
Roman Alvarez

STAT3010

R Project

10 March 2023

Univariate Quantitative Analysis

Quantitative Variables in the data set: Hours of TV Watched Per Day, Age, Number of Children

Table 1: Descriptive Statistics for Quantitative Variables (n=397)

	Overall (N=397)
TV HOURS	
Mean (SD)	3.10 (2.48)
Median [Min, Max]	3.00 [0, 22.0]
AGE	
Mean (SD)	46.4 (17.8)
Median [Min, Max]	45.0 [19.0, 89.0]
CHILDREN	
Mean (SD)	2.51 (0.926)
Median [Min, Max]	2.00 [0, 7.00]

Figure 1: Histogram of Hours of TV Watched (n=397)

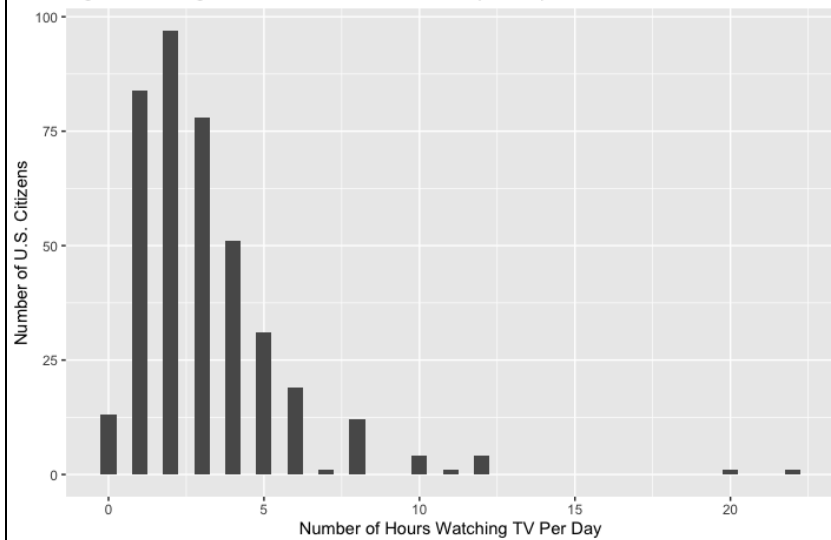


Figure 2: Boxplot of Hours of TV Watched (n=397)

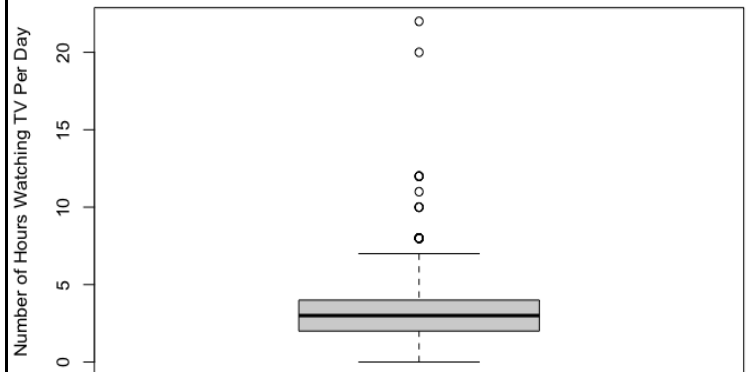


Figure 3: Histogram of Age From Sample Population (n=397)

