

# Fielding Optim BigelowBricklayers

Jiyue (Tom) Yang

3/8/2021

## Load Packages

```
library(dplyr)
library(ggplot2)
library(tibble)
```

## File Import

This imports the data into a dataframe and calls `head()` in order to get a front-sided view of the data.

```
fielding_data_df <- readr::read_csv("/Users/proje/Desktop/CollegeStuff/Math - Baseball/Baseball Fielding")

## Parsed with column specification:
## cols(
##   .default = col_double(),
##   BatSide = col_character(),
##   PitcherHand = col_character(),
##   PitchTypeCode = col_character(),
##   Trajectory = col_character()
## )

## See spec(...) for full column specifications.

fielding_data_df %>% head()

## # A tibble: 6 x 31
##   Game_PK pitcherid batterid BatSide PitcherHand inning PitchTypeCode
##   <dbl>     <dbl>     <dbl> <chr>    <chr>      <dbl> <chr>
## 1  566294     621242    614177 R         R           9  SL
## 2  530509     548389    546991 R         R           4  KC
## 3  566391     474463    592387 R         L           4  FF
## 4  530365     572140    435064 R         L           3  FT
## 5  529436     533167    543760 R         R           1  FF
## 6  566786     571946    608596 R         R           7  CU
## # ... with 24 more variables: ExitVelocity <dbl>, VertAngle <dbl>,
## #   HorizAngle <dbl>, Distance <dbl>, FlightTime <dbl>, Trajectory <chr>,
## #   HitValue <dbl>, RunValue <dbl>, X3 <dbl>, Y3 <dbl>, X4 <dbl>, Y4 <dbl>,
## #   X5 <dbl>, Y5 <dbl>, X6 <dbl>, Y6 <dbl>, X7 <dbl>, Y7 <dbl>, X8 <dbl>,
## #   Y8 <dbl>, X9 <dbl>, Y9 <dbl>, ballpos_x <dbl>, ballpos_y <dbl>
```

## Glimpsing the Data

This calls `glimpse()` in order to get all variables and their datatypes.

```
fielding_data_df %>% glimpse()
```

