Tyler Beede Project

Nick Hanhan

5/26/2020

Tyler Beede Project

```
#open up baseballr
library(baseballr)
```

```
## Registered S3 method overwritten by 'xts':
##   method     from
##   as.zoo.xts zoo
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

```
library(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
```

```
library(tidyr)
#lookup player id for beede
playerid_lookup("Beede")
```

```
## Be patient, this may take a few seconds...
```

```
## Data courtesy of the Chadwick Bureau Register (https://github.com/chadwickbureau/register)
```

```
## # A tibble: 3 x 11
##   first_name last_name given_name name_suffix nick_name birth_year
##   <chr>      <chr>     <chr>      <lgl>       <chr>          <dbl>
## 1 George     Beede     George L.  NA          <NA>              NA
## 2 Tyler      Beede     Tyler Jos… NA          <NA>            1993
## 3 Walter     Beede     Walter A.  NA          <NA>            1963
## # … with 5 more variables: mlb_played_first <dbl>, mlbam_id <dbl>,
## #   retrosheet_id <chr>, bbref_id <chr>, fangraphs_id <dbl>
```

```
#create a dataframe called beede by scraping statcast data from 2019
beede <- scrape_statcast_savant(start_date = "2019-04-06", end_date = "2019-10-15", playerid = 595881, player_typ
e='pitcher')
```

```
## 2019-04-06 is not a date. Attempting to coerce...
```

```
## https://baseballsavant.mlb.com/statcast_search/csv?all=true&hfPT=&hfAB=&hfBBT=&hfPR=&hfZ=&stadium=&hfBBL=&hfNewZones=&hfGT=R%7CPO%7CS%7C&hfC&hfSea=2019%7C&hfSit=&hfOuts=&opponent=&pitcher_throws=&batter_stands=&hfSA=&player_type=pitcher&hfInfield=&team=&position=&hfOutfield=&hfRO=&home_road=&pitchers_lookup%5B%5D=595881&game_date_gt=2019-04-06&game_date_lt=2019-10-15&hfFlag=&hfPull=&metric_1=&hfInn=&min_pitches=0&min_results=0&group_by=name&sort_col=pitches&player_event_sort=h_launch_speed&sort_order=desc&min_abs=0&type=details
```

```
## These data are from BaseballSavant and are property of MLB Advanced Media, L.P. All rights reserved.
```

```
## Grabbing data, this may take a minute...
```

```
## URL read and payload acquired successfully.
```

```
## Warning in function_list[[k]](value): NAs introduced by coercion

## Warning in function_list[[k]](value): NAs introduced by coercion

## Warning in function_list[[k]](value): NAs introduced by coercion

## Warning in function_list[[k]](value): NAs introduced by coercion

## Warning in function_list[[k]](value): NAs introduced by coercion

## Warning in function_list[[k]](value): NAs introduced by coercion

## Warning in function_list[[k]](value): NAs introduced by coercion

## Warning in function_list[[k]](value): NAs introduced by coercion
```

```
beede <- beede %>% rename(px = plate_x)
beede <- beede %>% rename(pz = plate_z)
```

For this project, I used the baseballr package and look at Tyler Beede's 2019 season stats. I wanted to specifically look at Tyler Beede because I have had a few interactions with him while preparing mounds at Oracle while with the groundscrew. Thus, I wanted to take a deep look at why he shows signs of being a good pitcher, but also has rough stretches.

```
#lookup how many rows beede has

nrow(beede)
```

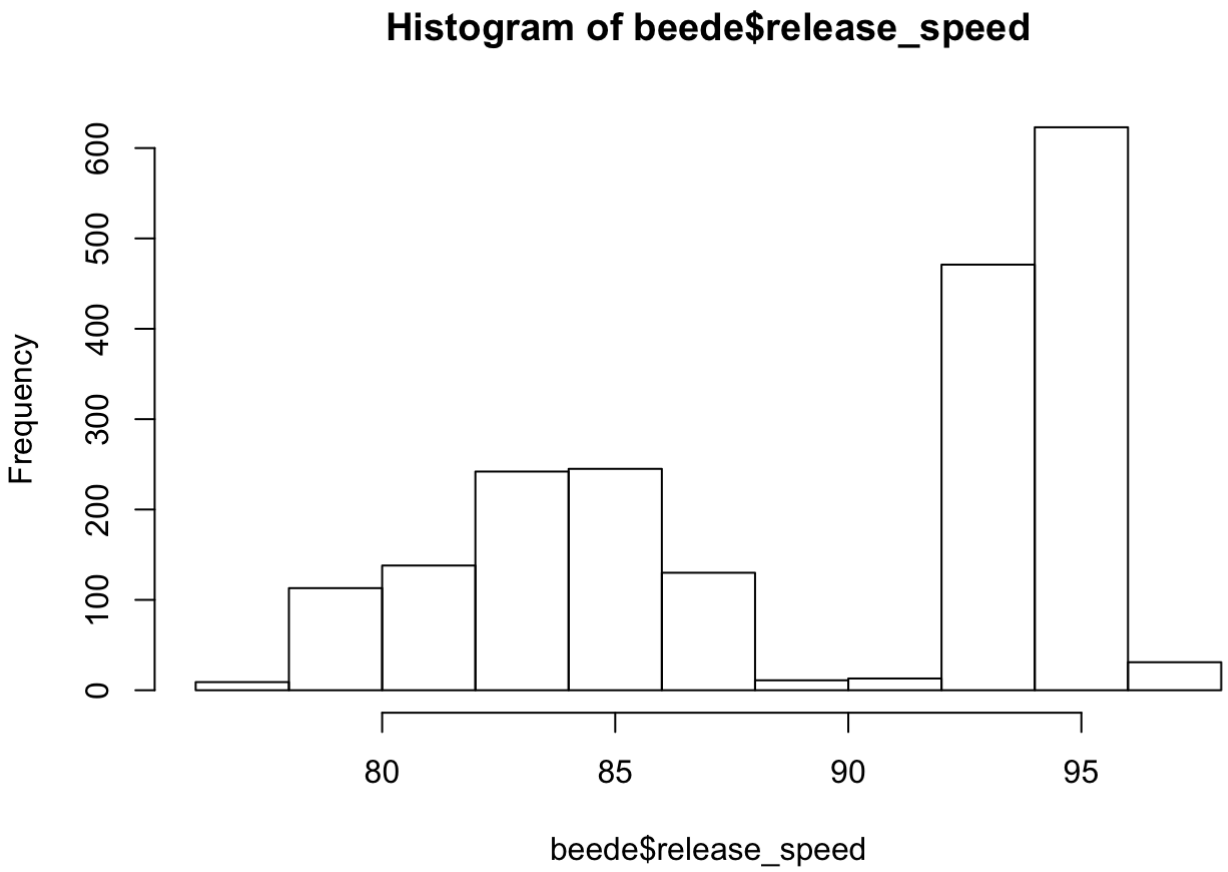
```
## [1] 2028
```

```
#subset dataframe to have only complete cases of pitch speed
beede <- subset(beede, !is.na(release_speed))
#make game data into a date column
beede$game_date <- as.Date(beede$game_date)
library(dplyr)
library(tidyr)
#seperate the date column into a year month and day column
beede <- separate(beede, game_date,
                  into = c("year", "month", "day"),
                  sep = "-", FALSE)
#make these columns numeric
beede$year <- as.numeric(beede$year)
beede$day <- as.numeric(beede$day)
beede$month <- as.numeric(beede$month)
#make a new column of beede in month july or other month
beede$july <- ifelse(beede$month == 7, "july", "other")
summary(factor(beede$july))
```

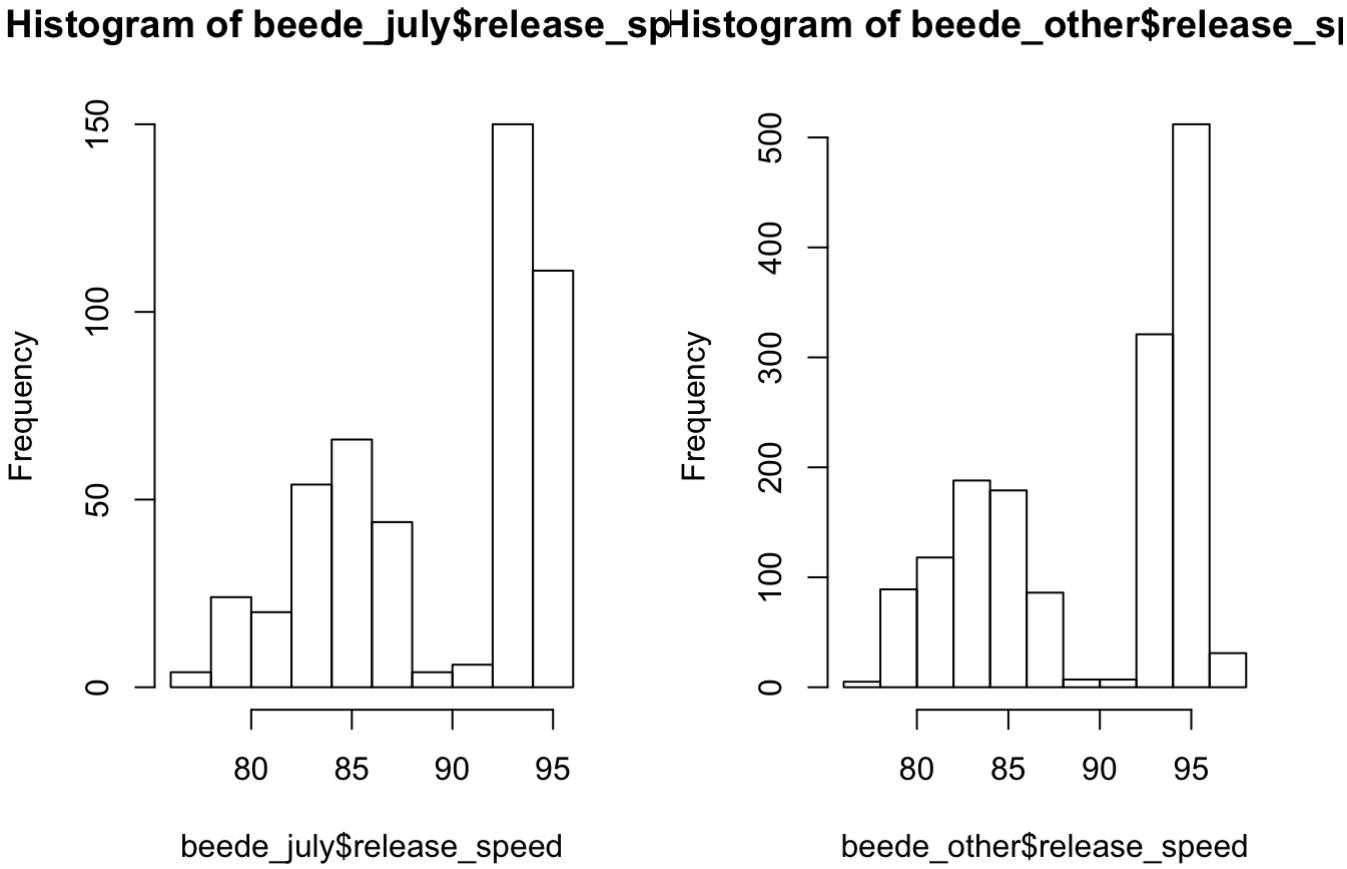
```
##  july other
##   483  1543
```

All we did here was get rid of rows that didn't have values. I also seperated the game dates into a month, day, and year column, while making them numeric values. Next line I made a new variable indicating if Beede pitched in July. I chose July because statistically he had the best month of the season going 2-2 with a 3.34 era in 5 starts.

```
#plot release speed for beede
hist(beede$release_speed)
```

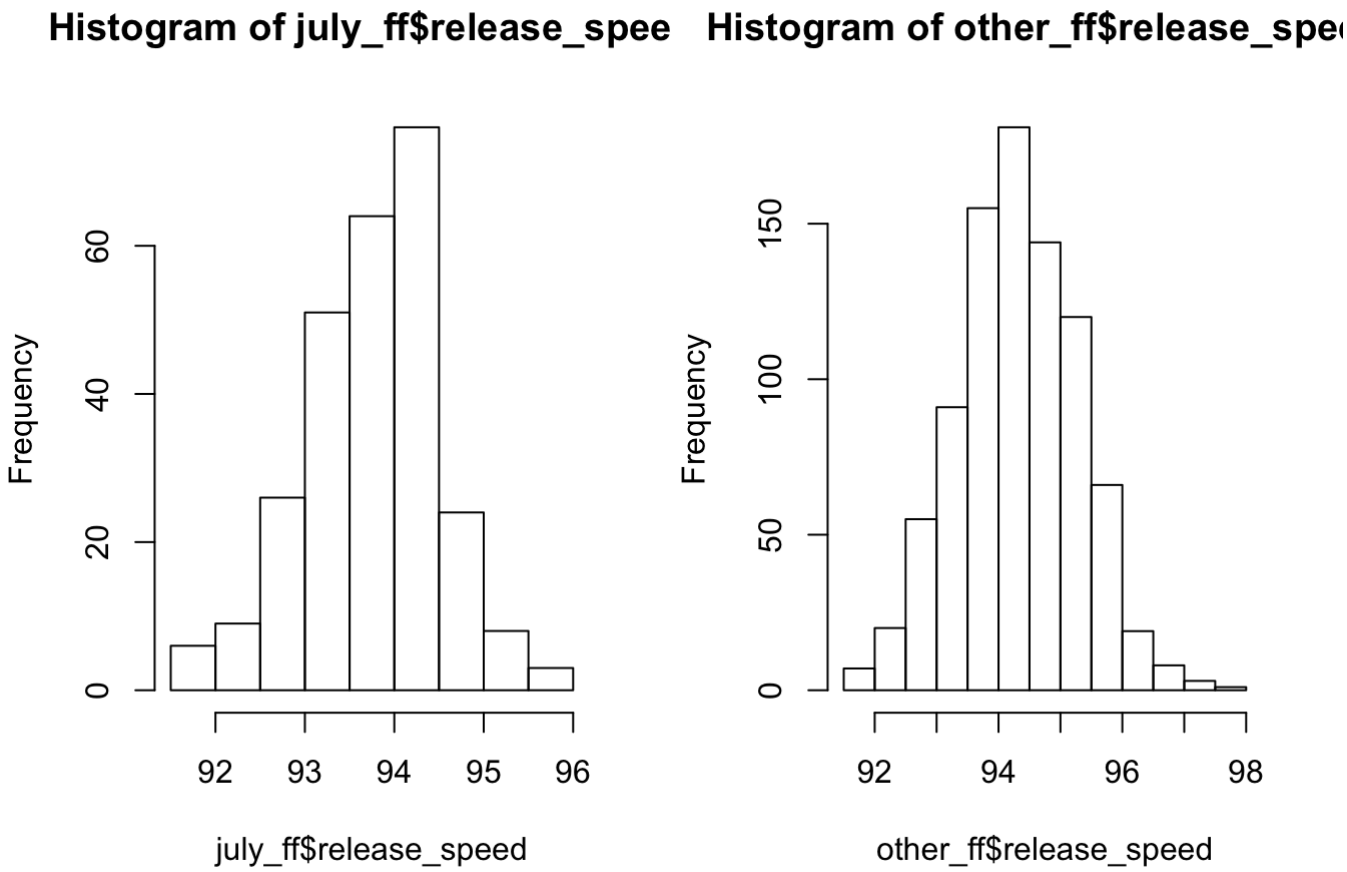



