ANA600 Final Exam

{Randall Crawford}

2024-10-27

INSTRUCTIONS

Perform basic exploratory data analysis (EDA) for the final exam. EDA consists of the procedures and concepts we have practiced throughout the course. You will be exploring your dataset, reviewing the variables, modeling an variable of interest, and interpreting results. Each section requires a complete paragraph explaining and interpreting the results within the R-Markdown above the code block required to analyze that section. All writing and code should be written as if presenting a report to your supervisor.

PREPARATION (10 Points)

Import the required file to a new dataframe and load necessary libraries

---

INTRODUCTION AND RESEARCH QUESTION (5 Points)

1. Introduce your purpose and scope, creating a story for the data generation process that might be responsible for the variation in income output variable.
2. Describe the research question, which is to examine income based on the sex variable.

{When considering the potential factors that influence an individual's income, it seems logical that many factors could contribute to it, and not just Sex. Although, recent focus on inequality in compensation between sexes may drive this type of investigation. Per your request, it is my task to create a linear model that estimates income by sex. In trying to better understand the relationship of the outcome variable Income, and the explanatory variable Sex, a word equation should be put together to help generalize what we are after with creating a linear model.}

```
{Income = B0 + B1(Sex) + error}
```

QUESTION #1 (10 Points)

1. Enter code to produce the structure of your dataframe.
2. Recode the Sex variable to 0=Female, 1=Male.
3. Produce a crosstab table of observations for the race and sex variables.
4. View the top five records in the dataframe.
5. Write one paragraph describing the structure of the data frame and interpreting the produced table

{Regarding the structure of the dataset, there are 1000 observations and 10 variables. Age, Income, and HoursWk are quantitative variables. Sex, Married, USCitizen, HealthInsurance, and Language are binary categorical variables. Race and the recently added Sex_str are string categorical variables. Regarding the tabled observations, there are 53.7% female and 46.3% male, and 76.1% are of the white race.}

```
#Enter code to produce the structure of your dataframe.
str(ACS_data)

## 'data.frame':    1000 obs. of  9 variables:
##  $ Sex            : int  0 1 1 0 1 1 1 1 0 0 ...
##  $ Age            : int  31 31 75 80 64 14 78 35 70 18 ...
##  $ Married        : int  0 0 0 0 1 0 1 0 1 0 ...
##  $ Income         : num  60 0.36 0 0 0 ...
##  $ HoursWk        : num  40 12 40 13.2 32.7 ...
##  $ Race           : chr  "white" "black" "white" "white" ...
##  $ USCitizen      : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ HealthInsurance: int  1 1 1 1 1 1 1 1 1 1 ...
##  $ Language       : int  1 0 0 0 0 0 0 1 0 0 ...

#Recode the Sex variable to 0=Female, 1=Male.
ACS_data$Sex <- as.character(factor(ACS_data$Sex, levels = c(0,1), labels =
c("Female", "Male")))
str(ACS_data)

## 'data.frame':    1000 obs. of  9 variables:
##  $ Sex            : chr  "Female" "Male" "Male" "Female" ...
##  $ Age            : int  31 31 75 80 64 14 78 35 70 18 ...
##  $ Married        : int  0 0 0 0 1 0 1 0 1 0 ...
##  $ Income         : num  60 0.36 0 0 0 ...
##  $ HoursWk        : num  40 12 40 13.2 32.7 ...
##  $ Race           : chr  "white" "black" "white" "white" ...
##  $ USCitizen      : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ HealthInsurance: int  1 1 1 1 1 1 1 1 1 1 ...
##  $ Language       : int  1 0 0 0 0 0 0 1 0 0 ...

#Produce a crosstab table of observations for the race and sex variables.
xtabs(~ Race + Sex, data=ACS_data)

##        Sex
## Race     Female Male
```

```
##    asian      30   40
##    black      60   46
##    other      33   30
##    white     414  347
```

```
#View the top five records in the dataframe.
head(ACS_data, 5)
```

```
##       Sex Age Married Income  HoursWk  Race USCitizen HealthInsurance
Language
## 1 Female  31       0  60.00 40.00000 white         1               1
1
## 2   Male  31       0   0.36 12.00000 black         1               1
0
## 3   Male  75       0   0.00 39.99126 white         1               1
0
## 4 Female  80       0   0.00 13.15004 white         1               1
0
## 5   Male  64       1   0.00 32.71688 white         1               1
0
```

QUESTION #2 (10 Points)

1. Recode the income variable to value x 1,000
2. Calculate the minimum, maximum, mean, median, IQR, and range for income
3. Calculate the mean of income each for males and for females
4. Write one paragraph explaining and interpreting the descriptive statistics

{At least 25% or 250 of the observations have $0 income. This brings the mean and median incomes down considerably, which currently reside at $22,785.13 and $13,000 respectively for 1000 observations. The maximum income is $563,000, which is far removed from the vast majority of observations. The $28,573.61 male mean income was $10,779.29 higher than the $17,794.32 female mean income.}

```
#1. Recode the income variable to value x 1,000
ACS_data$Income <- ACS_data$Income * 1000
str(ACS_data)
```

```
## 'data.frame':    1000 obs. of  9 variables:
##  $ Sex            : chr  "Female" "Male" "Male" "Female" ...
##  $ Age            : int  31 31 75 80 64 14 78 35 70 18 ...
##  $ Married        : int  0 0 0 0 1 0 1 0 1 0 ...
##  $ Income         : num  60000 360 0 0 0 ...
##  $ HoursWk        : num  40 12 40 13.2 32.7 ...
##  $ Race           : chr  "white" "black" "white" "white" ...
##  $ USCitizen      : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ HealthInsurance: int  1 1 1 1 1 1 1 1 1 1 ...
##  $ Language       : int  1 0 0 0 0 0 0 1 0 0 ...
```

```
#2. Calculate the minimum, maximum, mean, median, IQR, and range for income
Income.stats = favstats(ACS_data$Income)
Income.stats

##  min Q1 median       Q3    max     mean       sd    n missing
##    0  0  13000 31841.09 563000 22785.13 39141.88 1000       0

minimum = Income.stats[1,1]
cat("Minimum =", dollar(minimum), '\n')

## Minimum = $0

maximum = Income.stats[1,5]
cat("Maximum =", dollar(maximum), '\n')

## Maximum = $563,000

mean = Income.stats[1,6]
cat("Mean =", dollar(mean), '\n')

## Mean = $22,785.13

median = Income.stats[1,3]
cat("Median =", dollar(median), '\n')

## Median = $13,000

IQR = Income.stats[1,4] - Income.stats[1,2]
cat("IQR =", dollar(IQR), '\n')

## IQR = $31,841.09

range = Income.stats[1,5] - Income.stats[1,1]
cat("Range =", dollar(range))

## Range = $563,000

#3. Calculate the mean of income each for males and for females.
Income_Sex.stats = favstats(Income ~ Sex, data = ACS_data)
Income_Sex.stats

##      Sex min Q1   median       Q3    max     mean       sd   n missing
## 1 Female   0  0  9000.00 27810.91 382000 17794.32 27990.08 537       0
## 2   Male   0  0 16260.22 37241.34 563000 28573.61 48388.30 463       0

female_mean = Income_Sex.stats[1,7]
cat("Female Mean Income =", dollar(female_mean), '\n')

## Female Mean Income = $17,794.32

male_mean = Income_Sex.stats[2,7]
cat("Male Mean Income =", dollar(male_mean), '\n')

## Male Mean Income = $28,573.61
```

