FinalProject

In our project, we are comparing the relative stats and draft order of players in a certain positional side (Offense vs Defense) and how it relates to the money they make in their first contract. For each side of the ball we have at 2 statistics (draft order and average salary) that we will compare against first contract earned to see if there is any relation between draft order or performance that relates to earning a contract. In total, we have 6 positional groups (3 on each side of the ball) to analyze, and we have pulled data from draft years 2010 to 2013. In our final write-up we will go more in depth by position and analyze their skill statistics versus their average salaries after their rookie year.

```
#setwd("~/Spring/Stat/Final")
setwd("~/Desktop/BokeyZhangFinalR")
QB_Data <- read.csv("QB_Data.csv")
RB_Data <- read.csv("RB_Data.csv")
WR_Data <- read.csv("WR_Data.csv")
DL_Data <- read.csv("DL_Data.csv")
LB_Data <- read.csv("LB_Data.csv")
DB_Data <- read.csv("DB_Data.csv")
Offense <- read.csv("Offense.csv")
Defense <- read.csv("Defense.csv")
```

```
Mode <- function(x) {
  ux <- unique(x)
  ux[which.max(tabulate(match(x, ux)))]
}

FindOutliers <- function(x) {
  lowerq = quantile(x)[2]
  upperq = quantile(x)[4]
  iqr = upperq - lowerq
  extreme.threshold.upper = (iqr * 3) + upperq
  extreme.threshold.lower = lowerq - (iqr * 3)
  result <- which(x > extreme.threshold.upper | x < extreme.threshold.lower)
}
```

The first positional side we are discussing will be the Offensive side (3 positions - QB, RB, and WR)

```
OffenseDraftOrders <- Offense$Draft.Order
OffenseSalary <- Offense$Avg.Salary..million.
```

The measures of central tendencies of Offensive salaries are as follows (in the order: Mean, Median, Mode)

```
AvgOffenseSalary <- mean(OffenseSalary)
MedianOffenseSalary <- median(OffenseSalary)
ModeOffenseSalary <- Mode(OffenseSalary)
AvgOffenseSalary
```

```
## [1] 6.045944
```

```
MedianOffenseSalary
```

```
## [1] 1.8935
```

```
ModeOffenseSalary
```

```
## [1] 8
```

The measures of variation for Offensive salaries are as follows (in the order: Range, Variation, Standard Deviation)

```
RangeOffenseSalary <- range(OffenseSalary)
VarianceOffenseSalary <- var(OffenseSalary)
StdOffenseSalary <- sqrt(VarianceOffenseSalary)
RangeOffenseSalary
```

```
## [1]  0.116 28.000
```

```
VarianceOffenseSalary
```

```
## [1] 58.10113
```

```
StdOffenseSalary
```

```
## [1] 7.62241
```

The Correlation test between Offensive Draft Position and Offensive Salary is as follows:

```
cor.test(OffenseDraftOrders, OffenseSalary)
```

```
##
##  Pearson's product-moment correlation
##
## data:  OffenseDraftOrders and OffenseSalary
## t = -0.29752, df = 16, p-value = 0.7699
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.5229353  0.4067825
## sample estimates:
##         cor
## -0.07417432
```

