**STT 811**

**In-Class Assignment 11**

This problem will use the OJ dataset. Note that you will use Purchase as the target (no need to convert to 0/1)

1. Split the data into training and test datasets (with a 75/25 split).

oj <- ISLR2::OJ

split\_pct <- 0.75

n <- length(oj$Purchase)\*split\_pct # train size

row\_samp <- sample(1:length(oj$Purchase), n, replace = FALSE)

train <- oj[row\_samp,]

test <- oj[-row\_samp,]

1. Build a KNN model for your target based on PriceDiff and LoyalCH, varying K. How does the accuracy compare on the test dataset? Which K works best here?

train.Y = train$Purchase

test.Y = test$Purchase

train\_scale = scale(train[,c(2,3)])

test\_scale = scale(test[,c(2,3)])

knn\_mod<-knn(train = train\_scale, test = test\_scale, cl = train.Y, k=5)

cm<-confusionMatrix(knn\_mod, reference = as.factor(test.Y))

cm$table