***Name: Aaditi Rokade******Date: 09/15/2018***

***R Lab 1: Measures of Central Tendency***

***Measures of Central Tendency: (2006 Data)***

Please answer the following questions: (*These questions are based on the variable number of hours worked (‘hrs1).* **Please copy and paste your R output in the appropriate section.**

1. *What is the modal number of hours worked (‘hrs1’) for all respondents?  
     
   > Mode <- function(x) {*  
   *+ ux <- na.omit(unique(x) )*  
   *+ tab <- tabulate(match(x, ux)); ux[tab == max(tab) ]*  
   *+ }*  
   *> mode\_hrs\_worked <- Mode(GSS2006$hrs1) %>%*  
   *+ print()*  
   *[1] "40"*  
     
   *Modal number of hours worked for all respondents: 40*
2. *What is the median number of hours worked for all respondents?  
     
   > mean\_hrs\_worked <- summary(as.numeric(GSS2006$hrs1, na.rm=TRUE)) %>%*  
   *+ print()*  
   *Min. 1st Qu. Median Mean 3rd Qu. Max. NA's*   
   *1.00 38.00 40.00 41.73 50.00 88.00 1791*   
     
   *Median number of hours worked for all respondents: 40*
3. *What is the mean number of hours worked for all respondents?  
     
   > mean\_hrs\_worked <- summary(as.numeric(GSS2006$hrs1, na.rm=TRUE)) %>%*  
   *+ print()*  
   *Min. 1st Qu. Median Mean 3rd Qu. Max. NA's*   
   *1.00 38.00 40.00 41.73 50.00 88.00 1791*   
     
   *Mean number of hours worked for all respondents: 41.73*

*4. What is the mode, median, and mean for each of the following variables? If a measure does not apply, indicate so by writing NA in the space provided. Be sure to write numeric responses only for interval/ratio variables and text responses for nominal and ordinal variables.*

**Variable**

SPEDUC   
  
*> speduc\_sm <- summary(GSS2006$speduc) %>%*

*+ print()*

*Min. 1st Qu. Median Mean 3rd Qu. Max. NA's*

*0.00 12.00 13.00 13.41 16.00 20.00 3111*

*> mode\_speduc <- Mode(GSS2006$speduc) %>%*

*+ print()*

*[1] 12*  
  
*Mode for SPEDUC: 12*  
*Median for SPEDUC: 13.00*  
*Mean for SPEDUC: 13.41*  
  
  
  
HLTHCARE  
  
*> hlthcare\_sm <- summary(GSS2006$hlthcare) %>%*

*+ print()*

*Min. 1st Qu. Median Mean 3rd Qu. Max. NA's*

*1.000 1.000 1.000 1.563 2.000 4.000 3004*

*>*

*> # mode*

*> hlthcare\_mode <- Mode(GSS2006$hlthcare) %>%*

*+ print()*

*[1] 1*  
  
*Mode for HLTHCARE: 1  
Median for HLTHCARE: 1.000  
Mean for HLTHCARE: 1.563*

HIVTEST  
  
*> hivtest\_sm <- summary(GSS2006$hivtest) %>%*

*+ print()*

*Min. 1st Qu. Median Mean 3rd Qu. Max. NA's*

*1.000 1.000 2.000 1.607 2.000 2.000 2118*

*>*

*>*

*> # mode*

*> hivtest\_mode <- Mode(GSS2006$hivtest) %>%*

*+ print()*

*[1] 2*  
  
Mode for HIVTEST: 2  
Median for HIVTEST: 2.000  
Mean for HIVTEST: 1.607

*5. Split the data by sex in order to answer the following questions. Be sure to*

*indicate the values for each group.*

**Question: Males vs Females**

1. What is the mean years of educ (‘educ’)?   
     
   *> female\_educ\_mean <- mean(female\_data$educ, na.rm = TRUE) %>%*  
   *+ print()*  
   *[1] 13.2487*  
   *>*   
   *> male\_educ\_mean <- mean(male\_data$educ, na.rm = TRUE) %>%*  
   *+ print()*  
   *[1] 13.34935*  
     
   *Mean years of educ for females: 13.2487*  
   *Mean years of educ for males: 13.34935*
2. What is the modal response for ‘abpoor’?  
     
