Project Exam

Brennan Poole

2024-12-01

#Libraries  
  
library(MASS)  
library(car)

## Warning: package 'car' was built under R version 4.3.3

## Loading required package: carData

#Data set with structure and summary  
NFLData\_raw <- read.csv("ProjectData.csv", header=TRUE, sep=",")  
NFLData <- NFLData\_raw[-c(53,54,66,81,85,98,115,164,173,189,197),-c(2,4,5,9,12,13,14,16,17,20,21,22)]  
NFLData

## Age Pos Tgt Rec Yds TD X1D Ctch. PB AP  
## 1 24 WR 181 135 1749 12 80 74.6 y y  
## 2 29 WR 171 119 1799 13 83 69.6 y y  
## 3 24 WR 164 119 1515 10 75 72.6 y y  
## 4 29 TE 143 114 963 4 47 79.7 y n  
## 5 26 WR 156 109 1152 4 53 69.9 n n  
## 6 31 WR 150 108 1243 7 57 72.0 y n  
## 7 30 WR 160 107 1183 8 58 66.9 y n  
## 8 26 WR 158 106 1456 7 70 67.1 y y  
## 9 22 WR 160 105 1486 6 68 65.6 y y  
## 10 31 WR 175 103 1144 8 59 58.9 n n  
## 11 33 WR 137 103 1014 4 56 75.2 n n  
## 12 23 WR 145 100 1216 7 63 69.0 y n  
## 13 26 WR 136 96 1364 8 64 70.6 n n  
## 14 26 TE 127 95 960 5 48 74.8 n n  
## 15 23 WR 168 95 1042 3 47 56.5 n n  
## 16 34 TE 121 93 984 5 50 76.9 y n  
## 17 23 WR 138 87 1123 5 57 63.0 n n  
## 18 22 TE 120 86 889 10 48 71.7 y y  
## 19 27 WR 130 83 1024 2 53 63.8 n n  
## 20 24 TE 106 81 825 3 42 76.4 n n  
## 21 27 TE 123 81 882 6 35 65.9 y n  
## 22 25 WR 112 81 1066 7 47 72.3 n n  
## 23 24 WR 109 80 1297 8 53 73.4 n n  
## 24 30 WR 136 79 1255 13 54 58.1 y y  
## 25 31 WR 122 79 894 5 47 64.8 n n  
## 26 28 WR 132 79 1002 4 47 59.8 n n  
## 27 23 WR 102 79 938 7 44 77.5 n n  
## 28 23 WR 108 77 858 5 41 71.3 n n  
## 29 22 RB 95 76 591 4 28 80.0 n n  
## 30 29 WR 136 76 1016 8 49 55.9 n n  
## 31 25 WR 105 75 1342 7 61 71.4 n y  
## 32 31 WR 137 75 1057 7 48 54.7 n n  
## 33 28 RB 86 75 466 1 23 87.2 n n  
## 34 24 TE 91 73 673 2 29 80.2 n n  
## 35 24 TE 90 73 719 6 36 81.1 n n  
## 36 29 WR 128 72 1250 5 50 56.3 y n  
## 37 25 WR 104 72 1014 4 47 69.2 n n  
## 38 24 TE 102 71 761 5 40 69.6 y n  
## 39 27 WR 106 71 807 8 46 67.0 n n  
## 40 21 WR 108 70 911 10 38 64.8 n n  
## 41 22 WR 110 69 905 2 45 62.7 n n  
## 42 22 WR 98 68 771 2 30 69.4 n n  
## 43 24 WR 100 68 1074 5 51 68.0 n n  
## 44 29 WR 98 67 667 2 28 68.4 n n  
## 45 27 RB 83 67 564 7 31 80.7 y y  
## 46 26 WR 119 66 1114 8 52 55.5 y n  
## 47 30 TE 90 65 1020 6 42 72.2 y y  
## 48 23 WR 94 64 793 8 32 68.1 n n  
## 49 24 RB 70 64 549 3 23 91.4 n n  
## 50 22 WR 106 63 1140 5 42 59.4 n n  
## 51 21 WR 93 63 628 4 29 67.7 n n  
## 52 27 WR 91 62 613 4 26 68.1 n n  
## 55 22 WR 78 60 525 1 29 76.9 n n  
## 56 27 WR 89 60 892 7 34 67.4 n n  
## 57 23 WR 96 59 674 8 41 61.5 n n  
## 58 28 TE 83 59 592 3 33 71.1 n n  
## 59 30 WR 95 59 737 5 33 62.1 n n  
## 60 23 WR 104 59 640 2 29 56.7 n n  
## 61 27 TE 88 59 635 5 35 67.0 n n  
## 62 28 WR 90 59 772 10 37 65.6 n n  
## 63 24 RB 73 58 476 1 23 79.5 n n  
## 64 21 RB 86 58 487 4 24 67.4 n n  
## 65 27 WR 85 57 787 3 34 67.1 n n  
## 67 32 TE 78 55 496 4 21 70.5 n n  
## 68 30 WR 81 54 657 8 38 66.7 n n  
## 69 24 WR 87 54 758 2 28 62.1 n n  
## 70 24 TE 77 54 528 1 25 70.1 n n  
## 71 23 TE 90 53 667 3 35 58.9 n n  
## 72 21 RB 71 52 316 1 20 73.2 y n  
## 73 27 RB 64 52 376 3 14 81.3 n n  
## 74 22 RB 70 52 385 1 17 74.3 n n  
## 75 31 TE 74 52 552 1 25 70.3 n n  
## 76 26 WR 101 51 574 4 32 50.5 n n  
## 77 28 RB 74 51 436 1 20 68.9 n n  
## 78 28 RB 65 51 313 2 12 78.5 n n  
## 79 29 TE 70 51 411 3 25 72.9 n n  
## 80 27 WR 87 51 717 5 37 58.6 n n  
## 82 26 WR 79 50 770 4 29 63.3 n n  
## 83 28 TE 70 50 582 3 26 71.4 n n  
## 84 23 WR 83 49 518 4 27 59.0 n n  
## 86 25 RB 59 48 389 2 19 81.4 n n  
## 87 26 WR 75 48 540 3 28 64.0 n n  
## 88 24 WR 75 47 709 7 34 62.7 n n  
## 89 30 TE 70 47 495 2 26 67.1 n n  
## 90 24 TE 67 47 455 4 22 70.1 n n  
## 91 23 RB 58 47 228 2 10 81.0 n n  
## 92 25 WR 75 46 719 5 27 61.3 y y  
## 93 28 TE 61 45 544 6 27 73.8 n n  
## 94 24 WR 81 45 746 7 34 55.6 n n  
## 95 24 RB 54 44 445 4 16 81.5 y n  
## 96 24 RB 63 44 319 5 16 69.8 n n  
## 97 24 RB 49 44 244 2 12 89.8 n n  
## 99 29 TE 61 42 419 6 29 68.9 n n  
## 100 24 WR 76 42 656 5 33 55.3 n n  
## 101 23 WR 184 128 1809 8 80 69.6 y y  
## 102 28 WR 170 119 1710 7 77 70.0 y y  
## 103 33 TE 152 110 1338 12 78 72.4 y y  
## 104 29 WR 154 108 1429 11 74 70.1 y y  
## 105 27 RB 127 107 722 5 36 84.3 n n  
## 106 23 WR 156 107 1359 9 67 68.6 y y  
## 107 23 WR 146 106 1161 6 68 72.6 y n  
## 108 26 WR 142 104 1023 3 53 73.2 n n  
## 109 30 WR 180 100 1516 14 65 55.6 y y  
## 110 25 WR 141 99 925 4 51 70.2 n n  
## 111 24 WR 136 95 1196 7 57 69.9 n n  
## 112 25 WR 141 90 1048 6 47 63.8 n n  
## 113 25 WR 145 88 1496 11 59 60.7 y y  
## 114 22 WR 134 87 1046 9 58 64.9 y n  
## 116 25 TE 129 86 914 6 44 66.7 y n  
## 117 25 TE 43 26 395 3 17 60.5 n n  
## 118 25 TE 86 60 519 3 27 69.8 n n  
## 119 26 RB 108 85 741 5 36 78.7 y n  
## 120 26 RB 43 33 277 1 10 76.7 n n  
## 121 26 RB 65 52 464 4 26 80.0 n n  
## 122 26 WR 133 84 1108 8 55 63.2 n n  
## 123 30 WR 117 84 1033 9 50 71.8 n n  
## 124 22 WR 147 83 1103 4 56 56.5 n n  
## 125 27 WR 121 82 823 5 44 67.8 n n  
## 126 24 WR 114 78 1015 8 47 68.4 n n  
## 127 28 WR 132 78 1160 9 61 59.1 n n  
## 128 26 WR 101 78 933 3 47 77.2 n n  
## 129 29 WR 127 77 1124 6 53 60.6 n n  
## 130 27 WR 120 77 1191 5 56 64.2 y n  
## 131 29 WR 98 75 812 6 42 76.5 n n  
## 132 24 WR 117 75 1356 8 60 64.1 n n  
## 133 23 WR 109 74 1029 7 47 67.9 n n  
## 134 27 TE 113 73 847 5 48 64.6 y n  
## 135 28 TE 98 73 766 4 33 74.5 n n  
## 136 27 RB 83 73 523 3 26 88.0 n n  
## 137 29 TE 108 72 620 3 32 66.7 n n  
## 138 21 WR 117 72 866 4 48 61.5 n n  
## 139 22 WR 119 72 1042 4 48 60.5 n n  
## 140 23 WR 107 72 769 3 39 67.3 n n  
## 141 32 WR 107 70 716 6 45 65.4 n n  
## 142 24 RB 88 69 421 1 13 78.4 n n  
## 143 25 WR 107 67 709 3 33 62.6 n n  
## 144 23 WR 100 67 972 6 41 67.0 n n  
## 145 26 WR 96 67 804 6 38 69.8 n n  
## 146 30 WR 89 66 752 4 39 74.2 n n  
## 147 30 WR 96 64 717 3 35 66.7 n n  
## 148 26 WR 92 64 656 4 36 69.6 n n  
## 149 27 WR 109 64 829 2 38 58.7 n n  
## 150 25 WR 91 63 623 3 32 69.2 n n  
## 151 24 TE 98 63 732 2 37 64.3 n n  
## 152 25 WR 118 63 888 7 44 53.4 n n  
## 153 28 WR 93 63 895 4 39 67.7 n n  
## 154 23 WR 96 61 839 3 33 63.5 n n  
## 155 29 TE 86 60 765 11 36 69.8 y y  
## 156 27 WR 100 60 788 6 45 60.0 n n  
## 157 26 RB 75 60 441 2 17 80.0 n n  
## 158 25 WR 90 60 650 5 32 66.7 n n  
## 159 28 RB 72 59 395 5 19 81.9 n n  
## 160 28 WR 82 58 762 5 42 70.7 n n  
## 161 27 TE 87 58 552 3 25 66.7 n n  
## 162 28 TE 87 58 555 4 29 66.7 n n  
## 163 26 TE 80 58 628 4 36 72.5 n n  
## 165 29 WR 93 57 699 3 29 61.3 n n  
## 166 29 WR 94 57 690 4 34 60.6 n n  
## 167 27 WR 70 57 569 4 29 81.4 n n  
## 168 27 RB 77 57 490 2 17 74.0 n n  
## 169 26 TE 89 57 577 5 27 64.0 n n  
## 170 30 RB 71 56 512 9 28 78.9 n n  
## 171 26 WR 94 56 632 2 25 59.6 n n  
## 172 27 TE 69 55 702 3 40 79.7 n n  
## 174 28 TE 67 53 470 2 20 79.1 n n  
## 175 30 WR 91 53 527 2 31 58.2 n n  
## 176 24 WR 64 52 467 2 19 81.3 n n  
## 177 29 TE 68 52 414 2 26 76.5 n n  
## 178 21 WR 84 52 801 4 38 61.9 n n  
## 179 26 WR 70 51 426 5 31 72.9 n n  
## 180 25 TE 63 50 486 4 24 79.4 n n  
## 181 23 TE 69 50 544 7 27 72.5 n n  
## 182 23 RB 58 50 290 2 13 86.2 n n  
## 183 23 WR 93 48 836 7 35 51.6 n n  
## 184 26 TE 65 48 517 6 25 73.8 y n  
## 185 29 WR 74 48 548 2 24 64.9 n n  
## 186 28 WR 75 48 458 2 23 64.0 n n  
## 187 23 RB 70 48 389 3 17 68.6 n n  
## 188 32 TE 69 47 406 4 24 68.1 n n  
## 190 29 WR 65 46 538 3 29 70.8 n n  
## 191 27 RB 58 46 300 1 13 79.3 n n  
## 192 24 RB 58 46 353 2 18 79.3 n n  
## 193 32 WR 81 46 529 3 29 56.8 n n  
## 194 25 WR 71 46 724 2 30 64.8 n n  
## 195 24 WR 79 46 451 1 24 58.2 n n  
## 196 24 WR 50 32 311 1 17 64.0 n n  
## 198 26 WR 74 43 555 3 25 58.1 n n  
## 199 22 WR 67 42 425 3 20 62.7 n n  
## 200 26 TE 65 42 508 7 25 64.6 n n