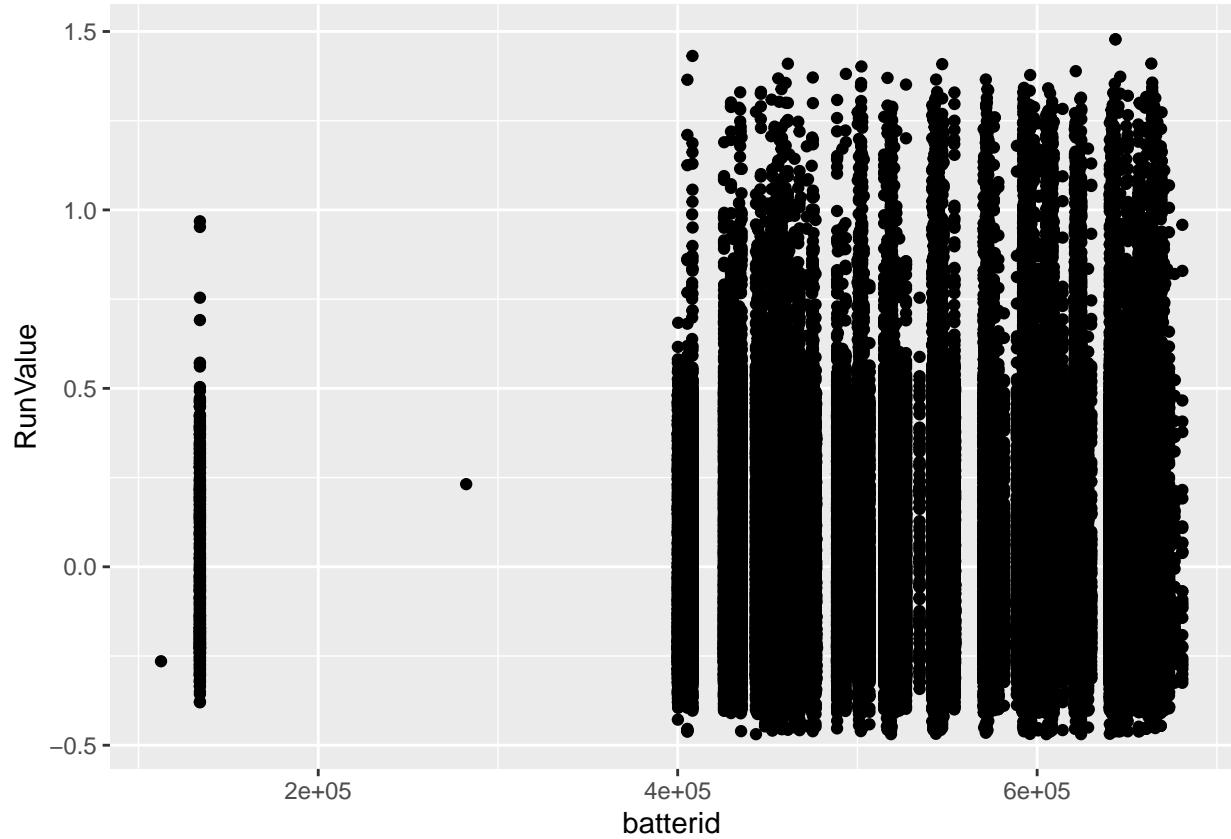
```
## Rows: 258,088
## Columns: 31
## $ Game_PK      <dbl> 566294, 530509, 566391, 530365, 529436, 566786, 56595...
## $ pitcherid    <dbl> 621242, 548389, 474463, 572140, 533167, 571946, 60545...
## $ batterid     <dbl> 614177, 546991, 592387, 435064, 543760, 608596, 60615...
## $ BatSide      <chr> "R", "R", "R", "R", "R", "R", "R", "L", "L", "L"...
## $ PitcherHand   <chr> "R", "R", "L", "L", "R", "R", "R", "R", "R", "R"...
## $ inning        <dbl> 9, 4, 4, 3, 1, 7, 2, 8, 9, 5, 8, 9, 9, 2, 3, 4, 9, 1, ...
## $ PitchTypeCode <chr> "SL", "KC", "FF", "FT", "FF", "CU", "SL", "CH", "SL", ...
## $ ExitVelocity  <dbl> 102.22041, 57.86483, 103.44940, 88.49261, 100.06711, ...
## $ VertAngle      <dbl> 11.0336704, 49.6903725, 9.7220182, 23.2390881, 21.229...
## $ HorizAngle    <dbl> 99.73433, 59.35676, 89.14812, 94.60698, 84.66924, 111...
## $ Distance       <dbl> 203.868, 138.340, 241.327, 324.016, 348.109, 266.253, ...
## $ FlightTime    <dbl> 1.648, 3.973, 2.024, 4.071, 3.794, 3.150, 0.450, 4.91...
## $ Trajectory    <chr> "L", "P", "L", "L", "L", "L", "G", "F", "P", "F", "L"...
## $ HitValue       <dbl> 1, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 1, 2, ...
## $ RunValue       <dbl> 0.4383, -0.1256, 0.4308, -0.1581, 0.2825, 0.1377, -0....
## $ X3             <dbl> 58.7550, 60.4970, 57.7930, 63.2240, 55.4730, 59.6030, ...
## $ Y3             <dbl> 105.0660, 98.6230, 99.6620, 103.2320, 101.4620, 98.79...
## $ X4             <dbl> 16.9300, 18.5170, 30.7560, 27.4560, 21.1760, 30.0930, ...
## $ Y4             <dbl> 142.9930, 156.9040, 150.1480, 155.5090, 154.4670, 151...
## $ X5             <dbl> -71.2920, -65.8170, -83.6160, -72.8840, -73.0440, -70...
## $ Y5             <dbl> 83.7640, 91.1210, 104.2720, 93.6510, 94.0250, 92.7400...
## $ X6             <dbl> -40.4420, -42.3550, -44.6200, -55.1240, -51.3480, -49...
## $ Y6             <dbl> 130.4720, 147.0490, 146.4070, 139.6090, 141.4430, 143...
## $ X7             <dbl> -135.8070, -130.7820, -142.5570, -127.2930, -130.5630...
## $ Y7             <dbl> 283.6480, 261.0520, 272.3400, 265.8880, 294.9030, 266...
## $ X8             <dbl> 27.6290, 34.3070, -5.2660, 16.8030, 39.9160, -11.2190...
## $ Y8             <dbl> 344.4420, 321.5150, 334.8550, 323.2660, 326.9590, 329...
## $ X9             <dbl> 151.6090, 159.7350, 140.4290, 139.4220, 157.1300, 141...
## $ Y9             <dbl> 272.3630, 238.7830, 270.6160, 241.6650, 237.4160, 265...
## $ ballpos_x      <dbl> -47.1386419, 103.9761927, 22.9863063, -11.6780055, 93...
## $ ballpos_y      <dbl> 198.343219, 91.251918, 240.229647, 323.805846, 335.45...
```

## Viewing the range of RunValue

This is a plotting of the range of *RunValue* in order to get an understanding of what kind of spread is available.

```
fielding_data_df %>%
  ggplot(mapping = aes(x = batterid, y = RunValue)) +
  geom_point()
```

```
## Warning: Removed 2 rows containing missing values (geom_point).
```