   *> female\_abpoor\_mode <- Mode(female\_data$abpoor) %>%*  
   *+ print()*  
   *[1] 2*  
   *>*   
   *> male\_abpoor\_mode <- Mode(male\_data$abpoor) %>%*  
   *+ print()*  
   *[1] 2*  
     
   *Modal response for abpoor for females: 2*  
   *Modal response for abpoor for males: 2*
3. What is the median response for mother’s highest degree (‘madeg’)?   
     
   *> female\_madeg\_mid <- summary(female\_data$madeg) %>%*  
   *+ print()*  
   *Min. 1st Qu. Median Mean 3rd Qu. Max. NA's*   
   *1.000 1.000 2.000 2.006 2.000 5.000 972*   
   *>*   
   *> male\_madeg\_mid <- summary(male\_data$madeg) %>%*  
   *+ print()*  
    *Min. 1st Qu. Median Mean 3rd Qu. Max. NA's*   
   *1.000 1.000 2.000 2.083 2.000 5.000 823*   
     
   *Median response for mother’s highest degree: 2*

***Measures of Central Tendency: (2016 Data)***

Please answer the following questions: (*These questions are based on the variable number of hours worked (‘hrs1).* **Please copy and paste your R output in the appropriate section.**

1. *What is the modal number of hours worked (‘hrs1’) for all respondents?  
     
   > Mode <- function(x)   
   + ux <- na.omit(unique(x))  
   + tab <- tabulate(match(x, ux)); ux[tab == max(tab) ]  
   + }  
   > mode\_hrs\_worked <- Mode(GSS2016$hrs1)  
   > print(mode\_hrs\_worked)  
   [1] "40"*

*Modal number of hours worked for all respondents: 40*

1. *What is the median number of hours worked for all respondents?  
     
   > mean\_hrs\_worked <- summary(as.numeric(GSS2016$hrs1, na.rm=TRUE))  
   > print(mean\_hrs\_worked)  
    Min. 1st Qu. Median Mean 3rd Qu. Max. NA's   
   1.00 35.00 40.00 40.62 48.00 86.00 1231   
     
   Median number of hours worked for all respondents: 40*
2. *What is the mean number of hours worked for all respondents?  
     
   > mean\_hrs\_worked <- summary(as.numeric(GSS2016$hrs1, na.rm=TRUE))  
   > print(mean\_hrs\_worked)  
    Min. 1st Qu. Median Mean 3rd Qu. Max. NA's   
   1.00 35.00 40.00 40.62 48.00 86.00 1231   
     
   Mean number of hours worked for all respondents: 40.62*

*4. What is the mode, median, and mean for each of the following variables? If a*

*measure does not apply, indicate so by writing NA in the space provided. Be sure to write numeric responses only for interval/ratio variables and text responses for nominal and ordinal variables.*

