Injuries

R Markdown

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

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

```
Injury<-read.csv("~/Downloads/Injury.csv")
Premier.League<-read.csv("~/Downloads/Premier League.csv")
```

```
dim(Injury)
```

```
## [1] 315  10
```

```
summary(Injury$Age)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   16.00   23.00   27.00   26.56   30.00   37.00
```

```
which(is.na(Injury))
```

```
## integer(0)
```

```
which(is.na(Premier.League))
```

```
##  [1]  967  974  975  976  977  978  979  982  983  985  989  992  993  996 1003
## [16] 1004 1005 1006 1007 1008 1011 1012 1013 1014 1015 1016 1017 1018 1019 1021
## [31] 1023 1025 1026 1028 1029 1030 1032 1034 1035 1036 1037 1038 1039 1044 1045
## [46] 1046 1047 1048 1049 1052 1059 1060 1065 1067 1068 1070 1071 1072 1073 1074
## [61] 1076 1077 1079 1082 1084 1085 1087 1088 1090 1091 1096 1097 1098 1100 1101
## [76] 1102 1291
```

```
str(Injury)
```

```
## 'data.frame':    315 obs. of  10 variables:
##  $ Player            : chr  "David Luis Moreira Moarinho" "Hector Bellerin" "Granit Xhaka" "Pablo Ma
##  $ Injury            : chr  "Hamstring" "Knock" "Groin/Hip/Pelvis" "Ankle/Foot" ...
##  $ Date.of.Injury    : chr  "5/2/21" "5/12/21" "5/6/21" "5/8/21" ...
##  $ Potential.Return  : chr  "5/19/21" "No Return Date" "5/19/21" "5/19/21" ...
```

```
##  $ Age               : int  34 26 28 27 20 23 28 26 20 25 ...
##  $ Team              : chr  "Arsenal" "Arsenal" "Arsenal" "Arsenal" ...
##  $ Total.Days        : int  33 23 13 4 3 22 6 55 65 81 ...
##  $ Position          : chr  "Defender" "Defender" "Midfielder" "Defender" ...
##  $ Reoccurring.Injury: chr  "Yes" "No" "Yes" "Yes" ...
##  $ League            : chr  "Premier League  " "Premier League  " "Premier League  " "Premier League
```

```
Injury$Date.of.Injury<-as.Date(Injury$Date.of.Injury, format= "%m/%d/%y")
head(Injury$Date.of.Injury)
```

```
## [1] "2021-05-02" "2021-05-12" "2021-05-06" "2021-05-08" "2021-05-09"
## [6] "2021-05-13"
```

```
## Loading required package: Hmisc
```

```
## Loading required package: lattice
```

```
## Loading required package: survival
```

```
## Loading required package: Formula
```

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

```
##
## Attaching package: 'Hmisc'
```

```
## The following objects are masked from 'package:base':
##
##     format.pval, units
```

```
## funModeling v.1.9.4 :)
## Examples and tutorials at livebook.datascienceheroes.com
##  / Now in Spanish: librovivodecienciadedatos.ai
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v tibble  3.1.2     v dplyr   1.0.7
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1
## v purrr   0.3.4
```

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

```
## Warning in freq_logic(data = data, input = input[i], plot, na.rm, path_out =
## path_out): Skipping plot for variable 'Player' (more than 100 categories)
```

	Player	frequency	percentage	cumulative_perc
1	Ben Davies	2	0.63	0.63
2	Dale Stephens	2	0.63	1.26
3	Aaron Cresswell	1	0.32	1.58
4	Adam Forshaw	1	0.32	1.90
5	Adam Lallana	1	0.32	2.22
6	Adama Traore	1	0.32	2.54
7	Aihen Munoz	1	0.32	2.86
8	Alejandro Pozo	1	0.32	3.18
9	Aleksandar Mitrovic	1	0.32	3.50
10	Alex Fernandez	1	0.32	3.82
11	Alex Oxlade-Chamberlain	1	0.32	4.14
12	Allan Saint-Maximin	1	0.32	4.46
13	Amin Younes	1	0.32	4.78
14	Ander Capa	1	0.32	5.10
15	Andreas Christensen	1	0.32	5.42
16	Angel Rodriguez	1	0.32	5.74
17	Angelo Ogbonna	1	0.32	6.06
18	Ansu Fati	1	0.32	6.38
19	Anthony Martial	1	0.32	6.70
20	Anthony Ujah	1	0.32	7.02
21	Antonee Robinson	1	0.32	7.34
22	Antonio Valera	1	0.32	7.66
23	Aridane Hernandez	1	0.32	7.98
24	Arthur Masuaku	1	0.32	8.30
25	Asier Illarramendi	1	0.32	8.62
26	Axel Witsel	1	0.32	8.94
27	Bartosz Bialek	1	0.32	9.26
28	Ben Mee	1	0.32	9.58
29	Benjamin Hubner	1	0.32	9.90
30	Benjamin Stambouli	1	0.32	10.22
31	Benno Schmitz	1	0.32	10.54
32	Billy Sharp	1	0.32	10.86
33	Branislav Ivanovic	1	0.32	11.18
34	Bruno Gonzalez	1	0.32	11.50
35	Burgui	1	0.32	11.82
36	Cala	1	0.32	12.14
37	Callum Wilson	1	0.32	12.46
38	Caoimhin Kelleher	1	0.32	12.78
39	Carlos Akapo	1	0.32	13.10
40	Carlos Fernandez Luna	1	0.32	13.42
41	Carlos Neva	1	0.32	13.74
42	Cedric Teuchert	1	0.32	14.06
43	Cengiz Under	1	0.32	14.38
44	Che Adams	1	0.32	14.70
45	Cheikhou Kouyate	1	0.32	15.02
46	Chris Richards	1	0.32	15.34
47	Chris Wood	1	0.32	15.66
48	Christian Benteke	1	0.32	15.98
49	Christoph Baumgartner	1	0.32	16.30
50	Connor Wickham	1	0.32	16.62
51	Cucho	1	0.32	16.94
52	Daley Sinkgraven	1	0.32	17.26
53	Dan-Axel Zagadou	1	0.32	17.58

```
## 54                  Dani Garcia         1       0.32          17.90
## 55                  Dani Martin         1       0.32          18.22
## 56                    Dani Olmo         1       0.32          18.54
## 57               Daniel Carvajal        1       0.32          18.86
## 58                  Daniel James        1       0.32          19.18
## 59                Daniel Podence        1       0.32          19.50
## 60                    Danny Ings        1       0.32          19.82
## 61                 Danny Welbeck        1       0.32          20.14
## 62                  Dario Poveda        1       0.32          20.46
## 63  David Luis Moreira Moarinho         1       0.32          20.78
## 64                  Davy Propper        1       0.32          21.10
## 65                  Declan Rice         1       0.32          21.42
## 66                  Denis Vavro         1       0.32          21.74
## 67                 Dennis Geiger        1       0.32          22.06
## 68                Diego Llorente        1       0.32          22.38
## 69                    Diogo Jota        1       0.32          22.70
## 70                  Divock Origi        1       0.32          23.02
## 71                 Djene Dakonam        1       0.32          23.34
## 72                  Djibril Sow         1       0.32          23.66
## 73                Dominic Revan         1       0.32          23.98
## 74            Dominik Szoboszlai        1       0.32          24.30
## 75                 Douglas Costa        1       0.32          24.62
## 76                  Eberechi Eze        1       0.32          24.94
## 77                  Eden Hazard         1       0.32          25.26
## 78                Edouard Mendy         1       0.32          25.58
## 79              Elliot Anderson         1       0.32          25.90
## 80              Elvis Rexhbecaj         1       0.32          26.22
## 81                  Emil Krafth         1       0.32          26.54
## 82             Emile Smith Rowe         1       0.32          26.86
## 83  Emmanuel Bonaventure Dennis         1       0.32          27.18
## 84                      Emre Mor        1       0.32          27.50
## 85                  Erick Cabaco        1       0.32          27.82
## 86                Ermin Bicakcic        1       0.32          28.14
## 87                  Ethan Ampadu        1       0.32          28.46
## 88             Eugeni Valderrama        1       0.32          28.78
## 89             Exequiel Palacios        1       0.32          29.10
## 90               Fabian Orellana        1       0.32          29.42
## 91              Facundo Ferreyra        1       0.32          29.74
## 92                 Fedrik Jensen        1       0.32          30.06
## 93                Felix Uduokhai        1       0.32          30.38
## 94                 Ferland Mendy        1       0.32          30.70
## 95                Fernando Marcal        1       0.32          31.02
## 96                Florian Hubner         1       0.32          31.34
## 97                  Florin Andone        1       0.32          31.66
## 98                           Fred        1       0.32          31.98
## 99               Frederik Ronnow         1       0.32          32.30
## 100                 Gareth Bale         1       0.32          32.62
## 101               George Baldock        1       0.32          32.94
## 102             Giovani Lo Celso        1       0.32          33.26
## 103                 Granit Xhaka        1       0.32          33.58
## 104               Grischa Promel        1       0.32          33.90
## 105                 Hamza Mendyl        1       0.32          34.22
## 106                Harrison Reed        1       0.32          34.54
## 107                Harry Maguire        1       0.32          34.86
```

```
## 108          Harvard Nordtveit        1       0.32          35.18
## 109              Harvey Barnes        1       0.32          35.50
## 110            Hector Bellerin        1       0.32          35.82
## 111               Helder Costa        1       0.32          36.14
## 112                 Hugo Mallo        1       0.32          36.46
## 113                 Iago Aspas        1       0.32          36.78
## 114            Ibrahima Traore        1       0.32          37.10
## 115              Igor Zubeldia        1       0.32          37.42
## 116               Iker Muniain        1       0.32          37.74
## 117              Inigo Cordoba        1       0.32          38.06
## 118                Inigo Lekue        1       0.32          38.38
## 119               Isaac Hayden        1       0.32          38.70
## 120              Jabez Makanda        1       0.32          39.02
## 121              Jack Grealish        1       0.32          39.34
## 122            Jack O'Connell        1       0.32          39.66
## 123                 Jakob Busk        1       0.32          39.98
## 124          Jamaal Lascelles        1       0.32          40.30
## 125              James Justin        1       0.32          40.62
## 126            James McArthur        1       0.32          40.94
## 127              James Milner        1       0.32          41.26
## 128            James Rodriguez        1       0.32          41.58
## 129              James Tomkins        1       0.32          41.90
## 130               Jan Bednarek        1       0.32          42.22
## 131                Jan Moravek        1       0.32          42.54
## 132           Japhet Tanganga        1       0.32          42.86
## 133      Javier Lopez Carballo        1       0.32          43.18
## 134       Jean-Philippe Gbamin        1       0.32          43.50
## 135             Jeison Murillo        1       0.32          43.82
## 136              Jesus Vallejo        1       0.32          44.14
## 137               Jhon Cordoba        1       0.32          44.46
## 138          Joachim Andersen        1       0.32          44.78
## 139                 Joe Gomez        1       0.32          45.10
## 140                Joel Matip        1       0.32          45.42
## 141              Joel Veltman        1       0.32          45.74
## 142                 Joelinton        1       0.32          46.06
## 143              Johan Mojica        1       0.32          46.38
## 144               Jon Pacheco        1       0.32          46.70
## 145              Jonas Hector        1       0.32          47.02
## 146         Jonathan Burkardt        1       0.32          47.34
## 147          Jonathan Calleri        1       0.32          47.66
## 148         Jonny Castro Otto        1       0.32          47.98
## 149               Jonny Evans        1       0.32          48.30
## 150          Jordan Henderson        1       0.32          48.62
## 151        Jose Angel Tasende        1       0.32          48.94
## 152               Jose Fontan        1       0.32          49.26
## 153              Joseph Aidoo        1       0.32          49.58
## 154            Joseph Willock        1       0.32          49.90
## 155               Joshua King        1       0.32          50.22
## 156             Josip Brekalo        1       0.32          50.54
## 157                Juan Foyth        1       0.32          50.86
## 158      Julian Baumgartlinger        1       0.32          51.18
## 159                Kai Havertz        1       0.32          51.50
## 160           Kalvin Phillips        1       0.32          51.82
## 161               Karl Darlow        1       0.32          52.14
```

```
## 162                Kenan Kodro        1       0.32          52.46
## 163                     Kenedy        1       0.32          52.78
## 164             Kevin Akpoguma        1       0.32          53.10
## 165            Kevin De Bruyne        1       0.32          53.42
## 166                 Kevin Long        1       0.32          53.74
## 167               Kevin Stoger        1       0.32          54.06
## 168         Klaas Jan Huntelaar        1       0.32          54.38
## 169              Koen Casteels        1       0.32          54.70
## 170     Konstantinos Mavropanos        1       0.32          55.02
## 171     Konstantinos Stafylidis        1       0.32          55.34
## 172           Krzysztof Piatek        1       0.32          55.66
## 173                Lars Bender        1       0.32          55.98
## 174             Lazar Samardzic        1       0.32          56.30
## 175    Leandro Barreiro Martins        1       0.32          56.62
## 176                Leon Bailey         1       0.32          56.94
## 177                Leon Dajaku        1       0.32          57.26
## 178              Leon Goretzka        1       0.32          57.58
## 179                Liam Cooper         1       0.32          57.90
## 180              Lilian Egloff        1       0.32          58.22
## 181                  Luca Netz        1       0.32          58.54
## 182              Luca Plogmann        1       0.32          58.86
## 183              Luca Sangalli        1       0.32          59.18
## 184               Lucas Alario        1       0.32          59.50
## 185              Lucas Tousart        1       0.32          59.82
## 186              Lucas Vazquez        1       0.32          60.14
## 187                 Luis Milla        1       0.32          60.46
## 188                   Luisinho        1       0.32          60.78
## 189              Luismi Quezada        1       0.32          61.10
## 190                Lukas Kubler        1       0.32          61.42
## 191            Lukasz Fabianski        1       0.32          61.74
## 192            Lukasz Piszczek        1       0.32          62.06
## 193              Mahmoud Dahoud        1       0.32          62.38
## 194            Mamadou Doucoure        1       0.32          62.70
## 195               Mamadou Sakho        1       0.32          63.02
## 196              Manuel Lanzini        1       0.32          63.34
## 197       Marc-Andre ter Stegen        1       0.32          63.66
## 198          Marcel Halstenberg        1       0.32          63.98
## 199            Marcel Schmelzer        1       0.32          64.30
## 200             Marcin Kaminski        1       0.32          64.62
## 201                  Mark Noble        1       0.32          64.94
## 202              Martin Dubravka        1       0.32          65.26
## 203              Martin Montoya        1       0.32          65.58
## 204                 Marvin Park        1       0.32          65.90
## 205                 Marwin Hitz        1       0.32          66.22
## 206              Mateo Klimowicz        1       0.32          66.54
## 207                Mateo Kovacic        1       0.32          66.86
## 208                  Mateu Morey        1       0.32          67.18
## 209                Matheus Cunha        1       0.32          67.50
## 210              Matheus Pereira        1       0.32          67.82
## 211              Matija Nastasic        1       0.32          68.14
## 212            Matteo Guendouzi        1       0.32          68.46
## 213                  Matty Cash        1       0.32          68.78
## 214     Maximillian Mittelstadt        1       0.32          69.10
## 215             Michy Batshuayi        1       0.32          69.42
```

