Stat Distributions

Braeden Kuether

7/23/2019

Statistical Conditions

Using statistics from Basketball Reference. Player has played over 20 games played and over 6 minutes per game.

Loading Packages

```
lapply(
  c('heatmaply', # devtools::install_github('talgalili/heatmaply')
    'mclust', #install.packages('mclust')
    'dplyr', # devtools::install_github('hadley/dplyr')
    'plotly', # devtools::install_github('ropensci/plotly')
    'nbastatR', # devtools::install_github('abresler/nbastatR')
    'purrr', # devtools::install_github('hadley/purrr')
    'tidyr' # devtools::install_github('hadley/tidyr')
  ),
  library,
  character.only = T
)

options(digits = 4)
```

Using statistics from Basketball Reference

Points Per Game Distribution

```
stats_2019 <- stat_2019[order(stat_2019$ptsPerGame),]

#Mean
mean(stats_2019$ptsPerGame)
```

```
## [1] 10.08
```

```
#Standard Deviation
sd(stats_2019$ptsPerGame)
```

```
## [1] 5.935
```

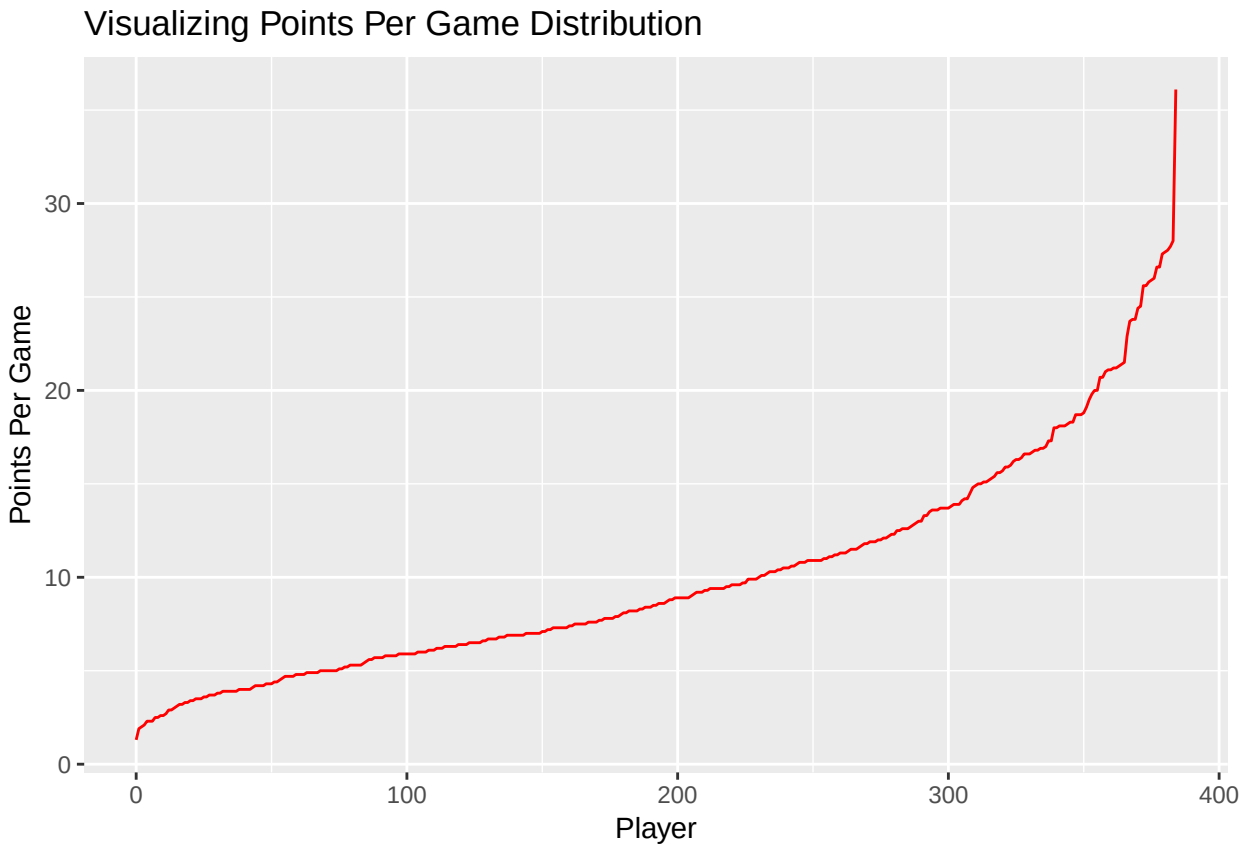
```
#Spread
quantile(stats_2019$ptsPerGame)
```

```
##   0%  25%  50%  75% 100%
##  1.3  5.8  8.5 12.9 36.1
```

```
x <- seq(0, size, 1)
y <- dnorm(stats_2019$ptsPerGame, mean(stats_2019$ptsPerGame), sd(stats_2019$ptsPerGame))

#Graph
ggplot(stats_2019) +
```

```
  geom_line(aes(x = x, y = ptsPerGame), color = "red") +
  xlab("Player") +
  ylab("Points Per Game") +
  ggtitle("Visualizing Points Per Game Distribution")
```