QUESTION 3 (10 Points)

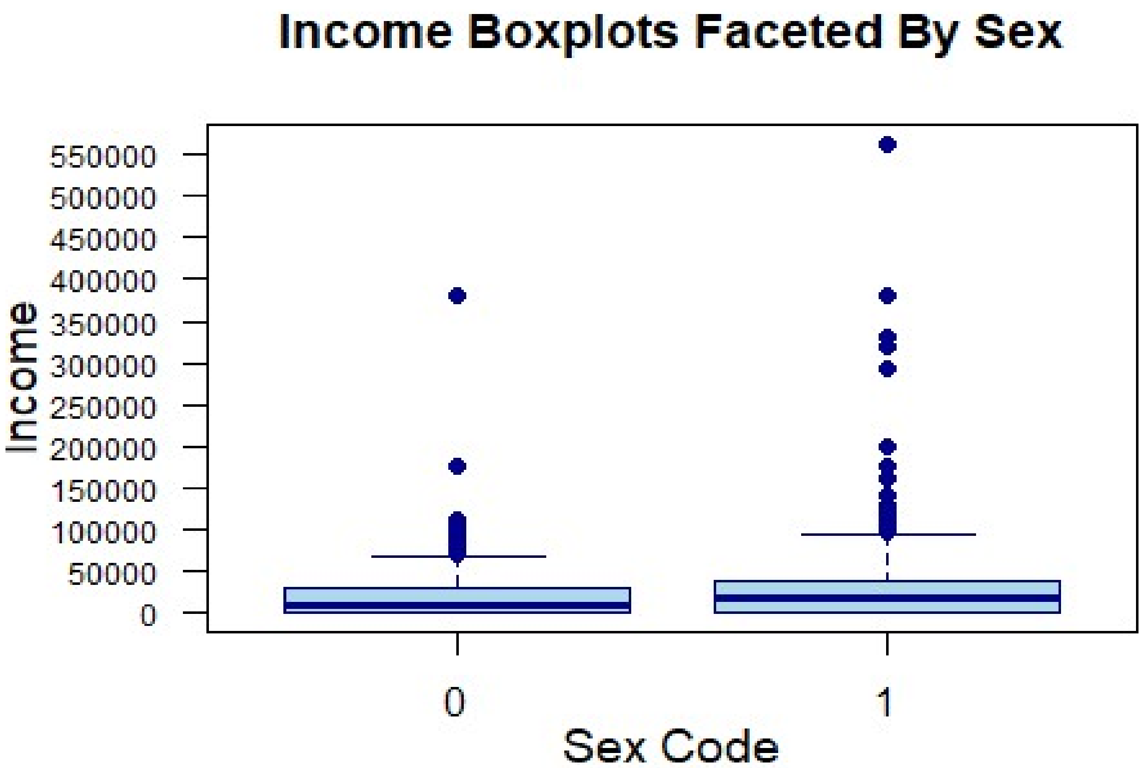
1. Create a new variable SexL such that Sex = 0 is "Female", Sex = 1 is "Male", and else is "Undefined"
2. Create an appropriate visualization for income and the new SexL variable
3. Write one paragraph explaining and interpreting the visualization

{The boxplot visualization for Income faceted by Sex reaffirms much of what was seen in our descriptive statistics. Females(0) account for the majority of $0 income observations, though there are a notable amount of Males(1) in the same predicament. Each light blue IQR box with a dark blue median line rests on $0 income. Both sexes have a number of outliers depicted with dark blue dots, mainly due to the shift of each income set toward $0 income. Males(1) seem to possess some higher income levels though.}

```r
#1. Create a new variable SexL such that Sex = 0 is "Female", Sex = 1 is
"Male", and else is "Undefined".
ACS_data$SexL <- as.integer(recode(ACS_data$Sex, "Female" = 0, "Male" = 1))
str(ACS_data)
```

```
## 'data.frame':    1000 obs. of  10 variables:
##  $ Sex            : chr  "Female" "Male" "Male" "Female" ...
##  $ Age            : int  31 31 75 80 64 14 78 35 70 18 ...
##  $ Married        : int  0 0 0 0 1 0 1 0 1 0 ...
##  $ Income         : num  60000 360 0 0 0 ...
##  $ HoursWk        : num  40 12 40 13.2 32.7 ...
##  $ Race           : chr  "white" "black" "white" "white" ...
##  $ USCitizen      : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ HealthInsurance: int  1 1 1 1 1 1 1 1 1 1 ...
##  $ Language       : int  1 0 0 0 0 0 0 1 0 0 ...
##  $ SexL           : int  0 1 1 0 1 1 1 1 0 0 ...
```

```r
#2. Create an appropriate visualization for income and the new SexL variable.
boxplot(Income ~ SexL, data = ACS_data, col = "lightblue", border =
"darkblue", outpch = 21, outbg = "darkblue", linewidth = 0.25, main = "Income
Boxplots Faceted By Sex", ylab="", xlab = "", yaxt = "n")
axis(2, at = seq(0, 600000, 50000), las = 2, cex.axis=0.75)
mtext("Sex Code", side=1, line=2, cex=1.2)
mtext("Income", side=2, line=3.3, cex=1.2)
```

