Player Development Associate Position

## R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

# Depicting data is essential for data scientist to understand what the data is explaining. Throughout this assignment I show outcomes of pitches in relation to a generic strike zone and the result of each ball in play on the field.

# Here I load the data set off my GitHub into the Global Evironment of R

url <- 'https://raw.githubusercontent.com/Chrisboatto/Player-Development-Data/main/TrackMan%20Data.csv?token=AKUDE7BOIJEBF3JX27NDIWLAG2BGU'  
  
MetsPlayerData <- read.csv(url)

# These packages are needed to complete this assignment

require("dplyr")

## Loading required package: dplyr

##   
## Attaching package: 'dplyr'

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

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

require("ggplot2")

## Loading required package: ggplot2

require("devtools")

## Loading required package: devtools

## Loading required package: usethis

library(dplyr)  
library(ggplot2)  
library(devtools)

# First, I check both the structure and summary of the data set to gain a preliminary understanding of the data. The data consists of 85 attributes and 300 observations. I noticed that the gameId, date, and stadium, HomeTeam and AwayTeam have exactly the same results within their respective columns. Therefore I can make the conclusion that this data is from a single home game for the Mets against the Phillies.

str(MetsPlayerData)

## 'data.frame': 300 obs. of 85 variables:  
## $ pitchid : int 1 2 3 4 5 6 7 8 9 10 ...  
## $ date : chr "2020-09-06" "2020-09-06" "2020-09-06" "2020-09-06" ...  
## $ time : chr "13:10:54" "13:11:08" "13:11:51" "13:12:12" ...  
## $ pa\_number : int 1 1 2 2 2 2 3 3 3 4 ...  
## $ pa\_pitch\_number : int 1 2 1 2 3 4 1 2 3 1 ...  
## $ pitcher : chr "deGrom, Jacob" "deGrom, Jacob" "deGrom, Jacob" "deGrom, Jacob" ...  
## $ pitcherid : int 594798 594798 594798 594798 594798 594798 594798 594798 594798 594798 ...  
## $ throws : chr "Right" "Right" "Right" "Right" ...  
## $ field\_team : chr "New York Mets" "New York Mets" "New York Mets" "New York Mets" ...  
## $ batter : chr "McCutchen, Andrew" "McCutchen, Andrew" "Hoskins, Rhys" "Hoskins, Rhys" ...  
## $ batterid : int 457705 457705 656555 656555 656555 656555 547180 547180 547180 544369 ...  
