Predicting NFL Statistic

Rich L

March 27, 2017

Predicting NFL Stats

The goal of this capstone project was to set a baseline linear regression for predicting NFL statistics. The use of the analysis would be to project player performance and see if the team I am working with needs to consider making adjustments given various factors of the upcoming game/season.

Where to get the data

I went to the website http://armchairanalysis.com/data.php (www.armchairanlysis.com). I have a subscription to the database, so I connected into it via SQL. I downloaded the historical database onto my hard drive, and mapped it in MySQL.

I then queried the DB to get the fields I would need. This operation took extensive time, so once it ran, I exported to a csv file, then read the csv into R.

```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 3.3.2
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(tidyr)
```

```
## Warning: package 'tidyr' was built under R version 3.3.2
```

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.3.2
```

```
nfl_data <- read.csv("NFL_offense.csv")
```

The data is pretty clean from armchairanalyis, fivethirtyeight.com (www.fivethirtyeight.com) uses this website for its sports data, so it is a pretty reputable site.

I felt there were pieces of data either missing, or needing cleaning up. This brought on the fun process of cleaning and tidying the data.

Weather and field conditions

From a qualitative perspective, we know that field turf and ideal temperatures are the least inhibitive towards speed, according to players themselves. I wanted to identify extremes and hinderences.

I replaced all “NULL” temp fields with a generic “room” temperature assumption.

cold_weather and hot_weather were fields created to identify extreme ends of the temperature spectrum, and see if they have an impact on play

```
#make all null temperatures at game time "room" temperature
nfl_data$temp[nfl_data$temp == "NULL"] <- 70
nfl_data$temp <- as.integer(nfl_data$temp)

#highlight temp extremes
nfl_data <- mutate(nfl_data, cold_weather= ifelse(temp < 45, 1,0))
nfl_data <- mutate(nfl_data, hot_weather=  ifelse(temp > 85, 1,0))

#weather factors
nfl_data <- mutate(nfl_data, grass_1 = ifelse(surf == "DD GrassMaster" | surf =
= "Grass",
                                              1,0))
nfl_data <- mutate(nfl_data, bad_weather_1 = ifelse(cond == "Light Rain" |
                                                    cond == "Rain" |
                                                    cond == "Flurries" |
                                                    cond == "Snow" |
                                                    cond == "Foggy" |
                                                    cond == "Windy" |
                                                    cond == "Hazy" |
                                                    cond == "Thunderstorms"|
                                                    cond == "Light Snow" |
                                                    cond == "Light Showers" ,
1,0))
```

Home field advantage

Do players play better at home?

```
#identify home team
nfl_data$h <- as.character(nfl_data$h)
nfl_data$team <- as.character(nfl_data$team)
nfl_data <- mutate(nfl_data, home_team_1=  ifelse(h == team, 1,0))
```

Positions

Ignoring player stats, does the position matter

```
#identify position
nfl_data <- mutate(nfl_data, is_WR =  ifelse(pos1 == "WR", 1,0))
nfl_data <- mutate(nfl_data, is_TE =  ifelse(pos1 == "TE", 1,0))
nfl_data <- mutate(nfl_data, is_RB =  ifelse(pos1 == "RB", 1,0))
nfl_data <- mutate(nfl_data, is_QB =  ifelse(pos1 == "QB", 1,0))
```

Age

Every year players get older, so we want to know “Does father time impact player performance?”

```
#age
nfl_data <- mutate(nfl_data, age = year - yob)
```

Combine cleanup

The NFL combine is an event where prospective new players work out for the entire league to see. Their physical measurements are taken, and people find merit in this event. I wanted to see if these stats had any impact on player performance. Not all players attend the combine. For the fields where there are zeroes for the combine stat, I took the average for all non-zero stats for that position. This basically implies if you didn’t attend the combine, your stats are middle of the road.

```
#replace 0 forty with avg for position
nfl_data <- nfl_data %>%
  group_by(pos1)%>%
  mutate(forty1 = ifelse(forty == 0, mean(forty[forty>0]), forty))
#replace 0 vertical with average for position
nfl_data <- nfl_data %>%
  group_by(pos1)%>%
  mutate(vertical1 = ifelse(vertical == 0, mean(vertical[vertical>0]), vertica
l))
#replace 0 arm length with formula for 40% of height is arm
nfl_data$arm <- ifelse(nfl_data$arm == 0, nfl_data$height*0.4, nfl_data$arm)
nfl_data <- nfl_data %>%
  group_by(pos1)%>%
  mutate(shuttle1 = ifelse(shuttle == 0, mean(shuttle[shuttle>0]), shuttle))
nfl_data <- nfl_data %>%
  group_by(pos1)%>%
  mutate(cone1 = ifelse(cone == 0, mean(cone[cone>0]), cone))
```

NFL Teams

I created fields for teams (1 if player plays for that team in the header 0 if it doesn't). I also cleaned up one team: The St Louis/LA Rams. The Rams moved in 2016 to LA, so the conditions of stadium changed. I combined the field into a single field.

```
#clean teams and give each team a field
nfl_data <- mutate(nfl_data, Teams =  ifelse(team == "STL" | team == "LA", "ST
L/LA",team))
nfl_data <- mutate(nfl_data, ARI = ifelse(Teams == "ARI",1,0))
nfl_data <- mutate(nfl_data, ATL = ifelse(Teams == "ATL",1,0))
nfl_data <- mutate(nfl_data, BAL = ifelse(Teams == "BAL",1,0))
nfl_data <- mutate(nfl_data, BUF = ifelse(Teams == "BUF",1,0))
nfl_data <- mutate(nfl_data, CAR = ifelse(Teams == "CAR",1,0))
nfl_data <- mutate(nfl_data, CHI = ifelse(Teams == "CHI",1,0))
nfl_data <- mutate(nfl_data, CIN = ifelse(Teams == "CIN",1,0))
nfl_data <- mutate(nfl_data, CLE = ifelse(Teams == "CLE",1,0))
nfl_data <- mutate(nfl_data, DAL = ifelse(Teams == "DAL",1,0))
nfl_data <- mutate(nfl_data, DEN = ifelse(Teams == "DEN",1,0))
nfl_data <- mutate(nfl_data, DET = ifelse(Teams == "DET",1,0))
nfl_data <- mutate(nfl_data, GB = ifelse(Teams == "GB",1,0))
nfl_data <- mutate(nfl_data, HOU = ifelse(Teams == "HOU",1,0))
nfl_data <- mutate(nfl_data, IND = ifelse(Teams == "IND",1,0))
nfl_data <- mutate(nfl_data, JAC = ifelse(Teams == "JAC",1,0))
nfl_data <- mutate(nfl_data, KC = ifelse(Teams == "KC",1,0))
nfl_data <- mutate(nfl_data, MIA = ifelse(Teams == "MIA",1,0))
nfl_data <- mutate(nfl_data, MINN = ifelse(Teams == "MIN",1,0))
nfl_data <- mutate(nfl_data, NE = ifelse(Teams == "NE",1,0))
nfl_data <- mutate(nfl_data, NOR = ifelse(Teams == "NO",1,0))
nfl_data <- mutate(nfl_data, NYG = ifelse(Teams == "NYG",1,0))
nfl_data <- mutate(nfl_data, NYJ = ifelse(Teams == "NYJ",1,0))
nfl_data <- mutate(nfl_data, OAK = ifelse(Teams == "OAK",1,0))
nfl_data <- mutate(nfl_data, PHI = ifelse(Teams == "PHI",1,0))
nfl_data <- mutate(nfl_data, PIT = ifelse(Teams == "PIT",1,0))
nfl_data <- mutate(nfl_data, SD = ifelse(Teams == "SD",1,0))
nfl_data <- mutate(nfl_data, SEA = ifelse(Teams == "SEA",1,0))
nfl_data <- mutate(nfl_data, SF = ifelse(Teams == "SF",1,0))
nfl_data <- mutate(nfl_data, STL = ifelse(Teams == "STL/LA",1,0))
nfl_data <- mutate(nfl_data, TB = ifelse(Teams == "TB",1,0))
nfl_data <- mutate(nfl_data, TEN = ifelse(Teams == "TEN",1,0))
nfl_data <- mutate(nfl_data, WAS = ifelse(Teams == "WAS",1,0))
```

Receiving Stats

For receiving, I wanted to get every players average: * yards * receptions * targets * touchdowns

I also wanted to get every position average, and average for team. Rationale for at least having that info is this: Compare player to team to league wide position

```
#calculate the averages by player, position, and team
#receiving
nfl_data <- nfl_data %>%
              group_by(player.1)%>%
                mutate(avg_recy_plyr = mean(recy))
nfl_data <- nfl_data %>%
              group_by(pos1)%>%
                mutate(avg_recy_pos = mean(recy))
nfl_data <- nfl_data %>%
              group_by(Teams)%>%
                mutate(avg_recy_team = mean(recy))
nfl_data <- nfl_data %>%
              group_by(player.1)%>%
                mutate(avg_rec_plyr = mean(rec))
nfl_data <- nfl_data %>%
              group_by(pos1)%>%
                mutate(avg_rec_pos = mean(rec))
nfl_data <- nfl_data %>%
              group_by(Teams)%>%
                mutate(avg_rec_team = mean(rec))
nfl_data <- nfl_data %>%
              group_by(player.1)%>%
                mutate(avg_trg_plyr = mean(trg))
nfl_data <- nfl_data %>%
              group_by(pos1)%>%
                mutate(avg_trg_pos = mean(trg))
nfl_data <- nfl_data %>%
              group_by(Teams)%>%
                mutate(avg_trg_team = mean(trg))
nfl_data <- nfl_data %>%
              group_by(player.1)%>%
                mutate(avg_rectd_plyr = mean(tdrec))
nfl_data <- nfl_data %>%
              group_by(pos1)%>%
                mutate(avg_rectd_pos = mean(tdrec))
nfl_data <- nfl_data %>%
              group_by(Teams)%>%
                mutate(avg_rectd_team = mean(tdrec))
```

Running Stats

I followed a similar process from up above.
The stats I was looking for the mean for were: * rushing attempts * rushing yards * fumbles