```
#create two dataframes one if beede is pitching in july or not
#Chose July because he pitched the best in July
beede_july <- subset(beede, beede$july=="july")
beede_other <- subset(beede, beede$july == "other")
par(mfrow = c(1, 2))
#plot both speeds, maybe there is some indicators for his performance
hist(beede_july$release_speed)
hist(beede_other$release_speed)
```



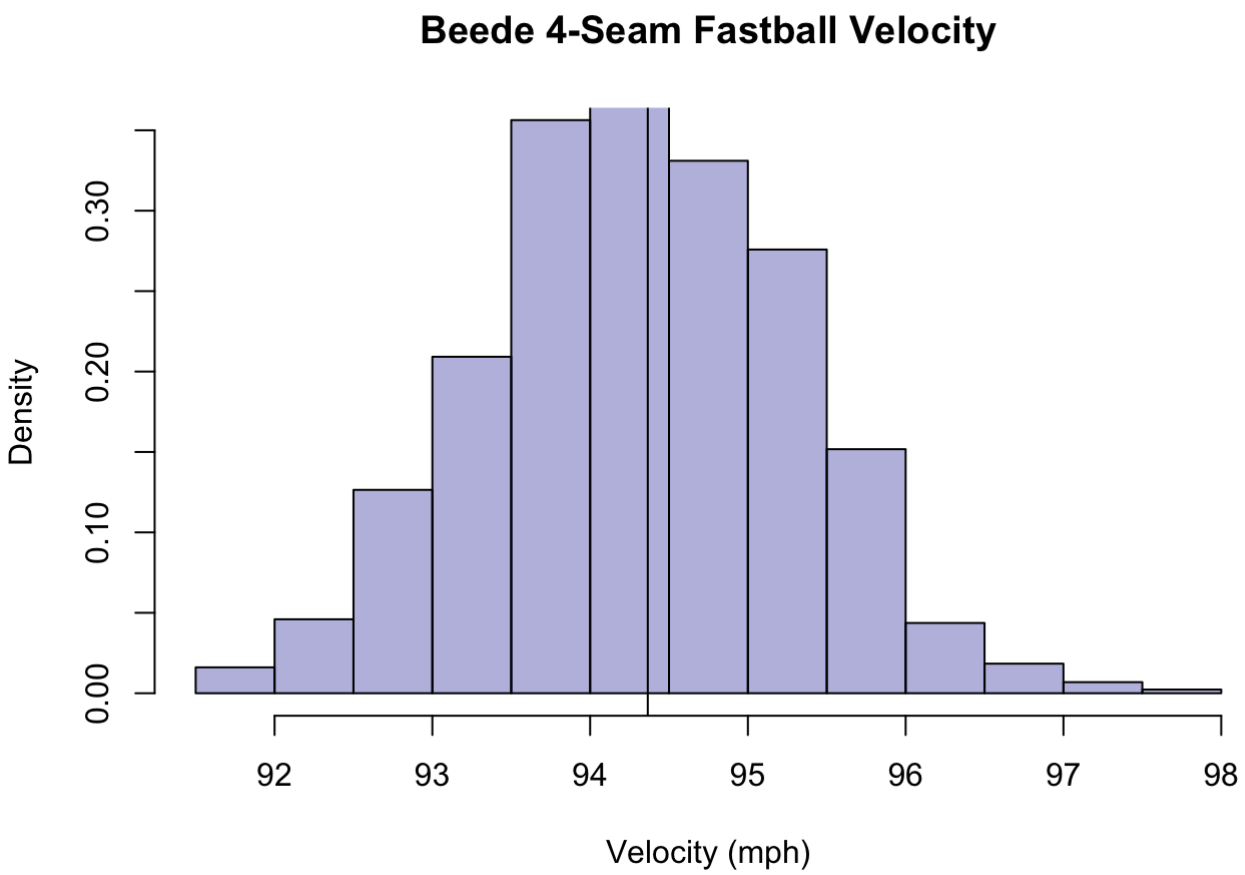
These are histograms of Beede's release speeds in July vs other months of 2019. Proportionally, looks like Beede threw with a lower mean velocity.

```
#new datasets looking at beede's fastballs for july and other months
july_ff <- subset(beede_july, pitch_type == "FF")
other_ff <- subset(beede_other, pitch_type == "FF")
par(mfrow=c(1,2))
hist(july_ff$release_speed)
hist(other_ff$release_speed)
```

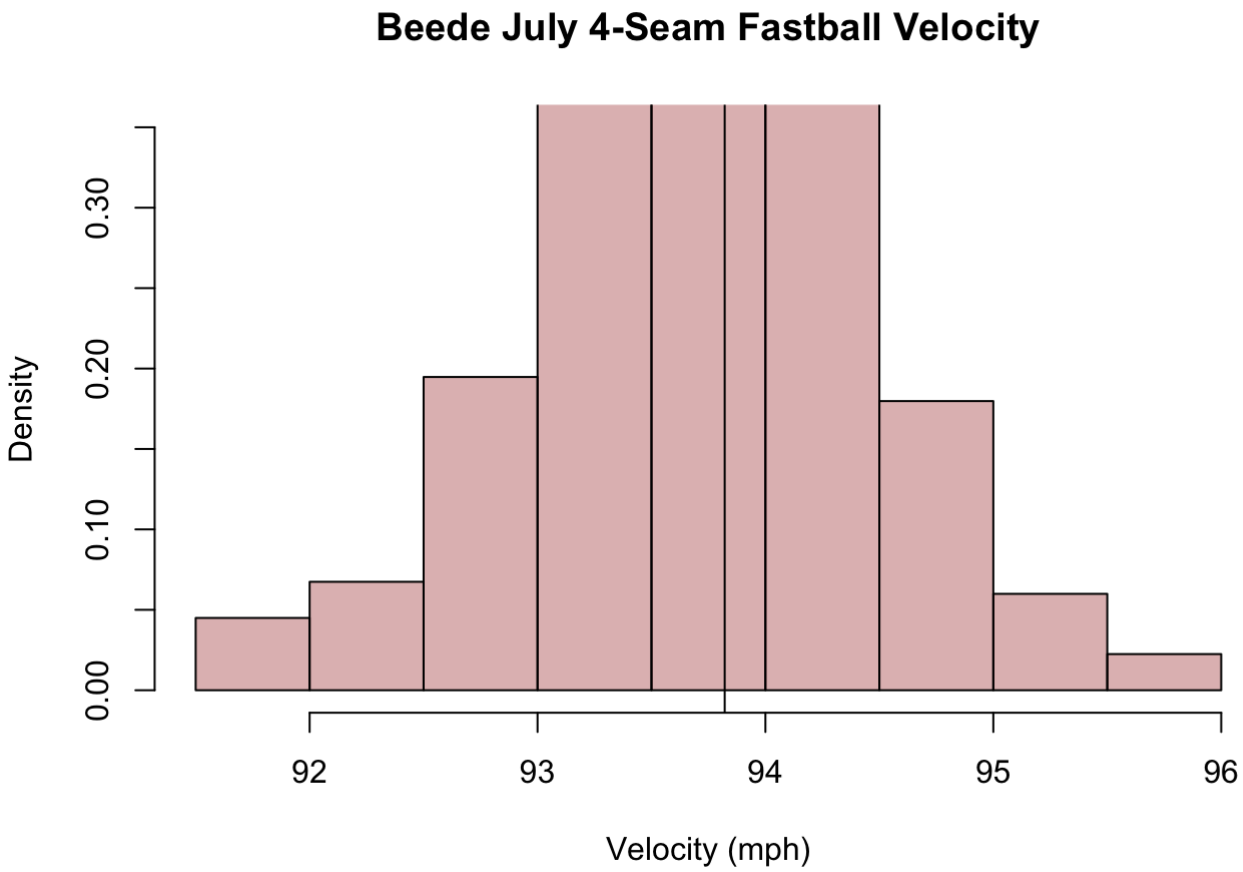


Here is a side by side chart of fastball speeds in july vs the other months in 2019. Beede threw a lot more pitches, but never threw over 96 in July, while he did reach 97 and 98 in a different month. We can break this down more to look at the averages to confirm if he threw his fastball harder in July.

```
#plot the release speeds of his fastballs with a mean line on both
hist(other_ff$release_speed,
     col = "#00009950", freq = FALSE,
     ylim = c(0, .35), xlab = "Velocity (mph)",
     main = "Beede 4-Seam Fastball Velocity")
abline(v=mean(other_ff$release_speed,
              col = "#0000950", freq=FALSE))
```




```
#mean fastball higher in other months, lower in July
hist(july_ff$release_speed,
     col = "#99000050", freq = FALSE,
     ylim = c(0, .35), xlab = "Velocity (mph)",
     main = "Beede July 4-Seam Fastball Velocity")
abline(v=mean(july_ff$release_speed,
              col = "#0000950", freq=FALSE))
```



These are the charts with a line of the average velocity for their respective months. In July, he threw his fastball with an average of 93.8 mph, while in the other months he averaged 94.4 mph. Although only a slight difference, this is more likely due to some wear and tear in the arm during the All star Break

```
#look at mean pitch speeds and mean ff speeds for july and other months
tapply(beede$release_speed, beede$july, mean)
```

```
##     july    other
## 89.36004 89.51523
```

```
beede_ff <- subset(beede, beede$pitch_type== "FF")
head(beede_ff)
```

```
## # A tibble: 6 x 94
##   pitch_type game_date   year month   day release_speed release_pos_x
##   <chr>      <date>     <dbl> <dbl> <dbl>         <dbl>         <dbl>
## 1 FF         2019-09-26  2019     9    26          94.8         -2.51
## 2 FF         2019-09-26  2019     9    26          95.9         -2.58
## 3 FF         2019-09-26  2019     9    26          94.3         -2.70
## 4 FF         2019-09-26  2019     9    26          94.8         -2.65
## 5 FF         2019-09-26  2019     9    26          95.5         -2.45
## 6 FF         2019-09-26  2019     9    26          94.7         -2.42
## # … with 87 more variables: release_pos_z <dbl>, player_name <chr>,
## #   batter <dbl>, pitcher...8 <dbl>, events <chr>, description <chr>,
## #   spin_dir <lgl>, spin_rate_deprecated <lgl>, break_angle_deprecated <lgl>,
## #   break_length_deprecated <lgl>, zone <dbl>, des <chr>, game_type <chr>,
## #   stand <chr>, p_throws <chr>, home_team <chr>, away_team <chr>, type <chr>,
## #   hit_location <chr>, bb_type <chr>, balls <dbl>, strikes <dbl>,
## #   game_year <dbl>, pfx_x <dbl>, pfx_z <dbl>, px <dbl>, pz <dbl>, on_3b <dbl>,
## #   on_2b <dbl>, on_1b <dbl>, outs_when_up <dbl>, inning <dbl>,
## #   inning_topbot <chr>, hc_x <dbl>, hc_y <dbl>, tfs_deprecated <lgl>,
## #   tfs_zulu_deprecated <lgl>, fielder_2...42 <dbl>, umpire <lgl>, sv_id <chr>,
## #   vx0 <dbl>, vy0 <dbl>, vz0 <dbl>, ax <dbl>, ay <dbl>, az <dbl>,
## #   sz_top <dbl>, sz_bot <dbl>, hit_distance_sc <dbl>, launch_speed <dbl>,
## #   launch_angle <dbl>, effective_speed <dbl>, release_spin_rate <dbl>,
## #   release_extension <dbl>, game_pk <dbl>, pitcher...60 <dbl>,
## #   fielder_2...61 <dbl>, fielder_3 <dbl>, fielder_4 <dbl>, fielder_5 <dbl>,
## #   fielder_6 <dbl>, fielder_7 <dbl>, fielder_8 <dbl>, fielder_9 <dbl>,
## #   release_pos_y <dbl>, estimated_ba_using_speedangle <chr>,
## #   estimated_woba_using_speedangle <chr>, woba_value <chr>, woba_denom <chr>,
## #   babip_value <chr>, iso_value <chr>, launch_speed_angle <chr>,
## #   at_bat_number <dbl>, pitch_number <dbl>, pitch_name <chr>,
## #   home_score <dbl>, away_score <dbl>, bat_score <dbl>, fld_score <dbl>,
## #   post_away_score <dbl>, post_home_score <dbl>, post_bat_score <dbl>,
## #   post_fld_score <dbl>, if_fielding_alignment <chr>,
## #   of_fielding_alignment <chr>, barrel <dbl>, july <chr>
```

```
ff_velo_month <- tapply(beede_ff$release_speed, beede_ff$july, mean)
print(ff_velo_month)
```

```
##     july    other
## 93.82135 94.36540
```

```
library(dplyr)
```

Found mean fastball speeds for july and other months, July being 93.8 like stated above and other months he averaged out at 94.4 mph.

```
ff_dt <- data.frame(tapply(beede_ff$release_speed, beede_ff$game_date, mean))
print(ff_dt)
```

