Data Exploration

```
batters <- read.csv("Batters_War.csv")
win_percentage <- read.csv("Team_Win%.csv")
team_batters <- read.csv("Team_Total_Batter_War.csv")
```

```
names <- c("ATL", "HOU", "TBR", "PHI", "LAD", "TOR", "MIN", "TEX", "MIL", "SEA",
           "SDP", "BAL", "ARI", "NYY", "CHC", "BOS", "MIA", "CLE", "NYM", "SFG",
           "STL", "DET", "CIN", "PIT", "OAK", "KCR", "LAA", "CHW", "WSN", "COL")
win_percentage$names <- names
```

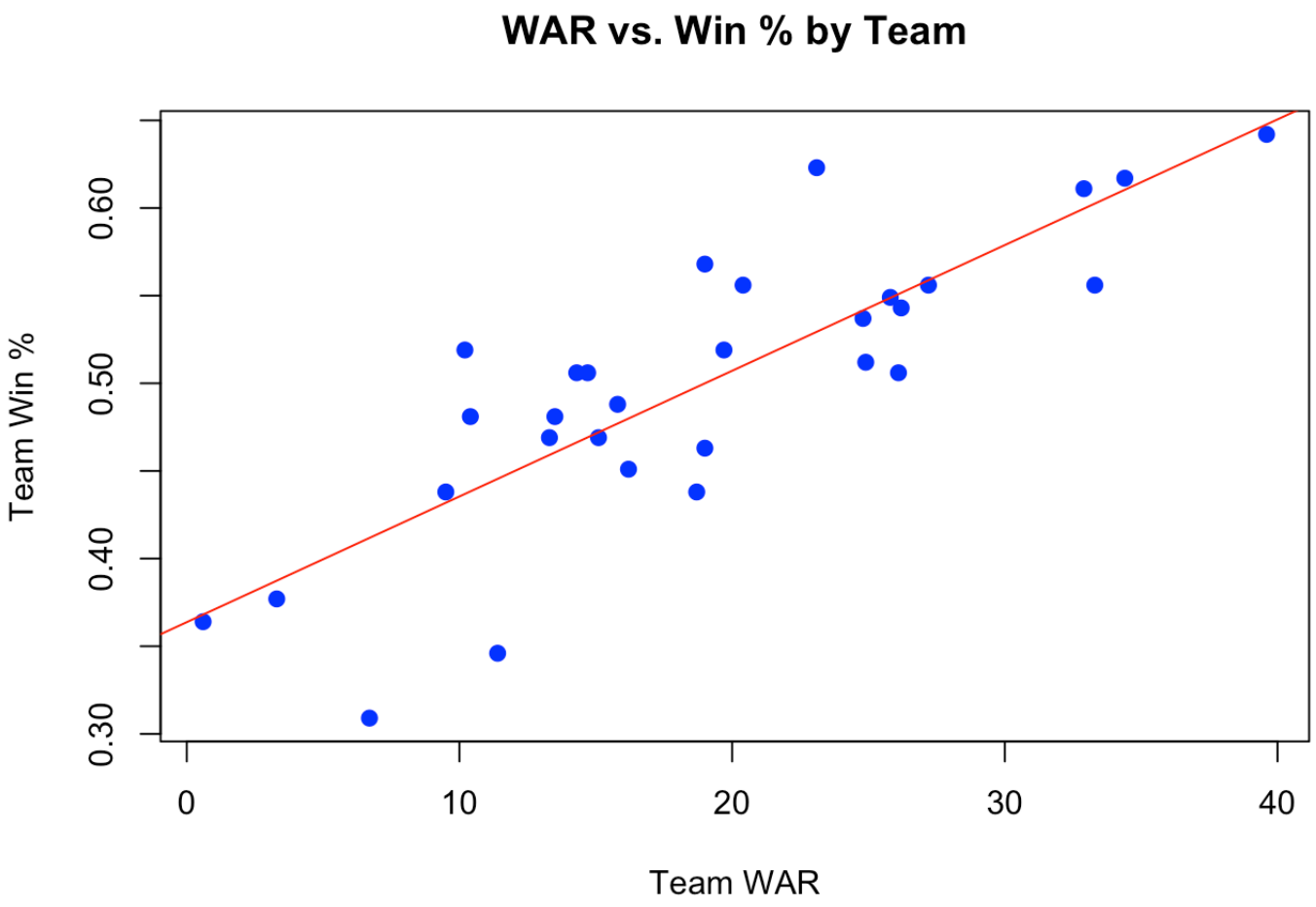
```
# sort(unique(batters$Team))
```

```
win_percent_name_ordered <- win_percentage[order(win_percentage$names),]$Win.
team_war_name_ordered <- team_batters[order(team_batters$Team),]$WAR
```