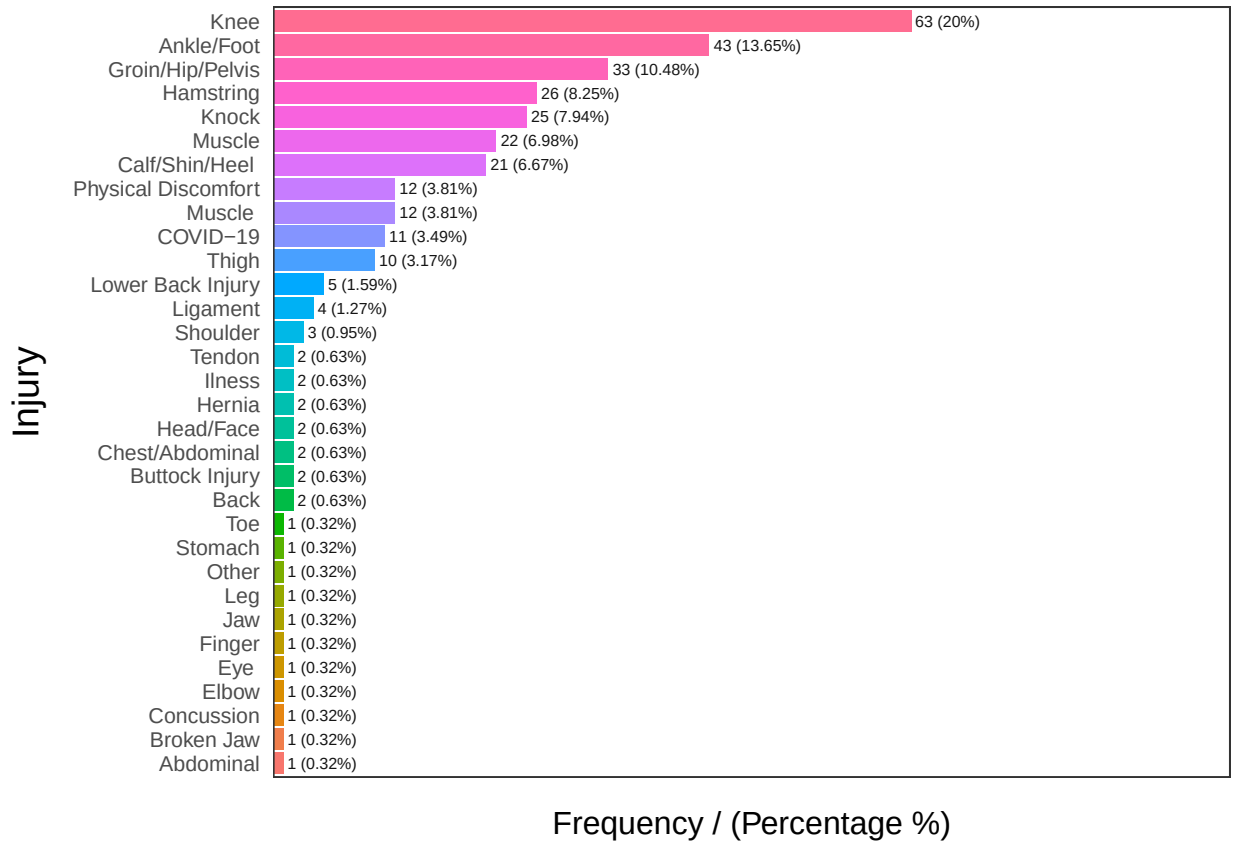
```
## 216                Mikel Merino         1       0.32          69.74
## 217            Milos Veljikovic         1       0.32          70.06
## 218                 Momo Cisse          1       0.32          70.38
## 219              Morgaon Sanson         1       0.32          70.70
## 220              Moussa Sissoko         1       0.32          71.02
## 221              Nabil Bentaleb         1       0.32          71.34
## 222                  Naby Keita         1       0.32          71.66
## 223            Nassim Boujellab         1       0.32          71.98
## 224             Nathan Ferguson         1       0.32          72.30
## 225             Nathaniel Clyne         1       0.32          72.62
## 226             Nemanja Radoja          1       0.32          72.94
## 227              Neyder Lozano          1       0.32          73.26
## 228                Ngolo Kante         1       0.32          73.58
## 229                  Nick Pope         1       0.32          73.90
## 230              Nick Woltemade         1       0.32          74.22
## 231            Nicolas Gonzalez         1       0.32          74.54
## 232             Nikola Vukcevic         1       0.32          74.86
## 233        Nils-Jonathan Korber         1       0.32          75.18
## 234    Noah-Joel Sarenren-Bazee         1       0.32          75.50
## 235                Oier Zarraga         1       0.32          75.82
## 236              Oliver Baumann         1       0.32          76.14
## 237                Oliver Burke         1       0.32          76.46
## 238             Oliver McBurnie         1       0.32          76.78
## 239              Omar Mascarell         1       0.32          77.10
## 240                Orel Mangala         1       0.32          77.42
## 241                 Oriol Romeu         1       0.32          77.74
## 242              Owen Otasowie          1       0.32          78.06
## 243                  Ozan Kabak         1       0.32          78.38
## 244                 Pablo Insua         1       0.32          78.70
## 245          Pablo Mari Villar          1       0.32          79.02
## 246              Paul Akouokou          1       0.32          79.34
## 247            Pedro Lomba Neto         1       0.32          79.66
## 248             Peru Nolaskoain         1       0.32          79.98
## 249                  Phil Jones         1       0.32          80.30
## 250            Phillip Bardsley         1       0.32          80.62
## 251          Phillippe Coutinho         1       0.32          80.94
## 252              Raphael Varane         1       0.32          81.26
## 253                 Raul Garcia         1       0.32          81.58
## 254                Raul Jimenez         1       0.32          81.90
## 255              Rayan Ait Nouri         1       0.32          82.22
## 256             Renato Steffen           1       0.32          82.54
## 257                Renato Tapia         1       0.32          82.86
## 258                Robbie Brady         1       0.32          83.18
## 259            Robert Snodgrass         1       0.32          83.50
## 260                  Robin Koch         1       0.32          83.82
## 261               Robin Zentner         1       0.32          84.14
## 262                 Rodrigo Ely         1       0.32          84.46
## 263                Ruben Blanco         1       0.32          84.78
## 264              Ruben Martinez         1       0.32          85.10
## 265     Rune Almenning Jarstein         1       0.32          85.42
## 266               Ryan Bertrand         1       0.32          85.74
## 267                 Ryan Fraser         1       0.32          86.06
## 268               Salvi Sanchez         1       0.32          86.38
## 269                 Sami Khedira         1       0.32          86.70
```

```
## 270            Samuel Chukwueze         1       0.32          87.02
## 271                Sander Berge         1       0.32          87.34
## 272              Santiago Arias         1       0.32          87.66
## 273                  Scott Dann         1       0.32          87.98
## 274             Scott McTominay         1       0.32          88.30
## 275              Sead Kolasinac         1       0.32          88.62
## 276              Seamus Coleman         1       0.32          88.94
## 277              Sebastian Rode         1       0.32          89.26
## 278                Serge Aurier         1       0.32          89.58
## 279               Sergi Roberto         1       0.32          89.90
## 280               Sergio Aguero         1       0.32          90.22
## 281              Sergio Alvarez         1       0.32          90.54
## 282              Sergio Postigo         1       0.32          90.86
## 283                Sergio Ramos         1       0.32          91.18
## 284           Silas Wamangituka         1       0.32          91.50
## 285              Solomon March          1       0.32          91.82
## 286           Tanguy Coulibaly          1       0.32          92.14
## 287             Tanguy Ndombele         1       0.32          92.46
## 288               Tariq Lamptey         1       0.32          92.78
## 289             Terence Kongolo         1       0.32          93.10
## 290                Thomas Lemar         1       0.32          93.42
## 291         Timothy Fosu-Mensah         1       0.32          93.74
## 292                Tom Cairney          1       0.32          94.06
## 293                 Toni Kroos          1       0.32          94.38
## 294               Torben Musel          1       0.32          94.70
## 295                  Trezeguet          1       0.32          95.02
## 296                Tyler Adams          1       0.32          95.34
## 297               Tyler Onyango         1       0.32          95.66
## 298              Tyrick Mitchell         1       0.32          95.98
## 299                Tyrone Mings          1       0.32          96.30
## 300               Urko Gonzalez          1       0.32          96.62
## 301              Vicente Iborra          1       0.32          96.94
## 302             Victor Camarasa          1       0.32          97.26
## 303                 Victor Diaz          1       0.32          97.58
## 304             Virgil van Dijk          1       0.32          97.90
## 305                  Wes Morgan          1       0.32          98.22
## 306               Wesley Fofana         1       0.32          98.54
## 307                     William         1       0.32          98.86
## 308           William Smallbone         1       0.32          99.18
## 309                  Willy Boly          1       0.32          99.50
## 310               Yeray Alvarez          1       0.32          99.82
## 311                  Yerry Mina          1       0.32         100.14
## 312           Youssoufa Moukoko          1       0.32         100.46
## 313              Yuri Berchiche          1       0.32         100.00
```

```
## Warning: ‘guides(<scale> = FALSE)‘ is deprecated. Please use ‘guides(<scale> =
## "none")‘ instead.
```

```
##                 Injury frequency percentage cumulative_perc
## 1                 Knee        63      20.00           20.00
## 2           Ankle/Foot        43      13.65           33.65
## 3     Groin/Hip/Pelvis        33      10.48           44.13
## 4            Hamstring        26       8.25           52.38
## 5                Knock        25       7.94           60.32
```