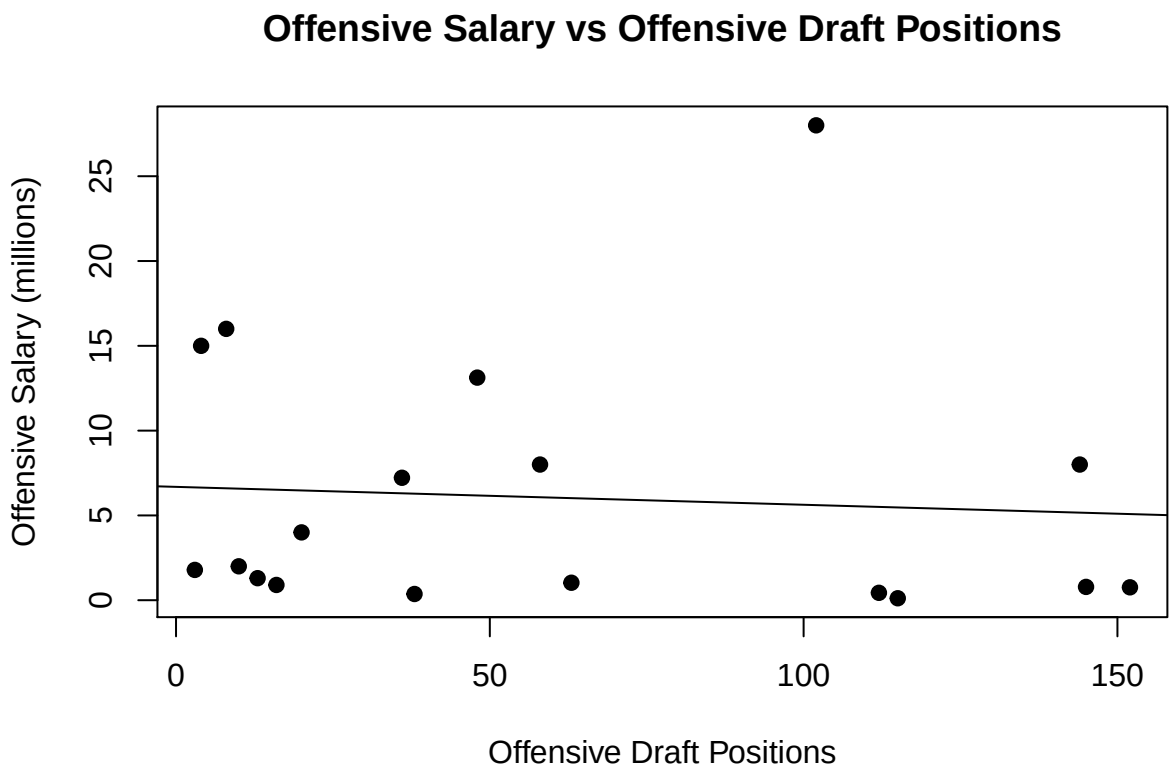
The linear regression between Offensive Draft Position and Offensive Salary is as follows:

```
lm(OffenseSalary ~ OffenseDraftOrders)
```

```
##
## Call:
## lm(formula = OffenseSalary ~ OffenseDraftOrders)
##
## Coefficients:
##        (Intercept)  OffenseDraftOrders
##            6.68330            -0.01055
```

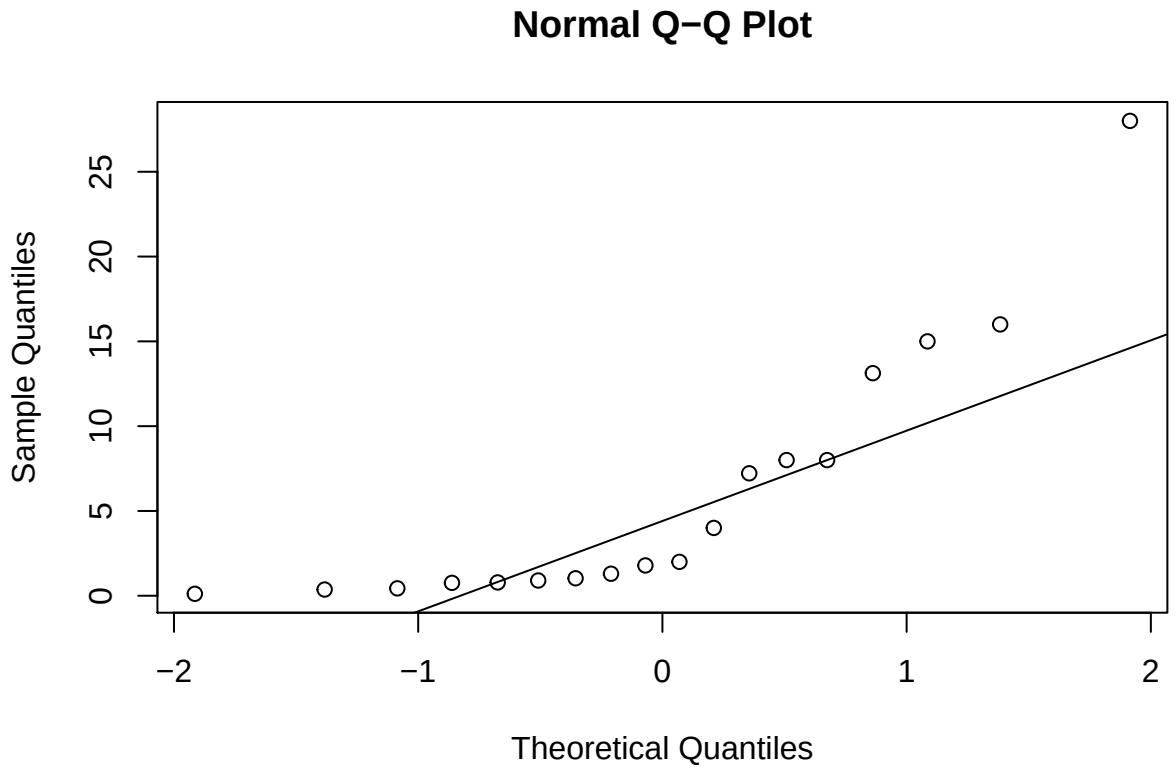
Below is a scatterplot between the draft order and average salary of Offensive players and the Correlation coefficient

```
plot(OffenseDraftOrders, OffenseSalary, main="Offensive Salary vs Offensive Draft Positions",
   xlab="Offensive Draft Positions", ylab="Offensive Salary (millions)", pch=19)
abline(lm(OffenseSalary ~ OffenseDraftOrders))
```



