

Week 3/3 Programming Workbook for R

Hyeonsu B. Kang
hyk149@eng.ucsd.edu

January 2015

1 In-class Practice Questions

1.1 Data Management with dplyr

- Exercise 1. Download the Gapminder data from [here](#).
- Exercise 2. What is the (absolute) difference between the average GDP per capita of Americas and Asia in year 2002?
- Exercise 3. Which continent had the highest GDP per capita in year 2002?
- Exercise 4. Which continent had the lowest GDP per capita in year 2002?
- Exercise 5. Which country had the highest GDP per capita in year 1967?
- Exercise 6. Which country had the highest life expectancy in year 2002 and what was the value of it?
- Exercise 7. What does the distribution of the log of GDP per capita look like in year 1967?
- Exercise 8. What does the distribution of the life expectancy in year 2002 look like?
- Exercise 9. Add a log(GDP per capita in 1967) column to the data frame that contains GDP per capita by countries in 1967.
- Exercise 10. Graph a scatter plot with an **x** axis of “log(GDP per capita in 1967)”, and a **y** axis of “life Expectancy in 2002”.

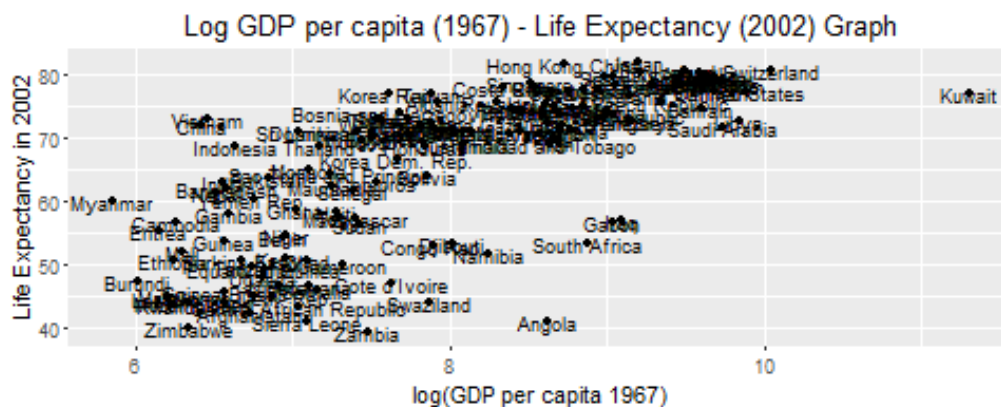


Figure 1: The target plot for Exercise 10

- Exercise 11. Graph a scatter plot with an **x** axis of “log(GDP per capita per continent in 1967)”, and a **y** axis of “Life expectancy per continent in 1967”.
- Exercise 12. Graph a line plot with an **x** axis of “year”, and a **y** axis of “log(GDP per capita per continent in 1967)”.

