Assignment 2
Miguel Perez-Luna
September 15, 2020
SES 5215

In this assignment, I found the sample mean, sample standard deviation, 95-percent confidence interval, and interquartile range for each of the continuous variables in my data set. I also found the 95-percent confidence interval for the proportion of each of the unique values in my categorical variables.

It should also be noted that I further filtered my original data set, selecting only values greater than or equal to zero for number of vehicles and income, and values greater than zero for travel time to work, gross rent, and age. This essentially narrows my sample down to renters and commuters. This reduced the number of observations to 6, 914.
I rounded all decimal values in my continuous variables to the hundredths place.

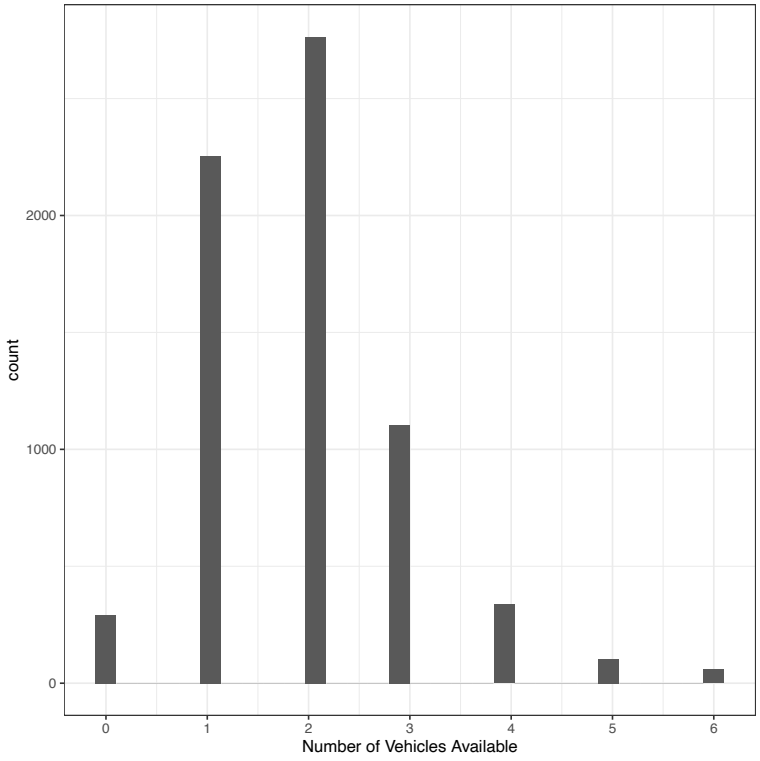
The variables I am considering are:

1. Sex (SEX) (categorical)
2. Means of Transportation to Work (JWTR) (categorical)
3. Educational Attainment (SCHL) (categorical)
4. Vehicles (1 ton or less) available (vehicle) (continuous)
5. Income (PINCP) (continuous)
6. Travel Time to Work (JWMNP) (continuous)
7. Gross Rent (GRNTP) (continuous)
8. Age (AGEP) (continuous)

Number of Vehicles Available (continuous)

Sample Mean	1.93
Sample Standard Deviation	1.06
Interquartile Range	1 to 2
95% Confidence Interval	1.90 to 1.95

