Simulated VulfenSarah (Nunez) hits landed with first 50 Wolfey (Mazvidal) as a Testing Set

Janis Corona

11/17/2019

This script was edited months afterwards to correct the regex error with the knee to body getting selected as a knee strike, so now the counts are accurate. Updated 2/11/2020.

Put in the first 50 samples of Wolfey as the testing set to see the prediction accuracy of hits landed with VulfenSarah hits landed comparison

Vulfen1 <- read.csv('Nunez-ML-Analysis-editedRegex.csv',   
 sep=',', header=TRUE,   
 na.strings=c('','NA'))

Wolfey <- read.csv('wolfey\_addedFeatures.csv',   
 sep=',', header=TRUE, nrows=50,  
 na.strings=c('','NA'))

Wolfey <- Wolfey[,c(1:7,8:15,48:155)]#omit all X1 landed and x2 received  
Vulfen <- Vulfen1[,c(1:7,8:15,48:155)]#omit all X1 landed and x2 received

library(caret)  
library(randomForest)  
library(MASS)  
library(gbm)  
library(dplyr)

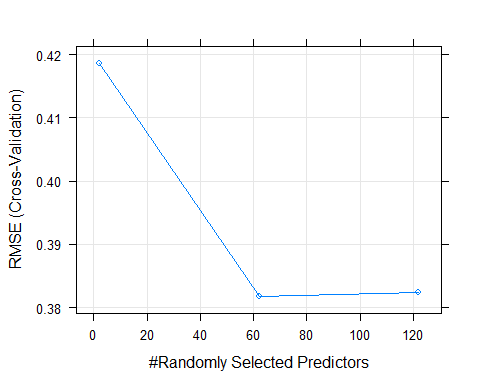
set.seed(189678345)

trainingSet <- Vulfen  
testingSet <- Wolfey

system.time(rfMod <- train(TotLandsX1~., method='rf', data=(trainingSet),   
 trControl=trainControl(method='cv'), number=5))

## user system elapsed   
## 30.69 0.20 32.56

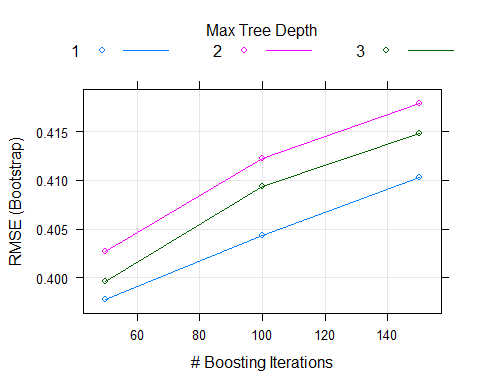
plot(rfMod)



system.time(gbmMod <- train(TotLandsX1~., method='gbm', data=trainingSet, verbose=FALSE ))

## user system elapsed   
## 17.83 0.08 19.26

plot(gbmMod)



predRF <- round(predict(rfMod, testingSet))  
predGbm <- round(predict(gbmMod, testingSet))  
  
predDF <- data.frame(predRF, predGbm, type=testingSet$TotLandsX1)  
predDF

## predRF predGbm type  
## 1 0 0 0  
## 2 0 0 1  
## 3 0 0 0  
## 4 0 0 0  
## 5 0 0 0  
## 6 0 0 0  
## 7 0 0 0  
## 8 0 0 0  
## 9 0 0 0  
## 10 0 0 0  
## 11 0 0 0  
## 12 0 0 2  
## 13 0 0 0  
## 14 0 0 0  
## 15 0 0 0  
## 16 0 0 0  
## 17 0 0 1  
## 18 0 0 0  
## 19 0 0 0  
## 20 0 0 0  
## 21 0 0 0  
## 22 0 0 0  
## 23 0 0 0  
## 24 1 0 1  
## 25 0 0 0  
## 26 1 0 0  
## 27 0 0 0  
## 28 1 0 0  
## 29 0 0 0  
## 30 1 0 0  
## 31 1 0 0  
## 32 0 0 1  
## 33 0 0 0  
## 34 0 0 0  
## 35 0 0 0  
## 36 1 0 0  
## 37 0 0 0  
## 38 0 0 0  
## 39 1 0 0  
## 40 1 0 0  
## 41 0 0 0  
## 42 0 0 0  
## 43 0 0 0  
## 44 0 0 0  
## 45 0 0 0  
## 46 0 0 0  
## 47 0 0 1  
## 48 0 0 0  
## 49 0 0 1  
## 50 0 0 0

Accuracy of the random forest algorithm:

sum <- sum(predRF==testingSet$TotLandsX1)  
length <- length(testingSet$TotLandsX1)  
accuracy\_rfMod <- (sum/length)   
accuracy\_rfMod

## [1] 0.74

Accuracy of the Generalized Boosted Machines algorithm:

sum <- sum(predGbm==testingSet$TotLandsX1)  
accuracy\_Gbm <- (sum/length)   
accuracy\_Gbm

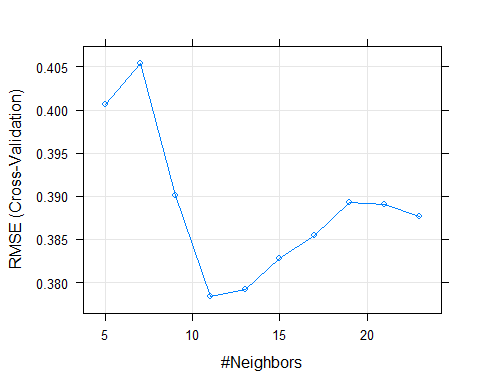
## [1] 0.86

Now, use the K-nearest neighbor or KNN algorithm.

system.time(knnMod <- train(TotLandsX1 ~ .,  
 method='knn', preProcess=c('center','scale'),  
 tuneLength=10, trControl=trainControl(method='cv'), data=trainingSet))

## user system elapsed   
## 3.92 0.06 5.16

plot(knnMod)

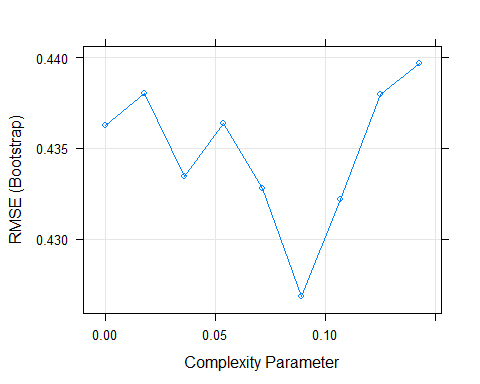
 From the above plot n=11 seems to have the lowest Root mean Squared error.

Now, use the recursive partitioning Trees alogorithm, a type of decision trees methos.

system.time(rpartMod <- train(TotLandsX1~ ., method='rpart', tuneLength=9, data=trainingSet))

## user system elapsed   
## 5.95 0.00 6.24

plot(rpartMod)



Now, use the generalized linear machines algorithm that encompasses linear and logistic regression models.

system.time(glmMod <- train(TotLandsX1~ .,   
 method='glm', data=trainingSet))

## user system elapsed   
## 2.11 0.00 2.14

predKNN <- round(predict(knnMod, testingSet))  
predRPART <- round(predict(rpartMod, testingSet))  
predGLM <- round(predict(glmMod, testingSet))  
  
df3 <- cbind(predKNN, predRPART, predGLM,testingSet$TotLandsX1)  
colnames(df3)[4] <- 'TrueValue'  
head(df3);tail(df3)

## predKNN predRPART predGLM TrueValue  
## 1 0 0 0 0  
## 2 0 0 0 1  
## 3 0 0 0 0  
## 4 0 0 0 0  
## 5 0 0 0 0  
## 6 0 0 0 0

## predKNN predRPART predGLM TrueValue  
## 45 1 0 0 0  
## 46 0 0 0 0  
## 47 1 1 1 1  
## 48 1 1 0 0  
## 49 1 1 1 1  
## 50 1 1 1 0

The above output shows the predicted values for the KNN, Rpart, and GLM models as well as the actual or true value.

length=length(testingSet$TotLandsX1)  
  
sumKNN <- sum(predKNN==testingSet$TotLandsX1)  
sumRPart <- sum(predRPART==testingSet$TotLandsX1)  
sumGLM <- sum(predGLM==testingSet$TotLandsX1)

The accuracy in prediction for Random Forest, GBM, KNN, Rpart, and GLM are (respectively):

accuracy\_KNN <- sumKNN/length   
accuracy\_RPART <- sumRPart/length   
accuracy\_GLM <- sumGLM/length   
  
accuracy\_rfMod; accuracy\_Gbm; accuracy\_KNN; accuracy\_RPART; accuracy\_GLM

## [1] 0.74

## [1] 0.86

## [1] 0.58

## [1] 0.86

## [1] 0.62

From the above algorithm accuracies, GBM and Rpart scored the best with 86%, then random forest with 74%.

predDF3 <- data.frame(predRF,predGbm,df3)  
head(predDF3);tail(predDF3)

## predRF predGbm predKNN predRPART predGLM TrueValue  
## 1 0 0 0 0 0 0  
## 2 0 0 0 0 0 1  
## 3 0 0 0 0 0 0  
## 4 0 0 0 0 0 0  
## 5 0 0 0 0 0 0  
## 6 0 0 0 0 0 0

## predRF predGbm predKNN predRPART predGLM TrueValue  
## 45 0 0 1 0 0 0  
## 46 0 0 0 0 0 0  
## 47 0 0 1 1 1 1  
## 48 0 0 1 1 0 0  
## 49 0 0 1 1 1 1  
## 50 0 0 1 1 1 0

colnames(predDF3)