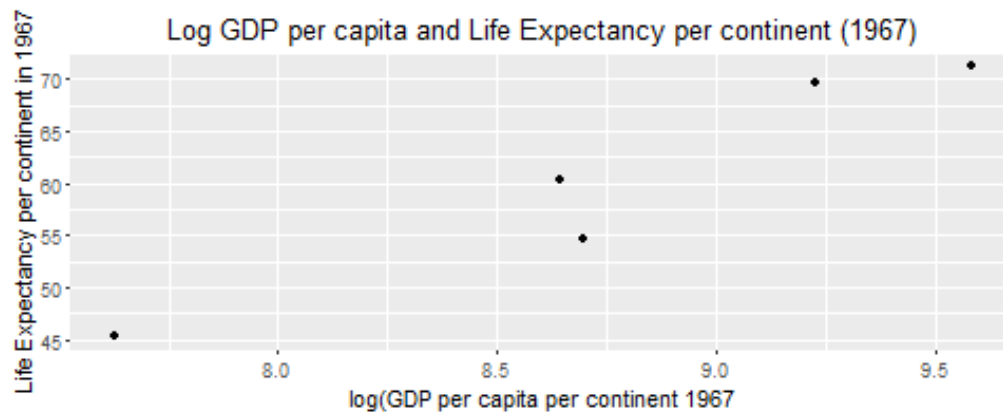


Figure 2: The target plot for Exercise 11

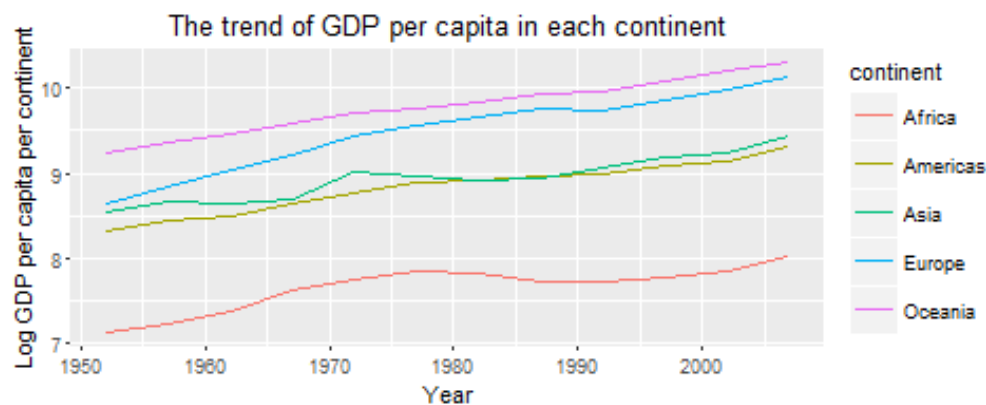


Figure 3: The target plot for Exercise 12

1.2 Data Shaping with tidyr

Exercise 1. Load the revenue.csv file into your R project.

Exercise 2. By using `gather()`, 'gather' the quarter revenues into `Quarter` and `Revenue`.

Exercise 3. By using `separate()`, 'separate' each of the `Quarter` strings into `Time_Interval` and `Interval_ID`.

Exercise 4. By using `unite()`, re-unite the `Time_Interval` and `Interval_ID` and re-create the original `Quarter` variable.

Exercise 5. By using `spread()`, restore the original structure of the dataframe.

2 Weekly Challenges (30pts) Due by 01/28, 11:59pm

For the submission method, please look at Appendix.

2.1 (40ts) Problem 1. Is the world's wealth & health gap widening?

- (1) (10pts) Using the gapminder data we utilized in class, calculate the difference between the poorest country and the wealthiest country's GDP per capita over the years.

HINT: Please fill up the blank in the box below.

```
gdp_grouped_byyear <- gdp_bycountries_byyear %>% group_by(year)
gdpDiff_byyear <- gdp_grouped_byyear %>%
  summarize()
print(gdpDiff_byyear)
```

- (2) (10pts) Using the gapminder data we utilized in class, calculate the difference between the unhealthiest country and the healthiest country's life expectancy over the years.
- (3) (10pts) Create a line plot that shows both the country of the highest GDP per capita and the country of the lowest GDP per capita over the years. Color the highest and lowest GDP per capita lines differently.
- (4) (10pts) Create a line plot that shows both the country of the longest life expectancy and the country of the shortest life expectancy over the years. Color the longest and shortest life expectancy lines differently.

2.2 (10pts) Problem 2. The health trend of each continent

Graph a line plot that shows the average life expectancy over the years. Color each continent's trend line differently.

APPENDIX

How to submit your challenge solutions:

- (1) Create an `.r` file for each problem: `w3p1.r` for Problem 1 and `w3p2.r` for Problem2. Please DO NOT modify these file names
- (2) Type the solution in the script file. Problem 1 will have 4 different parts. Make sure to comment which part of your code solves which part of the question. E.g. “`# This is Problem 1-(1)`”.
- (3) IMPORTANT: For Problem 1-(1) and Problem 1-(2), make sure to add an extra `print()` function at the end of your script. E.g. `print(YOUR_VARIABLE_NAME)`
- (4) For Problem 1-(3), Problem 1-(4) and Problem 2, save the graphs as `w3p1-(3).png`, `w3p1-(4).png` and `w3p2.png`, respectively.
- (5) Save the script files and collect the graph images.
- (6) Attach the (2) script files and the (3) image files and send an email to `challenges.gps.programming.ta@gmail.com`.

NOTE: Please use the same email address throughout the quarter (your email address will be the key for you to find out the grades for the assignments)