# 560_proj

Group 1

12/4/2020

```
library(ggplot2)
```

```
Data <- read.csv("D:/D drive contents/Fall 2020/CS560_Data Driven/MSR2/Final/RQ2test_data.csv")
head(Data)
```

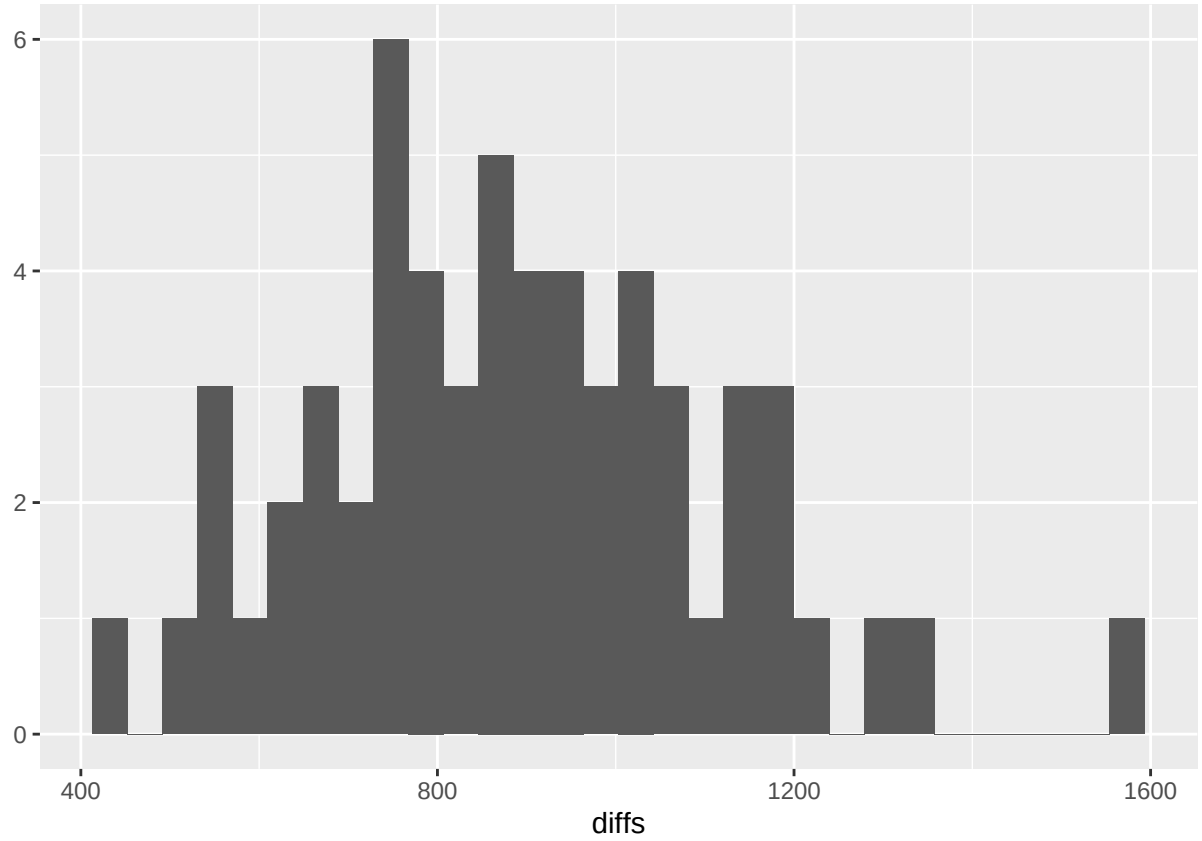
```
##   ï..Date Code Non_Code
## 1 2015-01  268      911
## 2 2015-02  433     1447
## 3 2015-03  228      878
## 4 2015-04  265     1150
## 5 2015-05  273     1219
## 6 2015-06  248     1822
```

```
diffs <- with(Data, Non_Code-Code)
diffs
```

```
##  [1]  643 1014  650  885  946 1574 1167 1224 1152 1337 1068 1039 1001  894 1109
## [16]  938  847  543  913 1200  937 1068  783  795 1034  673 1041  817  653  741
## [31]  857  758  759  981  778  594  710  779  764  882 1173 1044  914  936  871
## [46] 1317  762  433  766 1148  829  978  921  513  618  823  538 1137  690  564
```

```
qplot(diffs, geom="histogram")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
t.test(diffs)
```

```
##
##  One Sample t-test
##
## data:  diffs
## t = 30.689, df = 59, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  833.8871 950.2129
## sample estimates:
## mean of x
##    892.05
```