DATA607 TidyVerse Assignment

Peter Gatica

April 12 2021

Tidyverse Assignment Requirements - Tidyverse CREATE Assignment (25 points)

1. Clone the provided repository (1 point) *
2. Write a vignette using one TidyVerse package (15 points) *
3. Write a vignette using more than one TidyVerse packages (+ 2 points) *
4. Make a pull request on the shared repository (1 point)
5. Update the README.md file with your example (2 points)
6. Submit your GitHub handle name & link to Peergrade (1 point)
7. Grade your 3 peers and provide the feedback in Peergrade (2 points)
8. Submit the best peer link & your link to Blackboard (1 point)

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.3     v purrr   0.3.4
## v tibble  3.0.6     v dplyr   1.0.2
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(httr)
library(jsonlite)
```

```
##
## Attaching package: 'jsonlite'

## The following object is masked from 'package:purrr':
##
##     flatten
```

```
library(RCurl)
```

```
##
## Attaching package: 'RCurl'

## The following object is masked from 'package:tidyr':
##
##     complete
```

Vignette 1 - jsonlite and hittr

These are two packages within tidyverse that I really enjoyed using to extract data from a New York Times API.

It is important to code defensively by capturing and processing error codes. To test if the code will fail if the status code is anything other than 200, uncomment the get_status code

```
api_call_return <- GET(url)
(get_status <- api_call_return$status_code)
```

```
## [1] 200
```

```
# Uncomment to test stop_for_status for API call error
#get_status<- 404
```

Test the return status and extract the api result set for processing

```
if (get_status != 200) {
  stop_for_status(get_status)
}
api_call_header <- headers(api_call_return)
api_call_parsed <- content(api_call_return,"parse")
results_list <- api_call_parsed[["results"]]
head(results_list,1)
```

```
## [[1]]
## [[1]]$display_title
## [1] "Shiva Baby"
##
## [[1]]$mpaa_rating
## [1] ""
##
## [[1]]$critics_pick
## [1] 1
##
## [[1]]$byline
## [1] "Jason Bailey"
##
## [[1]]$headline
## [1] "'Shiva Baby' Review: It's Complicated"
##
## [[1]]$summary_short
## [1] "The potential land mines of a young woman's life are set to explode simultaneously in this tens
##
## [[1]]$publication_date
## [1] "2021-04-08"
##
## [[1]]$opening_date
## [1] "2021-04-02"
##
## [[1]]$date_updated
```

```
## [1] "2021-04-08 15:29:03"
##
## [[1]]$link
## [[1]]$link$type
## [1] "article"
##
## [[1]]$link$url
## [1] "https://www.nytimes.com/2021/04/08/movies/shiva-baby-review.html"
##
## [[1]]$link$suggested_link_text
## [1] "Read the New York Times Review of Shiva Baby"
##
##
## [[1]]$multimedia
## [[1]]$multimedia$type
## [1] "mediumThreeByTwo210"
##
## [[1]]$multimedia$src
## [1] "https://static01.nyt.com/images/2021/04/10/arts/shiva1/shiva1-mediumThreeByTwo440.jpg"
##
## [[1]]$multimedia$height
## [1] 140
##
## [[1]]$multimedia$width
## [1] 210
```

Vignette 2 - ggplot, rCURL

These are two packages within tidyverse. rCURL allow a source to be retrieved from a website while ggplot gives the ability to create numerous types of graphs to visualize your data with.

For this vignette I am using a subset of the baseball.csv from https://www.kaggle.com/danielmontilla/baseball-databank. The subset file contains 62 players from 2019 MLB season and I will calculate batting average by position for 2019 as well as visualize the top player batting averages for 2019 by position type (outfield, infield and catcher). The catcher position is not considered an infield nor outfield position. The catcher is the only postion positioned in foul territory. Plus the catcher is the only player that can see the entire field. How many of you new this?

```
# Source the subset file from the Kaggle Website from my github repository
filename <- getURL("https://raw.githubusercontent.com/audiorunner13/Masters-Coursework/main/DATA607%20S
batting_by_posit <- read.csv(text = filename,na.strings = "")
batting_by_posit
```

```
##     playerID yearID stint teamID lgID    posit   G  AB   H   avg   R X2B X3B HR
## 1  sanchga02   2019     1    NYA   AL  catcher 106 396  92 0.232  62  12   1 34
## 2  chiriro01   2019     1    HOU   AL  catcher 114 366  87 0.238  57  22   1 17
## 3  mathije01   2019     1    TEX   AL  catcher  88 228  36 0.158  17   9   0  2
## 4  contrwi01   2019     1    CHN   NL  catcher 105 360  98 0.272  57  18   2 24
## 5  vazquch01   2019     1    BOS   AL  catcher 138 482 133 0.276  66  26   1 23
## 6  mccanbr01   2019     1    ATL   NL  catcher  85 277  69 0.249  28   9   0 12
## 7  martiru01   2019     1    LAN   NL  catcher  83 209  46 0.220  29   5   0  6
## 8  severpe01   2019     1    BAL   AL  catcher  96 305  76 0.249  37  13   0 13
```

