

# Assignment 2: Report

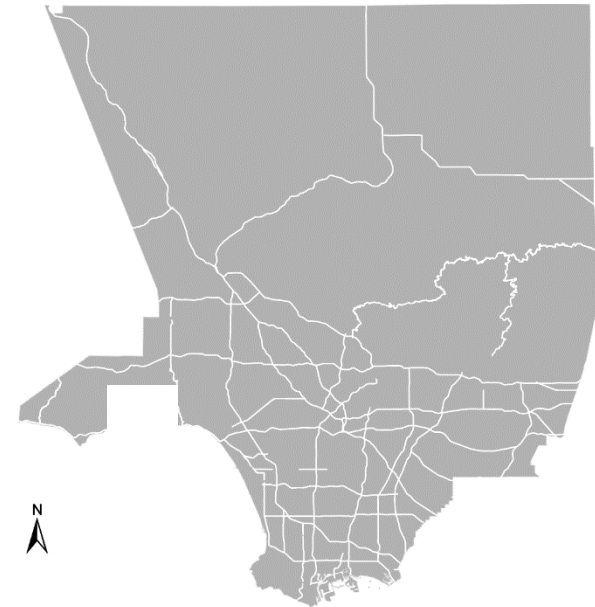
Catherine Saint

9/10/2020

Instructions: For each continuous variable in your dataset, calculate the sample mean, the sample standard deviation, the 95-percent confidence interval for the population mean, and the interquartile range. Create a histogram to illustrate the distribution of each variable and describe the distribution in a sentence or two.

## List of variables (from Assignment 1)

1. Median Age (continuous) : med\_ageE
2. Part of the total population unemployed in civilian labor force (age 16 years and above) (continuous): pct\_unemp\_labor
3. Part of the total population who are high school graduates (or equivalent) in labor force (continuous): pct\_hs\_grad
4. Majority race of the population (categorical) : maj\_race\_tract
5. Whether the majority of the population is foreign born or not(categorical): maj\_foreign