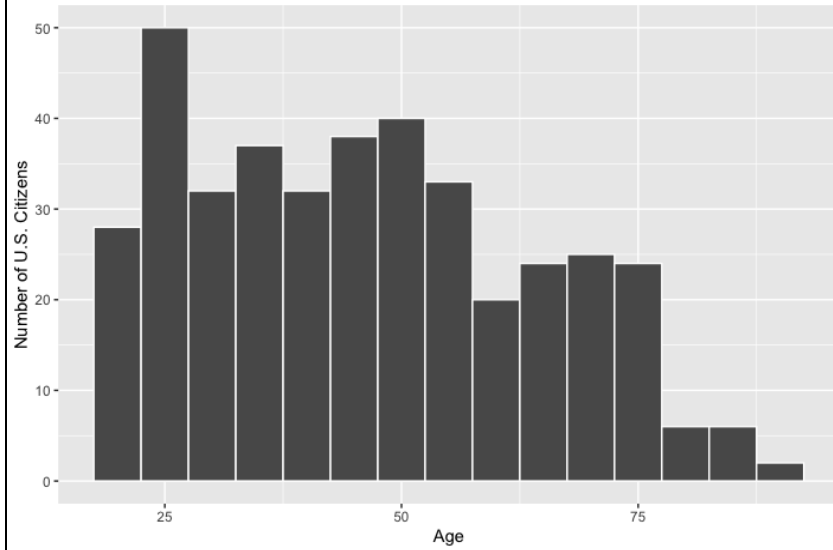


Figure 4: Boxplot of Age From Sample Population (n=397)

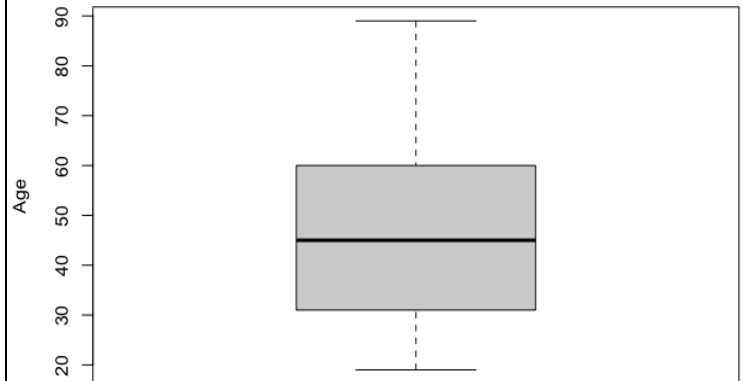


Figure 5: Histogram of the Amount of Ideal Childern Per Each Person (n=397)

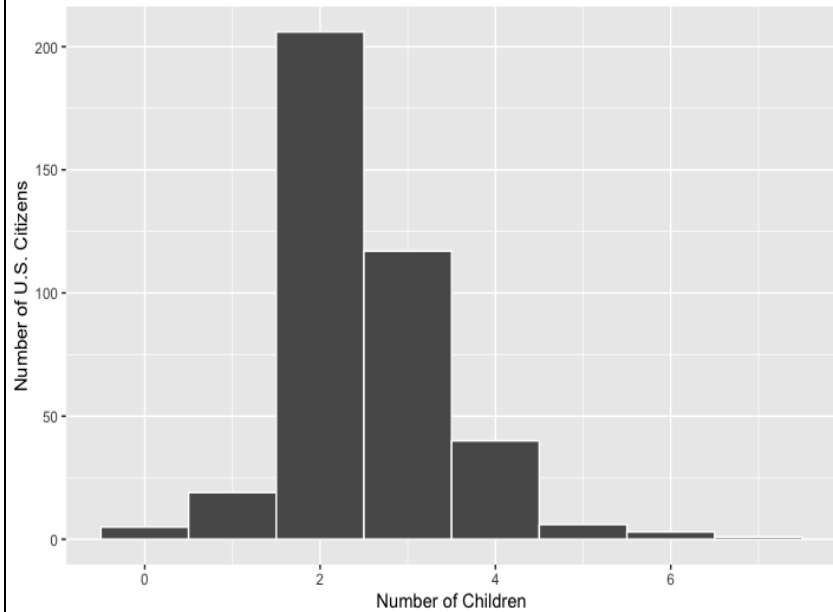
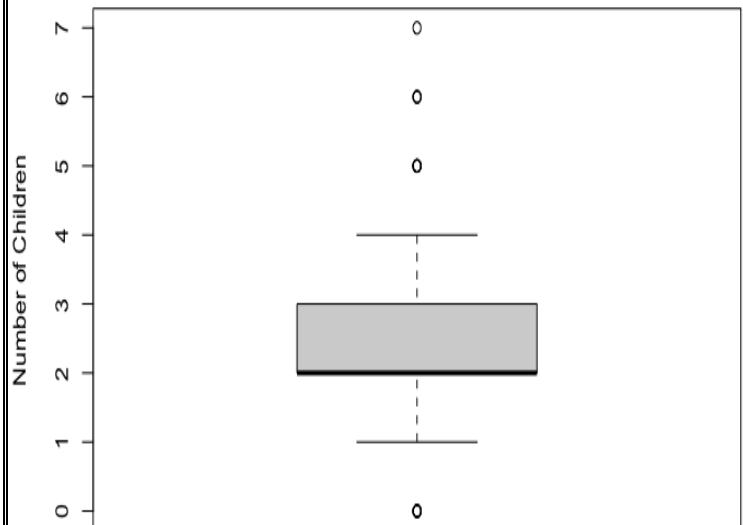


