# **ENSF 592 Spring 2021 – Final Project Report**

Authors: Bhavyai Gupta, Brandon Attai

Date: 16th June, 2021

Course: ENSF 592

## **Summary**

This program allows a user to request, select, or visualize various statistics surrounding data provided by United Nations to give insight into population and wealth trends for UN Regions, Sub-Regions, and Countries. To achieve this, a total of four Datasets are used as follows:

* UN Region, Sub-Region and Country
* Population Growth, Fertility, Life Expectancy and Mortality
* Population in the Capital City, Urban and Rural Areas
* GDP and GDP per Capita

The program makes use of four python source files to provide an interactive command line user interface. launch.py serves as the starting point of the program, as it controls the execution flow by using menu options.

As soon as the program is launched, it will initialize via a five-step process. These five steps are (1) importing the data into pandas’ DataFrames, (2) merging DataFrames together, (3) adding computed columns to the merged DataFrame, (4) performing checks for missing values, and (4) export of the final dataset into the project directory. User is given real time updates on each step and its status.

Once user enters the program menu after the initialization, the user would have 7 options to select from. These are summarized in *Table 1*. For example, the user can use menu option number 4 to print the aggregate stats on GDP per capita (US dollars) grouped by either UN Regions or UN Sub-Regions.

|  |  |
| --- | --- |
| **Option Number** | **Option Description** |
| 1 | Print the imported dataframes |
| 2 | Re-export the entire merged hierarchical dataset into Excel |
| 3 | Print aggregate stats for the entire dataset |
| 4 | Print aggregation stats grouped by UN Region/UN Sub-Region and available years |
| 5 | Print the list of countries that have higher GDP per capita than USA, and the year |
| 6 | Compare four different countries on various statistical data and plot graphs |
| 0 | Exit |

*Table 1*

One thing to note is that some countries may not get accepted as a valid input by the program. Examples include Russia and Iran. This is because these countries do not have adequate data, which results in NaN in the merged dataset. NaN values are dropped during the initialization process.

The program makes use of object-oriented programming with the help of classes, to efficiently handle data analysis and calculations. Finally, exception handling is performed throughout the program to ensure the program does not terminate if the user enters an invalid input.

A matrix summarizing how the requirements are met is shown in Appendix 1 on page 3.

## **References**

1. UN Region, Sub-Region and Country, Development Data Section of the Development Data and Outreach Branch within the Statistics Division of the Department of Economic and Social Affairs (UN DESA) of the UN Secretariat, June 2019. [Online]. Available:

<https://data.un.org/_Docs/SYB/CSV/SYB63_1_202105_Population,%20Surface%20Area%20and%20Density.csv>

1. Population Growth, Fertility, Life Expectancy and Mortality, Development Data Section of the Development Data and Outreach Branch within the Statistics Division of the Department of Economic and Social Affairs (UN DESA) of the UN Secretariat, Aug. 2019. [Online]. Available:

<https://data.un.org/_Docs/SYB/CSV/SYB62_246_201907_Population%20Growth,%20Fertility%20and%20Mortality%20Indicators.csv>

1. Population in the Capital City, Urban and Rural Areas, Development Data Section of the Development Data and Outreach Branch within the Statistics Division of the Department of Economic and Social Affairs (UN DESA) of the UN Secretariat, May 2018. [Online]. Available:

<https://data.un.org/_Docs/SYB/CSV/SYB61_253_Population%20Growth%20Rates%20in%20Urban%20areas%20and%20Capital%20cities.csv>

1. GDP and GDP per Capita, Development Data Section of the Development Data and Outreach Branch within the Statistics Division of the Department of Economic and Social Affairs (UN DESA) of the UN Secretariat, Nov. 2020. [Online]. Available:

<https://data.un.org/_Docs/SYB/CSV/SYB63_230_202009_GDP%20and%20GDP%20Per%20Capita.csv>