Problem Set 02

WRITE YOUR NAME HERE

2021-09-20

Learning goals

- Learn to "knit" R Markdown documents and create a PDF
- Recreate Figure 2.1 in ModernDive.
- Get familiar with the 6 R Markdown Fixes

Honor code

For this problem set I worked with:

Getting started

- At the top of this document replace `"WRITE YOUR NAME HERE"` with your name, including the quotation marks. Ex: `"Albert Y. Kim"`
- In the Honor code section above, indicate who (if anyone) you worked with.

Setup

Load necessary packages:

```
library(ggplot2)
library(dplyr)
library(gapminder)
```

The `gapminder` package includes a data frame called `gapminder`, containing information about different countries from 1952 to 2007. Run `View(gapminder)` in your console to explore this data.

We're only going to only focus on data for 2007 and save this in a new data frame `gapminder_2007`. We use data wrangling using code from the `dplyr` package. Don't worry if this doesn't make sense for now, we'll cover data wrangling in Chapter 3 of ModernDive. For now, run the following code:

```
gapminder_2007 <- gapminder %>%
  filter(year == 2007)
```

Run `View(gapminder_2007)` in your console to explore this data. An alternative method for exploring a data frame is by using the `glipmse()` function:

```
glimpse(gapminder_2007)
```

```
## Rows: 142
## Columns: 6
## $ country   <fct> "Afghanistan", "Albania", "Algeria", "Angola", "Argentina", ~
## $ continent <fct> Asia, Europe, Africa, Africa, Americas, Oceania, Europe, Asi~
```

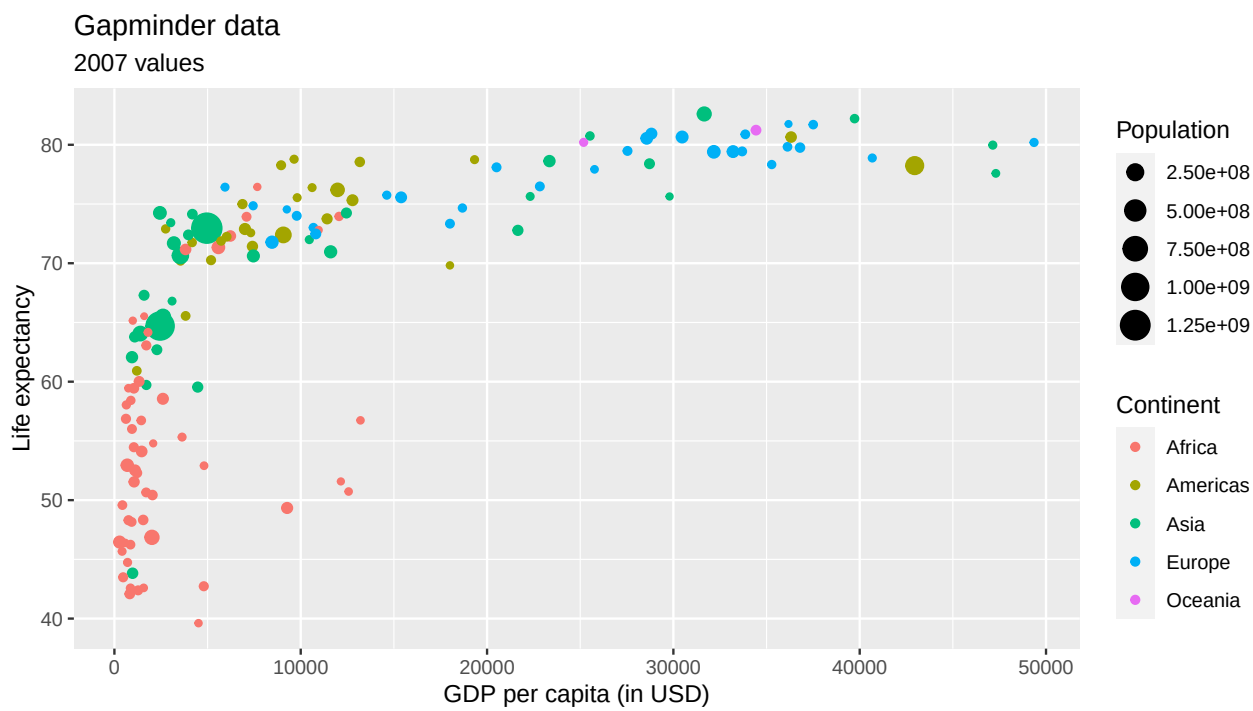
```
## $ year      <int> 2007, 2007, 2007, 2007, 2007, 2007, 2007, 2007, 2007, 2007, ~
## $ lifeExp   <dbl> 43.828, 76.423, 72.301, 42.731, 75.320, 81.235, 79.829, 75.6~
## $ pop       <int> 31889923, 3600523, 33333216, 12420476, 40301927, 20434176, 8~
## $ gdpPercap <dbl> 974.5803, 5937.0295, 6223.3675, 4797.2313, 12779.3796, 34435~
```

This data is in fact the same data as in Table 2.1 of ModernDive.

Question 1

Write the `ggplot2` that recreates Figure 2.1 of ModernDive. Do not worry about labeling the axes and legends for now.

```r
# Write your code below:
ggplot(data = gapminder_2007,
       mapping = aes(x = gdpPercap, y = lifeExp, size = pop, color = continent)) +
  geom_point() +
  # Note: this was not required for this problem set:
  labs(x = "GDP per capita (in USD)",
       y = "Life expectancy",
       size = "Population",
       color = "Continent",
       title = "Gapminder data",
       subtitle = "2007 values")
```



Question 2

Q: What country had the highest GDP per capita in 2007? Use the sorting arrows next to variable names in RStudio's `View()` function.

A: Norway