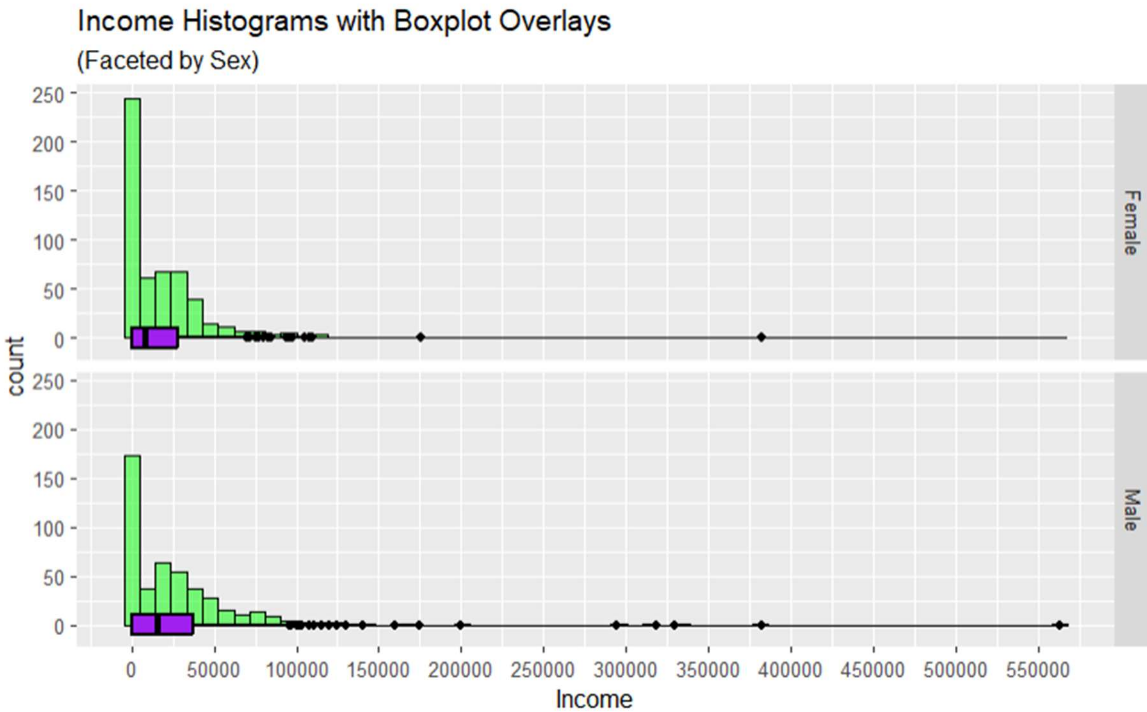


QUESTION 4 (10 Points)

1. Create a histogram of income by sex with facet grid
2. Write one paragraph explaining and interpreting the visualization

{Much of what was surmised in the boxplots can be seen in these histograms, especially the exceedingly large number of low income observations. Both histograms are unimodal with an apex at the mode, $0 income, which is less than the median that is in turn less than the mean, indicating a pronounce right skew. Some level of kurtosis is present with a significant number of outliers, more so in the Male histogram. And again, males seem to possess some higher income levels.}

```
#1. Create a histogram of income by sex with facet grid
gf_histogram(~ Income, data = ACS_data, fill = "green", color = "black",
linewidth = 0.5, bins = 60, title = "Income Histograms with Boxplot
Overlays", subtitle = "(Faceted by Sex)") %>%
  gf_facet_grid(Sex ~ .) %>%
  gf_boxplot(fill = "purple", width = 20, color = "black", linewidth = 0.75)
+ scale_x_continuous("Income", breaks = breaks_pretty(n=12))
```



QUESTION 5 (10 Points)

1. Create a model of income for females
2. Write one paragraph explaining and interpreting the model

{An empty model was created using income from a female filtered data set. Standard Equation Used: Y = b0 + e, with Y = the estimated female income, b0 = $17,794 (the mean female income coefficient), and e = error. After an ANOVA evaluation, the Sum of Squared Error was 4.1993e+11. This value is large due to the range of income values in the data set and the number of $0 income observations, but is the smallest possible outcome given that mean is being used.}

```
#1. Create a model of income for females.
ACS_data_female <- filter(ACS_data, SexL == 0)
str(ACS_data_female)

## 'data.frame':    537 obs. of  10 variables:
##  $ Sex            : chr  "Female" "Female" "Female" "Female" ...
##  $ Age            : int  31 80 70 18 61 52 36 20 25 48 ...
##  $ Married        : int  0 0 1 0 1 0 1 0 1 1 ...
##  $ Income         : num  60000 0 0 0 0 48000 2000 500 0 22000 ...
##  $ HoursWk        : num  40 13.2 1 29.8 20 ...
##  $ Race           : chr  "white" "white" "white" "white" ...
##  $ USCitizen      : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ HealthInsurance: int  1 1 1 1 0 1 1 1 1 1 ...
##  $ Language       : int  1 0 0 0 0 0 0 0 0 0 ...
##  $ SexL           : int  0 0 0 0 0 0 0 0 0 0 ...
```

```r
IncFem_model <- lm(Income ~ NULL, data = ACS_data_female)
IncFem_model

##
## Call:
## lm(formula = Income ~ NULL, data = ACS_data_female)
##
## Coefficients:
## (Intercept)
##       17794

anova(IncFem_model)

## Analysis of Variance Table
##
## Response: Income
##            Df     Sum Sq   Mean Sq F value Pr(>F)
## Residuals 536 4.1993e+11 783444330
```

QUESTION 6 (20 Points)

1. Create a model of income by sex
2. Write one paragraph explaining and interpreting the model

{A linear model was created using income as an outcome variable and the binary SexL as an explanatory variable. Standard Equation Used: Y = b0 + b1(X1) + e, with Y = the estimated income, b0 = $17,794 (the mean female income coefficient), b1 = $10,779 (difference between the mean male income and the mean female income), X1 = 0 (representing a female income observation) or 1 (representing a male income observation), and e = error.

```r
IncSex_model <- lm(Income ~ SexL, data = ACS_data)
IncSex_model

##
## Call:
## lm(formula = Income ~ SexL, data = ACS_data)
##
## Coefficients:
## (Intercept)         SexL
##       17794        10779
```

QUESTION 7 (20 Points)

1. Calculate the predicted value and residual value for each observation using the income by sex model
2. Calculate the sum of squared deviations and sum of absolute deviations
3. Write one paragraph explaining and interpreting the results