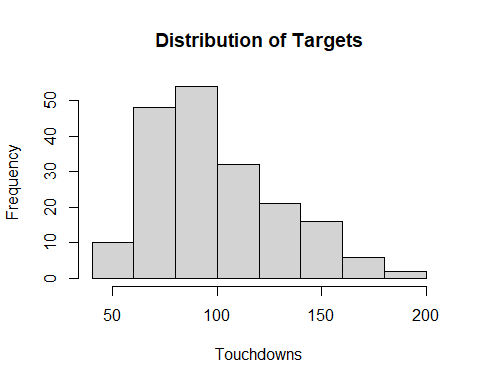
#Changing Pos, PB, and AP into factor variables  
NFLData$Pos <- as.factor(NFLData$Pos)  
NFLData$PB <- as.factor(NFLData$PB)  
NFLData$AP <- as.factor(NFLData$AP)  
  
  
str(NFLData)

## 'data.frame': 189 obs. of 10 variables:  
## $ Age : int 24 29 24 29 26 31 30 26 22 31 ...  
## $ Pos : Factor w/ 3 levels "RB","TE","WR": 3 3 3 2 3 3 3 3 3 3 ...  
## $ Tgt : int 181 171 164 143 156 150 160 158 160 175 ...  
## $ Rec : int 135 119 119 114 109 108 107 106 105 103 ...  
## $ Yds : int 1749 1799 1515 963 1152 1243 1183 1456 1486 1144 ...  
## $ TD : int 12 13 10 4 4 7 8 7 6 8 ...  
## $ X1D : int 80 83 75 47 53 57 58 70 68 59 ...  
## $ Ctch.: num 74.6 69.6 72.6 79.7 69.9 72 66.9 67.1 65.6 58.9 ...  
## $ PB : Factor w/ 2 levels "n","y": 2 2 2 2 1 2 2 2 2 1 ...  
## $ AP : Factor w/ 2 levels "n","y": 2 2 2 1 1 1 1 2 2 1 ...

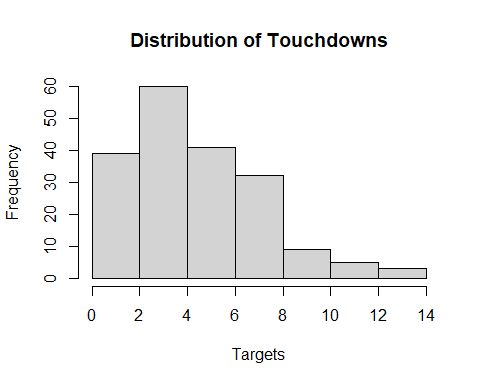
summary(NFLData)

## Age Pos Tgt Rec Yds   
## Min. :21.00 RB: 30 Min. : 43.00 Min. : 26.00 Min. : 228.0   
## 1st Qu.:24.00 TE: 43 1st Qu.: 75.00 1st Qu.: 52.00 1st Qu.: 528.0   
## Median :26.00 WR:116 Median : 93.00 Median : 63.00 Median : 732.0   
## Mean :26.05 Mean : 99.73 Mean : 67.88 Mean : 785.1   
## 3rd Qu.:28.00 3rd Qu.:119.00 3rd Qu.: 78.00 3rd Qu.:1014.0   
## Max. :34.00 Max. :184.00 Max. :135.00 Max. :1809.0   
## TD X1D Ctch. PB AP   
## Min. : 1.000 Min. :10.00 Min. :50.5 n:153 n:170   
## 1st Qu.: 3.000 1st Qu.:26.00 1st Qu.:63.5 y: 36 y: 19   
## Median : 4.000 Median :35.00 Median :68.4   
## Mean : 4.852 Mean :37.72 Mean :68.8   
## 3rd Qu.: 7.000 3rd Qu.:47.00 3rd Qu.:73.4   
## Max. :14.000 Max. :83.00 Max. :91.4

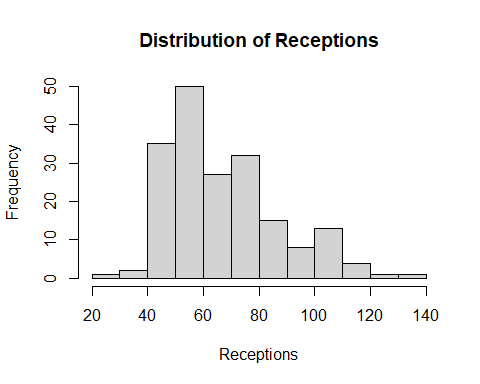
#Histograms and Added Variable Plot  
  
hist(NFLData$Tgt, main = "Distribution of Targets", xlab="Touchdowns")



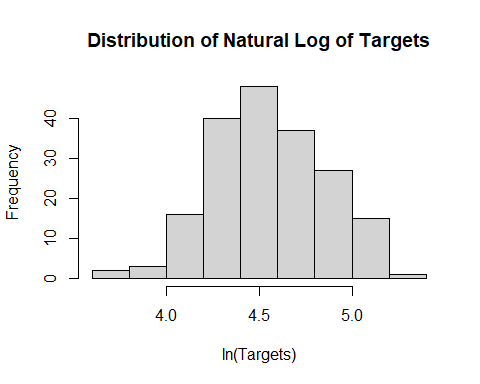
hist(NFLData$TD, main = "Distribution of Touchdowns", xlab="Targets")



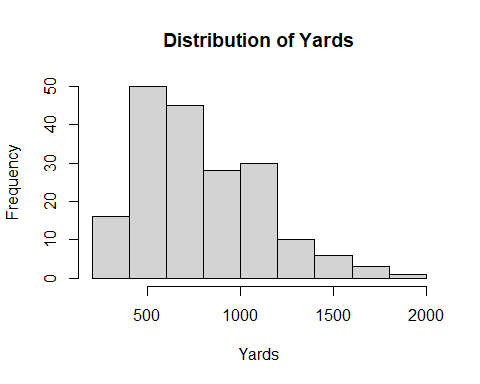
#Performing Logrithmic transformation on Tgt and TD and removing those variables from the dataset  
NFLData$lnTD <- log(NFLData$TD)  
NFLData$lnTgt <- log(NFLData$Tgt)  
NFLData <- NFLData[,-c(3,6)]  
  
#Histograms of all continuous variables  
hist(NFLData$Rec, main = "Distribution of Receptions", xlab="Receptions")