```
##            tapply.beede_ff.release_speed..beede_ff.game_date..mean.
## 2019-05-03                                                 95.35714
## 2019-05-09                                                 94.78485
## 2019-05-14                                                 94.43235
## 2019-05-30                                                 93.63962
## 2019-06-05                                                 94.68600
## 2019-06-11                                                 94.64909
## 2019-06-17                                                 94.88000
## 2019-06-22                                                 94.02200
## 2019-06-27                                                 93.61475
## 2019-07-02                                                 93.94000
## 2019-07-14                                                 93.83750
## 2019-07-19                                                 93.82000
## 2019-07-24                                                 94.05208
## 2019-07-30                                                 93.42500
## 2019-08-04                                                 94.31316
## 2019-08-09                                                 94.01667
## 2019-08-14                                                 93.69545
## 2019-08-20                                                 93.37708
## 2019-08-26                                                 93.97407
## 2019-09-02                                                 94.38947
## 2019-09-07                                                 94.78947
## 2019-09-13                                                 94.58140
## 2019-09-20                                                 95.41489
## 2019-09-26                                                 95.18182
```

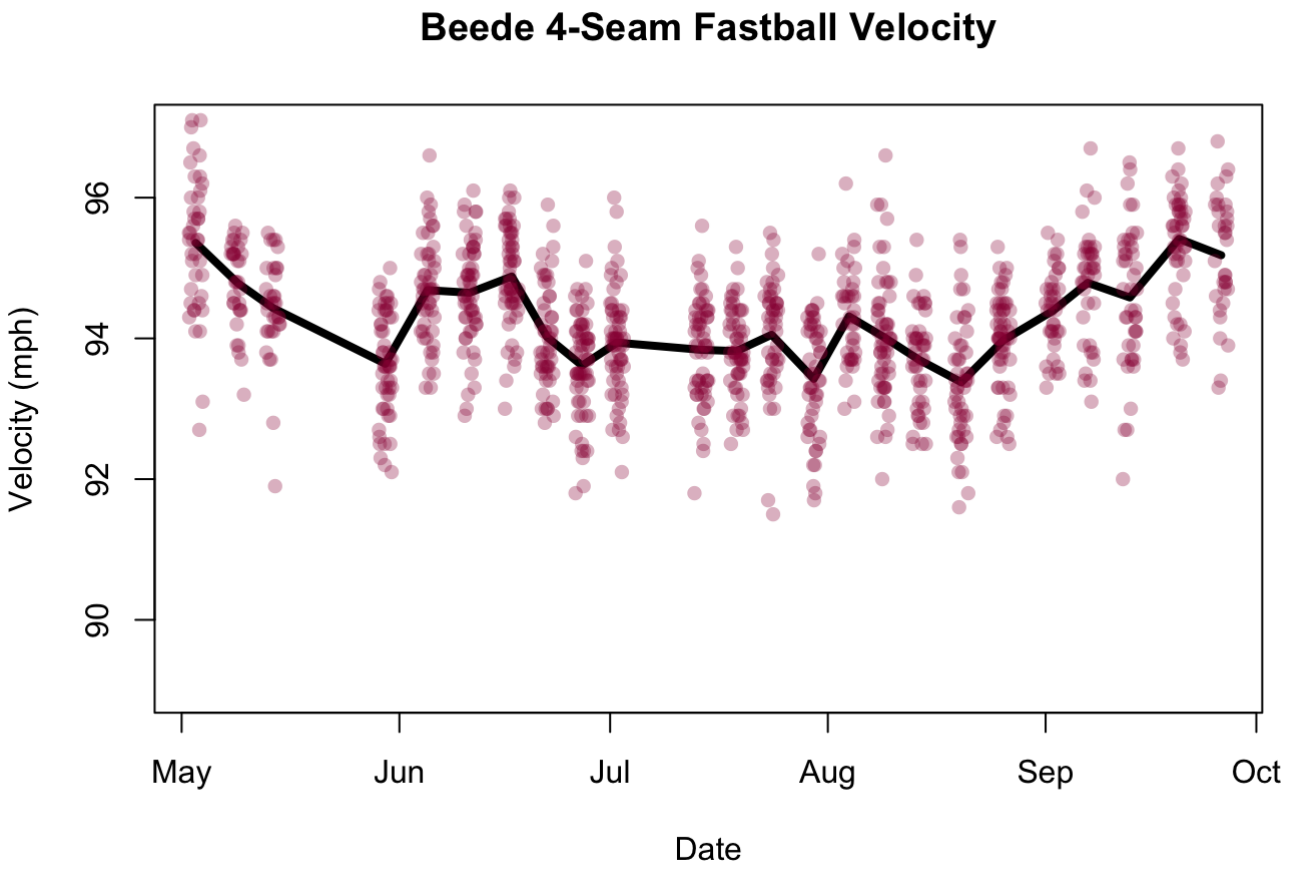
In this, I looked at mean fastball speeds for all his starts, the highest coming on his september 20th game, and his first start on May 3rd. Keep in mind, on May 3rd, he got roughed up a little bit playing against the reds and allowing 7 runs in 2.1/3 IP. In July, while performing at his best, was barely able to crack 94mph on average in his starts.

```
#Create game date in ff dataframe
ff_dt$game_date <- as.Date(row.names(ff_dt), "%Y-%m-%d")
#Rename first column
colnames(ff_dt)[1] <- "release_speed"
#Remove row  names
row.names(ff_dt) <- NULL
head(ff_dt)
```

```
##   release_speed  game_date
## 1      95.35714 2019-05-03
## 2      94.78485 2019-05-09
## 3      94.43235 2019-05-14
## 4      93.63962 2019-05-30
## 5      94.68600 2019-06-05
## 6      94.64909 2019-06-11
```

Means by game date shown above.

```
plot(ff_dt$release_speed ~ ff_dt$game_date,
     lwd =4, type = "l", ylim = c(89,97),
     main = "Beede 4-Seam Fastball Velocity",
     xlab = "Date", ylab = "Velocity (mph)")

points(beede_ff$release_speed ~ jitter(as.numeric(beede_ff$game_date)),
       pch = 16, col = "#99004450")
```



This is a plot of the averages as well as heatmaps of the fastballs thrown. His largest average came in May and September, smallest averages were July and beginning and middle of August.

```
table(beede$pitch_type)
```

```
##
##   CH   CU   FF   FT   SL
##  378  278 1137    1  232
```

```
table(beede$pitch_type, beede$month)
```

```
##
##        5   6   7   8   9
##   CH  66 111  75  73  53
##   CU  47  82  44  46  59
##   FF 162 271 267 238 199
##   FT   1   0   0   0   0
##   SL   0   6  97  59  70
```

```
prop.table(table(beede$pitch_type, beede$month), margin =2)
```

```
##
##                5           6           7           8           9
##   CH 0.239130435 0.236170213 0.155279503 0.175480769 0.139107612
##   CU 0.170289855 0.174468085 0.091097308 0.110576923 0.154855643
##   FF 0.586956522 0.576595745 0.552795031 0.572115385 0.522309711
##   FT 0.003623188 0.000000000 0.000000000 0.000000000 0.000000000
##   SL 0.000000000 0.012765957 0.200828157 0.141826923 0.183727034
```

```
type_tab2 <- prop.table(table(beede_july$pitch_type, beede_july$month), margin =2)
```

This is a proportion table of the pitches Beede threw each month. His percentage in fastball stays pretty consistent, but falters off in July and September. Something intriguing to me was that he stopped throwing his changeup as much after June. Having not thrown a single slider in May, looks like he developed that pitch more and got more comfortable with it. He threw it the most in July and September, his two best months, while also throwing his fastball less. Tyler also had a really high hard hit percentage in 2019, his best two months were July and September, where he threw more offspeed pitches.

Tyler's best pitch is obviously his curveball, hitters hit .159 off of it in 2019. In Kerry Crowley's article, he mentions the use of analytics in Stephen Strasburg's dominant postseason, especially against the Astros. Strasburg usually throws his fastball around 28% of his total pitches, which he hovered over in the postseason as well. Before investigating more, Beede may find more success throwing his offspeed pitches off of his fastball, especially in different counts.

Something else that caught my eye was Beede's uptick in 4 seam velocity in September, which usually doesn't happen. This may have been due to throwing it less and having the ability to throw it harder still. If Beede takes an approach like Stephen Strasburg, throwing more offspeed for strikes, this may give him an ability to catch hitters off guard when he does throw heat, hitters hit .285 off of it in 2019.

```
type_tab <- table(beede$pitch_type, beede$july)
print(type_tab)
```

```
##
##       july other
##   CH   75   303
##   CU   44   234
##   FF  267   870
##   FT    0     1
##   SL   97   135
```

```
type_prop <- round(prop.table(type_tab, margin = 2), 3)
print(type_prop)
```

```
##
##       july other
##   CH 0.155 0.196
##   CU 0.091 0.152
##   FF 0.553 0.564
##   FT 0.000 0.001
##   SL 0.201 0.087
```

```
#September
beede$september <- ifelse(beede$month == 9, "september", "other")

type_tab <- table(beede$pitch_type, beede$september)
print(type_tab)
```

```
##
##       other september
##   CH   325        53
##   CU   219        59
##   FF   938       199
##   FT     1         0
##   SL   162        70
```

```
type_prop <- round(prop.table(type_tab, margin = 2), 3)
print(type_prop)
```

```
##
##      other september
##   CH 0.198     0.139
##   CU 0.133     0.155
##   FF 0.570     0.522
##   FT 0.001     0.000
##   SL 0.098     0.184
```

In September, he used his curveball, statistically his best pitch a lot more. His breaking pitches got a 50% swing and miss rate in 2019, 2nd behind Blake Snell. Overall, he threw much more off-speed in July and September compared to other months.

```
ff_prop <- type_prop[3,]
print(ff_prop)
```

```
##     other september
##     0.570     0.522
```

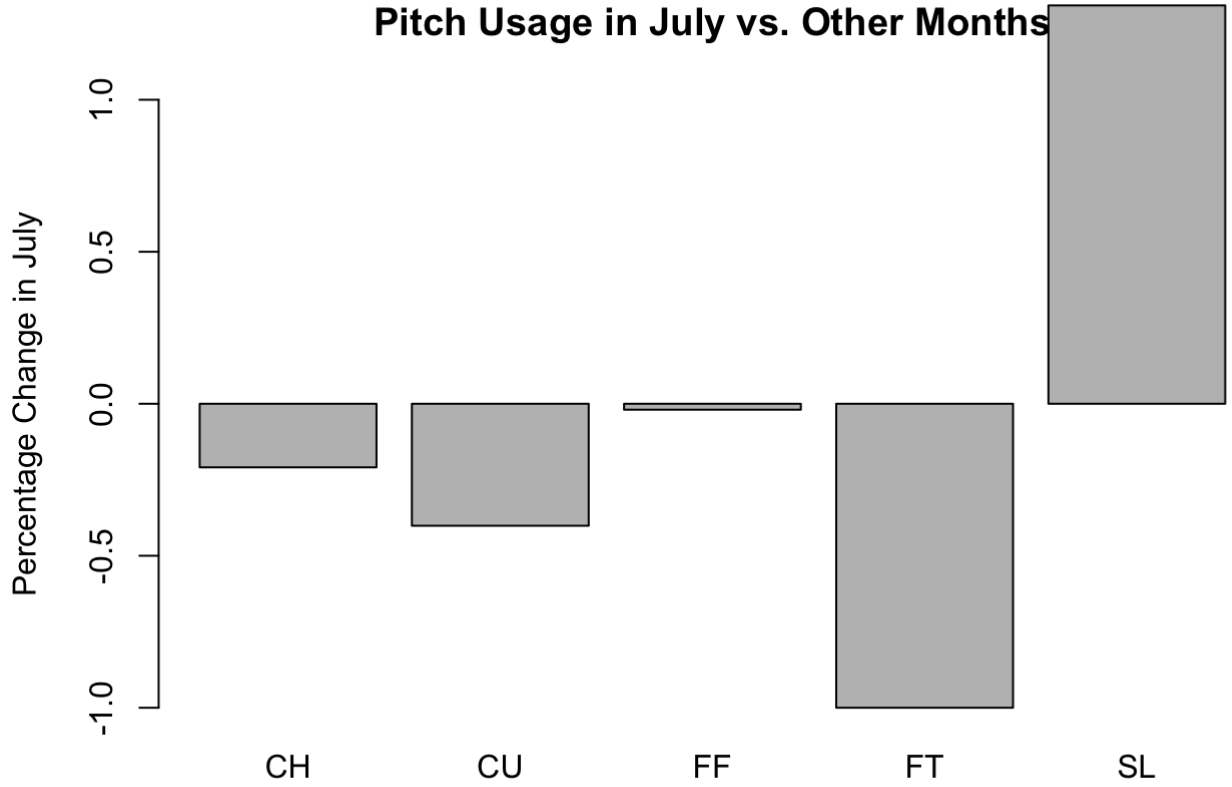
```
print(ff_velo_month)
```

```
##     july    other
## 93.82135 94.36540
```

```
# Pitch mix tables-changes in pitching rates
# Create the Difference column
Pitch <- c("CH", "CU", "FF", "FT", "SL")
july <- c(0.155, 0.091,0.553,0.000,0.201)
other <- c(0.196, 0.152, 0.564,0.001,0.087)
type_prop_df <- data.frame(Pitch, july, other)
type_prop_df$Difference <- (type_prop_df$july - type_prop_df$other) / type_prop_df$other

september <- c(0.139, 0.155, 0.522, 0.000, 0.184)
other2 <- c(0.198, 0.133, 0.570, 0.001, 0.098)
type_prop_df2 <- data.frame(Pitch, september, other2)
type_prop_df2$Difference <- (type_prop_df2$september - type_prop_df2$other) / type_prop_df2$other


# Print the type_prop
type_prop_df
```

```
##   Pitch  july other  Difference
## 1    CH 0.155 0.196 -0.20918367
## 2    CU 0.091 0.152 -0.40131579
## 3    FF 0.553 0.564 -0.01950355
## 4    FT 0.000 0.001 -1.00000000
## 5    SL 0.201 0.087  1.31034483
```

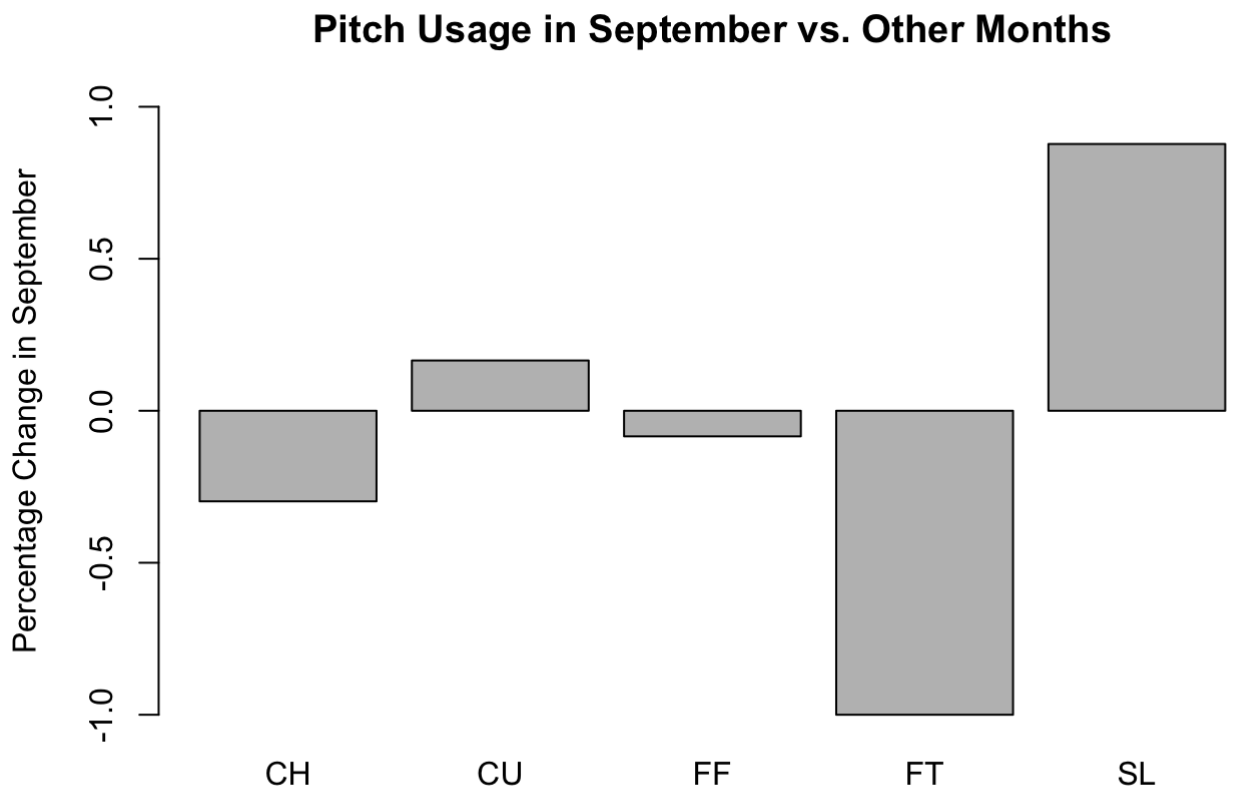
```
# Plot a barplot
barplot(type_prop_df$Difference, names.arg = type_prop_df$Pitch,
        main = "Pitch Usage in July vs. Other Months",
        ylab = "Percentage Change in July",
        ylim = c(-1, 1))
```




```
type_prop_df2
```

```
##   Pitch september other2  Difference
## 1    CH     0.139  0.198 -0.29797980
## 2    CU     0.155  0.133  0.16541353
## 3    FF     0.522  0.570 -0.08421053
## 4    FT     0.000  0.001 -1.00000000
## 5    SL     0.184  0.098  0.87755102
```

```
barplot(type_prop_df2$Difference, names.arg = type_prop_df2$Pitch,
        main = "Pitch Usage in September vs. Other Months",
        ylab = "Percentage Change in September",
        ylim = c(-1, 1))
```



July: Visually shown here is a percentage change in the pitches he threw in July compared to the other months. This is just reinforcing what we saw with the proportions. He threw his 4-seam a very small amount less, as well as throwing his changeup and curveball less. He threw his slider significantly more in July.

September: In September, He threw his fastball and changeup less, but threw his breaking pitches: curveball and slider more. Throwing his offspeed pitches more have shown to have more success, especially when throwing them for strikes, rather than throwing for swing and misses. In 2019, Beede allowed 16 of his 22 homeruns allowed on his fastball, above the league average of 14.

