STA303-Project-Part-2

Amrinder Sehmbi

2024-04-07

Imported Libraries

```
#oo <- options(repos = "https://cran.r-project.org/")
#install.packages("Matrix")
#install.packages("lme4")
#options(oo)


library(ggplot2)
library(knitr)
library(vcd)
```

```
## Loading required package: grid
```

```
library(lme4)
```

```
## Loading required package: Matrix
```

```
library(influence.ME)
```

```
##
## Attaching package: 'influence.ME'
```

```
## The following object is masked from 'package:stats':
##
##     influence
```

```
library(pROC)
```

```
## Type 'citation("pROC")' for a citation.
```

```
##
## Attaching package: 'pROC'
```

```
## The following objects are masked from 'package:stats':
##
##     cov, smooth, var
```

```
library(lattice)
library(mlmhelpr)
```

Load the data set in R

```
# Open CSV file with data
FullData <- read.csv("~/STA303/shotdetail_2022.csv")
```

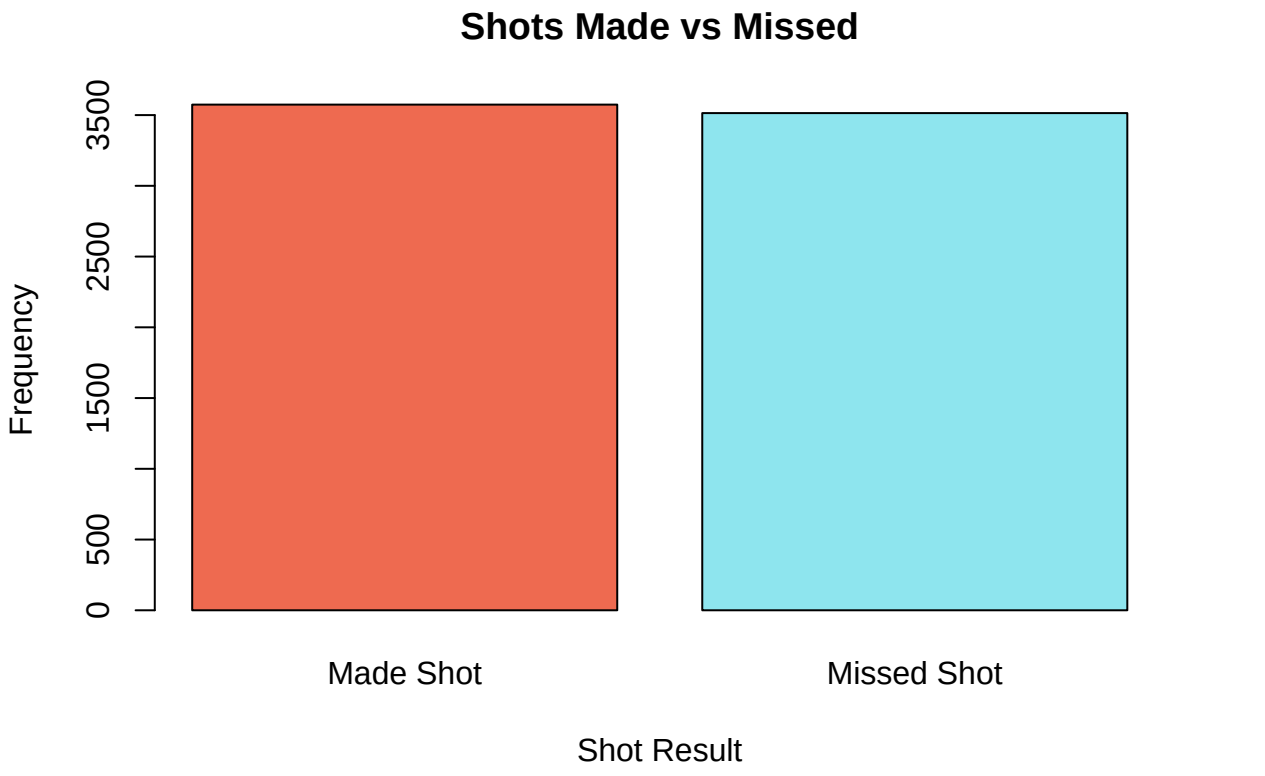
```r
# Remove ID columns
RemovedColumns <- FullData[, -c(which(names(FullData) %in%
                                        c('GRID_TYPE', 'GAME_ID',
                                          'GAME_EVENT_ID', 'PLAYER_ID',
                                          'TEAM_ID')))]
# Filter for Denver Nuggets Data
DenverNuggetsData <-subset(RemovedColumns, TEAM_NAME == "Denver Nuggets")

# Add column indicating if shot taken at home or away
DenverNuggetsData$HOME_AWAY <- ifelse(DenverNuggetsData$HTM == "DEN", "Home",
                                      "Away")
```

Data Summary

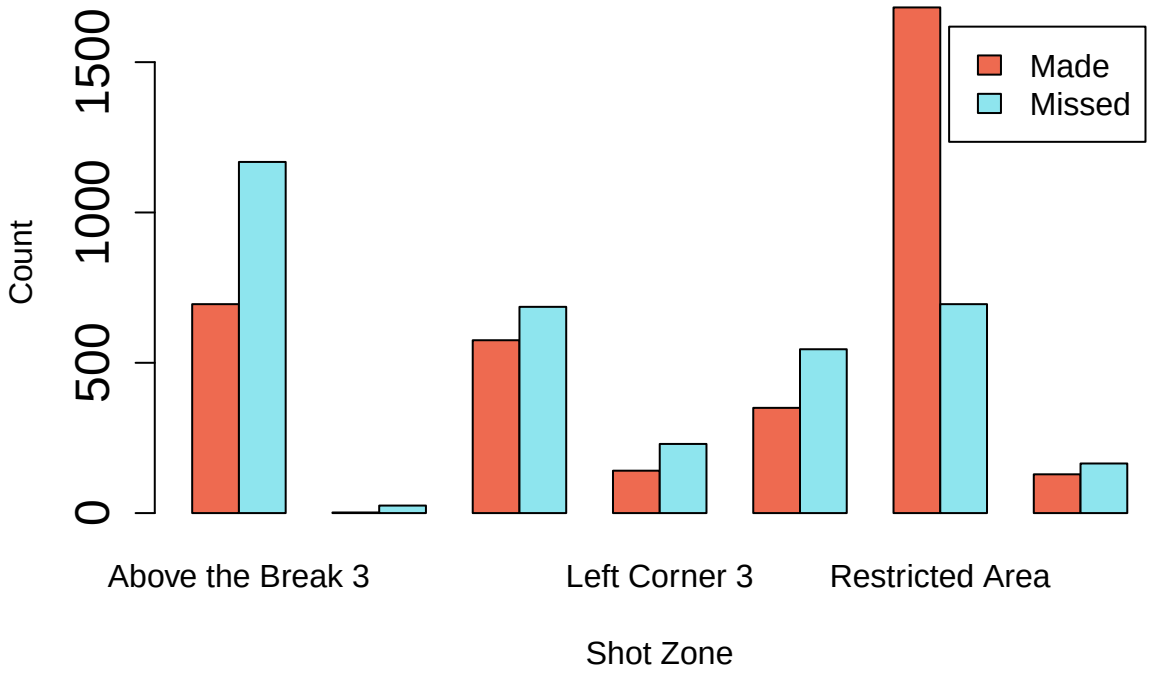
```r
# Summary of Data
summary(DenverNuggetsData)
```

```
##  PLAYER_NAME         TEAM_NAME            PERIOD      MINUTES_REMAINING
##  Length:7088        Length:7088        Min.   :1.000   Min.   : 0.000
##  Class :character   Class :character   1st Qu.:1.000   1st Qu.: 2.000
##  Mode  :character   Mode  :character   Median :2.000   Median : 5.000
##                                        Mean   :2.463   Mean   : 5.413
##                                        3rd Qu.:3.000   3rd Qu.: 8.000
##                                        Max.   :5.000   Max.   :11.000
##  SECONDS_REMAINING  EVENT_TYPE        ACTION_TYPE         SHOT_TYPE
##  Min.   : 0.00     Length:7088        Length:7088        Length:7088
##  1st Qu.:13.00     Class :character   Class :character   Class :character
##  Median :29.00     Mode  :character   Mode  :character   Mode  :character
##  Mean   :28.63
##  3rd Qu.:44.00
##  Max.   :59.00
##  SHOT_ZONE_BASIC    SHOT_ZONE_AREA     SHOT_ZONE_RANGE    SHOT_DISTANCE
##  Length:7088        Length:7088        Length:7088        Min.   : 0.00
##  Class :character   Class :character   Class :character   1st Qu.: 2.00
##  Mode  :character   Mode  :character   Mode  :character   Median :11.00
##                                                           Mean   :12.93
##                                                           3rd Qu.:24.00
##                                                           Max.   :82.00
##      LOC_X              LOC_Y        SHOT_ATTEMPTED_FLAG SHOT_MADE_FLAG
##  Min.   :-245.000   Min.   :-43.00   Min.   :1           Min.   :0.0000
##  1st Qu.: -50.000   1st Qu.: 14.00   1st Qu.:1           1st Qu.:0.0000
##  Median :   0.000   Median : 43.00   Median :1           Median :1.0000
##  Mean   :  -6.558   Mean   : 90.96   Mean   :1           Mean   :0.5042
##  3rd Qu.:  33.000   3rd Qu.:173.25   3rd Qu.:1           3rd Qu.:1.0000
##  Max.   : 245.000   Max.   :820.00   Max.   :1           Max.   :1.0000
##    GAME_DATE            HTM                VTM             HOME_AWAY
##  Min.   :20221019   Length:7088        Length:7088        Length:7088
##  1st Qu.:20221130   Class :character   Class :character   Class :character
##  Median :20230113   Mode  :character   Mode  :character   Mode  :character
##  Mean   :20226368
##  3rd Qu.:20230226
##  Max.   :20230409
```