The spread of this data mostly falls between 0.5 and -0.5, with some spread over 0.5 into 1.0

### Counting number of players

This gets a count of the number of different entires listed by player.

```
fielding_data_df %>%
  count(batterid) %>%
  arrange(desc(n))
```

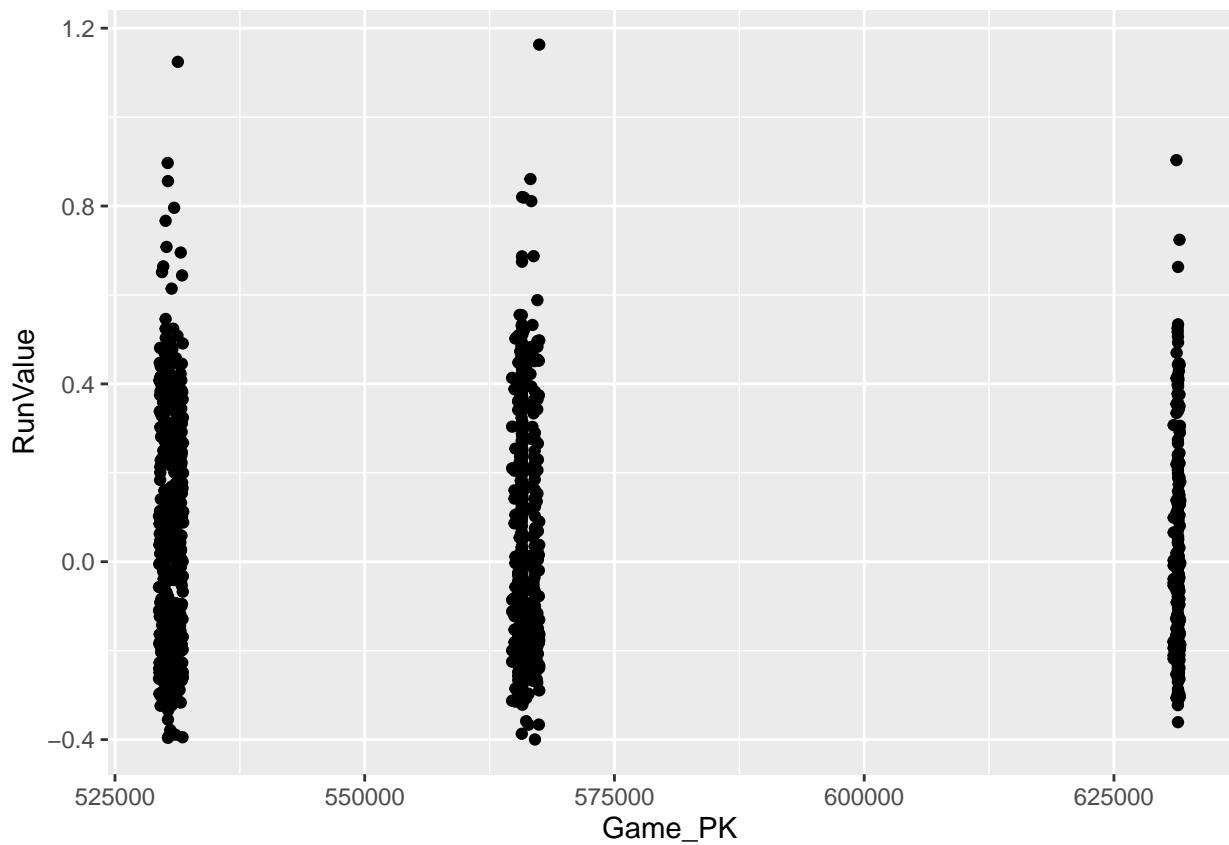
```
## # A tibble: 1,207 x 2
##   batterid     n
##       <dbl> <int>
## 1 593160    1187
## 2 543760    1088
## 3 488726    1061
## 4 596019    1049
## 5 516416    1039
## 6 514917    1036
## 7 493329    1034
## 8 453568    1033
## 9 518692    1032
## 10 592518   1020
## # ... with 1,197 more rows
```