{As expected, the prediction values are 17794.32, mean of female income, and 28573.61, mean of male income. Concerning the residual values, the most common residuals were -17794.32 and -28573.61, given the number of $0 income observations. The Sum of Squared Deviations calculation equaled 1.501666e+12, and the Sum of Absolute Deviations calculation equaled 22270304. These values are large due to the range of income values in the data set and the number of $0 income observations, but are the smallest possible outcomes given that means are being used.}

```
#1. Calculate the predicted value and residual value for each observation
using the income by sex model.
ACS_data$prediction <- predict(IncSex_model)
ACS_data$prediction

##    [1] 17794.32 28573.61 28573.61 17794.32 28573.61 28573.61 28573.61
28573.61
##    [9] 17794.32 17794.32 28573.61 17794.32 17794.32 17794.32 17794.32
28573.61
##   [17] 28573.61 28573.61 28573.61 17794.32 17794.32 28573.61 17794.32
17794.32
##   [25] 28573.61 17794.32 28573.61 17794.32 17794.32 17794.32 28573.61
17794.32
##   [33] 17794.32 28573.61 28573.61 28573.61 17794.32 28573.61 28573.61
17794.32
##   [41] 28573.61 28573.61 28573.61 28573.61 17794.32 28573.61 28573.61
17794.32
##   [49] 17794.32 17794.32 28573.61 28573.61 28573.61 28573.61 28573.61
28573.61
##   [57] 17794.32 28573.61 17794.32 17794.32 28573.61 28573.61 17794.32
17794.32
##   [65] 17794.32 28573.61 17794.32 17794.32 28573.61 28573.61 28573.61
28573.61
##   [73] 17794.32 28573.61 28573.61 17794.32 17794.32 17794.32 17794.32
17794.32
##   [81] 17794.32 28573.61 28573.61 28573.61 17794.32 28573.61 28573.61
28573.61
##   [89] 28573.61 28573.61 17794.32 28573.61 17794.32 28573.61 17794.32
17794.32
##   [97] 28573.61 17794.32 17794.32 17794.32 28573.61 17794.32 17794.32
17794.32
##  [105] 17794.32 17794.32 17794.32 28573.61 17794.32 28573.61 28573.61
17794.32
##  [113] 17794.32 28573.61 17794.32 17794.32 17794.32 17794.32 17794.32
28573.61
##  [121] 28573.61 28573.61 28573.61 28573.61 28573.61 17794.32 17794.32
17794.32
##  [129] 28573.61 17794.32 28573.61 17794.32 17794.32 17794.32 28573.61
28573.61
##  [137] 28573.61 17794.32 17794.32 17794.32 17794.32 17794.32 17794.32
17794.32
```

```
## [145] 28573.61 28573.61 17794.32 28573.61 17794.32 28573.61 28573.61
28573.61
## [153] 17794.32 28573.61 17794.32 17794.32 17794.32 28573.61 17794.32
17794.32
## [161] 17794.32 17794.32 28573.61 28573.61 17794.32 28573.61 28573.61
17794.32
## [169] 28573.61 17794.32 28573.61 28573.61 17794.32 17794.32 28573.61
17794.32
## [177] 17794.32 17794.32 17794.32 28573.61 28573.61 28573.61 28573.61
28573.61
## [185] 28573.61 17794.32 28573.61 28573.61 28573.61 28573.61 17794.32
28573.61
## [193] 17794.32 28573.61 28573.61 17794.32 28573.61 28573.61 17794.32
28573.61
## [201] 17794.32 17794.32 17794.32 17794.32 17794.32 17794.32 17794.32
17794.32
## [209] 28573.61 17794.32 17794.32 17794.32 17794.32 17794.32 17794.32
28573.61
## [217] 17794.32 17794.32 17794.32 17794.32 28573.61 17794.32 17794.32
28573.61
## [225] 17794.32 17794.32 28573.61 17794.32 28573.61 17794.32 17794.32
28573.61
## [233] 17794.32 28573.61 17794.32 28573.61 17794.32 17794.32 17794.32
28573.61
## [241] 17794.32 28573.61 28573.61 28573.61 28573.61 28573.61 17794.32
17794.32
## [249] 17794.32 28573.61 28573.61 17794.32 17794.32 28573.61 28573.61
28573.61
## [257] 17794.32 17794.32 28573.61 17794.32 28573.61 17794.32 17794.32
28573.61
## [265] 28573.61 17794.32 17794.32 28573.61 17794.32 17794.32 28573.61
17794.32
## [273] 28573.61 28573.61 28573.61 17794.32 28573.61 28573.61 17794.32
28573.61
## [281] 28573.61 17794.32 17794.32 17794.32 17794.32 28573.61 28573.61
28573.61
## [289] 17794.32 28573.61 28573.61 17794.32 28573.61 28573.61 17794.32
28573.61
## [297] 28573.61 28573.61 17794.32 28573.61 28573.61 17794.32 28573.61
17794.32
## [305] 17794.32 17794.32 17794.32 28573.61 28573.61 17794.32 28573.61
28573.61
## [313] 28573.61 28573.61 17794.32 17794.32 28573.61 28573.61 28573.61
17794.32
## [321] 17794.32 28573.61 28573.61 28573.61 17794.32 28573.61 28573.61
28573.61
## [329] 17794.32 17794.32 17794.32 17794.32 28573.61 17794.32 28573.61
17794.32
## [337] 17794.32 28573.61 17794.32 28573.61 17794.32 28573.61 17794.32
28573.61
```

```
## [345] 17794.32 28573.61 17794.32 28573.61 28573.61 28573.61 28573.61
28573.61
## [353] 17794.32 28573.61 28573.61 28573.61 28573.61 28573.61 17794.32
17794.32
## [361] 17794.32 17794.32 17794.32 28573.61 28573.61 17794.32 17794.32
17794.32
## [369] 28573.61 28573.61 28573.61 17794.32 17794.32 17794.32 17794.32
28573.61
## [377] 17794.32 28573.61 28573.61 17794.32 17794.32 17794.32 17794.32
28573.61
## [385] 17794.32 17794.32 28573.61 17794.32 17794.32 17794.32 17794.32
17794.32
## [393] 17794.32 28573.61 28573.61 28573.61 28573.61 17794.32 28573.61
17794.32
## [401] 28573.61 17794.32 17794.32 17794.32 17794.32 17794.32 17794.32
17794.32
## [409] 28573.61 17794.32 17794.32 28573.61 17794.32 28573.61 17794.32
17794.32
## [417] 28573.61 17794.32 28573.61 17794.32 28573.61 28573.61 28573.61
17794.32
## [425] 28573.61 28573.61 17794.32 17794.32 28573.61 28573.61 17794.32
28573.61
## [433] 28573.61 17794.32 28573.61 28573.61 28573.61 28573.61 28573.61
28573.61
## [441] 17794.32 28573.61 17794.32 17794.32 17794.32 17794.32 17794.32
17794.32
## [449] 28573.61 28573.61 17794.32 17794.32 28573.61 17794.32 17794.32
17794.32
## [457] 28573.61 17794.32 17794.32 28573.61 17794.32 28573.61 17794.32
17794.32
## [465] 28573.61 28573.61 17794.32 17794.32 17794.32 28573.61 17794.32
28573.61
## [473] 17794.32 17794.32 17794.32 17794.32 17794.32 17794.32 28573.61
17794.32
## [481] 17794.32 28573.61 17794.32 17794.32 28573.61 28573.61 28573.61
17794.32
## [489] 28573.61 17794.32 28573.61 28573.61 28573.61 17794.32 17794.32
17794.32
## [497] 17794.32 17794.32 17794.32 28573.61 17794.32 28573.61 28573.61
28573.61
## [505] 28573.61 17794.32 17794.32 28573.61 28573.61 17794.32 17794.32
28573.61
## [513] 17794.32 28573.61 28573.61 28573.61 17794.32 17794.32 28573.61
28573.61
## [521] 17794.32 28573.61 28573.61 28573.61 28573.61 17794.32 28573.61
17794.32
## [529] 17794.32 17794.32 17794.32 28573.61 17794.32 17794.32 28573.61
28573.61
## [537] 17794.32 28573.61 17794.32 17794.32 17794.32 17794.32 28573.61
28573.61
```

```
## [545] 28573.61 28573.61 17794.32 28573.61 17794.32 28573.61 17794.32
28573.61
## [553] 28573.61 28573.61 28573.61 28573.61 28573.61 28573.61 28573.61
28573.61
## [561] 28573.61 17794.32 28573.61 28573.61 17794.32 17794.32 17794.32
28573.61
## [569] 28573.61 28573.61 17794.32 28573.61 17794.32 28573.61 17794.32
28573.61
## [577] 17794.32 17794.32 17794.32 28573.61 17794.32 28573.61 17794.32
28573.61
## [585] 28573.61 17794.32 17794.32 28573.61 17794.32 17794.32 28573.61
17794.32