```
# Create bs_table
bs_table <- table(beede$balls, beede$strikes)


# Create bs_prop_table
bs_prop_table <- round(prop.table(bs_table),3)

# Print bs_prop_table
print(bs_prop_table)
```

```
##
##         0     1     2
##   0 0.258 0.132 0.059
##   1 0.089 0.102 0.097
##   2 0.030 0.052 0.088
##   3 0.010 0.024 0.059
```

```
# Print row sums
print(rowSums(bs_prop_table))
```

```
##     0     1     2     3
## 0.449 0.288 0.170 0.093
```

```
#Percentage of pitches thrown with different amounts of balls in the count

# Print column sums
print(colSums(bs_prop_table))
```

```
##     0     1     2
## 0.387 0.310 0.303
```

```
#Pitches Beede throws in different counts
#Percentage of pitches thrown with different amounts of balls in the count
```

In this chart, it shows the percentage of times he is throwing with different counts, 26.3% of pitches thrown in 2 and 3 ball counts, allowing for more hitter friendly counts. According to baseball savant, meatball % is a stat that calculates a percentage of pitches being classified as a “meatball”. Usually means a pitch right down the middle that is easily barrelled up by a hitter. Tyler Beede had a 8.4% meatball %, the league average was 7%, this could be due to being in a lot of hitter friendly counts.

```
# Create bs_count
beede$bs_count <- paste(beede$balls, beede$strikes, sep = "-")
#Combines balls and counts into single column
# Print the first 6 rows of beede
head(beede)
```

```
## # A tibble: 6 x 96
##   pitch_type game_date   year month   day release_speed release_pos_x
##   <chr>      <date>     <dbl> <dbl> <dbl>         <dbl>         <dbl>
## 1 FF         2019-09-26  2019     9    26          94.8         -2.51
## 2 FF         2019-09-26  2019     9    26          95.9         -2.58
## 3 CU         2019-09-26  2019     9    26          81.4         -2.73
## 4 FF         2019-09-26  2019     9    26          94.3         -2.70
## 5 FF         2019-09-26  2019     9    26          94.8         -2.65
## 6 CU         2019-09-26  2019     9    26          81.7         -2.74
## # … with 89 more variables: release_pos_z <dbl>, player_name <chr>,
## #   batter <dbl>, pitcher...8 <dbl>, events <chr>, description <chr>,
## #   spin_dir <lgl>, spin_rate_deprecated <lgl>, break_angle_deprecated <lgl>,
## #   break_length_deprecated <lgl>, zone <dbl>, des <chr>, game_type <chr>,
## #   stand <chr>, p_throws <chr>, home_team <chr>, away_team <chr>, type <chr>,
## #   hit_location <chr>, bb_type <chr>, balls <dbl>, strikes <dbl>,
## #   game_year <dbl>, pfx_x <dbl>, pfx_z <dbl>, px <dbl>, pz <dbl>, on_3b <dbl>,
## #   on_2b <dbl>, on_1b <dbl>, outs_when_up <dbl>, inning <dbl>,
## #   inning_topbot <chr>, hc_x <dbl>, hc_y <dbl>, tfs_deprecated <lgl>,
## #   tfs_zulu_deprecated <lgl>, fielder_2...42 <dbl>, umpire <lgl>, sv_id <chr>,
## #   vx0 <dbl>, vy0 <dbl>, vz0 <dbl>, ax <dbl>, ay <dbl>, az <dbl>,
## #   sz_top <dbl>, sz_bot <dbl>, hit_distance_sc <dbl>, launch_speed <dbl>,
## #   launch_angle <dbl>, effective_speed <dbl>, release_spin_rate <dbl>,
## #   release_extension <dbl>, game_pk <dbl>, pitcher...60 <dbl>,
## #   fielder_2...61 <dbl>, fielder_3 <dbl>, fielder_4 <dbl>, fielder_5 <dbl>,
## #   fielder_6 <dbl>, fielder_7 <dbl>, fielder_8 <dbl>, fielder_9 <dbl>,
## #   release_pos_y <dbl>, estimated_ba_using_speedangle <chr>,
## #   estimated_woba_using_speedangle <chr>, woba_value <chr>, woba_denom <chr>,
## #   babip_value <chr>, iso_value <chr>, launch_speed_angle <chr>,
## #   at_bat_number <dbl>, pitch_number <dbl>, pitch_name <chr>,
## #   home_score <dbl>, away_score <dbl>, bat_score <dbl>, fld_score <dbl>,
## #   post_away_score <dbl>, post_home_score <dbl>, post_bat_score <dbl>,
## #   post_fld_score <dbl>, if_fielding_alignment <chr>,
## #   of_fielding_alignment <chr>, barrel <dbl>, july <chr>, september <chr>,
## #   bs_count <chr>
```

Created a ball and strike count column

```
# Create bs_count_tab
bs_count_tab <- table(beede$bs_count, beede$july)
bs_count_tab2 <- table(beede$bs_count, beede$september)

# Create bs_month
bs_month <- round(prop.table(bs_count_tab,margin =2),3)
bs_month2 <- round(prop.table(bs_count_tab2, margin =2),3)

# Print bs_month
print(bs_month)
```

```
##
##        july other
##   0-0 0.280 0.251
##   0-1 0.139 0.130
##   0-2 0.062 0.058
##   1-0 0.095 0.087
##   1-1 0.083 0.108
##   1-2 0.095 0.097
##   2-0 0.029 0.030
##   2-1 0.043 0.055
##   2-2 0.089 0.087
##   3-0 0.004 0.012
##   3-1 0.014 0.027
##   3-2 0.066 0.056
```

```
print(bs_month2)
```

```
##
##       other september
##   0-0 0.258     0.260
##   0-1 0.134     0.126
##   0-2 0.059     0.060
##   1-0 0.088     0.094
##   1-1 0.102     0.102
##   1-2 0.101     0.079
##   2-0 0.028     0.039
##   2-1 0.050     0.060
##   2-2 0.089     0.084
##   3-0 0.009     0.013
##   3-1 0.023     0.029
##   3-2 0.060     0.052
```

```
#percentage change in rate at which Beede puts himself in these counts
```

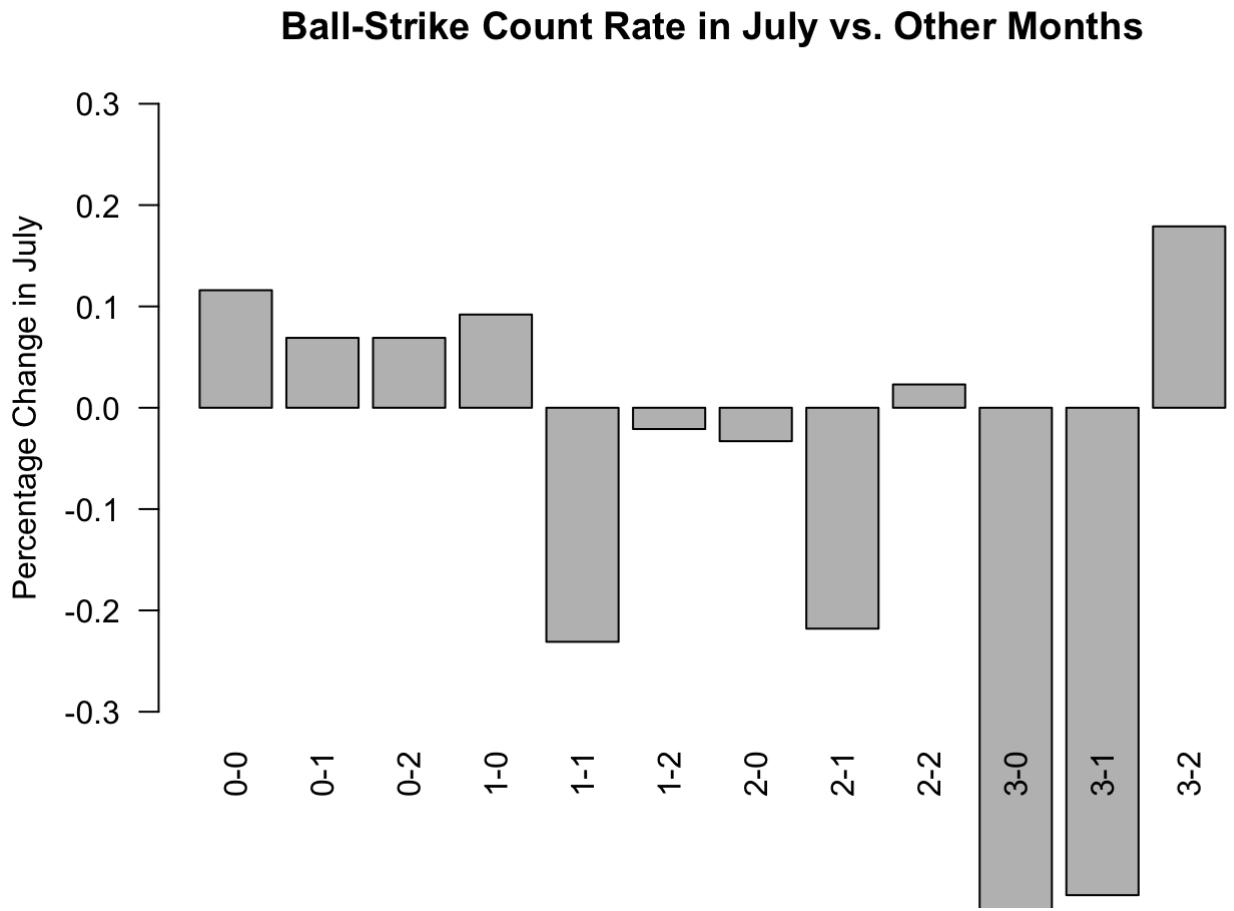
July: In July, Beede got into less 2 ball and 3 ball counts, usually considered great hitters counts. He did get a few more 0-1 and 0-2 counts, showing he was perhaps a bit more aggresive compared other months.

September: In September, Beede actually got into more hitter friendly counts, which is interesting considering his success, although this may be due to him throwing more offspeed in September. Offspeed pitches are more difficult to control compared to his fastball.

```
# Create diff_bs
diff_bs <- round(((bs_month[,1]-bs_month[,2])/bs_month[,2]),3)

# Print diff_bs
print(diff_bs)
```

```
##    0-0    0-1    0-2    1-0    1-1    1-2    2-0    2-1    2-2    3-0    3-1
##  0.116  0.069  0.069  0.092 -0.231 -0.021 -0.033 -0.218  0.023 -0.667 -0.481
##    3-2
##  0.179
```

```
#how common each ball-strike count was in July vs Other Months
# Create a bar plot of the changes
barplot(diff_bs, main = "Ball-Strike Count Rate in July vs. Other Months",
        ylab = "Percentage Change in July", ylim = c(-0.3, 0.3), las = 2)
```

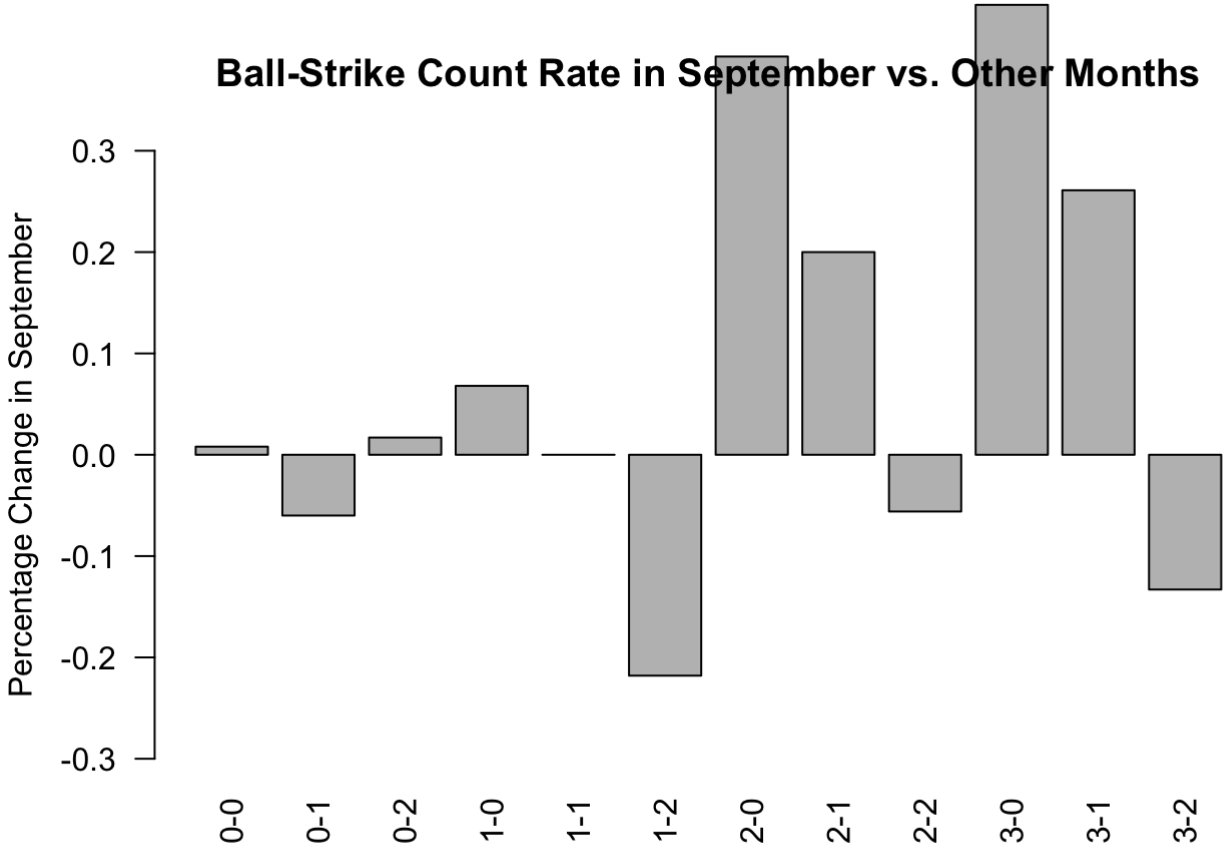


This is the visual plot of percentage changes in pitch counts, decline in 3 ball counts which is good news. There is also a decrease in 2 ball counts overall, leaving him with an advantage in the counts. The increase in 0-1 and 0-2 counts are also a reason for part of his successful July.