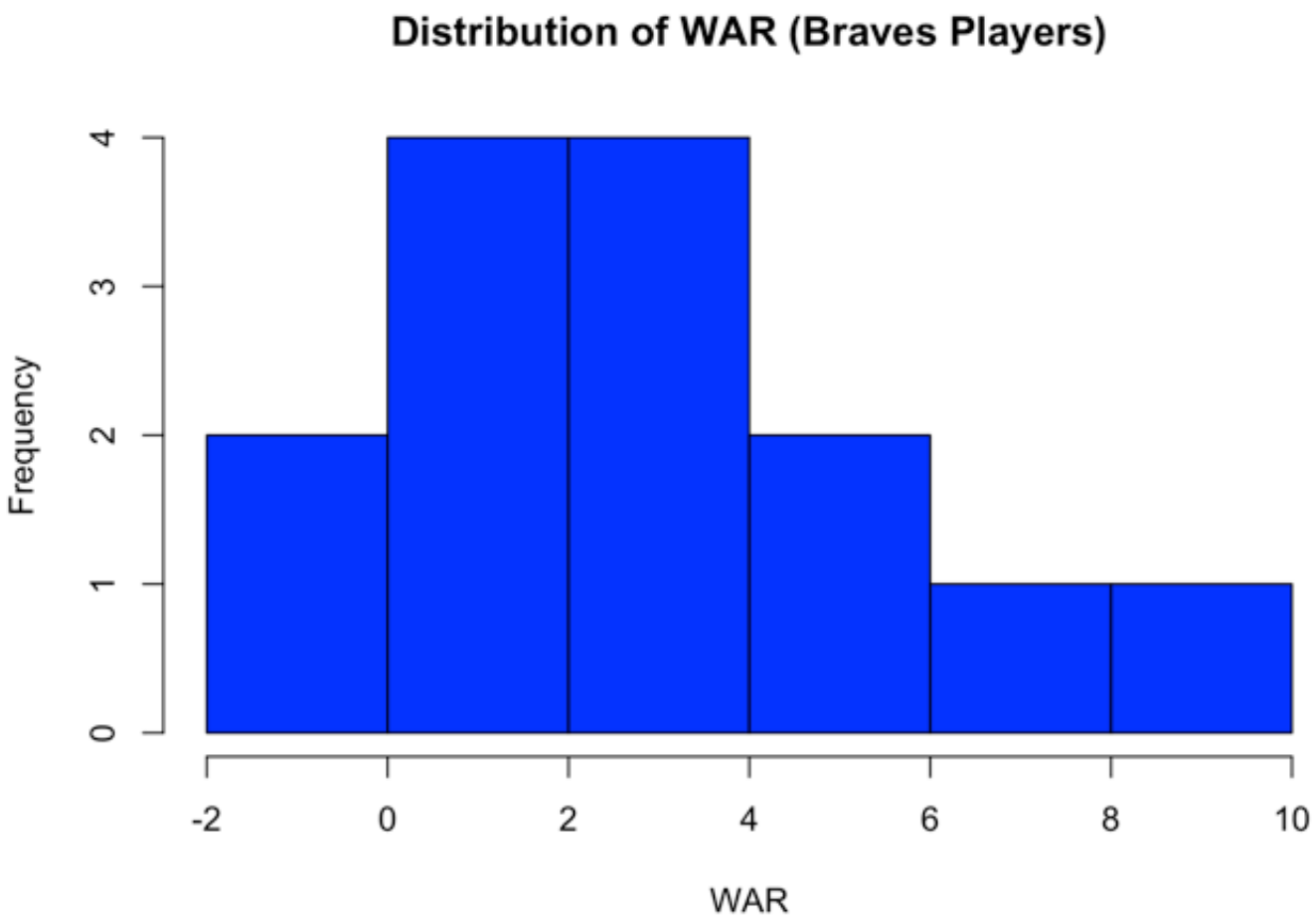
```
war_vs_wins <- lm(win_percent_name_ordered ~ team_war_name_ordered)
cor(win_percent_name_ordered, team_war_name_ordered)
```

```
## [1] 0.8282495
```

```
plot(team_war_name_ordered, win_percent_name_ordered,
     xlab = 'Team WAR', ylab = 'Team Win %',
     main = 'WAR vs. Win % by Team',
     col = 'blue',
     pch = 19)
abline(war_vs_wins, col = "red")
```




```
hist(batters[batters$Team == 'ATL' & batters$PA > 50,]$fWAR, xlab = 'WAR', main = "Di
stribution of WAR (Braves Players)", col = "blue")
```


Distribution of WAR (Braves Players)
Frequency
0
1
2
3
4
-2
0
2
4
6
8
10
WAR