## [1] "predRF" "predGbm" "predKNN" "predRPART" "predGLM" "TrueValue"

results <- c(round(accuracy\_rfMod,2),   
 round(accuracy\_Gbm,2),   
 round(accuracy\_KNN,2), round(accuracy\_RPART,2),  
 round(accuracy\_GLM,2),   
 round(100,2))  
  
results <- as.factor(results)  
results <- t(data.frame(results))  
colnames(results) <- colnames(predDF3)  
Results <- rbind(predDF3, results)   
head(Results);tail(Results)

## predRF predGbm predKNN predRPART predGLM TrueValue  
## 1 0 0 0 0 0 0  
## 2 0 0 0 0 0 1  
## 3 0 0 0 0 0 0  
## 4 0 0 0 0 0 0  
## 5 0 0 0 0 0 0  
## 6 0 0 0 0 0 0

## predRF predGbm predKNN predRPART predGLM TrueValue  
## 46 0 0 0 0 0 0  
## 47 0 0 1 1 1 1  
## 48 0 0 1 1 0 0  
## 49 0 0 1 1 1 1  
## 50 0 0 1 1 1 0  
## results 0.74 0.86 0.58 0.86 0.62 100

best <- order(results, decreasing=TRUE)  
bestResults <- Results[,best[1:3]]  
bestResults

## TrueValue predGbm predRPART  
## 1 0 0 0  
## 2 1 0 0  
## 3 0 0 0  
## 4 0 0 0  
## 5 0 0 0  
## 6 0 0 0  
## 7 0 0 0  
## 8 0 0 0  
## 9 0 0 0  
## 10 0 0 0  
## 11 0 0 0  
## 12 2 0 0  
## 13 0 0 0  
## 14 0 0 0  
## 15 0 0 0  
## 16 0 0 0  
## 17 1 0 0  
## 18 0 0 0  
## 19 0 0 0  
## 20 0 0 0  
## 21 0 0 0  
## 22 0 0 0  
## 23 0 0 0  
## 24 1 0 0  
## 25 0 0 0  
## 26 0 0 0  
## 27 0 0 0  
## 28 0 0 0  
## 29 0 0 0  
## 30 0 0 0  
## 31 0 0 0  
## 32 1 0 0  
## 33 0 0 0  
## 34 0 0 0  
## 35 0 0 0  
## 36 0 0 0  
## 37 0 0 0  
## 38 0 0 0  
## 39 0 0 0  
## 40 0 0 0  
## 41 0 0 0  
## 42 0 0 0  
## 43 0 0 0  
## 44 0 0 0  
## 45 0 0 0  
## 46 0 0 0  
## 47 1 0 1  
## 48 0 0 1  
## 49 1 0 1  
## 50 0 0 1  
## results 100 0.86 0.86

library(dplyr)  
BestPredictedHit <- subset(Results, Results$TrueValue == 1 | Results$TrueValue ==2)  
length=length(BestPredictedHit$TrueValue)  
  
sumRF <- sum(BestPredictedHit$predRF==BestPredictedHit$TrueValue)  
sumGbm <- sum(BestPredictedHit$predGbm==BestPredictedHit$TrueValue)  
sumKNN <- sum(BestPredictedHit$predKNN==BestPredictedHit$TrueValue)  
sumRPart <- sum(BestPredictedHit$predRPART==BestPredictedHit$TrueValue)  
sumGLM <- sum(BestPredictedHit$predGLM==BestPredictedHit$TrueValue)  
  
accuracy\_RF <- round(sumRF/length,2)  
accuracy\_Gbm <- round(sumGbm/length,2)   
accuracy\_KNN <- round(sumKNN/length,2)   
accuracy\_RPART <- round(sumRPart/length,2)   
accuracy\_GLM <- round(sumGLM/length,2)   
Truth <- round(sum(BestPredictedHit$TrueValue==BestPredictedHit$TrueValue)/length,2)  
  
HitAccuracy <- c(accuracy\_RF,accuracy\_Gbm,accuracy\_KNN,accuracy\_RPART,  
 accuracy\_GLM,Truth)  
HitAccuracy <- t(data.frame(as.factor(HitAccuracy)))  
colnames(HitAccuracy) <- colnames(BestPredictedHit)  
BestPredictedHit1 <- rbind(BestPredictedHit,HitAccuracy)  
row.names(BestPredictedHit1)[8] <- 'Accuracy'  
BestPredictedHit1

## predRF predGbm predKNN predRPART predGLM TrueValue  
## 2 0 0 0 0 0 1  
## 12 0 0 0 0 1 2  
## 17 0 0 0 0 1 1  
## 24 1 0 1 0 1 1  
## 32 0 0 1 0 1 1  
## 47 0 0 1 1 1 1  
## 49 0 0 1 1 1 1  
## Accuracy 0.14 0 0.57 0.29 0.71 1

KNN and GLM were more accurate in guessing which simulations would produce a hit landed by VulfenSarah.

testHits <- testingSet[row.names(BestPredictedHit1)[1:7],]  
Hits <-cbind(BestPredictedHit1[1:7,],testHits)  
Hits

## predRF predGbm predKNN predRPART predGLM TrueValue Round SecondsIntoRound  
## 2 0 0 0 0 0 1 1 4  
## 12 0 0 0 0 1 2 1 49  
## 17 0 0 0 0 1 1 1 67  
## 24 1 0 1 0 1 1 1 95  
## 32 0 0 1 0 1 1 1 139  
## 47 0 0 1 1 1 1 1 235  
## 49 0 0 1 1 1 1 1 245  
## lastAction SecondsLastRoundAction cmTotHitsR.X1 cmTotHitsL.X1 cmTotHitsM.X1  
## 2 1 3 0 1 0  
## 12 48 1 2 3 1  
## 17 66 1 2 4 3  
## 24 94 1 2 5 3  
## 32 138 1 2 6 7  
## 47 234 1 2 7 16  
## 49 238 7 2 8 17  
## TotLandsX1 TotMissedX1 TotReceivedX1 cmTotHitsR.X2 cmTotHitsL.X2  
## 2 1 0 0 1 0  
## 12 2 0 0 3 2  
## 17 1 0 0 4 2  
## 24 1 0 0 6 2  
## 32 1 1 0 7 2  
## 47 1 0 0 9 2  
## 49 1 0 0 10 2  
## cmTotHitsM.X2 TotLandsX2 TotMissedX2 Crossl.X2 Kneel.X2 Elbowl.X2 Hookl.X2  
## 2 0 0 0 0 0 0 0  
## 12 0 0 0 0 0 0 0  
## 17 0 0 0 0 0 0 0  
## 24 0 0 0 0 0 0 0  
## 32 1 0 0 0 0 0 0  
## 47 5 0 0 0 0 0 0  
## 49 5 0 0 0 0 0 0  
## Jabl.X2 Kickl.X2 upperl.X2 takedownl.X2 hammerl.X2 Cross2l.X2 Knee2l.X2  
## 2 0 0 0 0 0 0 0  
## 12 0 0 0 0 0 0 0  
## 17 0 0 0 0 0 0 0  
## 24 0 0 0 0 0 0 0  
## 32 0 0 0 0 0 0 0  
## 47 0 0 0 0 0 0 0  
## 49 0 0 0 0 0 0 0  
## Elbow2l.X2 Hook2l.X2 Jab2l.X2 Kick2l.X2 upper2l.X2 takedown2l.X2 hammer2l.X2  
## 2 0 0 0 0 0 0 0  
## 12 0 0 0 0 0 0 0  
## 17 0 0 0 0 0 0 0  
## 24 0 0 0 0 0 0 0  
## 32 0 0 0 0 0 0 0  
## 47 0 0 0 0 0 0 0  
## 49 0 0 0 0 0 0 0  
## Cross3l.X2 Knee3l.X2 Elbow3l.X2 Hook3l.X2 Jab3l.X2 Kick3l.X2 upper3l.X2  
## 2 0 0 0 0 0 0 0  
## 12 0 0 0 0 0 0 0  
## 17 0 0 0 0 0 0 0  
## 24 0 0 0 0 0 0 0  
## 32 0 0 0 0 0 0 0  
## 47 0 0 0 0 0 0 0  
## 49 0 0 0 0 0 0 0  
## takedown3l.X2 hammer3l.X2 Crossm.X1 Kneem.X1 Elbowm.X1 Hookm.X1 Jabm.X1  
## 2 0 0 0 0 0 0 0  
## 12 0 0 0 0 0 0 0  
## 17 0 0 0 0 0 0 0  
## 24 0 0 0 0 0 0 0  
## 32 0 0 0 0 0 0 0  
## 47 0 0 0 0 0 0 0  
## 49 0 0 0 0 0 0 0  
## Kickm.X1 upperm.X1 takedownm.X1 hammerm.X1 Cross2m.X1 Knee2m.X1 Elbow2m.X1  
## 2 0 0 0 0 0 0 0  
## 12 0 0 0 0 0 0 0  
## 17 0 0 0 0 0 0 0  
## 24 0 0 0 0 0 0 0  
## 32 0 0 0 0 0 0 0  
## 47 0 0 0 0 0 0 0  
## 49 0 0 0 0 0 0 0  
## Hook2m.X1 Jab2m.X1 Kick2m.X1 upper2m.X1 takedown2m.X1 hammer2m.X1 Cross3m.X1  
## 2 0 0 0 0 0 0 0  
## 12 0 0 0 0 0 0 0  
## 17 0 0 0 0 0 0 0  
## 24 0 0 0 0 0 0 0  
## 32 0 1 0 0 0 0 0  
## 47 0 0 0 0 0 0 0  
## 49 0 0 0 0 0 0 0  
## Knee3m.X1 Elbow3m.X1 Hook3m.X1 Jab3m.X1 Kick3m.X1 upper3m.X1 takedown3m.X1  
## 2 0 0 0 0 0 0 0  
## 12 0 0 0 0 0 0 0  
## 17 0 0 0 0 0 0 0  
## 24 0 0 0 0 0 0 0  
## 32 0 0 0 0 0 0 0  
## 47 0 0 0 0 0 0 0  
## 49 0 0 0 0 0 0 0  
## hammer3m.X1 Crossm.X2 Kneem.X2 Elbowm.X2 Hookm.X2 Jabm.X2 Kickm.X2 upperm.X2  
## 2 0 0 0 0 0 0 0 0  
## 12 0 0 0 0 0 0 0 0  
## 17 0 0 0 0 0 0 0 0  
## 24 0 0 0 0 0 0 0 0  
## 32 0 0 0 0 0 0 0 0  
## 47 0 0 0 0 0 0 0 0  
## 49 0 0 0 0 0 0 0 0  
## takedownm.X2 hammerm.X2 Cross2m.X2 Knee2m.X2 Elbow2m.X2 Hook2m.X2 Jab2m.X2  
## 2 0 0 0 0 0 0 0  
## 12 0 0 0 0 0 0 0  
## 17 0 0 0 0 0 0 0  
## 24 0 0 0 0 0 0 0  
## 32 0 0 0 0 0 0 0  
## 47 0 0 0 0 0 0 0  
## 49 0 0 0 0 0 0 0  
## Kick2m.X2 upper2m.X2 takedown2m.X2 hammer2m.X2 Cross3m.X2 Knee3m.X2  
## 2 0 0 0 0 0 0  
## 12 0 0 0 0 0 0  
## 17 0 0 0 0 0 0  
## 24 0 0 0 0 0 0  
## 32 0 0 0 0 0 0  
## 47 0 0 0 0 0 0  
## 49 0 0 0 0 0 0  
## Elbow3m.X2 Hook3m.X2 Jab3m.X2 Kick3m.X2 upper3m.X2 takedown3m.X2 hammer3m.X2  
## 2 0 0 0 0 0 0 0  
## 12 0 0 0 0 0 0 0  
## 17 0 0 0 0 0 0 0  
## 24 0 0 0 0 0 0 0  
## 32 0 0 0 0 0 0 0  
## 47 0 0 0 0 0 0 0  
## 49 0 0 0 0 0 0 0  
## Crossr.X1 Kneer.X1 Elbowr.X1 Hookr.X1 Jabr.X1 Kickr.X1 upperr.X1  
## 2 0 0 0 0 0 0 0  
## 12 0 0 0 0 0 0 0  
## 17 0 0 0 0 0 0 0  
## 24 0 0 0 0 0 0 0  
## 32 0 0 0 0 0 0 0  
## 47 0 0 0 0 0 0 0  
## 49 0 0 0 0 0 0 0  
## takedownr.X1 hammerr.X1 Cross2r.X1 Knee2r.X1 Elbow2r.X1 Hook2r.X1 Jab2r.X1  
## 2 0 0 0 0 0 0 0  
## 12 0 0 0 0 0 0 0  
## 17 0 0 0 0 0 0 0  
## 24 0 0 0 0 0 0 0  
## 32 0 0 0 0 0 0 0  
## 47 0 0 0 0 0 0 0  
## 49 0 0 0 0 0 0 0  
## Kick2r.X1 upper2r.X1 takedown2r.X1 hammer2r.X1 Cross3r.X1 Knee3r.X1  
## 2 0 0 0 0 0 0  
## 12 0 0 0 0 0 0  
## 17 0 0 0 0 0 0  
## 24 0 0 0 0 0 0  
## 32 0 0 0 0 0 0  
## 47 0 0 0 0 0 0  
## 49 0 0 0 0 0 0  
## Elbow3r.X1 Hook3r.X1 Jab3r.X1 Kick3r.X1 upper3r.X1 takedown3r.X1 hammer3r.X1  
## 2 0 0 0 0 0 0 0  
## 12 0 0 0 0 0 0 0  
## 17 0 0 0 0 0 0 0  
## 24 0 0 0 0 0 0 0  
## 32 0 0 0 0 0 0 0  
## 47 0 0 0 0 0 0 0  
## 49 0 0 0 0 0 0 0