```
# Create diff_bs
diff_bs2 <- round(((bs_month2[,2]-bs_month2[,1])/bs_month2[,1]),3)

# Print diff_bs
print(diff_bs2)
```

```
##    0-0    0-1    0-2    1-0    1-1    1-2    2-0    2-1    2-2    3-0    3-1
##  0.008 -0.060  0.017  0.068  0.000 -0.218  0.393  0.200 -0.056  0.444  0.261
##    3-2
## -0.133
```

```
#how common each ball-strike count was in July vs Other Months
# Create a bar plot of the changes
barplot(diff_bs2, main = "Ball-Strike Count Rate in September vs. Other Months",
        ylab = "Percentage Change in September", ylim = c(-0.3, 0.3), las = 2)
```



This is the plot for September, a big increase in 2 ball and 3 ball counts. Most likely, this is due to throwing more curveballs and sliders as these pitches are meant to get the batter to bite and get them off balance.

```
#Taking a look at Beede throwing certain pitches in certain counts more often or less often
# Create type_bs
type_bs <- table(beede$pitch_type, beede$bs_count)

# Print type_bs
print(type_bs)
```

```
##
##      0-0 0-1 0-2 1-0 1-1 1-2 2-0 2-1 2-2 3-0 3-1 3-2
##   CH  74  45  19  45  50  35  12  22  44   0   6  26
##   CU  56  64  29   7  34  47   1   5  31   1   0   3
##   FF 323 136  55  94 100  87  44  74  80  19  41  84
##   FT   0   0   0   0   0   1   0   0   0   0   0   0
##   SL  70  23  17  34  22  26   4   5  23   0   2   6
```

```
# Create type_bs_prop
type_bs_prop <- round(prop.table(type_bs,margin =2),3)

# Print type_bs_prop
print(type_bs_prop)
```

```
##
##        0-0   0-1   0-2   1-0   1-1   1-2   2-0   2-1   2-2   3-0   3-1   3-2
##   CH 0.141 0.168 0.158 0.250 0.243 0.179 0.197 0.208 0.247 0.000 0.122 0.218
##   CU 0.107 0.239 0.242 0.039 0.165 0.240 0.016 0.047 0.174 0.050 0.000 0.025
##   FF 0.618 0.507 0.458 0.522 0.485 0.444 0.721 0.698 0.449 0.950 0.837 0.706
##   FT 0.000 0.000 0.000 0.000 0.000 0.005 0.000 0.000 0.000 0.000 0.000 0.000
##   SL 0.134 0.086 0.142 0.189 0.107 0.133 0.066 0.047 0.129 0.000 0.041 0.050
```

This is the percentage of pitches thrown in each count for 2019. I am not a Major League hitter, but if I can work a 2 ball or 3 ball count, I'll just sit on the fastball. Also interesting that he throws a fastball 62% of the time in 0-0 counts, another 11% if include a cutter, a fastball variation. Comparatively, in 1-0 counts and 2-0 counts, he throws his changeup more compared to 0-0 counts.

```
# Create the late_in_game column
#Table indicates pitches thrown late in the game, past 5th inning
beede$late_in_game <- ifelse(beede$inning > 5, 1, 0)

# Convert late_in_game
beede$late_in_game <- factor(beede$late_in_game)

# Create type_late
type_late <- table(beede$pitch_type, beede$late_in_game)

# Create type_late_prop
type_late_prop <- round(prop.table(type_late, margin = 2), 3)

# Print type_late_prop
type_late_prop
```

```
##
##          0     1
##   CH 0.190 0.158
##   CU 0.143 0.095
##   FF 0.555 0.610
##   FT 0.001 0.000
##   SL 0.111 0.137
```

This chart denotes 0 for early in the game and 1 for late in the game. I wanted to see if Tyler was changing his pitches late in the game or not. For today's game of baseball, I consider past the 5th inning to be late. One notable difference is he throws his fastball 61% of the time late in the game, basically straying away from his off-speed. He also throws his best two pitches: his changeup and his curveball less late in the game.

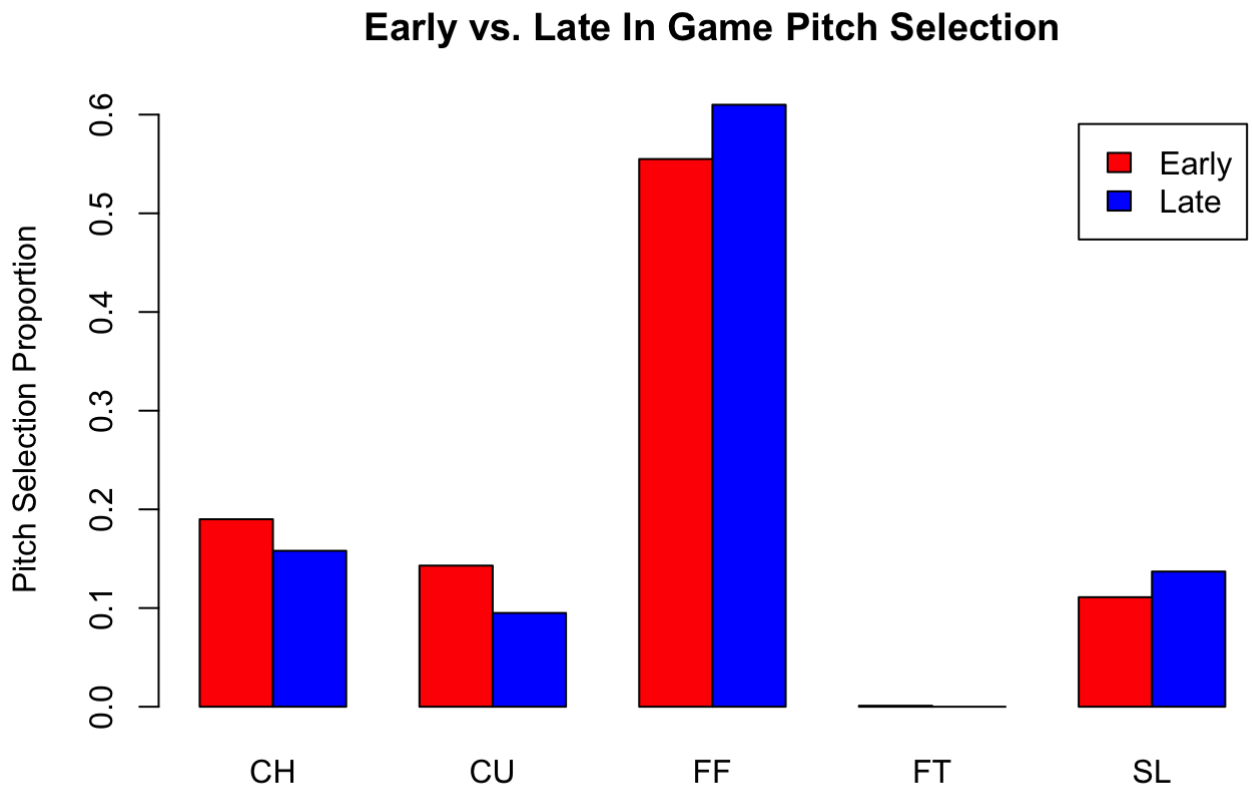
```
# Create t_type_late
t_type_late <- t(type_late_prop)

# Print dimensions of t_type_late
dim(t_type_late)
```

```
## [1] 2 5
```

```
# Print dimensions of type_late
dim(type_late)
```

```
## [1] 5 2
```

```
# Change row names
rownames(t_type_late) <- c("Early", "Late")

# Make barplot using t_type_late
barplot(t_type_late, beside = TRUE, col = c("red", "blue"),
        main = "Early vs. Late In Game Pitch Selection",
        ylab = "Pitch Selection Proportion",
        legend = rownames(t_type_late))
```



This is a proportion chart of early and late game pitch selections. As noted before, he throws his fastball more late in the game, and throws his off-speed less. Sticking with his better pitches like his curveball or his changeup can get batters guessing more rathering than sitting on his fastball.

```
# Calculate average pitch height in inches in July vs. other months
tapply(beede$pz, beede$july, mean) * 12
```

```
##     july    other
## 27.76324 26.89152
```

```
beede_lhb <- subset(beede, stand == "L")


beede_rhb <- subset(beede, stand == "R")

# Compute average px location for LHB
tapply(beede_lhb$px, beede_lhb$july, mean) * 12
```

```
##      july     other
## -2.997742 -3.563696
```

```
# Compute average px location for RHB
tapply(beede_rhb$px, beede_rhb$july, mean) * 12
```

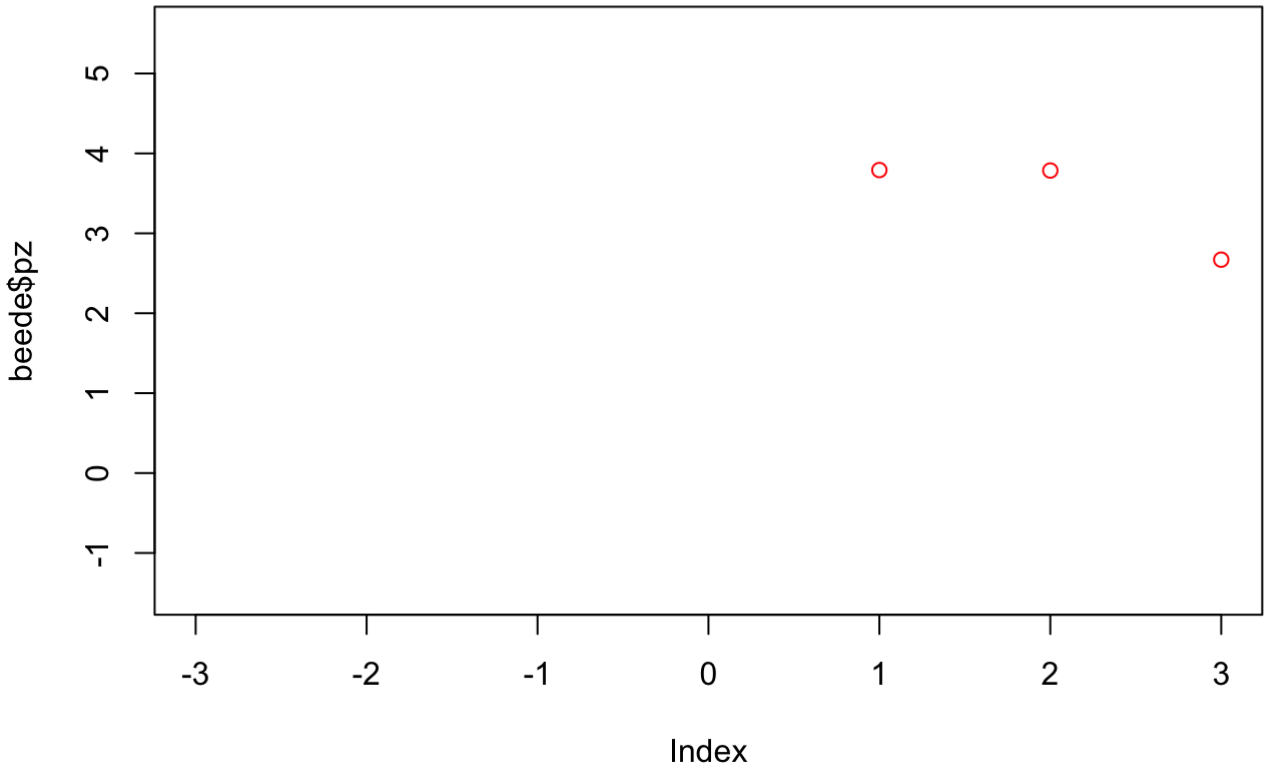
```
##     july    other
## 3.132771 4.345459
```

```
#Noting that negative px refers to outside half pitches to LHB, while positive px refers to outside half pitches
 to RHB,
```

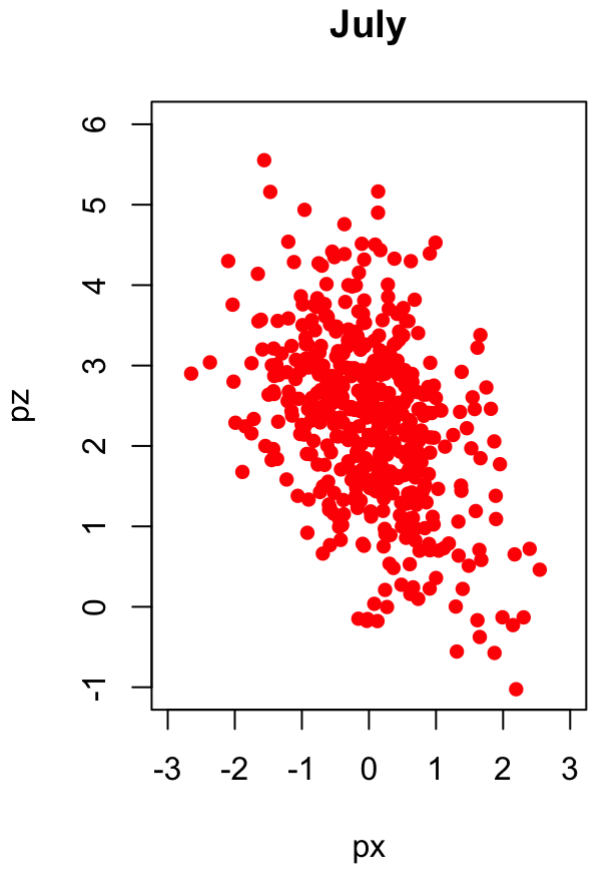
Essentially here we calculated the average px locations for each side hitter. Px is the horizontal location of the pitch, a negative px referes to outside pitches to left handed batters, while positive px refers to outside pitches to right handed batters. Averages came out different for July for average px locations, a big difference looking at outside pitches for righties in July at 3.13 compared to other months average of 4.35.

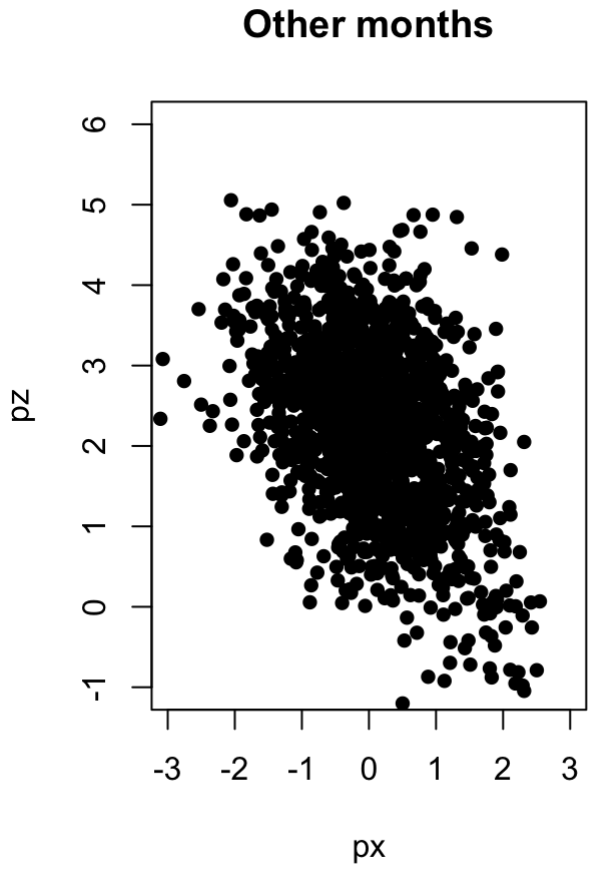
```
# Plot location of all pitches
plot(beede$pz ~ beede$pz,
     col = factor(beede$july),
     xlim = c(-3, 3))
```

