State Lab – Part I

This lab is about understanding aspects related to states within the United States.

**Data description**

Summarization of aspects related to each state such as income, size of the state, and population during the year of 1977**\***.

**\*U.S. Department of Commerce, Bureau of the Census (1977) *County and City Data Book*.**

The dataset contains details related to state per capita information for each state within the United States for the consensus year 1977.

**Instructions:**

• Read the lab all the way through

• Complete the lab

• Clearly label all .R scripts and submit relevant scripts, word documents and slides

Section 0: State Lab dataset

Open R studio, inside of the workspace install the **datasets** package. From there use the **library** statement to call the **datasets** package recently installed. Use the **data** function to call the **state** dataset. View the **state** dataset in your environment window.

**install.packages(“datasets”)**

**library(“datasets”)**

**data(state)**

Section 1: Explore the dataset and practice R, basic operations

1. Explore the data to learn the following information:
   1. Mean Area
   2. Median Illiteracy rate
   3. Minimum income
   4. The number of states that fall within each division
2. Create a new copy of your data frame, named "StatesDF" and use this data frame for the rest of the problems
3. Calculate the population per capita for the state of Utah
4. Use View(StatesDF) to visually inspect the data
5. **Challenge:** Find the state that has the highest life expectancy
   1. Hint: use the which.max() function to find the index of the state that has the highest life expectancy
   2. Hint: use the rownames() function and the index you found in 5a to locate the name of the state with the highest life expectancy

**PAUSE HERE, CHECK ANSWERS WITH CLASS, BRING FORWARD ANY QUESTIONS**

Section 2: Manipulate the dataset

**Use the StatesDF data frame for these exercises.**

1. Report the following for the state you located in section 1 question 5:
   1. What is the data type of this particular state’s population
   2. What is this state’s population
2. **Challenging:** Create a new column called PopulationGroup.
   1. In it create a category for states that follow the given criterion: population ≤ 2838 PopulationGroup=low; population > 2838 PopulationGroup =high
   2. What is the data type of PopulationGroup column just created?
3. Answer the following questions:
   1. StatesDF is a matrix. True or False? (if False what data object is it):
   2. The minimum value of income is 4755? True or False? (If False what is the minimum income):
   3. What state does the minimum income belong to?
   4. Return the value at row 39 column 6.
   5. What is the column name and what states does this value belong to?

**Save your work for use in a later lab: save("StateLabPt1.RData")**