

Continental Solutions

```
# load libraries
library(tidyverse)
library(ggthemes)
```

In this mini-assignment, you are going to use the following data set.

- The file `processed-20220221-owid-life-expectancy-vs-gdp-per-capita.csv` data set. (Max Roser & Ritchie, 2013)

```
D1 <- read_csv("processed-20220221-owid-life-expectancy-vs-gdp-per-capita.csv")
```

1. Create a subset of the data where you only choose the year range 2008-2018, and take the mean of the `life_expectancy` and `gdp_per_capita` variables for each country. Remove the rows with “Antarctica” in the `continent` variable. You can use the not-equal-to `!=` operator. Remove any rows with missing values.

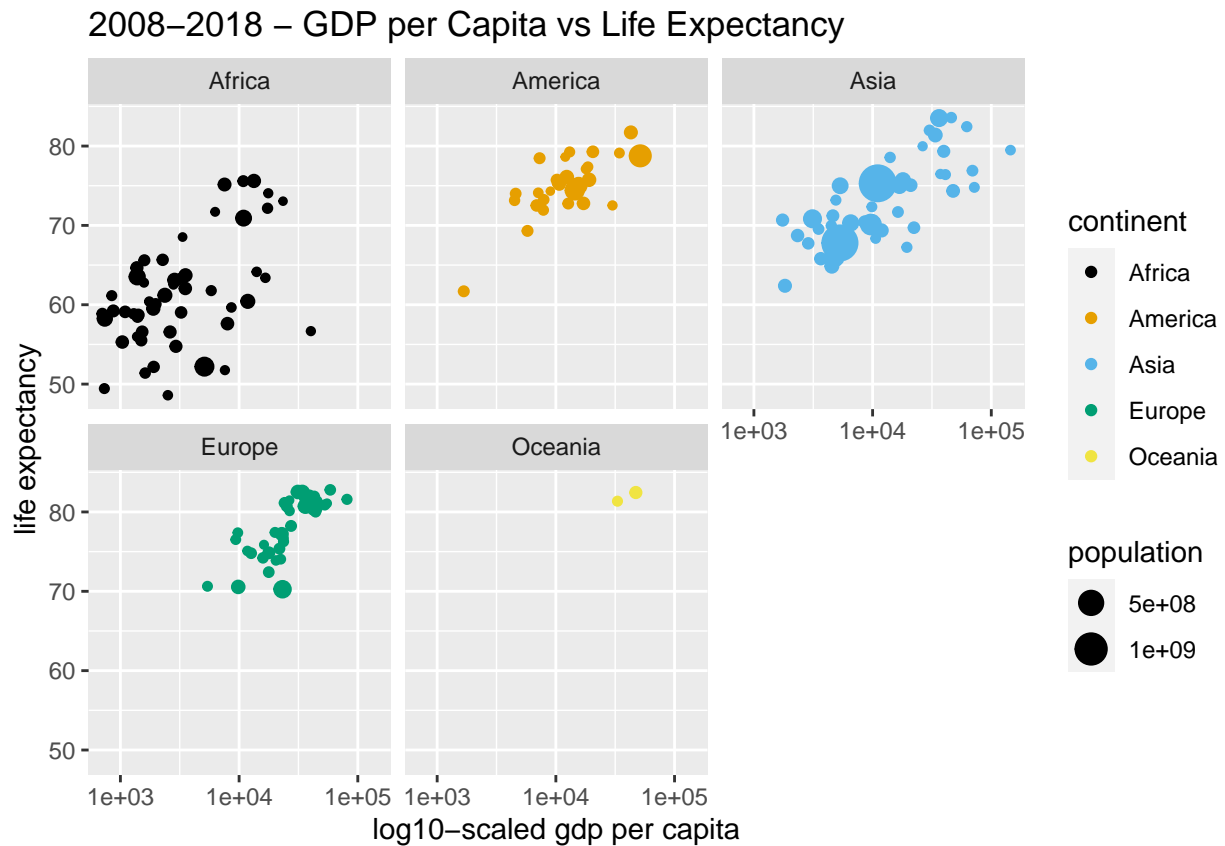
SOLUTIONS:

```
D1_sub <- D1 %>%
  # filter only years 2008-2018
  filter(year >= 2008 & year <= 2018) %>%
  # filter remove antarctica
  filter(continent != "Antarctica") %>%
  # take means for each country
  group_by(country, continent) %>%
  summarise(mean_le = mean(life_expectancy),
            mean_gdp = mean(gdp_per_capita),
            # we need to include this for the next problem
            max_pop = max(population_estimates),
            .groups = "drop") %>%
  # remove any rows with missing values
  drop_na()
```

- Using the subset data from problem 1, create a figure where the x-axis is the `gdp_per_capita` variable and the y-axis is the `life_expectancy` variable. Color each point according the `continent` variable, and size each point proportional to the `population_estimates` variable. These numerical variables are the summarised version from problem 1. Use the `facet_wrap()` function to separate each continents into sub-figures and use the `scale_colour_colorblind()` function to color the points. Use the `log10()` transformation on the x-axis. No need to write your observations.

SOLUTIONS:

```
ggplot(data = D1_sub, aes(x = mean_gdp, y = mean_le,
                          color = continent, size = max_pop)) +
  geom_point() +
  scale_x_log10() + # transform the x-axis using the log scale
  scale_colour_colorblind() + # apply the colorblind color palette
  facet_wrap(~continent) + # separate the plots by continent
  labs(x = "log10-scaled gdp per capita",
       y = "life expectancy",
       color = "continent",
       size = "population",
       title = "2008-2018 - GDP per Capita vs Life Expectancy")
```



3. Open the complement mini-assignment file named `complement-mini-assignment-20220224.Rmd` and finish the instructions written inside. Knit the Rmd into HTML, and upload the resulting files using the [Google form](#).

SOLUTIONS: See the [corresponding HTML file](#).

References

Max Roser, E. O.-O., & Ritchie, H. (2013). Life expectancy. *Our World in Data*. <https://ourworldindata.org/life-expectancy>