```
## Warning in plot.formula(beede$pz ~ beede$pz, col = factor(beede$july), xlim =
## c(-3, : the formula 'beede$pz ~ beede$pz' is treated as 'beede$pz ~ 1'
```

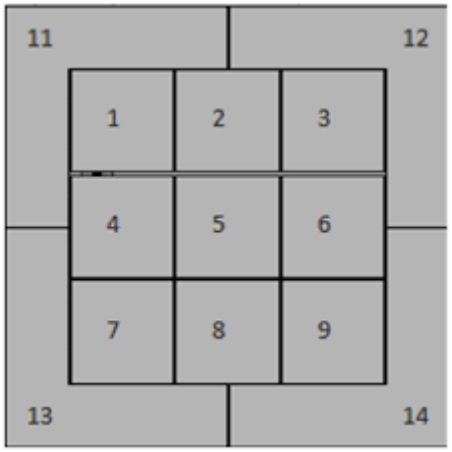

beede$pz
-1
0
1
2
3
4
5
-3
-2
-1
0
1
2
3
Index

```
# Formatting code, don't change this
par(mfrow = c(1, 2))

# Plot the pitch loctions for July
plot(pz ~ px, data = beede_july,
     col = "red", pch = 16,
     xlim = c(-3, 3), ylim = c(-1, 6),
     main = "July")

# Plot the pitch locations for other months
plot(pz ~ px, data = beede_other,
     col = "black", pch = 16,
     xlim = c(-3, 3), ylim = c(-1, 6),
     main = "Other months")
```





This was an interesting graph of pitches that Beede threw in July and other months. The boundaries of the strike zone are -2 to 2 for px and 1 to 4 for pz.



Baseball Savant Strike Zone Zones

```
# Create beede_sub
beede_sub <- subset(beede, px > -2 & px < 2 &
                        pz > 0 & pz < 5)
```

```
# Create zone_prop_july
zone_prop_july <- round(
  table(beede_sub$zone[beede_sub$july == "july"]) /
    nrow(subset(beede_sub, july == "july")), 3)

# Create zone_prop_other
zone_prop_other <- round(
  table(beede_sub$zone[beede_sub$july == "other"]) /
    nrow(subset(beede_sub, july == "other")), 3)

# Print zone_prop_july
zone_prop_july
```

```
##
##     1     2     3     4     5     6     7     8     9    11    12    13    14
## 0.050 0.035 0.037 0.070 0.098 0.044 0.028 0.083 0.061 0.144 0.070 0.105 0.176
```

```
# Print zone_prop_other
zone_prop_other
```

```
##
##     1     2     3     4     5     6     7     8     9    11    12    13    14
## 0.040 0.043 0.040 0.060 0.085 0.051 0.043 0.059 0.053 0.142 0.079 0.099 0.206
```

```
# Fix zone_prop_july vector, don't change this
zone_prop_july2 <- c(zone_prop_july[1:3], 0.00, zone_prop_july[4:14])
names(zone_prop_july2) <- c(1:14)

# Create zone_prop_diff
zone_prop_diff <- zone_prop_july - zone_prop_other

# Print zone_prop_diff
zone_prop_diff
```

```
##
##      1      2      3      4      5      6      7      8      9     11     12
##  0.010 -0.008 -0.003  0.010  0.013 -0.007 -0.015  0.024  0.008  0.002 -0.009
##     13     14
##  0.006 -0.030
```

These charts here denote the proportions of times Beede threw to each zone. Beede threw most to zone 13 and 14 in July which are low and away or low and in. He also threw more in zone 5 compared to other months, which is down the middle.

```
# Create beede_zone_tab
beede_zone_tab <- table(beede_sub$zone, beede_sub$bs_count)

# Create zone_count_prop
zone_count_prop <- round(prop.table(beede_zone_tab, margin = 2), 3)

# Print zone_count_prop
zone_count_prop
```

```
##
##        0-0   0-1   0-2   1-0   1-1   1-2   2-0   2-1   2-2   3-0   3-1   3-2
##   1  0.051 0.020 0.019 0.040 0.047 0.038 0.000 0.067 0.024 0.150 0.083 0.068
##   2  0.051 0.036 0.010 0.040 0.037 0.022 0.033 0.010 0.042 0.050 0.083 0.085
##   3  0.043 0.060 0.076 0.023 0.026 0.043 0.016 0.019 0.024 0.050 0.000 0.051
##   4  0.066 0.060 0.010 0.090 0.058 0.016 0.066 0.095 0.060 0.050 0.104 0.093
##   5  0.111 0.083 0.010 0.085 0.079 0.054 0.148 0.086 0.066 0.100 0.167 0.110
##   6  0.064 0.044 0.010 0.056 0.068 0.022 0.049 0.067 0.024 0.000 0.042 0.059
##   7  0.053 0.028 0.000 0.023 0.026 0.043 0.033 0.067 0.060 0.100 0.042 0.025
##   8  0.059 0.056 0.067 0.119 0.074 0.060 0.049 0.057 0.048 0.200 0.083 0.034
##   9  0.066 0.040 0.038 0.045 0.068 0.043 0.066 0.067 0.060 0.100 0.042 0.034
##   11 0.141 0.119 0.200 0.136 0.121 0.163 0.180 0.124 0.127 0.050 0.146 0.195
##   12 0.061 0.087 0.219 0.006 0.068 0.163 0.066 0.057 0.084 0.000 0.021 0.034
##   13 0.080 0.107 0.181 0.107 0.100 0.109 0.098 0.095 0.139 0.050 0.062 0.059
##   14 0.154 0.262 0.162 0.232 0.226 0.223 0.197 0.190 0.241 0.100 0.125 0.153
```

This is a percentage of pitches thrown in different zones in different counts. Beede throws a lot of pitches out of the zone on 2 strike counts, even 3-2 which is intereseting. I wanted to key in on 1-1 counts and 2-2 counts, these usually are key counts within a game or even inning. Looking at 1-1 counts it looks like he throws a lot up and in/ or up and away, or low and away/ low and in. Probably trying to throw a fastball up or drop a curveball/ changeup in the bottom half on these counts. To open an at bat, he throws a pitch down the middle 11% of the time, something to note if I'm going up there to hit off Beede. Beede's favorite zones are 11 and 14.

```
# Create zone_count_diff
#Looking at 0-2 counts with 3-0 counts
zone_count_diff <- zone_count_prop[, 3] - zone_count_prop[, 10]

# Print the table
zone_count_diff
```

```
##      1      2      3      4      5      6      7      8      9     11     12
## -0.131 -0.040  0.026 -0.040 -0.090  0.010 -0.100 -0.133 -0.062  0.150  0.219
##     13     14
##  0.131  0.062
```

I wanted to break it down for different counts now, and look at the difference between an 0-2 count and a 3-0 count. Clearly doesn't want to throw down the middle on 0-2, but he does extend the strike zone more on 0-2 denoted by the positive variables on zones 11-14.

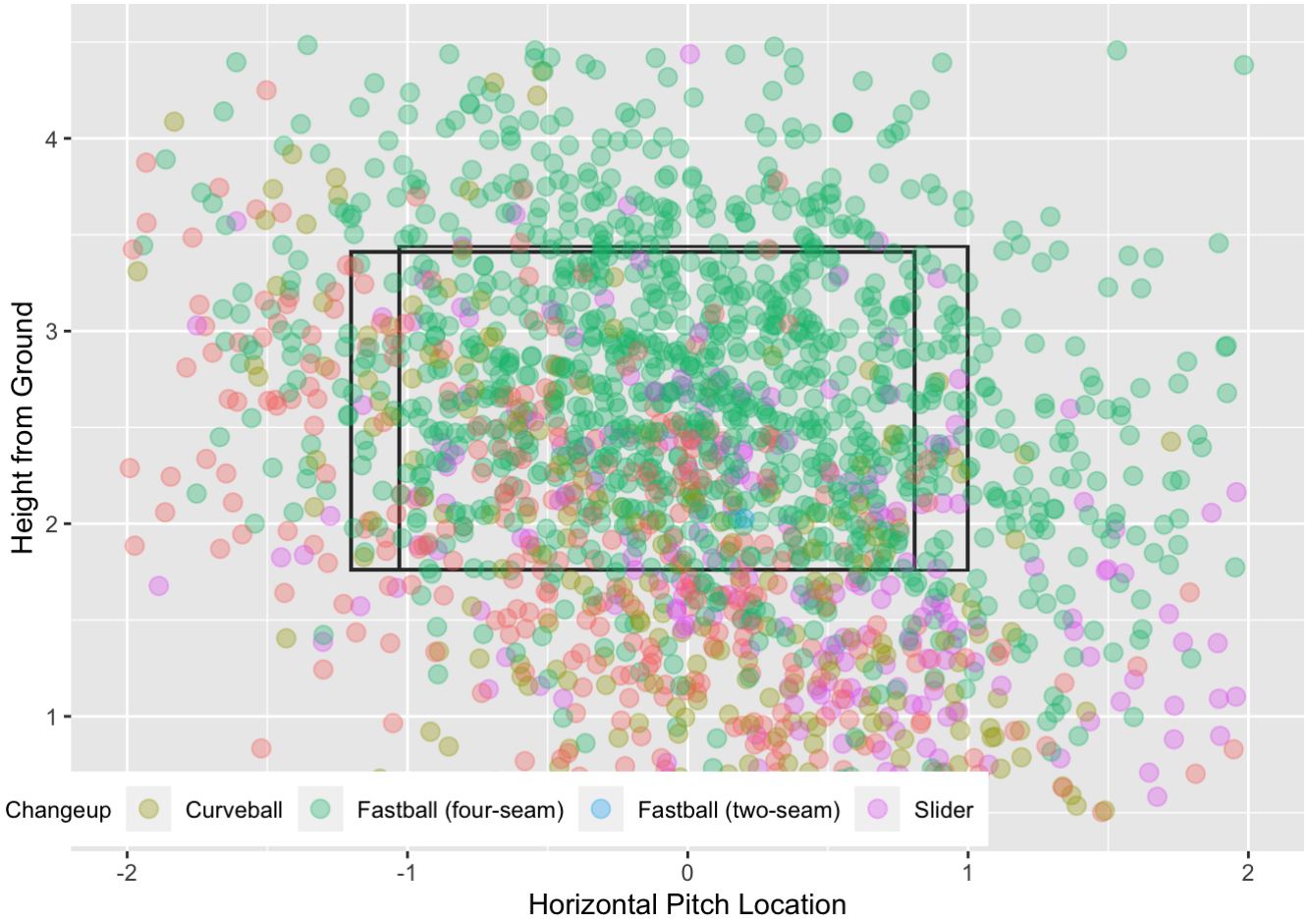
```
library(pitchRx)
```

```
## Loading required package: ggplot2
```

```
strikeFX(beede)
```

```
## Warning in strikeFX(beede): pitchRx assumes the height of each batter is
## recorded as 'b_height'. Since there is no such column, we will assume each
## batter has a height of 6'2''
```

```
## Warning: Removed 171 rows containing missing values (geom_point).
```



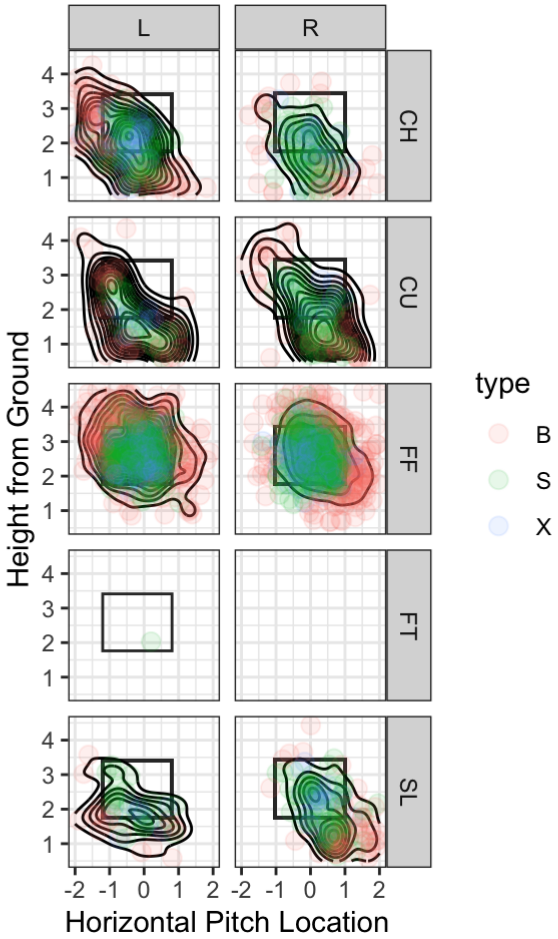
```
strikeFX(beede, color = "type", point.alpha = 0.1,
         adjust = TRUE, contour = TRUE) + facet_grid(pitch_type ~ stand) +
  theme(legend.position = "right", legend.direction = "vertical") +
  coord_equal() + theme_bw()
```

```
## Warning in strikeFX(beede, color = "type", point.alpha = 0.1, adjust = TRUE, :
## pitchRx assumes the height of each batter is recorded as 'b_height'. Since there
## is no such column, we will assume each batter has a height of 6'2''
```

```
## Warning: Removed 171 rows containing non-finite values (stat_density2d).
```

```
## Warning: Computation failed in `stat_density2d()`:
## missing value where TRUE/FALSE needed
```

```
## Warning: Removed 171 rows containing missing values (geom_point).
```



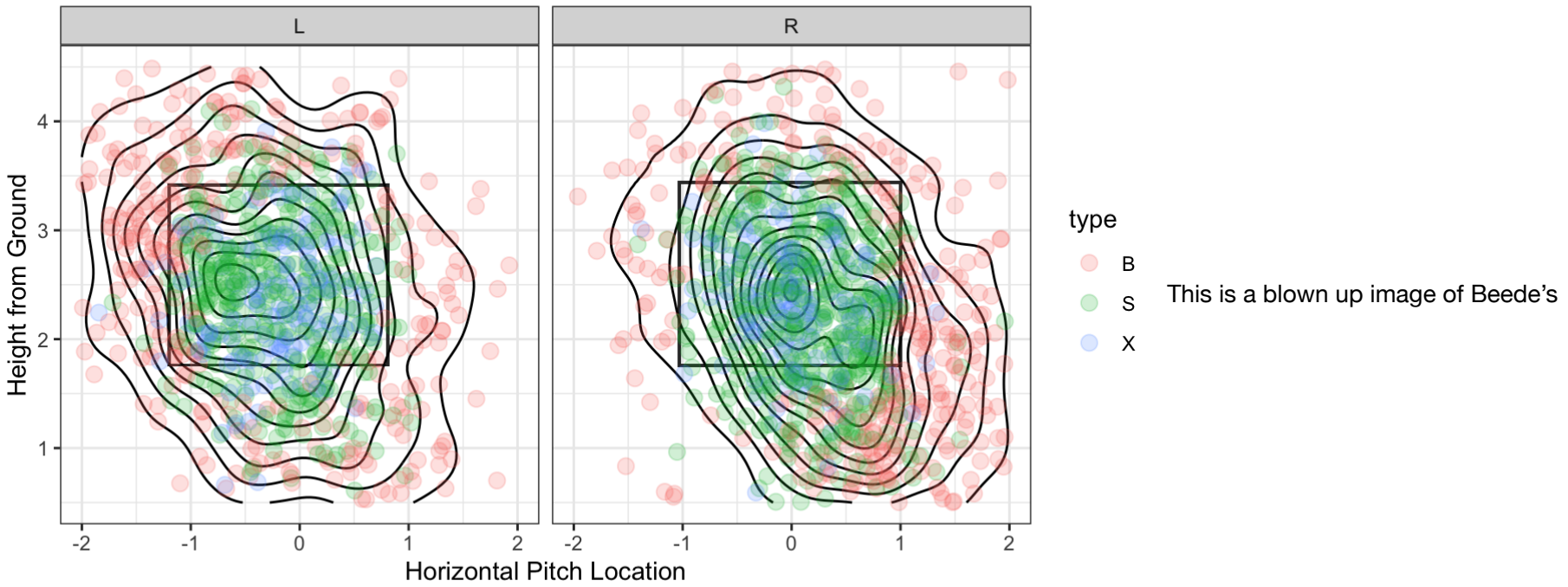
This next part is playing with heat maps on pitch locations, they denote a ball, strike, and ball in play(x). Shows visually where he throws his fastballs usually. Also shows that he mostly throws changeups and curveballs away to lefties, I assume this is because he plays at Oracle Park, which is quite difficult to hit a ball opposite field to left field as a lefty. With righties, he throws his slider away most of the time, while with lefties, it seems he throws all over the middle and lower portions of the zone.