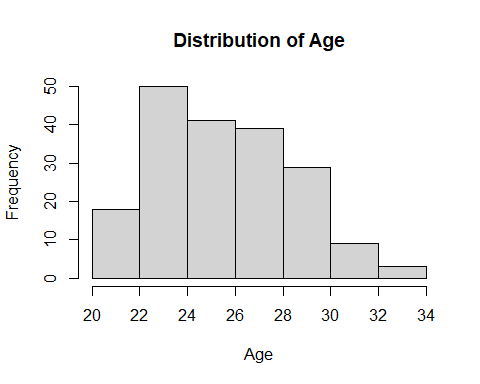
hist(NFLData$lnTgt, main = "Distribution of Natural Log of Targets", xlab="ln(Targets)")



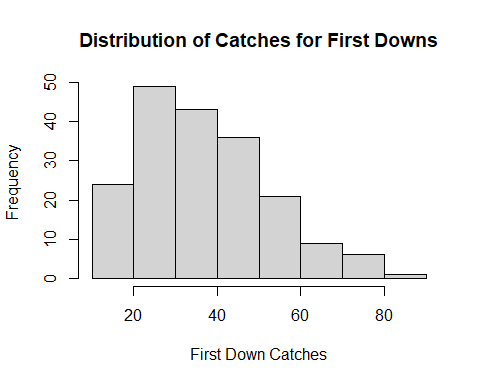
hist(NFLData$Yds, main = "Distribution of Yards", xlab="Yards")



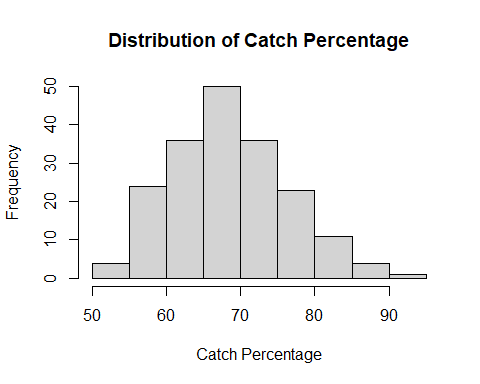
hist(NFLData$Age, main = "Distribution of Age", xlab="Age")



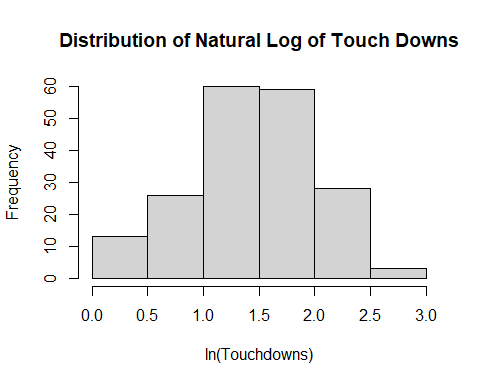
hist(NFLData$X1D, main = "Distribution of Catches for First Downs", xlab="First Down Catches")



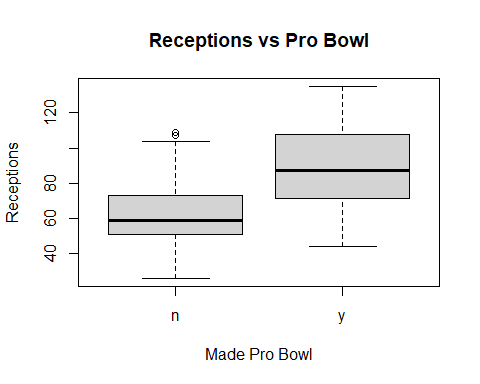
hist(NFLData$Ctch., main = "Distribution of Catch Percentage", xlab="Catch Percentage")



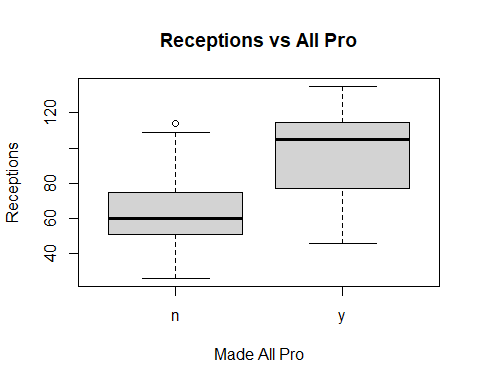
hist(NFLData$lnTD, main = "Distribution of Natural Log of Touch Downs", xlab="ln(Touchdowns)")



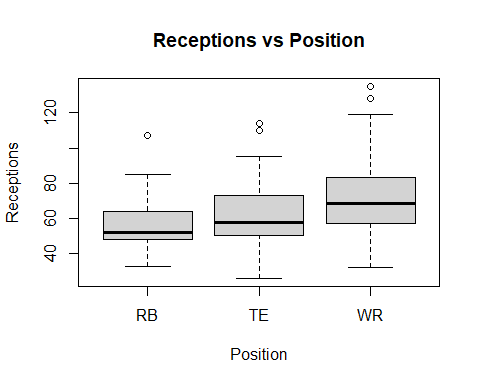
#Boxplots for factor variables  
plot(NFLData$PB,NFLData$Rec, ylab="Receptions", xlab="Made Pro Bowl", main="Receptions vs Pro Bowl")



plot(NFLData$AP,NFLData$Rec, ylab="Receptions", xlab="Made All Pro", main="Receptions vs All Pro")



plot(NFLData$Pos,NFLData$Rec, ylab="Receptions", xlab="Position", main="Receptions vs Position")



#Distributions of Factor variables in table format  
table(NFLData$Pos)

##   
## RB TE WR   
## 30 43 116

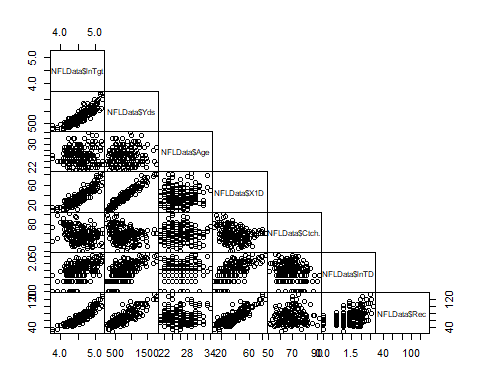
table(NFLData$PB)

##   
## n y   
## 153 36

table(NFLData$AP)

##   
## n y   
## 170 19

#Added Variable Plot and correlation table  
pairs(~ NFLData$lnTgt + NFLData$Yds + NFLData$Age + NFLData$X1D + NFLData$Ctch. + NFLData$lnTD + NFLData$Rec , upper.panel=NULL, gap=0)



cor(NFLData[,-c(2,7,8)])

## Age Rec Yds X1D Ctch. lnTD  
## Age 1.000000000 0.07221660 0.02951934 0.07925708 0.003333029 0.1193617  
## Rec 0.072216605 1.00000000 0.84171797 0.85898094 0.049505664 0.4964593  
## Yds 0.029519336 0.84171797 1.00000000 0.95994227 -0.292254531 0.6550475  
## X1D 0.079257084 0.85898094 0.95994227 1.00000000 -0.272856584 0.6719502  
## Ctch. 0.003333029 0.04950566 -0.29225453 -0.27285658 1.000000000 -0.2374472  
## lnTD 0.119361698 0.49645929 0.65504747 0.67195020 -0.237447197 1.0000000  
## lnTgt 0.065237454 0.90903031 0.88863942 0.89601362 -0.332632085 0.5568709  
## lnTgt  
## Age 0.06523745  
## Rec 0.90903031  
## Yds 0.88863942  
## X1D 0.89601362  
## Ctch. -0.33263209  
## lnTD 0.55687092  
## lnTgt 1.00000000

#Full Model with all Summary Measures, assumption plots, and VIF table  
fullmodel <- lm(Rec ~ Age + Pos + lnTgt + Yds + lnTD + X1D + Ctch. + PB + AP, data=NFLData)  
summary(fullmodel)

##   
## Call:  
## lm(formula = Rec ~ Age + Pos + lnTgt + Yds + lnTD + X1D + Ctch. +   
## PB + AP, data = NFLData)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -7.1891 -1.5979 -0.4132 0.9465 17.2772   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -2.704e+02 9.356e+00 -28.906 < 2e-16 \*\*\*  
## Age 1.749e-02 8.853e-02 0.198 0.843613   
## PosTE -5.308e-01 9.582e-01 -0.554 0.580349   
## PosWR -9.785e-01 1.134e+00 -0.863 0.389413   
## lnTgt 5.823e+01 2.005e+00 29.037 < 2e-16 \*\*\*  
## Yds 1.581e-03 2.956e-03 0.535 0.593486   
## lnTD -1.606e+00 5.457e-01 -2.942 0.003693 \*\*   
## X1D 2.169e-01 6.348e-02 3.416 0.000786 \*\*\*  
## Ctch. 9.533e-01 4.258e-02 22.389 < 2e-16 \*\*\*  
## PBy 1.834e+00 8.851e-01 2.072 0.039677 \*   
## APy 1.856e+00 1.184e+00 1.568 0.118551   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.323 on 178 degrees of freedom  
## Multiple R-squared: 0.9751, Adjusted R-squared: 0.9737   
## F-statistic: 697.3 on 10 and 178 DF, p-value: < 2.2e-16