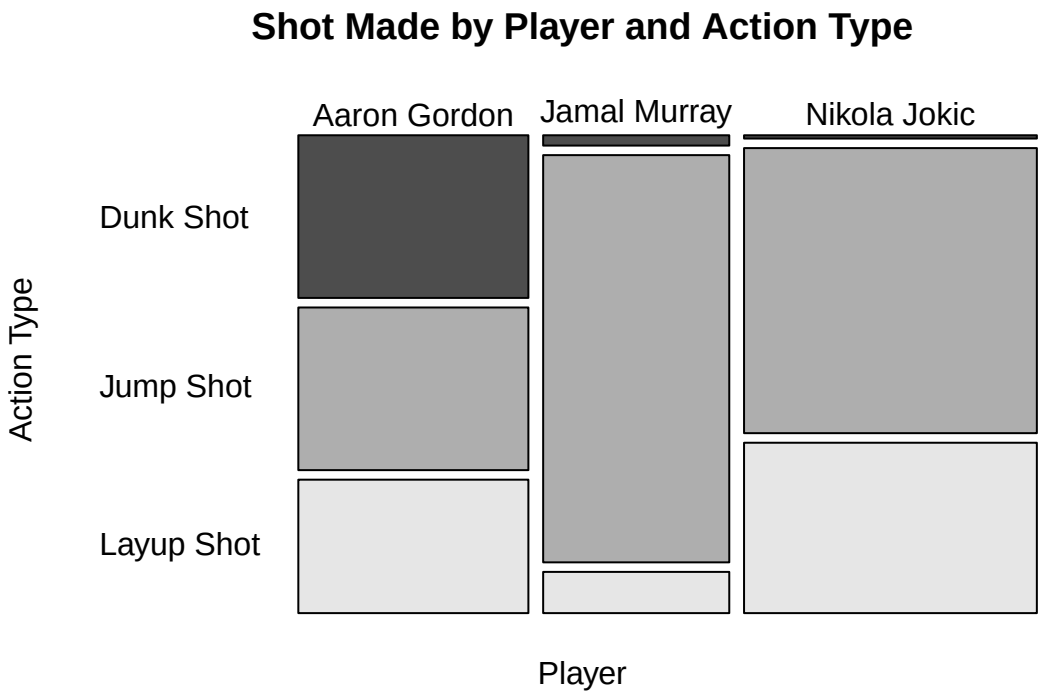
```r
# Visual 1: Distribution of EVENT_TYPE
barplot(table(DenverNuggetsData$EVENT_TYPE),
        main = "Shots Made vs Missed",
        xlab = "Shot Result",
        ylab = "Frequency",
        col = c("coral2", "cadetblue2"))
```




```r
# Visual 2: Shot Zone vs Shot Outcome
barplot(table(DenverNuggetsData$EVENT_TYPE, DenverNuggetsData$SHOT_ZONE_BASIC),
        beside = TRUE,
        legend.text = c("Made", "Missed"),
        main = "Shot Outcome vs Shot Zone",
        xlab = "Shot Zone",
        ylab = "Count",
        col = c("coral2", "cadetblue2"),
        cex.axis = 1.5)
```




```r
# Visual 3: Shots Made by Shot Type and Player Shooting
ShotsMadeData = subset(DenverNuggetsData, EVENT_TYPE == "Made Shot")
desired_categories <- c("Aaron Gordon", "Jamal Murray", "Nikola Jokic")
ShotsMadeData = subset(ShotsMadeData, PLAYER_NAME %in% desired_categories)
desired_categories <- c("Dunk Shot", "Jump Shot", "Layup Shot")
ShotsMadeData = subset(ShotsMadeData, ACTION_TYPE %in% desired_categories)

mosaicplot(table(ShotsMadeData$PLAYER_NAME, ShotsMadeData$ACTION_TYPE),
           main = "Shot Made by Player and Action Type",
           xlab = "Player",
           ylab = "Action Type",
           color = TRUE,
           las = 1,
           cex.axis = 1)
```



Model Fitting

Model 1: Initial Model with all Variables

```
# Model 1: Initial Model with all Variables
model1 <- glmer(SHOT_MADE_FLAG ~ SHOT_ZONE_BASIC + SHOT_DISTANCE
                + MINUTES_REMAINING + factor(HOME_AWAY) + factor(PERIOD)
                + (1|ACTION_TYPE)+(1|ACTION_TYPE:PLAYER_NAME),
                family = "binomial", data = DenverNuggetsData,
                control=glmerControl(optimizer="bobyqa",optCtrl=list(maxfun=2e5)))
summary(model1)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula:
## SHOT_MADE_FLAG ~ SHOT_ZONE_BASIC + SHOT_DISTANCE + MINUTES_REMAINING +
##     factor(HOME_AWAY) + factor(PERIOD) + (1 | ACTION_TYPE) +
##     (1 | ACTION_TYPE:PLAYER_NAME)
##    Data: DenverNuggetsData
## Control: glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 2e+05))
##
##      AIC      BIC   logLik deviance df.resid
##   8999.6   9109.4  -4483.8   8967.6     7072
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.3468 -0.8182  0.2980  0.9358  4.2737
##
## Random effects:
##  Groups                 Name        Variance Std.Dev.
```

```
##  ACTION_TYPE:PLAYER_NAME (Intercept) 0.0777   0.2787
##  ACTION_TYPE             (Intercept) 0.3265   0.5714
## Number of obs: 7088, groups:  ACTION_TYPE:PLAYER_NAME, 500; ACTION_TYPE, 48
##
## Fixed effects:
##                                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)                           0.423348   0.356813   1.186  0.23544
## SHOT_ZONE_BASICBackcourt             -1.024690   0.833909  -1.229  0.21915
## SHOT_ZONE_BASICIn The Paint (Non-RA) -0.039717   0.257722  -0.154  0.87752
## SHOT_ZONE_BASICLeft Corner 3         -0.050119   0.128590  -0.390  0.69672
## SHOT_ZONE_BASICMid-Range             -0.202317   0.162756  -1.243  0.21384
## SHOT_ZONE_BASICRestricted Area        0.857938   0.319940   2.682  0.00733 **
## SHOT_ZONE_BASICRight Corner 3         0.224001   0.137979   1.623  0.10449
## SHOT_DISTANCE                        -0.037102   0.014156  -2.621  0.00877 **
## MINUTES_REMAINING                     0.001484   0.007533   0.197  0.84380
## factor(HOME_AWAY)Home                 0.047821   0.051795   0.923  0.35586
## factor(PERIOD)2                      -0.109496   0.071626  -1.529  0.12634
## factor(PERIOD)3                      -0.053919   0.071598  -0.753  0.45141
## factor(PERIOD)4                      -0.179013   0.073846  -2.424  0.01534 *
## factor(PERIOD)5                       0.150878   0.436887   0.345  0.72983
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Correlation matrix not shown by default, as p = 14 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it
```

```
mlmhelpr::icc(model1)
```

```
##                                  grps   icc
## 1 ACTION_TYPE:PLAYER_NAME (Intercept) 0.023
## 2             ACTION_TYPE (Intercept) 0.090
```

```
auc(roc(DenverNuggetsData$SHOT_MADE_FLAG, predict(model1, type = "response")))
```

```
## Setting levels: control = 0, case = 1
```

```
## Setting direction: controls < cases
```

```
## Area under the curve: 0.7102
```