## Distribution of Continuous Variables

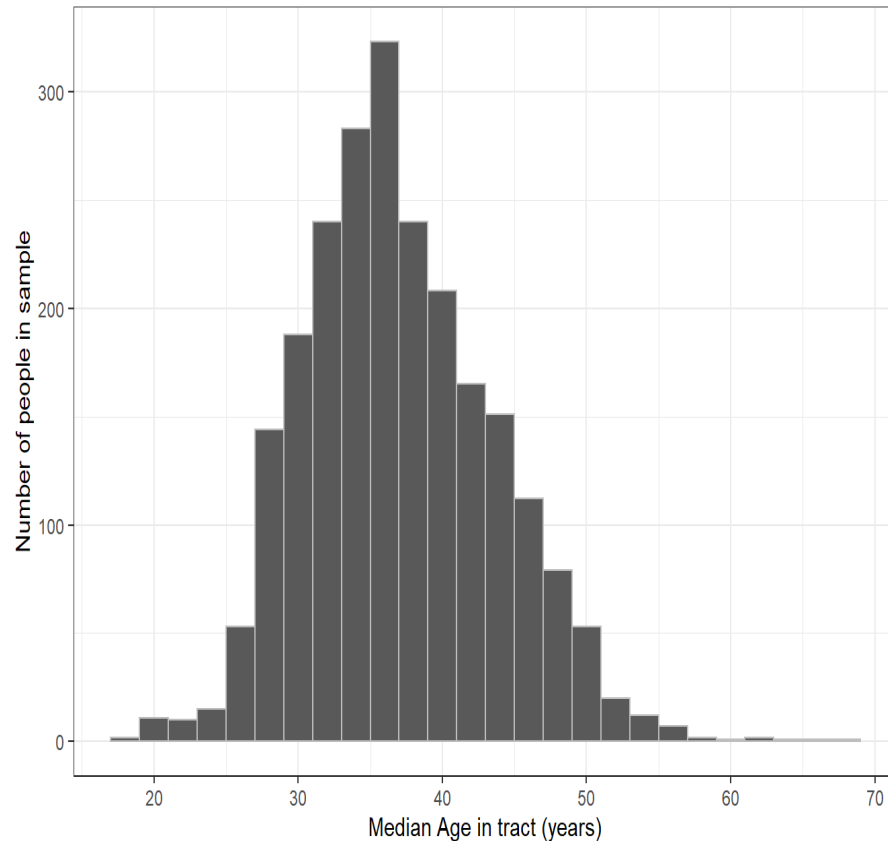
### Descriptive Statistics

#### Characteristics of Census Tracts in Los Angeles County

<i><b>Statistic</b></i>	<i><b>Median Age (years)</b></i>	<i><b>Percent unemployed</b></i>	<i><b>Percent high-school graduate workers</b></i>
<i><b>Sample Mean</b></i>	37.1	3.55%	8.4%
<i><b>Median</b></i>	36.4	3.32%	8.67%
<i><b>Standard Deviation</b></i>	6.59	1.71%	4.25%
<i><b>Interquartile Range</b></i>	32.4 to 41.3	2.38% to 4.51%	5.69% to 11%
<i><b>Population Mean</b></i>	36.8 to 37.3	3.48% to 3.62%	8.23% to 8.58%

## Distribution of Continuous Variables (cont.d)

### Histograms



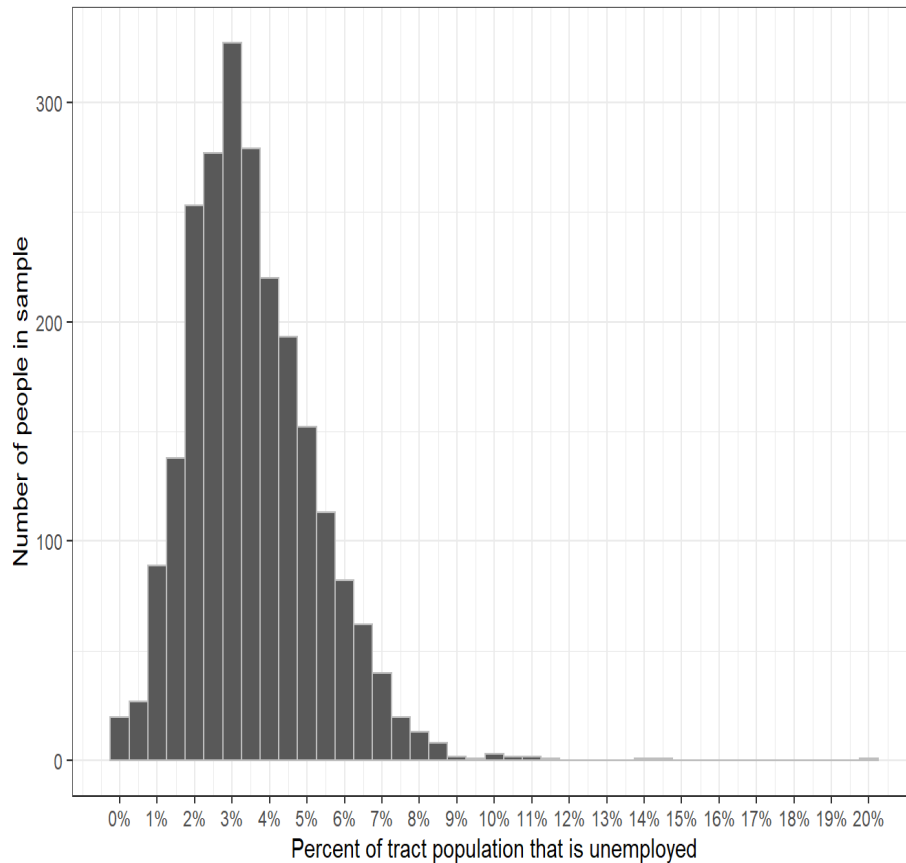
### Median Age

The distribution looks relatively normal, with a slightly positive, right-skew. The median value is less than the corresponding mean value, because there are a greater number of outliers within the older population, affecting the mean value.

Statistic	Median Age (years)	Percent unemployed	Percent high-school graduate workers
Sample Mean	37.1	3.55%	8.4%
Median	36.4	3.32%	8.67%
Standard Deviation	6.59	1.71%	4.25%
Interquartile Range	32.4 to 41.3	2.38% to 4.51%	5.69% to 11%
Population Mean	36.8 to 37.3	3.48% to 3.62%	8.23% to 8.58%