```
strikeFX(beede, color = "type", point.alpha = 0.2,
         adjust = TRUE, contour = TRUE) + facet_grid(. ~ stand) +
  theme(legend.position = "right", legend.direction = "vertical") +
  coord_equal() + theme_bw()
```

```
## Warning in strikeFX(beede, color = "type", point.alpha = 0.2, adjust = TRUE, :
## pitchRx assumes the height of each batter is recorded as 'b_height'. Since there
## is no such column, we will assume each batter has a height of 6'2''
```

```
## Warning: Removed 171 rows containing non-finite values (stat_density2d).
```

```
## Warning: Removed 171 rows containing missing values (geom_point).
```



This is a blown up image of Beede's heat maps ve lefties and righties. The middle circle shows the higher proportion of pitches going there, and expand outwards. When he faces lefties, the majority of pitches are outside, shown by the circle on the graph, while with righties he's more down the middle or outside.

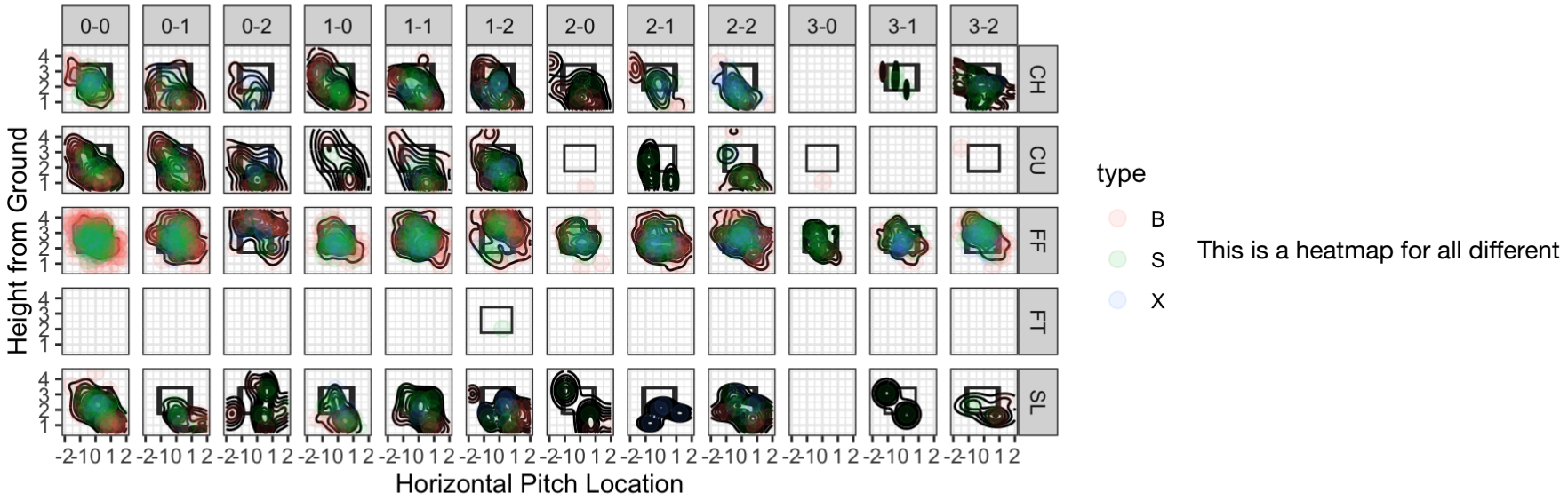
```
strikeFX(beede, color = "type", point.alpha = 0.1,
         adjust = TRUE, contour = TRUE) + facet_grid(pitch_type ~ bs_count) +
  theme(legend.position = "right", legend.direction = "vertical") +
  coord_equal() + theme_bw()
```

```
## Warning in strikeFX(beede, color = "type", point.alpha = 0.1, adjust = TRUE, :
## pitchRx assumes the height of each batter is recorded as 'b_height'. Since there
## is no such column, we will assume each batter has a height of 6'2''
```

```
## Warning: Removed 171 rows containing non-finite values (stat_density2d).
```

```
## Warning: Computation failed in `stat_density2d()`:
## missing value where TRUE/FALSE needed

## Warning: Computation failed in `stat_density2d()`:
## missing value where TRUE/FALSE needed

## Warning: Computation failed in `stat_density2d()`:
## missing value where TRUE/FALSE needed

## Warning: Computation failed in `stat_density2d()`:
## missing value where TRUE/FALSE needed
```

```
## Warning: Removed 171 rows containing missing values (geom_point).
```



This is a heatmap for all different counts and all pitches. This can be broken down for scouting reports i.e. for hitters stance. His curveball seems to be all over the zone, but with 2-2 counts, looks like he really works to keep the pitch low in the zone to get a swing.

```
# Create batter_swing
no_swing <- c("ball", "called_strike", "swinging_strike", "hit_by_pitch")
beede_ff$batter_swing <- ifelse(beede_ff$description %in% no_swing, 0, 1)

# Create swing_ff
swing_ff <- subset(beede_ff, batter_swing == 1)

# Create the contact variable
no_contact <- c("swinging_strike")
swing_ff$contact <- ifelse(swing_ff$description %in% no_contact, 0, 1)

# Create velo_bin: add one line for "Fast"
swing_ff$velo_bin <- ifelse(swing_ff$release_speed < 93.5, "Slow", NA)

swing_ff$velo_bin <- ifelse(swing_ff$release_speed >= 93.5 & swing_ff$release_speed < 95.5,
  "Medium", swing_ff$velo_bin)

swing_ff$velo_bin <- ifelse(swing_ff$release_speed >= 95.5,
  "Fast", swing_ff$velo_bin)

beede$batter_swing <- ifelse(beede$description %in% no_swing, 0,1)
swings <- subset(beede, batter_swing == 1 )

# Aggregate contact rate by velocity bin
tapply(swing_ff$release_speed, swing_ff$velo_bin, mean)
```

```
##     Fast   Medium     Slow
## 95.93261 94.35872 92.94253
```

```
# Create a new function called bin_pitch_speed() for use in calculating velo_bin.
bin_pitch_speed <- function(x) {
  cut(x, breaks = quantile(x, probs = c(0,1/3,2/3,1)), labels = FALSE)
  }


# Create the subsets for each pitch type
swing_ff <- subset(swings, pitch_type == "FF")
swing_ch <- subset(swings, pitch_type == "CH")
swing_cu <- subset(swings, pitch_type == "CU")
swing_ft <- subset(swings, pitch_type == "FT")
swing_sl <- subset(swings, pitch_type == "SL")

swing_ff$velo_bin <- bin_pitch_speed(swing_ff$release_speed)
swing_ch$velo_bin <- bin_pitch_speed(swing_ch$release_speed)
swing_cu$velo_bin <- bin_pitch_speed(swing_cu$release_speed)
swing_sl$velo_bin <- bin_pitch_speed(swing_sl$release_speed)


# Print quantile levels for each pitch
thirds <- c(0, 1/3, 2/3, 1)
quantile(swing_ff$release_speed, probs = thirds)
```

```
##        0% 33.33333% 66.66667%      100%
##      91.7      93.9      94.6      97.0
```

```
quantile(swing_ch$release_speed, probs = thirds)
```

```
##        0% 33.33333% 66.66667%      100%
##      81.7      83.5      84.3      87.0
```

```
quantile(swing_cu$release_speed, probs = thirds)
```

```
##        0% 33.33333% 66.66667%      100%
##  78.10000  80.06667  81.03333  84.40000
```

```
quantile(swing_sl$release_speed, probs = thirds)
```

```
##        0% 33.33333% 66.66667%      100%
##  82.10000  85.83333  86.66667  88.40000
```

```
#Took out two seam fastballs, very minimal amount thrown, not helpful
```

```
library(dplyr)
library(tidyr)
nl_west <-c("LAD", "SD", "COL", "ARI")
beede_nl_west <- ifelse(beede$home_team %in% nl_west, 0,1) & ifelse(beede$away_team %in% nl_west,0,1)
beede_nl_west <- subset(beede, beede_nl_west == 0)
beede_away <- subset(beede, beede$away_team == "SF")
beede_home <- subset(beede, beede$home_team == "SF")
```

Created new datasets of Beede vs the NL West, Giants playing at home and Giants playing on the road

```
beede_away_ff <- subset(beede_away, beede_away$pitch_type == "FF")
beede_home_ff <- subset(beede_home, beede_home$pitch_type == "FF")
mean(beede_away_ff$release_speed)
```

```
## [1] 94.30579
```

```
mean(beede_home_ff$release_speed)
```

```
## [1] 94.1447
```

```
ff_velo_home <- tapply(beede_home_ff$release_speed, beede_home_ff$game_date, mean)
print(ff_velo_home)
```

```
## 2019-05-14 2019-06-11 2019-06-27 2019-07-19 2019-07-24 2019-08-09 2019-08-14
##   94.43235   94.64909   93.61475   93.82000   94.05208   94.01667   93.69545
## 2019-08-26 2019-09-13 2019-09-26
##   93.97407   94.58140   95.18182
```

Broke it down ever more, Beede's velocity really took a dip in late June and early July, as well as in August. He bounced back really well and even threw harder in September.

```
ff_velo_away <- tapply(beede_away_ff$release_speed, beede_away_ff$game_date, mean)
print(ff_velo_away)
```

```
## 2019-05-03 2019-05-09 2019-05-30 2019-06-05 2019-06-17 2019-06-22 2019-07-02
##   95.35714   94.78485   93.63962   94.68600   94.88000   94.02200   93.94000
## 2019-07-14 2019-07-30 2019-08-04 2019-08-20 2019-09-02 2019-09-07 2019-09-20
##   93.83750   93.42500   94.31316   93.37708   94.38947   94.78947   95.41489
```

Not really much of a difference, but still the same trend with what we saw with velocity dipping in July and increasing in September

```
# Create bs_count_tab
bs_count_home <- table(beede_home$bs_count, beede_home$month)

# Create bs_month
bs_month_home <- round(prop.table(bs_count_home,margin =2),3)

# Print bs_month
print(bs_month_home)
```

```
##
##           5     6     7     8     9
##   0-0 0.230 0.243 0.289 0.266 0.229
##   0-1 0.082 0.153 0.144 0.125 0.143
##   0-2 0.066 0.062 0.077 0.060 0.064
##   1-0 0.098 0.062 0.088 0.101 0.064
##   1-1 0.098 0.102 0.082 0.117 0.114
##   1-2 0.131 0.073 0.098 0.117 0.093
##   2-0 0.016 0.028 0.031 0.028 0.029
##   2-1 0.066 0.068 0.041 0.048 0.071
##   2-2 0.082 0.107 0.062 0.065 0.114
##   3-0 0.000 0.000 0.005 0.012 0.007
##   3-1 0.049 0.017 0.015 0.020 0.029
##   3-2 0.082 0.085 0.067 0.040 0.043
```

```
#percentage change in rate at which Beede puts himself in these counts

# Create bs_count_tab
bs_count_away <- table(beede_away$bs_count, beede_away$month)

# Create bs_month
bs_month_away <- round(prop.table(bs_count_away,margin =2),3)

# Print bs_month
print(bs_month_away)
```

```
##
##           5     6     7     8     9
##   0-0 0.247 0.246 0.273 0.244 0.278
##   0-1 0.140 0.133 0.135 0.125 0.116
##   0-2 0.070 0.044 0.052 0.054 0.058
##   1-0 0.079 0.089 0.100 0.077 0.112
##   1-1 0.102 0.116 0.083 0.107 0.095
##   1-2 0.121 0.092 0.093 0.101 0.071
##   2-0 0.023 0.038 0.028 0.018 0.046
##   2-1 0.033 0.058 0.045 0.060 0.054
##   2-2 0.079 0.085 0.107 0.125 0.066
##   3-0 0.009 0.024 0.003 0.006 0.017
##   3-1 0.023 0.038 0.014 0.024 0.029
##   3-2 0.074 0.038 0.066 0.060 0.058
```

```
#percentage change in rate at which Beede puts himself in these counts
```

These two charts are proportions of counts Beede is in. The top is the home chart and the bottom is the away games chart. These are also broken down by month. There may be a trend in Beede getting in more friendly counts by looking at it monthly, but in baseball, I wanted to account for as much different game characteristics as possible like home and away games.

```
table(beede_home$pitch_type)
```

```
##
##  CH  CU  FF  FT  SL
## 110 133 481   1  95
```

```
table(beede_home$pitch_type, beede_home$month)
```

```
##
##        5   6   7   8   9
##   CH  12  22  28  35  13
##   CU  14  34  18  39  28
##   FF  34 116 103 152  76
##   FT   1   0   0   0   0
##   SL   0   5  45  22  23
```

```
prop.table(table(beede_home$pitch_type, beede_home$month), margin =2)
```

```
##
##               5          6          7          8          9
##   CH 0.19672131 0.12429379 0.14432990 0.14112903 0.09285714
##   CU 0.22950820 0.19209040 0.09278351 0.15725806 0.20000000
##   FF 0.55737705 0.65536723 0.53092784 0.61290323 0.54285714
##   FT 0.01639344 0.00000000 0.00000000 0.00000000 0.00000000
##   SL 0.00000000 0.02824859 0.23195876 0.08870968 0.16428571
```

```
type_tab2 <- prop.table(table(beede_home$pitch_type, beede_home$month), margin =2)
```

This is a chart with totals in pitches thrown broken down by months for only home games. This also shows proportionally the pitches thrown in home game only.

