

Caitlin Sizemore

DS-160-01 Spring 2023

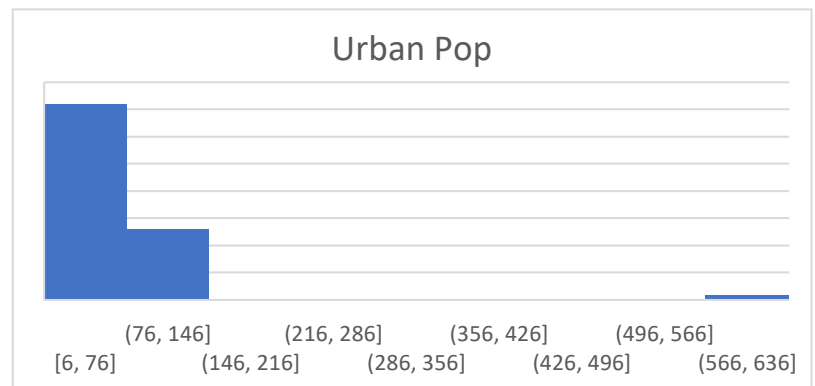
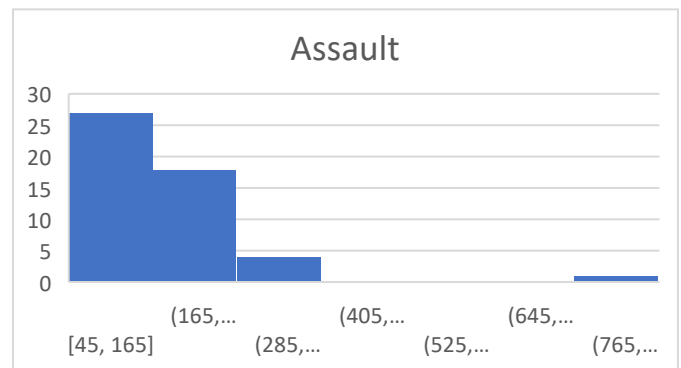
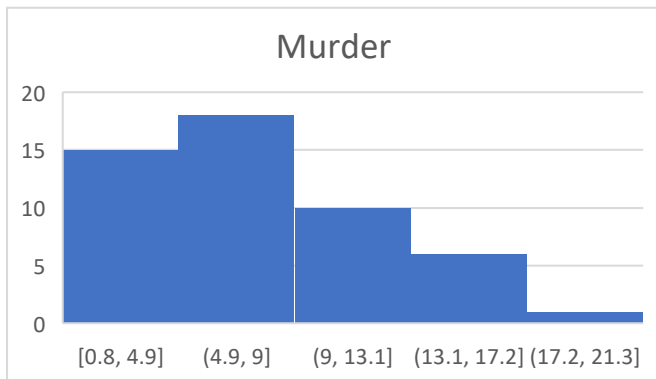
In-Class Assignment 1

1. Data Analysis: hands on, exploring and working with actual data and looking at what actually happened

Data Analytics: broad term that covers all activities relating to data, data analysis is part of analytics, uses data for prediction

2. Variable types: numerical
3. Interval data do not have an absolute zero and can go below zero. Ratio data has an absolute zero and cannot go below zero.
4. Descriptive Analysis: The initial analysis performed – consists of distribution and its dispersion and measure of centrality – gives an overall look at the data

- a. Murder: Positively skewed
- b. Assault: Positively skewed
- c. Urban Pop: Positively skewed



5. Measure of Dispersion: shows spread of data and consists of range and interquartile range
 - a. Murder
 - i. Min: 0.8
 - ii. Q1: 4.075

- iii. Q2: 7.25
- iv. Q3: 11.25
- v. Max: 17.4
- vi. IQR: 7.175

b. Assault

- i. Min: 45
- ii. Q1: 109
- iii. Q2: 159
- iv. Q3: 249
- v. Max: 879
- vi. IQR: 140

c. Urban Pop

- i. Min: 6
- ii. Q1: 53.25
- iii. Q2: 66
- iv. Q3: 77.75
- v. Max: 570
- vi. IQR: 24.5

6. Measure of Centrality: Mean, Median, Mode – describes the middle of the distribution

a. Murder:

- i. Mean: 7.788
- ii. Median: 7.25
- iii. Mode: 13.2

b. Assault:

- i. Mean: 182.1836
- ii. Median: 159
- iii. Mode: 120

c. Urban Pop:

- i. Mean: 74.2
- ii. Median: 66
- iii. Mode: 80

7. Diagnostic Analysis: The why – using the data to find why something happened
 - a. Murder/Assault: 64.94%
 - b. Murder/Urban Pop: 18.62%
 - c. Assault/Urban Pop: 14.07%