

## exercise5

Mohammad Imtiaz Nur

3/6/2021

ID: 1878074

```
soilData <- read.csv("SiltDataset.csv")

#ANOVA test
siteFactor <- factor(soilData$site)
analysis <- aov(soilData$silt ~ siteFactor)
summary(analysis)
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## siteFactor    7   600.1    85.73    3.432 0.00293 **
## Residuals   80  1998.4    24.98
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Since the p-value 0.00293 is less than 0.05, the difference in means is statistically significant. Hence there is a difference in silt content among the soils from different sites.

```
#boxplot
boxplot(soilData$silt~siteFactor,xlab="site",
        ylab="Silt",
        main="Boxplots of silt vs site variety")
means <- tapply(soilData$silt, siteFactor, mean)
points( means, pch=8, col="red")
```

**Boxplots of silt vs site variety**