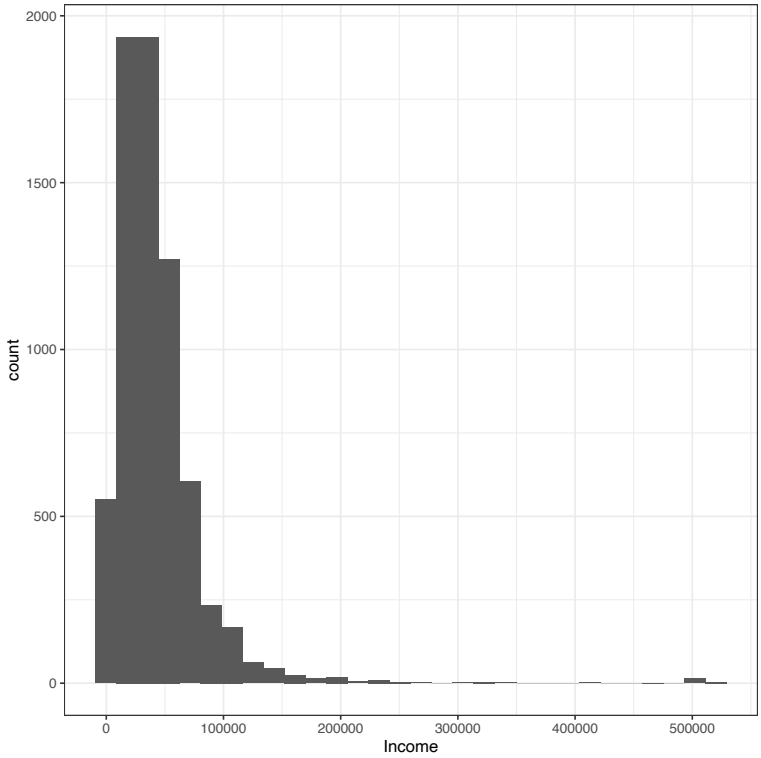
On average, a person in my sample has access to 1.93 cars. Most of the data is concentrated around the mean, and it skews right which means few people have access to 5 or more cars. We can be certain that the population mean lies somewhere between 1.90 and 1.95 cars.



Income (continuous)

Sample Mean	$42,726
Sample Standard Deviation	$40,789.26
Interquartile Range	$20,185 to $53,500
95% Confidence Interval	$41,764.79 to $43,688.04

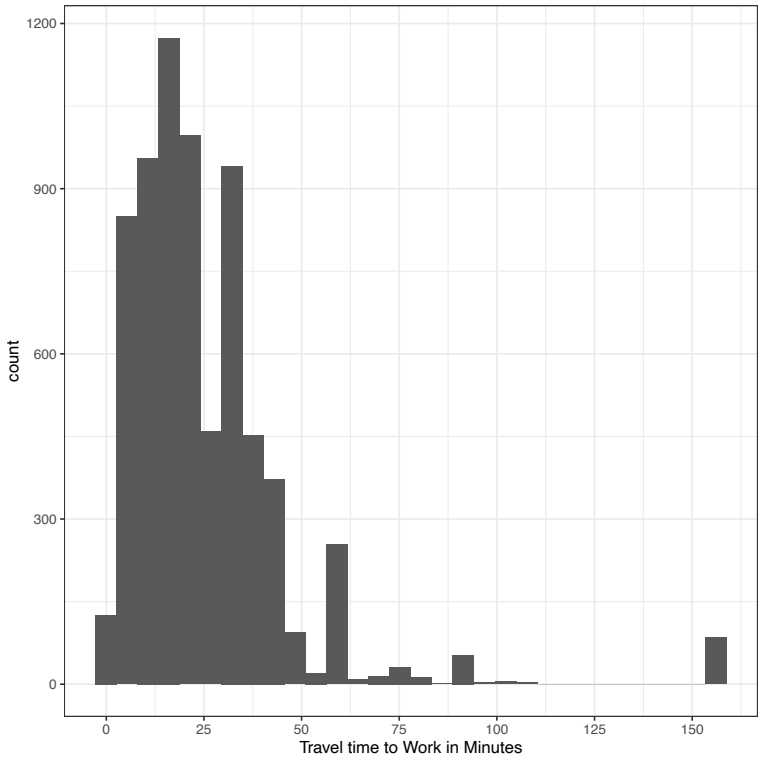
On average, a person in my sample earns an income of $42,726. The data is concentrated around the mean, but is skewed right by some large outliers. We can be certain that the population mean lies somewhere between $41,764.79 and $43,688.04.