```
#running
nfl_data <- nfl_data %>%
  group_by(player.1)%>%
  mutate(avg_rbra_plyr = mean(ra))

nfl_data <- nfl_data %>%
  group_by(Teams)%>%
  mutate(avg_rbra_team = mean(ra))

nfl_data <- nfl_data %>%
  group_by(pos1)%>%
  mutate(avg_rbra_pos = mean(ra))

nfl_data <- nfl_data %>%
  group_by(player.1)%>%
  mutate(avg_rbry_plyr = mean(ry))

nfl_data <- nfl_data %>%
  group_by(Teams)%>%
  mutate(avg_rbry_team = mean(ry))

nfl_data <- nfl_data %>%
  group_by(pos1)%>%
  mutate(avg_rbry_pos = mean(ry))

nfl_data <- nfl_data %>%
  group_by(player.1)%>%
  mutate(avg_fuml_plyr = mean(fuml))

nfl_data <- nfl_data %>%
  group_by(Teams)%>%
  mutate(avg_fuml_team = mean(fuml))

nfl_data <- nfl_data %>%
  group_by(pos1)%>%
  mutate(avg_fuml_pos = mean(fuml))

nfl_data <- nfl_data %>%
  group_by(player.1)%>%
  mutate(avg_tdr_plyr = mean(tdr))

nfl_data <- nfl_data %>%
  group_by(pos1)%>%
  mutate(avg_tdr_pos = mean(tdr))

nfl_data <- nfl_data %>%
  group_by(Teams)%>%
  mutate(avg_tdr_team = mean(tdr))
```

Passing

I followed a similar process from up above.
The stats I was looking for the mean for were: * passing yards * passing attempts * passing completions * passing touchdowns * interceptions

```
#passing
nfl_data <- nfl_data %>%
  group_by(player.1)%>%
  mutate(avg_qbpy_plyr = mean(py))

nfl_data <- nfl_data %>%
  group_by(Teams)%>%
  mutate(avg_qbpy_team = mean(py))

nfl_data <- nfl_data %>%
  group_by(pos1)%>%
  mutate(avg_qbpy_pos = mean(py))


nfl_data <- nfl_data %>%
  group_by(player.1)%>%
  mutate(avg_qbpc_plyr = mean(pc))

nfl_data <- nfl_data %>%
  group_by(Teams)%>%
  mutate(avg_qbpc_team = mean(pc))

nfl_data <- nfl_data %>%
  group_by(pos1)%>%
  mutate(avg_qbpc_pos = mean(pc))

nfl_data <- nfl_data %>%
  group_by(player.1)%>%
  mutate(avg_qbints_plyr = mean(ints))

nfl_data <- nfl_data %>%
  group_by(Teams)%>%
  mutate(avg_qbints_team = mean(ints))

nfl_data <- nfl_data %>%
  group_by(pos1)%>%
  mutate(avg_qbints_pos = mean(ints))

nfl_data <- nfl_data %>%
  group_by(player.1)%>%
  mutate(avg_qbtdp_plyr = mean(tdp))

nfl_data <- nfl_data %>%
  group_by(Teams)%>%
  mutate(avg_qbtdp_team = mean(tdp))

nfl_data <- nfl_data %>%
  group_by(pos1)%>%
```

```
  mutate(avg_qbtdp_pos = mean(tdp))

nfl_data <- nfl_data %>%
  group_by(player.1)%>%
  mutate(avg_qbpa_plyr = mean(pa))

nfl_data <- nfl_data %>%
  group_by(Teams)%>%
  mutate(avg_qbpa_team = mean(pa))

nfl_data <- nfl_data %>%
  group_by(pos1)%>%
  mutate(avg_qbpa_pos = mean(pa))
```

Age

There is an age component to a graph I show later, but here is the code right here

Receiving Regressions

**Receiving Yards first run

```
linRegrecy <- lm(recy ~ height+ weight + cold_weather + hot_weather + home_team
_1+ temp+ forty1 + vertical1  + shuttle1+ cone1 + ARI + ATL + BAL + BUF + CAR
+ CHI+CIN + CLE + DAL + DEN + DET + GB + HOU + IND + JAC + KC + MIA + MINN + N
E + NOR + NYG+
                    NYJ + OAK + PHI + PIT +SD + SEA + STL + TB + TEN + WAS + avg
_recy_plyr+avg_recy_pos +
                    avg_recy_team + avg_rec_plyr +avg_rec_pos + avg_rec_team +av
g_trg_plyr + avg_trg_pos +
                    avg_trg_team + avg_rectd_plyr + avg_rectd_pos +avg_rectd_tea
m+
                    avg_tdr_plyr + avg_tdr_pos + avg_tdr_team +
                    avg_rbra_plyr + avg_rbra_pos +avg_rbra_team +
                    avg_rbry_plyr + avg_rbry_pos +avg_rbry_team +
                    avg_fuml_plyr + avg_fuml_pos +avg_fuml_team +
                    avg_qbpy_plyr + avg_qbpy_pos +avg_qbpy_team +
                    avg_qbpa_plyr + avg_qbpa_pos +avg_qbpa_team+
                    avg_qbpc_plyr + avg_qbpc_pos +avg_qbpc_team +
                    avg_qbints_plyr + avg_qbints_pos +avg_qbints_team +
                    avg_qbtdp_plyr + avg_qbtdp_pos +avg_qbtdp_team +
                    grass_1 + bad_weather_1, data = nfl_data)

summary(linRegrecy)
```

```
##
## Call:
## lm(formula = recy ~ height + weight + cold_weather + hot_weather +
##     home_team_1 + temp + forty1 + vertical1 + shuttle1 + cone1 +
##     ARI + ATL + BAL + BUF + CAR + CHI + CIN + CLE + DAL + DEN +
##     DET + GB + HOU + IND + JAC + KC + MIA + MINN + NE + NOR +
##     NYG + NYJ + OAK + PHI + PIT + SD + SEA + STL + TB + TEN +
##     WAS + avg_recy_plyr + avg_recy_pos + avg_recy_team + avg_rec_plyr +
##     avg_rec_pos + avg_rec_team + avg_trg_plyr + avg_trg_pos +
##     avg_trg_team + avg_rectd_plyr + avg_rectd_pos + avg_rectd_team +
##     avg_tdr_plyr + avg_tdr_pos + avg_tdr_team + avg_rbra_plyr +
##     avg_rbra_pos + avg_rbra_team + avg_rbry_plyr + avg_rbry_pos +
##     avg_rbry_team + avg_fuml_plyr + avg_fuml_pos + avg_fuml_team +
##     avg_qbpy_plyr + avg_qbpy_pos + avg_qbpy_team + avg_qbpa_plyr +
##     avg_qbpa_pos + avg_qbpa_team + avg_qbpc_plyr + avg_qbpc_pos +
##     avg_qbpc_team + avg_qbints_plyr + avg_qbints_pos + avg_qbints_team +
##     avg_qbtdp_plyr + avg_qbtdp_pos + avg_qbtdp_team + grass_1 +
##     bad_weather_1, data = nfl_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -98.419 -12.000  -1.734   6.465 232.390
##
## Coefficients: (18 not defined because of singularities)
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       2.057970   9.933129   0.207 0.835868
## height           -0.038325   0.085704  -0.447 0.654750
## weight            0.002289   0.011435   0.200 0.841339
## cold_weather     -0.682754   0.517743  -1.319 0.187273
## hot_weather      -1.647985   2.045951  -0.805 0.420544
## home_team_1       0.218408   0.269888   0.809 0.418373
## temp              0.012260   0.014751   0.831 0.405896
## forty1           -0.210646   1.361048  -0.155 0.877005
## vertical1         0.006676   0.055595   0.120 0.904415
## shuttle1         -0.328503   1.174926  -0.280 0.779790
## cone1             0.052920   0.857337   0.062 0.950781
## ARI               1.926183   0.985713   1.954 0.050696 .
## ATL               2.645292   0.986540   2.681 0.007335 **
## BAL               2.808884   0.957663   2.933 0.003358 **
## BUF               2.583370   0.996958   2.591 0.009566 **
## CAR               1.921807   0.975734   1.970 0.048891 *
## CHI               2.658690   0.996948   2.667 0.007660 **
## CIN               2.431601   0.988469   2.460 0.013899 *
## CLE               3.255099   0.978635   3.326 0.000881 ***
## DAL               2.071710   0.998979   2.074 0.038102 *
## DEN               2.344790   0.961115   2.440 0.014706 *
## DET               1.950222   0.991697   1.967 0.049242 *
## GB                3.247170   0.966687   3.359 0.000783 ***
```