Figure 6: Boxplot of Ideal Amount of Children Per Each Person (n=397)



Hours of TV Watched: Figure 1 shows the data is skewed right, with most of the sample population consuming 0-5 hours of TV per day. From Table 1 we can conclude the average amount of time watching TV is 3.10 hours. There are outliers who consume 20+ hours a day, but this number is small. Because of the skewness, this data set's best measure of central tendency is the median, and interquartile range for measuring the dispersion.

Age of Sample Population: Figure 3 shows the data is uniform, with a slightly larger amount of people being in their mid-twenties. From Table 1 we can conclude that the average age is 46 with the youngest being 19 and the oldest being 89. Because the data is evenly distributed, the best

measure of central tendency is the mean and the standard deviation is the best measure of dispersion.

Ideal Number of Children: Figure 5 shows the data is symmetric with a slight skewness towards 2 children. From Table 1 we can conclude the average number of children is 2.5 with the minimum being 0 and the maximum being 7. Since the data is symmetric, the mean is the best measure of central tendency and the standard deviation is the best measure of dispersion.

Univariate Categorical Analysis

Categorical Variables: Sex, Race, Political Party

Table 2: Descriptive Statistics for Categorical Variables (n=397)

	Overall (N=397)
Sex	
Female	177 (44.6%)
Male	220 (55.4%)
Race	
AfrAm	54 (13.6%)
Hispanic	7 (1.8%)
Other	15 (3.8%)
White	321 (80.9%)
Political Party	
Democrat	139 (35.0%)
Independent	137 (34.5%)
Other	8 (2.0%)
Republican	113 (28.5%)

Figure 7: Pie Chart of Sex in Sample Population (n=397)

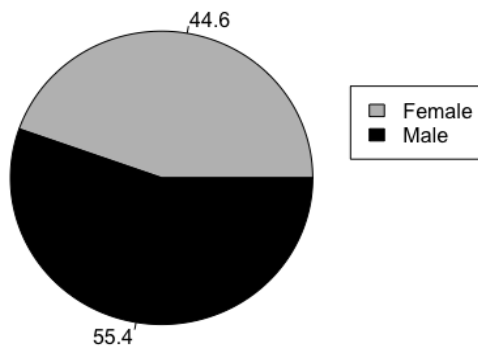


Figure 8: Bar Chart of Sex in Sample Population (n=397)