summary(fullmodel)$r.squared

## [1] 0.9751087

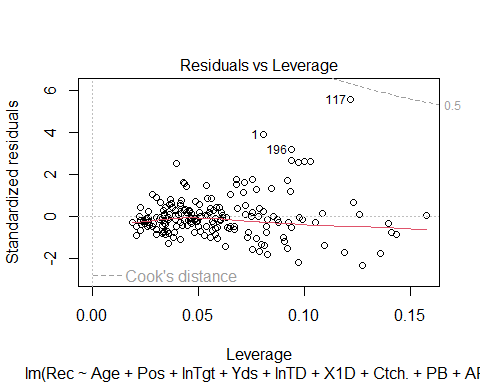
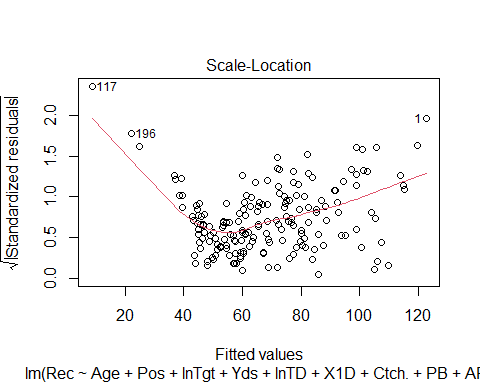
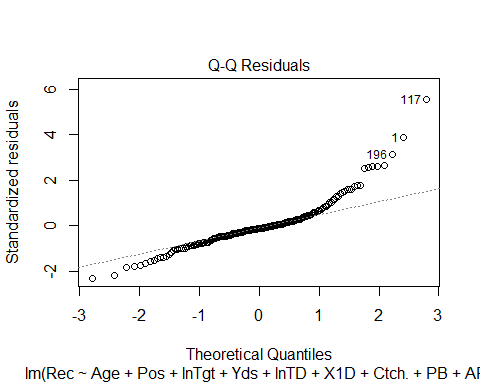
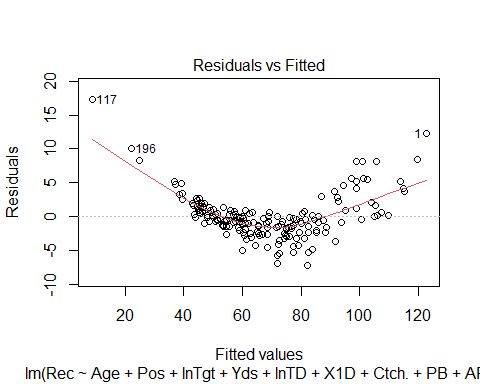
summary(fullmodel)$adj.r.squared

## [1] 0.9737103

summary(fullmodel)$sigma

## [1] 3.323353

plot(fullmodel)



vif(fullmodel)

## GVIF Df GVIF^(1/(2\*Df))  
## Age 1.088921 1 1.043514  
## Pos 3.280311 2 1.345794  
## lnTgt 6.431010 1 2.535944  
## Yds 16.198282 1 4.024709  
## lnTD 1.954849 1 1.398159  
## X1D 16.863383 1 4.106505  
## Ctch. 1.940222 1 1.392919  
## PB 2.067321 1 1.437818  
## AP 2.167355 1 1.472194

#Performing logarithmic transformation on Rec and removing it from the dataset  
NFLData$lnRec <- log(NFLData$Rec)  
NFLData <- NFLData[,-c(3)]  
  
  
#New Full Model with transformation on target variable  
fullmodel2 <- lm(lnRec ~ Age + Pos + lnTgt + Yds + lnTD + X1D + Ctch. + PB + AP, data=NFLData)  
summary(fullmodel2)

##   
## Call:  
## lm(formula = lnRec ~ Age + Pos + lnTgt + Yds + lnTD + X1D + Ctch. +   
## PB + AP, data = NFLData)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.033488 -0.002141 0.002286 0.004519 0.016553   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.385e+00 2.271e-02 -61.012 < 2e-16 \*\*\*  
## Age -1.212e-04 2.149e-04 -0.564 0.573563   
## PosTE 1.135e-02 2.326e-03 4.882 2.33e-06 \*\*\*  
## PosWR 9.490e-03 2.752e-03 3.448 0.000705 \*\*\*  
## lnTgt 9.965e-01 4.867e-03 204.725 < 2e-16 \*\*\*  
## Yds -1.568e-06 7.174e-06 -0.219 0.827279   
## lnTD -1.164e-03 1.324e-03 -0.879 0.380512   
## X1D 1.469e-04 1.541e-04 0.953 0.341807   
## Ctch. 1.472e-02 1.033e-04 142.443 < 2e-16 \*\*\*  
## PBy 7.910e-04 2.148e-03 0.368 0.713160   
## APy 7.820e-04 2.872e-03 0.272 0.785738   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.008066 on 178 degrees of freedom  
## Multiple R-squared: 0.9993, Adjusted R-squared: 0.9992   
## F-statistic: 2.438e+04 on 10 and 178 DF, p-value: < 2.2e-16

summary(fullmodel2)$r.squared

## [1] 0.9992703

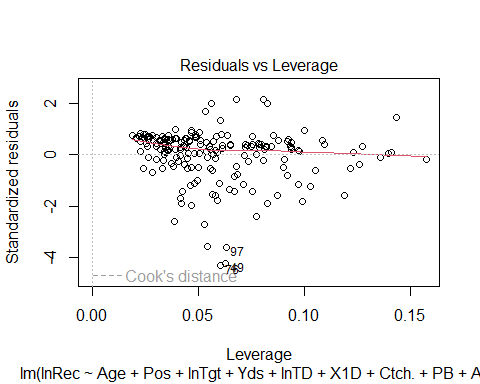
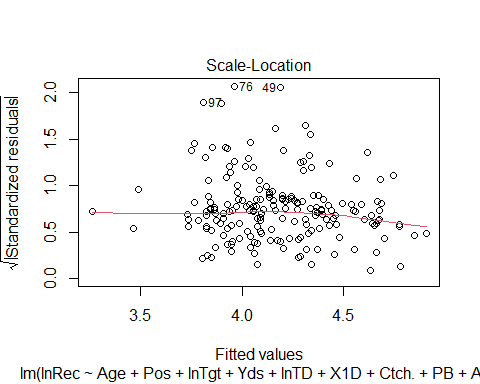
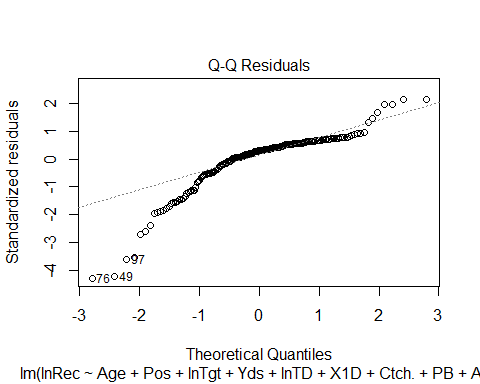
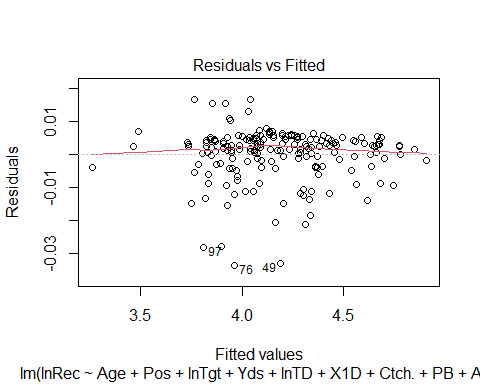
summary(fullmodel2)$adj.r.squared

## [1] 0.9992293

summary(fullmodel2)$sigma

## [1] 0.008065906

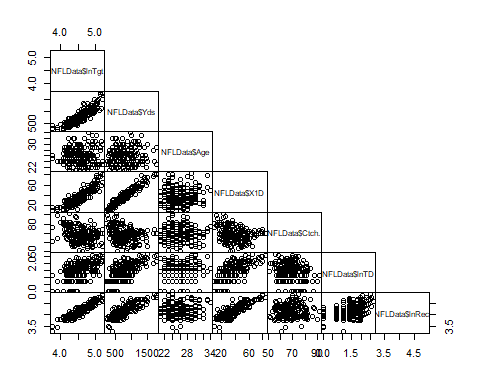
#Correlation table, assumption plots, VIF table, and added variable plot for new full model  
plot(fullmodel2)



cor(NFLData[,-c(2,6,7)])

## Age Yds X1D Ctch. lnTD lnTgt  
## Age 1.000000000 0.02951934 0.07925708 0.003333029 0.1193617 0.06523745  
## Yds 0.029519336 1.00000000 0.95994227 -0.292254531 0.6550475 0.88863942  
## X1D 0.079257084 0.95994227 1.00000000 -0.272856584 0.6719502 0.89601362  
## Ctch. 0.003333029 -0.29225453 -0.27285658 1.000000000 -0.2374472 -0.33263209  
## lnTD 0.119361698 0.65504747 0.67195020 -0.237447197 1.0000000 0.55687092  
## lnTgt 0.065237454 0.88863942 0.89601362 -0.332632085 0.5568709 1.00000000  
## lnRec 0.071207258 0.82689321 0.84314247 0.044497031 0.4967189 0.92691228  
## lnRec  
## Age 0.07120726  
## Yds 0.82689321  
## X1D 0.84314247  
## Ctch. 0.04449703  
## lnTD 0.49671885  
## lnTgt 0.92691228  
## lnRec 1.00000000

pairs(~ NFLData$lnTgt + NFLData$Yds + NFLData$Age + NFLData$X1D + NFLData$Ctch. + NFLData$lnTD + NFLData$lnRec , upper.panel=NULL, gap=0)



vif(fullmodel2)

## GVIF Df GVIF^(1/(2\*Df))  
## Age 1.088921 1 1.043514  
## Pos 3.280311 2 1.345794  
## lnTgt 6.431010 1 2.535944  
## Yds 16.198282 1 4.024709  
## lnTD 1.954849 1 1.398159  
## X1D 16.863383 1 4.106505  
## Ctch. 1.940222 1 1.392919  
## PB 2.067321 1 1.437818  
## AP 2.167355 1 1.472194

#Standardized residual, leverage, and cook's distance for all 3 outliers  
  
#Model with observation 49  
outliermodel1 <- lm(lnRec~.,data=NFLData, subset=-c(76,97))  
  
rstandard(outliermodel1)[49]