Model 2: Excluding MINUTES_REMAINING

```
# Model 2: Excluding MINUTES_REMAINING
model2 <- glmer(SHOT_MADE_FLAG ~ SHOT_ZONE_BASIC + SHOT_DISTANCE
                + factor(HOME_AWAY) + factor(PERIOD)
                + (1|ACTION_TYPE)+(1|ACTION_TYPE:PLAYER_NAME),
                family = "binomial", data = DenverNuggetsData,
                control=glmerControl(optimizer="bobyqa",optCtrl=list(maxfun=2e5)))



summary(model2)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
```

```
##  Family: binomial  ( logit )
## Formula:
## SHOT_MADE_FLAG ~ SHOT_ZONE_BASIC + SHOT_DISTANCE + factor(HOME_AWAY) +
##     factor(PERIOD) + (1 | ACTION_TYPE) + (1 | ACTION_TYPE:PLAYER_NAME)
##    Data: DenverNuggetsData
## Control: glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 2e+05))
##
##      AIC      BIC   logLik deviance df.resid
##   8997.6   9100.6  -4483.8   8967.6     7073
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.3555 -0.8182  0.2981  0.9349  4.2753
##
## Random effects:
##  Groups                  Name        Variance Std.Dev.
##  ACTION_TYPE:PLAYER_NAME (Intercept) 0.07771  0.2788
##  ACTION_TYPE             (Intercept) 0.32663  0.5715
## Number of obs: 7088, groups:  ACTION_TYPE:PLAYER_NAME, 500; ACTION_TYPE, 48
##
## Fixed effects:
##                                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)                          0.43280    0.35342   1.225  0.22072
## SHOT_ZONE_BASICBackcourt            -1.03038    0.83299  -1.237  0.21610
## SHOT_ZONE_BASICIn The Paint (Non-RA) -0.04017    0.25759  -0.156  0.87607
## SHOT_ZONE_BASICLeft Corner 3        -0.05008    0.12858  -0.390  0.69689
## SHOT_ZONE_BASICMid-Range            -0.20247    0.16269  -1.244  0.21332
## SHOT_ZONE_BASICRestricted Area       0.85666    0.31971   2.679  0.00737 **
## SHOT_ZONE_BASICRight Corner 3        0.22427    0.13796   1.626  0.10404
## SHOT_DISTANCE                       -0.03717    0.01414  -2.628  0.00859 **
## factor(HOME_AWAY)Home                0.04795    0.05179   0.926  0.35457
## factor(PERIOD)2                     -0.10970    0.07162  -1.532  0.12560
## factor(PERIOD)3                     -0.05400    0.07159  -0.754  0.45071
## factor(PERIOD)4                     -0.17893    0.07384  -2.423  0.01539 *
## factor(PERIOD)5                      0.14578    0.43603   0.334  0.73814
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it
```

```
mlmhelpr::icc(model2)
```

```
##                                  grps   icc
## 1 ACTION_TYPE:PLAYER_NAME (Intercept) 0.023
## 2             ACTION_TYPE (Intercept) 0.090
```

```
auc(roc(DenverNuggetsData$SHOT_MADE_FLAG, predict(model2, type = "response")))
```

```
## Setting levels: control = 0, case = 1
## Setting direction: controls < cases
## Area under the curve: 0.7102
```

Model 3: Excluding MINUTES_REMAINING, HOME_AWAY

```
# Model 3: Excluding MINUTES_REMAINING, HOME_AWAY
model3 <- glmer(SHOT_MADE_FLAG ~ SHOT_ZONE_BASIC + SHOT_DISTANCE + factor(PERIOD)
                + (1|ACTION_TYPE)+(1|ACTION_TYPE:PLAYER_NAME),
                family = "binomial", data = DenverNuggetsData,
                control=glmerControl(optimizer="bobyqa",optCtrl=list(maxfun=2e5)))



summary(model3)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: SHOT_MADE_FLAG ~ SHOT_ZONE_BASIC + SHOT_DISTANCE + factor(PERIOD) +
##     (1 | ACTION_TYPE) + (1 | ACTION_TYPE:PLAYER_NAME)
##    Data: DenverNuggetsData
## Control: glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 2e+05))
##
##      AIC      BIC   logLik deviance df.resid
##   8996.4   9092.6  -4484.2   8968.4     7074
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.4004 -0.8183  0.2995  0.9297  4.2019
##
## Random effects:
##  Groups                   Name        Variance Std.Dev.
##  ACTION_TYPE:PLAYER_NAME (Intercept) 0.07759  0.2785
##  ACTION_TYPE             (Intercept) 0.32858  0.5732
## Number of obs: 7088, groups:  ACTION_TYPE:PLAYER_NAME, 500; ACTION_TYPE, 48
##
## Fixed effects:
##                                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)                           0.44857    0.35319   1.270  0.20407
## SHOT_ZONE_BASICBackcourt             -1.03986    0.83420  -1.247  0.21257
## SHOT_ZONE_BASICIn The Paint (Non-RA) -0.03366    0.25757  -0.131  0.89603
## SHOT_ZONE_BASICLeft Corner 3         -0.04748    0.12855  -0.369  0.71188
## SHOT_ZONE_BASICMid-Range             -0.20065    0.16275  -1.233  0.21762
## SHOT_ZONE_BASICRestricted Area        0.86452    0.31970   2.704  0.00685 **
## SHOT_ZONE_BASICRight Corner 3         0.22596    0.13795   1.638  0.10144
## SHOT_DISTANCE                        -0.03658    0.01414  -2.588  0.00966 **
## factor(PERIOD)2                      -0.11025    0.07162  -1.539  0.12370
## factor(PERIOD)3                      -0.05423    0.07158  -0.758  0.44873
## factor(PERIOD)4                      -0.17976    0.07383  -2.435  0.01490 *
## factor(PERIOD)5                       0.15534    0.43582   0.356  0.72152
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##                     (Intr) SHOT_ZONE_BASICB SHOTP( SHOT_ZONE_BASICLC3
## SHOT_ZONE_BASICB     0.380
## SHOT_ZONTP(         -0.912 -0.363
```

```
## SHOT_ZONE_BASICLC3 -0.317 -0.111            0.322
## SHOT_ZONE_BASICM   -0.799 -0.326            0.827  0.327
## SHOT_ZONE_A        -0.934 -0.373            0.940  0.317
## SHOT_ZONE_BASICRC3 -0.289 -0.096            0.292  0.265
## SHOT_DISTAN        -0.892 -0.443            0.863  0.312
## fc(PERIOD)2        -0.090  0.002           -0.008 -0.015
## fc(PERIOD)3        -0.079  0.005           -0.012 -0.008
## fc(PERIOD)4        -0.086  0.016           -0.008 -0.023
## fc(PERIOD)5        -0.020  0.001            0.004  0.012
##                    SHOT_ZONE_BASICM SHOT_A SHOT_ZONE_BASICRC3 SHOT_D f(PERIOD)2
## SHOT_ZONE_BASICB
## SHOT_ZONTP(
## SHOT_ZONE_BASICLC3
## SHOT_ZONE_BASICM
## SHOT_ZONE_A         0.805
## SHOT_ZONE_BASICRC3  0.297            0.287
## SHOT_DISTAN         0.779            0.877  0.275
## fc(PERIOD)2        -0.007           -0.012  0.018             -0.008
## fc(PERIOD)3        -0.005           -0.019 -0.014             -0.020  0.488
## fc(PERIOD)4        -0.003           -0.009  0.009             -0.016  0.481
## fc(PERIOD)5         0.002            0.008  0.003              0.001  0.080
##                    f(PERIOD)3 f(PERIOD)4
## SHOT_ZONE_BASICB
## SHOT_ZONTP(
## SHOT_ZONE_BASICLC3
## SHOT_ZONE_BASICM
## SHOT_ZONE_A
## SHOT_ZONE_BASICRC3
## SHOT_DISTAN
## fc(PERIOD)2
## fc(PERIOD)3
## fc(PERIOD)4         0.476
## fc(PERIOD)5         0.082      0.077
```

```
mlmhelpr::icc(model3)
```

```
##                                 grps   icc
## 1 ACTION_TYPE:PLAYER_NAME (Intercept) 0.023
## 2             ACTION_TYPE (Intercept) 0.091
```

```
auc(roc(DenverNuggetsData$SHOT_MADE_FLAG, predict(model3, type = "response")))
```

```
## Setting levels: control = 0, case = 1
```

```
## Setting direction: controls < cases
```

```
## Area under the curve: 0.7101
```

