

# Homework 1

## Overview:

In this assignment, you'll be using R/RStudio and the GSS data set (2014) to report univariate statistics for two dependent variables of interest. These variables must be different than the ones I present in the example below (unless you've discussed it with me).

## What you will need:

The following can be downloaded from our course website

- **hw1\_script.R**
- **GSS2014\_final.csv** aka GSS Data Set (2014)
- **LABELS\_script.R**
- **GSS Variables and Descriptions.pdf**
- **GSS\_Codebook.pdf**

## Tasks:

- Open up the Homework 1 script/code file (**hw1\_script.R**). This is the file you'll be working from.
- Set your working directory
- Load the required packages/libraries
- Download the data set (**GSS2014\_final.csv**) from under "R Resources" on our course website, save the data set on your computer, and pull in the data set by locating that file's directory. (Alternatively, you can pull in the data set using using URL method.)
- Download the cleaner/labels file (**LABELS\_script.R**) from under "R Resources" on our course website, save the labels file on your computer, and pull in the labels file by locating that file's directory. (Alternatively, you can pull in the labels file using using URL method.)
- Select two (2) interval/continuous dependent variables that interest you. To do so, you will need to look through the list of variables (**GSS Variables and Descriptions.pdf**). Once you've selected two variables, locate them in the codebook (**GSS\_Codebook.pdf**) to figure out how those variables are coded.

## What you will turn in:

- A brief paragraph describing:
  - A sentence or two about the data set you're using for these variables.
  - For each variable, a sentence or two describing the variable name and its description.

- A table with that lists the five univariate statistics for each variable (e.g. Mean, Standard Deviation, Median, Minimum, and Maximum)
- The text of your R code/script on a separate page.

# Example Homework 1

## 1: Data Set Description

The data for this assignment come from the General Social Survey administered in 2014. The data set has 2,538 observations, with individuals as the unit of analysis.

## 2: Variable Descriptions

The two outcome variables I selected were **POLVIEWS** and **OPPSEGOV**. The variable **POLVIEWS** is a continuous measure of conservatism that ranges from 1 to 7, with higher scores on this variable indicating higher levels of conservatism. The text of the variable is as follows:

“We hear a lot of talk these days about liberals and conservatives. I’m going to show you a seven-point scale on which the political views that people might hold are arranged from extremely liberal–point 1–to extremely conservative– point 7. Where would you place yourself on this scale?”

The variable **OPPSEGOV** is a continuous measure on the importance of civil disobedience that ranges from 1 to 7, with higher scores indicating that the respondent believes civil disobedience is more important. The text of the variable is as follows:

“That citizens may engage in acts of civil disobedience when they oppose government actions.”

## 3: Univariate Statistics Table

<i>Variable</i>	<i>Mean</i>	<i>SD</i>	<i>Median</i>	<i>Minimum</i>	<i>Maximum</i>
<i>POLVIEWS</i>	4.089016	1.43409	4	1	7
<i>OPPSEGOV</i>	4.250441	2.087003	4	1	7

#### 4: R Code/Script for Homework 1

```
setwd("/Users/burrelvannjr/Dropbox/methods_csuf")

library(psych)
library(Hmisc)
library(RCurl)

DATA1<-read.csv("resources/data/GSS2014_cleaned_nm.csv",header=TRUE,sep=",")
source("resources/data/LABELS_script.R")

DATA1$polviews
DATA1$oppsegov

mean(DATA1$polviews, na.rm=TRUE)
sd(DATA1$polviews, na.rm=TRUE)
min(DATA1$polviews, na.rm=TRUE)
max(DATA1$polviews, na.rm=TRUE)
median(DATA1$polviews, na.rm=TRUE)

mean(DATA1$oppsegov, na.rm=TRUE)
sd(DATA1$oppsegov, na.rm=TRUE)
median(DATA1$oppsegov, na.rm=TRUE)
min(DATA1$oppsegov, na.rm=TRUE)
max(DATA1$oppsegov, na.rm=TRUE)
```