```
## 9  realmjt01   2019     1    PHI   NL  catcher 145 538 148 0.275  92  36   3 25
## 10 gomesya01   2019     1    WAS   NL  catcher  97 314  70 0.223  36  16   0 12
## 11 suzukku01   2019     1    WAS   NL  catcher  85 280  74 0.264  37  11   0 17
## 12 flowety01   2019     1    ATL   NL  catcher  85 271  62 0.229  36  11   3 11
## 13 kellyca02   2019     1    ARI   NL  catcher 111 314  77 0.245  46  19   0 18
## 14 perezro02   2019     1    CLE   AL  catcher 119 389  93 0.239  46   9   1 24
## 15 judgeaa01   2019     1    NYA   AL outfield 102 378 103 0.272  75  18   1 27
## 16 springe01   2019     1    HOU   AL outfield 122 479 140 0.292  96  20   3 39
## 17 gardnbr01   2019     1    NYA   AL outfield 141 491 123 0.251  86  26   7 28
## 18 brantmi02   2019     1    HOU   AL outfield 148 575 179 0.311  88  40   2 22
## 19 reddijo01   2019     1    HOU   AL outfield 141 501 138 0.275  57  19   3 14
## 20  choosh01   2019     1    TEX   AL outfield 151 563 149 0.265  93  31   2 24
## 21 santada01   2019     1    TEX   AL outfield 130 474 134 0.283  81  23   6 28
## 22 almoral01   2019     1    CHN   NL outfield 130 339  80 0.236  41  11   1 12
## 23 heywaja01   2019     1    CHN   NL outfield 147 513 129 0.251  78  20   4 21
## 24 schwaky01   2019     1    CHN   NL outfield 155 529 132 0.250  82  29   3 38
## 25 beninan01   2019     1    BOS   AL outfield 138 541 144 0.266  72  40   5 13
## 26 bettsmo01   2019     1    BOS   AL outfield 150 597 176 0.295 135  40   5 29
## 27 bradlja02   2019     1    BOS   AL outfield 147 494 111 0.225  69  28   3 21
## 28 acunaro01   2019     1    ATL   NL outfield 156 626 175 0.280 127  22   2 41
## 29 culbech01   2019     1    ATL   NL outfield 108 135  35 0.259  14   5   2  5
## 30 joycema01   2019     1    ATL   NL outfield 129 200  59 0.295  32  10   0  7
## 31 markani01   2019     1    ATL   NL outfield 116 414 118 0.285  61  25   2  9
## 32 pederjo01   2019     1    LAN   NL outfield 149 450 112 0.249  83  16   3 36
## 33 verdual01   2019     1    LAN   NL outfield 106 343 101 0.294  43  22   2 12
## 34 mancitr01   2019     1    BAL   AL outfield 154 602 175 0.291 106  38   2 35
## 35 eatonad02   2019     1    WAS   NL outfield 151 566 158 0.279 103  25   7 15
## 36 roblevi01   2019     1    WAS   NL outfield 155 546 139 0.255  86  33   3 17
## 37  sotoju01   2019     1    WAS   NL outfield 150 542 153 0.282 110  32   5 34
## 38 jonesad01   2019     1    ARI   NL outfield 137 485 126 0.260  66  25   1 16
## 39  baezja01   2019     1    CHN   NL  infield 138 531 149 0.281  89  38   4 29
## 40  boteda01   2019     1    CHN   NL  infield 127 303  78 0.257  47  17   0 11
## 41 bryankr01   2019     1    CHN   NL  infield 147 543 153 0.282 108  35   1 31
## 42 rizzoan01   2019     1    CHN   NL  infield 146 512 150 0.293  89  29   3 27
## 43 lemahdj01   2019     1    NYA   AL  infield 145 602 197 0.327 109  33   2 26
## 44 torregl01   2019     1    NYA   AL  infield 144 546 152 0.278  96  26   0 38
## 45 urshegi01   2019     1    NYA   AL  infield 132 442 139 0.314  73  34   0 21
## 46  voitlu01   2019     1    NYA   AL  infield 118 429 113 0.263  72  21   1 21
## 47 altuvjo01   2019     1    HOU   AL  infield 124 500 149 0.298  89  27   3 31
## 48 bregmal01   2019     1    HOU   AL  infield 156 554 164 0.296 122  37   2 41
## 49 gourryu01   2019     1    HOU   AL  infield 144 564 168 0.298  85  40   2 31
## 50 andruel01   2019     1    TEX   AL  infield 147 600 165 0.275  81  27   4 12
## 51 forsylo01   2019     1    TEX   AL  infield 101 317  72 0.227  38  17   1  7
## 52  odorro01   2019     1    TEX   AL  infield 145 522 107 0.205  77  30   1 30
## 53 bogaexa01   2019     1    BOS   AL  infield 155 614 190 0.309 110  52   0 33
## 54 deverra01   2019     1    BOS   AL  infield 156 647 201 0.311 129  54   4 32
## 55 albieoz01   2019     1    ATL   NL  infield 160 640 189 0.295 102  43   8 24
## 56 donaljo02   2019     1    ATL   NL  infield 155 549 142 0.259  96  33   0 37
## 57 freemfr01   2019     1    ATL   NL  infield 158 597 176 0.295 113  34   2 38
## 58 swansda01   2019     1    ATL   NL  infield 127 483 121 0.251  77  26   3 17
## 59 adamsma01   2019     1    WAS   NL  infield 111 310  70 0.226  42  14   0 20
## 60 doziebr01   2019     1    WAS   NL  infield 135 416  99 0.238  54  20   0 20
## 61 kendrho01   2019     1    WAS   NL  infield 121 334 115 0.344  61  23   1 17
## 62 rendoan01   2019     1    WAS   NL  infield 146 545 174 0.319 117  44   3 34
```