```
## HOU              2.989320   0.981071   3.047 0.002313 **
## IND              1.221505   0.980764   1.245 0.212969
## JAC              2.873811   0.993145   2.894 0.003810 **
## KC               2.628836   0.969487   2.712 0.006699 **
## MIA              1.336018   0.991161   1.348 0.177688
## MINN             1.791633   0.993568   1.803 0.071360 .
## NE               2.795808   0.975699   2.865 0.004166 **
## NOR              2.817817   0.980680   2.873 0.004064 **
## NYG              3.343864   0.981834   3.406 0.000661 ***
## NYJ              1.586744   0.982028   1.616 0.106149
## OAK              3.014011   0.967162   3.116 0.001832 **
## PHI              1.236107   0.983721   1.257 0.208920
## PIT              1.454012   0.972049   1.496 0.134708
## SD               3.029863   0.992367   3.053 0.002266 **
## SEA              0.587397   0.957559   0.613 0.539594
## STL              2.040074   0.990183   2.060 0.039376 *
## TB               2.489462   0.985848   2.525 0.011567 *
## TEN              2.507334   0.972806   2.577 0.009958 **
## WAS              2.353757   0.969766   2.427 0.015223 *
## avg_recy_plyr    1.017963   0.035248  28.880  < 2e-16 ***
## avg_recy_pos    -0.160468   4.531690  -0.035 0.971753
## avg_recy_team          NA         NA      NA       NA
## avg_rec_plyr    -0.179048   0.543808  -0.329 0.741969
## avg_rec_pos      0.042182  10.659064   0.004 0.996842
## avg_rec_team           NA         NA      NA       NA
## avg_trg_plyr    -0.008485   0.398804  -0.021 0.983026
## avg_trg_pos      1.490538  51.274041   0.029 0.976809
## avg_trg_team           NA         NA      NA       NA
## avg_rectd_plyr  -0.620218   1.884906  -0.329 0.742124
## avg_rectd_pos   -7.414988 234.910227  -0.032 0.974819
## avg_rectd_team         NA         NA      NA       NA
## avg_tdr_plyr    -0.033980   2.324859  -0.015 0.988339
## avg_tdr_pos      5.284057 114.995513   0.046 0.963350
## avg_tdr_team           NA         NA      NA       NA
## avg_rbra_plyr   -0.082981   0.224684  -0.369 0.711888
## avg_rbra_pos     0.017673   8.199564   0.002 0.998280
## avg_rbra_team          NA         NA      NA       NA
## avg_rbry_plyr    0.021557   0.052194   0.413 0.679592
## avg_rbry_pos    -0.082653   0.747991  -0.111 0.912013
## avg_rbry_team          NA         NA      NA       NA
## avg_fuml_plyr    0.098736   2.782721   0.035 0.971696
## avg_fuml_pos     1.045914  32.845700   0.032 0.974597
## avg_fuml_team          NA         NA      NA       NA
## avg_qbpy_plyr    0.017355   0.044076   0.394 0.693762
## avg_qbpy_pos           NA         NA      NA       NA
## avg_qbpy_team          NA         NA      NA       NA
## avg_qbpa_plyr    0.024448   0.343151   0.071 0.943204
## avg_qbpa_pos           NA         NA      NA       NA
## avg_qbpa_team          NA         NA      NA       NA
```

```
## avg_qbpc_plyr    -0.167068   0.619851  -0.270 0.787524
## avg_qbpc_pos            NA         NA      NA       NA
## avg_qbpc_team           NA         NA      NA       NA
## avg_qbints_plyr  -0.425143   1.984999  -0.214 0.830409
## avg_qbints_pos          NA         NA      NA       NA
## avg_qbints_team         NA         NA      NA       NA
## avg_qbtdp_plyr   -0.874542   2.173354  -0.402 0.687397
## avg_qbtdp_pos           NA         NA      NA       NA
## avg_qbtdp_team          NA         NA      NA       NA
## grass_1          -0.784849   0.278239  -2.821 0.004793 **
## bad_weather_1    -1.032719   0.553706  -1.865 0.062175 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 23.99 on 39189 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.4333, Adjusted R-squared:  0.4324
## F-statistic: 468.2 on 64 and 39189 DF,  p-value: < 2.2e-16
```

**Second run at Receiving Yards

```
linRegrecy2 <- lm(recy ~ ATL + BAL + BUF + CHI+
                    CIN + CLE + GB + HOU  + JAC + KC + NE + NOR + NYG+
                    OAK + SD + TB + TEN + WAS+ avg_recy_plyr+grass_1+bad_weathe
r_1, data = nfl_data)

summary(linRegrecy2)
```

```
##
## Call:
## lm(formula = recy ~ ATL + BAL + BUF + CHI + CIN + CLE + GB +
##     HOU + JAC + KC + NE + NOR + NYG + OAK + SD + TB + TEN + WAS +
##     avg_recy_plyr + grass_1 + bad_weather_1, data = nfl_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -98.025 -12.009  -1.667   6.541 233.065
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -0.211685   0.261695  -0.809  0.41858
## ATL            1.173657   0.704988   1.665  0.09596 .
## BAL            1.129004   0.682446   1.654  0.09806 .
## BUF            0.963714   0.734433   1.312  0.18946
## CHI            1.048220   0.741523   1.414  0.15749
## CIN            0.751482   0.719641   1.044  0.29638
## CLE            1.654023   0.721499   2.292  0.02188 *
## GB             1.473707   0.683379   2.157  0.03105 *
## HOU            1.532794   0.708139   2.165  0.03043 *
## JAC            1.513912   0.729942   2.074  0.03808 *
## KC             1.039219   0.711806   1.460  0.14430
## NE             1.039777   0.690682   1.505  0.13222
## NOR            1.308918   0.685438   1.910  0.05619 .
## NYG            1.658276   0.712309   2.328  0.01992 *
## OAK            1.574307   0.702652   2.241  0.02506 *
## SD             1.555423   0.732265   2.124  0.03367 *
## TB             1.158182   0.726134   1.595  0.11072
## TEN            1.070310   0.717197   1.492  0.13561
## WAS            0.906280   0.708146   1.280  0.20063
## avg_recy_plyr  1.000167   0.005805 172.298  < 2e-16 ***
## grass_1       -0.832766   0.259501  -3.209  0.00133 **
## bad_weather_1 -1.380465   0.542598  -2.544  0.01096 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 23.98 on 39233 degrees of freedom
## Multiple R-squared:  0.433,  Adjusted R-squared:  0.4327
## F-statistic:  1426 on 21 and 39233 DF,  p-value: < 2.2e-16
```

Modest gains in r-square and residual standard error.

**Receptions

```
linRegrec <- lm(rec ~ height+ weight+cold_weather + hot_weather + home_team_1+
temp+ forty1 + vertical1 + shuttle1+ cone1 + ARI + ATL + BAL + BUF + CAR + CHI+
                   CIN + CLE + DAL + DEN + DET + GB + HOU + IND + JAC + KC + MI
A + MINN + NE + NOR + NYG+
                   NYJ + OAK + PHI + PIT +SD + SEA + STL + TB + TEN + WAS + avg
_recy_plyr+avg_recy_pos +
                  avg_recy_team + avg_rec_plyr +avg_rec_pos + avg_rec_team +avg
_trg_plyr + avg_trg_pos +
                  avg_trg_team + avg_rectd_plyr + avg_rectd_pos +avg_rectd_team
+
                  avg_tdr_plyr + avg_tdr_pos + avg_tdr_team +
                  avg_rbra_plyr + avg_rbra_pos +avg_rbra_team +
                  avg_rbry_plyr + avg_rbry_pos +avg_rbry_team +
                  avg_fuml_plyr + avg_fuml_pos +avg_fuml_team +
                  avg_qbpy_plyr + avg_qbpy_pos +avg_qbpy_team +
                  avg_qbpa_plyr + avg_qbpa_pos +avg_qbpa_team+
                  avg_qbpc_plyr + avg_qbpc_pos +avg_qbpc_team +
                  avg_qbints_plyr + avg_qbints_pos +avg_qbints_team +
                  avg_qbtdp_plyr + avg_qbtdp_pos +avg_qbtdp_team + grass_1 + ba
d_weather_1, data = nfl_data)

summary(linRegrec)
```

```
##
## Call:
## lm(formula = rec ~ height + weight + cold_weather + hot_weather +
##     home_team_1 + temp + forty1 + vertical1 + shuttle1 + cone1 +
##     ARI + ATL + BAL + BUF + CAR + CHI + CIN + CLE + DAL + DEN +
##     DET + GB + HOU + IND + JAC + KC + MIA + MINN + NE + NOR +
##     NYG + NYJ + OAK + PHI + PIT + SD + SEA + STL + TB + TEN +
##     WAS + avg_recy_plyr + avg_recy_pos + avg_recy_team + avg_rec_plyr +
##     avg_rec_pos + avg_rec_team + avg_trg_plyr + avg_trg_pos +
##     avg_trg_team + avg_rectd_plyr + avg_rectd_pos + avg_rectd_team +
##     avg_tdr_plyr + avg_tdr_pos + avg_tdr_team + avg_rbra_plyr +
##     avg_rbra_pos + avg_rbra_team + avg_rbry_plyr + avg_rbry_pos +
##     avg_rbry_team + avg_fuml_plyr + avg_fuml_pos + avg_fuml_team +
##     avg_qbpy_plyr + avg_qbpy_pos + avg_qbpy_team + avg_qbpa_plyr +
##     avg_qbpa_pos + avg_qbpa_team + avg_qbpc_plyr + avg_qbpc_pos +
##     avg_qbpc_team + avg_qbints_plyr + avg_qbints_pos + avg_qbints_team +
##     avg_qbtdp_plyr + avg_qbtdp_pos + avg_qbtdp_team + grass_1 +
##     bad_weather_1, data = nfl_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.4192 -0.9846 -0.0923  0.6720 12.9113
##
## Coefficients: (18 not defined because of singularities)
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.0938807  0.7039114   0.133 0.893901
## height          -0.0042403  0.0060734  -0.698 0.485069
## weight           0.0002323  0.0008104   0.287 0.774409
## cold_weather    -0.0220358  0.0366898  -0.601 0.548112
## hot_weather     -0.1970381  0.1449864  -1.359 0.174151
## home_team_1     -0.0254156  0.0191256  -1.329 0.183895
## temp             0.0017991  0.0010453   1.721 0.085234 .
## forty1           0.0017486  0.0964507   0.018 0.985536
## vertical1        0.0005648  0.0039398   0.143 0.886002
## shuttle1        -0.0271131  0.0832612  -0.326 0.744699
## cone1            0.0060101  0.0607552   0.099 0.921200
## ARI              0.1197501  0.0698525   1.714 0.086476 .
## ATL              0.1953800  0.0699112   2.795 0.005198 **
## BAL              0.2263529  0.0678648   3.335 0.000853 ***
## BUF              0.2195255  0.0706494   3.107 0.001890 **
## CAR              0.1062628  0.0691454   1.537 0.124350
## CHI              0.2570649  0.0706488   3.639 0.000274 ***
## CIN              0.2251935  0.0700479   3.215 0.001306 **
## CLE              0.3016798  0.0693510   4.350 1.36e-05 ***
## DAL              0.1397207  0.0707926   1.974 0.048427 *
## DEN              0.1536972  0.0681094   2.257 0.024037 *
## DET              0.1571431  0.0702766   2.236 0.025353 *
## GB               0.2185524  0.0685043   3.190 0.001422 **
```

