# Homework (HW03) Descriptive Stats and Functions

### General Instructions

For this homework you will upload 1 R file into blackboard.

Reminder:

* All HW must start with an Identification Block like this sample…

################################################

# IST 387/687, Standard Homework Heading

#

# Student name:

# Homework number:

# Date due:

#

# Attribution statement: (choose the statements that are true)

# 1. I did this work by myself, with help from the book and the professor

# 2. I did this work with help from the book and the professor and these Internet sources: <provide the urls>

# 3. I did this work with coaching from <Name of another student> but did not cut and paste any code

# Run these three functions to get a clean test of homework code

dev.off() # Clear the graph window

cat('\014') # Clear the console

rm(list=ls()) # Clear all user objects from the environment!!!

# Set working directory

# Change to the folder containing your homework data files

setwd("~/MyDesktop/ISTX87/Homework")

### HW03

**Step 1: Using your R code from the weekly Prep Exercise find the states with the largest and smallest populations.**

1. Create a function, call readStates(), that returns the cleaned dataframe of the states population (i.e. executes all of your R-code from the Prep work, but within a function so that we can load the states dataset anytime we want).
2. Use your newly created function to create a new dataframe, called dfStates.
3. Using the Pop2011 column within dfStates, calculate the minimum, mean, and maximum population of the states.
4. Use which.max() on the data from the Pop2011 column to find the row that contains the largest state. How can you use the results of your which.max() expression to output just the name of the state?
5. Use which.min() on the data from the Pop2011 column to find the row that contains the smallest state. How can you use the results of your which.min() expression to output just the name of the state?

**Step 2: Sort the data frame by population size using the Pop2011 numeric column**

1. Store the results in a new dataframe called dfStatesOrdered.

**Step 3: Creating a Plot**

1. Create a histogram of one of the numeric columns from the sorted dataframe, what do you observe?

***You must submit all Homework to blackboard prior to the deadline specified for each assignment.***

Late HW assignments will not be accepted for credit.

*Materials provided by the instructors of this course are copyrighted 2018 by Jeffrey Saltz and Jeffrey Stanton, or by the respective instructor that produced them. These materials are provided to enrolled students for their exclusive use during the semester of instruction. Uploading instructor-produced materials to Internet sites (e.g., Coursehero) without permission is considered a breach of professional ethics and will be reported as a possible academic integrity violation.*