```
##    RBI SB CS  BB  SO IBB HBP SH SF GIDP
## 1   77  0  1  40 125   3   9  0  1    3
## 2   58  1  2  51 125   1  13  2  5   11
## 3   12  1  0  15  87   1   0  0  1    2
## 4   64  1  2  38 102   2   9  0  2    4
## 5   72  4  2  33 101   3   0  3  3   17
## 6   45  0  0  31  53   1   2  0  6   10
## 7   20  1  0  30  60   3   8  0  2    1
## 8   44  3  1  29  73   0   4  1  2    5
## 9   83  9  1  41 123   2   5  0  8   12
## 10  43  2  0  38  84   6   5  0  1    7
## 11  63  0  1  20  36   1   6  0  3   10
## 12  34  0  0  31 105   3   6  0  2    8
## 13  47  0  0  48  79  10   2  0  1   11
## 14  63  0  0  45 127   1   4  7  4   12
## 15  55  3  2  64 141   4   3  0  1   11
## 16  96  6  2  67 113   1   6  0  4   12
## 17  74 10  2  52 108   0   4  0  3    6
## 18  90  3  2  51  66   3   7  0  4   21
## 19  56  5  2  36  66   1   0  1  9    9
## 20  61 15  1  78 165   3  18  0  1    6
## 21  81 21  6  25 151   2   6  0  5    8
## 22  32  2  1  16  62   4   1  5  2    8
## 23  62  8  3  68 110   5   5  0  3   12
## 24  92  2  3  70 156   5   5  0  6    6
## 25  68 10  3  59 140   1   7  3  5    6
## 26  80 16  3  97 101   6   3  0  9   11
## 27  62  8  6  56 155   3  12  3  2    6
## 28 101 37  9  76 188   4   9  0  1    8
## 29  20  0  1   6  44   0   1  1  1    5
## 30  23  0  0  38  45   0   0  0  0    3
## 31  62  2  0  47  59   1   2  0  6   11
## 32  74  1  1  50 111   2  12  0  2    4
## 33  44  4  1  26  49   1   2  0  6    8
## 34  97  1  0  63 143   3   9  0  5   22
## 35  49 15  3  65 106   0  13  9  3    8
## 36  65 28  9  35 140   3  25  6  5    6
## 37 110 12  1 108 132   3   3  0  6   11
## 38  67  2  1  31 101   2   8  0  3   15
## 39  85 11  7  28 156   3   0  0  2   16
## 40  41  5  1  44  93   4   7  0  2   11
## 41  77  4  0  74 145   1  15  0  2   10
## 42  94  5  2  71  86   3  27  0  3   15
## 43 102  5  2  46  90   0   2  1  4   14
## 44  90  5  2  48 129   3   3  1  6   10
## 45  74  1  1  25  87   1   5  0  4   13
## 46  62  0  0  71 142   2   9  0  1   12
## 47  74  6  5  41  82   0   3  1  3   19
## 48 112  5  1 119  83   2   9  0  8    9
## 49 104  5  3  37  65   2   5  0  6   12
## 50  72 31  8  34  96   1   4  0 10   16
## 51  39  2  0  44 100   0   3  0  2    8
## 52  93 11  9  52 178   2   5  1  1    4
## 53 117  4  2  76 122   2   2  0  6   11
```

```
## 54 115  8  8  48 119   7   4  1  2    8
## 55  86 15  4  54 112   6   4  0  4    2
## 56  94  4  2 100 155   2   8  0  2   13
## 57 121  6  3  87 127  11   6  0  2   17
## 58  65 10  5  51 124   2   5  1  5    7
## 59  56  0  0  20 115   1   2  0  1    7
## 60  50  3  4  61 105   2   4  0  1   11
## 61  62  2  1  27  49   1   4  0  5   11
## 62 126  5  1  80  86   8  12  0  9   13
```

The next four sections calculate the batting average for each position type

