S.J.S.U Crime Stats (2014-2019)

Project Task 4 Update
GR 4
Kevin Thai, Austin Tom, Antonio Perez, Shruti Hardasani, & Ching-Fen Juan
4/14/2023

1

Descriptive Data

```
> summary(Crime_data)
 Sexual Assault    Aggravated Assault     Theft
 Min.    :13.00    Min.    : 2.000    Min.    :27.0
 1st Qu.:16.75    1st Qu.: 5.250    1st Qu.:39.0
 Median :19.50    Median : 9.000    Median :45.5
 Mean   :19.67    Mean   : 8.333    Mean   :45.5
 3rd Qu.:21.50    3rd Qu.:12.000    3rd Qu.:55.0
 Max.    :28.00    Max.    :13.000    Max.    :60.0
> summary(C_data)
   Offenses              Crime_Type
 Length:18          Min.    : 2.00
 Class  :character    1st Qu.:12.25
 Mode   :character    Median :19.50
                      Mean   :24.50
                      3rd Qu.:35.50
                      Max.    :60.00
>
```
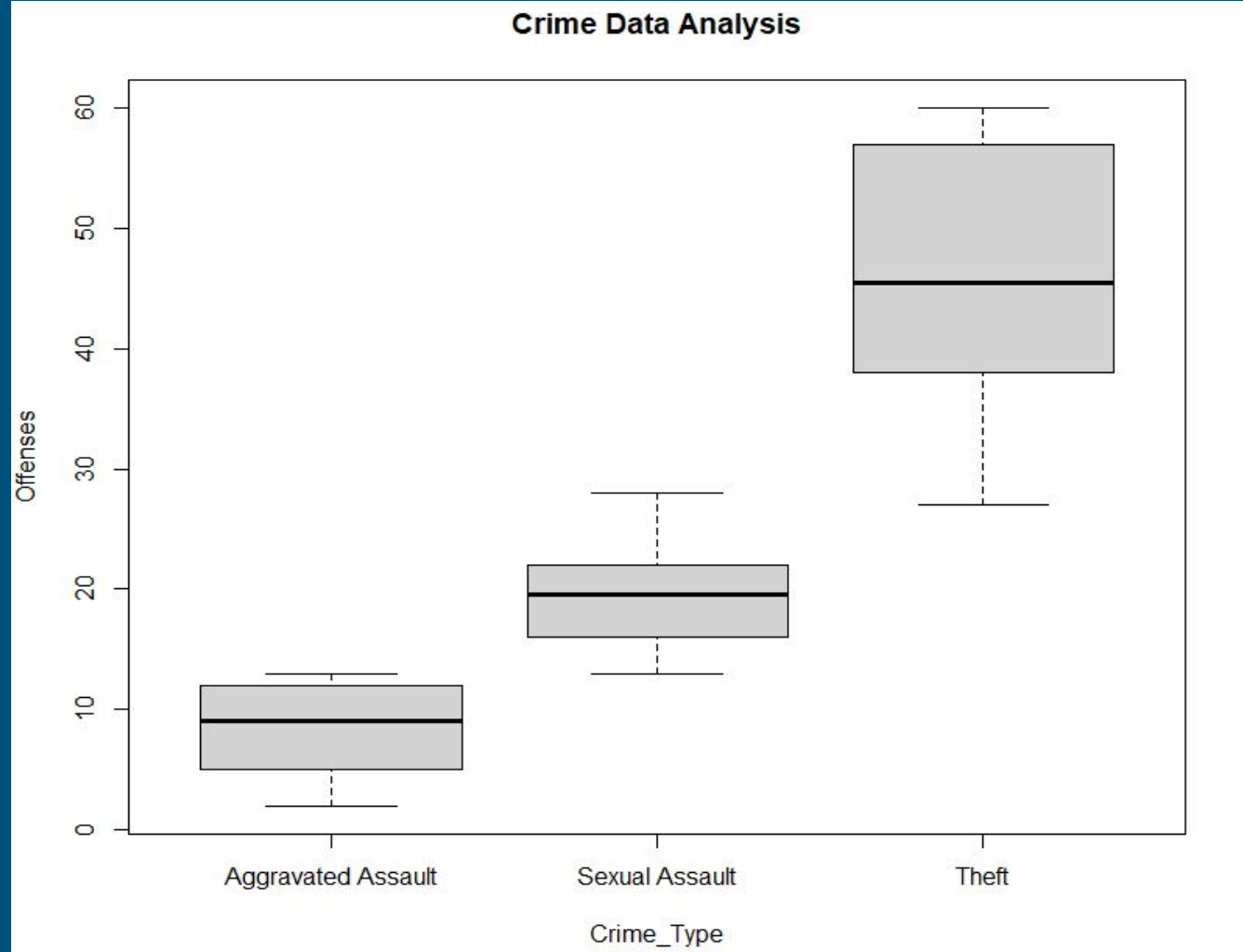
Data Visualization

Crime Data Analysis

Shapiro Wilk Test

```
> shapiro.test(Crime_data$'Sexual Assault')

        Shapiro-Wilk normality test

data:   Crime_data$"Sexual Assault"
W = 0.97763, p-value = 0.9392

> shapiro.test(Crime_data$'Aggravated Assault')

        Shapiro-Wilk normality test

data:   Crime_data$"Aggravated Assault"
W = 0.86897, p-value = 0.2221

> shapiro.test(Crime_data$Theft)

        Shapiro-Wilk normality test

data:   Crime_data$Theft
W = 0.9643, p-value = 0.8522
```

Analytical Approach: *Anova Analysis* (Observational Study)

PARAMETERS

U1 = SA (*Sexual Assault*)

U2 = AA (*Aggravated Assault*)

U3 = Theft

NULL & ALT HYPOTHESIS

H0 = U1 = U2 = U3

Ha = Not all equal

Preliminary Results
Assumptions

- Level of Significance α = 0.05
- Each observation is independent of each other
- MSE: 66.94

```
> anova <- aov(Crime_Type ~ offenses, data = C_data)
> summary(anova)
            Df Sum Sq Mean Sq F value  Pr(>F)
offenses     2   4354  2177.2   32.52 3.51e-06 ***
Residuals   15   1004    66.9
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Pairwise Comparison (Bonferroni)

```
> pairwise.t.test(C_data$Crime_Type, C_data$Offenses, p.adj = "bonferroni")

        Pairwise comparisons using t tests with pooled SD

data:   C_data$Crime_Type and C_data$Offenses

              Aggravated Assault Sexual Assault
Sexual Assault 0.08962            -
Theft          3.2e-06            0.00019

P value adjustment method: bonferroni
```