## [593] 28573.61 28573.61 17794.32 28573.61 17794.32 28573.61 17794.32
28573.61
## [601] 17794.32 28573.61 17794.32 17794.32 17794.32 28573.61 17794.32
17794.32
## [609] 28573.61 17794.32 17794.32 28573.61 17794.32 17794.32 28573.61
17794.32
## [617] 17794.32 17794.32 17794.32 17794.32 17794.32 28573.61 28573.61
28573.61
## [625] 17794.32 28573.61 17794.32 17794.32 28573.61 28573.61 17794.32
17794.32
## [633] 17794.32 17794.32 17794.32 28573.61 17794.32 17794.32 17794.32
17794.32
## [641] 28573.61 17794.32 28573.61 17794.32 17794.32 28573.61 17794.32
17794.32
## [649] 17794.32 28573.61 28573.61 28573.61 28573.61 28573.61 28573.61
17794.32
## [657] 28573.61 28573.61 17794.32 17794.32 28573.61 28573.61 17794.32
28573.61
## [665] 17794.32 28573.61 17794.32 17794.32 17794.32 28573.61 17794.32
28573.61
## [673] 17794.32 17794.32 28573.61 17794.32 28573.61 17794.32 28573.61
28573.61
## [681] 17794.32 17794.32 17794.32 28573.61 17794.32 17794.32 28573.61
17794.32
## [689] 28573.61 17794.32 17794.32 28573.61 28573.61 17794.32 17794.32
17794.32
## [697] 28573.61 17794.32 17794.32 17794.32 17794.32 17794.32 28573.61
28573.61
## [705] 17794.32 17794.32 28573.61 17794.32 17794.32 28573.61 17794.32
17794.32
## [713] 17794.32 17794.32 28573.61 17794.32 17794.32 17794.32 17794.32
17794.32
## [721] 28573.61 17794.32 28573.61 28573.61 28573.61 17794.32 17794.32
17794.32
## [729] 17794.32 28573.61 28573.61 17794.32 28573.61 17794.32 28573.61
17794.32
## [737] 28573.61 28573.61 28573.61 28573.61 17794.32 17794.32 28573.61
28573.61
```

```
## [745] 28573.61 17794.32 28573.61 17794.32 28573.61 17794.32 28573.61
17794.32
## [753] 28573.61 17794.32 28573.61 28573.61 17794.32 28573.61 17794.32
17794.32
## [761] 28573.61 28573.61 28573.61 17794.32 28573.61 28573.61 17794.32
17794.32
## [769] 28573.61 17794.32 17794.32 17794.32 17794.32 17794.32 28573.61
17794.32
## [777] 28573.61 28573.61 17794.32 28573.61 28573.61 17794.32 28573.61
28573.61
## [785] 17794.32 17794.32 28573.61 17794.32 28573.61 28573.61 17794.32
17794.32
## [793] 28573.61 17794.32 17794.32 28573.61 28573.61 17794.32 17794.32
28573.61
## [801] 17794.32 17794.32 28573.61 17794.32 17794.32 17794.32 17794.32
17794.32
## [809] 17794.32 28573.61 17794.32 28573.61 28573.61 17794.32 17794.32
17794.32
## [817] 17794.32 28573.61 28573.61 28573.61 17794.32 17794.32 28573.61
28573.61
## [825] 17794.32 17794.32 28573.61 17794.32 17794.32 17794.32 28573.61
28573.61
## [833] 17794.32 28573.61 17794.32 17794.32 17794.32 17794.32 28573.61
17794.32
## [841] 17794.32 17794.32 17794.32 17794.32 17794.32 28573.61 17794.32
28573.61
## [849] 28573.61 17794.32 17794.32 28573.61 17794.32 17794.32 28573.61
28573.61
## [857] 17794.32 28573.61 17794.32 28573.61 28573.61 28573.61 28573.61
17794.32
## [865] 17794.32 28573.61 28573.61 17794.32 17794.32 17794.32 17794.32
28573.61
## [873] 28573.61 28573.61 17794.32 28573.61 17794.32 17794.32 17794.32
17794.32
## [881] 28573.61 28573.61 17794.32 28573.61 17794.32 28573.61 17794.32
28573.61
## [889] 17794.32 17794.32 17794.32 28573.61 17794.32 17794.32 17794.32
17794.32
## [897] 17794.32 17794.32 28573.61 17794.32 28573.61 17794.32 28573.61
17794.32
## [905] 17794.32 17794.32 28573.61 17794.32 17794.32 28573.61 28573.61
28573.61
## [913] 17794.32 17794.32 17794.32 17794.32 28573.61 17794.32 28573.61
17794.32
## [921] 17794.32 17794.32 28573.61 28573.61 28573.61 28573.61 28573.61
17794.32
## [929] 17794.32 28573.61 17794.32 28573.61 28573.61 17794.32 17794.32
28573.61
## [937] 28573.61 17794.32 28573.61 17794.32 17794.32 17794.32 17794.32
17794.32
```

```
## [945] 17794.32 17794.32 17794.32 17794.32 28573.61 17794.32 28573.61
28573.61
## [953] 17794.32 28573.61 28573.61 28573.61 28573.61 17794.32 28573.61
28573.61
## [961] 28573.61 17794.32 17794.32 17794.32 28573.61 28573.61 17794.32
17794.32
## [969] 17794.32 28573.61 17794.32 28573.61 17794.32 28573.61 28573.61
28573.61
## [977] 17794.32 28573.61 28573.61 28573.61 17794.32 28573.61 17794.32
28573.61
## [985] 17794.32 28573.61 17794.32 17794.32 28573.61 28573.61 17794.32
28573.61
## [993] 28573.61 17794.32 17794.32 17794.32 17794.32 28573.61 17794.32
17794.32
```

```
ACS_data$residual <- resid(IncSex_model)
ACS_data$residual
```

```
##    [1]  42205.68172 -28213.60814 -28573.60814 -17794.31828 -28573.60814
##    [6]   8641.87690 -28573.60814  58426.39186 -17794.31828 -17794.31828
##   [11]   3426.39186 -17794.31828  30205.68172 -15794.31828 -17294.31828
##   [16] -28573.60814   3026.93177 -28573.60814  -6973.60814 -17794.31828
##   [21]   4205.68172 -15014.31345 -10994.31828 -17794.31828   1426.39186
##   [26] -15294.31828 -27073.60814 -17794.31828 -14794.31828  57205.68172
##   [31] -28573.60814   1205.68172  32205.68172 101426.39186 -28573.60814
##   [36] -28573.60814   6350.02336 -14198.13708 -28573.60814 -17794.31828
##   [41] -28573.60814  18426.39186  11095.78897 -27073.60814  -6394.31828
##   [46] -13573.60814 -15573.60814  13205.68172  -6731.77015 -17794.31828
##   [51] 111426.39186 -28573.60814 -15277.33454 -28573.60814 -27573.60814
##   [56] -28573.60814 -17794.31828 -28573.60814  -7294.31828   -694.31828
##   [61] -28573.60814 -28573.60814  27205.68172 -16694.31828   -794.31828
##   [66] -13473.60814 -17794.31828  30105.68172 -28573.60814  -8257.79072
##   [71]  21426.39186 -27073.60814   7913.09656 -27773.60814 -15573.60814
##   [76]   5205.68172  12405.68172 -17794.31828 -16794.31828 -17794.31828
##   [81]  -7794.31828 -17020.77328  31426.39186  46426.39186 -13994.31828
##   [86]  18426.39186  23426.39186  28426.39186 -28573.60814    326.39186
##   [91]  -2794.31828  43426.39186 -11794.31828   2851.01970   9928.03844
##   [96] -17794.31828 -14573.60814  16599.48016 -17794.31828  20205.68172
##  [101] -28013.60814 -12894.31828  -2794.31828 -13694.31828  32205.68172
##  [106]  34205.68172  19521.75610  56426.39186  11905.68172 -28573.60814
##  [111]  36426.39186 -17794.31828  -5596.39107 -28573.60814  -1794.31828
##  [116] -17794.31828 -17794.31828 -17794.31828 -12294.31828   7426.39186
##  [121] -20773.60814 -28573.60814 -28573.60814 -28573.60814 -28573.60814
##  [126]   -294.31828 -17794.31828  16924.03033 534426.39186 -17794.31828
##  [131]  -7950.37992 -17794.31828  16713.18849  -3794.31828  -3573.60814
##  [136] -28573.60814 -28573.60814 -15794.31828  29205.68172 -17794.31828
##  [141]  15205.68172  14352.46274  16606.06413  76205.68172  68426.39186
##  [146] -28573.60814 -17794.31828  31426.39186  52205.68172  61426.39186
##  [151]  31426.39186  31426.39186   2205.68172  51426.39186 -17794.31828
##  [156] -17794.31828 -17794.31828 -28573.60814    173.99745 -17794.31828
```

```