```
## HOU              0.2088800  0.0695237   3.004 0.002662 **
## IND              0.1201685  0.0695019   1.729 0.083818 .
## JAC              0.2439920  0.0703792   3.467 0.000527 ***
## KC               0.2208369  0.0687027   3.214 0.001308 **
## MIA              0.1037127  0.0702386   1.477 0.139797
## MINN             0.1565144  0.0704092   2.223 0.026227 *
## NE               0.1415596  0.0691429   2.047 0.040631 *
## NOR              0.2432929  0.0694959   3.501 0.000464 ***
## NYG              0.2506883  0.0695777   3.603 0.000315 ***
## NYJ              0.1014262  0.0695915   1.457 0.145000
## OAK              0.2556766  0.0685380   3.730 0.000191 ***
## PHI              0.0569294  0.0697114   0.817 0.414137
## PIT              0.0941025  0.0688842   1.366 0.171917
## SD               0.2398785  0.0703241   3.411 0.000648 ***
## SEA              0.0139364  0.0678575   0.205 0.837278
## STL              0.1932653  0.0701693   2.754 0.005885 **
## TB               0.1944023  0.0698621   2.783 0.005394 **
## TEN              0.1973571  0.0689379   2.863 0.004201 **
## WAS              0.1976704  0.0687225   2.876 0.004025 **
## avg_recy_plyr    0.0015752  0.0024978   0.631 0.528282
## avg_recy_pos     0.0086644  0.3211383   0.027 0.978476
## avg_recy_team           NA         NA      NA       NA
## avg_rec_plyr     0.9902182  0.0385370  25.695  < 2e-16 ***
## avg_rec_pos      0.0769220  0.7553548   0.102 0.918888
## avg_rec_team            NA         NA      NA       NA
## avg_trg_plyr    -0.0064921  0.0282613  -0.230 0.818313
## avg_trg_pos     -0.1284888  3.6335359  -0.035 0.971791
## avg_trg_team            NA         NA      NA       NA
## avg_rectd_plyr  -0.0009410  0.1335739  -0.007 0.994379
## avg_rectd_pos    0.2897527 16.6469179   0.017 0.986113
## avg_rectd_team          NA         NA      NA       NA
## avg_tdr_plyr     0.0195821  0.1647512   0.119 0.905388
## avg_tdr_pos      0.0477189  8.1491593   0.006 0.995328
## avg_tdr_team            NA         NA      NA       NA
## avg_rbra_plyr   -0.0062248  0.0159222  -0.391 0.695837
## avg_rbra_pos     0.0165374  0.5810623   0.028 0.977295
## avg_rbra_team           NA         NA      NA       NA
## avg_rbry_plyr    0.0015785  0.0036987   0.427 0.669546
## avg_rbry_pos    -0.0042079  0.0530064  -0.079 0.936727
## avg_rbry_team           NA         NA      NA       NA
## avg_fuml_plyr   -0.0096984  0.1971976  -0.049 0.960775
## avg_fuml_pos     0.1110329  2.3276112   0.048 0.961954
## avg_fuml_team           NA         NA      NA       NA
## avg_qbpy_plyr    0.0020097  0.0031234   0.643 0.519948
## avg_qbpy_pos            NA         NA      NA       NA
## avg_qbpy_team           NA         NA      NA       NA
## avg_qbpa_plyr    0.0044174  0.0243174   0.182 0.855854
## avg_qbpa_pos            NA         NA      NA       NA
## avg_qbpa_team           NA         NA      NA       NA
```

```
## avg_qbpc_plyr   -0.0258841  0.0439258  -0.589 0.555684
## avg_qbpc_pos            NA         NA      NA       NA
## avg_qbpc_team           NA         NA      NA       NA
## avg_qbints_plyr -0.0655407  0.1406670  -0.466 0.641270
## avg_qbints_pos          NA         NA      NA       NA
## avg_qbints_team         NA         NA      NA       NA
## avg_qbtdp_plyr  -0.0451540  0.1540147  -0.293 0.769386
## avg_qbtdp_pos           NA         NA      NA       NA
## avg_qbtdp_team          NA         NA      NA       NA
## grass_1         -0.0478848  0.0197174  -2.429 0.015164 *
## bad_weather_1   -0.0976921  0.0392384  -2.490 0.012789 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.7 on 39189 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.4514, Adjusted R-squared:  0.4505
## F-statistic: 503.8 on 64 and 39189 DF,  p-value: < 2.2e-16
```

**Second run at Receptions

```
linRegrec2 <- lm(rec ~ temp+ ATL + BAL + BUF + CHI+
                   CIN + CLE + GB + HOU + JAC + KC + NOR + NYG+
                   OAK +SD + WAS + avg_rec_plyr+ grass_1+
                   bad_weather_1, data = nfl_data)

summary(linRegrec2)
```

```
##
## Call:
## lm(formula = rec ~ temp + ATL + BAL + BUF + CHI + CIN + CLE +
##     GB + HOU + JAC + KC + NOR + NYG + OAK + SD + WAS + avg_rec_plyr +
##     grass_1 + bad_weather_1, data = nfl_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.4197 -0.9880 -0.0884  0.6727 12.9019
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -0.1506914  0.0371362  -4.058 4.96e-05 ***
## temp           0.0024483  0.0005763   4.248 2.16e-05 ***
## ATL            0.0803292  0.0497909   1.613 0.106680
## BAL            0.0996549  0.0481277   2.071 0.038399 *
## BUF            0.0937706  0.0518060   1.810 0.070298 .
## CHI            0.1349074  0.0522877   2.580 0.009881 **
## CIN            0.0989237  0.0507972   1.947 0.051491 .
## CLE            0.1820672  0.0508968   3.577 0.000348 ***
## GB             0.1020118  0.0483521   2.110 0.034885 *
## HOU            0.0936476  0.0499888   1.873 0.061024 .
## JAC            0.1176936  0.0515057   2.285 0.022315 *
## KC             0.1016323  0.0501548   2.026 0.042734 *
## NOR            0.1201051  0.0483883   2.482 0.013065 *
## NYG            0.1254061  0.0502777   2.494 0.012626 *
## OAK            0.1314350  0.0494564   2.658 0.007873 **
## SD             0.1110372  0.0516142   2.151 0.031459 *
## WAS            0.0833186  0.0498704   1.671 0.094789 .
## avg_rec_plyr   1.0000105  0.0055956 178.715  < 2e-16 ***
## grass_1       -0.0557042  0.0181608  -3.067 0.002162 **
## bad_weather_1 -0.1035875  0.0391033  -2.649 0.008075 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.7 on 39235 degrees of freedom
## Multiple R-squared:  0.451,  Adjusted R-squared:  0.4507
## F-statistic:  1696 on 19 and 39235 DF,  p-value: < 2.2e-16
```

We had a modest r-square improvement

**Targets

```
linRegtrg <- lm(trg ~ height+ weight+cold_weather + hot_weather + home_team_1+
temp+forty1 + vertical1  + shuttle1+ cone1 + ARI + ATL + BAL + BUF + CAR + CHI+
                  CIN + CLE + DAL + DEN + DET + GB + HOU + IND + JAC + KC + MI
A + MINN + NE + NOR + NYG+
                  NYJ + OAK + PHI + PIT +SD + SEA + STL + TB + TEN + WAS +avg_r
ecy_plyr+avg_recy_pos +
                  avg_recy_team + avg_rec_plyr +avg_rec_pos + avg_rec_team +avg
_trg_plyr + avg_trg_pos +
                  avg_trg_team + avg_rectd_plyr + avg_rectd_pos +avg_rectd_team
+
                  avg_tdr_plyr + avg_tdr_pos + avg_tdr_team +
                  avg_rbra_plyr + avg_rbra_pos +avg_rbra_team +
                  avg_rbry_plyr + avg_rbry_pos +avg_rbry_team +
                  avg_fuml_plyr + avg_fuml_pos +avg_fuml_team +
                  avg_qbpy_plyr + avg_qbpy_pos +avg_qbpy_team +
                  avg_qbpa_plyr + avg_qbpa_pos +avg_qbpa_team+
                  avg_qbpc_plyr + avg_qbpc_pos +avg_qbpc_team +
                  avg_qbints_plyr + avg_qbints_pos +avg_qbints_team +
                  avg_qbtdp_plyr + avg_qbtdp_pos +avg_qbtdp_team + grass_1 + ba
d_weather_1 , data = nfl_data)

