# Homework (HW07) Data Viz: Using GGPLOT and GGMAP

### 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")

This homework builds on our efforts from the Prep Exercise and depends on a careful read of Chapter 13 of *An Introduction to Data Science*. As usual we’ll use the dataframe we created in our Prep Ex. Reminder: our dataframe contains data about median incomes in zipcodes. We also used the merge function to add state names. In this homework we continue to use ggplot, together with a companion package called ggmap. This companion package enhances the capabilities of ggplot by adding the capability to draw geographic outlines (polygons), shading, labeling, and other map markings.

### HW06

**Step 1:** **Plot points for each zipcode (don’t forget to library ggplot2 and ggmap)**

1. Code and execute the following lines of code

us <- map\_data("state")  
dotmap<- ggplot(dfNew, aes(map\_id = stateNname))  
dotmap<- dotmap + geom\_map(map = us)  
dotmap<- dotmap + geom\_point(aes(x=longitude,y=latitude,color=Mean))  
dotmap

1. Comment each line of code explaining what it does.
2. Add a block comment that criticizes the resulting map. It’s not very good (but you need to explain why it is not very good).

**Step 2: Use Tidyverse to create a Data Frame of State-by-State Income**

1. Library the tidyverse() package (if you have not already done so), and then run the following command to create a new data frame:

summaryDF <- dfNew %>%

group\_by(stateName) %>%

summarize(totalPop=sum(Pop), Income=sum(Mean\*Pop))

1. Add a comment after each line, describing what each line of code does. Make sure to describe how many rows there are in the new dataframe, and how the new dataframe was created.
2. Create a new variable on this data frame called meanIncome. Divide Pop by Income to find the average income for each state.

**Step 3: Create a map of the U.S. showing average income**

1. Create a map visualization, where the color of each state represents the mean income for that state.
2. If you need help creating the map, review Chapter 13, and how Figure 13.2 was created.
3. You will notice some states are grey. In a block comment explain why they are grey
4. Fix this issue so that all states have an appropriate shade of blue – i.e., generate the map visualization, where the color of each state represents the mean income for that state without any state being grey (hint: look if there are NAs).

**Step 4: Show the population of each state on the map**

1. Use the stateNameDF (that was created and used in the prework) and merge that dataframe with the summaryDF dataframe, so that the center.x and center.y information is merged with the summaryDF dataframe information in your new dataframe.
2. Create a new map visualization, which is the same visualization created in step 3, but now add a new layer, with a point, at the center of each state, where the point size represents the total population of that state.
3. If you need a hint on how to do this visualization, you need to have a center.x and center.y in your summaryDF (i.e., you need to create a new summaryDF with center.x and center.y).

***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.*