Below is a normality plot for Offensive Salary to determine if the sample is drawn from a normal distribution

```
qqnorm(OffenseSalary)
qqline(OffenseSalary)
```



Outlier Check (The outliers of the Offense Salaries are listed below)

```
OffenseOutliers <- FindOutliers(OffenseSalary)
OffenseOutliers
```

```
## integer(0)
```

As you can see there are no outliers.

In the offense, we are combining 3 different positions; however, when we compare each position we must also consider skills independent to the position itself - not the entire offensive side of the ball. These statistics are not part of the correlation test, they are just for the readers reference as they will be used in arguments in the final project. For each QB: QBR RB: Average Touchdowns Scored WR: Average Touchdowns Scored

QBs

```
QB_Qbr <- QB_Data$QBR
mean(QB_Qbr)
```

```
## [1] 74.255
```

RBs

```
RB_TDs <- RB_Data$Touchdowns
mean(RB_TDs)
```

```
## [1] 2.291667
```

WRs

```
WR_TDs <- WR_Data$Touchdowns
mean(WR_TDs)
```

```
## [1] 5.975
```

The second positional side we are discussing will be the Defensive side (3 positions - DB, DL, and LB)

```
DefensiveDraftOrders <- Defense$Draft.Order
DefenseSalary <- Defense$Avg.Salary
```

The measures of central tendencies of Defensive salaries are as follows (in the order: Mean, Median, Mode)

```
AvgDefenseSalary <- mean(DefenseSalary)
MedianDefenseSalary <- median(DefenseSalary)
ModeDefenseSalary <- Mode(DefenseSalary)
AvgDefenseSalary
```

```
## [1] 6.225158
```

```
MedianDefenseSalary
```

```
## [1] 5.6
```

```
ModeDefenseSalary
```

```
## [1] 0.73
```

The measures of variation for Defensive salaries are as follows (in the order: Range, Variation, Standard Deviation)

```
RangeDefenseSalary <- range(DefenseSalary)
VarianceDefenseSalary <- var(DefenseSalary)
StdDefenseSalary <- sqrt(VarianceDefenseSalary)
RangeDefenseSalary
```

```
## [1]  0.346 19.000
```

```
VarianceDefenseSalary
```

```
## [1] 36.82356
```

```
StdDefenseSalary
```

```
## [1] 6.068242
```

The Correlation test between Defensive Draft Position and Defensive Salary is as follows:

```
cor.test(DefensiveDraftOrders, DefenseSalary)
```

```
##
##  Pearson's product-moment correlation
##
## data:  DefensiveDraftOrders and DefenseSalary
## t = -2.1415, df = 17, p-value = 0.04701
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.756713055 -0.008496887
## sample estimates:
##        cor
## -0.4609273
```