The above table shows the 7 simulations of another fighter in our trained model and the results of those hits that were landed in this testing set against the prediction of a hit landed with those algorithms for machine learning: random forest (rf), global boosted machines (gbm), k nearest neighbors (KNN), recursive partitioning and regression trees (rpart), generalized linear models (glm), and the true testing set value aiming to predict for hits landed.

Aside, put aside the above, some time later as in months. What about looking at those instances where Vulfen lands 2 hits in one second and plotting this against the number of hits and seconds that passed? Lets do that.

library(tidyr)

Aside: This is the un-altered table, everything works as it should, the code to grab and extract each action is good. I originally thought there was a calculation problem, but the fields for X1’s landed,missed, received actions were omitted in the beginning of this script. Carry on.

X1\_2plus <- subset(Vulfen1, Vulfen1$TotLandsX1 > 1)   
X1\_2plus\_lands\_X1 <- X1\_2plus[,c(2,4:8,20:47)]  
X1\_2plus\_lands\_X1

## SecondsIntoRound SecondsLastRoundAction cmTotHitsR.X1 cmTotHitsL.X1  
## 46 164 1 2 10  
## 160 44 1 0 13  
## cmTotHitsM.X1 TotLandsX1 Notes Crossl.X1 Kneel.X1 Elbowl.X1 Hookl.X1  
## 46 30 2 Tate 0 0 0 0  
## 160 22 2 Rousey 0 0 0 0  
## Jabl.X1 Kickl.X1 upperl.X1 takedownl.X1 hammerl.X1 Cross2l.X1 Knee2l.X1  
## 46 1 0 0 0 0 1 0  
## 160 1 0 0 0 0 1 0  
## Elbow2l.X1 Hook2l.X1 Jab2l.X1 Kick2l.X1 upper2l.X1 takedown2l.X1  
## 46 0 0 0 0 0 0  
## 160 0 0 0 0 0 0  
## hammer2l.X1 Cross3l.X1 Knee3l.X1 Elbow3l.X1 Hook3l.X1 Jab3l.X1 Kick3l.X1  
## 46 0 0 0 0 0 0 0  
## 160 0 0 0 0 0 0 0  
## upper3l.X1 takedown3l.X1 hammer3l.X1  
## 46 0 0 0  
## 160 0 0 0

x1\_2plus\_lands\_tidy <- gather(X1\_2plus\_lands\_X1, 'actionReaction', 'actionCounts', 8:34)  
x1\_2plus\_lands\_tidy

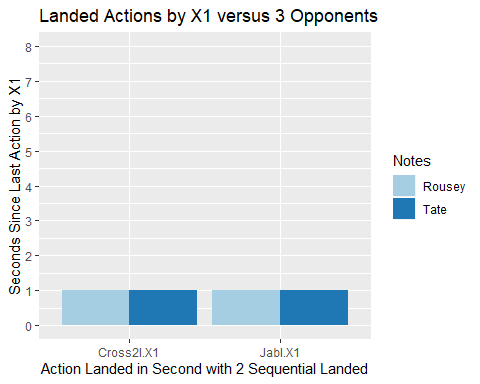
## SecondsIntoRound SecondsLastRoundAction cmTotHitsR.X1 cmTotHitsL.X1  
## 1 164 1 2 10  
## 2 44 1 0 13  
## 3 164 1 2 10  
## 4 44 1 0 13  
## 5 164 1 2 10  
## 6 44 1 0 13  
## 7 164 1 2 10  
## 8 44 1 0 13  
## 9 164 1 2 10  
## 10 44 1 0 13  
## 11 164 1 2 10  
## 12 44 1 0 13  
## 13 164 1 2 10  
## 14 44 1 0 13  
## 15 164 1 2 10  
## 16 44 1 0 13  
## 17 164 1 2 10  
## 18 44 1 0 13  
## 19 164 1 2 10  
## 20 44 1 0 13  
## 21 164 1 2 10  
## 22 44 1 0 13  
## 23 164 1 2 10  
## 24 44 1 0 13  
## 25 164 1 2 10  
## 26 44 1 0 13  
## 27 164 1 2 10  
## 28 44 1 0 13  
## 29 164 1 2 10  
## 30 44 1 0 13  
## 31 164 1 2 10  
## 32 44 1 0 13  
## 33 164 1 2 10  
## 34 44 1 0 13  
## 35 164 1 2 10  
## 36 44 1 0 13  
## 37 164 1 2 10  
## 38 44 1 0 13  
## 39 164 1 2 10  
## 40 44 1 0 13  
## 41 164 1 2 10  
## 42 44 1 0 13  
## 43 164 1 2 10  
## 44 44 1 0 13  
## 45 164 1 2 10  
## 46 44 1 0 13  
## 47 164 1 2 10  
## 48 44 1 0 13  
## 49 164 1 2 10  
## 50 44 1 0 13  
## 51 164 1 2 10  
## 52 44 1 0 13  
## 53 164 1 2 10  
## 54 44 1 0 13  
## cmTotHitsM.X1 TotLandsX1 Notes actionReaction actionCounts  
## 1 30 2 Tate Crossl.X1 0  
## 2 22 2 Rousey Crossl.X1 0  
## 3 30 2 Tate Kneel.X1 0  
## 4 22 2 Rousey Kneel.X1 0  
## 5 30 2 Tate Elbowl.X1 0  
## 6 22 2 Rousey Elbowl.X1 0  
## 7 30 2 Tate Hookl.X1 0  
## 8 22 2 Rousey Hookl.X1 0  
## 9 30 2 Tate Jabl.X1 1  
## 10 22 2 Rousey Jabl.X1 1  
## 11 30 2 Tate Kickl.X1 0  
## 12 22 2 Rousey Kickl.X1 0  
## 13 30 2 Tate upperl.X1 0  
## 14 22 2 Rousey upperl.X1 0  
## 15 30 2 Tate takedownl.X1 0  
## 16 22 2 Rousey takedownl.X1 0  
## 17 30 2 Tate hammerl.X1 0  
## 18 22 2 Rousey hammerl.X1 0  
## 19 30 2 Tate Cross2l.X1 1  
## 20 22 2 Rousey Cross2l.X1 1  
## 21 30 2 Tate Knee2l.X1 0  
## 22 22 2 Rousey Knee2l.X1 0  
## 23 30 2 Tate Elbow2l.X1 0  
## 24 22 2 Rousey Elbow2l.X1 0  
## 25 30 2 Tate Hook2l.X1 0  
## 26 22 2 Rousey Hook2l.X1 0  
## 27 30 2 Tate Jab2l.X1 0  
## 28 22 2 Rousey Jab2l.X1 0  
## 29 30 2 Tate Kick2l.X1 0  
## 30 22 2 Rousey Kick2l.X1 0  
## 31 30 2 Tate upper2l.X1 0  
## 32 22 2 Rousey upper2l.X1 0  
## 33 30 2 Tate takedown2l.X1 0  
## 34 22 2 Rousey takedown2l.X1 0  
## 35 30 2 Tate hammer2l.X1 0  
## 36 22 2 Rousey hammer2l.X1 0  
## 37 30 2 Tate Cross3l.X1 0  
## 38 22 2 Rousey Cross3l.X1 0  
## 39 30 2 Tate Knee3l.X1 0  
## 40 22 2 Rousey Knee3l.X1 0  
## 41 30 2 Tate Elbow3l.X1 0  
## 42 22 2 Rousey Elbow3l.X1 0  
## 43 30 2 Tate Hook3l.X1 0  
## 44 22 2 Rousey Hook3l.X1 0  
## 45 30 2 Tate Jab3l.X1 0  
## 46 22 2 Rousey Jab3l.X1 0  
## 47 30 2 Tate Kick3l.X1 0  
## 48 22 2 Rousey Kick3l.X1 0  
## 49 30 2 Tate upper3l.X1 0  
## 50 22 2 Rousey upper3l.X1 0  
## 51 30 2 Tate takedown3l.X1 0  
## 52 22 2 Rousey takedown3l.X1 0  
## 53 30 2 Tate hammer3l.X1 0  
## 54 22 2 Rousey hammer3l.X1 0