## $ batside : chr "Right" "Right" "Right" "Right" ...  
## $ bat\_team : chr "Philadelphia Phillies" "Philadelphia Phillies" "Philadelphia Phillies" "Philadelphia Phillies" ...  
## $ pitcher\_set : logi NA NA NA NA NA NA ...  
## $ inning : int 1 1 1 1 1 1 1 1 1 1 ...  
## $ inning\_half : chr "Top" "Top" "Top" "Top" ...  
## $ outs : int 0 0 1 1 1 1 2 2 2 2 ...  
## $ balls : int 0 1 0 0 0 1 0 0 0 0 ...  
## $ strikes : int 0 0 0 1 2 2 0 1 2 0 ...  
## $ pitch\_type\_tagged : chr "Undefined" "Undefined" "Undefined" "Undefined" ...  
## $ pitch\_type\_auto : chr "Fastball" "Fastball" "Fastball" "Fastball" ...  
## $ pitch\_call : chr "BallCalled" "InPlay" "StrikeSwinging" "FoulBall" ...  
## $ k\_or\_bb : chr "Undefined" "Undefined" "Undefined" "Undefined" ...  
## $ hit\_type : chr "Undefined" "Undefined" "Undefined" "Undefined" ...  
## $ play\_result : chr "Undefined" "Out" "Undefined" "Undefined" ...  
## $ outs\_on\_play : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ runs\_scored : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ notes : chr "statcast" "statcast" "statcast" "statcast" ...  
## $ rel\_speed : num 97.9 97.6 99.1 99.2 99.8 ...  
## $ vert\_rel\_angle : num -3.16 -2.711 -2.582 -1.431 -0.652 ...  
## $ horz\_rel\_angle : num -3.68 -2.28 -2.36 -2.9 -3.34 ...  
## $ spin\_rate : num 2451 2417 2505 2543 2564 ...  
## $ spin\_axis : num 210 208 216 221 205 ...  
## $ tilt : chr "NULL" "NULL" "NULL" "NULL" ...  
## $ rel\_height : num 5.51 5.43 5.43 5.46 5.45 ...  
## $ rel\_side : num 1.33 1.3 1.27 1.27 1.09 ...  
## $ extension : num 6.82 7.07 6.87 6.87 6.93 ...  
## $ vert\_break : num -10.7 -12.1 -11 -11.7 -10 ...  
## $ induced\_vert\_break : num 17.6 16.4 16.7 16.1 17.4 ...  
## $ horz\_break : num 6.51 7.95 10.35 9.91 9.7 ...  
## $ plate\_loc\_z : num 1.73 1.96 2.16 3.18 4.02 ...  
## $ plate\_loc\_x : num -1.4866 -0.1131 -0.0194 -0.5506 -1.145 ...  
## $ zone\_speed : num 90.5 90 91.1 90.6 91.2 ...  
## $ vert\_appr\_angle : num -5.16 -4.98 -4.65 -3.64 -2.55 ...  
## $ horz\_appr\_angle : num -2.454 -0.781 -0.401 -1.018 -1.501 ...  
## $ zone\_time : num 0.383 0.384 0.379 0.379 0.377 ...  
## $ exit\_speed : chr "NULL" "88.6981" "NULL" "82.7058" ...  
## $ angle : chr "NULL" "53.705" "NULL" "57.217" ...  