Model 4: Excluding MINUTES_REMAINING, HOME_AWAY, SHOT_DISTANCE

```
# Model 4: Excluding MINUTES_REMAINING, HOME_AWAY, SHOT_DISTANCE
model4 <- glmer(SHOT_MADE_FLAG ~ SHOT_ZONE_BASIC  + factor(PERIOD)
                + (1|ACTION_TYPE)+(1|ACTION_TYPE:PLAYER_NAME),
                family = "binomial", data = DenverNuggetsData,
                control=glmerControl(optimizer="bobyqa",optCtrl=list(maxfun=2e5)))
```

```r
summary(model4)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula:
## SHOT_MADE_FLAG ~ SHOT_ZONE_BASIC + factor(PERIOD) + (1 | ACTION_TYPE) +
##     (1 | ACTION_TYPE:PLAYER_NAME)
##    Data: DenverNuggetsData
## Control: glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 2e+05))
##
##      AIC      BIC   logLik deviance df.resid
##   9001.2   9090.5  -4487.6   8975.2     7075
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.4015 -0.8200  0.3012  0.9330  3.7159
##
## Random effects:
##  Groups                  Name        Variance Std.Dev.
##  ACTION_TYPE:PLAYER_NAME (Intercept) 0.07716  0.2778
##  ACTION_TYPE             (Intercept) 0.33602  0.5797
## Number of obs: 7088, groups:  ACTION_TYPE:PLAYER_NAME, 500; ACTION_TYPE, 48
##
## Fixed effects:
##                                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)                         -0.36604    0.16140  -2.268  0.02334 *
## SHOT_ZONE_BASICBackcourt            -2.03572    0.74437  -2.735  0.00624 **
## SHOT_ZONE_BASICIn The Paint (Non-RA)  0.54170    0.13009   4.164 3.13e-05 ***
## SHOT_ZONE_BASICLeft Corner 3          0.05663    0.12212   0.464  0.64287
## SHOT_ZONE_BASICMid-Range              0.12671    0.10191   1.243  0.21376
## SHOT_ZONE_BASICRestricted Area        1.59055    0.15407  10.324  < 2e-16 ***
## SHOT_ZONE_BASICRight Corner 3         0.32477    0.13260   2.449  0.01431 *
## factor(PERIOD)2                     -0.11175    0.07159  -1.561  0.11851
## factor(PERIOD)3                     -0.05804    0.07155  -0.811  0.41722
## factor(PERIOD)4                     -0.18300    0.07379  -2.480  0.01314 *
## factor(PERIOD)5                      0.15683    0.43585   0.360  0.71898
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##                   (Intr) SHOT_ZONE_BASICB SHOTP( SHOT_ZONE_BASICLC3
## SHOT_ZONE_BASICB  -0.033
## SHOT_ZONTP(       -0.623  0.038
## SHOT_ZONE_BASICLC3 -0.088  0.031            0.111
## SHOT_ZONE_BASICM  -0.365  0.031            0.487  0.142
## SHOT_ZONE_A       -0.699  0.034            0.752  0.095
## SHOT_ZONE_BASICRC3 -0.099  0.029            0.113  0.196
## fc(PERIOD)2       -0.211 -0.003           -0.003 -0.013
## fc(PERIOD)3       -0.211 -0.004            0.011 -0.002
## fc(PERIOD)4       -0.221  0.009            0.013 -0.019
## fc(PERIOD)5       -0.041  0.002            0.006  0.013
```

```
##                     SHOT_ZONE_BASICM SHOT_A SHOT_ZONE_BASICRC3 f(PERIOD)2
## SHOT_ZONE_BASICB
## SHOT_ZONTP(
## SHOT_ZONE_BASICLC3
## SHOT_ZONE_BASICM
## SHOT_ZONE_A          0.401
## SHOT_ZONE_BASICRC3   0.137            0.098
## fc(PERIOD)2         -0.001           -0.012  0.021
## fc(PERIOD)3          0.017           -0.003 -0.009              0.488
## fc(PERIOD)4          0.015            0.012  0.014              0.481
## fc(PERIOD)5          0.002            0.015  0.002              0.080
##                     f(PERIOD)3 f(PERIOD)4
## SHOT_ZONE_BASICB
## SHOT_ZONTP(
## SHOT_ZONE_BASICLC3
## SHOT_ZONE_BASICM
## SHOT_ZONE_A
## SHOT_ZONE_BASICRC3
## fc(PERIOD)2
## fc(PERIOD)3
## fc(PERIOD)4          0.476
## fc(PERIOD)5          0.082      0.077
```

```
mlmhelpr::icc(model4)
```

```
##                                  grps   icc
## 1 ACTION_TYPE:PLAYER_NAME (Intercept) 0.023
## 2             ACTION_TYPE (Intercept) 0.093
```

```
auc(roc(DenverNuggetsData$SHOT_MADE_FLAG, predict(model4, type = "response")))
```

```
## Setting levels: control = 0, case = 1
```

```
## Setting direction: controls < cases
```

```
## Area under the curve: 0.71
```

Model 5: Final Model excluding MINUTES_REMAINING, HOME_AWAY, SHOT_DISTANCE, PERIOD

```
# Model 5: Final Model
final.model <- glmer(SHOT_MADE_FLAG ~ SHOT_ZONE_BASIC
                + (1|ACTION_TYPE)+(1|ACTION_TYPE:PLAYER_NAME),
                family = "binomial", data = DenverNuggetsData,
                control=glmerControl(optimizer="bobyqa",optCtrl=list(maxfun=2e5)))



summary(final.model)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula:
## SHOT_MADE_FLAG ~ SHOT_ZONE_BASIC + (1 | ACTION_TYPE) + (1 | ACTION_TYPE:PLAYER_NAME)
##    Data: DenverNuggetsData
```

```
## Control: glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 2e+05))
##
##      AIC      BIC   logLik deviance df.resid
##   9000.2   9062.0  -4491.1   8982.2     7079
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.2595 -0.8217  0.3068  0.9267  3.7552
##
## Random effects:
##  Groups                   Name        Variance Std.Dev.
##  ACTION_TYPE:PLAYER_NAME (Intercept) 0.08007  0.2830
##  ACTION_TYPE             (Intercept) 0.33467  0.5785
## Number of obs: 7088, groups:  ACTION_TYPE:PLAYER_NAME, 500; ACTION_TYPE, 48
##
## Fixed effects:
##                                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)                         -0.45375    0.15555  -2.917  0.00353 **
## SHOT_ZONE_BASICBackcourt            -2.01905    0.74449  -2.712  0.00669 **
## SHOT_ZONE_BASICIn The Paint (Non-RA)  0.54470    0.13007   4.188 2.82e-05 ***
## SHOT_ZONE_BASICLeft Corner 3          0.04923    0.12206   0.403  0.68672
## SHOT_ZONE_BASICMid-Range              0.12924    0.10189   1.268  0.20464
## SHOT_ZONE_BASICRestricted Area        1.59447    0.15396  10.356  < 2e-16 ***
## SHOT_ZONE_BASICRight Corner 3         0.33255    0.13248   2.510  0.01207 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##                     (Intr) SHOT_ZONE_BASICB SHOTP( SHOT_ZONE_BASICLC3
## SHOT_ZONE_BASICB    -0.034
## SHOT_ZONTP(         -0.644  0.038
## SHOT_ZONE_BASICLC3 -0.096  0.031            0.111
## SHOT_ZONE_BASICM    -0.375  0.031            0.487  0.142
## SHOT_ZONE_A         -0.725  0.034            0.752  0.095
## SHOT_ZONE_BASICRC3 -0.100  0.029            0.113  0.197
##                     SHOT_ZONE_BASICM SHOT_A
## SHOT_ZONE_BASICB
## SHOT_ZONTP(
## SHOT_ZONE_BASICLC3
## SHOT_ZONE_BASICM
## SHOT_ZONE_A          0.401
## SHOT_ZONE_BASICRC3  0.138            0.098
```

```
mlmhelpr::icc(final.model)
```

```
##                                  grps   icc
## 1 ACTION_TYPE:PLAYER_NAME (Intercept) 0.024
## 2             ACTION_TYPE (Intercept) 0.092
```

```
auc(roc(DenverNuggetsData$SHOT_MADE_FLAG, predict(final.model, type = "response")))
```

```
## Setting levels: control = 0, case = 1
```

```
## Setting direction: controls < cases
```

```
## Area under the curve: 0.7093
```