summary(linRegtrg)
```

```
##
## Call:
## lm(formula = trg ~ height + weight + cold_weather + hot_weather +
##     home_team_1 + temp + forty1 + vertical1 + shuttle1 + cone1 +
##     ARI + ATL + BAL + BUF + CAR + CHI + CIN + CLE + DAL + DEN +
##     DET + GB + HOU + IND + JAC + KC + MIA + MINN + NE + NOR +
##     NYG + NYJ + OAK + PHI + PIT + SD + SEA + STL + TB + TEN +
##     WAS + avg_recy_plyr + avg_recy_pos + avg_recy_team + avg_rec_plyr +
##     avg_rec_pos + avg_rec_team + avg_trg_plyr + avg_trg_pos +
##     avg_trg_team + avg_rectd_plyr + avg_rectd_pos + avg_rectd_team +
##     avg_tdr_plyr + avg_tdr_pos + avg_tdr_team + avg_rbra_plyr +
##     avg_rbra_pos + avg_rbra_team + avg_rbry_plyr + avg_rbry_pos +
##     avg_rbry_team + avg_fuml_plyr + avg_fuml_pos + avg_fuml_team +
##     avg_qbpy_plyr + avg_qbpy_pos + avg_qbpy_team + avg_qbpa_plyr +
##     avg_qbpa_pos + avg_qbpa_team + avg_qbpc_plyr + avg_qbpc_pos +
##     avg_qbpc_team + avg_qbints_plyr + avg_qbints_pos + avg_qbints_team +
##     avg_qbtdp_plyr + avg_qbtdp_pos + avg_qbtdp_team + grass_1 +
##     bad_weather_1, data = nfl_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.8392 -1.2671 -0.1191  0.8976 15.9457
##
## Coefficients: (18 not defined because of singularities)
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.2805420  0.9609969   0.292 0.770343
## height          -0.0050085  0.0082916  -0.604 0.545816
## weight           0.0001882  0.0011063   0.170 0.864933
## cold_weather    -0.0458794  0.0500899  -0.916 0.359703
## hot_weather     -0.2172190  0.1979389  -1.097 0.272472
## home_team_1     -0.0972821  0.0261107  -3.726 0.000195 ***
## temp            -0.0007050  0.0014271  -0.494 0.621291
## forty1          -0.0071155  0.1316769  -0.054 0.956905
## vertical1        0.0006616  0.0053787   0.123 0.902110
## shuttle1        -0.0409339  0.1136702  -0.360 0.718766
## cone1            0.0066546  0.0829445   0.080 0.936055
## ARI              0.2403120  0.0953644   2.520 0.011742 *
## ATL              0.2866087  0.0954444   3.003 0.002676 **
## BAL              0.3775828  0.0926507   4.075 4.60e-05 ***
## BUF              0.3935459  0.0964523   4.080 4.51e-05 ***
## CAR              0.2493833  0.0943990   2.642 0.008250 **
## CHI              0.3678562  0.0964514   3.814 0.000137 ***
## CIN              0.3416093  0.0956311   3.572 0.000354 ***
## CLE              0.5036674  0.0946796   5.320 1.05e-07 ***
## DAL              0.2107326  0.0966478   2.180 0.029232 *
## DEN              0.2696729  0.0929846   2.900 0.003731 **
## DET              0.2720134  0.0959433   2.835 0.004583 **
## GB               0.3200362  0.0935237   3.422 0.000622 ***
```

```
## HOU              0.3616955  0.0949154   3.811 0.000139 ***
## IND              0.2467854  0.0948857   2.601 0.009302 **
## JAC              0.4402984  0.0960834   4.582 4.61e-06 ***
## KC               0.3218716  0.0937946   3.432 0.000601 ***
## MIA              0.2334368  0.0958915   2.434 0.014922 *
## MINN             0.2619964  0.0961244   2.726 0.006421 **
## NE               0.2339584  0.0943956   2.478 0.013198 *
## NOR              0.3255428  0.0948775   3.431 0.000602 ***
## NYG              0.4347323  0.0949892   4.577 4.74e-06 ***
## NYJ              0.2333530  0.0950079   2.456 0.014048 *
## OAK              0.4152517  0.0935697   4.438 9.11e-06 ***
## PHI              0.1568084  0.0951717   1.648 0.099435 .
## PIT              0.1139187  0.0940424   1.211 0.225767
## SD               0.3244898  0.0960082   3.380 0.000726 ***
## SEA              0.0376568  0.0926406   0.406 0.684390
## STL              0.3736498  0.0957968   3.900 9.62e-05 ***
## TB               0.3788240  0.0953775   3.972 7.14e-05 ***
## TEN              0.3174885  0.0941157   3.373 0.000743 ***
## WAS              0.2768352  0.0938216   2.951 0.003173 **
## avg_recy_plyr    0.0030372  0.0034101   0.891 0.373123
## avg_recy_pos     0.0318712  0.4384258   0.073 0.942049
## avg_recy_team           NA         NA      NA       NA
## avg_rec_plyr     0.0119431  0.0526116   0.227 0.820421
## avg_rec_pos      0.1448771  1.0312286   0.140 0.888274
## avg_rec_team            NA         NA      NA       NA
## avg_trg_plyr     0.9667014  0.0385830  25.055  < 2e-16 ***
## avg_trg_pos     -0.4069845  4.9605914  -0.082 0.934612
## avg_trg_team            NA         NA      NA       NA
## avg_rectd_plyr   0.0307513  0.1823584   0.169 0.866088
## avg_rectd_pos    1.6199949 22.7267761   0.071 0.943174
## avg_rectd_team          NA         NA      NA       NA
## avg_tdr_plyr     0.0388035  0.2249223   0.173 0.863030
## avg_tdr_pos     -0.5249072 11.1254300  -0.047 0.962369
## avg_tdr_team            NA         NA      NA       NA
## avg_rbra_plyr   -0.0087116  0.0217374  -0.401 0.688595
## avg_rbra_pos     0.0557440  0.7932803   0.070 0.943979
## avg_rbra_team           NA         NA      NA       NA
## avg_rbry_plyr    0.0020798  0.0050495   0.412 0.680429
## avg_rbry_pos    -0.0052586  0.0723656  -0.073 0.942071
## avg_rbry_team           NA         NA      NA       NA
## avg_fuml_plyr   -0.0109521  0.2692189  -0.041 0.967551
## avg_fuml_pos     0.1504232  3.1777113   0.047 0.962245
## avg_fuml_team           NA         NA      NA       NA
## avg_qbpy_plyr    0.0033452  0.0042642   0.784 0.432757
## avg_qbpy_pos            NA         NA      NA       NA
## avg_qbpy_team           NA         NA      NA       NA
## avg_qbpa_plyr   -0.0092676  0.0331987  -0.279 0.780127
## avg_qbpa_pos            NA         NA      NA       NA
## avg_qbpa_team           NA         NA      NA       NA
```

```
## avg_qbpc_plyr   -0.0170934  0.0599685  -0.285 0.775616
## avg_qbpc_pos            NA         NA      NA       NA
## avg_qbpc_team           NA         NA      NA       NA
## avg_qbints_plyr -0.0865833  0.1920420  -0.451 0.652096
## avg_qbints_pos          NA         NA      NA       NA
## avg_qbints_team         NA         NA      NA       NA
## avg_qbtdp_plyr  -0.0799650  0.2102647  -0.380 0.703720
## avg_qbtdp_pos           NA         NA      NA       NA
## avg_qbtdp_team          NA         NA      NA       NA
## grass_1         -0.0415842  0.0269187  -1.545 0.122401
## bad_weather_1    0.0103342  0.0535692   0.193 0.847029
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.321 on 39189 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.5394, Adjusted R-squared:  0.5387
## F-statistic: 717.1 on 64 and 39189 DF,  p-value: < 2.2e-16
```

** Second run at targets

```
linRegtrg2 <- lm(trg ~ home_team_1+ ARI + ATL + BAL + BUF + CAR + CHI + CIN + C
LE +
                   DAL + DEN + DET + GB + HOU + IND + JAC + KC + MIA + MINN + N
E + NOR + NYG+
                   NYJ + OAK + SD + STL + TB + TEN + WAS + avg_trg_plyr,
                 data = nfl_data)

summary(linRegtrg2)
```

```
##
## Call:
## lm(formula = trg ~ home_team_1 + ARI + ATL + BAL + BUF + CAR +
##     CHI + CIN + CLE + DAL + DEN + DET + GB + HOU + IND + JAC +
##     KC + MIA + MINN + NE + NOR + NYG + NYJ + OAK + SD + STL +
##     TB + TEN + WAS + avg_trg_plyr, data = nfl_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.7643 -1.2701 -0.1182  0.9013 15.9577
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -0.169857   0.038321  -4.433 9.34e-06 ***
## home_team_1  -0.104879   0.025714  -4.079 4.54e-05 ***
## ARI           0.158549   0.074505   2.128 0.033342 *
## ATL           0.220675   0.074292   2.970 0.002976 **
## BAL           0.292681   0.071444   4.097 4.20e-05 ***
## BUF           0.318000   0.075959   4.186 2.84e-05 ***
## CAR           0.167939   0.074438   2.256 0.024071 *
## CHI           0.274797   0.076461   3.594 0.000326 ***
## CIN           0.268641   0.074635   3.599 0.000319 ***
## CLE           0.407410   0.074616   5.460 4.79e-08 ***
## DAL           0.145570   0.075861   1.919 0.055002 .
## DEN           0.175741   0.072298   2.431 0.015071 *
## DET           0.199522   0.075596   2.639 0.008310 **
## GB            0.242803   0.071315   3.405 0.000663 ***
## HOU           0.283416   0.074612   3.799 0.000146 ***
## IND           0.174557   0.074405   2.346 0.018979 *
## JAC           0.338296   0.075461   4.483 7.38e-06 ***
## KC            0.227502   0.073579   3.092 0.001990 **
## MIA           0.141292   0.075833   1.863 0.062442 .
## MINN          0.186664   0.074910   2.492 0.012712 *
## NE            0.170836   0.071943   2.375 0.017572 *
## NOR           0.261349   0.071912   3.634 0.000279 ***
## NYG           0.361206   0.073965   4.883 1.05e-06 ***
## NYJ           0.154073   0.074116   2.079 0.037642 *
## OAK           0.314571   0.072775   4.323 1.55e-05 ***
## SD            0.230201   0.075435   3.052 0.002277 **
## STL           0.287395   0.075574   3.803 0.000143 ***
## TB            0.281585   0.075177   3.746 0.000180 ***
## TEN           0.226852   0.074222   3.056 0.002242 **
## WAS           0.200823   0.073533   2.731 0.006316 **
## avg_trg_plyr  1.000674   0.004689 213.428  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.32 on 39224 degrees of freedom
```