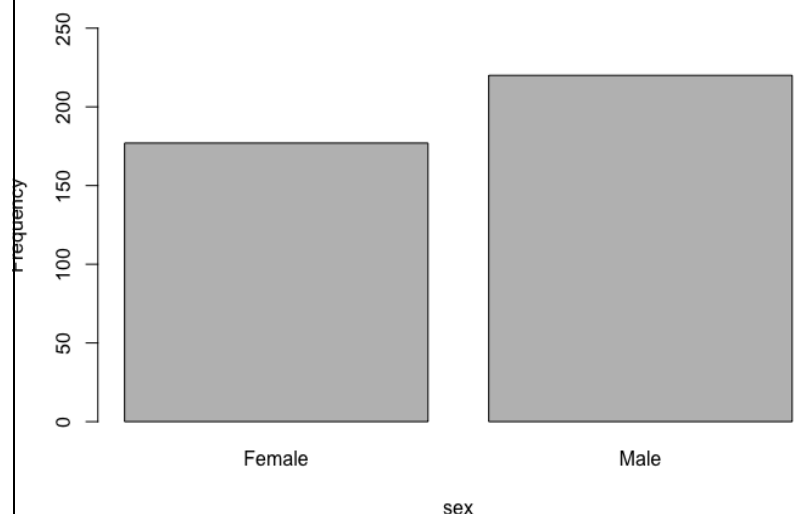


Figure 9: Pie Chart of Race in Sample Population (n=397)

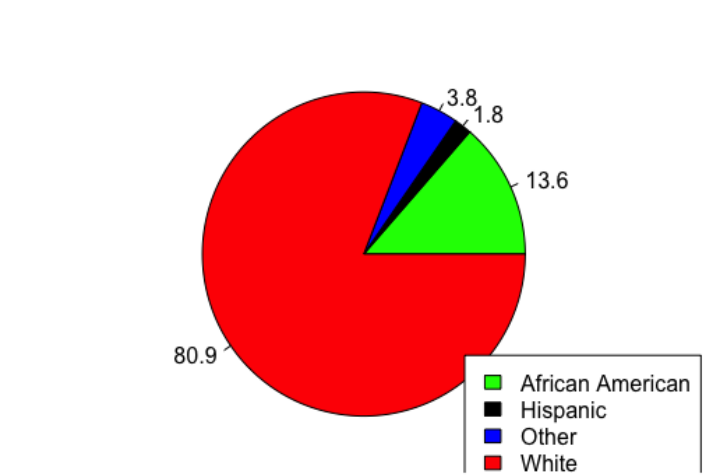


Figure 10: Bar Chart of Race in Sample Population (n=397)

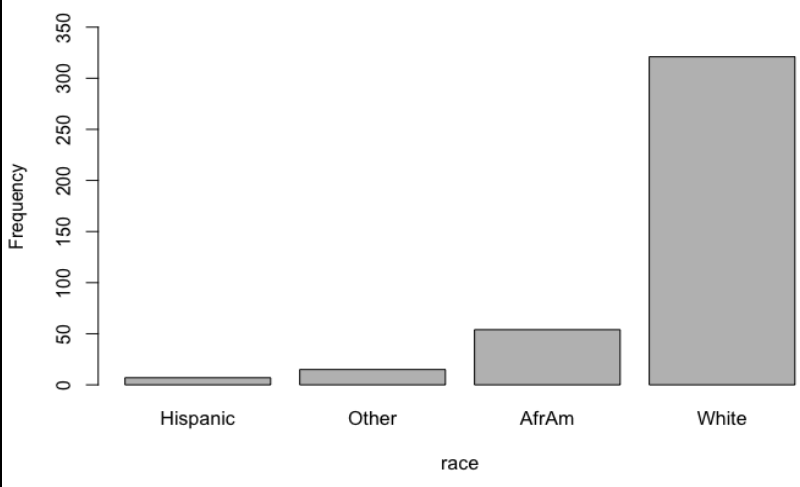


Figure 11: Pie Chart of Political Party in Sample Population (n=397)