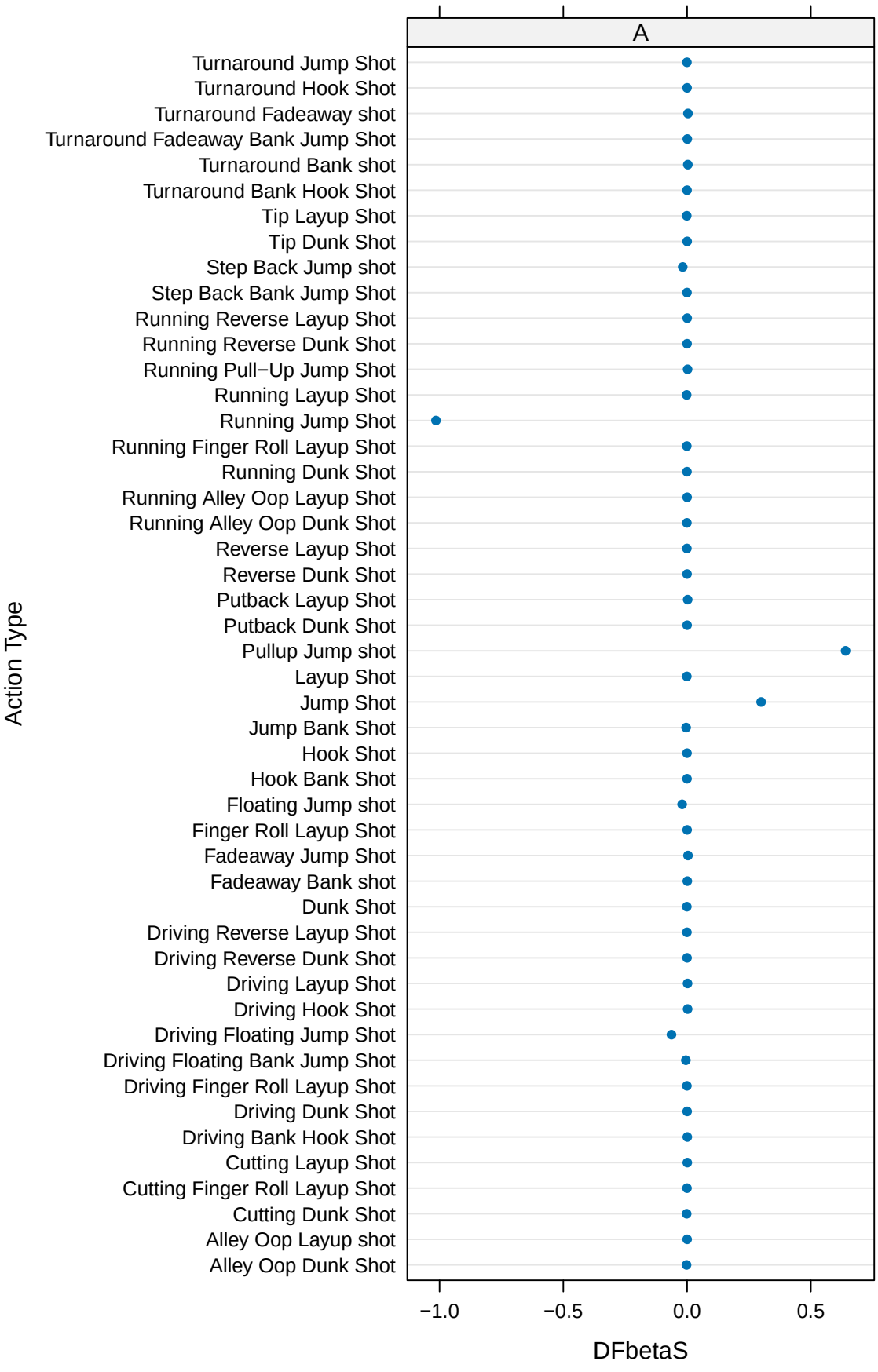
Model Diagnostics

Dfbetas

```
infl <- influence(final.model,"ACTION_TYPE")

plot(infl,which="dfbetas",parameters=c(1),xlab="DFbetaS",ylab="Action Type")
```


A
Turnaround Jump Shot
Turnaround Hook Shot
Turnaround Fadeaway shot
Turnaround Fadeaway Bank Jump Shot
Turnaround Bank shot
Turnaround Bank Hook Shot
Tip Layup Shot
Tip Dunk Shot
Step Back Jump shot
Step Back Bank Jump Shot
Running Reverse Layup Shot
Running Reverse Dunk Shot
Running Pull−Up Jump Shot
Running Layup Shot
Running Jump Shot
Running Finger Roll Layup Shot
Running Dunk Shot
Running Alley Oop Layup Shot
Running Alley Oop Dunk Shot
Reverse Layup Shot
Reverse Dunk Shot
Putback Layup Shot
Putback Dunk Shot
Pullup Jump shot
Layup Shot
Jump Shot
Jump Bank Shot
Hook Shot
Hook Bank Shot
Floating Jump shot
Finger Roll Layup Shot
Fadeaway Jump Shot
Fadeaway Bank shot
Dunk Shot
Driving Reverse Layup Shot
Driving Reverse Dunk Shot
Driving Layup Shot
Driving Hook Shot
Driving Floating Jump Shot
Driving Floating Bank Jump Shot
Driving Finger Roll Layup Shot
Driving Dunk Shot
Driving Bank Hook Shot
Cutting Layup Shot
Cutting Finger Roll Layup Shot
Cutting Dunk Shot
Alley Oop Layup shot
Alley Oop Dunk Shot
Action Type
−0.3
−0.2
−0.1
0.0
0.1
0.2
DFbetaS

```r
plot(infl,which="dfbetas",parameters=c(2),xlab="DFbetaS",ylab="Action Type")
```


A
Turnaround Jump Shot
Turnaround Hook Shot
Turnaround Fadeaway shot
Turnaround Fadeaway Bank Jump Shot
Turnaround Bank shot
Turnaround Bank Hook Shot
Tip Layup Shot
Tip Dunk Shot
Step Back Jump shot
Step Back Bank Jump Shot
Running Reverse Layup Shot
Running Reverse Dunk Shot
Running Pull−Up Jump Shot
Running Layup Shot
Running Jump Shot
Running Finger Roll Layup Shot
Running Dunk Shot
Running Alley Oop Layup Shot
Running Alley Oop Dunk Shot
Reverse Layup Shot
Reverse Dunk Shot
Putback Layup Shot
Putback Dunk Shot
Pullup Jump shot
Layup Shot
Jump Shot
Jump Bank Shot
Hook Shot
Hook Bank Shot
Floating Jump shot
Finger Roll Layup Shot
Fadeaway Jump Shot
Fadeaway Bank shot
Dunk Shot
Driving Reverse Layup Shot
Driving Reverse Dunk Shot
Driving Layup Shot
Driving Hook Shot
Driving Floating Jump Shot
Driving Floating Bank Jump Shot
Driving Finger Roll Layup Shot
Driving Dunk Shot
Driving Bank Hook Shot
Cutting Layup Shot
Cutting Finger Roll Layup Shot
Cutting Dunk Shot
Alley Oop Layup shot
Alley Oop Dunk Shot
Action Type
−1.0
−0.5
0.0
0.5
DFbetaS