```
# create a batting_by_posit_totals data.frame and aggregate the avg by position type
batting_by_posit_totals <- aggregate.data.frame(x = batting_by_posit$avg,   # Sum by group
          by = list(batting_by_posit$posit),
          FUN = sum)
# rename fields in df
(batting_by_posit_totals <- batting_by_posit_totals %>%
   dplyr::rename("position" = Group.1, "total_bavg" = x))
```

```
##   position total_bavg
## 1  catcher      3.369
## 2  infield      6.741
## 3 outfield      6.501
```

```
# calc a prercentage field and append to repsective rows
catcher_total_bavg_rec <- batting_by_posit_totals %>% filter(batting_by_posit_totals$position == "catch
catcher_bavg <- catcher_total_bavg_rec$total_bavg/14
bat_avg <- round(catcher_bavg,3)
(catcher_total_bavg_rec <- cbind(catcher_total_bavg_rec,bat_avg))
```

```
##   position total_bavg bat_avg
## 1  catcher      3.369   0.241
```

```
# calc a prercentage field and append to repsective rows
outfield_total_bavg_rec <- batting_by_posit_totals %>% filter(batting_by_posit_totals$position == "outf
outfield_bavg <- outfield_total_bavg_rec$total_bavg/24
bat_avg <- round(outfield_bavg,3)
(outfield_total_bavg_rec <- cbind(outfield_total_bavg_rec,bat_avg))
```

```
##   position total_bavg bat_avg
## 1 outfield      6.501   0.271
```

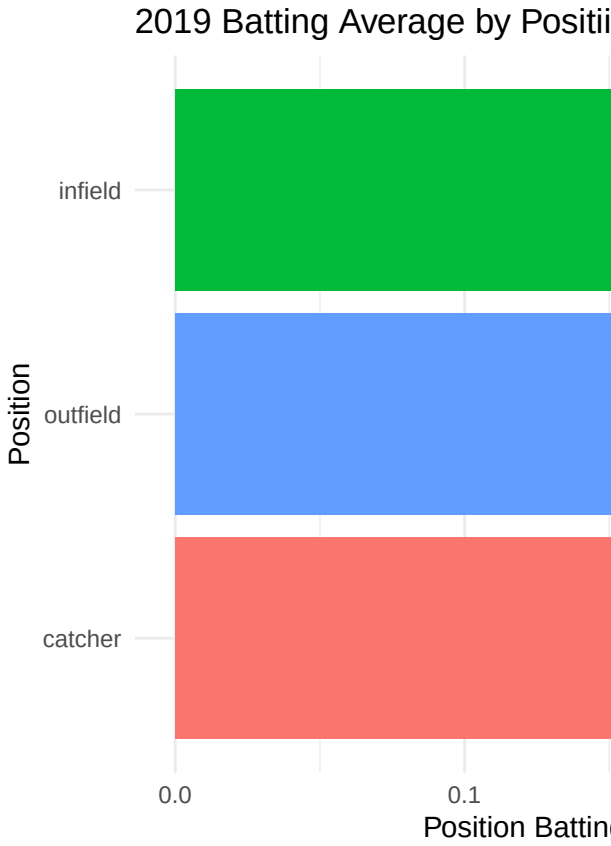
```
# calc a prercentage field and append to repsective rows
infield_total_bavg_rec <- batting_by_posit_totals %>% filter(batting_by_posit_totals$position == "infie
infield_bavg <- infield_total_bavg_rec$total_bavg/24
bat_avg <- round(infield_bavg,3)
(infield_total_bavg_rec <- cbind(infield_total_bavg_rec,bat_avg))
```

```
##   position total_bavg bat_avg
## 1  infield      6.741   0.281
```

```
final_batting_by_posit <- data.frame(c())
final_batting_by_posit <- rbind(final_batting_by_posit,catcher_total_bavg_rec)
final_batting_by_posit <- rbind(final_batting_by_posit,outfield_total_bavg_rec)
final_batting_by_posit <- rbind(final_batting_by_posit,infield_total_bavg_rec)
final_batting_by_posit
```

```
##   position total_bavg bat_avg
## 1  catcher      3.369   0.241
## 2 outfield      6.501   0.271
## 3  infield      6.741   0.281
```

