```r
# Read in data
baseball = read.csv("baseball.csv")
str(baseball)

# Subset to only include moneyball years
moneyball = subset(baseball, Year < 2002)
str(moneyball)

# Compute Run Difference
moneyball$RD = moneyball$RS - moneyball$RA
str(moneyball)

# Scatterplot to check for linear relationship
plot(moneyball$RD, moneyball$W)

# Regression model to predict wins
WinsReg = lm(W ~ RD, data=moneyball)
summary(WinsReg)


str(moneyball)

# Regression model to predict runs scored
RunsReg = lm(RS ~ OBP + SLG + BA, data=moneyball)
summary(RunsReg)

RunsReg = lm(RS ~ OBP + SLG, data=moneyball)
summary(RunsReg)
```