

Assignment1

R Programming for Public Policy

ByungKoo Kim

The assignment is due Feb 27th before the lecture. Please use R Markdown to compile your assignment pdf. You can work with your classmates, but your answers must be your own. Do NOT use any generative AIs, and do NOT copy from other people's work.

Part 1. Vectors and Data Frames

- Import the data in `assignment1_vectors.csv` file and save the dataset into the object name `df`.
- How many observations are there in this dataset? How many variables?
- What are the averages of each variable?
- What are the minimum and maximum values of each variable?
- Create a subset of the data frame that meets the condition `v1 < 10`.
- Create a subset of the data frame that meets the conditions `v1 < 10` and `v2 > 5`.

Part 2. Visualizations

```
install.packages("gapminder")
```

For this exercise, we will use the `gapminder` dataset in `gapminder` package.

```
library(gapminder)
```

- Describe the dataset (use `?gapminder`).
- Create a subset `gapminder07` from `gapminder` that only contains the data for 2007.
- How many different countries are in the data?
- Visualize the distribution of life expectancy in 2007.
- Describe the relationship between GDP per capita and life expectancy in 2007 using visualization.
- Visualize the relationship between GDP per capita and life expectancy in 2007, with points colored by continents (use `points()`). Add legends to the plot to explain what colors correspond to what continents.
- Compute the average life expectancy for each continent in 2007.
- Pick three countries of your choice in `gapminder` data. Visualize the over time change of life expectancy in those countries.