Sandbox

Eren Akgunduz

20 April 2023

The data

```r
library(tidyverse)
```

Clean it

```r
df <- read.csv("data/all.csv", stringsAsFactors = FALSE)
df$Damage <- gsub(".*\\((\\d+)\\).*", "\\1", df$Damage)
df$Damage <- gsub("(\\d+)\\s*\\((\\d+)\\)", "pmax(\\1, \\2)", df$Damage)
df$Damage <- gsub("(\\d+)\\*(\\d+)", "\\1+\\2", df$Damage)
df$Damage <- gsub("(\\d+)x(\\d+)", "\\1*\\2", df$Damage)
df$Damage <- sapply(df$Damage, function(x) eval(parse(text = x)))
df$Damage <- as.numeric(df$Damage)
df$Stun <- gsub(".*\\((\\d+)\\).*", "\\1", df$Stun)
df$Stun <- gsub("(\\d+)\\s*\\((\\d+)\\)", "pmax(\\1, \\2)", df$Stun)
df$Stun <- gsub("(\\d+)\\*(\\d+)", "\\1+\\2", df$Stun)
df$Stun <- gsub("(\\d+)x(\\d+)", "\\1*\\2", df$Stun)
df$Stun <- sapply(df$Stun, function(x) eval(parse(text = x)))
df$Stun <- as.numeric(df$Stun)
df$health <- as.numeric(df$health)
df$stun <- as.numeric(df$stun)
df$vgauge1 <- as.numeric(df$vgauge1)
df$vgauge2 <- as.numeric(df$vgauge2)
df$fDash <- as.numeric(df$fDash)
df$bDash <- as.numeric(df$bDash)
df$fWalk <- as.numeric(df$fWalk)
df$bWalk <- as.numeric(df$bWalk)
df$throwHurt <- as.numeric(df$throwHurt)
df$throwRange <- as.numeric(df$throwRange)
na_rows <- df[!complete.cases(df), ]
na_rows
write.csv(df, "data/all.csv", row.names = FALSE)
```

```r
df <- read.csv("data/all.csv", stringsAsFactors = FALSE)

# Replace values in parentheses with the minimum value
regx <- ".*?(\\-?\\d+)\\((\\-?\\d+)\\).*"
df$onBlock <- gsub(regx, "\\1", df$onBlock)
df$onBlock <- gsub("(\\-?\\d+)[\\*/](\\-?\\d+)", "\\1", df$onBlock)

# Use ifelse to replace NA values w/ NA and non-NA values with the minimum value
df$onBlock <- ifelse(
```

```r
  is.na(df$onBlock), # condition: check if value is NA
  NA, # value to replace with if condition is TRUE
  # value to replace with if condition is FALSE
  pmin(as.numeric(df$onBlock), as.numeric(gsub(regx, "\\2", df$onBlock)))
)

# Convert column to numeric
df$onBlock <- as.numeric(df$onBlock)

write.csv(df, "data/all.csv", row.names = FALSE)
```

```r
df <- read.csv("data/all.csv", stringsAsFactors = FALSE)
# Check for missing values
na_rows <- df[!complete.cases(df), ]
na_rows
```

```
##      Character                                Move onBlock
## 402        Dan                  Haoh Gadoken (hold)     NA
## 516         Ed                        Psycho Cannon     NA
## 539      Falke                             Luftjagd     NA
## 821       Kage              Senha Kassatsu (hold)      NA
## 1122     Menat                    Divine Retribution     NA
## 1141     Menat                     Soul Sphere Ankh     NA
## 1315    Rashid                EX Eagle Spike (wind)     NA
## 1338    R.Mika                              Jump LP     NA
## 1340    R.Mika                              Jump HP     NA
## 1341    R.Mika                              Jump LK     NA
## 1342    R.Mika                              Jump MK     NA
## 1343    R.Mika                              Jump HK     NA
## 1554      Seth Install Art | [Menat] Soul