

DATA 605 Discussion Week 11 (12)

CUNY Spring 2021

Philip Tanofsky

13 April 2021

```
library(tidyverse)
library(stats)
library(broom)
library(ggplot2)

nba_data <- read_csv("https://raw.githubusercontent.com/fivethirtyeight/nba-player-advanced-metrics/master/data/players/players.csv")

nba_data <- nba_data %>% filter(year_id >= 2016) %>% select( 'P/36', 'A/36', 'R/36', 'Raptor 0')

nba_data <- rename_with(nba_data, ~ gsub(" ", "_", .x, fixed = TRUE))
nba_data <- rename_with(nba_data, ~ gsub("/", "_", .x, fixed = TRUE))

nba_data$P_36 <- as.numeric(nba_data$P_36)
nba_data$A_36 <- as.numeric(nba_data$A_36)
nba_data$R_36 <- as.numeric(nba_data$R_36)

nba_data <- na.omit(nba_data)

nba_data <- nba_data %>% filter(P_36 <= 50)

dim(nba_data)
```

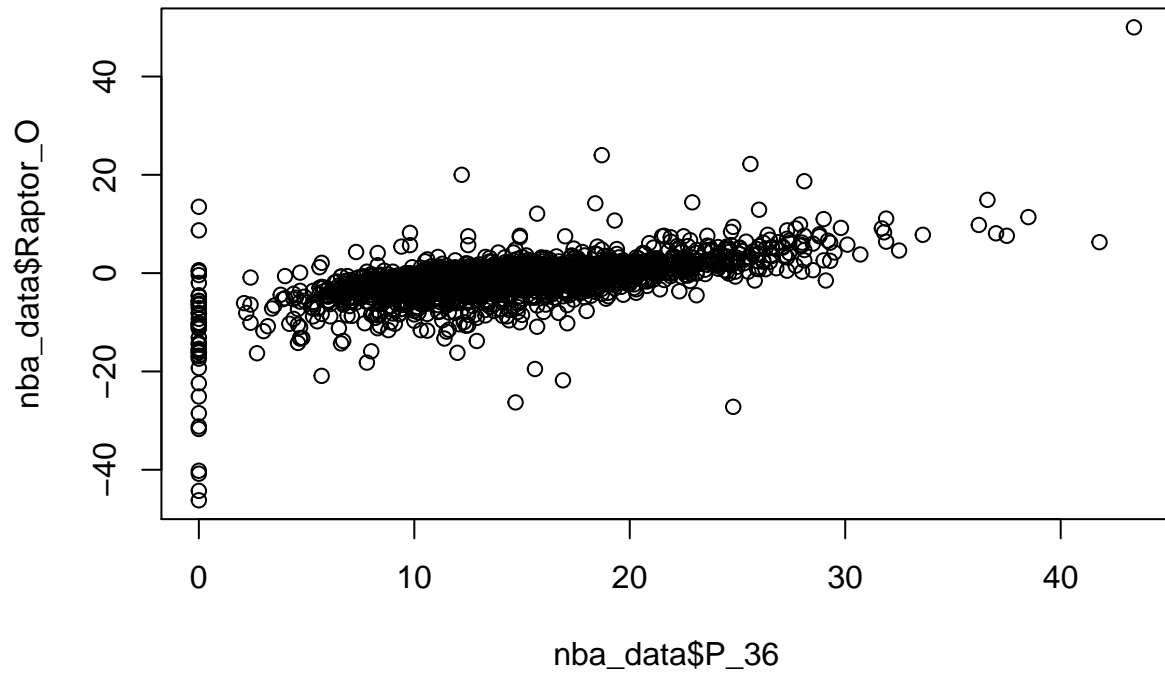
```
## [1] 2294    4
```

```
summary(nba_data)
```

| ## | P_36 | A_36 | R_36 | Raptor_0 |
|-------------|--------|----------------|----------------|-----------------|
| ## Min. | : 0.00 | Min. : 0.000 | Min. : 0.000 | Min. : -46.200 |
| ## 1st Qu.: | 11.20 | 1st Qu.: 1.700 | 1st Qu.: 4.200 | 1st Qu.: -2.800 |
| ## Median | :14.10 | Median : 2.500 | Median : 5.900 | Median : -1.100 |
| ## Mean | :14.45 | Mean : 3.156 | Mean : 6.721 | Mean : -1.338 |
| ## 3rd Qu.: | 17.40 | 3rd Qu.: 4.200 | 3rd Qu.: 8.700 | 3rd Qu.: 0.500 |
| ## Max. | :43.40 | Max. :12.800 | Max. :37.000 | Max. : 50.000 |