x1\_2plus\_lands\_counts <- subset(x1\_2plus\_lands\_tidy, x1\_2plus\_lands\_tidy$actionCounts > 0)  
x1\_2plus\_lands\_counts

## SecondsIntoRound SecondsLastRoundAction cmTotHitsR.X1 cmTotHitsL.X1  
## 9 164 1 2 10  
## 10 44 1 0 13  
## 19 164 1 2 10  
## 20 44 1 0 13  
## cmTotHitsM.X1 TotLandsX1 Notes actionReaction actionCounts  
## 9 30 2 Tate Jabl.X1 1  
## 10 22 2 Rousey Jabl.X1 1  
## 19 30 2 Tate Cross2l.X1 1  
## 20 22 2 Rousey Cross2l.X1 1

Now, for plotting the actions that were landed for those observations or seconds that had more than one sequence of actions land by X1.

library(ggplot2)  
  
ggplot(data = x1\_2plus\_lands\_counts, aes(x=actionReaction, y=SecondsLastRoundAction, fill=Notes)) +  
 geom\_bar(stat='identity', position=position\_dodge())+  
 scale\_y\_continuous(breaks = seq(0, 8, by=1), limits=c(0,8))+  
 scale\_fill\_brewer(palette='Paired') +   
 ggtitle('Landed Actions by X1 versus 3 Opponents')+  
 ylab('Seconds Since Last Action by X1')+  
 xlab('Action Landed in Second with 2 Sequential Landed')



The above bar chart shows that the combo used by Nunez was a jab and cross for the double hits landed on Rousey and Tate, but not on Pennington.

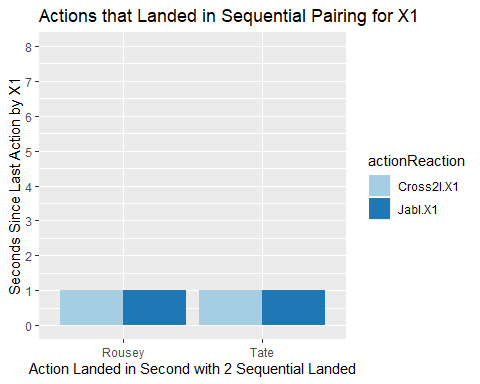
X1\_2plus\_lands\_X1[,c(2,6,7:34)]

## SecondsLastRoundAction TotLandsX1 Notes Crossl.X1 Kneel.X1 Elbowl.X1  
## 46 1 2 Tate 0 0 0  
## 160 1 2 Rousey 0 0 0  
## Hookl.X1 Jabl.X1 Kickl.X1 upperl.X1 takedownl.X1 hammerl.X1 Cross2l.X1  
## 46 0 1 0 0 0 0 1  
## 160 0 1 0 0 0 0 1  
## Knee2l.X1 Elbow2l.X1 Hook2l.X1 Jab2l.X1 Kick2l.X1 upper2l.X1 takedown2l.X1  
## 46 0 0 0 0 0 0 0  
## 160 0 0 0 0 0 0 0  
## hammer2l.X1 Cross3l.X1 Knee3l.X1 Elbow3l.X1 Hook3l.X1 Jab3l.X1 Kick3l.X1  
## 46 0 0 0 0 0 0 0  
## 160 0 0 0 0 0 0 0  
## upper3l.X1 takedown3l.X1 hammer3l.X1  
## 46 0 0 0  
## 160 0 0 0

plot2plus <- x1\_2plus\_lands\_counts # originally REGEX error x1\_2plus\_lands\_counts$Notes != 'Pennington')  
plot2plus

## SecondsIntoRound SecondsLastRoundAction cmTotHitsR.X1 cmTotHitsL.X1  
## 9 164 1 2 10  
## 10 44 1 0 13  
## 19 164 1 2 10  
## 20 44 1 0 13  
## cmTotHitsM.X1 TotLandsX1 Notes actionReaction actionCounts  
## 9 30 2 Tate Jabl.X1 1  
## 10 22 2 Rousey Jabl.X1 1  
## 19 30 2 Tate Cross2l.X1 1  
## 20 22 2 Rousey Cross2l.X1 1

ggplot(data = plot2plus, aes(x=Notes, y=SecondsLastRoundAction, fill=actionReaction)) +  
 geom\_bar(stat='identity', position=position\_dodge())+  
 scale\_y\_continuous(breaks = seq(0, 8, by=1), limits=c(0,8))+  
 scale\_fill\_brewer(palette='Paired') +   
 ggtitle('Actions that Landed in Sequential Pairing for X1')+  
 ylab('Seconds Since Last Action by X1')+  
 xlab('Action Landed in Second with 2 Sequential Landed')



Keep in mind the opponents for this data collection were Pennington, Rousey, and Tate against Nunez (Vulfen) but that the only second observations with more than one hit landed by Nunez were the observations with Rousey or Tate, and each of these rounds were less than two minutes out of a five minute first round with Nunez. The combo favored for quick knockouts or technical knockouts by Nunez is the jab and cross (or strong right hand punch to face). We could rearrange the visual to show the cumulative hits missed with all three opponents for the first minute to see if Nunez gives it away that she plans on fighting all five of the five minute rounds or wants to damage the opponent.

Why don’t we do that and put some skills to use.

oneMinute <- subset(Vulfen1, (Vulfen1$SecondsIntoRound < 60) & (Vulfen1$TotLandsX1 > 0 | (Vulfen1$TotMissedX1 > 0)))  
dim(oneMinute)

## [1] 36 182

head(colnames(oneMinute),20)

## [1] "Round" "SecondsIntoRound"   
## [3] "lastAction" "SecondsLastRoundAction"   
## [5] "cmTotHitsR.X1" "cmTotHitsL.X1"   
## [7] "cmTotHitsM.X1" "TotLandsX1"   
## [9] "TotMissedX1" "TotReceivedX1"   
## [11] "cmTotHitsR.X2" "cmTotHitsL.X2"   
## [13] "cmTotHitsM.X2" "TotLandsX2"   
## [15] "TotMissedX2" "TotReceivedX2"   
## [17] "Time" "FighterActionReactions.X1"   
## [19] "FightersActionsReactions.X2" "Notes"

Lets use the SecondsIntoRound, SecondsLastRoundAction, cmTotHits columns for X1, TotLandsX1, and cumulative actions of X2 as well as the Notes which is the opponent name, and also the sequential action columns.