## Distribution of Continuous Variables (cont.d)

### Histograms



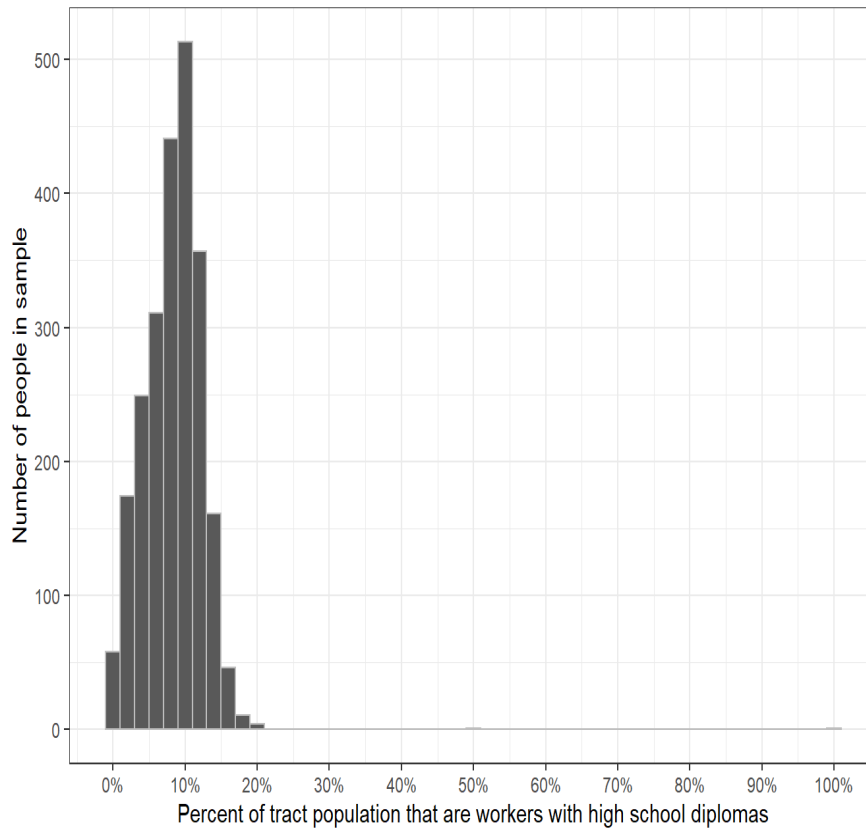
### Unemployment

The distribution of unemployment is similar to median age in that it is relatively normal, with a slightly positive, right-skew. The median value is less than the corresponding mean value, because there are a greater number of outliers located along our x axis, where unemployment percentage increases.

Statistic	Median Age (years)	Percent unemployed	Percent high-school graduate workers
Sample Mean	37.1	3.55%	8.4%
Median	36.4	3.32%	8.67%
Standard Deviation	6.59	1.71%	4.25%
Interquartile Range	32.4 to 41.3	2.38% to 4.51%	5.69% to 11%
Population Mean	36.8 to 37.3	3.48% to 3.62%	8.23% to 8.58%

## Distribution of Continuous Variables (cont.d)

### Histograms



#### Worker Education

The data skews negatively to the left. Although there are significant outliers located to the right of the mean and median, there still remains a greater number of observations closer to 0%. This explains why the median value is greater than the mean.

Statistic	Median Age (years)	Percent unemployed	Percent high-school graduate workers
Sample Mean	37.1	3.55%	8.4%
Median	36.4	3.32%	8.67%
Standard Deviation	6.59	1.71%	4.25%
Interquartile Range	32.4 to 41.3	2.38% to 4.51%	5.69% to 11%
Population Mean	36.8 to 37.3	3.48% to 3.62%	8.23% to 8.58%

## Distribution of Categorical Variables

### Descriptive Statistics

#### Race/Ethnicity Statistics for Census Tracts in Los Angeles County

<i>Majority Race</i>	<i>Sample Proportion</i>	<i>95-percent confidence interval</i>
<i>Hispanic/Latino</i>	46.2%	44.2% to 48.2%
<i>White (not Hispanic/Latino)</i>	23.2%	21.5% to 24.9%
<i>No Majority</i>	23%	21.3% to 24.7%
<i>Asian (not Hispanic/Latino)</i>	5.29%	4.38% to 6.19%
<i>Black (not Hispanic/Latino)</i>	2.26%	1.66% to 2.86%

#### Nativity Statistics for Census Tracts in Los Angeles County

<i>Majority Nativity</i>	<i>Sample Proportion</i>	<i>95-percent confidence interval</i>
<i>Native Born</i>	88.3%	87% to 89.6%
<i>Foreign Born</i>	11.7%	10.4% to 13%