## 49   
## -4.19518

hatvalues(outliermodel1)[49]

## 49   
## 0.06313881

cooks.distance(outliermodel1)[49]

## 49   
## 0.1078275

#Model with observation 76  
outliermodel2 <- lm(lnRec~.,data=NFLData, subset=-c(49,97))  
  
rstandard(outliermodel1)[76]

## 80   
## -0.5210353

hatvalues(outliermodel1)[76]

## 80   
## 0.02456713

cooks.distance(outliermodel1)[76]

## 80   
## 0.0006215823

#Model with observation 97  
outliermodel3 <- lm(lnRec~.,data=NFLData, subset=-c(49,76))  
  
rstandard(outliermodel1)[97]

## 105   
## -1.139965

hatvalues(outliermodel1)[97]

## 105   
## 0.09764723

cooks.distance(outliermodel1)[97]

## 105   
## 0.01278421

qf(.95,10,187)

## [1] 1.88162

#Confidence Interval  
  
samplemean <- mean(exp(NFLData$lnRec))  
samplesd <- sd(exp(NFLData$lnRec))  
samplesize <- length(NFLData$lnRec)  
samplese <- samplesd/sqrt(samplesize)  
  
t <- qt(.975, df=(samplesize-1))  
  
lowerbound <- samplemean - (t\*samplese)  
upperbound <- samplemean + (t\*samplese)  
  
interval <- c(lowerbound, upperbound)  
interval

## [1] 64.94253 70.82467

#Forward Selection  
nullmodel <- lm(lnRec ~ 1, data=NFLData)  
  
#AIC as selection criteria  
stepAIC(nullmodel, scope = list(lower=nullmodel, upper=fullmodel2), k=2, direction = "forward")

## Start: AIC=-466.21  
## lnRec ~ 1  
##   
## Df Sum of Sq RSS AIC  
## + lnTgt 1 13.6354 2.2351 -834.68  
## + X1D 1 11.2821 4.5883 -698.75  
## + Yds 1 10.8515 5.0190 -681.79  
## + lnTD 1 3.9157 11.9547 -517.76  
## + PB 1 3.4242 12.4463 -510.14  
## + AP 1 2.5995 13.2710 -498.02  
## + Pos 2 1.4651 14.4053 -480.51  
## <none> 15.8704 -466.21  
## + Age 1 0.0805 15.7900 -465.17  
## + Ctch. 1 0.0314 15.8390 -464.58  
##   
## Step: AIC=-834.68  
## lnRec ~ lnTgt  
##   
## Df Sum of Sq RSS AIC  
## + Ctch. 1 2.22134 0.01376 -1794.81  
## + Pos 2 0.73795 1.49714 -906.42  
## + PB 1 0.07567 2.15943 -839.19  
## + AP 1 0.04215 2.19295 -836.28  
## <none> 2.23509 -834.68  
## + X1D 1 0.01281 2.22228 -833.77  
## + lnTD 1 0.00870 2.22639 -833.42  
## + Age 1 0.00184 2.23325 -832.84  
## + Yds 1 0.00077 2.23432 -832.75  
##   
## Step: AIC=-1794.81  
## lnRec ~ lnTgt + Ctch.  
##   
## Df Sum of Sq RSS AIC  
## + Pos 2 0.00199383 0.011761 -1820.4  
## + X1D 1 0.00055657 0.013199 -1800.6  
## + Yds 1 0.00026649 0.013489 -1796.5  
## <none> 0.013755 -1794.8  
## + AP 1 0.00004675 0.013708 -1793.5  
## + lnTD 1 0.00003682 0.013718 -1793.3  
## + PB 1 0.00003056 0.013725 -1793.2  
## + Age 1 0.00001039 0.013745 -1793.0  
##   
## Step: AIC=-1820.4  
## lnRec ~ lnTgt + Ctch. + Pos  
##   
## Df Sum of Sq RSS AIC  
## <none> 0.011761 -1820.4  
## + X1D 1 8.8371e-05 0.011673 -1819.8  
## + Yds 1 5.0363e-05 0.011711 -1819.2  
## + AP 1 4.6974e-05 0.011714 -1819.2  
## + PB 1 4.4348e-05 0.011717 -1819.1  
## + Age 1 2.1714e-05 0.011740 -1818.8  
## + lnTD 1 3.9500e-06 0.011757 -1818.5

##   
## Call:  
## lm(formula = lnRec ~ lnTgt + Ctch. + Pos, data = NFLData)  
##   
## Coefficients:  
## (Intercept) lnTgt Ctch. PosTE PosWR   
## -1.40937 1.00098 0.01476 0.01175 0.01014

forwardmodel <- stepAIC(nullmodel, scope = list(lower=nullmodel, upper=fullmodel2), k=2, direction = "forward")

## Start: AIC=-466.21  
## lnRec ~ 1  
##   
## Df Sum of Sq RSS AIC  
## + lnTgt 1 13.6354 2.2351 -834.68  
## + X1D 1 11.2821 4.5883 -698.75  
## + Yds 1 10.8515 5.0190 -681.79  
## + lnTD 1 3.9157 11.9547 -517.76  
## + PB 1 3.4242 12.4463 -510.14  
## + AP 1 2.5995 13.2710 -498.02  
## + Pos 2 1.4651 14.4053 -480.51  
## <none> 15.8704 -466.21  
## + Age 1 0.0805 15.7900 -465.17  
## + Ctch. 1 0.0314 15.8390 -464.58  
##   
## Step: AIC=-834.68  
## lnRec ~ lnTgt  
##   
## Df Sum of Sq RSS AIC  
## + Ctch. 1 2.22134 0.01376 -1794.81  
## + Pos 2 0.73795 1.49714 -906.42  
## + PB 1 0.07567 2.15943 -839.19  
## + AP 1 0.04215 2.19295 -836.28  
## <none> 2.23509 -834.68  
## + X1D 1 0.01281 2.22228 -833.77  
## + lnTD 1 0.00870 2.22639 -833.42  
## + Age 1 0.00184 2.23325 -832.84  
## + Yds 1 0.00077 2.23432 -832.75  
##   
## Step: AIC=-1794.81  
## lnRec ~ lnTgt + Ctch.  
##   
## Df Sum of Sq RSS AIC  
## + Pos 2 0.00199383 0.011761 -1820.4  
## + X1D 1 0.00055657 0.013199 -1800.6  
## + Yds 1 0.00026649 0.013489 -1796.5  
## <none> 0.013755 -1794.8  
## + AP 1 0.00004675 0.013708 -1793.5  
## + lnTD 1 0.00003682 0.013718 -1793.3  
## + PB 1 0.00003056 0.013725 -1793.2  
## + Age 1 0.00001039 0.013745 -1793.0  
##   
## Step: AIC=-1820.4  
## lnRec ~ lnTgt + Ctch. + Pos  
##   
## Df Sum of Sq RSS AIC  
## <none> 0.011761 -1820.4  
## + X1D 1 8.8371e-05 0.011673 -1819.8  
## + Yds 1 5.0363e-05 0.011711 -1819.2  
## + AP 1 4.6974e-05 0.011714 -1819.2  
## + PB 1 4.4348e-05 0.011717 -1819.1  
## + Age 1 2.1714e-05 0.011740 -1818.8  
## + lnTD 1 3.9500e-06 0.011757 -1818.5

forwardmodel

##   
## Call:  
## lm(formula = lnRec ~ lnTgt + Ctch. + Pos, data = NFLData)  
##   
## Coefficients:  
## (Intercept) lnTgt Ctch. PosTE PosWR   
## -1.40937 1.00098 0.01476 0.01175 0.01014

summary(forwardmodel)

##   
## Call:  
## lm(formula = lnRec ~ lnTgt + Ctch. + Pos, data = NFLData)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.033914 -0.002325 0.002629 0.004426 0.016349   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.409e+00 1.227e-02 -114.840 < 2e-16 \*\*\*  
## lnTgt 1.001e+00 2.232e-03 448.556 < 2e-16 \*\*\*  
## Ctch. 1.476e-02 9.682e-05 152.440 < 2e-16 \*\*\*  
## PosTE 1.175e-02 2.109e-03 5.574 8.76e-08 \*\*\*  
## PosWR 1.014e-02 2.328e-03 4.356 2.20e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.007995 on 184 degrees of freedom  
## Multiple R-squared: 0.9993, Adjusted R-squared: 0.9992   
## F-statistic: 6.203e+04 on 4 and 184 DF, p-value: < 2.2e-16

summary(forwardmodel)$r.squared

## [1] 0.9992589

summary(forwardmodel)$adj.r.squared

## [1] 0.9992428

summary(forwardmodel)$sigma

## [1] 0.007995028

#BIC as selection criteria  
stepAIC(nullmodel, scope = list(lower=nullmodel, upper=fullmodel2), k=log(187), direction = "forward")