```
final_batting_by_posit %>%
  ggplot(aes(y=reorder(position,bat_avg),x=bat_avg,fill=position)) +
      geom_bar(stat = 'identity',position=position_dodge()) +
      geom_text(aes(label=bat_avg), vjust=1.0, color="black",
                position = position_dodge(0.9), size=3.0) +
      labs(y = ("Position"),x = ("Position Batting Average"),
           title = ("2019 Batting Average by Positiion")) +
      theme_minimal()
```



Use ggplot to create the following graphs to visualize the data.

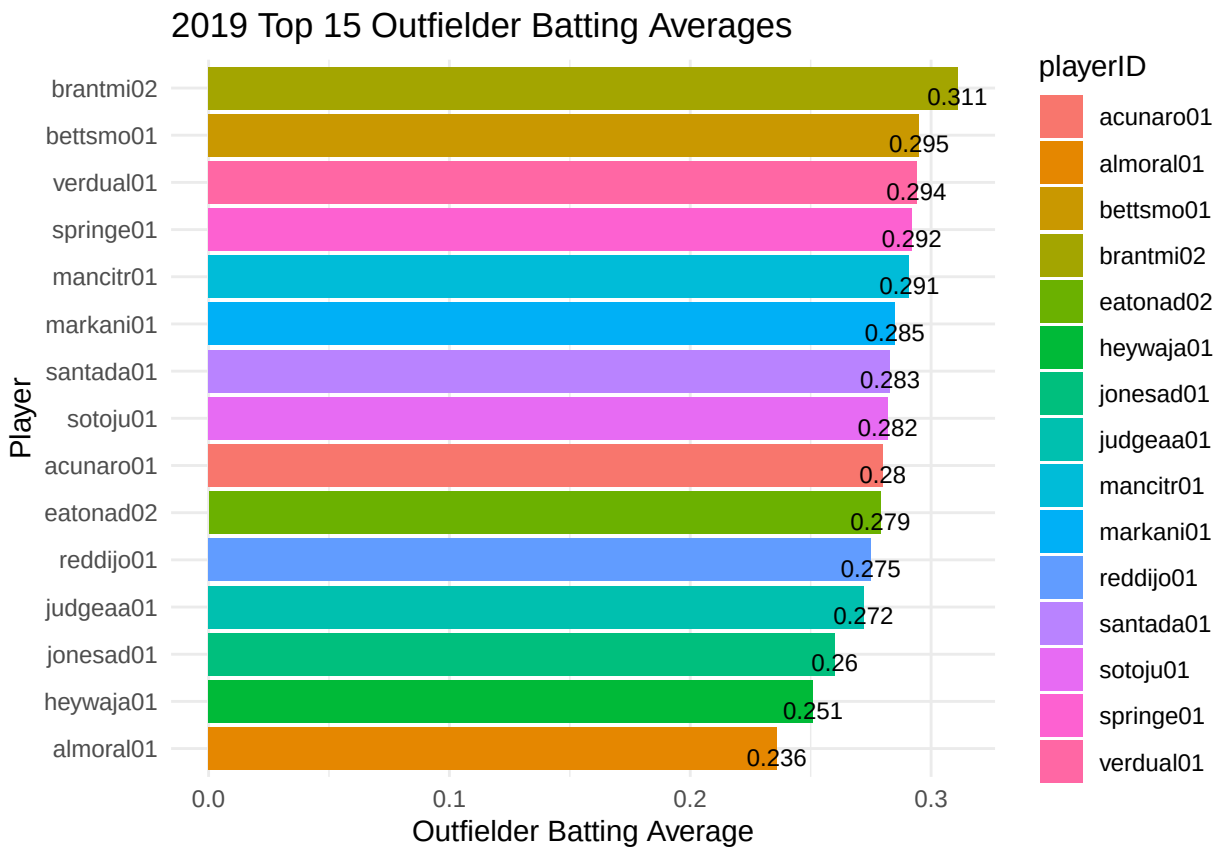
```
(outfield_players <- batting_by_posit %>% filter(batting_by_posit$posit == "outfield"))
```