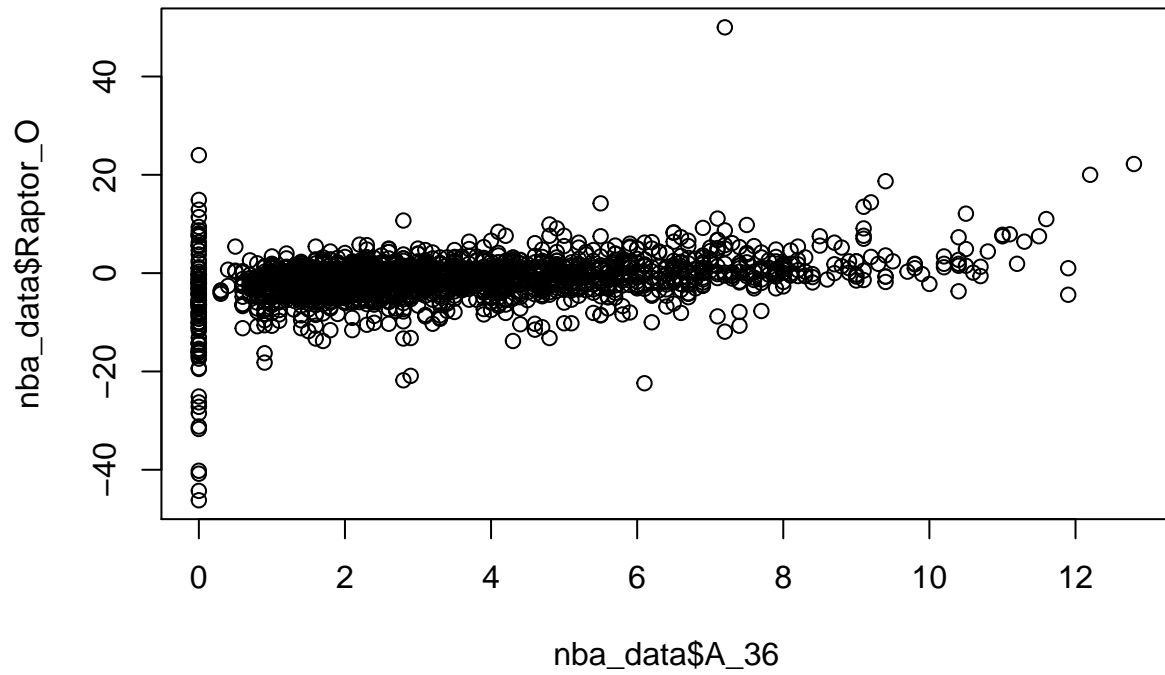
```
plot(x = nba_data$P_36, y = nba_data$Raptor_0, main = "Raptor Offensive Rating by Points per 36")
```