## $ direction : chr "NULL" "17.7199" "NULL" "107.195" ...  
## $ hit\_spin\_rate : chr "NULL" "4691.83" "NULL" "NULL" ...  
## $ pos\_at\_100\_x : chr "NULL" "NULL" "NULL" "NULL" ...  
## $ pos\_at\_100\_y : chr "NULL" "NULL" "NULL" "NULL" ...  
## $ pos\_at\_100\_z : chr "NULL" "NULL" "NULL" "NULL" ...  
## $ distance : chr "NULL" "235.743" "NULL" "225.167" ...  
## $ last\_tracked\_distance: chr "NULL" "NULL" "NULL" "NULL" ...  
## $ bearing : chr "NULL" "32.9276" "NULL" "122.651" ...  
## $ hang\_time : chr "NULL" "6.08034" "NULL" "5.17727" ...  
## $ pfx\_x : num -4.46 -4.92 -6.29 -6.23 -6.23 ...  
## $ pfx\_z : num 10.46 9.67 9.86 9.23 9.72 ...  
## $ x0 : num -1.098 -1.164 -1.125 -1.086 -0.884 ...  
## $ y0 : num 50 50 50 50 50 ...  
## $ z0 : num 5.3 5.26 5.26 5.36 5.4 ...  
## $ vx0 : num 8.94 5.4 5.6 6.97 8.14 ...  
## $ vy0 : num -142 -142 -144 -144 -145 ...  
## $ vz0 : num -8.22 -7.13 -6.87 -3.99 -1.99 ...  
## $ ax0 : num -9.25 -10.15 -13.32 -13.15 -13.32 ...  
## $ ay0 : num 29 29.7 31.6 33.2 33.3 ...  
## $ az0 : num -10.5 -12.2 -11.3 -12.7 -11.4 ...  
## $ home\_team : int 121 121 121 121 121 121 121 121 121 121 ...  
## $ away\_team : int 143 143 143 143 143 143 143 143 143 143 ...  
## $ stadium : chr "CITI" "CITI" "CITI" "CITI" ...  
## $ levelid : chr "MLB" "MLB" "MLB" "MLB" ...  
## $ leagueid : chr "NL" "NL" "NL" "NL" ...  
## $ trackman\_gameid : chr "20200906-CITI-1" "20200906-CITI-1" "20200906-CITI-1" "20200906-CITI-1" ...  
## $ pitch\_uid : chr "355a9188-648d-491d-814d-5f575e722ff7" "e42681f1-bab9-498e-bc69-ae5ee1759ede" "3a57f821-c945-4d0b-a1f1-24ed302df3b3" "b0d8947b-75c8-422c-b621-5d13c5e8c6a2" ...  
## $ effective\_velo : chr "NULL" "NULL" "NULL" "NULL" ...  
## $ max\_height : chr "NULL" "NULL" "NULL" "NULL" ...  
## $ measured\_duration : chr "NULL" "NULL" "NULL" "NULL" ...  
## $ speed\_drop : chr "NULL" "NULL" "NULL" "NULL" ...  
## $ pitch\_last\_measured\_x: chr "NULL" "NULL" "NULL" "NULL" ...  
## $ pitch\_last\_measured\_y: chr "NULL" "NULL" "NULL" "NULL" ...  
## $ pitch\_last\_measured\_z: chr "NULL" "NULL" "NULL" "NULL" ...  
## $ contact\_position\_x : chr "NULL" "NULL" "NULL" "NULL" ...  
## $ contact\_position\_y : chr "NULL" "NULL" "NULL" "NULL" ...  
## $ contact\_position\_z : chr "NULL" "NULL" "NULL" "NULL" ...