## [161]  -1794.31828 -17794.31828 -28573.60814   4726.39186 364205.68172
## [166] -28573.60814  61426.39186  15994.89626  11426.39186  27205.68172
## [171] -15613.99784 -28573.60814  37205.68172  52205.68172 -28573.60814
## [176]   6205.68172  17176.79259    626.89375 -17794.31828  11426.39186
## [181]    -55.97382 -16023.02816  27426.39186  56426.39186 -19573.60814
## [186] -17794.31828  53426.39186  33426.39186 -27873.60814 -28573.60814
## [191]  41205.68172  -1881.17946  -9794.31828 -26573.60814 -28573.60814
## [196] -17794.31828  -6473.60814   1426.39186   7205.68172  -8121.87780
## [201] -15794.31828 -17794.31828  13405.68172  33205.68172 -17794.31828
## [206] -17794.31828 -17794.31828  21250.78901 -12573.60814 -17794.31828
## [211] -17794.31828  11105.68172  -4110.76471 -16994.31828  16837.09084
## [216] -12313.39194  87205.68172   9205.68172  -7794.31828 -17794.31828
## [221] -28573.60814 -17794.31828  18205.68172   6634.33059 -17794.31828
## [226] -17794.31828  11426.39186  65205.68172 -24973.60814   1027.99523
## [231]   2205.68172  -4378.53617 -17794.31828  43426.39186 -17794.31828
## [236] -28573.60814 -17394.31828 -17794.31828 -17794.31828 -28573.60814
## [241] -11894.31828 -26573.60814    285.63843  51426.39186 -28573.60814
## [246]  -6173.60814  12276.76034  -8794.31828 -17794.31828 -28573.60814
## [251]  10349.34493 -10094.31828 -17794.31828  -9573.60814 -20573.60814
## [256] -28073.60814   9066.87672  42205.68172 -27673.60814 -17794.31828
## [261]   8426.39186 -17794.31828 -17794.31828  11426.39186  -7525.74833
## [266]  23205.68172  -3394.31828 -28573.60814   1518.07405 -17794.31828
## [271]   1426.39186  -4418.94629 -28573.60814  38426.39186  86426.39186
## [276]  37205.68172 -28573.60814 -23573.60814 -17794.31828  -8373.60814
## [281] -20873.60814 -17794.31828 -17794.31828  -2794.31828 -17794.31828
## [286] -28573.60814 -28573.60814 -13573.60814 -17794.31828 -28573.60814
## [291] -28573.60814 -17794.31828 -28573.60814 -28573.60814  13819.52401
## [296] -20573.60814  23426.39186   2464.64140 -17794.31828  10464.92142
## [301]  31426.39186  39205.68172 -13573.60814 -17794.31828 -16294.31828
## [306] -17794.31828   3205.68172  63426.39186 -28573.60814  13667.49512
## [311]  36426.39186 -13274.73967  -1519.46861   7426.39186 -17794.31828
## [316]  42205.68172    -55.36332  13426.39186  79426.39186 -17794.31828
## [321]  23205.68172 -23573.60814 -28573.60814   8426.39186   6350.32162
## [326] -28573.60814  -5973.30131 -28573.60814  15857.84239 -17794.31828
## [331] -17794.31828 -17794.31828 -28573.60814 -17794.31828  -5565.08242
## [336] -17794.31828 -17794.31828   2673.91162 -17794.31828  -6126.62880
## [341] -17794.31828  31426.39186 -17304.31828  -6973.60814   8242.72975
## [346] -28573.60814  -7620.36334 -18573.60814 -28573.60814 -28573.60814
## [351]  21426.39186 -13182.60334 -15794.31828  27426.39186  51426.39186
## [356]   -573.60814 -28573.60814  17426.39186  -1794.31828 -17794.31828
## [361] -15694.31828 -15894.31828 -17794.31828   -642.72852 -28573.60814
## [366]  28205.68172   6905.68172  -4005.26998 -28573.60814  11426.39186
## [371] -11892.75438 -17794.31828  -3394.31828  11205.68172  14205.68172
## [376] -28573.60814 -12794.31828  23426.39186 -12360.50348 -17794.31828
## [381]  -7794.31828  -4694.31828 -17794.31828  -9273.60814  -8794.31828
## [386]   -794.31828  31426.39186 -17794.31828 -17794.31828 -15994.31828
## [391] -17794.31828 -17794.31828  -6794.31828 -17756.80032  18426.39186
## [396] -28573.60814 -23273.60814  -7168.76514  -8713.04787 -17794.31828
## [401] -28273.60814    205.68172 -17794.31828   9107.18459 -17794.31828
## [406]  16050.52868   7249.74264 -17794.31828  40426.39186 -17794.31828
```

```
## [411]  47205.68172 -10163.99495  92205.68172  -1573.60814 -17794.31828
## [416]  17505.68172 -28573.60814 -17794.31828 -27873.60814 -17794.31828
## [421]   7215.07258 -28573.60814   5853.63725  14425.06234 289426.39186
## [426] -13091.03222   5205.68172  19965.43391 -18347.33040 -28573.60814
## [431] -17794.31828 -28573.60814 -28573.60814   4205.68172   5950.55849
## [436] -28573.60814  19426.39186 -27793.60814 -28573.60814 -28573.60814
## [441] -17794.31828 -28573.60814   6477.37215 -10794.31828  14205.68172
## [446]   -992.29379 -17794.31828 -17794.31828  -9373.60814 -28573.60814
## [451] -13994.31828 -17794.31828  11426.39186    205.68172 -17794.31828
## [456] -17794.31828  -4573.60814    -47.12770 -17794.31828  13126.39186
## [461] -17794.31828   6426.39186 -17794.31828 -17794.31828  -8902.78448
## [466]   7378.50714 -17794.31828  30505.68172 -17794.31828 -28573.60814
## [471] -17794.31828   3529.28410   4753.59536  18205.68172 -17794.31828
## [476]   9205.68172 -17794.31828 -17794.31828 -18273.60814 -12794.31828
## [481]  36205.68172  -6073.60814 -17794.31828  -1294.31828  16426.39186
## [486] -28573.60814 -28573.60814 -17794.31828    426.39186 -17794.31828
## [491]  -8862.48836  38426.39186   8373.84840  14205.68172 -17294.31828
## [496]   7909.16563  15131.57996 -17794.31828 -17694.31828 -28573.60814
## [501] -17794.31828  -5780.17952 -12308.84907 -26573.60814 -13573.60814
## [506] -17794.31828  -2257.46188   6426.39186  -7981.08059 -17254.31828
## [511]  19225.12779 -26573.60814  22030.72440 -28573.60814 -28573.60814
## [516] -24873.60814 -17794.31828 -17794.31828    819.52271  49426.39186
## [521] -17094.31828 -28573.60814  -3573.60814  51426.39186 -10373.60814
## [526]  22205.68172  17126.39186  19969.11557 -15894.31828  -9794.31828
## [531] -12794.31828 -28573.60814   9890.22506 -17794.31828  -7473.60814
## [536]  18426.39186 -17794.31828 -13842.29718    485.83172 -17794.31828
## [541]  14205.68172 -17794.31828  21426.39186 -28573.60814 -28573.60814
## [546] -12573.60814 -17794.31828  10685.57043 -17794.31828   7560.94335
## [551]   7205.68172   8693.59089 -22973.60814 -28573.60814 -28573.60814
## [556]  -1573.60814  11426.39186  13426.39186 301426.39186  23426.39186
## [561]  16426.39186 -17794.31828 -11279.81775   6426.39186 -17794.31828
## [566] -17794.31828 -17794.31828  21426.39186  41426.39186 -17203.02664
## [571]  18205.68172   4265.93584 -11794.31828 -28573.60814   1205.68172
## [576] -28573.60814 -17794.31828 -17794.31828 -17794.31828  -3573.60814
## [581] -10794.31828 -10363.01591 -17794.31828  -6573.60814 -10573.60814
## [586]  15016.70280 -17794.31828 -28573.60814  10016.58798   4905.68172
## [591]  -4638.59676  92205.68172  -2573.60814   3426.39186 -14294.31828
## [596]  -7180.47666  -3024.56065  96426.39186 -17794.31828 -28573.60814
## [601] -17794.31828  67426.39186  22205.68172 -17794.31828 -17794.31828
## [606]  51426.39186 -17794.31828 -11794.31828  17026.39186  17529.40950
## [611] -17794.31828  -5773.57596  -1794.31828  -4794.31828   7426.39186
## [616]   2359.33008 -17794.31828  30205.68172  37205.68172 -17794.31828
## [621]   -523.50492  19426.39186 -17833.25606  11426.39186  22205.68172
## [626]  51426.39186  -5382.75166  -4794.31828 -28573.60814 -14523.83678
## [631]   5805.68172   9105.68172  -9294.31828   8205.68172    772.90306
## [636]  -5570.27790 -17794.31828  10098.89521 -17794.31828  -2794.31828
## [641]  -8573.60814 -17794.31828  -3573.60814  22205.68172  12341.19048
## [646] -28573.60814 158205.68172 -11194.31828  -8194.31828 -23573.60814
## [651] -28573.60814 265426.39186   6259.73337 -13573.60814 -28573.60814
## [656]  -5294.31828  54426.39186 -28573.60814  -7794.31828  10205.68172
```