```
##     playerID yearID stint teamID lgID    posit   G  AB   H   avg   R X2B X3B HR
## 1  judgeaa01   2019     1    NYA   AL outfield 102 378 103 0.272  75  18   1 27
## 2  springe01   2019     1    HOU   AL outfield 122 479 140 0.292  96  20   3 39
## 3  gardnbr01   2019     1    NYA   AL outfield 141 491 123 0.251  86  26   7 28
## 4  brantmi02   2019     1    HOU   AL outfield 148 575 179 0.311  88  40   2 22
## 5  reddijo01   2019     1    HOU   AL outfield 141 501 138 0.275  57  19   3 14
## 6   choosh01   2019     1    TEX   AL outfield 151 563 149 0.265  93  31   2 24
## 7  santada01   2019     1    TEX   AL outfield 130 474 134 0.283  81  23   6 28
## 8  almoral01   2019     1    CHN   NL outfield 130 339  80 0.236  41  11   1 12
## 9  heywaja01   2019     1    CHN   NL outfield 147 513 129 0.251  78  20   4 21
## 10 schwaky01   2019     1    CHN   NL outfield 155 529 132 0.250  82  29   3 38
## 11 beninan01   2019     1    BOS   AL outfield 138 541 144 0.266  72  40   5 13
## 12 bettsmo01   2019     1    BOS   AL outfield 150 597 176 0.295 135  40   5 29
## 13 bradlja02   2019     1    BOS   AL outfield 147 494 111 0.225  69  28   3 21
## 14 acunaro01   2019     1    ATL   NL outfield 156 626 175 0.280 127  22   2 41
## 15 culbech01   2019     1    ATL   NL outfield 108 135  35 0.259  14   5   2  5
## 16 joycema01   2019     1    ATL   NL outfield 129 200  59 0.295  32  10   0  7
## 17 markani01   2019     1    ATL   NL outfield 116 414 118 0.285  61  25   2  9
## 18 pederjo01   2019     1    LAN   NL outfield 149 450 112 0.249  83  16   3 36
## 19 verdual01   2019     1    LAN   NL outfield 106 343 101 0.294  43  22   2 12
## 20 mancitr01   2019     1    BAL   AL outfield 154 602 175 0.291 106  38   2 35
## 21 eatonad02   2019     1    WAS   NL outfield 151 566 158 0.279 103  25   7 15
## 22 roblevi01   2019     1    WAS   NL outfield 155 546 139 0.255  86  33   3 17
## 23  sotoju01   2019     1    WAS   NL outfield 150 542 153 0.282 110  32   5 34
## 24 jonesad01   2019     1    ARI   NL outfield 137 485 126 0.260  66  25   1 16
##    RBI SB CS  BB  SO IBB HBP SH SF GIDP
## 1   55  3  2  64 141   4   3  0  1   11
## 2   96  6  2  67 113   1   6  0  4   12
## 3   74 10  2  52 108   0   4  0  3    6
## 4   90  3  2  51  66   3   7  0  4   21
## 5   56  5  2  36  66   1   0  1  9    9
## 6   61 15  1  78 165   3  18  0  1    6
## 7   81 21  6  25 151   2   6  0  5    8
## 8   32  2  1  16  62   4   1  5  2    8
## 9   62  8  3  68 110   5   5  0  3   12
## 10  92  2  3  70 156   5   5  0  6    6
## 11  68 10  3  59 140   1   7  3  5    6
## 12  80 16  3  97 101   6   3  0  9   11
## 13  62  8  6  56 155   3  12  3  2    6
## 14 101 37  9  76 188   4   9  0  1    8
## 15  20  0  1   6  44   0   1  1  1    5
## 16  23  0  0  38  45   0   0  0  0    3
## 17  62  2  0  47  59   1   2  0  6   11
## 18  74  1  1  50 111   2  12  0  2    4
## 19  44  4  1  26  49   1   2  0  6    8
## 20  97  1  0  63 143   3   9  0  5   22
## 21  49 15  3  65 106   0  13  9  3    8
## 22  65 28  9  35 140   3  25  6  5    6
## 23 110 12  1 108 132   3   3  0  6   11
## 24  67  2  1  31 101   2   8  0  3   15
```

```
outfield_players %>%
  top_n(15) %>%
  ggplot(aes(y=reorder(playerID,avg),x=avg,fill=playerID)) +
```

```
      geom_bar(stat = 'identity',position=position_dodge()) +
      geom_text(aes(label=avg), vjust=1.0, color="black",
                position = position_dodge(0.9), size=3.0) +
      labs(y = ("Player"),x = ("Outfielder Batting Average"),
           title = ("2019 Top 15 Outfielder Batting Averages")) +
      theme_minimal()
```

