0
49175848.0 38258629.0
48683865.0 38248076.0
48202500.0 38230364.0
47812950.0 38204570.0
47451600.0 38182222.0
47105150.0 38165445.0
46787750.0 38141267.0
46509350.0 38120560.0
46258200.0 38125759.0
46053300.0 38151603.0
45870700.0 38042794.0
45706100.0 38063255.0
45593300.0 38063164.0
45489600.0 38040196.0
45271947.0 38011735.0
45154029.0 37986412.0
```

---

3 Amount of the data

The amount of the collected data is evaluated with the following formula: $AD = (EGI + SOI) \cdot NY \cdot NC$, where

- **EGI** - Economy and growth indicators;
- **SOI** - Some other indicators;
- **NY** - Number of years;
- **NC** - Number of countries;
- **AD** - Amount of the data.

For this research it is $(7 + 4) \cdot 26 \cdot 2 \approx 572$ numbers (percentages, years, money etc.).
The amount of data is approximate because some indicators do not have data for some years.