War Analysis

Evan Meade

1/17/2021

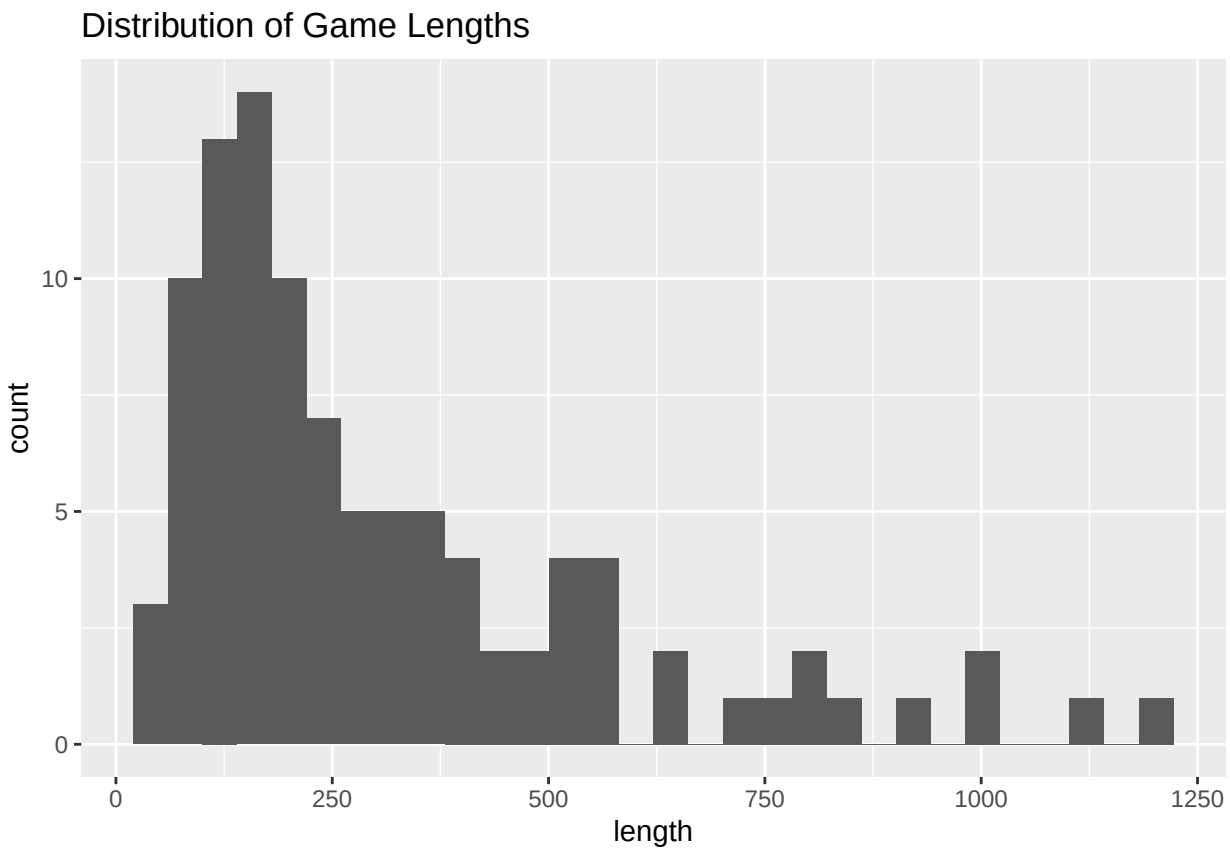
Data Loading

```
library(ggplot2)
games <- read.csv("war_game_data.csv")
turns <- read.csv("war_turn_data.csv")
```

Introductory Plots

```
ggplot(data = games) +
  geom_histogram(mapping = aes(x = length)) +
  labs(title = "Distribution of Game Lengths")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



This appears to be a long tailed distribution, perhaps a gamma distribution or something similar. Higher sampling is required to have a better idea of its shape.

Feature Engineering

```
higher <- function(row) {
  v0 <- as.integer(row[8])
  v1 <- as.integer(row[10])
  if (v0 > v1) {
    return(0)
  } else if (v1 > v0) {
    return(1)
  } else {
    return(-1)
  }
}

games$higher_start <- apply(games, 1, higher)
games$higher_start <- as.integer(games$higher_start)
```

```
higher_match <- function(row) {
  actual <- as.integer(row[5])
  pred <- as.integer(row[11])
```

```
  if (actual == pred) {
    return(1)
  } else {
    return(0)
  }
}

games$higher_match <- apply(games, 1, higher_match)
games$higher_match <- as.integer(games$higher_match)
```