Minute <- oneMinute[,c(2,4:13,20:182)]  
Minute\_tidy <- gather(Minute, 'action','actionCount',13:174)  
Minute\_tidy1 <- subset(Minute\_tidy, Minute\_tidy$actionCount > 0)  
Minute\_tidy1

## SecondsIntoRound SecondsLastRoundAction cmTotHitsR.X1 cmTotHitsL.X1  
## 20 18 2 0 3  
## 27 30 1 0 7  
## 32 39 1 0 10  
## 34 43 2 0 11  
## 165 22 1 0 4  
## 167 24 1 0 5  
## 170 29 2 0 6  
## 174 35 1 0 8  
## 179 44 1 0 13  
## 180 45 1 0 14  
## 341 11 1 0 2  
## 355 36 1 0 9  
## 359 44 1 0 13  
## 484 10 1 0 1  
## 1951 34 12 0 0  
## 1957 6 1 0 0  
## 1958 8 2 0 0  
## 1960 10 1 0 1  
## 1962 14 3 0 2  
## 1968 25 1 0 5  
## 1969 27 2 0 5  
## 1977 41 1 0 10  
## 2089 16 3 0 0  
## 2092 55 7 0 0  
## 2099 59 14 0 0  
## 2100 5 5 0 0  
## 2105 11 1 0 2  
## 2110 23 1 0 4  
## 2116 31 1 0 7  
## 2119 36 1 0 9  
## 2126 18 2 0 0  
## 2127 38 1 0 0  
## 2129 59 4 0 0  
## 2130 22 22 0 0  
## 2132 37 3 0 0  
## 2133 40 3 0 0  
## 2134 45 5 0 0  
## 2139 9 1 0 0  
## 2143 16 2 0 2  
## 2189 34 3 0 7  
## 2289 22 1 0 4  
## 2290 23 1 0 4  
## 2296 31 1 0 7  
## 2298 35 1 0 8  
## 2430 14 3 0 2  
## 2508 25 1 0 5  
## 2610 14 3 0 2  
## 2932 10 1 0 1  
## 2943 30 1 0 7  
## 3073 6 1 0 0  
## 3086 29 2 0 6  
## 3090 35 1 0 8  
## 3113 11 1 0 2  
## 3116 18 2 0 3  
## 3404 18 2 0 3  
## 3412 31 1 0 7  
## 4880 18 2 0 3  
## 4887 30 1 0 7  
## 4892 39 1 0 10  
## 4894 43 2 0 11  
## 5025 22 1 0 4  
## 5027 24 1 0 5  
## 5030 29 2 0 6  
## 5034 35 1 0 8  
## 5039 44 1 0 13  
## 5040 45 1 0 14  
## 5201 11 1 0 2  
## 5215 36 1 0 9  
## 5219 44 1 0 13  
## 5344 10 1 0 1  
## cmTotHitsM.X1 TotLandsX1 TotMissedX1 TotReceivedX1 cmTotHitsR.X2  
## 20 10 1 0 0 3  
## 27 16 1 0 0 7  
## 32 21 1 0 0 10  
## 34 22 1 0 0 11  
## 165 11 1 1 0 4  
## 167 13 1 0 0 5  
## 170 16 1 0 0 6  
## 174 20 1 1 0 8  
## 179 22 2 0 0 13  
## 180 22 1 0 0 14  
## 341 6 1 1 0 2  
## 355 21 1 1 0 9  
## 359 22 2 0 0 13  
## 484 5 1 1 0 1  
## 1951 2 0 1 0 0  
## 1957 2 0 1 0 0  
## 1958 3 0 1 0 0  
## 1960 5 1 1 0 1  
## 1962 9 0 3 0 2  
## 1968 15 0 2 0 5  
## 1969 16 0 1 0 5  
## 1977 22 0 1 0 10  
## 2089 1 0 1 0 0  
## 2092 4 0 1 0 0  
## 2099 6 0 1 0 0  
## 2100 1 0 1 0 0  
## 2105 6 1 1 0 2  
## 2110 13 0 2 0 4  
## 2116 18 0 2 0 7  
## 2119 21 1 1 0 9  
## 2126 2 0 1 0 0  
## 2127 3 0 1 0 0  
## 2129 5 0 1 0 0  
## 2130 1 0 1 0 0  
## 2132 3 0 1 0 0  
## 2133 4 0 1 0 0  
## 2134 5 0 1 0 0  
## 2139 4 0 1 0 0  
## 2143 10 0 1 0 2  
## 2189 19 0 1 0 7  
## 2289 11 1 1 0 4  
## 2290 13 0 2 0 4  
## 2296 18 0 2 0 7  
## 2298 20 1 1 0 8  
## 2430 9 0 3 0 2  
## 2508 15 0 2 0 5  
## 2610 9 0 3 0 2  
## 2932 5 1 1 0 1  
## 2943 16 1 0 0 7  
## 3073 2 0 1 0 0  
## 3086 16 1 0 0 6  
## 3090 20 1 1 0 8  
## 3113 6 1 1 0 2  
## 3116 10 1 0 0 3  
## 3404 10 1 0 0 3  
## 3412 18 0 2 0 7  
## 4880 10 1 0 0 3  
## 4887 16 1 0 0 7  
## 4892 21 1 0 0 10  
## 4894 22 1 0 0 11  
## 5025 11 1 1 0 4  
## 5027 13 1 0 0 5  
## 5030 16 1 0 0 6  
## 5034 20 1 1 0 8  
## 5039 22 2 0 0 13  
## 5040 22 1 0 0 14  
## 5201 6 1 1 0 2  
## 5215 21 1 1 0 9  
## 5219 22 2 0 0 13  
## 5344 5 1 1 0 1  
## cmTotHitsL.X2 cmTotHitsM.X2 Notes action actionCount  
## 20 0 5 Rousey Crossl.X1 1  
## 27 0 8 Rousey Crossl.X1 1  
## 32 0 10 Rousey Crossl.X1 1  
## 34 0 10 Rousey Crossl.X1 1  
## 165 0 6 Rousey Jabl.X1 1  
## 167 0 6 Rousey Jabl.X1 1  
## 170 0 7 Rousey Jabl.X1 1  
## 174 0 10 Rousey Jabl.X1 1  
## 179 0 10 Rousey Jabl.X1 1  
## 180 0 10 Rousey Jabl.X1 1  
## 341 0 3 Rousey Cross2l.X1 1  
## 355 0 10 Rousey Cross2l.X1 1  
## 359 0 10 Rousey Cross2l.X1 1  
## 484 0 2 Rousey Jab2l.X1 1  
## 1951 0 0 Pennington Crossm.X1 1  
## 1957 0 1 Rousey Crossm.X1 1  
## 1958 0 1 Rousey Crossm.X1 1  
## 1960 0 2 Rousey Crossm.X1 1  
## 1962 0 3 Rousey Crossm.X1 1  
## 1968 0 6 Rousey Crossm.X1 1  
## 1969 0 6 Rousey Crossm.X1 1  
## 1977 0 10 Rousey Crossm.X1 1  
## 2089 0 2 Tate Jabm.X1 1  
## 2092 0 5 Tate Jabm.X1 1  
## 2099 0 0 Pennington Jabm.X1 1  
## 2100 0 0 Rousey Jabm.X1 1  
## 2105 0 3 Rousey Jabm.X1 1  
## 2110 0 6 Rousey Jabm.X1 1  
## 2116 0 9 Rousey Jabm.X1 1  
## 2119 0 10 Rousey Jabm.X1 1  
## 2126 0 2 Tate Kickm.X1 1  
## 2127 0 3 Tate Kickm.X1 1  
## 2129 0 5 Tate Kickm.X1 1  
## 2130 0 0 Pennington Kickm.X1 1  
## 2132 0 0 Pennington Kickm.X1 1  
## 2133 0 0 Pennington Kickm.X1 1  
## 2134 0 0 Pennington Kickm.X1 1  
## 2139 0 1 Rousey Kickm.X1 1  
## 2143 0 3 Rousey Kickm.X1 1  
## 2189 0 9 Rousey upperm.X1 1  
## 2289 0 6 Rousey Cross2m.X1 1  
## 2290 0 6 Rousey Cross2m.X1 1  
## 2296 0 9 Rousey Cross2m.X1 1  
## 2298 0 10 Rousey Cross2m.X1 1  
## 2430 0 3 Rousey Jab2m.X1 1  
## 2508 0 6 Rousey upper2m.X1 1  
## 2610 0 3 Rousey Cross3m.X1 1  
## 2932 0 2 Rousey Crossm.X2 1  
## 2943 0 8 Rousey Crossm.X2 1  
## 3073 0 1 Rousey Jabm.X2 1  
## 3086 0 7 Rousey Jabm.X2 1  
## 3090 0 10 Rousey Jabm.X2 1  
## 3113 0 3 Rousey Kickm.X2 1  
## 3116 0 5 Rousey Kickm.X2 1  
## 3404 0 5 Rousey Jab2m.X2 1  
## 3412 0 9 Rousey Jab2m.X2 1  
## 4880 0 5 Rousey Crossr.X2 1  
## 4887 0 8 Rousey Crossr.X2 1  
## 4892 0 10 Rousey Crossr.X2 1  
## 4894 0 10 Rousey Crossr.X2 1  
## 5025 0 6 Rousey Jabr.X2 1  
## 5027 0 6 Rousey Jabr.X2 1  
## 5030 0 7 Rousey Jabr.X2 1  
## 5034 0 10 Rousey Jabr.X2 1  
## 5039 0 10 Rousey Jabr.X2 1  
## 5040 0 10 Rousey Jabr.X2 1  
## 5201 0 3 Rousey Cross2r.X2 1  
## 5215 0 10 Rousey Cross2r.X2 1  
## 5219 0 10 Rousey Cross2r.X2 1  
## 5344 0 2 Rousey Jab2r.X2 1

We have a subset of data above that is when an action as an attempt was made by either X1 or X2 for observations less than one minute into the round. The opponent name is in the Notes column. Cumulative hits landed and missed for both are available as well as the seconds that have passed since the last action in the round.