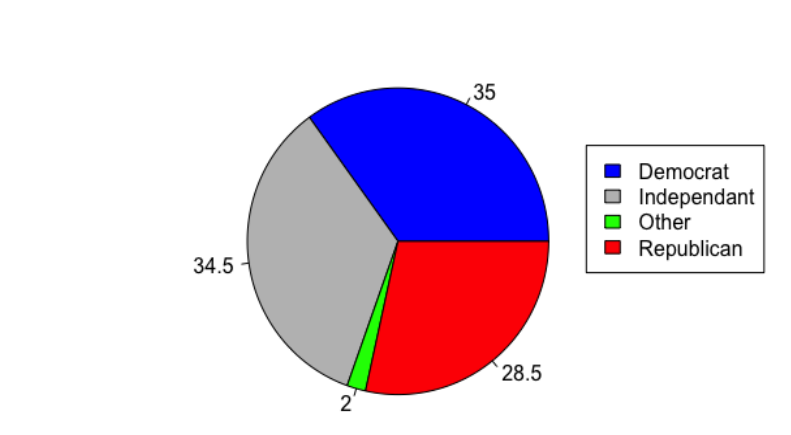
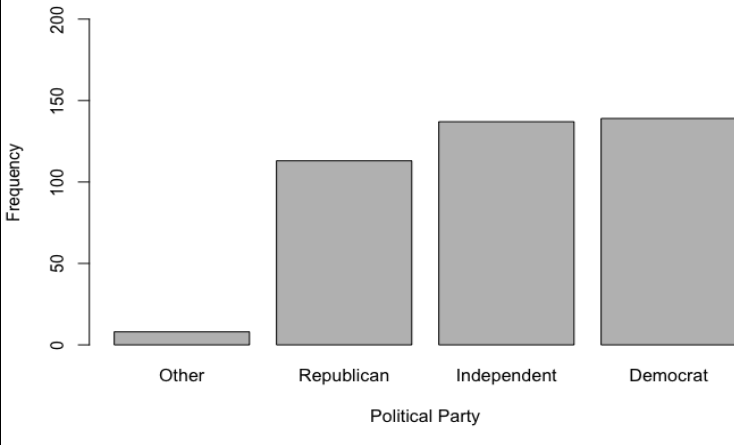


Figure 12: Bar Chart of Political Party in Sample Population (n=397)



Sex: From Figures 7 and 8 we can conclude that this sample population has a relatively proportionate amount of men and women, with 55.4% or 220 being men and 44.6% or 177 being women.

Race: From Figures 9 and 10 we can conclude that the majority of the sample population is White, making up 80.9%, totaling 321 people. There are 54 African Americans that make up 13.6% of the sample size, 7 Hispanics make up 1.8%, and 15 or 3.8% of individuals classified as “other”.

Political Party: From Figures 11 and 12 we can see that the political parties that make up the sample population are distributed relatively evenly amongst the three

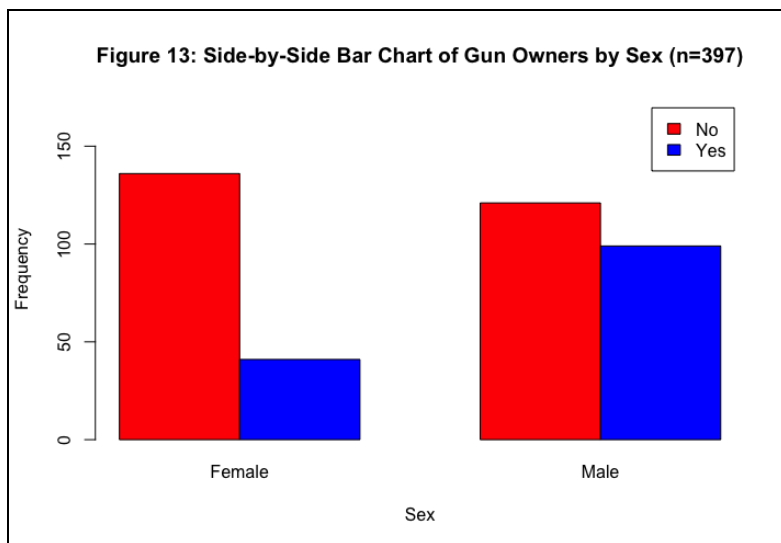
parties. Democrats make up 35% or 139 people, Republicans make up 28.5% or 113 people, and Independents make up 34.5% or 137 people. The outlier of the group that is classified as “other” only consists of 2% of the sample with 8 people.