The linear regression between Defensive Draft Position and Defensive Salary is as follows:

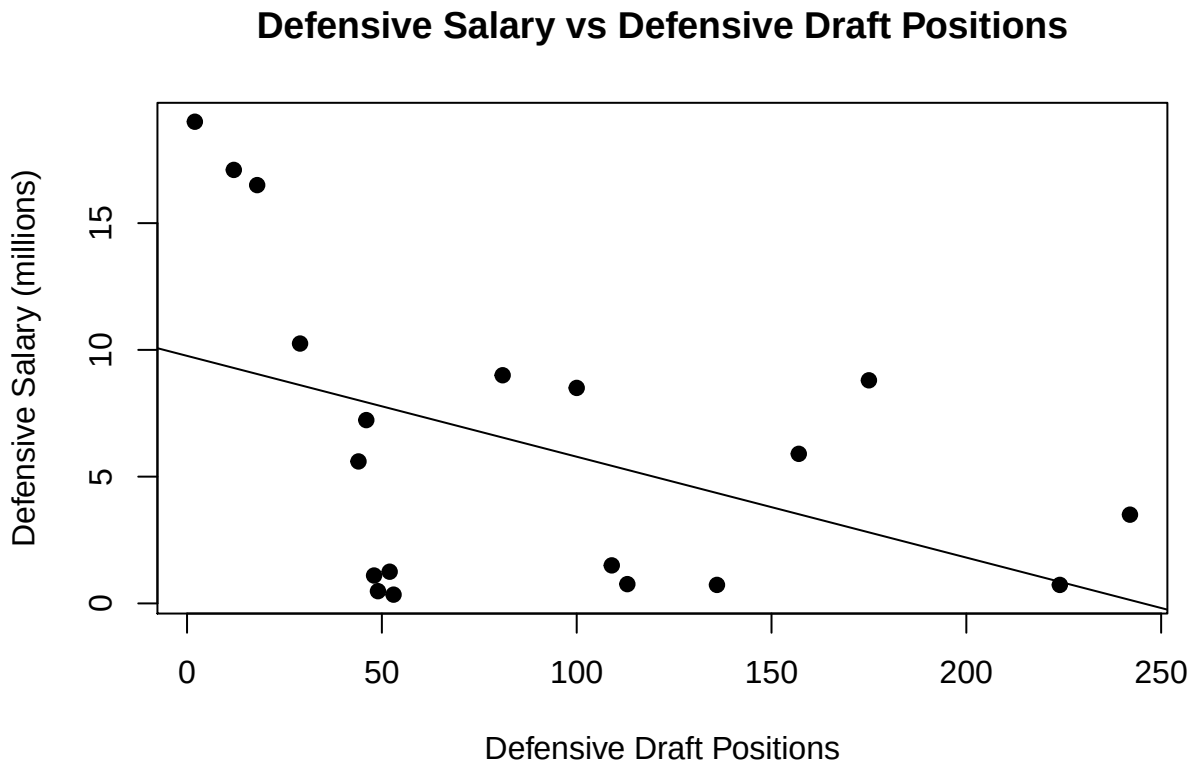
```
lm(DefenseSalary ~ DefensiveDraftOrders)
```

```
##
## Call:
```

```
## lm(formula = DefenseSalary ~ DefensiveDraftOrders)
##
## Coefficients:
##          (Intercept)  DefensiveDraftOrders
##              9.76359              -0.03978
```