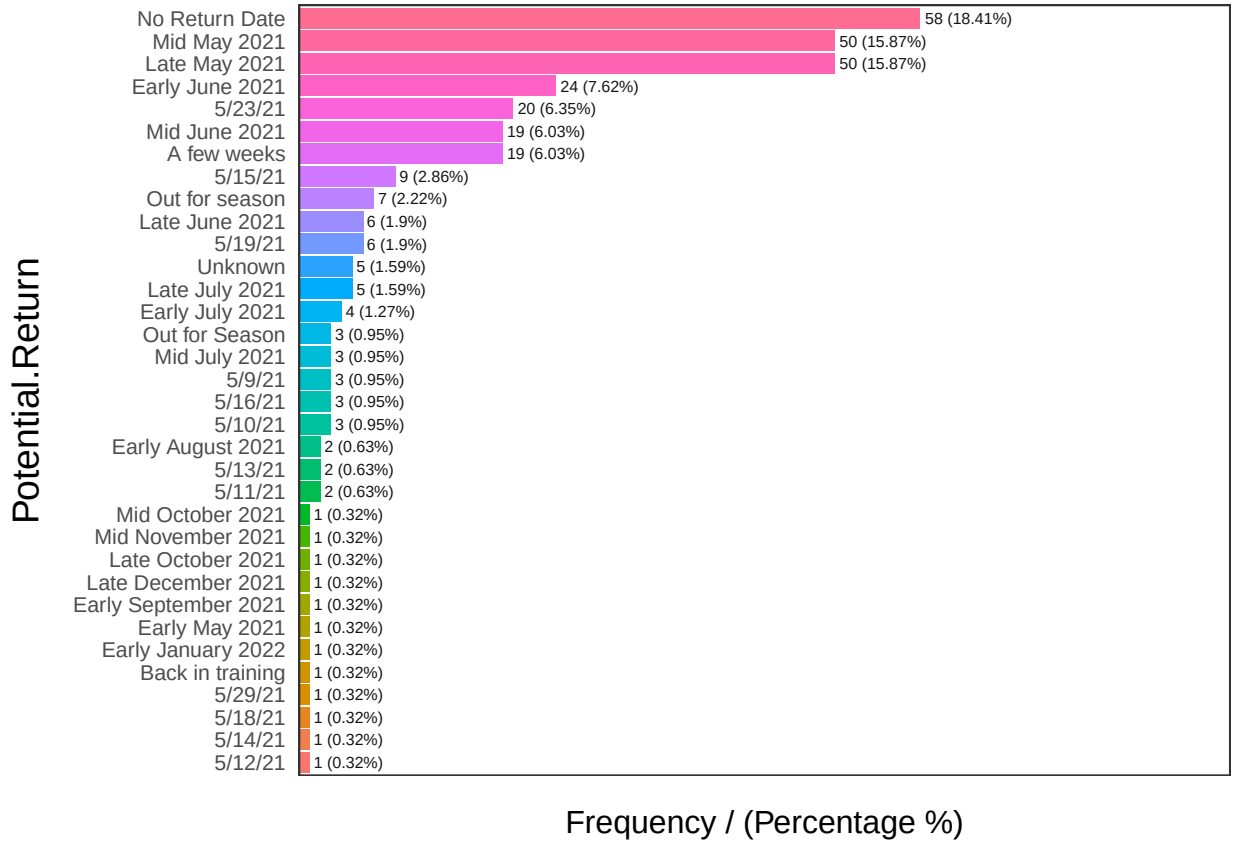
```
## 6               Muscle        22       6.98          67.30
## 7      Calf/Shin/Heel         21       6.67          73.97
## 8               Muscle        12       3.81          77.78
## 9  Physical Discomfort        12       3.81          81.59
## 10            COVID-19        11       3.49          85.08
## 11               Thigh        10       3.17          88.25
## 12   Lower Back Injury         5       1.59          89.84
## 13             Ligament         4       1.27          91.11
## 14             Shoulder         3       0.95          92.06
## 15                 Back         2       0.63          92.69
## 16      Buttock Injury          2       0.63          93.32
## 17     Chest/Abdominal         2       0.63          93.95
## 18           Head/Face         2       0.63          94.58
## 19              Hernia         2       0.63          95.21
## 20              Ilness         2       0.63          95.84
## 21              Tendon         2       0.63          96.47
## 22           Abdominal         1       0.32          96.79
## 23          Broken Jaw         1       0.32          97.11
## 24          Concussion         1       0.32          97.43
## 25               Elbow         1       0.32          97.75
## 26                Eye          1       0.32          98.07
## 27              Finger         1       0.32          98.39
## 28                 Jaw         1       0.32          98.71
## 29                 Leg         1       0.32          99.03
## 30               Other         1       0.32          99.35
## 31             Stomach         1       0.32          99.67
## 32                 Toe         1       0.32         100.00

## Warning: ‘guides(<scale> = FALSE)‘ is deprecated. Please use ‘guides(<scale> =
## "none")‘ instead.
```




```
##        Potential.Return frequency percentage cumulative_perc
## 1         No Return Date        58      18.41           18.41
## 2          Late May 2021        50      15.87           34.28
## 3           Mid May 2021        50      15.87           50.15
## 4        Early June 2021        24       7.62           57.77
## 5                5/23/21        20       6.35           64.12
## 6            A few weeks        19       6.03           70.15
## 7          Mid June 2021        19       6.03           76.18
## 8                5/15/21         9       2.86           79.04
## 9         Out for season         7       2.22           81.26
## 10               5/19/21         6       1.90           83.16
## 11         Late June 2021         6       1.90           85.06
## 12         Late July 2021         5       1.59           86.65
## 13                Unknown         5       1.59           88.24
## 14       Early July 2021         4       1.27           89.51
## 15               5/10/21         3       0.95           90.46
## 16               5/16/21         3       0.95           91.41
## 17                5/9/21         3       0.95           92.36
## 18         Mid July 2021         3       0.95           93.31
## 19         Out for Season        3       0.95           94.26
## 20               5/11/21         2       0.63           94.89
## 21               5/13/21         2       0.63           95.52
## 22    Early August 2021          2       0.63           96.15
## 23               5/12/21         1       0.32           96.47
## 24               5/14/21         1       0.32           96.79
## 25               5/18/21         1       0.32           97.11
```

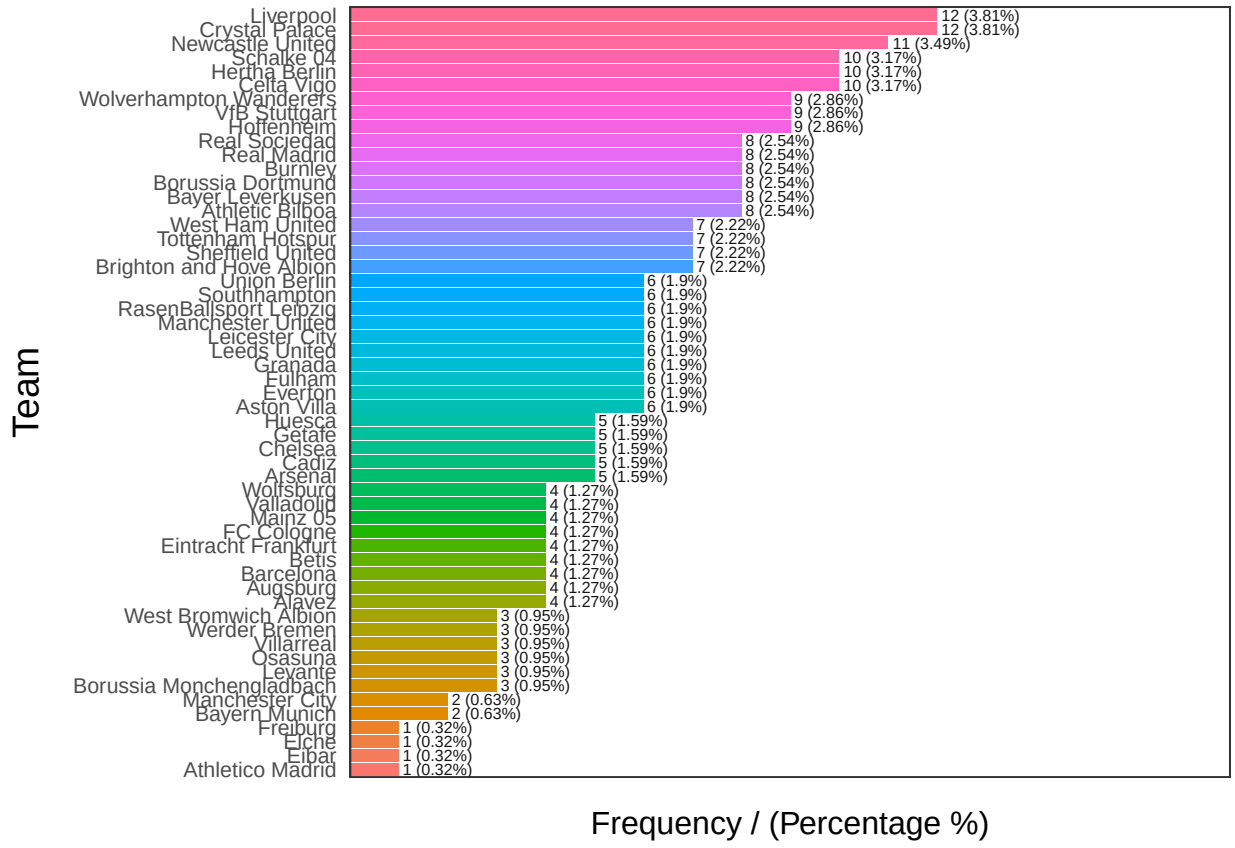
```
## 26             5/29/21         1       0.32          97.43
## 27     Back in training         1       0.32          97.75
## 28   Early January 2022         1       0.32          98.07
## 29       Early May 2021         1       0.32          98.39
## 30 Early September 2021         1       0.32          98.71
## 31   Late December 2021         1       0.32          99.03
## 32    Late October 2021         1       0.32          99.35
## 33    Mid November 2021         1       0.32          99.67
## 34     Mid October 2021         1       0.32         100.00
```

```
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.
```