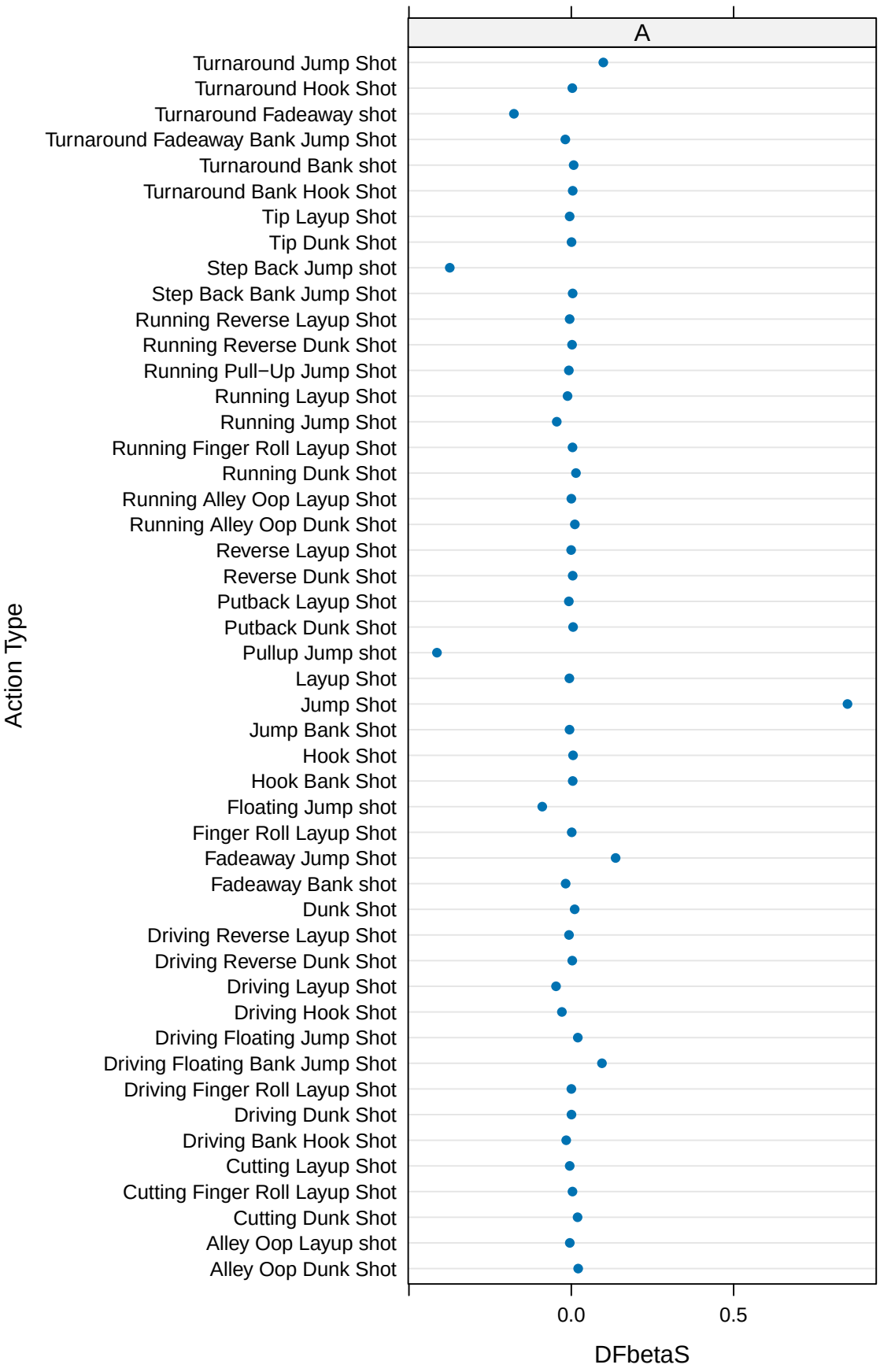
```r
plot(infl,which="dfbetas",parameters=c(3),xlab="DFbetaS",ylab="Action Type")
```


A
Turnaround Jump Shot
Turnaround Hook Shot
Turnaround Fadeaway shot
Turnaround Fadeaway Bank Jump Shot
Turnaround Bank shot
Turnaround Bank Hook Shot
Tip Layup Shot
Tip Dunk Shot
Step Back Jump shot
Step Back Bank Jump Shot
Running Reverse Layup Shot
Running Reverse Dunk Shot
Running Pull−Up Jump Shot
Running Layup Shot
Running Jump Shot
Running Finger Roll Layup Shot
Running Dunk Shot
Running Alley Oop Layup Shot
Running Alley Oop Dunk Shot
Reverse Layup Shot
Reverse Dunk Shot
Putback Layup Shot
Putback Dunk Shot
Pullup Jump shot
Layup Shot
Jump Shot
Jump Bank Shot
Hook Shot
Hook Bank Shot
Floating Jump shot
Finger Roll Layup Shot
Fadeaway Jump Shot
Fadeaway Bank shot
Dunk Shot
Driving Reverse Layup Shot
Driving Reverse Dunk Shot
Driving Layup Shot
Driving Hook Shot
Driving Floating Jump Shot
Driving Floating Bank Jump Shot
Driving Finger Roll Layup Shot
Driving Dunk Shot
Driving Bank Hook Shot
Cutting Layup Shot
Cutting Finger Roll Layup Shot
Cutting Dunk Shot
Alley Oop Layup shot
Alley Oop Dunk Shot
Action Type
−0.2
−0.1
0.0
0.1
0.2
0.3
DFbetaS

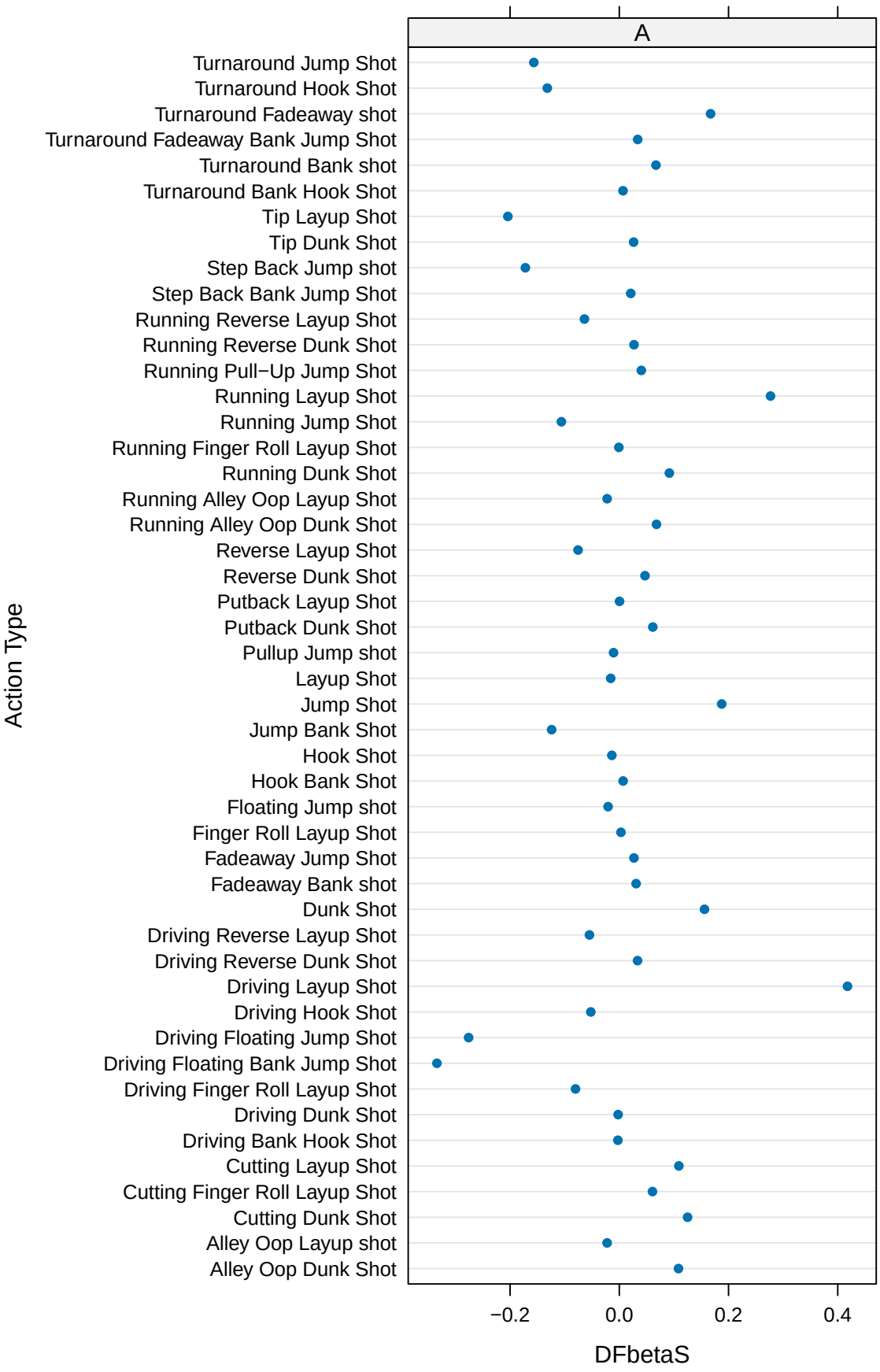
```r
plot(infl,which="dfbetas",parameters=c(4),xlab="DFbetaS",ylab="Action Type")
```