Travel Time to Work in Minutes (continuous)

Sample Mean	24.79
Sample Standard Deviation	21.88
Interquartile Range	10 to 30
95% Confidence Interval	24.27 to 25.30

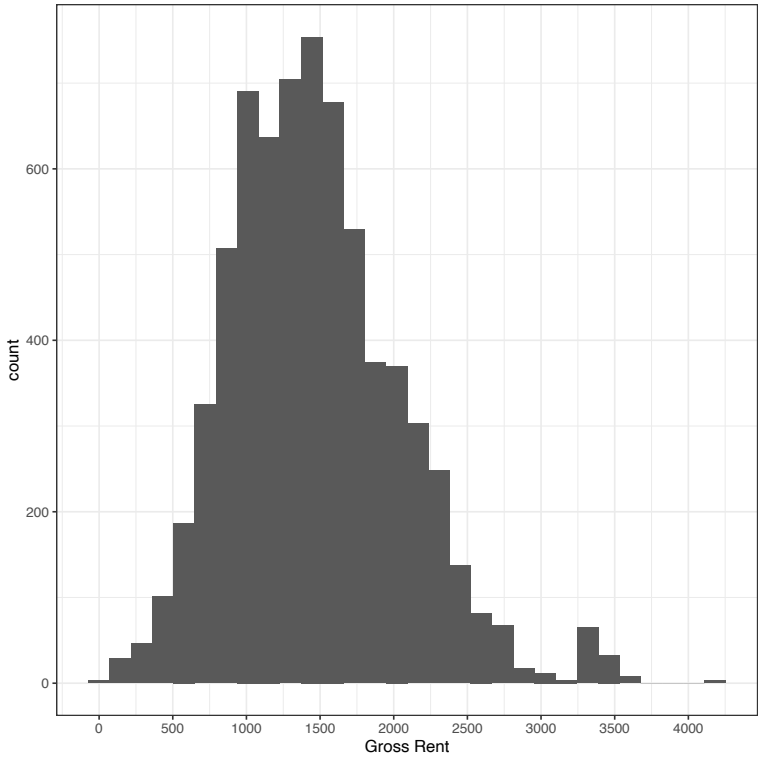
On average, a person in my sample spends 24.79 minutes traveling to work. The data is concentrated around the mean, but is skewed right by some outliers. We can be certain the population mean lies somewhere between 24.27 and 25.30 minutes.



Gross Rent (continuous)

Sample Mean	$1,466
Sample Standard Deviation	$581.28
Interquartile Range	$1047 to $1801
95% Confidence Interval	$1452.36 to $1479.77

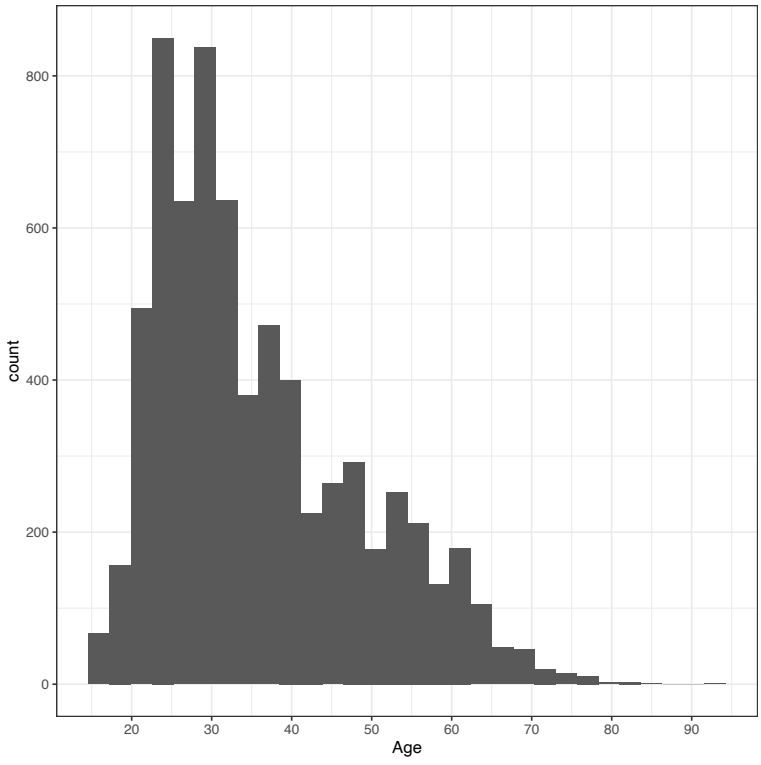
On average, a person in my sample spends $1,466 on monthly rent. The data is nearly normally distributed with some outliers paying higher rents. We can be confident that the population mean lies somewhere between $1,452.36 and $1,479.77.