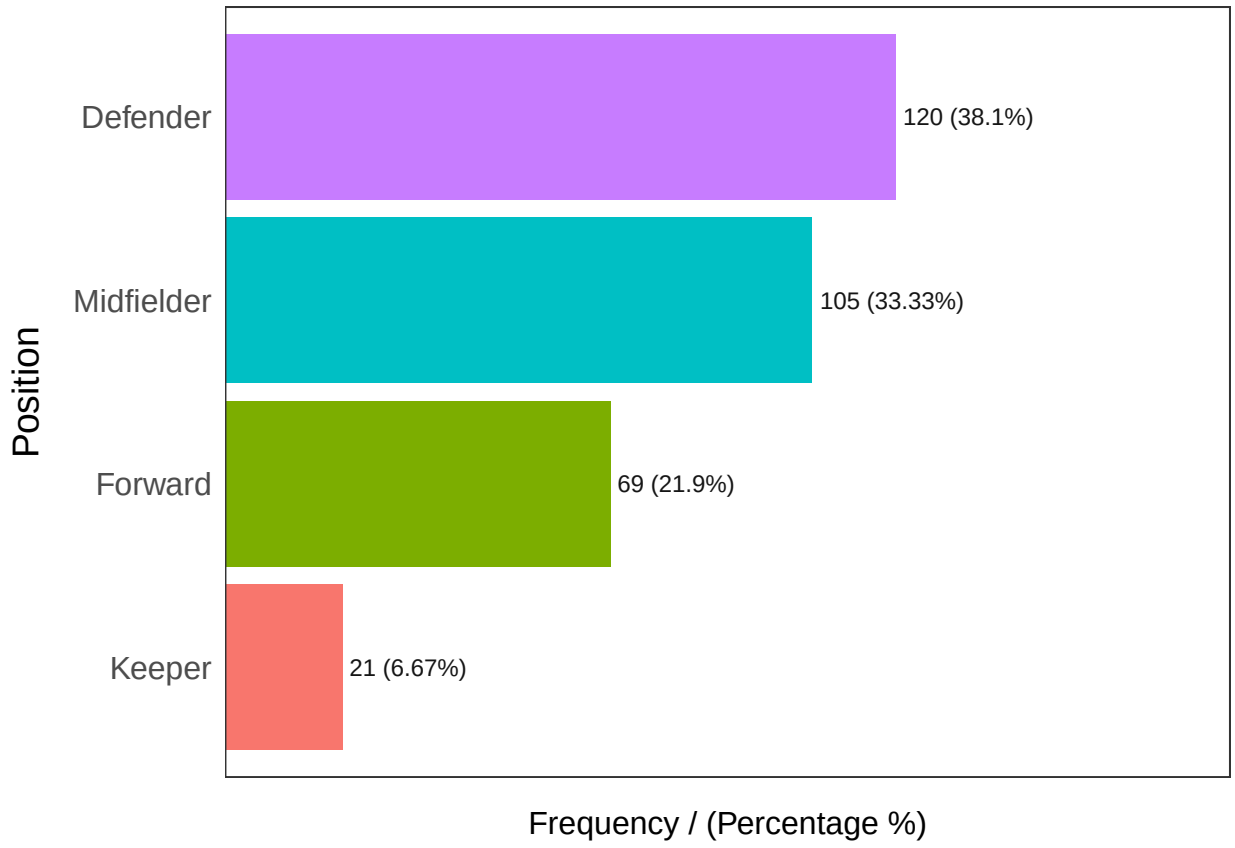

```
##                    Team frequency percentage cumulative_perc
## 1          Crystal Palace        12       3.81            3.81
## 2               Liverpool        12       3.81            7.62
## 3         Newcastle United        11       3.49           11.11
## 4               Celta Vigo        10       3.17           14.28
## 5            Hertha Berlin        10       3.17           17.45
## 6               Schalke 04        10       3.17           20.62
## 7               Hoffenheim         9       2.86           23.48
## 8            VfB Stuttgart         9       2.86           26.34
## 9  Wolverhampton Wanderers         9       2.86           29.20
## 10         Athletic Bilboa         8       2.54           31.74
## 11        Bayer Leverkusen         8       2.54           34.28
```

```
## 12        Borussia Dortmund        8       2.54          36.82
## 13                  Burnley        8       2.54          39.36
## 14             Real Madrid         8       2.54          41.90
## 15           Real Sociedad         8       2.54          44.44
## 16 Brighton and Hove Albion        7       2.22          46.66
## 17         Sheffield United        7       2.22          48.88
## 18        Tottenham Hotspur        7       2.22          51.10
## 19          West Ham United        7       2.22          53.32
## 20              Aston Villa        6       1.90          55.22
## 21                  Everton        6       1.90          57.12
## 22                   Fulham        6       1.90          59.02
## 23                  Granada        6       1.90          60.92
## 24             Leeds United        6       1.90          62.82
## 25           Leicester City        6       1.90          64.72
## 26        Manchester United        6       1.90          66.62
## 27   RasenBallsport Leipzig        6       1.90          68.52
## 28             Southhampton        6       1.90          70.42
## 29             Union Berlin        6       1.90          72.32
## 30                  Arsenal        5       1.59          73.91
## 31                    Cadiz        5       1.59          75.50
## 32                  Chelsea        5       1.59          77.09
## 33                   Getafe        5       1.59          78.68
## 34                   Huesca        5       1.59          80.27
## 35                   Alavez        4       1.27          81.54
## 36                 Augsburg        4       1.27          82.81
## 37                Barcelona        4       1.27          84.08
## 38                    Betis        4       1.27          85.35
## 39      Eintracht Frankfurt        4       1.27          86.62
## 40                FC Cologne        4       1.27          87.89
## 41                  Mainz 05        4       1.27          89.16
## 42                Valladolid        4       1.27          90.43
## 43                 Wolfsburg        4       1.27          91.70
## 44 Borussia Monchengladbach        3       0.95          92.65
## 45                   Levante        3       0.95          93.60
## 46                   Osasuna        3       0.95          94.55
## 47                Villarreal        3       0.95          95.50
## 48             Werder Bremen        3       0.95          96.45
## 49     West Bromwich Albion         3       0.95          97.40
## 50             Bayern Munich        2       0.63          98.03
## 51           Manchester City        2       0.63          98.66
## 52         Athletico Madrid         1       0.32          98.98
## 53                     Eibar        1       0.32          99.30
## 54                     Elche        1       0.32          99.62
## 55                  Freiburg        1       0.32         100.00

## Warning: ‘guides(<scale> = FALSE)‘ is deprecated. Please use ‘guides(<scale> =
## "none")‘ instead.
```



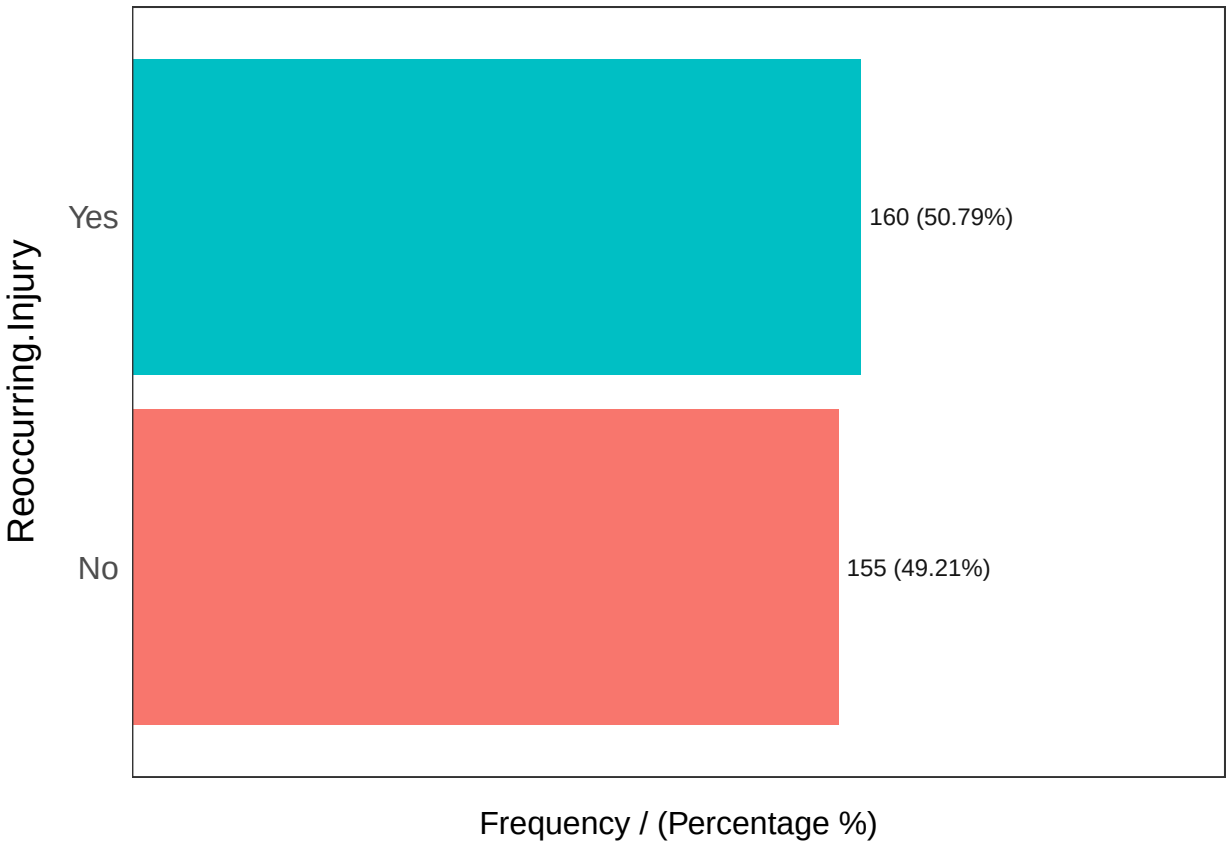

```
##     Position frequency percentage cumulative_perc
## 1   Defender       120      38.10           38.10
## 2 Midfielder       105      33.33           71.43
## 3    Forward        69      21.90           93.33
## 4     Keeper        21       6.67          100.00
```