Sphere: Ankh     NA
## 1717   Zangief              Super Russian Kick (hold)     NA
##                              plnCmd airmove followUp projectile moveType
## 402              HP+HK (VT1) Hold    False    False       True vtrigger
## 516                   HP+HK (VT1)    False    False       True vtrigger
## 539                    d+MK (air)     True    False      False   normal
## 821              MP+MK (VS1) (hold)   False     True      False   vskill
## 1122                        f+MK    False     True      False   normal
## 1141                 qcf+P (recall)   False     True       True  special
## 1315                  qcb+KK (wind)    True    False      False  special
## 1338                        u+LP     True    False      False   normal
## 1340                        u+HP     True    False      False   normal
## 1341                        u+LK     True    False      False   normal
## 1342                        u+MK     True    False      False   normal
## 1343                        u+HK     True    False      False   normal
## 1554 MP+MK (Menat Soul Sphere: Ankh)   False     True       True   vskill
## 1717             MP+MK (VS2) (hold)   False    False      False   vskill
##      health stun vgauge1 vgauge2 fDash bDash  fWalk  bWalk throwHurt throwRange
## 402    1025  950     300     600    16    21 0.0470 0.0320     0.250      0.800
## 516    1025 1050     600     900    16    23 0.0470 0.0340     0.250      0.800
## 539     975 1000     600     600    20    25 0.0400 0.0300     0.250      0.850
## 821     925  950     600     900    16    21 0.0520 0.0350     0.250      0.800
## 1122    950  950     900     600    20    24 0.0400 0.0320     0.250      0.850
## 1141    950  950     900     600    20    24 0.0400 0.0320     0.250      0.850
## 1315    950  950     600     600    15    24 0.0400 0.0360     0.402      0.852
```
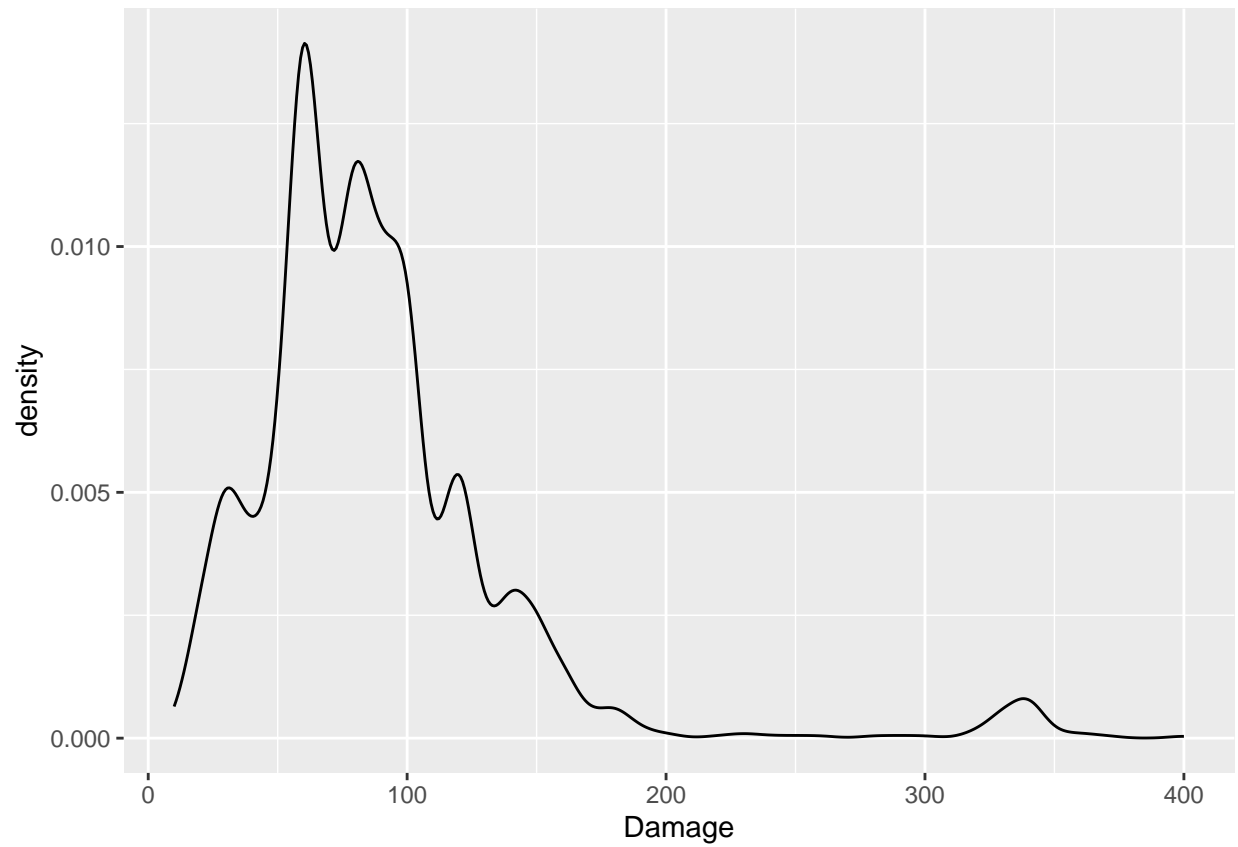
```
## 1338      950 1000      600      900      18      24 0.0420 0.0300      0.250      0.850
## 1340      950 1000      600      900      18      24 0.0420 0.0300      0.250      0.850
## 1341      950 1000      600      900      18      24 0.0420 0.0300      0.250      0.850
## 1342      950 1000      600      900      18      24 0.0420 0.0300      0.250      0.850
## 1343      950 1000      600      900      18      24 0.0420 0.0300      0.250      0.850
## 1554      900  900      600      600      16      23 0.0470 0.0320      0.250      0.800
## 1717     1075 1100      900      900      25      25 0.0305 0.0235      0.400      0.900
##      Damage Stun
## 402      220  150
## 516      140  175
## 539       60  100
## 821       90  150
## 1122      60  100
## 1141      40   70
## 1315     160  200
## 1338      40   70
## 1340      90  150
## 1341      40   70
## 1342      70  100
## 1343      90  150
## 1554      40   70
## 1717     100  120
```

KDE

```r
ggplot(df, aes(Damage)) +
  geom_density()
```

```
ggplot(df, aes(Stun)) +
  geom_density()
```