```
## Selecting by GIDP
```



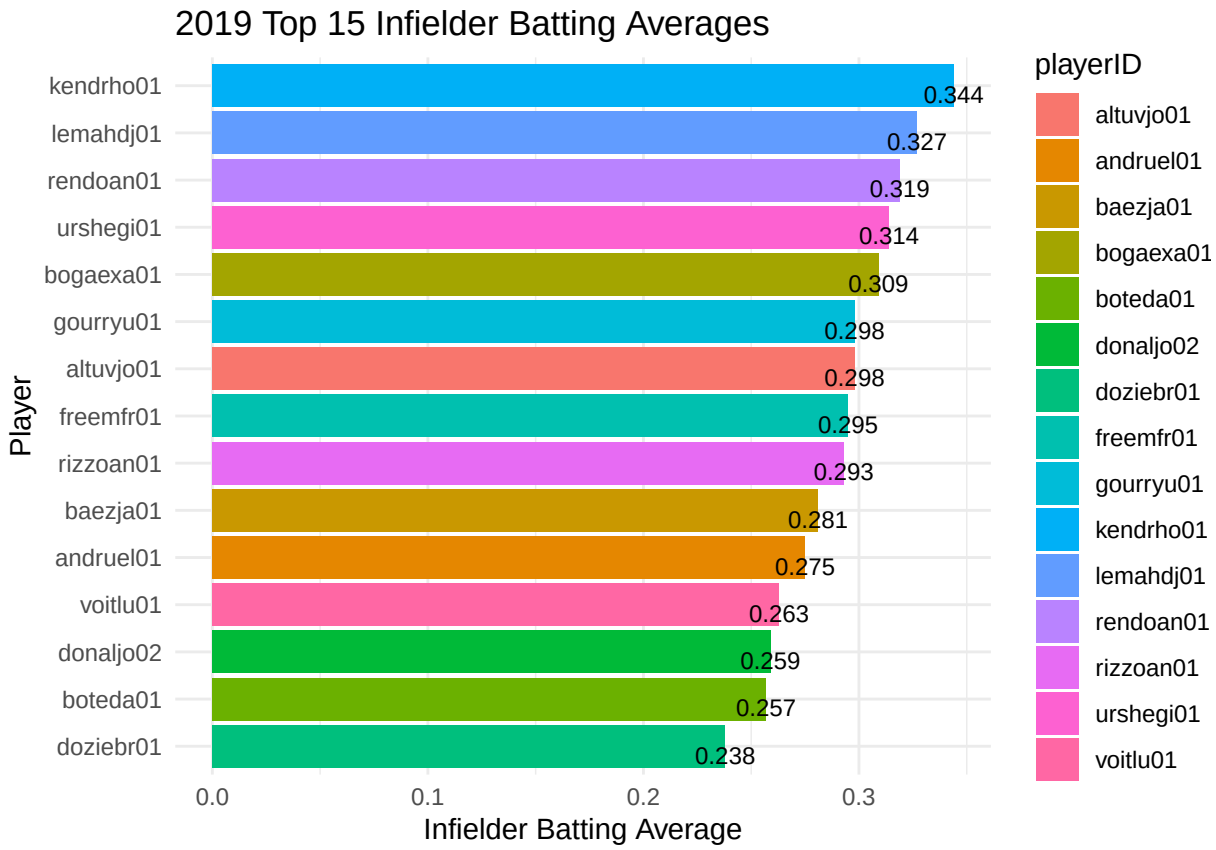

```
(infield_players <- batting_by_posit %>% filter(batting_by_posit$posit == "infield"))
```

```
##      playerID yearID stint teamID lgID   posit   G  AB   H   avg   R X2B X3B HR
## 1   baezja01   2019     1    CHN   NL infield 138 531 149 0.281  89  38   4 29
## 2   boteda01   2019     1    CHN   NL infield 127 303  78 0.257  47  17   0 11
## 3  bryankr01   2019     1    CHN   NL infield 147 543 153 0.282 108  35   1 31
## 4  rizzoan01   2019     1    CHN   NL infield 146 512 150 0.293  89  29   3 27
## 5  lemahdj01   2019     1    NYA   AL infield 145 602 197 0.327 109  33   2 26
## 6  torregl01   2019     1    NYA   AL infield 144 546 152 0.278  96  26   0 38
## 7  urshegi01   2019     1    NYA   AL infield 132 442 139 0.314  73  34   0 21
## 8   voitlu01   2019     1    NYA   AL infield 118 429 113 0.263  72  21   1 21
## 9  altuvjo01   2019     1    HOU   AL infield 124 500 149 0.298  89  27   3 31
## 10 bregmal01   2019     1    HOU   AL infield 156 554 164 0.296 122  37   2 41
## 11 gourryu01   2019     1    HOU   AL infield 144 564 168 0.298  85  40   2 31
## 12 andruel01   2019     1    TEX   AL infield 147 600 165 0.275  81  27   4 12
## 13 forsylo01   2019     1    TEX   AL infield 101 317  72 0.227  38  17   1  7
```

```
## 14  odorro01   2019     1    TEX   AL infield 145 522 107 0.205  77  30   1 30
## 15 bogaexa01   2019     1    BOS   AL infield 155 614 190 0.309 110  52   0 33
## 16 deverra01   2019     1    BOS   AL infield 156 647 201 0.311 129  54   4 32
## 17 albieoz01   2019     1    ATL   NL infield 160 640 189 0.295 102  43   8 24
## 18 donaljo02   2019     1    ATL   NL infield 155 549 142 0.259  96  33   0 37
## 19 freemfr01   2019     1    ATL   NL infield 158 597 176 0.295 113  34   2 38
## 20 swansda01   2019     1    ATL   NL infield 127 483 121 0.251  77  26   3 17
## 21 adamsma01   2019     1    WAS   NL infield 111 310  70 0.226  42  14   0 20
## 22 doziebr01   2019     1    WAS   NL infield 135 416  99 0.238  54  20   0 20
## 23 kendrho01   2019     1    WAS   NL infield 121 334 115 0.344  61  23   1 17
## 24 rendoan01   2019     1    WAS   NL infield 146 545 174 0.319 117  44   3 34
##    RBI SB CS  BB  SO IBB HBP SH SF GIDP
## 1   85 11  7  28 156   3   0  0  2   16
## 2   41  5  1  44  93   4   7  0  2   11
## 3   77  4  0  74 145   1  15  0  2   10
## 4   94  5  2  71  86   3  27  0  3   15
## 5  102  5  2  46  90   0   2  1  4   14
## 6   90  5  2  48 129   3   3  1  6   10
## 7   74  1  1  25  87   1   5  0  4   13
## 8   62  0  0  71 142   2   9  0  1   12
## 9   74  6  5  41  82   0   3  1  3   19
## 10 112  5  1 119  83   2   9  0  8    9
## 11 104  5  3  37  65   2   5  0  6   12
## 12  72 31  8  34  96   1   4  0 10   16
## 13  39  2  0  44 100   0   3  0  2    8
## 14  93 11  9  52 178   2   5  1  1    4
## 15 117  4  2  76 122   2   2  0  6   11
## 16 115  8  8  48 119   7   4  1  2    8
## 17  86 15  4  54 112   6   4  0  4    2
## 18  94  4  2 100 155   2   8  0  2   13
## 19 121  6  3  87 127  11   6  0  2   17
## 20  65 10  5  51 124   2   5  1  5    7
## 21  56  0  0  20 115   1   2  0  1    7
## 22  50  3  4  61 105   2   4  0  1   11
## 23  62  2  1  27  49   1   4  0  5   11
## 24 126  5  1  80  86   8  12  0  9   13
```