summary(MetsPlayerData)

## pitchid date time pa\_number   
## Min. : 1.00 Length:300 Length:300 Min. : 1.000   
## 1st Qu.: 75.75 Class :character Class :character 1st Qu.: 2.000   
## Median :150.50 Mode :character Mode :character Median : 3.000   
## Mean :150.50 Mean : 3.193   
## 3rd Qu.:225.25 3rd Qu.: 4.000   
## Max. :300.00 Max. :10.000   
## pa\_pitch\_number pitcher pitcherid throws   
## Min. :1.00 Length:300 Min. :453284 Length:300   
## 1st Qu.:1.00 Class :character 1st Qu.:594798 Class :character   
## Median :3.00 Mode :character Median :594798 Mode :character   
## Mean :2.85 Mean :594556   
## 3rd Qu.:4.00 3rd Qu.:605400   
## Max. :8.00 Max. :661440   
## field\_team batter batterid batside   
## Length:300 Length:300 Min. :429664 Length:300   
## Class :character Class :character 1st Qu.:544369 Class :character   
## Mode :character Mode :character Median :624413 Mode :character   
## Mean :586533   
## 3rd Qu.:643446   
## Max. :665926   
## bat\_team pitcher\_set inning inning\_half   
## Length:300 Mode:logical Min. :1.0 Length:300   
## Class :character NA's:300 1st Qu.:3.0 Class :character   
## Mode :character Median :5.0 Mode :character   
## Mean :5.2   
## 3rd Qu.:7.0   
## Max. :9.0   
## outs balls strikes pitch\_type\_tagged   
## Min. :0.00 Min. :0.0000 Min. :0.0000 Length:300   
## 1st Qu.:0.00 1st Qu.:0.0000 1st Qu.:0.0000 Class :character   
## Median :1.00 Median :1.0000 Median :1.0000 Mode :character   
## Mean :0.96 Mean :0.8467 Mean :0.9067   
## 3rd Qu.:2.00 3rd Qu.:1.0000 3rd Qu.:2.0000   
## Max. :2.00 Max. :3.0000 Max. :2.0000   
## pitch\_type\_auto pitch\_call k\_or\_bb hit\_type   
## Length:300 Length:300 Length:300 Length:300   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
## play\_result outs\_on\_play runs\_scored notes   
## Length:300 Min. :0 Min. :0 Length:300   
## Class :character 1st Qu.:0 1st Qu.:0 Class :character   
## Mode :character Median :0 Median :0 Mode :character   
## Mean :0 Mean :0   
## 3rd Qu.:0 3rd Qu.:0   
## Max. :0 Max. :0   
## rel\_speed vert\_rel\_angle horz\_rel\_angle spin\_rate   
## Min. : 75.09 Min. :-5.146 Min. :-6.528 Min. : 815.8   
## 1st Qu.: 85.19 1st Qu.:-1.932 1st Qu.:-3.272 1st Qu.:1895.2   
## Median : 91.46 Median :-1.211 Median :-2.366 Median :2194.7   
## Mean : 90.08 Mean :-1.082 Mean :-2.187 Mean :2156.6   
## 3rd Qu.: 93.13 3rd Qu.:-0.235 3rd Qu.:-1.444 3rd Qu.:2513.5   
## Max. :100.82 Max. : 2.521 Max. : 4.651 Max. :3555.5   
## spin\_axis tilt rel\_height rel\_side   
## Min. : 31.29 Length:300 Min. :4.627 Min. :-2.212   
## 1st Qu.:158.65 Class :character 1st Qu.:4.904 1st Qu.: 1.119   
## Median :214.36 Mode :character Median :5.431 Median : 1.410   
## Mean :194.55 Mean :5.356 Mean : 1.462   
## 3rd Qu.:235.19 3rd Qu.:5.518 3rd Qu.: 1.895   
## Max. :330.73 Max. :6.445 Max. : 3.682   
## extension vert\_break induced\_vert\_break horz\_break   
## Min. :5.119 Min. :-61.014 Min. :-14.068 Min. :-18.501   
## 1st Qu.:6.512 1st Qu.:-33.311 1st Qu.: 3.110 1st Qu.: -3.834   
## Median :6.762 Median :-25.668 Median : 6.817 Median : 8.749   
## Mean :6.623 Mean :-26.645 Mean : 7.299 Mean : 5.294   
## 3rd Qu.:6.904 3rd Qu.:-17.763 3rd Qu.: 13.872 3rd Qu.: 13.499   
## Max. :7.515 Max. : -9.958 Max. : 19.252 Max. : 19.934   
