

# Writing About the Data

(Regression Modeling in Practice Week 1 Assignment – Step 1)

## Expected Activities

Describe the sample - Provide enough detail so that the reader can clearly understand the population that the study sample came from. Use meaningful labels. Do not use abbreviations (“PPM100”) or variable names.

a) Describe the study population (who or what was studied).

*The social and economic aspects of the data provided by **Gapminder** through their dataset, made me interested in it and I have decided to choose the same for further analysis and research. Gapminder dataset contains **various statistical data related to social, economic, and environmental development of all the countries of the world.***

*Note: Gapminder (<https://www.gapminder.org/>) is a non-profit venture that seeks to promote sustainable global development and achievement of the United Nations Millennium Development Goals. It seeks to increase the use and understanding of statistics about social, economic, and environmental development at local, national, and global levels.*

b) Report the level of analysis studied (individual, group, or aggregate)

*Aggregate - The data is captured at country level.*

c) Report the number of observations in the data set.

*Gapminder dataset contains data for all 192 UN members, aggregating data for Serbia and Montenegro. Additionally, it includes data for 24 other areas, generating **a total of 215 areas.***

d) Describe your data analytic sample (the sample you are using for your analyses).

*I will be using the complete set of records for the analysis here. However, after examining the codebook for the Gapminder dataset, I got particularly interested in the following parameters and intending to perform regression analysis using those.*

- Life Expectancy (“lifeexpectancy”)
- Urbanization Rate (“urbanrate”)
- Income per Person (“incomeperperson”)