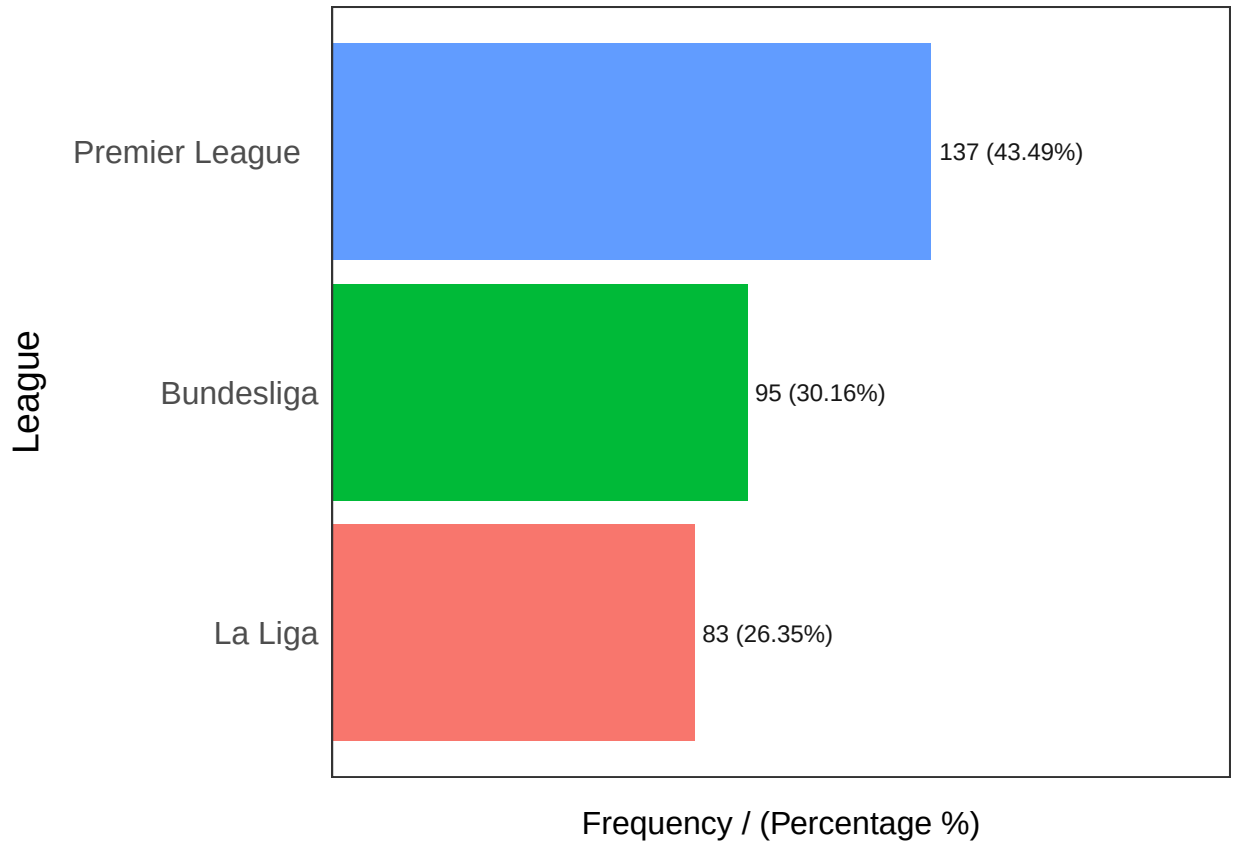
```
## Warning: ‘guides(<scale> = FALSE)‘ is deprecated. Please use ‘guides(<scale> =
## "none")‘ instead.
```




```
##   Reoccurring.Injury frequency percentage cumulative_perc
## 1                Yes       160      50.79           50.79
## 2                 No       155      49.21          100.00
```

```
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.
```