## Start: AIC=-462.98  
## lnRec ~ 1  
##   
## Df Sum of Sq RSS AIC  
## + lnTgt 1 13.6354 2.2351 -828.22  
## + X1D 1 11.2821 4.5883 -692.28  
## + Yds 1 10.8515 5.0190 -675.33  
## + lnTD 1 3.9157 11.9547 -511.29  
## + PB 1 3.4242 12.4463 -503.68  
## + AP 1 2.5995 13.2710 -491.55  
## + Pos 2 1.4651 14.4053 -470.82  
## <none> 15.8704 -462.98  
## + Age 1 0.0805 15.7900 -458.71  
## + Ctch. 1 0.0314 15.8390 -458.12  
##   
## Step: AIC=-828.22  
## lnRec ~ lnTgt  
##   
## Df Sum of Sq RSS AIC  
## + Ctch. 1 2.22134 0.01376 -1785.11  
## + Pos 2 0.73795 1.49714 -893.49  
## + PB 1 0.07567 2.15943 -829.50  
## <none> 2.23509 -828.22  
## + AP 1 0.04215 2.19295 -826.59  
## + X1D 1 0.01281 2.22228 -824.07  
## + lnTD 1 0.00870 2.22639 -823.72  
## + Age 1 0.00184 2.23325 -823.14  
## + Yds 1 0.00077 2.23432 -823.05  
##   
## Step: AIC=-1785.11  
## lnRec ~ lnTgt + Ctch.  
##   
## Df Sum of Sq RSS AIC  
## + Pos 2 0.00199383 0.011761 -1804.2  
## + X1D 1 0.00055657 0.013199 -1787.7  
## <none> 0.013755 -1785.1  
## + Yds 1 0.00026649 0.013489 -1783.6  
## + AP 1 0.00004675 0.013708 -1780.5  
## + lnTD 1 0.00003682 0.013718 -1780.4  
## + PB 1 0.00003056 0.013725 -1780.3  
## + Age 1 0.00001039 0.013745 -1780.0  
##   
## Step: AIC=-1804.25  
## lnRec ~ lnTgt + Ctch. + Pos  
##   
## Df Sum of Sq RSS AIC  
## <none> 0.011761 -1804.2  
## + X1D 1 8.8371e-05 0.011673 -1800.4  
## + Yds 1 5.0363e-05 0.011711 -1799.8  
## + AP 1 4.6974e-05 0.011714 -1799.8  
## + PB 1 4.4348e-05 0.011717 -1799.7  
## + Age 1 2.1714e-05 0.011740 -1799.4  
## + lnTD 1 3.9500e-06 0.011757 -1799.1

##   
## Call:  
## lm(formula = lnRec ~ lnTgt + Ctch. + Pos, data = NFLData)  
##   
## Coefficients:  
## (Intercept) lnTgt Ctch. PosTE PosWR   
## -1.40937 1.00098 0.01476 0.01175 0.01014

forwardmodel2 <- stepAIC(nullmodel, scope = list(lower=nullmodel, upper=fullmodel2), k=log(187), direction = "forward")

## Start: AIC=-462.98  
## lnRec ~ 1  
##   
## Df Sum of Sq RSS AIC  
## + lnTgt 1 13.6354 2.2351 -828.22  
## + X1D 1 11.2821 4.5883 -692.28  
## + Yds 1 10.8515 5.0190 -675.33  
## + lnTD 1 3.9157 11.9547 -511.29  
## + PB 1 3.4242 12.4463 -503.68  
## + AP 1 2.5995 13.2710 -491.55  
## + Pos 2 1.4651 14.4053 -470.82  
## <none> 15.8704 -462.98  
## + Age 1 0.0805 15.7900 -458.71  
## + Ctch. 1 0.0314 15.8390 -458.12  
##   
## Step: AIC=-828.22  
## lnRec ~ lnTgt  
##   
## Df Sum of Sq RSS AIC  
## + Ctch. 1 2.22134 0.01376 -1785.11  
## + Pos 2 0.73795 1.49714 -893.49  
## + PB 1 0.07567 2.15943 -829.50  
## <none> 2.23509 -828.22  
## + AP 1 0.04215 2.19295 -826.59  
## + X1D 1 0.01281 2.22228 -824.07  
## + lnTD 1 0.00870 2.22639 -823.72  
## + Age 1 0.00184 2.23325 -823.14  
## + Yds 1 0.00077 2.23432 -823.05  
##   
## Step: AIC=-1785.11  
## lnRec ~ lnTgt + Ctch.  
##   
## Df Sum of Sq RSS AIC  
## + Pos 2 0.00199383 0.011761 -1804.2  
## + X1D 1 0.00055657 0.013199 -1787.7  
## <none> 0.013755 -1785.1  
## + Yds 1 0.00026649 0.013489 -1783.6  
## + AP 1 0.00004675 0.013708 -1780.5  
## + lnTD 1 0.00003682 0.013718 -1780.4  
## + PB 1 0.00003056 0.013725 -1780.3  
## + Age 1 0.00001039 0.013745 -1780.0  
##   
## Step: AIC=-1804.25  
## lnRec ~ lnTgt + Ctch. + Pos  
##   
## Df Sum of Sq RSS AIC  
## <none> 0.011761 -1804.2  
## + X1D 1 8.8371e-05 0.011673 -1800.4  
## + Yds 1 5.0363e-05 0.011711 -1799.8  
## + AP 1 4.6974e-05 0.011714 -1799.8  
## + PB 1 4.4348e-05 0.011717 -1799.7  
## + Age 1 2.1714e-05 0.011740 -1799.4  
## + lnTD 1 3.9500e-06 0.011757 -1799.1

forwardmodel2

##   
## Call:  
## lm(formula = lnRec ~ lnTgt + Ctch. + Pos, data = NFLData)  
##   
## Coefficients:  
## (Intercept) lnTgt Ctch. PosTE PosWR   
## -1.40937 1.00098 0.01476 0.01175 0.01014

summary(forwardmodel2)

##   
## Call:  
## lm(formula = lnRec ~ lnTgt + Ctch. + Pos, data = NFLData)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.033914 -0.002325 0.002629 0.004426 0.016349   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.409e+00 1.227e-02 -114.840 < 2e-16 \*\*\*  
## lnTgt 1.001e+00 2.232e-03 448.556 < 2e-16 \*\*\*  
## Ctch. 1.476e-02 9.682e-05 152.440 < 2e-16 \*\*\*  
## PosTE 1.175e-02 2.109e-03 5.574 8.76e-08 \*\*\*  
## PosWR 1.014e-02 2.328e-03 4.356 2.20e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.007995 on 184 degrees of freedom  
## Multiple R-squared: 0.9993, Adjusted R-squared: 0.9992   
## F-statistic: 6.203e+04 on 4 and 184 DF, p-value: < 2.2e-16

summary(forwardmodel2)$r.squared

## [1] 0.9992589

summary(forwardmodel2)$adj.r.squared

## [1] 0.9992428

summary(forwardmodel2)$sigma

## [1] 0.007995028

#Backwards Selection  
  
#AIC as selection criteria  
stepAIC(fullmodel2, scope = list(lower=nullmodel, upper=fullmodel2), k=2, direction = "backward")

## Start: AIC=-1811.33  
## lnRec ~ Age + Pos + lnTgt + Yds + lnTD + X1D + Ctch. + PB + AP  
##   
## Df Sum of Sq RSS AIC  
## - Yds 1 0.00000 0.01158 -1813.28  
## - AP 1 0.00000 0.01159 -1813.26  
## - PB 1 0.00001 0.01159 -1813.19  
## - Age 1 0.00002 0.01160 -1813.00  
## - lnTD 1 0.00005 0.01163 -1812.52  
## - X1D 1 0.00006 0.01164 -1812.37  
## <none> 0.01158 -1811.33  
## - Pos 2 0.00156 0.01314 -1791.46  
## - Ctch. 1 1.32004 1.33162 -916.56  
## - lnTgt 1 2.72676 2.73835 -780.30  
##   
## Step: AIC=-1813.28  
## lnRec ~ Age + Pos + lnTgt + lnTD + X1D + Ctch. + PB + AP  
##   
## Df Sum of Sq RSS AIC  
## - AP 1 0.00000 0.01159 -1815.23  
## - PB 1 0.00001 0.01159 -1815.15  
## - Age 1 0.00002 0.01160 -1814.98  
## - lnTD 1 0.00005 0.01164 -1814.43  
## - X1D 1 0.00008 0.01166 -1814.02  
## <none> 0.01158 -1813.28  
## - Pos 2 0.00157 0.01315 -1793.33  
## - Ctch. 1 1.32033 1.33191 -918.52  
## - lnTgt 1 2.96278 2.97436 -766.67  
##   
## Step: AIC=-1815.23  
## lnRec ~ Age + Pos + lnTgt + lnTD + X1D + Ctch. + PB  
##   
## Df Sum of Sq RSS AIC  
## - PB 1 0.00002 0.01160 -1816.94  
## - Age 1 0.00002 0.01161 -1816.93  
## - lnTD 1 0.00005 0.01164 -1816.40  
## - X1D 1 0.00009 0.01168 -1815.69  
## <none> 0.01159 -1815.23  
## - Pos 2 0.00157 0.01316 -1795.21  
## - Ctch. 1 1.32139 1.33297 -920.37  
## - lnTgt 1 3.10432 3.11591 -759.89  
##   
## Step: AIC=-1816.94  
## lnRec ~ Age + Pos + lnTgt + lnTD + X1D + Ctch.  
##   
## Df Sum of Sq RSS AIC  
## - Age 1 0.00002 0.01162 -1818.64  
## - lnTD 1 0.00004 0.01165 -1818.21  
## <none> 0.01160 -1816.94  
## - X1D 1 0.00013 0.01174 -1816.80  
## - Pos 2 0.00155 0.01316 -1797.21  
## - Ctch. 1 1.32753 1.33914 -921.50  
## - lnTgt 1 3.10477 3.11638 -761.86  
##   
## Step: AIC=-1818.64  
## lnRec ~ Pos + lnTgt + lnTD + X1D + Ctch.  
##   
## Df Sum of Sq RSS AIC  
## - lnTD 1 0.00005 0.01167 -1819.83  
## <none> 0.01162 -1818.64  
## - X1D 1 0.00013 0.01176 -1818.47  
## - Pos 2 0.00154 0.01316 -1799.10  
## - Ctch. 1 1.32901 1.34064 -923.29  
## - lnTgt 1 3.10760 3.11922 -763.69  
##   
## Step: AIC=-1819.83  
## lnRec ~ Pos + lnTgt + X1D + Ctch.  
##   
## Df Sum of Sq RSS AIC  
## - X1D 1 0.0001 0.0118 -1820.40  
## <none> 0.0117 -1819.83  
## - Pos 2 0.0015 0.0132 -1800.61  
## - Ctch. 1 1.3504 1.3621 -922.29  
## - lnTgt 1 3.1861 3.1978 -760.99  
##   
## Step: AIC=-1820.4  
## lnRec ~ Pos + lnTgt + Ctch.  
##   
## Df Sum of Sq RSS AIC  
## <none> 0.0118 -1820.40  
## - Pos 2 0.0020 0.0138 -1794.81  
## - Ctch. 1 1.4854 1.4971 -906.42  
## - lnTgt 1 12.8610 12.8727 -499.77