Bivariate Analysis 2 Categorical

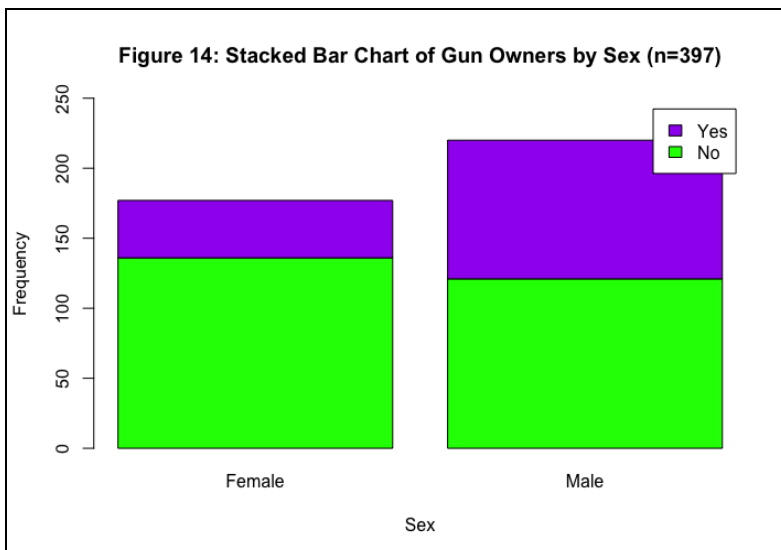
2 Categorical Variables that are being compared: Race and Political Party

Table 3: Descriptive Statistics for the Sex of Gun Owners (n=397)

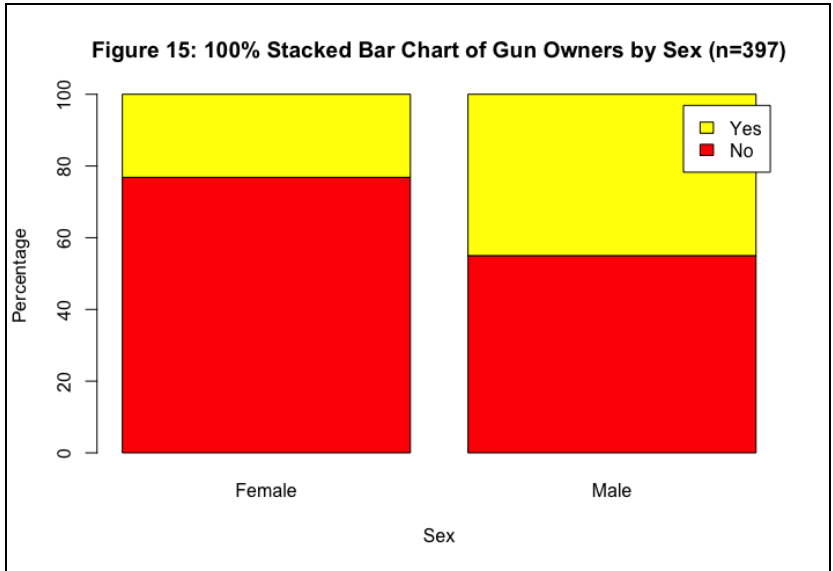
	Female (N=177)	Male (N=220)	Overall (N=397)
Own a Gun			
No	136 (76.8%)	121 (55.0%)	257 (64.7%)
Yes	41 (23.2%)	99 (45.0%)	140 (35.3%)



From Figure 13, it shows that the majority of the gun owners in the sample are Male. Additionally, it shows that the Male and Female proportions for owning a gun are different.



From Figure 14, we can see a better representation of how each population is divided from gun ownership.



From Figure 15 we can see how the percentages of each sex's population are divided through gun ownership. Male gun owners make up 45%, and female gun owners make up 23.2%.