## plate\_loc\_z plate\_loc\_x zone\_speed vert\_appr\_angle   
## Min. :-0.7011 Min. :-2.3939 Min. :70.11 Min. :-12.706   
## 1st Qu.: 1.6169 1st Qu.:-0.7959 1st Qu.:78.56 1st Qu.: -7.200   
## Median : 2.1616 Median :-0.1423 Median :84.51 Median : -5.949   
## Mean : 2.1454 Mean :-0.1037 Mean :83.22 Mean : -6.087   
## 3rd Qu.: 2.7674 3rd Qu.: 0.5455 3rd Qu.:86.23 3rd Qu.: -4.649   
## Max. : 4.6472 Max. : 2.2360 Max. :93.40 Max. : -2.147   
## horz\_appr\_angle zone\_time exit\_speed angle   
## Min. :-6.53282 Min. :0.3711 Length:300 Length:300   
## 1st Qu.:-2.41046 1st Qu.:0.4021 Class :character Class :character   
## Median :-1.04627 Median :0.4096 Mode :character Mode :character   
## Mean :-1.18598 Mean :0.4184   
## 3rd Qu.: 0.06968 3rd Qu.:0.4413   
## Max. : 3.12266 Max. :0.4999   
## direction hit\_spin\_rate pos\_at\_100\_x pos\_at\_100\_y   
## Length:300 Length:300 Length:300 Length:300   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
## pos\_at\_100\_z distance last\_tracked\_distance bearing   
## Length:300 Length:300 Length:300 Length:300   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
## hang\_time pfx\_x pfx\_z x0   
## Length:300 Min. :-12.197 Min. :-7.343 Min. :-3.5001   
## Class :character 1st Qu.: -8.292 1st Qu.: 1.827 1st Qu.:-1.7411   
## Mode :character Median : -5.238 Median : 3.842 Median :-1.2682   
## Mean : -3.457 Mean : 4.276 Mean :-1.3142   
## 3rd Qu.: 1.644 3rd Qu.: 8.001 3rd Qu.:-0.9768   
## Max. : 11.335 Max. :11.000 Max. : 1.9536   
## y0 z0 vx0 vy0   
## Min. :50.00 Min. :4.574 Min. :-9.746 Min. :-146.7   
## 1st Qu.:50.00 1st Qu.:4.907 1st Qu.: 3.425 1st Qu.:-135.5   
## Median :50.00 Median :5.303 Median : 5.153 Median :-133.1   
## Mean :50.00 Mean :5.271 Mean : 4.903 Mean :-131.1   
## 3rd Qu.:50.00 3rd Qu.:5.434 3rd Qu.: 7.364 3rd Qu.:-124.0   
## Max. :50.01 Max. :6.327 Max. :15.145 Max. :-109.4   
## vz0 ax0 ay0 az0   
## Min. :-12.393 Min. :-22.014 Min. :16.89 Min. :-41.35   
## 1st Qu.: -5.285 1st Qu.:-13.990 1st Qu.:23.43 1st Qu.:-29.35   
## Median : -3.597 Median :-10.393 Median :25.25 Median :-25.34   
## Mean : -3.358 Mean : -6.420 Mean :25.62 Mean :-23.84   
## 3rd Qu.: -1.391 3rd Qu.: 3.035 3rd Qu.:27.77 3rd Qu.:-17.01   
## Max. : 3.505 Max. : 17.305 Max. :33.31 Max. :-10.13   
## home\_team away\_team stadium levelid   
## Min. :121 Min. :143 Length:300 Length:300   
## 1st Qu.:121 1st Qu.:143 Class :character Class :character   
## Median :121 Median :143 Mode :character Mode :character   
## Mean :121 Mean :143   
## 3rd Qu.:121 3rd Qu.:143   
## Max. :121 Max. :143   
## leagueid trackman\_gameid pitch\_uid effective\_velo   
## Length:300 Length:300 Length:300 Length:300   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
## max\_height measured\_duration speed\_drop pitch\_last\_measured\_x  
## Length:300 Length:300 Length:300 Length:300   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
## pitch\_last\_measured\_y pitch\_last\_measured\_z contact\_position\_x  
## Length:300 Length:300 Length:300   
## Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character   
##   
##   
##   
## contact\_position\_y contact\_position\_z  
## Length:300 Length:300   
## Class :character Class :character   
## Mode :character Mode :character   
##   
##   
##