```
## [661] -16742.37189  -4573.60814  22205.68172   3013.79740 -17794.31828
## [666] -27073.60814  -5996.89712  12594.15815 -16944.31828  26426.39186
## [671] -12794.31828 -28573.60814  58205.68172  13820.08814 -28573.60814
## [676] -17794.31828  16426.39186  79205.68172 -22573.60814 -11073.60814
## [681]  14205.68172 -14594.31828  12205.68172 -28573.60814 -17794.31828
## [686]   2205.68172  -1573.60814 -17794.31828 -13016.14265  -5794.31828
## [691] -17794.31828  -4573.60814  64426.39186    705.68172 -17794.31828
## [696]  15919.41333  10226.39186 -17794.31828    905.68172 -11994.31828
## [701] -17794.31828   9324.99098 -28573.60814  42426.39186 -17794.31828
## [706]   1360.02553 -28573.60814  -6994.31828 -17794.31828 -28573.60814
## [711]  52205.68172 -17794.31828 -17794.31828  -5694.92810 -28573.60814
## [716]  92205.68172   -794.31828  10205.68172 -17344.31828  22205.68172
## [721]  -4573.60814 -15794.31828 -27613.60814  -8150.30212 -16573.60814
## [726]  -5794.31828  22205.68172   5576.35800 -17794.31828  21426.39186
## [731] -28573.60814   -294.31828 -28573.60814   6599.86989 -28573.60814
## [736]  77205.68172  46426.39186 -26773.60814  21426.39186  43426.39186
## [741] -17794.31828 -17794.31828 -28573.60814   1426.39186 -23973.60814
## [746] -10594.31828 -28573.60814 -17794.31828  31426.39186 -17794.31828
## [751]  55426.39186 -17794.31828 -28573.60814   -553.39841 -13573.60814
## [756]  11034.76451 -17794.31828   7719.36058 -17794.31828  54205.68172
## [761] -28573.60814 -28573.60814 -28573.60814 -16794.31828 -28573.60814
## [766]  19426.39186 -11794.31828  10805.68172 -28573.60814  21438.93726
## [771]    205.68172 -17794.31828 -17794.31828  37205.68172 -28573.60814
## [776] -17794.31828 -28573.60814 171426.39186  -1707.01910  -2573.60814
## [781]  16426.39186 -17794.31828   1426.39186 -28573.60814 -17794.31828
## [786]  12205.68172 -25773.60814  -5794.31828   -329.02530 -28573.60814
## [791]   3447.27897 -17794.31828 146426.39186    205.68172 -17794.31828
## [796] -28573.60814  -3573.60814  22205.68172 -17794.31828  20226.39186
## [801]  -7794.31828 -17794.31828  -4573.60814  50205.68172 -16694.31828
## [806]  30205.68172 -17794.31828  12646.48579   6398.78267 -28073.60814
## [811]  -3152.91634 353426.39186  31426.39186   6410.92923  21908.58882
## [816]   -794.31828 -17794.31828 -17773.60814  24426.39186 -28573.60814
## [821]  11205.68172  50205.68172 -12573.60814  73426.39186 -17794.31828
## [826]  62205.68172 -28573.60814 -10294.31828 -14794.31828  18205.68172
## [831] -18573.60814 -28573.60814  22205.68172 -28573.60814  67205.68172
## [836]  -7794.31828  57205.68172   8471.57385 -28573.60814 -17794.31828
## [841] -17794.31828  -8294.31828 -17794.31828 -17794.31828 -17794.31828
## [846]  -7764.78844  14031.39816 -14226.30145   4518.09676  27205.68172
## [851]  10205.68172 -28573.60814  -5794.31828  -5794.31828  57426.39186
## [856] -28573.60814   9705.68172  71426.39186 -17794.31828  41426.39186
## [861] -28573.60814   2398.40153  -4073.60814 -17794.31828  -4794.31828
## [866] -28573.60814 -28573.60814 -17794.31828  90205.68172 -17794.31828
## [871]    -94.31828 -28573.60814 -26073.60814 -28573.60814 -17794.31828
## [876] -28573.60814  79205.68172  62205.68172 -17794.31828   6205.68172
## [881] -27373.60814 -28573.60814  -3394.31828 -20173.60814  66205.68172
## [886]  -4573.60814 -17794.31828  -9873.60814 -13794.31828 -17794.31828
## [891] -13094.31828  -4573.60814   3105.68172  47205.68172 -17794.31828
## [896] -17794.31828   -394.31828 -14194.31828 -28573.60814   1705.68172
## [901] -16106.76303  33205.68172  21426.39186  -5794.31828 -17794.31828
## [906]  20205.68172  17426.39186 -17794.31828  36205.68172 -16573.60814
```

```
## [911] -28573.60814  37426.39186 -17794.31828 -17794.31828 -17794.31828
## [916]  13861.27110 -26573.60814  -7194.31828 -15087.11412   2205.68172
## [921]  59205.68172  22205.68172  -9988.63602   6426.39186 131426.39186
## [926] -16656.70680 -13329.67104  18205.68172   1958.99962 -28573.60814
## [931] -17794.31828  -8573.60814  74426.39186 -17794.31828 -17794.31828
## [936]  -1864.43416 -28573.60814 -17794.31828 -28573.60814   5205.68172
## [941]  10849.01490 -17794.31828 -17794.31828 -17794.31828 -17794.31828
## [946] -17794.31828   7205.68172  80205.68172 -28573.60814 -17794.31828
## [951] -14946.51249 -28573.60814  -4382.65528 -28573.60814   3313.59000
## [956]  -3573.60814 -28573.60814 -17794.31828 -28573.60814 -24773.60814
## [961] -28573.60814 -17794.31828  13457.67692 -17794.31828 -28573.60814
## [966]  41426.39186 -17794.31828 -17794.31828 -17794.31828 -28573.60814
## [971]    205.68172  -7917.27389   6060.02453 -12973.60814 -28573.60814
## [976]  -3573.60814 -16794.31828 -28573.60814  20826.39186    932.22443
## [981]   4205.68172 -14846.56205  15205.68172  -6673.60814  24205.68172
## [986]  41426.39186 -11794.31828 -17794.31828  -3821.39384  45426.39186
## [991]  14882.27555  82426.39186  91426.39186  14205.68172   2205.68172
## [996]   5905.68172  12205.68172 -13573.60814   8755.25914 -14594.31828
```

```
str(ACS_data)
```

```
## 'data.frame':    1000 obs. of  12 variables:
##  $ Sex            : chr  "Female" "Male" "Male" "Female" ...
##  $ Age            : int  31 31 75 80 64 14 78 35 70 18 ...
##  $ Married        : int  0 0 0 0 1 0 1 0 1 0 ...
##  $ Income         : num  60000 360 0 0 0 ...
##  $ HoursWk        : num  40 12 40 13.2 32.7 ...
##  $ Race           : chr  "white" "black" "white" "white" ...
##  $ USCitizen      : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ HealthInsurance: int  1 1 1 1 1 1 1 1 1 1 ...
##  $ Language       : int  1 0 0 0 0 0 0 1 0 0 ...
##  $ SexL           : int  0 1 1 0 1 1 1 1 0 0 ...
##  $ prediction     : num  17794 28574 28574 17794 28574 ...
##  $ residual       : num  42206 -28214 -28574 -17794 -28574 ...
```

```
#2. Calculate the sum of squared deviations and sum of absolute deviations.
SSD = sum(ACS_data$residual ^ 2)
cat("Sum of Squared Deviations(SSD) =", SSD, '\n')
```

```
## Sum of Squared Deviations(SSD) = 1.501666e+12
```

```
SAD = sum(abs(ACS_data$residual))
cat("Sum of Absolute Deviations Of Residuals(SAD) =", SAD, '\n')
```

```
## Sum of Absolute Deviations Of Residuals(SAD) = 22270304
```