A
Turnaround Jump Shot
Turnaround Hook Shot
Turnaround Fadeaway shot
Turnaround Fadeaway Bank Jump Shot
Turnaround Bank shot
Turnaround Bank Hook Shot
Tip Layup Shot
Tip Dunk Shot
Step Back Jump shot
Step Back Bank Jump Shot
Running Reverse Layup Shot
Running Reverse Dunk Shot
Running Pull−Up Jump Shot
Running Layup Shot
Running Jump Shot
Running Finger Roll Layup Shot
Running Dunk Shot
Running Alley Oop Layup Shot
Running Alley Oop Dunk Shot
Reverse Layup Shot
Reverse Dunk Shot
Putback Layup Shot
Putback Dunk Shot
Pullup Jump shot
Layup Shot
Jump Shot
Jump Bank Shot
Hook Shot
Hook Bank Shot
Floating Jump shot
Finger Roll Layup Shot
Fadeaway Jump Shot
Fadeaway Bank shot
Dunk Shot
Driving Reverse Layup Shot
Driving Reverse Dunk Shot
Driving Layup Shot
Driving Hook Shot
Driving Floating Jump Shot
Driving Floating Bank Jump Shot
Driving Finger Roll Layup Shot
Driving Dunk Shot
Driving Bank Hook Shot
Cutting Layup Shot
Cutting Finger Roll Layup Shot
Cutting Dunk Shot
Alley Oop Layup shot
Alley Oop Dunk Shot
Action Type
−0.3
−0.2
−0.1
0.0
0.1
DFbetaS

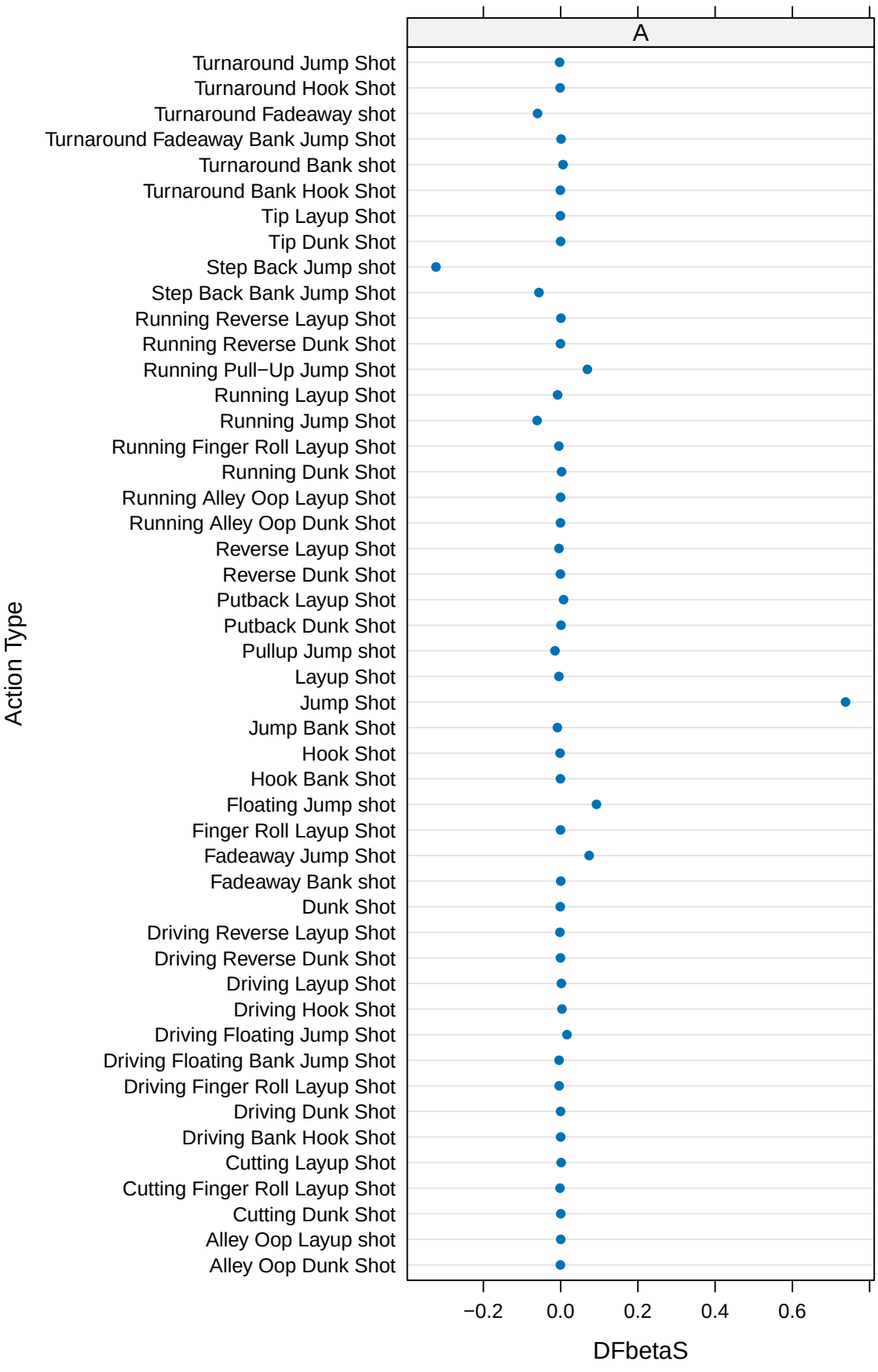
```r
plot(infl,which="dfbetas",parameters=c(5),xlab="DFbetaS",ylab="Action Type")
```


A
Turnaround Jump Shot
Turnaround Hook Shot
Turnaround Fadeaway shot
Turnaround Fadeaway Bank Jump Shot
Turnaround Bank shot
Turnaround Bank Hook Shot
Tip Layup Shot
Tip Dunk Shot
Step Back Jump shot
Step Back Bank Jump Shot
Running Reverse Layup Shot
Running Reverse Dunk Shot
Running Pull−Up Jump Shot
Running Layup Shot
Running Jump Shot
Running Finger Roll Layup Shot
Running Dunk Shot
Running Alley Oop Layup Shot
Running Alley Oop Dunk Shot
Reverse Layup Shot
Reverse Dunk Shot
Putback Layup Shot
Putback Dunk Shot
Pullup Jump shot
Layup Shot
Jump Shot
Jump Bank Shot
Hook Shot
Hook Bank Shot
Floating Jump shot
Finger Roll Layup Shot
Fadeaway Jump Shot
Fadeaway Bank shot
Dunk Shot
Driving Reverse Layup Shot
Driving Reverse Dunk Shot
Driving Layup Shot
Driving Hook Shot
Driving Floating Jump Shot
Driving Floating Bank Jump Shot
Driving Finger Roll Layup Shot
Driving Dunk Shot
Driving Bank Hook Shot
Cutting Layup Shot
Cutting Finger Roll Layup Shot
Cutting Dunk Shot
Alley Oop Layup shot
Alley Oop Dunk Shot
Action Type
0.0
0.5
DFbetaS

```r
plot(infl,which="dfbetas",parameters=c(6),xlab="DFbetaS",ylab="Action Type")
```