This relationship of the sex of gun owners shows that males are much more likely to own a gun in comparison to women. Figure 14 visualizes this relationship the best, portraying the discrepancy of gun ownership from the sex of the individual.

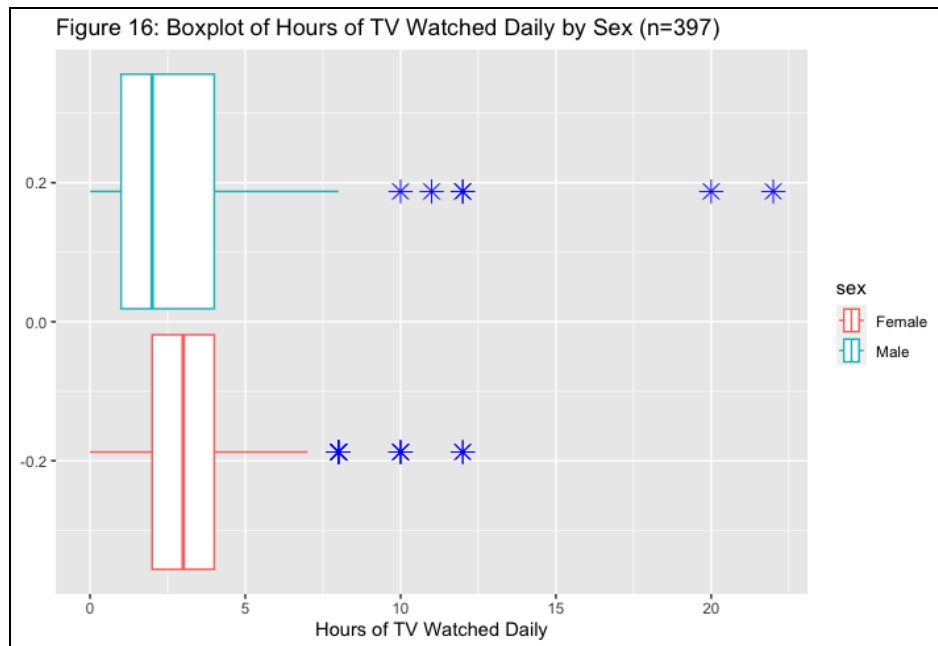
Bivariate Analysis 1 Cat 1 Quant

3 quantitative variables: Daily TV Hours Watched, Age, Ideal Number of Children

Categorical Variable: Sex

Table 4: Stratified Analysis of Hours of TV Watched Daily, Age, and Ideal Number of Children in Conjunction with Sex (n=397)

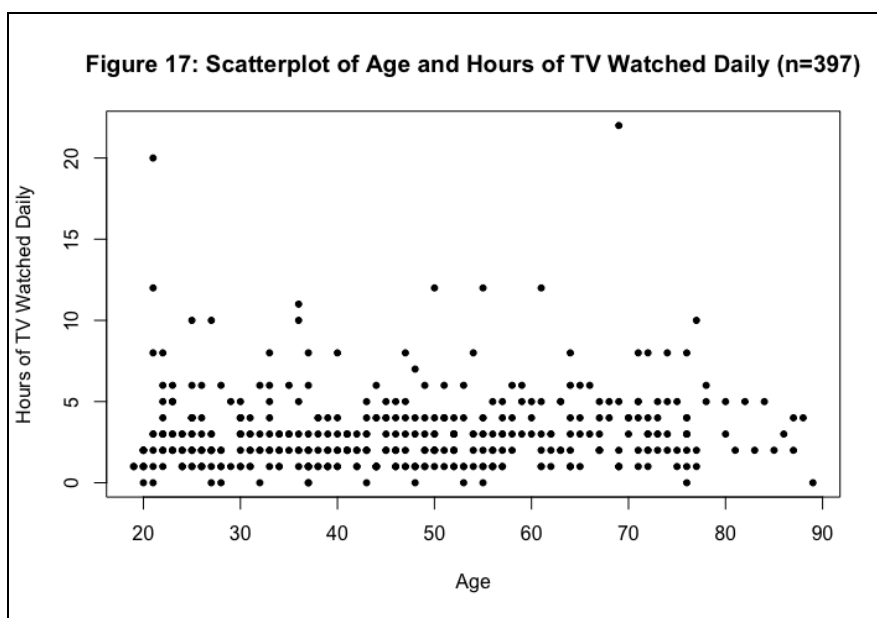
	Female (N=177)	Male (N=220)	Overall (N=397)
Hours of TV Watched Daily			
Mean (SD)	3.27 (2.25)	2.96 (2.65)	3.10 (2.48)
Median [Min, Max]	3.00 [0, 12.0]	2.00 [0, 22.0]	3.00 [0, 22.0]
Age			
Mean (SD)	48.1 (18.5)	45.0 (17.0)	46.4 (17.8)
Median [Min, Max]	48.0 [19.0, 88.0]	42.5 [19.0, 89.0]	45.0 [19.0, 89.0]
Ideal Number of Children			
Mean (SD)	2.56 (0.844)	2.47 (0.986)	2.51 (0.926)
Median [Min, Max]	2.00 [1.00, 7.00]	2.00 [0, 6.00]	2.00 [0, 7.00]