```
infield_players %>%
  top_n(15) %>%
  ggplot(aes(y=reorder(playerID,avg),x=avg,fill=playerID)) +
      geom_bar(stat = 'identity',position=position_dodge()) +
      geom_text(aes(label=avg), vjust=1.0, color="black",
                position = position_dodge(0.9), size=3.0) +
      labs(y = ("Player"),x = ("Infielder Batting Average"),
           title = ("2019 Top 15 Infielder Batting Averages")) +
      theme_minimal()
```

```
## Selecting by GIDP
```




```
(catchers <- batting_by_posit %>% filter(batting_by_posit$posit == "catcher"))
```

```
##     playerID yearID stint teamID lgID   posit   G  AB   H   avg  R X2B X3B HR
## 1  sanchga02   2019     1    NYA   AL catcher 106 396  92 0.232 62  12   1 34
## 2  chiriro01   2019     1    HOU   AL catcher 114 366  87 0.238 57  22   1 17
## 3  mathije01   2019     1    TEX   AL catcher  88 228  36 0.158 17   9   0  2
## 4  contrwi01   2019     1    CHN   NL catcher 105 360  98 0.272 57  18   2 24
## 5  vazquch01   2019     1    BOS   AL catcher 138 482 133 0.276 66  26   1 23
## 6  mccanbr01   2019     1    ATL   NL catcher  85 277  69 0.249 28   9   0 12
## 7  martiru01   2019     1    LAN   NL catcher  83 209  46 0.220 29   5   0  6
## 8  severpe01   2019     1    BAL   AL catcher  96 305  76 0.249 37  13   0 13
## 9  realmjt01   2019     1    PHI   NL catcher 145 538 148 0.275 92  36   3 25
## 10 gomesya01   2019     1    WAS   NL catcher  97 314  70 0.223 36  16   0 12
## 11 suzukku01   2019     1    WAS   NL catcher  85 280  74 0.264 37  11   0 17
## 12 flowety01   2019     1    ATL   NL catcher  85 271  62 0.229 36  11   3 11
## 13 kellyca02   2019     1    ARI   NL catcher 111 314  77 0.245 46  19   0 18
## 14 perezro02   2019     1    CLE   AL catcher 119 389  93 0.239 46   9   1 24
##    RBI SB CS BB  SO IBB HBP SH SF GIDP
## 1   77  0  1 40 125   3   9  0  1    3
## 2   58  1  2 51 125   1  13  2  5   11
## 3   12  1  0 15  87   1   0  0  1    2
## 4   64  1  2 38 102   2   9  0  2    4
## 5   72  4  2 33 101   3   0  3  3   17
## 6   45  0  0 31  53   1   2  0  6   10
## 7   20  1  0 30  60   3   8  0  2    1
## 8   44  3  1 29  73   0   4  1  2    5
```

```
## 9   83  9  1 41 123   2   5  0  8   12
## 10  43  2  0 38  84   6   5  0  1    7
## 11  63  0  1 20  36   1   6  0  3   10
## 12  34  0  0 31 105   3   6  0  2    8
## 13  47  0  0 48  79  10   2  0  1   11
## 14  63  0  0 45 127   1   4  7  4   12
```

```
catchers %>%
  top_n(10) %>%
  ggplot(aes(y=reorder(playerID,avg),x=avg,fill=playerID)) +
      geom_bar(stat = 'identity',position=position_dodge()) +
      geom_text(aes(label=avg), vjust=1.0, color="black",
                position = position_dodge(0.9), size=3.0) +
      labs(y = ("Player"),x = ("Catcher Batting Average"),
           title = ("2019 Top 10 Catcher Batting Averages")) +
      theme_minimal()
```

```
## Selecting by GIDP
```