# For my first depiction I plan on gaining an uderstanding of the strike zone the umprire for the game had. In order to do this I needed to understand how many unique pitch outcomes there were. Below shows there are six outcomes within the dataset.

unique(MetsPlayerData$pitch\_call)

## [1] "BallCalled" "InPlay" "StrikeSwinging" "FoulBall"   
## [5] "StrikeCalled" "HitByPitch"

pitchCall <- table(MetsPlayerData$pitch\_call)  
pitchCall

##   
## BallCalled FoulBall HitByPitch InPlay StrikeCalled   
## 102 55 2 44 35   
## StrikeSwinging   
## 62

# I then began to clean the data. I removed all the columns that were not helpful in creating my depictions along with any column that had the value “NULL” throughout. These columns in my mind made the data expanded for no neccessary reason. By removing these I was able to read the data a lot easier without compromising the data itself.

# I then removed the HitByPitch result because it has no affect on depicting the strike zone itself. If a pitcher hits a batter it means he missed the zone by a wide margin thus does not help my desired result.

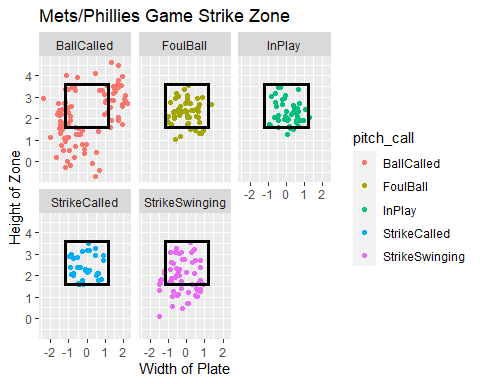
# Converting the columns positioned between 41:57 within the data set will help my offensive depiction. The attributes were set to ‘Character’ as one of their outputs was “NULL” result. When they were converted to ‘Numeric’ the “NULL” values changed to NA, thus the conversion performed successfully.

MetsPlayerData <- MetsPlayerData[, -c(2, 14, 20, 28, 34, 51:53, 55, 59, 69:73, 76:85)]  
MetsPlayerData <- filter(MetsPlayerData, pitch\_call != "HitByPitch")  
MetsPlayerData[41:57] <- lapply(MetsPlayerData[41:57], as.numeric)

## Warning in lapply(MetsPlayerData[41:57], as.numeric): NAs introduced by coercion  
  
## Warning in lapply(MetsPlayerData[41:57], as.numeric): NAs introduced by coercion  
  
## Warning in lapply(MetsPlayerData[41:57], as.numeric): NAs introduced by coercion  
  
## Warning in lapply(MetsPlayerData[41:57], as.numeric): NAs introduced by coercion  
  
## Warning in lapply(MetsPlayerData[41:57], as.numeric): NAs introduced by coercion  
  
## Warning in lapply(MetsPlayerData[41:57], as.numeric): NAs introduced by coercion  
  
## Warning in lapply(MetsPlayerData[41:57], as.numeric): NAs introduced by coercion

# The below depictions show the strike zone based off the result from each pitch thrown. Based on the depiction below you can make the assumption that the umpire had a fairly conservative strike zone as pitches placed in the zone were called balls yet none outside were called strikes. This allowed hitters to be more selective as you do not see many InPlay results outside of the strike zone. It is also interesting to see that no batters chased above the zone but did below.

ggplot(MetsPlayerData, aes(plate\_loc\_x, plate\_loc\_z, col= pitch\_call)) +   
 geom\_point(aes\_string(x="plate\_loc\_x", y="plate\_loc\_z"), size=1.4) +  
 geom\_rect(mapping = aes(ymax = 3.6, ymin = 1.6, xmax = -1.2, xmin = 1.2), alpha = 0, size=1.2, colour = "black") +  
 facet\_wrap(~pitch\_call) +   
 labs(title = "Mets/Phillies Game Strike Zone", x = "Width of Plate", y = "Height of Zone")

 # Below I created a chart to show how each pitcher in the game faired based on their velocity, spin rate, and break fro each pitch they threw. This chart would allow a person to scroll through and see what each pitchers’ metrics were on average throughout the game and if for some reason they deviated from their norm. If so, someone would be able to immediately notice the drop and begin the process of assessing the pitcher.

# I used the aggregate function muktiple times to determine the averages of each metric for each pitcher and pitch. I then bound the outputs together and removed the duplicate columns.