Lets look at the action to group by it using dplyr and get a count of the action that is attempted most by X1 only. This means we should remove the X2 actions.Remove any field with ‘X2’ in the name as it won’t be needed for now.

library(dplyr)  
  
X2actions <- grep('X2', colnames(Minute))  
Minute2 <- Minute[,-(X2actions)]  
  
minute2\_tidy <- gather(Minute2, 'action','actionCount',10:90)  
  
minute2\_tidy1 <- subset(minute2\_tidy, minute2\_tidy$actionCount > 0)  
  
actionsGrouped <- minute2\_tidy1 %>% group\_by(action) %>% count()  
mostAction <- actionsGrouped[order(actionsGrouped$n,decreasing=TRUE),]  
mostAction

## # A tibble: 12 x 2  
## # Groups: action [12]  
## action n  
## <chr> <int>  
## 1 Kickm.X1 9  
## 2 Crossm.X1 8  
## 3 Jabm.X1 8  
## 4 Jabl.X1 6  
## 5 Cross2m.X1 4  
## 6 Crossl.X1 4  
## 7 Cross2l.X1 3  
## 8 Cross3m.X1 1  
## 9 Jab2l.X1 1  
## 10 Jab2m.X1 1  
## 11 upper2m.X1 1  
## 12 upperm.X1 1

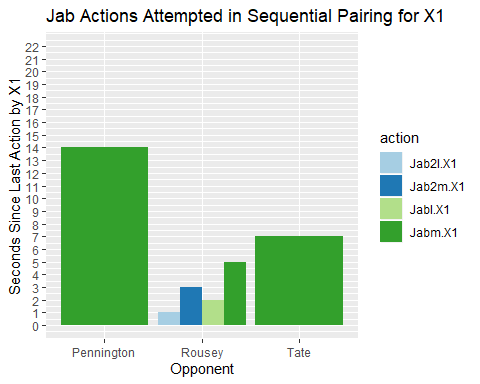
We can see the most attempted action is the kick but missed, then the cross that is missed, then the jab missed. But following those attempted actions that were missed, the most landed action by X1 is the jab, then the cross in the 2nd sequence or combo order. But the Cross is also a first attempted action. The landed hits most taken by X1 are the jab and cross from these samples.

Lets look at the jabs only first then the cross only. We will look at the seconds since last action, the opponent as Notes, and cumulative hits missed and landed by X1. But not in the same bar chart because that is beyond the three dimension capability.

jab <- minute2\_tidy1[grep('Jab', minute2\_tidy1$action),]  
jab

## SecondsIntoRound SecondsLastRoundAction cmTotHitsR.X1 cmTotHitsL.X1  
## 165 22 1 0 4  
## 167 24 1 0 5  
## 170 29 2 0 6  
## 174 35 1 0 8  
## 179 44 1 0 13  
## 180 45 1 0 14  
## 484 10 1 0 1  
## 1117 16 3 0 0  
## 1120 55 7 0 0  
## 1127 59 14 0 0  
## 1128 5 5 0 0  
## 1133 11 1 0 2  
## 1138 23 1 0 4  
## 1144 31 1 0 7  
## 1147 36 1 0 9  
## 1458 14 3 0 2  
## cmTotHitsM.X1 TotLandsX1 TotMissedX1 TotReceivedX1 Notes action  
## 165 11 1 1 0 Rousey Jabl.X1  
## 167 13 1 0 0 Rousey Jabl.X1  
## 170 16 1 0 0 Rousey Jabl.X1  
## 174 20 1 1 0 Rousey Jabl.X1  
## 179 22 2 0 0 Rousey Jabl.X1  
## 180 22 1 0 0 Rousey Jabl.X1  
## 484 5 1 1 0 Rousey Jab2l.X1  
## 1117 1 0 1 0 Tate Jabm.X1  
## 1120 4 0 1 0 Tate Jabm.X1  
## 1127 6 0 1 0 Pennington Jabm.X1  
## 1128 1 0 1 0 Rousey Jabm.X1  
## 1133 6 1 1 0 Rousey Jabm.X1  
## 1138 13 0 2 0 Rousey Jabm.X1  
## 1144 18 0 2 0 Rousey Jabm.X1  
## 1147 21 1 1 0 Rousey Jabm.X1  
## 1458 9 0 3 0 Rousey Jab2m.X1  
## actionCount  
## 165 1  
## 167 1  
## 170 1  
## 174 1  
## 179 1  
## 180 1  
## 484 1  
## 1117 1  
## 1120 1  
## 1127 1  
## 1128 1  
## 1133 1  
## 1138 1  
## 1144 1  
## 1147 1  
## 1458 1

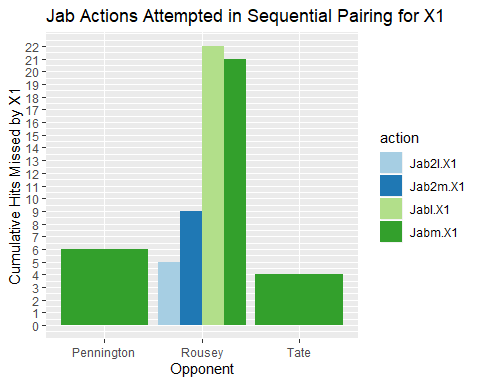
ggplot(data = jab, aes(x=Notes, y=SecondsLastRoundAction, fill=action)) +  
 geom\_bar(stat='identity', position=position\_dodge())+  
 scale\_y\_continuous(breaks = seq(0, 22, by=1), limits=c(0,22))+  
 scale\_fill\_brewer(palette='Paired') +   
 ggtitle('Jab Actions Attempted in Sequential Pairing for X1')+  
 ylab('Seconds Since Last Action by X1')+  
 xlab('Opponent')



Take note that from the outcomes of these fights and after looking at the plot above, Pennington went to decision and lasted all five rounds of five minutes each and has more seconds between actions attempted by Nunez looking at the most landed hit and attempted hit by Nunez, the jab. The Rousey fight had less seconds between actions attempted by Nunez and lasted under one minute with an outcome of technical knockout. The Tate fight is in the middle and lasted under five minutes of the first round.

Next, the jabs attempted and missed by Nunez on opponents for cumulative hits missed for Nunez at the time she makes the attempted hit with her most used strike landed, the jab.

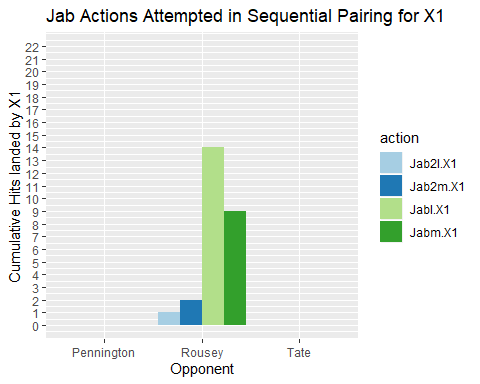
ggplot(data = jab, aes(x=Notes, y=cmTotHitsM.X1, fill=action)) +  
 geom\_bar(stat='identity', position=position\_dodge())+  
 scale\_y\_continuous(breaks = seq(0, 22, by=1), limits=c(0,22))+  
 scale\_fill\_brewer(palette='Paired') +   
 ggtitle('Jab Actions Attempted in Sequential Pairing for X1')+  
 ylab('Cumulative Hits Missed by X1')+  
 xlab('Opponent')



Nunez has more cumulative hits missed, which means she is taking more risks to get the fight over and done with when looking at her short lived fighting event with Rousey as you can see in the bar chart above.

Now, for the cumulative hits landed by Nunez against her opponent at each sequential second that she attempt a strike against her opponent.

ggplot(data = jab, aes(x=Notes, y=cmTotHitsL.X1, fill=action)) +  
 geom\_bar(stat='identity', position=position\_dodge())+  
 scale\_y\_continuous(breaks = seq(0, 22, by=1), limits=c(0,22))+  
 scale\_fill\_brewer(palette='Paired') +   
 ggtitle('Jab Actions Attempted in Sequential Pairing for X1')+  
 ylab('Cumulative Hits landed by X1')+  
 xlab('Opponent')



As we can see from the plot above, for the first minute of her fight with the three opponents, she has made more attempted hits that landed by way of the jab against Rousey, having had 14 cumulative jabs landed on Rousey in under one minute.

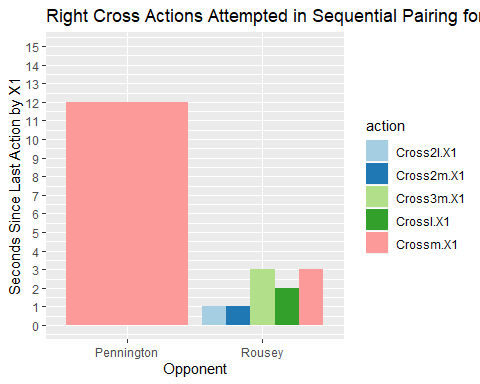
Now lets look at the power strike, Nunez’s cross, on her opponent, which is her 2nd most attempted strike against her opponent.

cross <- minute2\_tidy1[grep('Cross', minute2\_tidy1$action),]  
cross

## SecondsIntoRound SecondsLastRoundAction cmTotHitsR.X1 cmTotHitsL.X1  
## 20 18 2 0 3  
## 27 30 1 0 7  
## 32 39 1 0 10  
## 34 43 2 0 11  
## 341 11 1 0 2  
## 355 36 1 0 9  
## 359 44 1 0 13  
## 979 34 12 0 0  
## 985 6 1 0 0  
## 986 8 2 0 0  
## 988 10 1 0 1  
## 990 14 3 0 2  
## 996 25 1 0 5  
## 997 27 2 0 5  
## 1005 41 1 0 10  
## 1317 22 1 0 4  
## 1318 23 1 0 4  
## 1324 31 1 0 7  
## 1326 35 1 0 8  
## 1638 14 3 0 2  
## cmTotHitsM.X1 TotLandsX1 TotMissedX1 TotReceivedX1 Notes action  
## 20 10 1 0 0 Rousey Crossl.X1  
## 27 16 1 0 0 Rousey Crossl.X1  
## 32 21 1 0 0 Rousey Crossl.X1  
## 34 22 1 0 0 Rousey Crossl.X1  
## 341 6 1 1 0 Rousey Cross2l.X1  
## 355 21 1 1 0 Rousey Cross2l.X1  
## 359 22 2 0 0 Rousey Cross2l.X1  
## 979 2 0 1 0 Pennington Crossm.X1  
## 985 2 0 1 0 Rousey Crossm.X1  
## 986 3 0 1 0 Rousey Crossm.X1  
## 988 5 1 1 0 Rousey Crossm.X1  
## 990 9 0 3 0 Rousey Crossm.X1  
## 996 15 0 2 0 Rousey Crossm.X1  
## 997 16 0 1 0 Rousey Crossm.X1  
## 1005 22 0 1 0 Rousey Crossm.X1  
## 1317 11 1 1 0 Rousey Cross2m.X1  
## 1318 13 0 2 0 Rousey Cross2m.X1  
## 1324 18 0 2 0 Rousey Cross2m.X1  
## 1326 20 1 1 0 Rousey Cross2m.X1  
## 1638 9 0 3 0 Rousey Cross3m.X1  
## actionCount  
## 20 1  
## 27 1  
## 32 1  
## 34 1  
## 341 1  
## 355 1  
## 359 1  
## 979 1  
## 985 1  
## 986 1  
## 988 1  
## 990 1  
## 996 1  
## 997 1  
## 1005 1  
## 1317 1  
## 1318 1  
## 1324 1  
## 1326 1  
## 1638 1

The above table shows the cross actions attempted by X1.