Reoccurring.Injury
Yes
No
160 (50.79%)
155 (49.21%)
Frequency / (Percentage %)




```
##            League frequency percentage cumulative_perc
## 1 Premier League         137      43.49           43.49
## 2       Bundesliga         95      30.16           73.65
## 3          La Liga         83      26.35          100.00
```

```
## [1] "Variables processed: Player, Injury, Potential.Return, Team, Position, Reoccurring.Injury, Leagu
```

```
library(tm)
```

```
## Loading required package: NLP
```

```
##
## Attaching package: 'NLP'
```

```
## The following object is masked from 'package:ggplot2':
##
##     annotate
```

```
library(wordcloud)
```

```
## Loading required package: RColorBrewer
```

```
library(RColorBrewer)
text<-Premier.League$Details
docs<-Corpus(VectorSource(text))
docs<-tm_map(docs, content_transformer(tolower))
```

```
## Warning in tm_map.SimpleCorpus(docs, content_transformer(tolower)):
## transformation drops documents
```

```
docs<-tm_map(docs, removeWords, stopwords("english"))
```

```
## Warning in tm_map.SimpleCorpus(docs, removeWords, stopwords("english")):
## transformation drops documents
```

```
dtm<-TermDocumentMatrix(docs)
matrix<-as.matrix(dtm)
words<-sort(rowSums(matrix), decreasing=TRUE)
df<-data.frame(word=names(words), freq=words)
set.seed(1234)
wordcloud(words=df$word, freq = df$freq, min.freq = 1, max.words = 200, random.order = FALSE, rot.per =
```


```
glimpse(Injury)
```

```
## Rows: 315
```

```
## Columns: 10
## $ Player             <chr> "David Luis Moreira Moarinho", "Hector Bellerin", "~
## $ Injury             <chr> "Hamstring", "Knock", "Groin/Hip/Pelvis", "Ankle/Fo~
## $ Date.of.Injury     <date> 2021-05-02, 2021-05-12, 2021-05-06, 2021-05-08, 20~
## $ Potential.Return   <chr> "5/19/21", "No Return Date", "5/19/21", "5/19/21", ~
## $ Age                <int> 34, 26, 28, 27, 20, 23, 28, 26, 20, 25, 26, 33, 29,~
## $ Team               <chr> "Arsenal", "Arsenal", "Arsenal", "Arsenal", "Arsena~
## $ Total.Days         <int> 33, 23, 13, 4, 3, 22, 6, 55, 65, 81, 69, 15, 22, 12~
## $ Position           <chr> "Defender", "Defender", "Midfielder", "Defender", "~
## $ Reoccurring.Injury <chr> "Yes", "No", "Yes", "Yes", "No", "Yes", "No", "Yes"~
## $ League             <chr> "Premier League  ", "Premier League  ", "Premier Le~
```

```
names<-c(1,2,3,4,6,8,9,10)
Injury[,names]<-lapply(Injury[,names], factor)
glimpse(Injury)
```

```
## Rows: 315
## Columns: 10
## $ Player             <fct> David Luis Moreira Moarinho, Hector Bellerin, Grani~
## $ Injury             <fct> Hamstring, Knock, Groin/Hip/Pelvis, Ankle/Foot, Ham~
## $ Date.of.Injury     <fct> 2021-05-02, 2021-05-12, 2021-05-06, 2021-05-08, 202~
## $ Potential.Return   <fct> 5/19/21, No Return Date, 5/19/21, 5/19/21, No Retur~
## $ Age                <int> 34, 26, 28, 27, 20, 23, 28, 26, 20, 25, 26, 33, 29,~
## $ Team               <fct> Arsenal, Arsenal, Arsenal, Arsenal, Arsenal, Aston ~
## $ Total.Days         <int> 33, 23, 13, 4, 3, 22, 6, 55, 65, 81, 69, 15, 22, 12~
## $ Position           <fct> Defender, Defender, Midfielder, Defender, Midfielde~
## $ Reoccurring.Injury <fct> Yes, No, Yes, Yes, No, Yes, No, Yes, No, Yes, No, Y~
## $ League             <fct> Premier League  , Premier League  , Premier League ~
```

```
Injury<-read.csv("~/Downloads/Injury.csv")
Injury$Date.of.Injury<-as.Date(Injury$Date.of.Injury, format= "%m/%d/%y")
injury.df <- Injury %>% select(Date.of.Injury, Age, Total.Days, Reoccurring.Injury)
injury.df
```

```
##     Date.of.Injury Age Total.Days Reoccurring.Injury
## 1       2021-05-02  34         33                Yes
## 2       2021-05-12  26         23                 No
## 3       2021-05-06  28         13                Yes
## 4       2021-05-08  27          4                Yes
## 5       2021-05-09  20          3                 No
## 6       2021-05-13  23         22                Yes
## 7       2021-05-13  28          6                 No
## 8       2021-04-10  26         55                Yes
## 9       2021-04-10  20         65                 No
## 10      2021-02-21  25         81                Yes
## 11      2021-04-10  26         69                 No
## 12      2021-04-29  33         15                Yes
## 13      2021-05-13  29         22                Yes
## 14      2020-07-20  28         12                 No
## 15      2021-05-19  30         26                Yes
## 16      2021-02-04  26        135                Yes
## 17      2021-05-09  29         18                 No
## 18      2020-12-19  20        182                Yes
```

```
## 19      2021-05-13  31         32              No
## 20      2021-05-13  29         37             Yes
## 21      2021-05-13  35         37              No
## 22      2021-05-04  29          2              No
## 23      2021-05-06  31          4              No
## 24      2021-05-23  31         12             Yes
## 25      2021-03-30  29         76             Yes
## 26      2021-03-22  30         84              No
## 27      2021-05-19  30          7             Yes
## 28      2021-05-19  21          4             Yes
## 29      2021-04-13  27         35             Yes
## 30      2021-05-23  29          3              No
## 31      2021-05-09  25         12             Yes
## 32      2021-05-21  21          2              No
## 33      2021-05-14  30          5              No
## 34      2021-05-21  30          5              No
## 35      2021-05-21  34         14             Yes
## 36      2021-05-19  22         31             Yes
## 37      2021-01-25  32        111             Yes
## 38      2021-01-30  33        135             Yes
## 39      2021-04-09  28         56              No
## 40      2020-08-26  20        282              No
## 41      2021-01-09  31        146             Yes
## 42      2021-05-23  31         12              No
## 43      2021-05-23  27          3              No
## 44      2021-05-09  26          9             Yes
## 45      2021-05-07  29          9             Yes
## 46      2021-05-21  32          2             Yes
## 47      2021-05-16  29         29             Yes
## 48      2021-04-12  18         68              No
## 49      2021-04-11  25         64              No
## 50      2021-01-14  30        151             Yes
## 51      2021-05-16  23         14             Yes
## 52      2021-05-19  26         16             Yes
## 53      2021-05-07  26          6             Yes
## 54      2021-05-23  24          3              No
## 55      2021-04-16  27         64             Yes
## 56      2021-05-15  24         11              No
## 57      2021-05-21  27          5             Yes
## 58      2020-09-10  29        277             Yes
## 59      2021-05-01  27         34             Yes
## 60      2021-05-23  25         10             Yes
## 61      2021-05-08  29          5              No
## 62      2021-05-17  23         14              No
## 63      2021-02-10  23        129              No
## 64      2021-05-08  33         27             Yes
## 65      2021-02-28  23        111              No
## 66      2021-05-23  20         12             Yes
## 67      2021-01-22  37        113             Yes
## 68      2021-05-16  27          3              No
## 69      2021-05-16  24          7             Yes
## 70      2021-02-20  30         92              No
## 71      2021-01-27  29        143             Yes
## 72      2020-11-11  23        220             Yes
```

```
## 73      2021-05-08  21         23               No
## 74      2021-05-08  35         11              Yes
## 75      2021-05-08  26         27              Yes
## 76      2021-05-15  25         20              Yes
## 77      2021-02-28  22         74               No
## 78      2020-10-17  29        245              Yes
## 79      2021-03-15  26         59              Yes
## 80      2021-05-14  32          7              Yes
## 81      2021-05-08  29         10              Yes
## 82      2021-03-28  25         83              Yes
## 83      2021-04-11  23          6              Yes
## 84      2021-05-09  28         16              Yes
## 85      2020-06-12  29        372              Yes
## 86      2021-05-19  24          4              Yes
## 87      2021-05-19  28          5               No
## 88      2021-05-19  24          4              Yes
## 89      2021-05-19  32          4               No
## 90      2021-05-19  21          4               No
## 91      2021-05-19  24         16              Yes
## 92      2021-04-08  27         57              Yes
## 93      2021-04-05  27         51              Yes
## 94      2021-05-23  26          3              Yes
## 95      2021-04-13  30         52              Yes
## 96      2021-05-05  18         30               No
## 97      2021-05-07  29         35              Yes
## 98      2021-03-20  26         91              Yes
## 99      2021-05-19  28          4              Yes
## 100     2021-04-11  24         69               No
## 101     2021-05-02  20         29               No
## 102     2021-05-14  23         31               No
## 103     2021-05-14  24         36               No
## 104     2020-09-27  27        265              Yes
## 105     2021-03-20  35         86               No
## 106     2021-05-21  24          7               No
## 107     2021-02-24  29         86               No
## 108     2021-05-12  25          5               No
## 109     2021-03-21  28         20               No
## 110     2021-04-21  31         44              Yes
## 111     2021-01-17  21        153               No
## 112     2021-05-21  22         29               No
## 113     2021-05-18  31          1              Yes
## 114     2021-05-16  28          3               No
## 115     2021-03-23  28         69              Yes
## 116     2021-05-23  25         12              Yes
## 117     2021-05-23  24        340              Yes
## 118     2021-05-23  31          3               No
## 119     2021-05-21  25         14               No
## 120     2021-04-04  37         71              Yes
## 121     2021-04-29  33         46              Yes
## 122     2021-05-19  36          4               No
## 123     2021-04-05  22         39               No
## 124     2021-04-17  27         48              Yes
## 125     2021-05-03  34         12              Yes
## 126     2021-05-09  31          5              Yes
```

```
## 127     2021-05-09  28         26              Yes
## 128     2021-05-04  32          5              Yes
## 129     2021-05-16  25         34              Yes
## 130     2020-11-29  30        202               No
## 131     2021-04-05  30         41               No
## 132     2021-05-04  19          5               No
## 133     2021-04-05  27         75              Yes
## 134     2021-04-09  21         71               No
## 135     2021-02-20  32         85              Yes
## 136     2021-05-14  20         21               No
## 137     2021-05-23  25          3              Yes
## 138     2021-04-24  31         56              Yes
## 139     2021-04-08  22         55              Yes
## 140     2021-05-08  26          8              Yes
## 141     2021-05-06  28         13              Yes
## 142     2021-05-03  30          9              Yes
## 143     2021-04-26  29         37               No
## 144     2021-04-24  28         39              Yes
## 145     2021-04-01  22         79               No
## 146     2021-05-09  25         17              Yes
## 147     2021-05-17  29         33              Yes
## 148     2020-11-08  18        223               No
## 149     2021-05-12  29         21               No
## 150     2020-12-30  28        171               No
## 151     2021-04-27  22         19              Yes
## 152     2021-05-17  25          5              Yes
## 153     2021-05-16  26         17               No
## 154     2021-05-16  36         17              Yes
## 155     2020-10-18  24        227               No
## 156     2021-05-07  27         43              Yes
## 157     2021-05-16  32         34              Yes
## 158     2021-04-11  28         52               No
## 159     2021-03-26  29         68               No
## 160     2021-05-16  32         17               No
## 161     2021-05-16  27         17              Yes
## 162     2021-02-08  36        131              Yes
## 163     2021-05-22  30          4               No
## 164     2021-05-16  21         34               No
## 165     2021-05-16  31         10               No
## 166     2021-05-03  28         19               No
## 167     2021-04-10  29         53               No
## 168     2021-05-09  25         24              Yes
## 169     2021-05-08  35         26              Yes
## 170     2021-04-29  29         34              Yes
## 171     2021-05-21  24         29               No
## 172     2021-05-17  31         33              Yes
## 173     2021-05-03  24         18              Yes
## 174     2020-06-13  26        371               No
## 175     2021-04-08  23         72               No
## 176     2021-04-29  20         51               No
## 177     2021-04-29  20         13               No
## 178     2021-04-24  24         56              Yes
## 179     2021-05-21  31         12               No
## 180     2021-05-06  35         27              Yes
```

```
## 181     2021-02-12  32        110              Yes
## 182     2021-05-21  27         29               No
## 183     2020-12-21  27        163               No
## 184     2021-05-07  27         26              Yes
## 185     2021-04-29  19         51               No
## 186     2021-04-21  24         42              Yes
## 187     2021-05-16  23         34               No
## 188     2021-05-03  30         18              Yes
## 189     2021-01-02  22        151               No
## 190     2020-08-24  26        282               No
## 191     2021-05-12  28          9               No
## 192     2020-12-14  25        153               No
## 193     2021-05-06  32         44              Yes
## 194     2021-05-02  30         31              Yes
## 195     2021-04-09  28         54               No
## 196     2021-02-17  23        105               No
## 197     2020-05-22  34        376               No
## 198     2021-03-07  25         76              Yes
## 199     2021-05-01  28         21              Yes
## 200     2021-05-10  33          2               No
## 201     2021-05-16  30          6               No
## 202     2021-05-16  29          6               No
## 203     2021-05-10  25          2              Yes
## 204     2021-05-10  21          6               No
## 205     2021-05-01  25         21              Yes
## 206     2021-05-11  28         23              Yes
## 207     2021-05-17  34         33               No
## 208     2021-05-06  24         27               No
## 209     2021-03-14  22         63              Yes
## 210     2021-05-10  29          2               No
## 211     2021-03-16  26         57               No
## 212     2021-05-12  32         11               No
## 213     2021-03-23  26         71              Yes
## 214     2021-04-26  24         27              Yes
## 215     2021-04-26  27         85              Yes
## 216     2021-05-02  25         31              Yes
## 217     2021-04-26  24         37              Yes
## 218     2021-05-09  21         41               No
## 219     2021-05-09  23         26              Yes
## 220     2021-01-04  33        137               No
## 221     2021-05-08  33         42              Yes
## 222     2020-06-20  33        364              Yes
## 223     2020-06-20  32        159               No
## 224     2021-03-27  16         84               No
## 225     2021-05-22  25         16              Yes
## 226     2021-03-20  21         91              Yes
## 227     2021-05-01  21         49               No
## 228     2021-05-16  35          3              Yes
## 229     2021-03-16  21         83              Yes
## 230     2021-04-25  22         55               No
## 231     2021-05-20  33         15               No
## 232     2021-05-15  24         19               No
## 233     2021-03-23  19         88               No
## 234     2021-05-22  27         28              Yes
```

```
## 235     2021-05-15  30         35              Yes
## 236     2021-05-22  21         11               No
## 237     2021-05-20  27         30               No
## 238     2021-05-16  20         13               No
## 239     2021-05-16  26         34               No
## 240     2021-01-05  20        165               No
## 241     2021-05-05  22         26              Yes
## 242     2021-05-05  19         45               No
## 243     2021-05-13  29          9               No
## 244     2021-05-15  23         11               No
## 245     2021-05-16  24         34               No
## 246     2021-05-05  28         33               No
## 247     2021-02-24  21         96               No
## 248     2021-05-10  26         14               No
## 249     2021-05-04  26         46               No
## 250     2021-03-17  23         68              Yes
## 251     2021-05-05  30         26               No
## 252     2021-02-13  28        107              Yes
## 253     2021-05-15  28         11               No
## 254     2021-05-15  37         22              Yes
## 255     2021-05-22  27         28              Yes
## 256     2021-04-28  19         52               No
## 257     2021-03-06  29        105               No
## 258     2021-05-19  28         30              Yes
## 259     2021-05-16  22         31              Yes
## 260     2021-05-13  23          9              Yes
## 261     2021-05-03  23         47              Yes
## 262     2021-05-15  24         35              Yes
## 263     2021-05-22  31         28              Yes
## 264     2020-10-10  29        252              Yes
## 265     2021-02-28  23        111              Yes
## 266     2021-04-30  28         50               No
## 267     2021-05-08  23         42               No
## 268     2021-02-23  32         88              Yes
## 269     2021-01-23  33        115               No
## 270     2021-05-14  22         20              Yes
## 271     2021-05-08  25         42              Yes
## 272     2021-02-15  30        121              Yes
## 273     2021-05-08  26         23              Yes
## 274     2021-05-13  23         12               No
## 275     2021-05-15  23         16              Yes
## 276     2021-05-09  30         11               No
## 277     2021-05-15  26          7               No
## 278     2021-05-06  34         12              Yes
## 279     2021-04-04  36         57               No
## 280     2021-05-09  24         22               No
## 281     2021-05-06  22         44               No
## 282     2021-05-09  21         22               No
## 283     2021-03-25  18         86               No
## 284     2021-05-22  24         16               No
## 285     2021-05-12  25         10               No
## 286     2021-05-11  28         39              Yes
## 287     2021-05-22  24         28               No
## 288     2021-05-11  30         39               No
```

```
## 289     2021-02-23  27         97               Yes
## 290     2021-05-05  30         26               Yes
## 291     2020-09-27  31        265                No
## 292     2020-12-22  22        179               Yes
## 293     2020-05-13  21         13                No
## 294     2021-04-24  21         37                No
## 295     2020-09-10  31        282                No
## 296     2021-05-15  26          9                No
## 297     2021-05-06  30         18                No
## 298     2021-04-24  24         27               Yes
## 299     2021-05-06  26         18               Yes
## 300     2021-05-06  27         18                No
## 301     2020-08-02  30        321               Yes
## 302     2021-04-06  20         55                No
## 303     2020-04-20  29         28                No
## 304     2021-05-14  20         15                No
## 305     2021-05-06  18         44               Yes
## 306     2020-12-28  18        154               Yes
## 307     2021-04-07  23         61                No
## 308     2021-03-31  23         55               Yes
## 309     2021-03-20  21         91               Yes
## 310     2021-05-07  20         43                No
## 311     2021-05-15  23          9               Yes
## 312     2021-04-01  19         60                No
## 313     2021-03-11  21         81                No
## 314     2021-05-13  25         11                No
## 315     2021-05-22  28         28                No
```

```
# Train Models
library(tidymodels)
```

```