MetsPlayerDataAggVelo <- aggregate(rel\_speed ~ pitcher + pitch\_type\_auto, FUN = mean, data = MetsPlayerData)  
MetsPlayerDataAggSpin <- aggregate(spin\_rate ~ pitcher + pitch\_type\_auto, FUN = mean, data = MetsPlayerData)  
MetsPlayerDataAggVertBreak <- aggregate(vert\_break ~ pitcher + pitch\_type\_auto, FUN = mean, data = MetsPlayerData)  
MetsPlayerDataAggHorzBreak <- aggregate(horz\_break ~ pitcher + pitch\_type\_auto, FUN = mean, data = MetsPlayerData)  
MetsPlayerDataAgg <- cbind(MetsPlayerDataAggHorzBreak, MetsPlayerDataAggSpin, MetsPlayerDataAggVelo, MetsPlayerDataAggVertBreak)  
MetsPlayerDataAgg <-MetsPlayerDataAgg[, -c(4,5,7,8,10,11)]  
MetsPlayerDataAgg[order(-MetsPlayerDataAgg$rel\_speed),]

## pitcher pitch\_type\_auto horz\_break spin\_rate rel\_speed vert\_break  
## 11 deGrom, Jacob Fastball 7.358359 2411.5659 98.62832 -11.72864  
## 12 Hembree, Heath Fastball 12.112871 2510.9386 95.18256 -14.03439  
## 18 Llovera, Mauricio Sinker 11.283439 1941.5654 93.39660 -19.76320  
## 13 Nola, Aaron Fastball 13.550900 2265.1235 92.35852 -18.83951  
## 15 Shreve, Chasen Fastball -9.901249 2359.2050 92.17566 -15.40896  
## 21 deGrom, Jacob Slider -4.101522 2585.4400 92.06956 -27.22091  
## 17 Brach, Brad Sinker 9.403265 2122.6900 91.38180 -17.84685  
## 2 deGrom, Jacob Changeup 12.744641 1623.9423 91.25609 -28.55725  
## 19 Nola, Aaron Sinker 17.346791 2190.7904 91.24236 -24.96080  
## 14 Parker, Blake Fastball 3.184689 2065.4242 90.54210 -16.40197  
## 10 Brach, Brad Fastball 7.407770 2149.9000 90.20647 -19.73517  
## 9 Brach, Brad Cutter -3.953950 2310.5400 88.28380 -24.50470  
## 5 deGrom, Jacob Curveball -7.041470 2590.7400 88.27980 -31.99910  
## 3 Llovera, Mauricio Changeup 15.590440 1981.5473 85.64267 -32.17941  
## 4 Nola, Aaron Changeup 14.196950 1586.5206 84.28299 -36.47116  
## 25 Shreve, Chasen Splitter -15.000687 1184.5738 84.18220 -34.45239  
## 1 Brach, Brad Changeup 11.564400 1784.4050 84.15940 -37.96545  
## 22 Llovera, Mauricio Slider -11.604258 2366.9620 83.47654 -39.44884  
## 6 Hembree, Heath Curveball -9.759690 2861.1450 83.09335 -47.73445  
## 20 Brach, Brad Slider -6.506630 2494.8900 81.30620 -40.58440  
## 23 Shreve, Chasen Slider 6.414840 2360.8700 81.06460 -47.35070  
## 7 Llovera, Mauricio Curveball -8.663230 2148.9700 80.49140 -42.38510  
## 24 Parker, Blake Splitter 8.922557 917.2725 79.99977 -39.27142  
## 16 Nola, Aaron Other -14.702179 2461.4871 78.14725 -55.01999  
## 8 Parker, Blake Curveball -7.177614 2193.3480 75.87866 -58.96554

# I used the GeomMLBStadiums package found off Ben Dilday’s GitHub page to create the outline of Citi Field. Being that this is the home field of the New York Mets and where the game was played I thought it pertinent to this assignment to have Citi Field a part of this depiction.

devtools::install\_github("bdilday/GeomMLBStadiums")

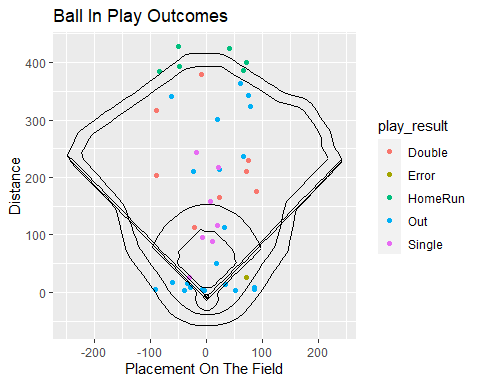
## Skipping install of 'GeomMLBStadiums' from a github remote, the SHA1 (68c11a99) has not changed since last install.  
## Use `force = TRUE` to force installation

library(GeomMLBStadiums)