Age (continuous)

Sample Mean	35.90
Sample Standard Deviation	12.76
Interquartile Range	26 to 44
95% Confidence Interval	35.59 to 36.20

On average, a person in my sample is 35.90 years old. The data is quite spread out and skews right. We can be confident that the population mean lies somewhere between 35.59 and 36.20 years.



Sex
(categorical)

The table below shows the proportion and lower and upper bounds of the 95-percent confidence interval for the proportion of each unique value in the Sex category. There are two unique values.

Proportions and 95-percent confidence intervals

Sex	Share	Low	High
Male	0.5410761	0.5293274	0.5528248
Female	0.4589239	0.4471752	0.4706726

Means of Transportation to Work
(categorical)

The table below shows the count, proportion, and 95-percent confidence interval for the proportion of each unique value in the Means of Transportation to Work category. There are ten unique values.

Proportions and 95-percent confidence intervals

Means of Travel	Share	Low	High
Car, Truck, or Van	0.8761932	0.8684278	0.8839586
Walked	0.0496095	0.0444900	0.0547289
Bicycle	0.0186578	0.0154675	0.0218481
Bus or Trolley Bus	0.0306624	0.0265977	0.0347272
Subway or Elevated	0.0049176	0.0032683	0.0065668
Other Method	0.0088227	0.0066179	0.0110275
Taxicab	0.0027480	0.0015138	0.0039823
Motorcycle	0.0023141	0.0011813	0.0034470
Railroad	0.0046283	0.0030280	0.0062286
Streetcar or Trolley Car	0.0014463	0.0005503	0.0023423

Educational Attainment (categorical)

The table below shows the proportion and lower and upper bounds of the 95-percent confidence interval for the proportion of each unique value in the Educational Attainment category. There are twenty-four unique values.

Proportions and 95-percent confidence intervals

Educational Attainment	Share	Low	High
Some college, but less than 1 year	0.0679780	0.0620435	0.0739126
Regular high school diploma	0.1706682	0.1617981	0.1795384
Bachelor's Degree	0.2652589	0.2548503	0.2756675
One year of college, no degree	0.1687880	0.1599568	0.1776191
Master's Degree	0.0828753	0.0763753	0.0893754
GED or alternative credential	0.0445473	0.0396832	0.0494114
Doctorate Degree	0.0141741	0.0113871	0.0169612
Associate's Degree	0.0877929	0.0811207	0.0944651
Grade 8	0.0034712	0.0020845	0.0048579
Grade 7	0.0021695	0.0010725	0.0032665
Grade 10	0.0131617	0.0104747	0.0158487
Grade 12 - no diploma	0.0161990	0.0132226	0.0191754
Grade 6	0.0075210	0.0054840	0.0095580
Grade 11	0.0216951	0.0182603	0.0251300
Professional Degree beyond Bachelor's	0.0156205	0.0126969	0.0185441
Grade 9	0.0080995	0.0059862	0.0102128
No schooling completed	0.0072317	0.0052340	0.0092294
Grade 5	0.0008678	0.0001736	0.0015621
Grade 4	0.0004339	-0.0000571	0.0009249
Nursery schoolor preschool	0.0001446	-0.0001389	0.0004282
Grade 3	0.0004339	-0.0000571	0.0009249
Grade 1	0.0002893	-0.0001117	0.0006902
Kindergarten	0.0002893	-0.0001117	0.0006902
Grade 2	0.0002893	-0.0001117	0.0006902