```
## Multiple R-squared:  0.5393, Adjusted R-squared:  0.5389
## F-statistic:  1530 on 30 and 39224 DF,  p-value: < 2.2e-16
```

Modest gains in second run

**Receiving TD's

```
linRegRecTD <- lm(trg ~ height+ weight+cold_weather + hot_weather + home_team_1
+ temp+ forty1 + vertical1  + shuttle1+ cone1 + ARI + ATL + BAL + BUF + CAR + C
HI+
                  CIN + CLE + DAL + DEN + DET + GB + HOU + IND + JAC + KC + MI
A + MINN + NE + NOR + NYG+
                  NYJ + OAK + PHI + PIT +SD + SEA + STL + TB + TEN + WAS +avg_r
ecy_plyr+avg_recy_pos +
                  avg_recy_team + avg_rec_plyr +avg_rec_pos + avg_rec_team +avg
_trg_plyr + avg_trg_pos +
                  avg_trg_team + avg_rectd_plyr + avg_rectd_pos +avg_rectd_team
+
                  avg_tdr_plyr + avg_tdr_pos + avg_tdr_team +
                  avg_rbra_plyr + avg_rbra_pos +avg_rbra_team +
                  avg_rbry_plyr + avg_rbry_pos +avg_rbry_team +
                  avg_fuml_plyr + avg_fuml_pos +avg_fuml_team +
                  avg_qbpy_plyr + avg_qbpy_pos +avg_qbpy_team +
                  avg_qbpa_plyr + avg_qbpa_pos +avg_qbpa_team+
                  avg_qbpc_plyr + avg_qbpc_pos +avg_qbpc_team +
                  avg_qbints_plyr + avg_qbints_pos +avg_qbints_team +
                  avg_qbtdp_plyr + avg_qbtdp_pos +avg_qbtdp_team + grass_1 + ba
d_weather_1 , data = nfl_data)

summary(linRegRecTD)
```

```
##
## Call:
## lm(formula = trg ~ height + weight + cold_weather + hot_weather +
##     home_team_1 + temp + forty1 + vertical1 + shuttle1 + cone1 +
##     ARI + ATL + BAL + BUF + CAR + CHI + CIN + CLE + DAL + DEN +
##     DET + GB + HOU + IND + JAC + KC + MIA + MINN + NE + NOR +
##     NYG + NYJ + OAK + PHI + PIT + SD + SEA + STL + TB + TEN +
##     WAS + avg_recy_plyr + avg_recy_pos + avg_recy_team + avg_rec_plyr +
##     avg_rec_pos + avg_rec_team + avg_trg_plyr + avg_trg_pos +
##     avg_trg_team + avg_rectd_plyr + avg_rectd_pos + avg_rectd_team +
##     avg_tdr_plyr + avg_tdr_pos + avg_tdr_team + avg_rbra_plyr +
##     avg_rbra_pos + avg_rbra_team + avg_rbry_plyr + avg_rbry_pos +
##     avg_rbry_team + avg_fuml_plyr + avg_fuml_pos + avg_fuml_team +
##     avg_qbpy_plyr + avg_qbpy_pos + avg_qbpy_team + avg_qbpa_plyr +
##     avg_qbpa_pos + avg_qbpa_team + avg_qbpc_plyr + avg_qbpc_pos +
##     avg_qbpc_team + avg_qbints_plyr + avg_qbints_pos + avg_qbints_team +
##     avg_qbtdp_plyr + avg_qbtdp_pos + avg_qbtdp_team + grass_1 +
##     bad_weather_1, data = nfl_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.8392 -1.2671 -0.1191  0.8976 15.9457
##
## Coefficients: (18 not defined because of singularities)
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.2805420  0.9609969   0.292 0.770343
## height          -0.0050085  0.0082916  -0.604 0.545816
## weight           0.0001882  0.0011063   0.170 0.864933
## cold_weather    -0.0458794  0.0500899  -0.916 0.359703
## hot_weather     -0.2172190  0.1979389  -1.097 0.272472
## home_team_1     -0.0972821  0.0261107  -3.726 0.000195 ***
## temp            -0.0007050  0.0014271  -0.494 0.621291
## forty1          -0.0071155  0.1316769  -0.054 0.956905
## vertical1        0.0006616  0.0053787   0.123 0.902110
## shuttle1        -0.0409339  0.1136702  -0.360 0.718766
## cone1            0.0066546  0.0829445   0.080 0.936055
## ARI              0.2403120  0.0953644   2.520 0.011742 *
## ATL              0.2866087  0.0954444   3.003 0.002676 **
## BAL              0.3775828  0.0926507   4.075 4.60e-05 ***
## BUF              0.3935459  0.0964523   4.080 4.51e-05 ***
## CAR              0.2493833  0.0943990   2.642 0.008250 **
## CHI              0.3678562  0.0964514   3.814 0.000137 ***
## CIN              0.3416093  0.0956311   3.572 0.000354 ***
## CLE              0.5036674  0.0946796   5.320 1.05e-07 ***
## DAL              0.2107326  0.0966478   2.180 0.029232 *
## DEN              0.2696729  0.0929846   2.900 0.003731 **
## DET              0.2720134  0.0959433   2.835 0.004583 **
## GB               0.3200362  0.0935237   3.422 0.000622 ***
```

```
## HOU              0.3616955  0.0949154   3.811 0.000139 ***
## IND              0.2467854  0.0948857   2.601 0.009302 **
## JAC              0.4402984  0.0960834   4.582 4.61e-06 ***
## KC               0.3218716  0.0937946   3.432 0.000601 ***
## MIA              0.2334368  0.0958915   2.434 0.014922 *
## MINN             0.2619964  0.0961244   2.726 0.006421 **
## NE               0.2339584  0.0943956   2.478 0.013198 *
## NOR              0.3255428  0.0948775   3.431 0.000602 ***
## NYG              0.4347323  0.0949892   4.577 4.74e-06 ***
## NYJ              0.2333530  0.0950079   2.456 0.014048 *
## OAK              0.4152517  0.0935697   4.438 9.11e-06 ***
## PHI              0.1568084  0.0951717   1.648 0.099435 .
## PIT              0.1139187  0.0940424   1.211 0.225767
## SD               0.3244898  0.0960082   3.380 0.000726 ***
## SEA              0.0376568  0.0926406   0.406 0.684390
## STL              0.3736498  0.0957968   3.900 9.62e-05 ***
## TB               0.3788240  0.0953775   3.972 7.14e-05 ***
## TEN              0.3174885  0.0941157   3.373 0.000743 ***
## WAS              0.2768352  0.0938216   2.951 0.003173 **
## avg_recy_plyr    0.0030372  0.0034101   0.891 0.373123
## avg_recy_pos     0.0318712  0.4384258   0.073 0.942049
## avg_recy_team           NA         NA      NA       NA
## avg_rec_plyr     0.0119431  0.0526116   0.227 0.820421
## avg_rec_pos      0.1448771  1.0312286   0.140 0.888274
## avg_rec_team            NA         NA      NA       NA
## avg_trg_plyr     0.9667014  0.0385830  25.055  < 2e-16 ***
## avg_trg_pos     -0.4069845  4.9605914  -0.082 0.934612
## avg_trg_team            NA         NA      NA       NA
## avg_rectd_plyr   0.0307513  0.1823584   0.169 0.866088
## avg_rectd_pos    1.6199949 22.7267761   0.071 0.943174
## avg_rectd_team          NA         NA      NA       NA
## avg_tdr_plyr     0.0388035  0.2249223   0.173 0.863030
## avg_tdr_pos     -0.5249072 11.1254300  -0.047 0.962369
## avg_tdr_team            NA         NA      NA       NA
## avg_rbra_plyr   -0.0087116  0.0217374  -0.401 0.688595
## avg_rbra_pos     0.0557440  0.7932803   0.070 0.943979
## avg_rbra_team           NA         NA      NA       NA
## avg_rbry_plyr    0.0020798  0.0050495   0.412 0.680429
## avg_rbry_pos    -0.0052586  0.0723656  -0.073 0.942071
## avg_rbry_team           NA         NA      NA       NA
## avg_fuml_plyr   -0.0109521  0.2692189  -0.041 0.967551
## avg_fuml_pos     0.1504232  3.1777113   0.047 0.962245
## avg_fuml_team           NA         NA      NA       NA
## avg_qbpy_plyr    0.0033452  0.0042642   0.784 0.432757
## avg_qbpy_pos            NA         NA      NA       NA
## avg_qbpy_team           NA         NA      NA       NA
## avg_qbpa_plyr   -0.0092676  0.0331987  -0.279 0.780127
## avg_qbpa_pos            NA         NA      NA       NA
## avg_qbpa_team           NA         NA      NA       NA
```

```
## avg_qbpc_plyr   -0.0170934  0.0599685  -0.285 0.775616
## avg_qbpc_pos            NA         NA      NA       NA
## avg_qbpc_team           NA         NA      NA       NA
## avg_qbints_plyr -0.0865833  0.1920420  -0.451 0.652096
## avg_qbints_pos          NA         NA      NA       NA
## avg_qbints_team         NA         NA      NA       NA
## avg_qbtdp_plyr  -0.0799650  0.2102647  -0.380 0.703720
## avg_qbtdp_pos           NA         NA      NA       NA
## avg_qbtdp_team          NA         NA      NA       NA
## grass_1         -0.0415842  0.0269187  -1.545 0.122401
## bad_weather_1    0.0103342  0.0535692   0.193 0.847029
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.321 on 39189 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.5394, Adjusted R-squared:  0.5387
## F-statistic: 717.1 on 64 and 39189 DF,  p-value: < 2.2e-16
```

**Second Run