## Registered S3 method overwritten by 'tune':
##   method                   from
##   required_pkgs.model_spec parsnip

## -- Attaching packages --------------------------------------- tidymodels 0.1.3 --

## v broom        0.7.8          v rsample      0.1.0
## v dials        0.0.9          v tune         0.1.5
## v infer        0.5.4          v workflows    0.2.2
## v modeldata    0.1.0          v workflowsets 0.0.2
## v parsnip      0.1.6.9000     v yardstick    0.0.8
## v recipes      0.1.16

## -- Conflicts ------------------------------------------ tidymodels_conflicts() --
## x NLP::annotate()      masks ggplot2::annotate()
## x scales::discard()    masks purrr::discard()
## x dplyr::filter()      masks stats::filter()
## x recipes::fixed()     masks stringr::fixed()
## x dplyr::lag()         masks stats::lag()
## x yardstick::spec()    masks readr::spec()
## x dplyr::src()         masks Hmisc::src()
## x recipes::step()      masks stats::step()
```

```
## x dplyr::summarize()   masks Hmisc::summarize()
## x parsnip::translate() masks Hmisc::translate()
## * Use tidymodels_prefer() to resolve common conflicts.
```

```
set.seed(1234)
injury.split<-injury.df %>% initial_split(strata = Reoccurring.Injury)
injury.split
```

```
## <Analysis/Assess/Total>
## <236/79/315>
```

```
injury_train<-training(injury.split)
injury_test<-testing(injury.split)
injury_train
```

```
##     Date.of.Injury Age Total.Days Reoccurring.Injury
## 5       2021-05-09  20          3                  No
## 7       2021-05-13  28          6                  No
## 9       2021-04-10  20         65                  No
## 14      2020-07-20  28         12                  No
## 19      2021-05-13  31         32                  No
## 21      2021-05-13  35         37                  No
## 22      2021-05-04  29          2                  No
## 23      2021-05-06  31          4                  No
## 30      2021-05-23  29          3                  No
## 32      2021-05-21  21          2                  No
## 34      2021-05-21  30          5                  No
## 39      2021-04-09  28         56                  No
## 42      2021-05-23  31         12                  No
## 43      2021-05-23  27          3                  No
## 48      2021-04-12  18         68                  No
## 49      2021-04-11  25         64                  No
## 61      2021-05-08  29          5                  No
## 62      2021-05-17  23         14                  No
## 63      2021-02-10  23        129                  No
## 65      2021-02-28  23        111                  No
## 70      2021-02-20  30         92                  No
## 77      2021-02-28  22         74                  No
## 90      2021-05-19  21          4                  No
## 96      2021-05-05  18         30                  No
## 101     2021-05-02  20         29                  No
## 103     2021-05-14  24         36                  No
## 105     2021-03-20  35         86                  No
## 106     2021-05-21  24          7                  No
## 107     2021-02-24  29         86                  No
## 114     2021-05-16  28          3                  No
## 118     2021-05-23  31          3                  No
## 119     2021-05-21  25         14                  No
## 122     2021-05-19  36          4                  No
## 131     2021-04-05  30         41                  No
## 132     2021-05-04  19          5                  No
## 134     2021-04-09  21         71                  No
## 136     2021-05-14  20         21                  No
```

```
## 143     2021-04-26  29         37              No
## 145     2021-04-01  22         79              No
## 148     2020-11-08  18        223              No
## 149     2021-05-12  29         21              No
## 150     2020-12-30  28        171              No
## 153     2021-05-16  26         17              No
## 155     2020-10-18  24        227              No
## 158     2021-04-11  28         52              No
## 159     2021-03-26  29         68              No
## 163     2021-05-22  30          4              No
## 164     2021-05-16  21         34              No
## 165     2021-05-16  31         10              No
## 166     2021-05-03  28         19              No
## 167     2021-04-10  29         53              No
## 171     2021-05-21  24         29              No
## 175     2021-04-08  23         72              No
## 182     2021-05-21  27         29              No
## 183     2020-12-21  27        163              No
## 185     2021-04-29  19         51              No
## 189     2021-01-02  22        151              No
## 190     2020-08-24  26        282              No
## 191     2021-05-12  28          9              No
## 192     2020-12-14  25        153              No
## 195     2021-04-09  28         54              No
## 197     2020-05-22  34        376              No
## 200     2021-05-10  33          2              No
## 204     2021-05-10  21          6              No
## 207     2021-05-17  34         33              No
## 208     2021-05-06  24         27              No
## 210     2021-05-10  29          2              No
## 211     2021-03-16  26         57              No
## 212     2021-05-12  32         11              No
## 218     2021-05-09  21         41              No
## 220     2021-01-04  33        137              No
## 223     2020-06-20  32        159              No
## 224     2021-03-27  16         84              No
## 227     2021-05-01  21         49              No
## 230     2021-04-25  22         55              No
## 231     2021-05-20  33         15              No
## 233     2021-03-23  19         88              No
## 236     2021-05-22  21         11              No
## 237     2021-05-20  27         30              No
## 238     2021-05-16  20         13              No
## 239     2021-05-16  26         34              No
## 242     2021-05-05  19         45              No
## 243     2021-05-13  29          9              No
## 244     2021-05-15  23         11              No
## 245     2021-05-16  24         34              No
## 246     2021-05-05  28         33              No
## 247     2021-02-24  21         96              No
## 249     2021-05-04  26         46              No
## 251     2021-05-05  30         26              No
## 253     2021-05-15  28         11              No
## 256     2021-04-28  19         52              No
```

```
## 257     2021-03-06  29        105               No
## 266     2021-04-30  28         50               No
## 267     2021-05-08  23         42               No
## 274     2021-05-13  23         12               No
## 277     2021-05-15  26          7               No
## 280     2021-05-09  24         22               No
## 281     2021-05-06  22         44               No
## 282     2021-05-09  21         22               No
## 284     2021-05-22  24         16               No
## 285     2021-05-12  25         10               No
## 288     2021-05-11  30         39               No
## 293     2020-05-13  21         13               No
## 294     2021-04-24  21         37               No
## 295     2020-09-10  31        282               No
## 296     2021-05-15  26          9               No
## 297     2021-05-06  30         18               No
## 300     2021-05-06  27         18               No
## 302     2021-04-06  20         55               No
## 303     2020-04-20  29         28               No
## 304     2021-05-14  20         15               No
## 307     2021-04-07  23         61               No
## 310     2021-05-07  20         43               No
## 312     2021-04-01  19         60               No
## 314     2021-05-13  25         11               No
## 315     2021-05-22  28         28               No
## 3       2021-05-06  28         13              Yes
## 4       2021-05-08  27          4              Yes
## 6       2021-05-13  23         22              Yes
## 10      2021-02-21  25         81              Yes
## 12      2021-04-29  33         15              Yes
## 13      2021-05-13  29         22              Yes
## 16      2021-02-04  26        135              Yes
## 18      2020-12-19  20        182              Yes
## 20      2021-05-13  29         37              Yes
## 24      2021-05-23  31         12              Yes
## 27      2021-05-19  30          7              Yes
## 29      2021-04-13  27         35              Yes
## 31      2021-05-09  25         12              Yes
## 35      2021-05-21  34         14              Yes
## 36      2021-05-19  22         31              Yes
## 37      2021-01-25  32        111              Yes
## 38      2021-01-30  33        135              Yes
## 41      2021-01-09  31        146              Yes
## 44      2021-05-09  26          9              Yes
## 45      2021-05-07  29          9              Yes
## 46      2021-05-21  32          2              Yes
## 47      2021-05-16  29         29              Yes
## 50      2021-01-14  30        151              Yes
## 51      2021-05-16  23         14              Yes
## 52      2021-05-19  26         16              Yes
## 53      2021-05-07  26          6              Yes
## 57      2021-05-21  27          5              Yes
## 59      2021-05-01  27         34              Yes
## 60      2021-05-23  25         10              Yes
```

```
## 64      2021-05-08  33         27              Yes
## 66      2021-05-23  20         12              Yes
## 67      2021-01-22  37        113              Yes
## 69      2021-05-16  24          7              Yes
## 72      2020-11-11  23        220              Yes
## 74      2021-05-08  35         11              Yes
## 75      2021-05-08  26         27              Yes
## 76      2021-05-15  25         20              Yes
## 79      2021-03-15  26         59              Yes
## 81      2021-05-08  29         10              Yes
## 85      2020-06-12  29        372              Yes
## 88      2021-05-19  24          4              Yes
## 91      2021-05-19  24         16              Yes
## 92      2021-04-08  27         57              Yes
## 93      2021-04-05  27         51              Yes
## 94      2021-05-23  26          3              Yes
## 95      2021-04-13  30         52              Yes
## 97      2021-05-07  29         35              Yes
## 99      2021-05-19  28          4              Yes
## 110     2021-04-21  31         44              Yes
## 115     2021-03-23  28         69              Yes
## 116     2021-05-23  25         12              Yes
## 120     2021-04-04  37         71              Yes
## 121     2021-04-29  33         46              Yes
## 124     2021-04-17  27         48              Yes
## 125     2021-05-03  34         12              Yes
## 126     2021-05-09  31          5              Yes
## 128     2021-05-04  32          5              Yes
## 135     2021-02-20  32         85              Yes
## 137     2021-05-23  25          3              Yes
## 138     2021-04-24  31         56              Yes
## 139     2021-04-08  22         55              Yes
## 140     2021-05-08  26          8              Yes
## 141     2021-05-06  28         13              Yes
## 142     2021-05-03  30          9              Yes
## 144     2021-04-24  28         39              Yes
## 147     2021-05-17  29         33              Yes
## 151     2021-04-27  22         19              Yes
## 157     2021-05-16  32         34              Yes
## 161     2021-05-16  27         17              Yes
## 168     2021-05-09  25         24              Yes
## 170     2021-04-29  29         34              Yes
## 172     2021-05-17  31         33              Yes
## 173     2021-05-03  24         18              Yes
## 178     2021-04-24  24         56              Yes
## 180     2021-05-06  35         27              Yes
## 181     2021-02-12  32        110              Yes
## 186     2021-04-21  24         42              Yes
## 188     2021-05-03  30         18              Yes
## 193     2021-05-06  32         44              Yes
## 198     2021-03-07  25         76              Yes
## 205     2021-05-01  25         21              Yes
## 206     2021-05-11  28         23              Yes
## 209     2021-03-14  22         63              Yes
```

```
## 213     2021-03-23  26         71                Yes
## 215     2021-04-26  27         85                Yes
## 217     2021-04-26  24         37                Yes
## 219     2021-05-09  23         26                Yes
## 221     2021-05-08  33         42                Yes
## 222     2020-06-20  33        364                Yes
## 225     2021-05-22  25         16                Yes
## 226     2021-03-20  21         91                Yes
## 228     2021-05-16  35          3                Yes
## 229     2021-03-16  21         83                Yes
## 234     2021-05-22  27         28                Yes
## 241     2021-05-05  22         26                Yes
## 250     2021-03-17  23         68                Yes
## 255     2021-05-22  27         28                Yes
## 258     2021-05-19  28         30                Yes
## 259     2021-05-16  22         31                Yes
## 260     2021-05-13  23          9                Yes
## 261     2021-05-03  23         47                Yes
## 262     2021-05-15  24         35                Yes
## 264     2020-10-10  29        252                Yes
## 265     2021-02-28  23        111                Yes
## 268     2021-02-23  32         88                Yes
## 270     2021-05-14  22         20                Yes
## 271     2021-05-08  25         42                Yes
## 272     2021-02-15  30        121                Yes
## 275     2021-05-15  23         16                Yes
## 286     2021-05-11  28         39                Yes
## 289     2021-02-23  27         97                Yes
## 290     2021-05-05  30         26                Yes
## 292     2020-12-22  22        179                Yes
## 298     2021-04-24  24         27                Yes
## 299     2021-05-06  26         18                Yes
## 301     2020-08-02  30        321                Yes
## 305     2021-05-06  18         44                Yes
## 306     2020-12-28  18        154                Yes
## 309     2021-03-20  21         91                Yes
## 311     2021-05-15  23          9                Yes
```

```
injury_test
```

```
##     Date.of.Injury Age Total.Days Reoccurring.Injury
## 1       2021-05-02  34         33                Yes
## 2       2021-05-12  26         23                 No
## 8       2021-04-10  26         55                Yes
## 11      2021-04-10  26         69                 No
## 15      2021-05-19  30         26                Yes
## 17      2021-05-09  29         18                 No
## 25      2021-03-30  29         76                Yes
## 26      2021-03-22  30         84                 No
## 28      2021-05-19  21          4                Yes
## 33      2021-05-14  30          5                 No
## 40      2020-08-26  20        282                 No
## 54      2021-05-23  24          3                 No
## 55      2021-04-16  27         64                Yes
```

```
## 56      2021-05-15  24         11              No
## 58      2020-09-10  29        277             Yes
## 68      2021-05-16  27          3              No
## 71      2021-01-27  29        143             Yes
## 73      2021-05-08  21         23              No
## 78      2020-10-17  29        245             Yes
## 80      2021-05-14  32          7             Yes
## 82      2021-03-28  25         83             Yes
## 83      2021-04-11  23          6             Yes
## 84      2021-05-09  28         16             Yes
## 86      2021-05-19  24          4             Yes
## 87      2021-05-19  28          5              No
## 89      2021-05-19  32          4              No
## 98      2021-03-20  26         91             Yes
## 100     2021-04-11  24         69              No
## 102     2021-05-14  23         31              No
## 104     2020-09-27  27        265             Yes
## 108     2021-05-12  25          5              No
## 109     2021-03-21  28         20              No
## 111     2021-01-17  21        153              No
## 112     2021-05-21  22         29              No
## 113     2021-05-18  31          1             Yes
## 117     2021-05-23  24        340             Yes
## 123     2021-04-05  22         39              No
## 127     2021-05-09  28         26             Yes
## 129     2021-05-16  25         34             Yes
## 130     2020-11-29  30        202              No
## 133     2021-04-05  27         75             Yes
## 146     2021-05-09  25         17             Yes
## 152     2021-05-17  25          5             Yes
## 154     2021-05-16  36         17             Yes
## 156     2021-05-07  27         43             Yes
## 160     2021-05-16  32         17              No
## 162     2021-02-08  36        131             Yes
## 169     2021-05-08  35         26             Yes
## 174     2020-06-13  26        371              No
## 176     2021-04-29  20         51              No
## 177     2021-04-29  20         13              No
## 179     2021-05-21  31         12              No
## 184     2021-05-07  27         26             Yes
## 187     2021-05-16  23         34              No
## 194     2021-05-02  30         31             Yes
## 196     2021-02-17  23        105              No
## 199     2021-05-01  28         21             Yes
## 201     2021-05-16  30          6              No
## 202     2021-05-16  29          6              No
## 203     2021-05-10  25          2             Yes
## 214     2021-04-26  24         27             Yes
## 216     2021-05-02  25         31             Yes
## 232     2021-05-15  24         19              No
## 235     2021-05-15  30         35             Yes
## 240     2021-01-05  20        165              No
## 248     2021-05-10  26         14              No
## 252     2021-02-13  28        107             Yes
```

```
## 254     2021-05-15  37         22                Yes
## 263     2021-05-22  31         28                Yes
## 269     2021-01-23  33        115                 No
## 273     2021-05-08  26         23                Yes
## 276     2021-05-09  30         11                 No
## 278     2021-05-06  34         12                Yes
## 279     2021-04-04  36         57                 No
## 283     2021-03-25  18         86                 No
## 287     2021-05-22  24         28                 No
## 291     2020-09-27  31        265                 No
## 308     2021-03-31  23         55                Yes
## 313     2021-03-11  21         81                 No
```

```
# Model Specification
lm_spec<- linear_reg() %>% set_engine(engine = "lm")
lm_spec
```

```
## Linear Regression Model Specification (regression)
##
## Computational engine: lm
```

```
lm_fit<-lm_spec %>% fit(Total.Days~., data=injury_train)
lm_fit
```

```
## parsnip model object
##
## Fit time:  4ms
##
## Call:
## stats::lm(formula = Total.Days ~ ., data = data)
##
## Coefficients:
##           (Intercept)         Date.of.Injury                    Age
##             13547.4790                -0.7204                -0.4304
## Reoccurring.InjuryYes
##                8.8635
```

```
tidy(lm_fit)
```

```
## # A tibble: 4 x 5
##   term                  estimate std.error statistic  p.value
##   <chr>                    <dbl>     <dbl>     <dbl>    <dbl>
## 1 (Intercept)           13547.     639.        21.2   3.63e-56
## 2 Date.of.Injury           -0.720    0.0341   -21.1   5.07e-56
## 3 Age                      -0.430    0.580     -0.741 4.59e- 1
## 4 Reoccurring.InjuryYes     8.86     5.06       1.75  8.14e- 2
```

```
# Random Forest Model
library(ranger)
rf_spec<-rand_forest(mode = "regression") %>% set_engine('ranger')
rf_spec
```

```
## Random Forest Model Specification (regression)
##
## Computational engine: ranger
```

```
rf_fit<- rf_spec %>% fit(Total.Days ~., data=injury_train)
rf_fit
```

```
## parsnip model object
##
## Fit time:  36ms
## Ranger result
##
## Call:
##  ranger::ranger(x = maybe_data_frame(x), y = y, num.threads = 1,      verbose = FALSE, seed = sample
##
## Type:                             Regression
## Number of trees:                  500
## Sample size:                      236
## Number of independent variables:  3
## Mtry:                             1
## Target node size:                 5
## Variable importance mode:         none
## Splitrule:                        variance
## OOB prediction error (MSE):       1620.098
## R squared (OOB):                  0.6162403
```

```
library(tidymodels)
library(stats)
library(dplyr)
library(magrittr)
```

```
##
## Attaching package: 'magrittr'
```