**Variable**

SPEDUC Mode [1] **12***> speduc\_sm <- summary(GSS2016$speduc)*

*> print(speduc\_sm)*

*Min. 1st Qu. Median Mean 3rd Qu. Max. NA's*

*0.00 12.00 14.00 13.89 16.00 20.00 672**> mode\_speduc <- Mode(GSS2016$speduc)*

*> print(mode\_speduc)*

*[1] 12*

*Mode for SPEDUC: 12  
Median for SPEDUC: 14.00  
Mean for SPEDUC: 13.89*

HLTHCARE Mode [1] **1***hlthcare\_sm <- summary(GSS2016$hlthcare) %>%*

*+ print()*

*Min. 1st Qu. Median Mean 3rd Qu. Max. NA's*

*1.000 1.000 2.000 1.682 2.000 4.000 1500   
  
> hlthcare\_mode <- Mode(GSS2016$hlthcare) %>%*

*+ print()*

*[1] 1  
  
Mode for HLTHCARE: 1  
Median for HLTHCARE: 2.000  
Mean for HLTHCARE: 1.682*

HIVTEST Mode [1] **2***> hivtest\_sm <- summary(GSS2016$hivtest)*

*> print(hivtest\_sm)*

*Min. 1st Qu. Median Mean 3rd Qu. Max. NA's*

*1.000 1.000 2.000 1.591 2.000 2.000 1119*

*> hivtest\_mode <- Mode(GSS2016$hivtest)*

*> print(hivtest\_mode)*

*[1] 2*  
  
Mode for HIVTEST: 2  
Median for HIVTEST: 2.000  
Mean for HIVTEST: 1.591

*5. Split the data by sex in order to answer the following questions. Be sure to*

*indicate the values for each group.*

**Question: Males vs Females**

1. What is the mean years of educ (‘educ’)?   
     
   *> female\_educ\_mean <- mean(female\_data$educ, na.rm = TRUE)*  
   *> print(female\_educ\_mean)*  
   *[1] 13.73723*  
   *>*   
   *> male\_educ\_mean <- mean(male\_data$educ, na.rm = TRUE)*  
   *> print(male\_educ\_mean)*  
   *[1] 13.73723*  
     
   *Mean years of educ for females: 13.73723*  
   *Mean years of educ for males: 13.73723*
2. What is the modal response for ‘abpoor’?   
     
   *> female\_abpoor\_mode <- Mode(female\_data$abpoor)*  
   *> print(female\_abpoor\_mode)*  
   *[1] 2*  
   *> male\_abpoor\_mode <- Mode(male\_data$abpoor)*  
   *> print(male\_abpoor\_mode)*  
   *[1] 2*  
     
   *Modal response for abpoor for females: 2*  
   *Modal response for abpoor for males: 2*
3. What is the median response for mother’s highest degree (‘madeg’)?   
     
   *> female\_madeg\_mid <- summary(female\_data$madeg) %>%*  
   *+ print()*  
     
   *Min. 1st Qu. Median Mean 3rd Qu. Max. NA's*   
   *1.000 1.000 2.000 2.161 2.000 5.000 200*   
     
   *> male\_madeg\_mid <- summary(male\_data$madeg) %>%*  
   *+ print()*  
   *Min. 1st Qu. Median Mean 3rd Qu. Max. NA's*   
   *1.000 1.000 2.000 2.161 2.000 5.000 200*   
     
   *Median response for mother’s highest degree: 2*

**Comparisons and Observations - 2006 vs 2016**   
  
1. (2006) Mean number of hours worked for all respondents: 41.73  
 (2016) Mean number of hours worked for all respondents: 40.62  
  
 Hence, the mean has decreased by 1.73 in 2016.  
  
2. (2006) Median for SPEDUC: 13.00 & Mean for SPEDUC: 13.41  
 (2016) Median for SPEDUC: 14.00 & Mean for SPEDUC: 13.89   
  
Median of ‘HIGHEST YEAR SCHOOL COMPLETED, SPOUSE’ increased by 1 and Mean increased by 0.48 in 2016.  
  
3. (2006) Median for HLTHCARE: 1.00 & Mean for HLTHCARE: 1.563   
 (2016) Median for HLTHCARE: 2.00 & Mean for HLTHCARE: 1.682  
  
Median of ‘GOVTS RESP: PROVIDE HLTH CARE FOR SICK’ increased by 1 and Mean increased by 0.119 in 2016.  
  
4. (2006) Mean for HIVTEST: 1.607  
 (2016) Mean for HIVTEST: 1.591   
  
Mean of ‘ HAVE YOU EVER BEEN TESTED FOR HIV’ decreased by 0.016 in 2016.  
  
5. Mean years of ‘HIGHEST YEAR SCHOOL COMPLETED’ for females is increased in 2016 by 0.48853 whereas for males, it’s increased by 0.38788