QUESTION 8 (20 Points)[i]

1. Run an analysis of variance on the model of income by sex
2. Write one paragraph explaining and interpreting the results Hint:

- H0: bi = 0
- Ha: bi <> 0

{The Sum of Squared Residuals matches my previous calculation. The Sum of Squared SexL looks significantly less than the Sum of Squared Residuals. With a F-statistic value of 19.2, the null hypothesis H0 should be rejected in favor of Ha. With a p-value near 0, less than 0.001 in this case, the model is viewed as statistically significant.}

```
Betelguese <- anova(IncSex_model)
Betelguese

## Analysis of Variance Table
##
## Response: Income
##             Df     Sum Sq    Mean Sq F value    Pr(>F)
## SexL         1 2.8889e+10 2.8889e+10    19.2 1.302e-05 ***
## Residuals  998 1.5017e+12 1.5047e+09
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

QUESTION 9 (10 Points)

1. Calculate the proportional reduction in error
2. Write one paragraph explaining and interpreting the improvement of the linear model by adding sex Hint:

- H0: bi = 0
- Ha: bi <> 0

{The Proportional Reduction In Error calculation indicates that 1.89% of the total error in the model has been reduced by adding Sex. Ideally, one would want this value to be much higher, if you intended to use Sex only to estimate income. The Cohen's D value calculation equaled 0.278. This value is closer to a small effect size(d~0.2) than a medium effect size(d~0.5), and indicates the size of group difference in standard deviation units.}

```
PRE = Betelguese["SexL","Sum Sq"]/(Betelguese["SexL","Sum Sq"] +
Betelguese["Residuals","Sum Sq"])
cat("The Proprortional Reduction In Error(PRE) =", PRE)

## The Proprortional Reduction In Error(PRE) = 0.01887499

CD = cohensD(Income ~ SexL, data = ACS_data)
cat("The Cohen's D Value =", CD)

## The Cohen's D Value = 0.2778874
```

QUESTION 10 (20 Points)

Write one paragraph interpreting and concluding the results of your analysis.

{Adding sex to the model to explain income was statistically significant. Although, it only corrected a small of amount of error in it. Recommending two courses of possible actions to proceed with creating a better model to explain income: One, if it is not necessary to estimate the income of those who are not earning anything, maybe $0 income observations should be filtered from the data set. Two, there are 7 other variables in the data set that could be potentially added to help explain income further. Would be happy to pursue both of these recommendations.}

{Special Note: Noticed that the Sum of Squared Deviations from the empty model used for female incomes only, 4.1993e+11, was less than that of the Income by Sex model, 1.5017e+12. This difference can attributed to the larger standard deviation associated with male incomes, and a higher mean income that produces a higher residual from the $0 income observations.}

END OF FINAL EXAM