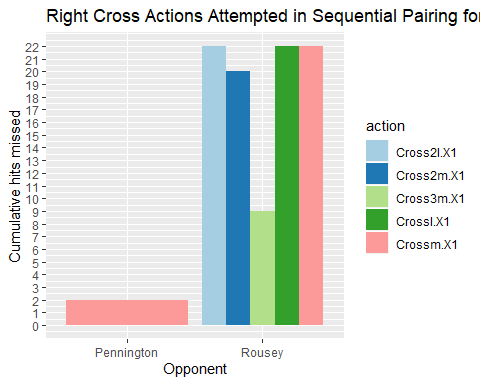
ggplot(data = cross, aes(x=Notes, y=SecondsLastRoundAction, fill=action)) +  
 geom\_bar(stat='identity', position=position\_dodge())+  
 scale\_y\_continuous(breaks = seq(0, 15, by=1), limits=c(0,15))+  
 scale\_fill\_brewer(palette='Paired') +   
 ggtitle('Right Cross Actions Attempted in Sequential Pairing for X1')+  
 ylab('Seconds Since Last Action by X1')+  
 xlab('Opponent')



The above bar chart shows that Tate is not in the list of opponents she attempted to hit with her right cross, but Rousey got all attempts in the 1st, 2nd, and 3rd sequences of at least one of the one-second observations under one minute of the first round. The seconds since the last action by Nunez were more frequent in the first round than with Pennington at 12 seconds since last action.

Now, look at the cumulative hits missed for x1-Nunez and cross attempts on her opponents.

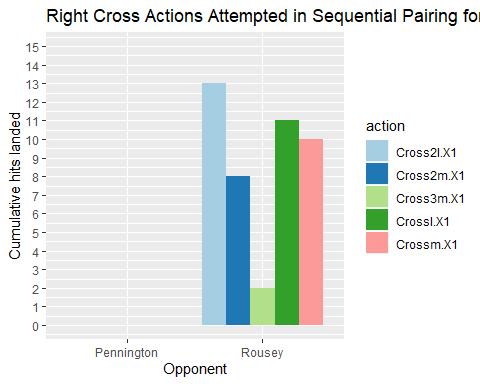
ggplot(data = cross, aes(x=Notes, y=cmTotHitsM.X1, fill=action)) +  
 geom\_bar(stat='identity', position=position\_dodge())+  
 scale\_y\_continuous(breaks = seq(0, 22, by=1), limits=c(0,22))+  
 scale\_fill\_brewer(palette='Paired') +   
 ggtitle('Right Cross Actions Attempted in Sequential Pairing for X1')+  
 ylab('Cumulative hits missed')+  
 xlab('Opponent')



Looking at the bar chart above for the amount of hits missed up to the point that X1 attempts a right cross on her opponent, Rousey has a much larger number of hits missed than Pennington.

Now, look at the cumulative hits landed for X1’s right cross on her opponents during the first minute of the first round.

ggplot(data = cross, aes(x=Notes, y=cmTotHitsL.X1, fill=action)) +  
 geom\_bar(stat='identity', position=position\_dodge())+  
 scale\_y\_continuous(breaks = seq(0, 15, by=1), limits=c(0,15))+  
 scale\_fill\_brewer(palette='Paired') +   
 ggtitle('Right Cross Actions Attempted in Sequential Pairing for X1')+  
 ylab('Cumulative hits landed')+  
 xlab('Opponent')



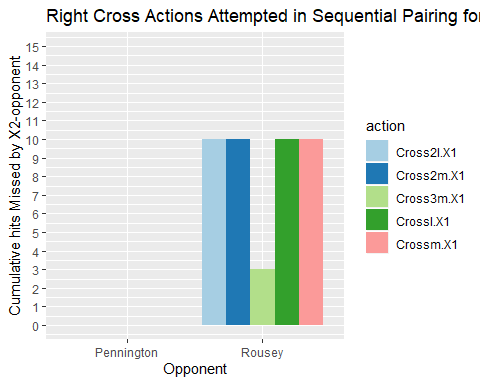
As we can see from the above bar chart, X1 had more right cross strikes against Rousey in the 2nd sequence of a combo and first sequence. But even the 3rd sequence had a hit when she had 2 hits landed. The fact that it is higher for her when using it as the 2nd strike of a combo when she had 13 hits landed, and as the first of a combo when she had 11 hits landed, shows she wanted to win, and most likely wasn’t deterred by her opponent. To assert that last assumption, we would have to look at the cumulative hits missed and/or landed by her opponent to confirm X1 wasn’t deterred when making her selection of strikes to take down her opponent.

X2actions <- grep('X2', colnames(Minute))  
X2act <- X2actions[4:84]  
Minute3 <- Minute[,-(X2act)]  
  
minute3\_tidy <- gather(Minute3, 'action','actionCount',13:93)  
  
minute3\_tidy1 <- subset(minute3\_tidy, minute3\_tidy$actionCount > 0)

cross1 <- minute3\_tidy1[grep('Cross', minute3\_tidy1$action),]  
cross1

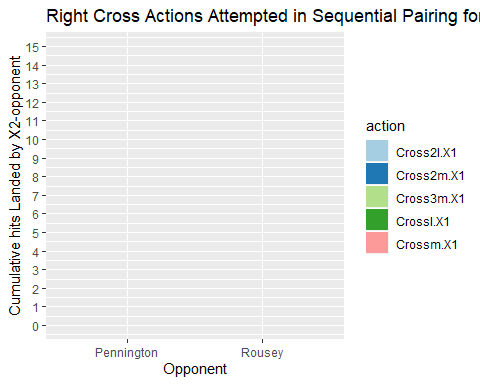
## SecondsIntoRound SecondsLastRoundAction cmTotHitsR.X1 cmTotHitsL.X1  
## 20 18 2 0 3  
## 27 30 1 0 7  
## 32 39 1 0 10  
## 34 43 2 0 11  
## 341 11 1 0 2  
## 355 36 1 0 9  
## 359 44 1 0 13  
## 979 34 12 0 0  
## 985 6 1 0 0  
## 986 8 2 0 0  
## 988 10 1 0 1  
## 990 14 3 0 2  
## 996 25 1 0 5  
## 997 27 2 0 5  
## 1005 41 1 0 10  
## 1317 22 1 0 4  
## 1318 23 1 0 4  
## 1324 31 1 0 7  
## 1326 35 1 0 8  
## 1638 14 3 0 2  
## cmTotHitsM.X1 TotLandsX1 TotMissedX1 TotReceivedX1 cmTotHitsR.X2  
## 20 10 1 0 0 3  
## 27 16 1 0 0 7  
## 32 21 1 0 0 10  
## 34 22 1 0 0 11  
## 341 6 1 1 0 2  
## 355 21 1 1 0 9  
## 359 22 2 0 0 13  
## 979 2 0 1 0 0  
## 985 2 0 1 0 0  
## 986 3 0 1 0 0  
## 988 5 1 1 0 1  
## 990 9 0 3 0 2  
## 996 15 0 2 0 5  
## 997 16 0 1 0 5  
## 1005 22 0 1 0 10  
## 1317 11 1 1 0 4  
## 1318 13 0 2 0 4  
## 1324 18 0 2 0 7  
## 1326 20 1 1 0 8  
## 1638 9 0 3 0 2  
## cmTotHitsL.X2 cmTotHitsM.X2 Notes action actionCount  
## 20 0 5 Rousey Crossl.X1 1  
## 27 0 8 Rousey Crossl.X1 1  
## 32 0 10 Rousey Crossl.X1 1  
## 34 0 10 Rousey Crossl.X1 1  
## 341 0 3 Rousey Cross2l.X1 1  
## 355 0 10 Rousey Cross2l.X1 1  
## 359 0 10 Rousey Cross2l.X1 1  
## 979 0 0 Pennington Crossm.X1 1  
## 985 0 1 Rousey Crossm.X1 1  
## 986 0 1 Rousey Crossm.X1 1  
## 988 0 2 Rousey Crossm.X1 1  
## 990 0 3 Rousey Crossm.X1 1  
## 996 0 6 Rousey Crossm.X1 1  
## 997 0 6 Rousey Crossm.X1 1  
## 1005 0 10 Rousey Crossm.X1 1  
## 1317 0 6 Rousey Cross2m.X1 1  
## 1318 0 6 Rousey Cross2m.X1 1  
## 1324 0 9 Rousey Cross2m.X1 1  
## 1326 0 10 Rousey Cross2m.X1 1  
## 1638 0 3 Rousey Cross3m.X1 1

ggplot(data = cross1, aes(x=Notes, y=cmTotHitsM.X2, fill=action)) +  
 geom\_bar(stat='identity', position=position\_dodge())+  
 scale\_y\_continuous(breaks = seq(0, 15, by=1), limits=c(0,15))+  
 scale\_fill\_brewer(palette='Paired') +   
 ggtitle('Right Cross Actions Attempted in Sequential Pairing for X1')+  
 ylab('Cumulative hits Missed by X2-opponent')+  
 xlab('Opponent')