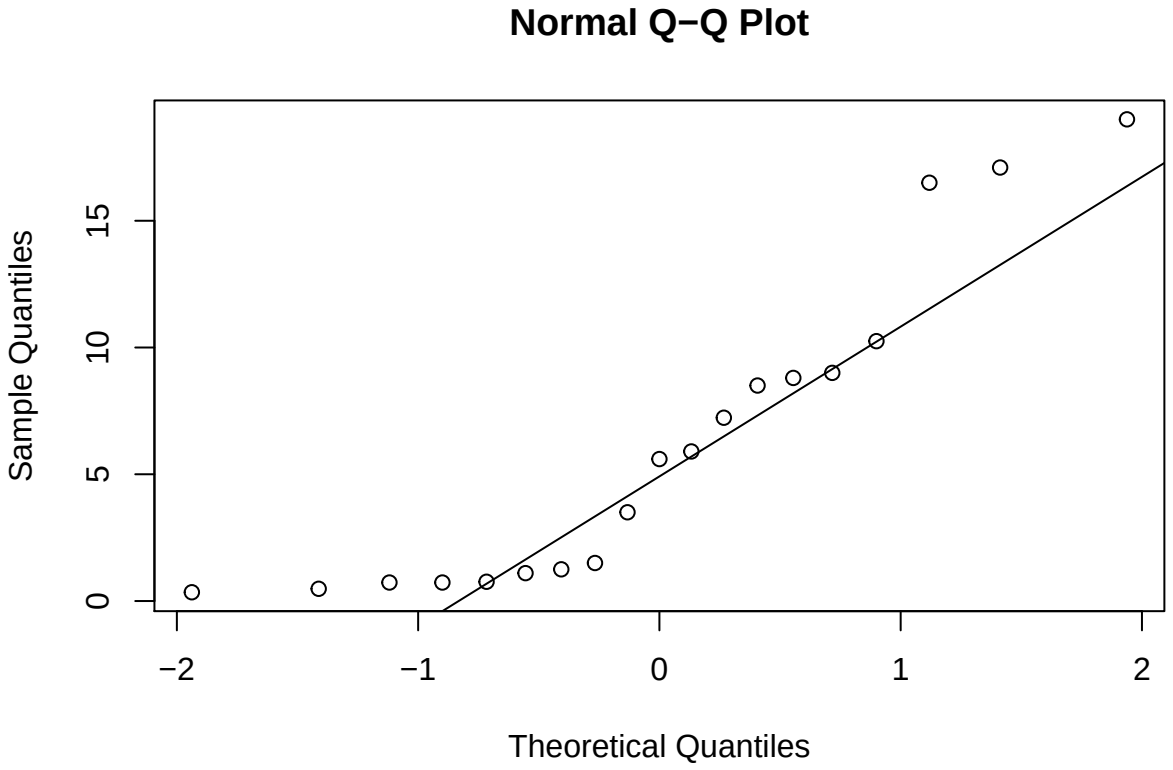
Below is a scatterplot between the draft order and average salary of Defensive players and the Correlation coefficient

```
plot(DefensiveDraftOrders, DefenseSalary, main="Defensive Salary vs Defensive Draft Positions",
   xlab="Defensive Draft Positions", ylab="Defensive Salary (millions)", pch=19)
abline(lm(DefenseSalary ~ DefensiveDraftOrders))
```



Below is a normality plot for Defensive Salary to determine if the sample is drawn from a normal distribution

```
qqnorm(DefenseSalary)
qqline(DefenseSalary)
```



Outlier Check (The outliers of the Defensive Salaries are listed below)

```
DefensiveOutliers <- FindOutliers(DefenseSalary)
DefensiveOutliers
```

```
## integer(0)
```

As you can see there are no outliers.

Just like the offense, we also want to list the descriptive statistics for some data points for each individual positional group on the defensive side of the ball.

DB: Interceptions (INTS) LB: Combined Tackles and Sacks DL: Combined Tackles and Sacks

DBs (Interceptions)

```
MeanDBInts <- mean(DB_Data$Ints)
MeanDBInts
```

```
## [1] 1.166667
```

DL (Tackles and Sacks)

```
MeanDLTackles <- mean(DL_Data$Tackles)
MeanDLSacks <- mean(DL_Data$Sacks)
MeanDLTackles
```

```
## [1] 34.83286
```

```
MeanDLSacks
```

```
## [1] 4.542143
```

LB (Tackles and Sacks)

```r
MeanLBTackles <- mean(LB_Data$Tackles)
MeanLBSacks <- mean(LB_Data$Sacks)
MeanLBTackles
```

```
## [1] 42.175
```

```r
MeanLBSacks
```

```
## [1] 3.691667
```