## Getting one player's entires

This graphs the *RunValue* of player 593160, as they had the most number of entires in the table to see a potential spread.

```
player593160_data <- fielding_data_df %>%
  filter(batterid == 593160)

player593160_data %>%
  ggplot(mapping = aes(x = Game_PK, y = RunValue)) +
  geom_point()
```



```
player593160_data %>%
  summary()

##      Game_PK      pitcherid      batterid      BatSide
##  Min.   :529408   Min.   :112526   Min.   :593160  Length:1187
##  1st Qu.:530974  1st Qu.:518617  1st Qu.:593160  Class  :character
##  Median :565706  Median :592644  Median :593160  Mode   :character
##  Mean   :563142  Mean   :564930  Mean   :593160
##  3rd Qu.:566903  3rd Qu.:621244  3rd Qu.:593160
##  Max.   :631645  Max.   :677976  Max.   :593160
##
##      PitcherHand      inning      PitchTypeCode      ExitVelocity
##  Length:1187      Min.   : 1.000  Length:1187      Min.   : 9.293
```

```

##  Class :character  1st Qu.: 2.000  Class :character  1st Qu.: 79.030
##  Mode  :character Median : 5.000  Mode  :character Median : 89.745
##               Mean   : 4.492  Mean   : 86.147
##               3rd Qu.: 7.000  3rd Qu.: 95.917
##               Max.   :14.000  Max.   :106.362
##
##    VertAngle        HorizAngle       Distance      FlightTime
##  Min.   :-72.726  Min.   :-88.65  Min.   : 1.651  Min.   :0.0090
##  1st Qu.: 1.327   1st Qu.: 80.08  1st Qu.: 47.373  1st Qu.:0.4042
##  Median : 15.024  Median : 92.93  Median :191.611  Median :2.1560
##  Mean   : 15.319  Mean   : 94.14  Mean   :177.705  Mean   :2.4848
##  3rd Qu.: 29.130  3rd Qu.:108.93  3rd Qu.:284.921  3rd Qu.:4.3495
##  Max.   : 82.918  Max.   :257.64  Max.   :408.470  Max.   :7.1680
##  NA's   :1
##
##    Trajectory        HitValue       RunValue      X3
##  Length:1187  Min.   :0.0000  Min.   :-0.39970  Min.   :34.30
##  Class :character  1st Qu.:0.0000  1st Qu.:-0.18465  1st Qu.:53.97
##  Mode  :character  Median :0.0000  Median :-0.02480  Median :57.94
##               Mean   :0.4583  Mean   : 0.03721  Mean   :56.98
##               3rd Qu.:1.0000  3rd Qu.: 0.23485  3rd Qu.:60.48
##               Max.   :3.0000  Max.   : 1.16290  Max.   :77.68
##
##    Y3              X4          Y4          X5
##  Min.   : 60.73  Min.   :-24.64  Min.   :100.1  Min.   :-85.85
##  1st Qu.: 89.84  1st Qu.: 20.23  1st Qu.:148.8  1st Qu.:-69.98
##  Median : 96.92  Median : 25.76  Median :151.8  Median :-64.80
##  Mean   : 91.55  Mean   : 24.89  Mean   :150.2  Mean   :-63.58
##  3rd Qu.:100.47  3rd Qu.: 30.69  3rd Qu.:154.3  3rd Qu.:-56.46
##  Max.   :110.93  Max.   : 63.83  Max.   :164.6  Max.   :-40.72
##
##    Y5              X6          Y6          X7
##  Min.   : 62.99  Min.   :-61.63  Min.   : 88.98  Min.   :-166.90
##  1st Qu.: 79.96  1st Qu.:-47.38  1st Qu.:139.99  1st Qu.:-137.58
##  Median : 89.42  Median :-43.80  Median :142.90  Median :-129.78
##  Mean   : 87.80  Mean   :-43.17  Mean   :142.02  Mean   :-130.07
##  3rd Qu.: 94.72  3rd Qu.:-39.64  3rd Qu.:145.64  3rd Qu.:-122.02
##  Max.   :109.82  Max.   : 17.11  Max.   :154.11  Max.   :-94.97
##
##    Y7              X8          Y8          X9
##  Min.   :233.3   Min.   :-46.68  Min.   :280.1  Min.   : 88.04
##  1st Qu.:265.3   1st Qu.: 11.08  1st Qu.:312.2  1st Qu.:133.75
##  Median :272.4   Median : 19.25  Median :318.3  Median :140.47
##  Mean   :272.4   Mean   : 16.07  Mean   :318.6  Mean   :140.61
##  3rd Qu.:279.8   3rd Qu.: 25.08  3rd Qu.:325.8  3rd Qu.:147.55
##  Max.   :317.4   Max.   : 54.62  Max.   :352.7  Max.   :180.02
##
##    Y9          ballpos_x      ballpos_y
##  Min.   :207.6   Min.   :-228.6721  Min.   :-89.79
##  1st Qu.:244.6   1st Qu.: -26.6350  1st Qu.: 41.88
##  Median :253.8   Median :  0.2523  Median :171.90
##  Mean   :252.7   Mean   : 12.4247  Mean   :162.86
##  3rd Qu.:260.9   3rd Qu.:  62.2363  3rd Qu.:258.05
##  Max.   :289.3   Max.   : 214.6590  Max.   :407.30
##

```