```
linRegRecTD2 <- lm(tdrec ~weight+home_team_1+ ATL+ DAL + DEN + GB + NE + NOR +
NYG+
                  SD + avg_recy_plyr+ avg_rec_plyr, data = nfl_data)

summary(linRegRecTD2)
```

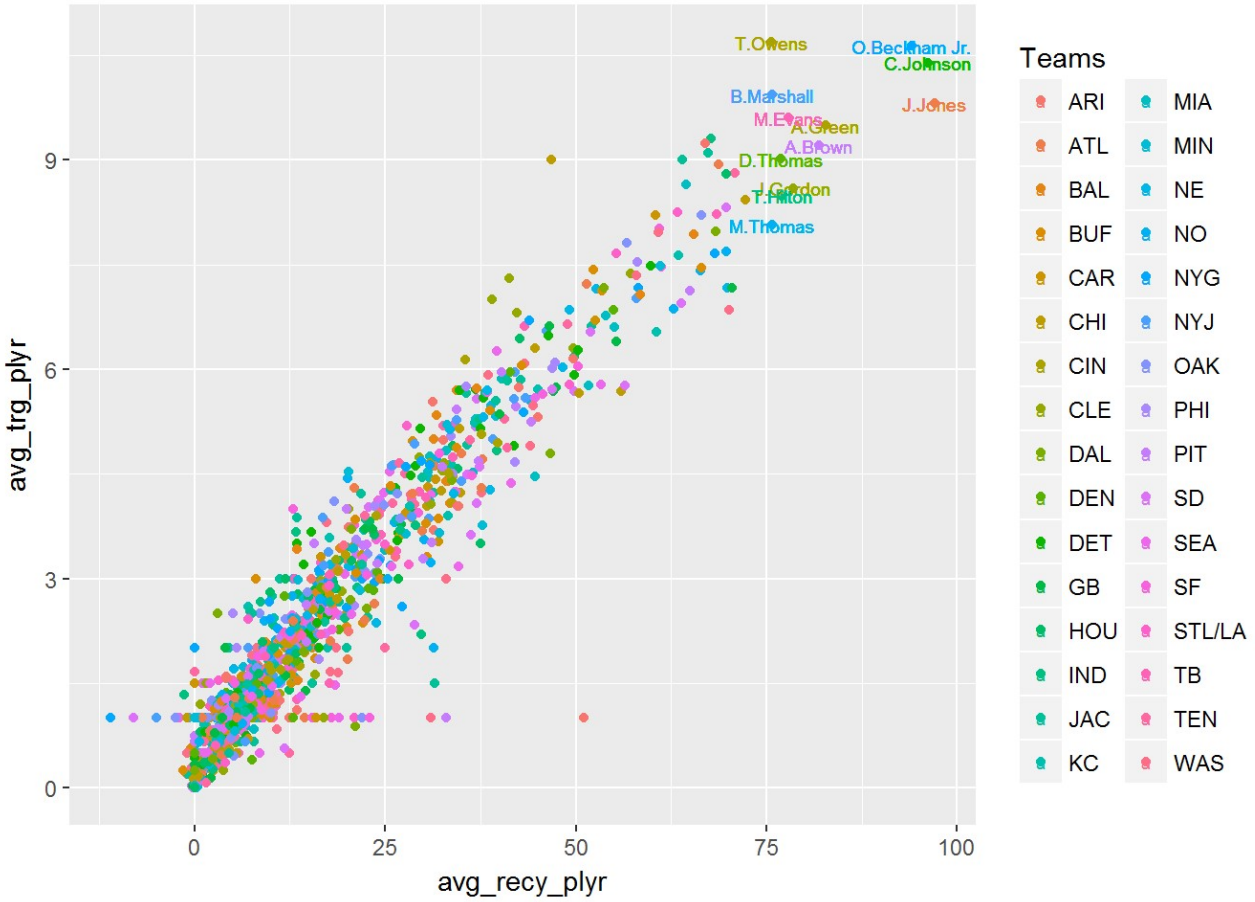
```
##
## Call:
## lm(formula = tdrec ~ weight + home_team_1 + ATL + DAL + DEN +
##     GB + NE + NOR + NYG + SD + avg_recy_plyr + avg_rec_plyr,
##     data = nfl_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.6874 -0.1763 -0.0684 -0.0023  3.8923
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -2.175e-01  1.798e-02 -12.096  < 2e-16 ***
## weight         9.066e-04  7.912e-05  11.458  < 2e-16 ***
## home_team_1    1.021e-02  4.007e-03   2.549  0.01080 *
## ATL            2.741e-02  1.105e-02   2.481  0.01310 *
## DAL            3.351e-02  1.132e-02   2.960  0.00308 **
## DEN            2.881e-02  1.076e-02   2.677  0.00743 **
## GB             6.169e-02  1.059e-02   5.827 5.68e-09 ***
## NE             4.574e-02  1.071e-02   4.272 1.94e-05 ***
## NOR            4.955e-02  1.066e-02   4.648 3.36e-06 ***
## NYG            2.220e-02  1.106e-02   2.008  0.04467 *
## SD             2.509e-02  1.132e-02   2.216  0.02669 *
## avg_recy_plyr  8.530e-03  3.167e-04  26.934  < 2e-16 ***
## avg_rec_plyr  -2.519e-02  4.294e-03  -5.865 4.52e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3784 on 39242 degrees of freedom
## Multiple R-squared:  0.1239, Adjusted R-squared:  0.1236
## F-statistic: 462.3 on 12 and 39242 DF,  p-value: < 2.2e-16
```

So, this goes on and on for each type of stat. I have the code saved, I feel this is getting a litte redundant.

Charts

WR targets by avg yards per player

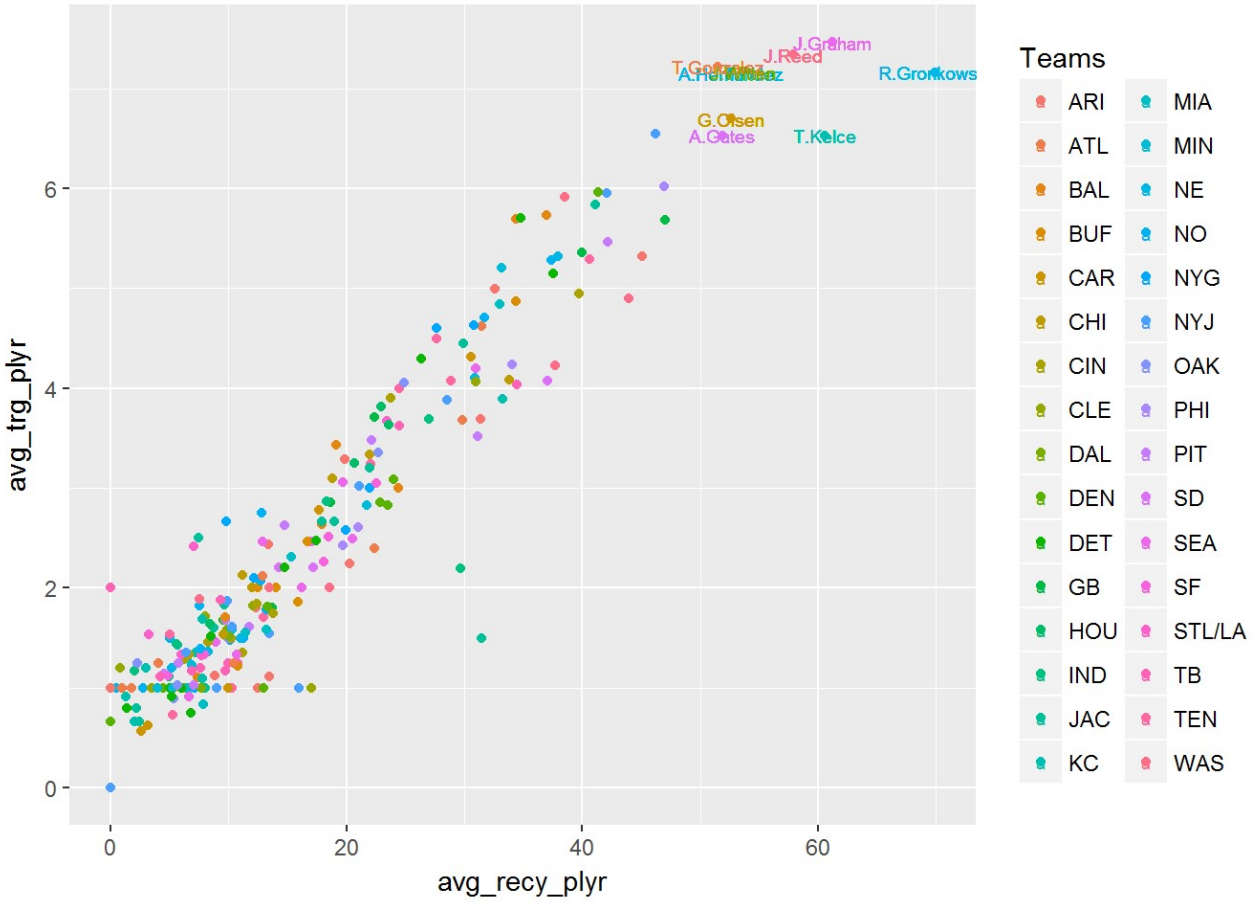
```
ggplot(data = nfl_data, aes(x = avg_recy_plyr, y = avg_trg_plyr, col = Teams ))
+
  geom_point()+
  geom_text(data = subset(nfl_data, avg_recy_plyr > 75), aes(label = pname), si