```
table(beede_away$pitch_type)
```

```
##
##  CH  CU  FF  SL
## 268 145 656 137
```

```
table(beede_away$pitch_type, beede_away$month)
```

```
##
##        5   6   7   8   9
##   CH  54  89  47  38  40
##   CU  33  48  26   7  31
##   FF 128 155 164  86 123
##   SL   0   1  52  37  47
```

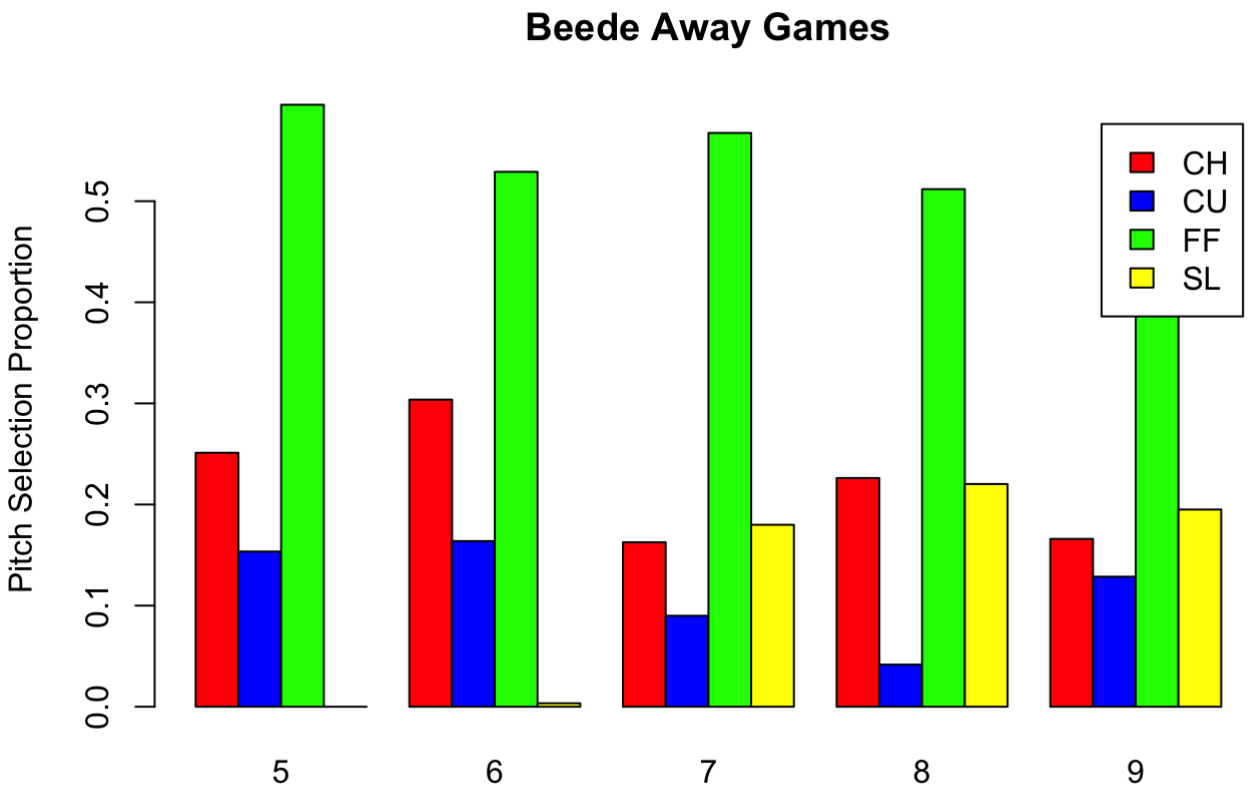
```
prop.table(table(beede_away$pitch_type, beede_away$month), margin =2)
```

```
##
##                5           6           7           8           9
##   CH 0.251162791 0.303754266 0.162629758 0.226190476 0.165975104
##   CU 0.153488372 0.163822526 0.089965398 0.041666667 0.128630705
##   FF 0.595348837 0.529010239 0.567474048 0.511904762 0.510373444
##   SL 0.000000000 0.003412969 0.179930796 0.220238095 0.195020747
```

```
type_tab3 <- prop.table(table(beede_away$pitch_type, beede_away$month), margin =2)
```

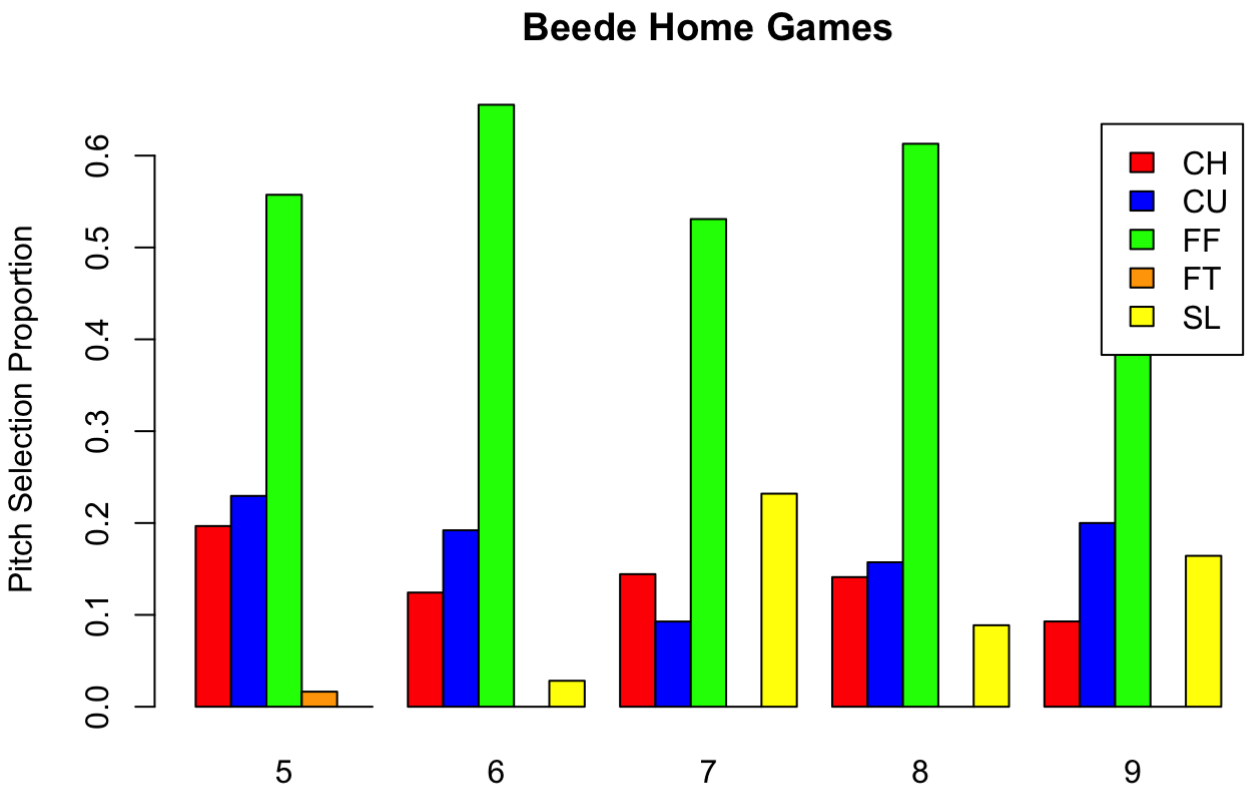
This is the same chart above, except now displaying away splits only. This just looks like a lot of numbers, so we'll break both of these down into charts.

```
barplot(type_tab3, beside = TRUE, col = c("red", "blue", "green", "yellow"),
        main = "Beede Away Games",
        ylab = "Pitch Selection Proportion",
        legend = rownames(type_tab3))
```



This is the graphic representation for the away games. Beede starts off the season with a very high fastball usage in his away games and it fluctuates, and by the time he gets to September, it is his lowest usage of his fastball. Beede says in Kerry Crowley's article that his might have explained his untapped velocity he found in September. After barely throwing a slider in May and June, throughout the summer he threw a lot more sliders.

```
barplot(type_tab2, beside = TRUE, col = c("red", "blue", "green", "orange","yellow"),
        main = "Beede Home Games",
        ylab = "Pitch Selection Proportion",
        legend = rownames(type_tab2))
```



This is the home games graph

showing the same, except he did throw his fastball more in home games than away games. This may be due to the Oracle Park factor with it being a notorious pitchers park in the National League. He also threw much more off-speed in away games when compared to home. Changeup's were the most notable difference. He threw his slider much more in July than any other month and did not even throw one in May.

```
beede <- mutate(beede, newvar=ifelse(beede$home_team == "SF", "home", "away"))
names(beede)[99] <- paste("location")
tab_type <- table(beede$location, beede$pitch_type)
type_tab_away <- prop.table(table(beede$pitch_type, beede$location), margin=2)
print(type_tab_away)
```

```
##
##             away        home
##   CH 0.222222222 0.134146341
##   CU 0.120232172 0.162195122
##   FF 0.543946932 0.586585366
##   FT 0.000000000 0.001219512
##   SL 0.113598673 0.115853659
```

```
print(tab_type)
```

```
##
##         CH  CU  FF  FT  SL
##   away 268 145 656   0 137
##   home 110 133 481   1  95
```

```
tab_type[,-4]
```

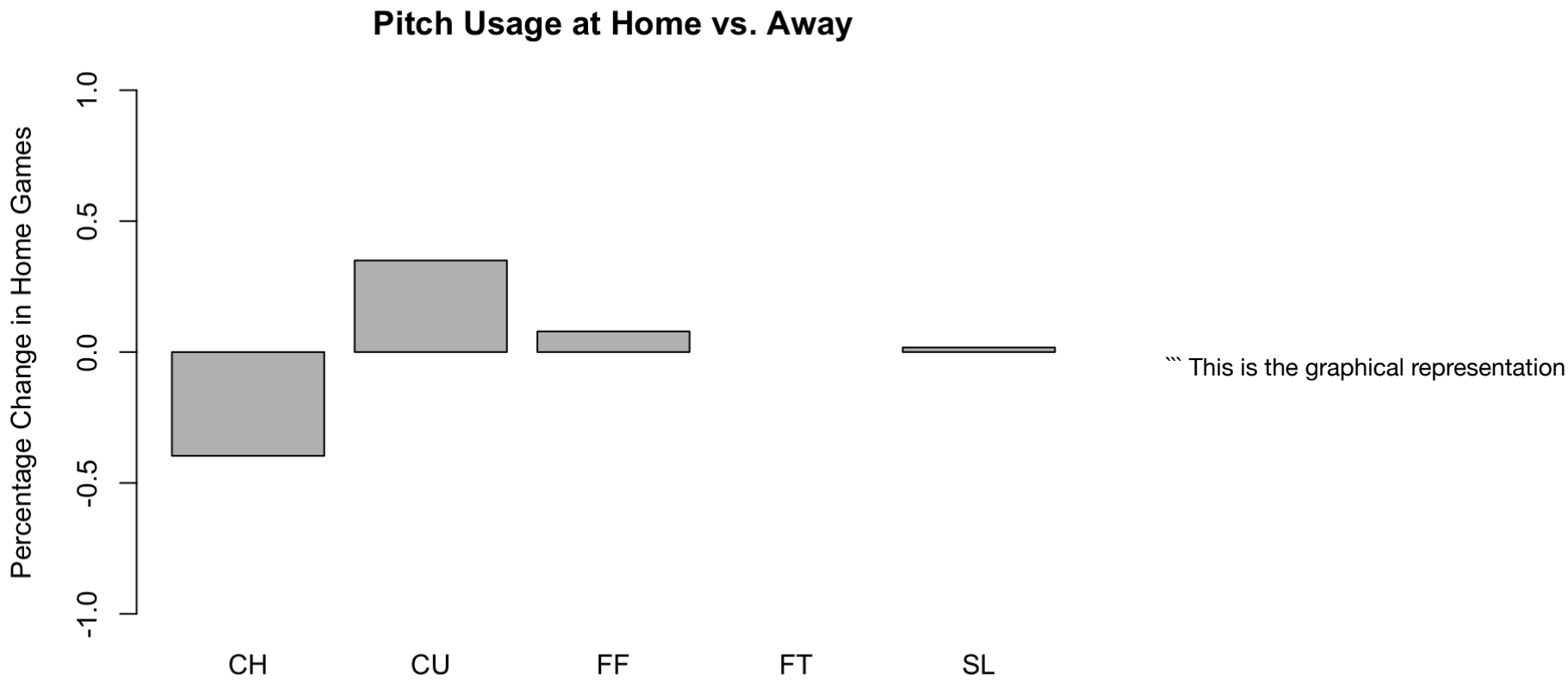
```
##
##         CH  CU  FF  SL
##   away 268 145 656 137
##   home 110 133 481  95
```

```
type_prop_dif <- round(prop.table(tab_type, margin = 2), 3)
print(type_prop_dif)
```

```
##
##           CH    CU    FF    FT    SL
##   away 0.709 0.522 0.577 0.000 0.591
##   home 0.291 0.478 0.423 1.000 0.409
```

These are the proportions comparing home and away games. The proportions on top are of his pitch arsenal during home and away games. His slider seems to be pretty much the same, while his fastball and curveball are thrown more at home than away. Also a huge difference in changeups between the home and away games.

```
home <- c(0.134, 0.162,0.587,0.0012,0.116)
away <- c(0.222, 0.120, 0.544,0.000,0.114)
type_prop_location <- data.frame(Pitch, home, away)
type_prop_location$Difference <- (type_prop_location$home - type_prop_location$away) / type_prop_location$away




# Plot a barplot
barplot(type_prop_location$Difference, names.arg = type_prop_location$Pitch,
        main = "Pitch Usage at Home vs. Away",
        ylab = "Percentage Change in Home Games",
        ylim = c(-1, 1))
```




``` This is the graphical representation

of the Pitch usage differences in Home games compared to away games. As mentioned, he threw his changeup a lot less and his curveball a lot more. There is minimal increases in Fastball and slider usage at home as well.

In comparison, using baseball savant, Beede was 1-4 with a 3.75 era in 50.1 IP, while he was 4-6 with a 6.08 era in 66.2 IP. That Oracle Park advantage is a huge factor for pitchers and is one of the reasons Beede pitches better there.

Pending any new information about the new 2020 season, Tyler Beede can have a big rebound season for the San Francisco Giants. Altering his pitch usages may be a huge factor for him to find more success, similar to a Stephen Strasburg. Beede also had more success when he was in more friendly counts and walking less batters. In July, Beede had a total of 4 walks with a 3.34 era. In contrast, Beede had 17 walks in June with a 5.76 era.

I'm sure I am as eager as many others to have some sort of baseball in any form this season, but as we wait patiently, the game can still be broken down and we all can make improvements. And if there is a season, I am really looking forward to being back out there on the Giants groundscrew and hopefully seeing Tyler Beede dominate batters all year long.
```

Credits: https://www.mercurynews.com/2020/02/19/stephen-strasburg-how-sf-giants-are-using-analytics-to-change-tyler-beedes-approach-on-the-mound/ (https://www.mercurynews.com/2020/02/19/stephen-strasburg-how-sf-giants-are-using-analytics-to-change-tyler-beedes-approach-on-the-mound/), written by: Kerry Crowley https://baseballsavant.mlb.com/savant-player/tyler-beede-595881?stats=splits-r-pitching-mlb&season=2019 (https://baseballsavant.mlb.com/savant-player/tyler-beede-595881?stats=splits-r-pitching-mlb&season=2019)