##   
## Call:  
## lm(formula = lnRec ~ Pos + lnTgt + Ctch., data = NFLData)  
##   
## Coefficients:  
## (Intercept) PosTE PosWR lnTgt Ctch.   
## -1.40937 0.01175 0.01014 1.00098 0.01476

backwardsmodel <- stepAIC(fullmodel2, scope = list(lower=nullmodel, upper=fullmodel2), k=2, direction = "backward")

## Start: AIC=-1811.33  
## lnRec ~ Age + Pos + lnTgt + Yds + lnTD + X1D + Ctch. + PB + AP  
##   
## Df Sum of Sq RSS AIC  
## - Yds 1 0.00000 0.01158 -1813.28  
## - AP 1 0.00000 0.01159 -1813.26  
## - PB 1 0.00001 0.01159 -1813.19  
## - Age 1 0.00002 0.01160 -1813.00  
## - lnTD 1 0.00005 0.01163 -1812.52  
## - X1D 1 0.00006 0.01164 -1812.37  
## <none> 0.01158 -1811.33  
## - Pos 2 0.00156 0.01314 -1791.46  
## - Ctch. 1 1.32004 1.33162 -916.56  
## - lnTgt 1 2.72676 2.73835 -780.30  
##   
## Step: AIC=-1813.28  
## lnRec ~ Age + Pos + lnTgt + lnTD + X1D + Ctch. + PB + AP  
##   
## Df Sum of Sq RSS AIC  
## - AP 1 0.00000 0.01159 -1815.23  
## - PB 1 0.00001 0.01159 -1815.15  
## - Age 1 0.00002 0.01160 -1814.98  
## - lnTD 1 0.00005 0.01164 -1814.43  
## - X1D 1 0.00008 0.01166 -1814.02  
## <none> 0.01158 -1813.28  
## - Pos 2 0.00157 0.01315 -1793.33  
## - Ctch. 1 1.32033 1.33191 -918.52  
## - lnTgt 1 2.96278 2.97436 -766.67  
##   
## Step: AIC=-1815.23  
## lnRec ~ Age + Pos + lnTgt + lnTD + X1D + Ctch. + PB  
##   
## Df Sum of Sq RSS AIC  
## - PB 1 0.00002 0.01160 -1816.94  
## - Age 1 0.00002 0.01161 -1816.93  
## - lnTD 1 0.00005 0.01164 -1816.40  
## - X1D 1 0.00009 0.01168 -1815.69  
## <none> 0.01159 -1815.23  
## - Pos 2 0.00157 0.01316 -1795.21  
## - Ctch. 1 1.32139 1.33297 -920.37  
## - lnTgt 1 3.10432 3.11591 -759.89  
##   
## Step: AIC=-1816.94  
## lnRec ~ Age + Pos + lnTgt + lnTD + X1D + Ctch.  
##   
## Df Sum of Sq RSS AIC  
## - Age 1 0.00002 0.01162 -1818.64  
## - lnTD 1 0.00004 0.01165 -1818.21  
## <none> 0.01160 -1816.94  
## - X1D 1 0.00013 0.01174 -1816.80  
## - Pos 2 0.00155 0.01316 -1797.21  
## - Ctch. 1 1.32753 1.33914 -921.50  
## - lnTgt 1 3.10477 3.11638 -761.86  
##   
## Step: AIC=-1818.64  
## lnRec ~ Pos + lnTgt + lnTD + X1D + Ctch.  
##   
## Df Sum of Sq RSS AIC  
## - lnTD 1 0.00005 0.01167 -1819.83  
## <none> 0.01162 -1818.64  
## - X1D 1 0.00013 0.01176 -1818.47  
## - Pos 2 0.00154 0.01316 -1799.10  
## - Ctch. 1 1.32901 1.34064 -923.29  
## - lnTgt 1 3.10760 3.11922 -763.69  
##   
## Step: AIC=-1819.83  
## lnRec ~ Pos + lnTgt + X1D + Ctch.  
##   
## Df Sum of Sq RSS AIC  
## - X1D 1 0.0001 0.0118 -1820.40  
## <none> 0.0117 -1819.83  
## - Pos 2 0.0015 0.0132 -1800.61  
## - Ctch. 1 1.3504 1.3621 -922.29  
## - lnTgt 1 3.1861 3.1978 -760.99  
##   
## Step: AIC=-1820.4  
## lnRec ~ Pos + lnTgt + Ctch.  
##   
## Df Sum of Sq RSS AIC  
## <none> 0.0118 -1820.40  
## - Pos 2 0.0020 0.0138 -1794.81  
## - Ctch. 1 1.4854 1.4971 -906.42  
## - lnTgt 1 12.8610 12.8727 -499.77

backwardsmodel

##   
## Call:  
## lm(formula = lnRec ~ Pos + lnTgt + Ctch., data = NFLData)  
##   
## Coefficients:  
## (Intercept) PosTE PosWR lnTgt Ctch.   
## -1.40937 0.01175 0.01014 1.00098 0.01476

summary(backwardsmodel)

##   
## Call:  
## lm(formula = lnRec ~ Pos + lnTgt + Ctch., data = NFLData)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.033914 -0.002325 0.002629 0.004426 0.016349   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.409e+00 1.227e-02 -114.840 < 2e-16 \*\*\*  
## PosTE 1.175e-02 2.109e-03 5.574 8.76e-08 \*\*\*  
## PosWR 1.014e-02 2.328e-03 4.356 2.20e-05 \*\*\*  
## lnTgt 1.001e+00 2.232e-03 448.556 < 2e-16 \*\*\*  
## Ctch. 1.476e-02 9.682e-05 152.440 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.007995 on 184 degrees of freedom  
## Multiple R-squared: 0.9993, Adjusted R-squared: 0.9992   
## F-statistic: 6.203e+04 on 4 and 184 DF, p-value: < 2.2e-16

summary(backwardsmodel)$r.squared

## [1] 0.9992589

summary(backwardsmodel)$adj.r.squared

## [1] 0.9992428

summary(backwardsmodel)$sigma

## [1] 0.007995028

#BIC as selection criteria  
stepAIC(fullmodel2, scope = list(lower=nullmodel, upper=fullmodel2), k=log(length(NFLData$lnRec)), direction = "backward")

## Start: AIC=-1775.68  
## lnRec ~ Age + Pos + lnTgt + Yds + lnTD + X1D + Ctch. + PB + AP  
##   
## Df Sum of Sq RSS AIC  
## - Yds 1 0.00000 0.01158 -1780.87  
## - AP 1 0.00000 0.01159 -1780.84  
## - PB 1 0.00001 0.01159 -1780.77  
## - Age 1 0.00002 0.01160 -1780.58  
## - lnTD 1 0.00005 0.01163 -1780.10  
## - X1D 1 0.00006 0.01164 -1779.95  
## <none> 0.01158 -1775.68  
## - Pos 2 0.00156 0.01314 -1762.28  
## - Ctch. 1 1.32004 1.33162 -884.14  
## - lnTgt 1 2.72676 2.73835 -747.88  
##   
## Step: AIC=-1780.87  
## lnRec ~ Age + Pos + lnTgt + lnTD + X1D + Ctch. + PB + AP  
##   
## Df Sum of Sq RSS AIC  
## - AP 1 0.00000 0.01159 -1786.06  
## - PB 1 0.00001 0.01159 -1785.97  
## - Age 1 0.00002 0.01160 -1785.80  
## - lnTD 1 0.00005 0.01164 -1785.25  
## - X1D 1 0.00008 0.01166 -1784.85  
## <none> 0.01158 -1780.87  
## - Pos 2 0.00157 0.01315 -1767.39  
## - Ctch. 1 1.32033 1.33191 -889.34  
## - lnTgt 1 2.96278 2.97436 -737.50  
##   
## Step: AIC=-1786.06  
## lnRec ~ Age + Pos + lnTgt + lnTD + X1D + Ctch. + PB  
##   
## Df Sum of Sq RSS AIC  
## - PB 1 0.00002 0.01160 -1791.01  
## - Age 1 0.00002 0.01161 -1790.99  
## - lnTD 1 0.00005 0.01164 -1790.47  
## - X1D 1 0.00009 0.01168 -1789.76  
## <none> 0.01159 -1786.06  
## - Pos 2 0.00157 0.01316 -1772.52  
## - Ctch. 1 1.32139 1.33297 -894.44  
## - lnTgt 1 3.10432 3.11591 -733.95  
##   
## Step: AIC=-1791.01  
## lnRec ~ Age + Pos + lnTgt + lnTD + X1D + Ctch.  
##   
## Df Sum of Sq RSS AIC  
## - Age 1 0.00002 0.01162 -1795.95  
## - lnTD 1 0.00004 0.01165 -1795.52  
## - X1D 1 0.00013 0.01174 -1794.11  
## <none> 0.01160 -1791.01  
## - Pos 2 0.00155 0.01316 -1777.76  
## - Ctch. 1 1.32753 1.33914 -898.81  
## - lnTgt 1 3.10477 3.11638 -739.17  
##   
## Step: AIC=-1795.95  
## lnRec ~ Pos + lnTgt + lnTD + X1D + Ctch.  
##   
## Df Sum of Sq RSS AIC  
## - lnTD 1 0.00005 0.01167 -1800.38  
## - X1D 1 0.00013 0.01176 -1799.02  
## <none> 0.01162 -1795.95  
## - Pos 2 0.00154 0.01316 -1782.89  
## - Ctch. 1 1.32901 1.34064 -903.84  
## - lnTgt 1 3.10760 3.11922 -744.24  
##   
## Step: AIC=-1800.38  
## lnRec ~ Pos + lnTgt + X1D + Ctch.  
##   
## Df Sum of Sq RSS AIC  
## - X1D 1 0.0001 0.0118 -1804.20  
## <none> 0.0117 -1800.38  
## - Pos 2 0.0015 0.0132 -1787.65  
## - Ctch. 1 1.3504 1.3621 -906.08  
## - lnTgt 1 3.1861 3.1978 -744.78  
##   
## Step: AIC=-1804.2  
## lnRec ~ Pos + lnTgt + Ctch.  
##   
## Df Sum of Sq RSS AIC  
## <none> 0.0118 -1804.20  
## - Pos 2 0.0020 0.0138 -1785.08  
## - Ctch. 1 1.4854 1.4971 -893.45  
## - lnTgt 1 12.8610 12.8727 -486.81