Raptor Offensive Rating by Points per 36



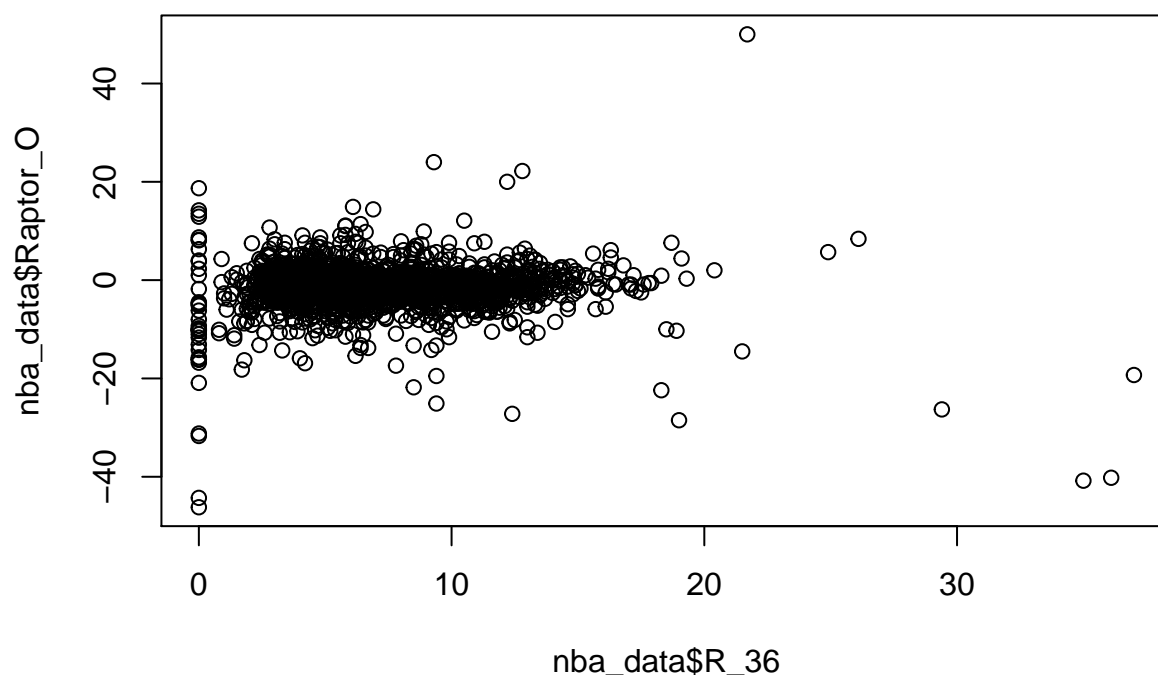
```
plot(x = nba_data$A_36, y = nba_data$Raptor_O, main = "Raptor Offensive Rating by Assists per 36")
```

Raptor Offensive Rating by Assists per 36



```
plot(x = nba_data$R_36, y = nba_data$Raptor_O, main = "Raptor Offensive Rating by Rebounds per 36")
```

Raptor Offensive Rating by Rebounds per 36



```
model <- lm(Raptor_0 ~ ., data=nba_data)
#model
summary(model)
```

```
##
## Call:
## lm(formula = Raptor_0 ~ ., data = nba_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -37.355  -1.165   0.252   1.581  37.252
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -8.84546    0.25165  -35.150  < 2e-16 ***
## P_36           0.46499    0.01367   34.004  < 2e-16 ***
## A_36           0.37830    0.03481   10.866  < 2e-16 ***
## R_36          -0.06041    0.02045   -2.954  0.00316 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.395 on 2290 degrees of freedom
## Multiple R-squared:  0.4093, Adjusted R-squared:  0.4086
## F-statistic:  529 on 3 and 2290 DF, p-value: < 2.2e-16
```

```
model.diag.metrics <- augment(model)
head(model.diag.metrics)
```

```
## # A tibble: 6 x 10
##   Raptor_0 P_36 A_36 R_36 .fitted .resid .std.resid   .hat .sigma   .cooksd
##   <dbl> <dbl> <dbl> <dbl>   <dbl> <dbl>   <dbl>   <dbl> <dbl>   <dbl>
## 1     2.4  21.4   9     4.2   4.26  -1.86   -0.548  0.00379   3.40  0.000285
## 2     0.4  12.3   3.6    4.1  -2.01   2.41    0.711  0.000719   3.40  0.0000908
## 3    -3.2  12.3   3.6    6.4  -2.15  -1.05   -0.309  0.000546   3.40  0.0000131
## 4     3.2  22.5   2.3   11.3   1.80   1.40    0.411  0.00206   3.40  0.0000873
## 5    -2.9  16.4   3.2    5.5  -0.341 -2.56   -0.754  0.000566   3.40  0.0000805
## 6    -0.9  14.9   1.9   10.3  -1.82   0.920    0.271  0.000916   3.40  0.0000169
```

```
#autoplot(model)
```