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

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

```
results.train<-lm_fit %>% predict(new_data= injury_train) %>% mutate(truth=injury_train$Total.Days, mod
results.train
```

```
## # A tibble: 472 x 3
##    .pred truth model
##    <dbl> <int> <chr>
##  1  26.9     3 lm
##  2  20.5     6 lm
##  3  47.8    65 lm
```

```
##  4 235.     12 lm
##  5  19.3    32 lm
##  6  17.5    37 lm
##  7  26.6     2 lm
##  8  24.3     4 lm
##  9  12.9     3 lm
## 10  17.8     2 lm
## # ... with 462 more rows
```

```
results.test<-lm_fit %>% predict(new_data= injury_test) %>% mutate(truth=injury_test$Total.Days, model=
results.test
```

```
## # A tibble: 158 x 3
##    .pred truth model
##    <dbl> <int> <chr>
##  1  34.8    33 lm
##  2  22.1    23 lm
##  3  54.0    55 lm
##  4  45.2    69 lm
##  5  24.2    26 lm
##  6  23.0    18 lm
##  7  60.7    76 lm
##  8  57.1    84 lm
##  9  28.1     4 lm
## 10  19.0     5 lm
## # ... with 148 more rows
```

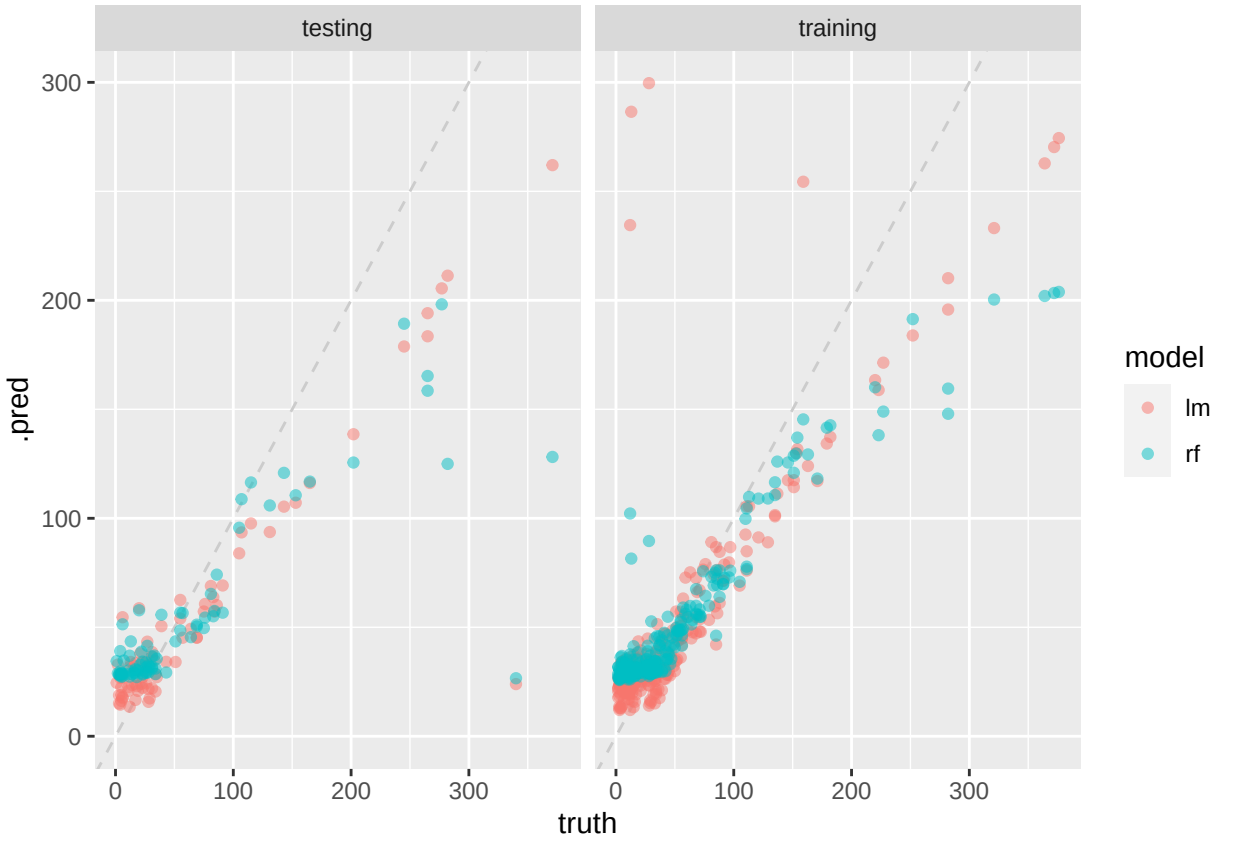
```
results.train %>% group_by(model) %>% rmse(truth=truth, estimate =.pred)
```

```
## # A tibble: 2 x 4
##   model .metric .estimator .estimate
##   <chr> <chr>   <chr>          <dbl>
## 1 lm    rmse    standard        37.8
## 2 rf    rmse    standard        31.4
```

```
results.test %>% group_by(model) %>% rmse(truth=truth, estimate =.pred)
```

```
## # A tibble: 2 x 4
##   model .metric .estimator .estimate
##   <chr> <chr>   <chr>          <dbl>
## 1 lm    rmse    standard        46.0
## 2 rf    rmse    standard        55.9
```

```
# rf higher on testing data which is not a good model

results.test %>% mutate(train="testing") %>% bind_rows(results.train %>% mutate(train="training")) %>% g
```




```
set.seed(1234)
injury_folds<-vfold_cv(injury_train, strata=Reoccurring.Injury)
library(multilevelmod)
```

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.