# I filtered out any result that was not a ball in play along with any NA values. Having any outcome that was an NA or ‘Undefined’ would not help me show how each ball landed within the field of play. I also had to multiply the ‘bearing’ attribute by two as after some research into the Trackman data and the package. Comparing Trackman to the specifications of the package from GitHub I found that to apply the Trackman data to properly I would have to multiple the ‘width placement’ by two for optimal visualization.

MetsDataOffense <- filter(MetsPlayerData, !is.na(exit\_speed) & play\_result != "Undefined")

ggplot(MetsDataOffense, aes(x = bearing \* 2, y = distance, color = play\_result)) +   
 geom\_spraychart(stadium\_ids = "mets", stadium\_transform\_coords = TRUE, stadium\_segments = "all") +   
 labs(title = "Ball In Play Outcomes", x = "Placement On The Field", y = "Distance")

 # Below I created the chart to determine what batters on average hit the ball the hardest and where their tendencies were. Using the same method as I did to create the pitchers chart, I created the hitters chart. I then created another attribute to simplify the depiction of where each batter’s tendencies were when he hit the ball. Based off the summary of the ‘direction’ attribute I showed what field the batter tended to hit the ball that way the manager could prepare his defense properly the next time they play the opponent.

MetsDataOffenseAggEV <- aggregate(exit\_speed ~ batter, data = MetsDataOffense, FUN = mean)  
MetsDataOffenseAggDir <- aggregate(direction ~ batter, data = MetsDataOffense, FUN = mean)  
MetsDataOffenseAgg <- cbind(MetsDataOffenseAggEV, MetsDataOffenseAggDir)

summary(MetsDataOffense$direction)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -46.420 -12.912 5.920 5.962 25.469 61.359

MetsDataOffenseAgg <- MetsDataOffenseAgg[, -c(3)]  
  
MetsDataOffenseAgg <- mutate(MetsDataOffenseAgg, FieldHit = case\_when(direction < -13 ~ "LF", direction > -13 & direction < 13 ~ "CF", FieldHit = direction > 13 ~ "RF"))  
  
MetsDataOffenseAgg

## batter exit\_speed direction FieldHit  
## 1 Alonso, Pete 103.37230 -20.313367 LF  
## 2 Cano, Robinson 100.99667 26.723000 RF  
## 3 Conforto, Michael 98.66115 16.103352 RF  
## 4 Davis, J.D. 79.61680 -21.794940 LF  
## 5 Garlick, Kyle 89.07610 5.239150 CF  
## 6 Gimenez, Andres 76.79835 34.985735 RF  
## 7 Gregorius, Didi 83.74053 33.133367 RF  
## 8 Guillorme, Luis 90.33080 -32.202500 LF  
## 9 Harper, Bryce 94.50380 23.676280 RF  
## 10 Haseley, Adam 72.41390 47.540200 RF  
## 11 Hoskins, Rhys 93.02710 -19.685650 LF  
## 12 Knapp, Andrew 92.35420 -5.767935 CF  
## 13 McCutchen, Andrew 79.51593 -17.277267 LF  
## 14 McNeil, Jeff 92.61580 17.010600 RF  
## 15 Nimmo, Brandon 93.37193 -4.958873 CF  
## 16 Ramos, Wilson 103.87800 17.546200 RF  
## 17 Segura, Jean 95.06270 2.496850 CF  
## 18 Smith, Dominic 101.90865 -2.053670 CF  
## 19 Walker, Neil 88.30110 38.737400 RF