##   
## Call:  
## lm(formula = lnRec ~ Pos + lnTgt + Ctch., data = NFLData)  
##   
## Coefficients:  
## (Intercept) PosTE PosWR lnTgt Ctch.   
## -1.40937 0.01175 0.01014 1.00098 0.01476

backwardsmodel2 <- stepAIC(fullmodel2, scope = list(lower=nullmodel, upper=fullmodel2), k=log(length(NFLData$lnRec)), direction = "backward")

## Start: AIC=-1775.68  
## lnRec ~ Age + Pos + lnTgt + Yds + lnTD + X1D + Ctch. + PB + AP  
##   
## Df Sum of Sq RSS AIC  
## - Yds 1 0.00000 0.01158 -1780.87  
## - AP 1 0.00000 0.01159 -1780.84  
## - PB 1 0.00001 0.01159 -1780.77  
## - Age 1 0.00002 0.01160 -1780.58  
## - lnTD 1 0.00005 0.01163 -1780.10  
## - X1D 1 0.00006 0.01164 -1779.95  
## <none> 0.01158 -1775.68  
## - Pos 2 0.00156 0.01314 -1762.28  
## - Ctch. 1 1.32004 1.33162 -884.14  
## - lnTgt 1 2.72676 2.73835 -747.88  
##   
## Step: AIC=-1780.87  
## lnRec ~ Age + Pos + lnTgt + lnTD + X1D + Ctch. + PB + AP  
##   
## Df Sum of Sq RSS AIC  
## - AP 1 0.00000 0.01159 -1786.06  
## - PB 1 0.00001 0.01159 -1785.97  
## - Age 1 0.00002 0.01160 -1785.80  
## - lnTD 1 0.00005 0.01164 -1785.25  
## - X1D 1 0.00008 0.01166 -1784.85  
## <none> 0.01158 -1780.87  
## - Pos 2 0.00157 0.01315 -1767.39  
## - Ctch. 1 1.32033 1.33191 -889.34  
## - lnTgt 1 2.96278 2.97436 -737.50  
##   
## Step: AIC=-1786.06  
## lnRec ~ Age + Pos + lnTgt + lnTD + X1D + Ctch. + PB  
##   
## Df Sum of Sq RSS AIC  
## - PB 1 0.00002 0.01160 -1791.01  
## - Age 1 0.00002 0.01161 -1790.99  
## - lnTD 1 0.00005 0.01164 -1790.47  
## - X1D 1 0.00009 0.01168 -1789.76  
## <none> 0.01159 -1786.06  
## - Pos 2 0.00157 0.01316 -1772.52  
## - Ctch. 1 1.32139 1.33297 -894.44  
## - lnTgt 1 3.10432 3.11591 -733.95  
##   
## Step: AIC=-1791.01  
## lnRec ~ Age + Pos + lnTgt + lnTD + X1D + Ctch.  
##   
## Df Sum of Sq RSS AIC  
## - Age 1 0.00002 0.01162 -1795.95  
## - lnTD 1 0.00004 0.01165 -1795.52  
## - X1D 1 0.00013 0.01174 -1794.11  
## <none> 0.01160 -1791.01  
## - Pos 2 0.00155 0.01316 -1777.76  
## - Ctch. 1 1.32753 1.33914 -898.81  
## - lnTgt 1 3.10477 3.11638 -739.17  
##   
## Step: AIC=-1795.95  
## lnRec ~ Pos + lnTgt + lnTD + X1D + Ctch.  
##   
## Df Sum of Sq RSS AIC  
## - lnTD 1 0.00005 0.01167 -1800.38  
## - X1D 1 0.00013 0.01176 -1799.02  
## <none> 0.01162 -1795.95  
## - Pos 2 0.00154 0.01316 -1782.89  
## - Ctch. 1 1.32901 1.34064 -903.84  
## - lnTgt 1 3.10760 3.11922 -744.24  
##   
## Step: AIC=-1800.38  
## lnRec ~ Pos + lnTgt + X1D + Ctch.  
##   
## Df Sum of Sq RSS AIC  
## - X1D 1 0.0001 0.0118 -1804.20  
## <none> 0.0117 -1800.38  
## - Pos 2 0.0015 0.0132 -1787.65  
## - Ctch. 1 1.3504 1.3621 -906.08  
## - lnTgt 1 3.1861 3.1978 -744.78  
##   
## Step: AIC=-1804.2  
## lnRec ~ Pos + lnTgt + Ctch.  
##   
## Df Sum of Sq RSS AIC  
## <none> 0.0118 -1804.20  
## - Pos 2 0.0020 0.0138 -1785.08  
## - Ctch. 1 1.4854 1.4971 -893.45  
## - lnTgt 1 12.8610 12.8727 -486.81

backwardsmodel2

##   
## Call:  
## lm(formula = lnRec ~ Pos + lnTgt + Ctch., data = NFLData)  
##   
## Coefficients:  
## (Intercept) PosTE PosWR lnTgt Ctch.   
## -1.40937 0.01175 0.01014 1.00098 0.01476

summary(backwardsmodel2)

##   
## Call:  
## lm(formula = lnRec ~ Pos + lnTgt + Ctch., data = NFLData)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.033914 -0.002325 0.002629 0.004426 0.016349   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.409e+00 1.227e-02 -114.840 < 2e-16 \*\*\*  
## PosTE 1.175e-02 2.109e-03 5.574 8.76e-08 \*\*\*  
## PosWR 1.014e-02 2.328e-03 4.356 2.20e-05 \*\*\*  
## lnTgt 1.001e+00 2.232e-03 448.556 < 2e-16 \*\*\*  
## Ctch. 1.476e-02 9.682e-05 152.440 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.007995 on 184 degrees of freedom  
## Multiple R-squared: 0.9993, Adjusted R-squared: 0.9992   
## F-statistic: 6.203e+04 on 4 and 184 DF, p-value: < 2.2e-16

summary(backwardsmodel2)$r.squared

## [1] 0.9992589

summary(backwardsmodel2)$adj.r.squared

## [1] 0.9992428

summary(backwardsmodel2)$sigma

## [1] 0.007995028

#Model Selection  
  
#Full Model  
AIC(fullmodel2)

## [1] -1272.976

BIC(fullmodel2)

## [1] -1234.075

summary(fullmodel2)$adj.r.squared

## [1] 0.9992293

#Stepwise Model  
AIC(forwardmodel)

## [1] -1282.046

BIC(forwardmodel)

## [1] -1262.596

summary(forwardmodel)$adj.r.squared

## [1] 0.9992428

#Handpicked Model with variables with high correlation  
model3 <- lm(lnRec ~ Yds + X1D + lnTD + lnTgt, data=NFLData)  
summary(model3)

##   
## Call:  
## lm(formula = lnRec ~ Yds + X1D + lnTD + lnTgt, data = NFLData)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.289380 -0.078617 0.006627 0.075517 0.266077   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.869e-01 2.381e-01 1.625 0.106   
## Yds -7.549e-05 8.862e-05 -0.852 0.395   
## X1D 3.384e-03 1.979e-03 1.710 0.089 .   
## lnTD -2.530e-02 1.751e-02 -1.445 0.150   
## lnTgt 8.243e-01 6.073e-02 13.575 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.109 on 184 degrees of freedom  
## Multiple R-squared: 0.8622, Adjusted R-squared: 0.8593   
## F-statistic: 287.9 on 4 and 184 DF, p-value: < 2.2e-16

AIC(model3)

## [1] -294.5035

BIC(model3)

## [1] -275.053

summary(model3)$adj.r.squared

## [1] 0.8592534

#Handpicked Model with added interaction term  
model4 <- lm(lnRec ~ Yds + X1D + lnTD + lnTgt + Yds\*lnTD, data=NFLData)  
summary(model4)

##   
## Call:  
## lm(formula = lnRec ~ Yds + X1D + lnTD + lnTgt + Yds \* lnTD, data = NFLData)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.295254 -0.072324 0.008254 0.072716 0.266838   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.239e-01 2.368e-01 1.368 0.17300   
## Yds -2.895e-04 1.262e-04 -2.294 0.02292 \*   
## X1D 3.273e-03 1.956e-03 1.674 0.09591 .   
## lnTD -8.402e-02 3.035e-02 -2.768 0.00621 \*\*   
## lnTgt 8.673e-01 6.270e-02 13.832 < 2e-16 \*\*\*  
## Yds:lnTD 9.877e-05 4.195e-05 2.355 0.01960 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.1077 on 183 degrees of freedom  
## Multiple R-squared: 0.8663, Adjusted R-squared: 0.8626   
## F-statistic: 237.1 on 5 and 183 DF, p-value: < 2.2e-16

AIC(model4)

## [1] -298.1445

BIC(model4)

## [1] -275.4523

summary(model4)$adj.r.squared

## [1] 0.8626456