ze = 2.5)
```



There are few anomalies in this graph, not surprising, the amount of targets correlates with the amount of yards a player gets. The top right corner is “ALL PRO” corner.

**Tight ends should not be compared to WR

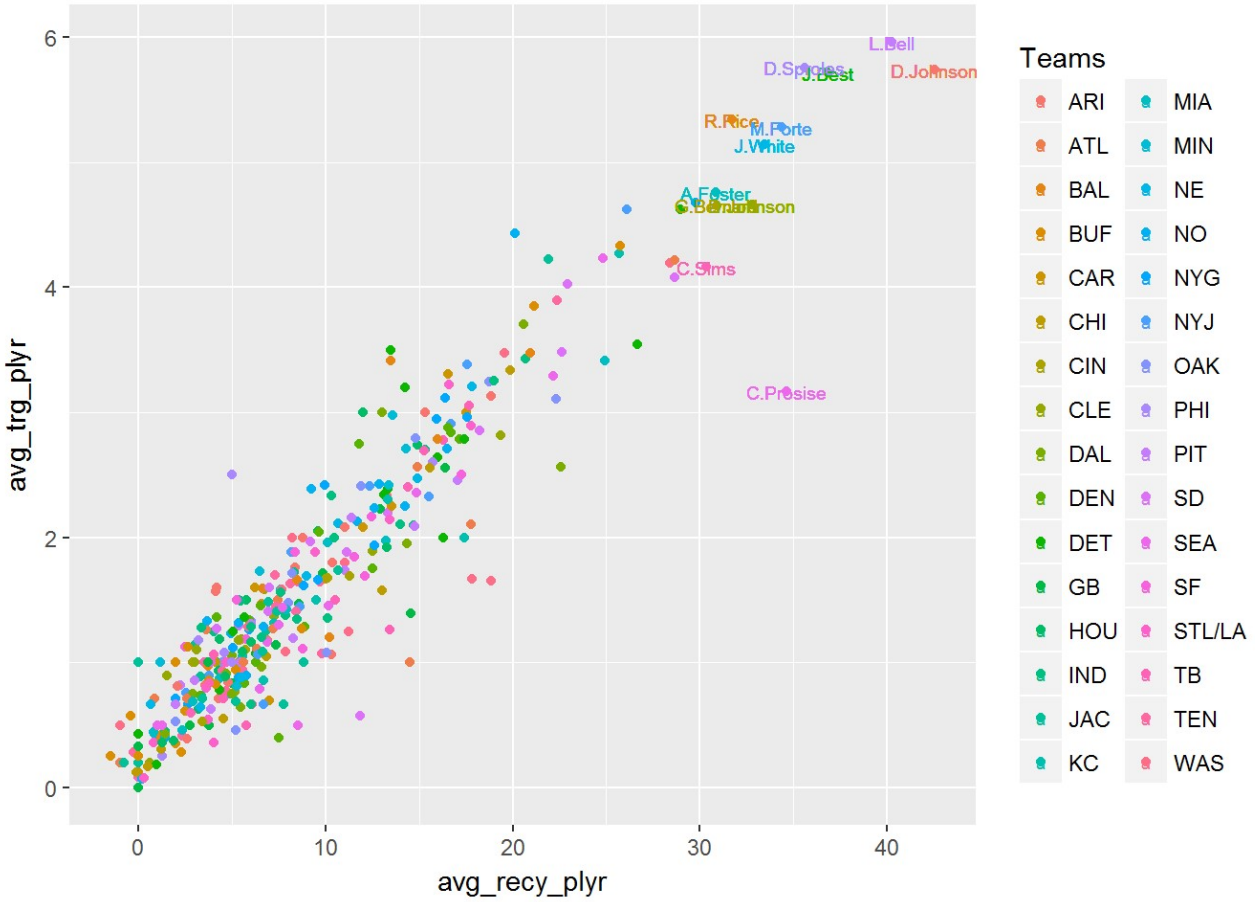
```
ggplot(data = nfl_data, aes(x = avg_recy_plyr, y = avg_trg_plyr, col = Teams ))
+
  geom_point(data = subset(nfl_data, pos1 == "TE"))+
  geom_text(data = subset(nfl_data, avg_recy_plyr > 50 & pos1 == "TE"), aes(lab
el = pname), size = 2.5)
```



I separated out the TE from the WR. TE are not “homerun” hitters, but are frequent targets of QB’s. Rob Gronkowski is the biggest anomaly here, he is widely considered the best position player to ever play.

** RB’s separated out

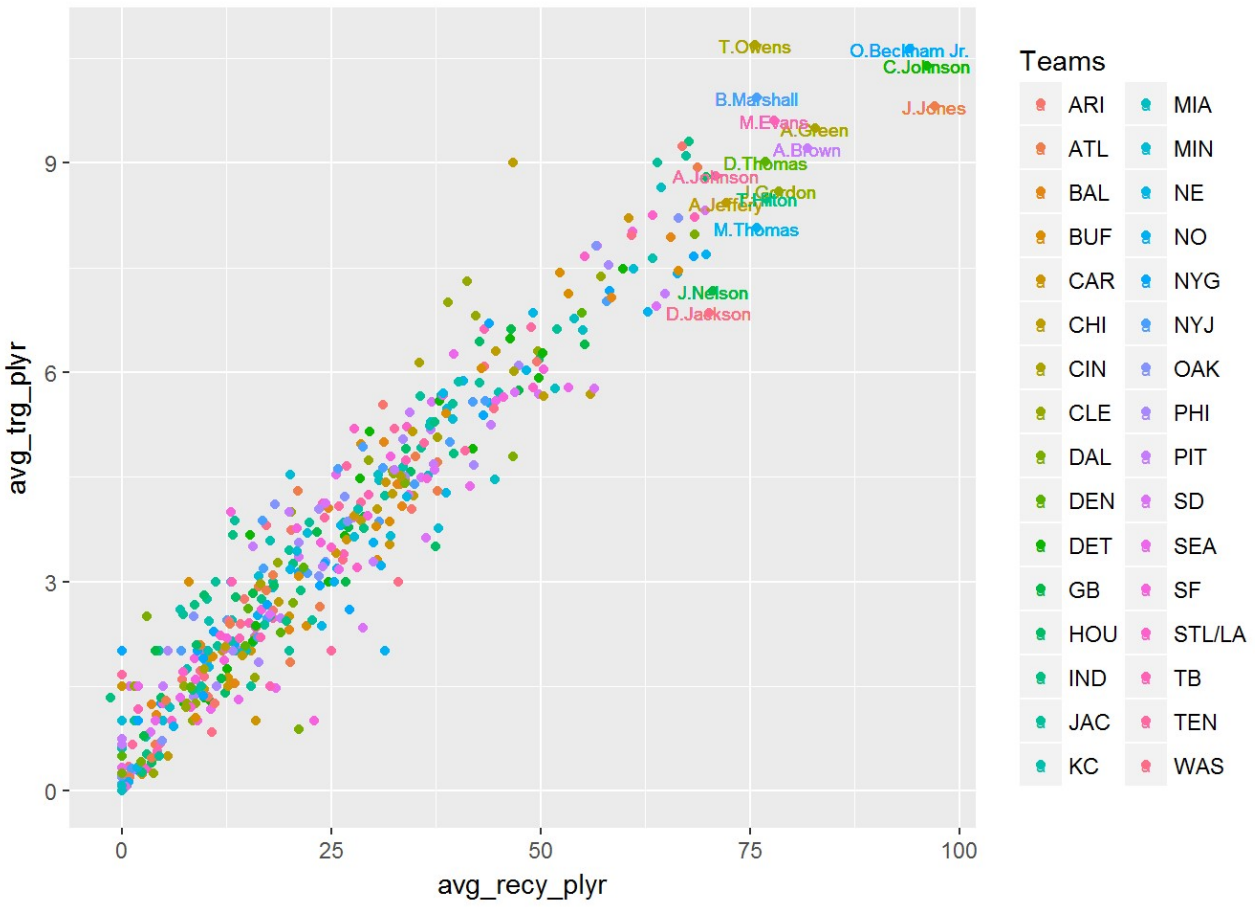
```
ggplot(data = nfl_data, aes(x = avg_recy_plyr, y = avg_trg_plyr, col = Teams ))
+
  geom_point(data = subset(nfl_data, pos1 == "RB"))+
  geom_text(data = subset(nfl_data, avg_recy_plyr > 30 & pos1 == "RB"), aes(lab
el = pname), size = 2.5)
```



CJ prosise was a rookie who had a couple of explosive games. He is a RB who played WR in college. He switched to RB his senior year of college and became an elite RB. This trend will regress somewhat, however, he is a very legit dual threat.

**WR only

```
ggplot(data = nfl_data, aes(x = avg_recy_plyr, y = avg_trg_plyr, col = Teams ))
+
  geom_point(data = subset(nfl_data, pos1 == "WR"))+
  geom_text(data = subset(nfl_data, avg_recy_plyr > 70 & pos1 == "WR"), aes(lab
el = pname), size = 2.5)
```


T.Owens
O.Beckham Jr.
C.Johnson
J.Jones
B.Marshall
M.Evans
A.Green
A.Brown
D.Thomas
A.Johnson
T.Gordon
A.Jeffery
T.Hilton
M.Thomas
J.Nelson
D.Jackson
avg_trg_plyr
avg_recy_plyr
0
3
6
9
0
25
50
75
100
Teams
ARI
ATL
BAL
BUF
CAR
CHI
CIN
CLE
DAL
DEN
DET
GB
HOU
IND
JAC
KC
MIA
MIN
NE
NO
NYG
NYJ
OAK
PHI
PIT
SD
SEA
SF
STL/LA
TB
TEN
WAS