Week 2 Homework

Problem 1: Exploring the gapminder data

1. Load in the gapminder data from the `dslabs` package using the data() function.

2. Use the class function to determine what type of object the data is.

3. Check the dimensions of the gapminder data. Report the total number of rows and columns within the data frame.

4. Print the column names to your console.
5. Let's say we wanted to check for any attributes that might be hidden in the data. Use the attributes function to further explore the life expectancy and fertility variables and report your findings.

6. Why do you think there aren't any hidden attributes in this dataframe?

7. We want to find the population size of Norway in `1962`. First, find this information manually using the View() function, then use bracket notation to find the population size by column and row.

8. Suppose we want to filter our data to only observe countries with life expectancies lower than 40. Use bracket notation to subset the data accordingly. Are there any countries that recorded this low of a life expectancy in the 21st century?

Problem 2: College Majors

Part A

1. Read in the data "recent-grads.csv" that can be found on Canvas. Name the data recent_grads.

2. Print out to your console the dimensions, class, and names of rows/columns of the recent_grads object.

3. Check to see if there are attributes for the dataframe. Then, check to see if there are attributes for the columns. If there isn't, why is this the case?

Part B

1. Find the major category with highest and lowest median income.

2. Find the average Median salary for both of those majory categories.

3. Are there any professions that both (a) are one of the top 10 professions with the highest median income and (b) are in the top ten professions with the highest share of women? Hint: start by making variables for top income and share of women.

4. According to this dataset, are women underrepresented in STEM fields?