From Figure 16 we can note that females watch more hours of TV on a daily basis with their average consumption being 3.27 hours in comparison to the 2.96 hours watched by males. However, males have more outliers and a larger maximum amount of hours at 22 in comparison to the female maximum of 12. Though men have a lower median amount of hours consumed, their interquartile range is larger, showing a larger spread of data than women.

Bivariate Analysis 2 Quant

2 Quantitative Variables that are being compared: Age and Hours of tv watched daily



The correlation coefficient for age and hours spent watching tv daily is 0.077, thus stating the age of the individual has a low effect on the time spent watching tv a day. From the scatterplot we can conclude this correlation is weak because of how spread the data points are. The correlation is neither strong nor negative and has a nonlinear structure.

Confidence Interval

Quantitative Variable: Age

At 90% Confidence

upper	mean	lower
-------	------	-------

47.85	46.38	44.91
-------	-------	-------

At 95% Confidence

upper	mean	lower
-------	------	-------

48.13	46.38	44.63
-------	-------	-------

At 99% Confidence

upper	mean	lower
-------	------	-------

48.69	46.38	44.08
-------	-------	-------

As the confidence level increases, the parameter of the population average also increases. At a 95% confidence level we can conclude that the average age for 95% of the population is likely to be 44.63 to 48.13 years old.

Variable Creation

New Categorical Variable: TV Consumption

Levels:

0-2 Hours a day- “Low”

3-4 Hours a day- “Medium”

5+ Hours a day- “High”

These levels were created from evaluating that majority of the sample population watches 0-5 hours of tv a day.

Table 5: Stratified Analysis of Age, and Ideal Number of Children in Conjunction with TV Consumption Category (n=397)

	High (N=74)	Low (N=194)	Medium (N=129)	Overall (N=397)
Age				
Mean (SD)	49.9 (19.4)	44.1 (17.0)	47.7 (17.6)	46.4 (17.8)
Median [Min, Max]	50.5 [21.0, 84.0]	42.0 [19.0, 89.0]	47.0 [21.0, 88.0]	45.0 [19.0, 89.0]
Ideal Number of Children				
Mean (SD)	2.74 (1.14)	2.51 (0.900)	2.39 (0.803)	2.51 (0.926)
Median [Min, Max]	3.00 [0, 6.00]	2.00 [0, 7.00]	2.00 [0, 5.00]	2.00 [0, 7.00]

From Table 5 we can conclude that as tv consumption rises, so does the average age of the individual. To compare, the average age of a “high” consumer of tv is 49.9, while the average age for a “low” consumer of tv is 44.1. There is little to no correlation in regards to the ideal amount of children and tv consumption.