The above chart shows that Rousey as X2 had up to 10 hits cumulatively missed at the moment that Nunez or X1 attempted a right cross in every three sequence of an observational second in the round 1 of this fight event.

ggplot(data = cross1, aes(x=Notes, y=cmTotHitsL.X2, fill=action)) +  
 geom\_bar(stat='identity', position=position\_dodge())+  
 scale\_y\_continuous(breaks = seq(0, 15, by=1), limits=c(0,15))+  
 scale\_fill\_brewer(palette='Paired') +   
 ggtitle('Right Cross Actions Attempted in Sequential Pairing for X1')+  
 ylab('Cumulative hits Landed by X2-opponent')+  
 xlab('Opponent')



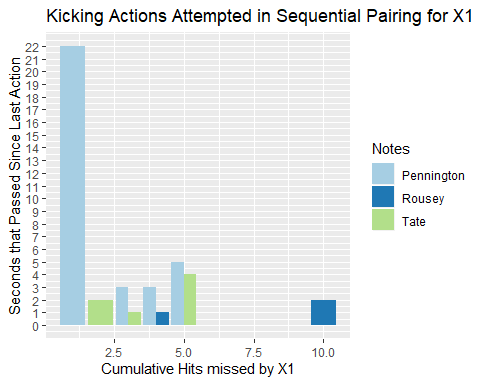
The above bar chart shows that The cumulative hits landed by the opponent or X2 were zero for the actions attempted by Nunez using her 2nd most attempted strike, the right cross.

Now, for more curiosity and in answering some questions about the kick, which is Nunez’s most attempted strike against her opponent in under one minute for these three opponents’ first round fight with said fighter. She didn’t land any kicks, but when the fight is just sizing up or opening up the opponent, she used the front kick at times or the muay thai kick to disable the opponent’s anchor or leverage against her.

kick <- minute3\_tidy1[grep('Kick', minute3\_tidy1$action),]  
kick

## SecondsIntoRound SecondsLastRoundAction cmTotHitsR.X1 cmTotHitsL.X1  
## 1154 18 2 0 0  
## 1155 38 1 0 0  
## 1157 59 4 0 0  
## 1158 22 22 0 0  
## 1160 37 3 0 0  
## 1161 40 3 0 0  
## 1162 45 5 0 0  
## 1167 9 1 0 0  
## 1171 16 2 0 2  
## cmTotHitsM.X1 TotLandsX1 TotMissedX1 TotReceivedX1 cmTotHitsR.X2  
## 1154 2 0 1 0 0  
## 1155 3 0 1 0 0  
## 1157 5 0 1 0 0  
## 1158 1 0 1 0 0  
## 1160 3 0 1 0 0  
## 1161 4 0 1 0 0  
## 1162 5 0 1 0 0  
## 1167 4 0 1 0 0  
## 1171 10 0 1 0 2  
## cmTotHitsL.X2 cmTotHitsM.X2 Notes action actionCount  
## 1154 0 2 Tate Kickm.X1 1  
## 1155 0 3 Tate Kickm.X1 1  
## 1157 0 5 Tate Kickm.X1 1  
## 1158 0 0 Pennington Kickm.X1 1  
## 1160 0 0 Pennington Kickm.X1 1  
## 1161 0 0 Pennington Kickm.X1 1  
## 1162 0 0 Pennington Kickm.X1 1  
## 1167 0 1 Rousey Kickm.X1 1  
## 1171 0 3 Rousey Kickm.X1 1

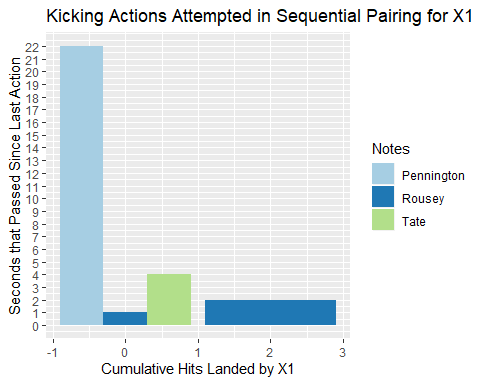
ggplot(data = kick, aes(x=cmTotHitsM.X1, y=SecondsLastRoundAction, fill=Notes)) +  
 geom\_bar(stat='identity', position=position\_dodge())+  
 scale\_y\_continuous(breaks = seq(0, 22, by=1), limits=c(0,22))+  
 scale\_fill\_brewer(palette='Paired') +   
 ggtitle('Kicking Actions Attempted in Sequential Pairing for X1')+  
 xlab('Cumulative Hits missed by X1')+  
 ylab('Seconds that Passed Since Last Action')



The above bar chart is a different arrangement, but shows that the time since last action and cumulative hits missed up to the point that X1-Nunez attempt a kick, and all kicks in the first minute were missed, against her opponent. We can see that after 22 seconds she made a kicking attempt with Pennington, and made more kick attempts against Pennington that lasted the whole event of 5-five minute rounds and a win for Nunez through a decision by the judges. Her attempts to kick Rousey in under a minute are actually less than Pennington and Tate. I recall that Rousey didn’t shake Nunez’s hand before the fight, and this could mean it was a grudge fight where Nunez looks to knock out her opponent with her upper body strikes than her lower body strikes.

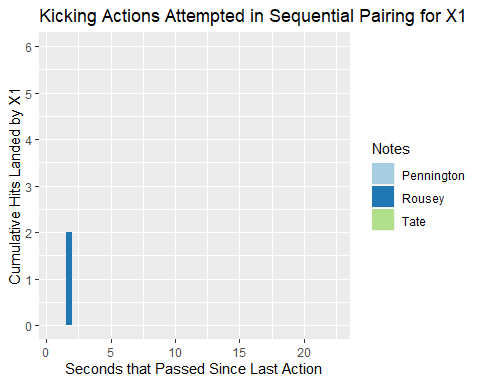
To confirm her upper body strikes are X1’s go to in ending a fight early, look at her cumulative hits landed in comparison to the seconds that passed since the last attempted action by X1.

ggplot(data = kick, aes(x=cmTotHitsL.X1, y=SecondsLastRoundAction, fill=Notes)) +  
 geom\_bar(stat='identity', position=position\_dodge())+  
 scale\_y\_continuous(breaks = seq(0, 22, by=1), limits=c(0,22))+  
 scale\_fill\_brewer(palette='Paired') +   
 ggtitle('Kicking Actions Attempted in Sequential Pairing for X1')+  
 xlab('Cumulative Hits Landed by X1')+  
 ylab('Seconds that Passed Since Last Action')

 The above bar chart isn’t making sense as it shows negative values for cumulative hits landed by X1. And the values are discreet, so the bars should be exactly on an integer of the x-axis but as we see above are not.

Here is a better chart by switching the x and y labels.

ggplot(data = kick, aes(y=cmTotHitsL.X1, x=SecondsLastRoundAction, fill=Notes)) +  
 geom\_bar(stat='identity', position=position\_dodge())+  
 scale\_y\_continuous(breaks = seq(0, 6, by=1), limits=c(0,6))+  
 scale\_fill\_brewer(palette='Paired') +   
 ggtitle('Kicking Actions Attempted in Sequential Pairing for X1')+  
 ylab('Cumulative Hits Landed by X1')+  
 xlab('Seconds that Passed Since Last Action')



Lets look at the data table and see if Rousey is the only one receiving attempted kicks.

kick2 <- kick[,c(2,4,12:14)]  
kick2

## SecondsLastRoundAction cmTotHitsL.X1 Notes action actionCount  
## 1154 2 0 Tate Kickm.X1 1  
## 1155 1 0 Tate Kickm.X1 1  
## 1157 4 0 Tate Kickm.X1 1  
## 1158 22 0 Pennington Kickm.X1 1  
## 1160 3 0 Pennington Kickm.X1 1  
## 1161 3 0 Pennington Kickm.X1 1  
## 1162 5 0 Pennington Kickm.X1 1  
## 1167 1 0 Rousey Kickm.X1 1  
## 1171 2 2 Rousey Kickm.X1 1

By looking at the chart above, only Rousey was the opponent that received any hits that landed when Nunez attempted a kick as a strike. Nunez made kick attempts at the other two opponents, but she hadn’t landed any hits by the time she attempted to kick those other opponents. Nunez, did however miss a number of strikes attempted on her other opponents when attempting a kick.

This was an interesting analysis of the fighter Nunez, than also compared results from Mazvidal earlier using machine learning models: random forest, recursive partitioning trees or decision trees, generalized linear models like liner and logistic regression, k-nearest neighber, and generalized boosted machines also a tree algorithm. The best predictors were the tree algorithms in using the samples of Mazvidal as the testing set for Nunez’s model built only on her sampled second observations in her first round. The comparison with the best algorithm scored 86% meaning Mazvidal is 86% similar in fighting style to Nunez.

The other comparisons were for analyzing and using visualizations to understand Nunez’s method in fighting. She uses more jabs and crosses, and leaves more seconds between actions of fights she wants to play out for longer as in with Pennington. If she wants to end the fight early or uses a grudge fight method, her style changes, hits become more frequent, and this is based on the attempted hits her opponent makes towards her, as Rousey made more attempts that coincidentally missed, while the seconds between fight actions on the part of Nunez on Rousey were much smaller.

If we were to use this understanding and see how her next opponent acts or behaves towards Nunez and other factors before the fight, we could determine how long the fight is predicted to last, but this was only the first round of three fights, when she has fought many fights and had many fights last more than one round. Measuring her exhaustion as a new feature, if the fight is a grudge fight as another feature, if the opponent is a dominating or aggressive bat out of the cave type as another feature, and so on.