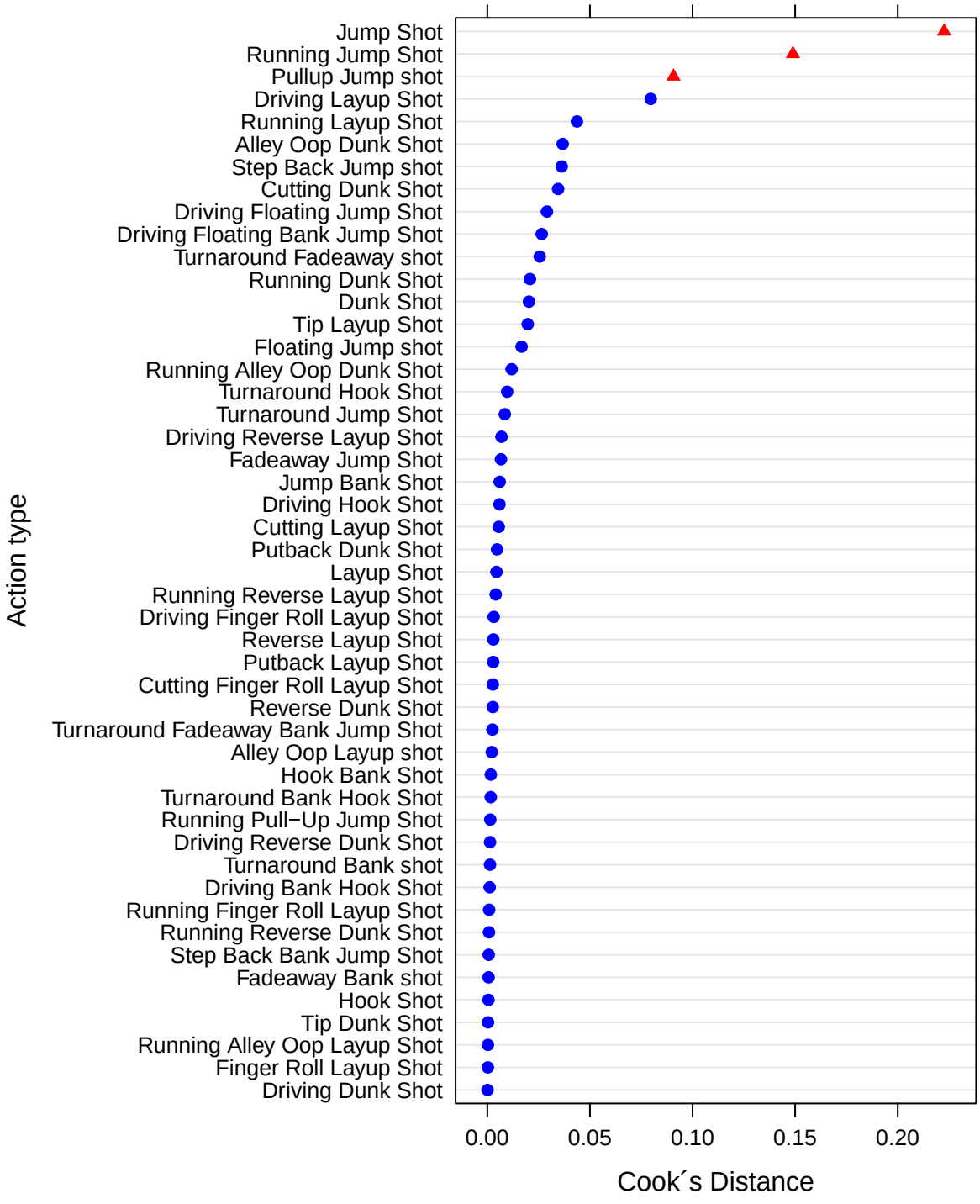

A
Turnaround Jump Shot
Turnaround Hook Shot
Turnaround Fadeaway shot
Turnaround Fadeaway Bank Jump Shot
Turnaround Bank shot
Turnaround Bank Hook Shot
Tip Layup Shot
Tip Dunk Shot
Step Back Jump shot
Step Back Bank Jump Shot
Running Reverse Layup Shot
Running Reverse Dunk Shot
Running Pull−Up Jump Shot
Running Layup Shot
Running Jump Shot
Running Finger Roll Layup Shot
Running Dunk Shot
Running Alley Oop Layup Shot
Running Alley Oop Dunk Shot
Reverse Layup Shot
Reverse Dunk Shot
Putback Layup Shot
Putback Dunk Shot
Pullup Jump shot
Layup Shot
Jump Shot
Jump Bank Shot
Hook Shot
Hook Bank Shot
Floating Jump shot
Finger Roll Layup Shot
Fadeaway Jump Shot
Fadeaway Bank shot
Dunk Shot
Driving Reverse Layup Shot
Driving Reverse Dunk Shot
Driving Layup Shot
Driving Hook Shot
Driving Floating Jump Shot
Driving Floating Bank Jump Shot
Driving Finger Roll Layup Shot
Driving Dunk Shot
Driving Bank Hook Shot
Cutting Layup Shot
Cutting Finger Roll Layup Shot
Cutting Dunk Shot
Alley Oop Layup shot
Alley Oop Dunk Shot
Action Type
−0.2
0.0
0.2
0.4
DFbetaS

```r
plot(infl,which="dfbetas",parameters=c(7),xlab="DFbetaS",ylab="Action Type")
```


A
Turnaround Jump Shot
Turnaround Hook Shot
Turnaround Fadeaway shot
Turnaround Fadeaway Bank Jump Shot
Turnaround Bank shot
Turnaround Bank Hook Shot
Tip Layup Shot
Tip Dunk Shot
Step Back Jump shot
Step Back Bank Jump Shot
Running Reverse Layup Shot
Running Reverse Dunk Shot
Running Pull−Up Jump Shot
Running Layup Shot
Running Jump Shot
Running Finger Roll Layup Shot
Running Dunk Shot
Running Alley Oop Layup Shot
Running Alley Oop Dunk Shot
Reverse Layup Shot
Reverse Dunk Shot
Putback Layup Shot
Putback Dunk Shot
Pullup Jump shot
Layup Shot
Jump Shot
Jump Bank Shot
Hook Shot
Hook Bank Shot
Floating Jump shot
Finger Roll Layup Shot
Fadeaway Jump Shot
Fadeaway Bank shot
Dunk Shot
Driving Reverse Layup Shot
Driving Reverse Dunk Shot
Driving Layup Shot
Driving Hook Shot
Driving Floating Jump Shot
Driving Floating Bank Jump Shot
Driving Finger Roll Layup Shot
Driving Dunk Shot
Driving Bank Hook Shot
Cutting Layup Shot
Cutting Finger Roll Layup Shot
Cutting Dunk Shot
Alley Oop Layup shot
Alley Oop Dunk Shot
Action Type
−0.2
0.0
0.2
0.4
0.6
DFbetaS

Cooks Distance

```
plot(infl, which="cook", cutoff=4/nrow(ranef(final.model)$ACTION_TYPE), sort=TRUE, xlab="Cook´s Distanc
```



Removing Outliers for ACTION_TYPE

```
# removing influential groups from data
remove_categories <- c("Jump Shot", "Running Jump Shot")
```

```r
sub = subset(DenverNuggetsData, !(ACTION_TYPE %in% remove_categories))

# fit final model without influential observations
final.model.removed.outliers <- glmer(SHOT_MADE_FLAG ~ SHOT_ZONE_BASIC
                + (1|ACTION_TYPE)+(1|ACTION_TYPE:PLAYER_NAME),
                family = "binomial", data = sub,
                control=glmerControl(optimizer="bobyqa",optCtrl=list(maxfun=2e5)))

summary(final.model.removed.outliers)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula:
## SHOT_MADE_FLAG ~ SHOT_ZONE_BASIC + (1 | ACTION_TYPE) + (1 | ACTION_TYPE:PLAYER_NAME)
##    Data: sub
## Control: glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 2e+05))
##
##      AIC      BIC   logLik deviance df.resid
##   6182.6   6241.2  -3082.3   6164.6     4956
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.3249 -0.8382  0.3516  0.7952  1.8819
##
## Random effects:
##  Groups                  Name        Variance Std.Dev.
##  ACTION_TYPE:PLAYER_NAME (Intercept) 0.1008   0.3175
##  ACTION_TYPE             (Intercept) 0.3627   0.6023
## Number of obs: 4965, groups:  ACTION_TYPE:PLAYER_NAME, 467; ACTION_TYPE, 46
##
## Fixed effects:
##                                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)                         -0.430134   0.187222  -2.297  0.02159 *
## SHOT_ZONE_BASICBackcourt            -0.670835   1.140867  -0.588  0.55653
## SHOT_ZONE_BASICIn The Paint (Non-RA)  0.527474   0.162181   3.252  0.00114 **
## SHOT_ZONE_BASICLeft Corner 3          0.236563   0.399509   0.592  0.55376
## SHOT_ZONE_BASICMid-Range              0.028299   0.133917   0.211  0.83264
## SHOT_ZONE_BASICRestricted Area        1.586870   0.182586   8.691  < 2e-16 ***
## SHOT_ZONE_BASICRight Corner 3         0.007594   0.408450   0.019  0.98517
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##                    (Intr) SHOT_ZONE_BASICB SHOTP( SHOT_ZONE_BASICLC3
## SHOT_ZONE_BASICB   -0.076
## SHOT_ZONTP(        -0.733  0.097
## SHOT_ZONE_BASICLC3 -0.119  0.018            0.140
## SHOT_ZONE_BASICM   -0.507  0.084            0.616  0.168
## SHOT_ZONE_A        -0.790  0.084            0.824  0.123
## SHOT_ZONE_BASICRC3 -0.139  0.020            0.157  0.069
##                    SHOT_ZONE_BASICM SHOT_A
## SHOT_ZONE_BASICB
## SHOT_ZONTP(
```

```
## SHOT_ZONE_BASICLC3
## SHOT_ZONE_BASICM
## SHOT_ZONE_A         0.538
## SHOT_ZONE_BASICRC3  0.175           0.141
```

```
mlmhelpr::icc(final.model.removed.outliers)
```

```
##                                grps   icc
## 1 ACTION_TYPE:PLAYER_NAME (Intercept) 0.030
## 2             ACTION_TYPE (Intercept) 0.099
```

```
auc(roc(sub$SHOT_MADE_FLAG,  predict(final.model.removed.outliers, type = "response")))
```

```
## Setting levels: control = 0, case = 1
```

```
## Setting direction: controls < cases
```

```
## Area under the curve: 0.7342
```