Points Per Minute Distribution

```
stats_2019 <- stat_2019[order(stat_2019$ptsPerMinute),]

#Mean
mean(stats_2019$ptsPerMinute)
```

```
## [1] 0.429
```

```
#Standard Deviation
sd(stats_2019$ptsPerMinute)
```

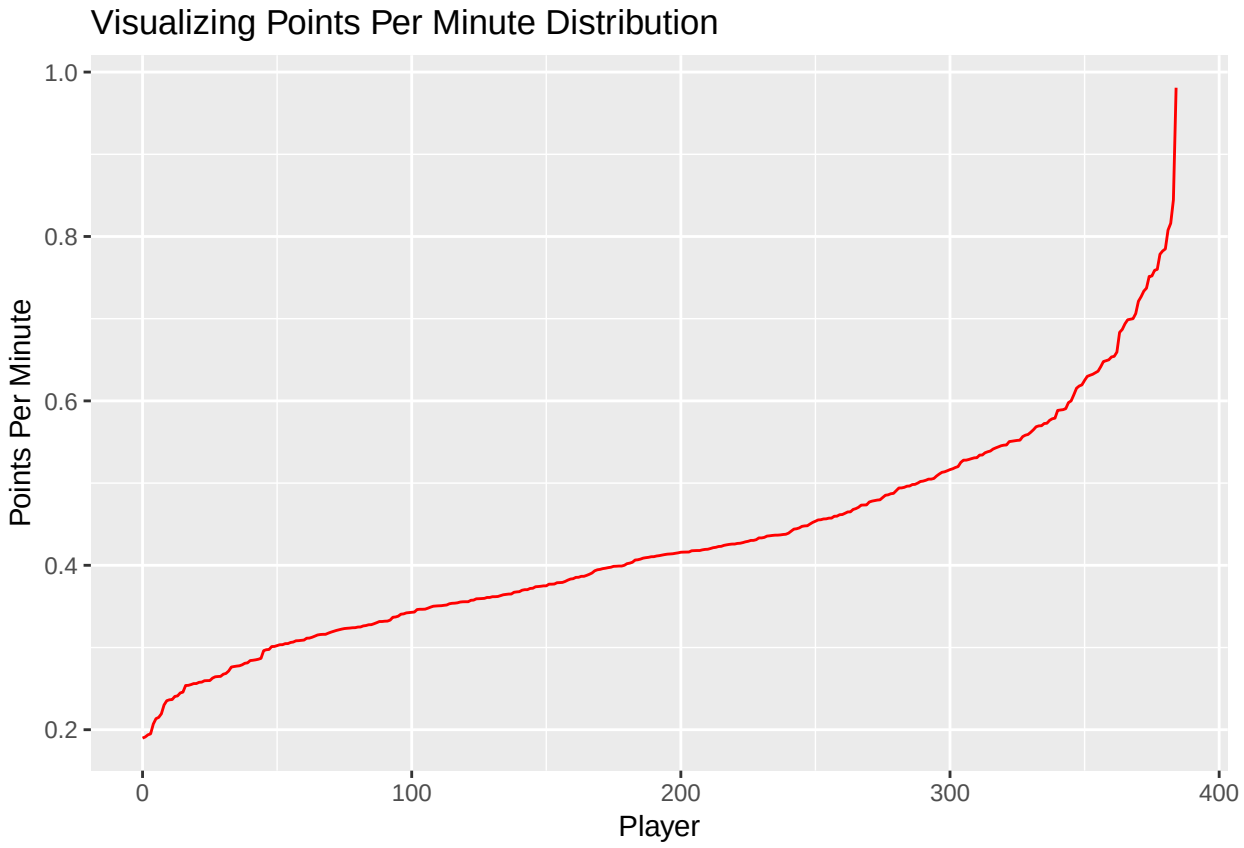
```
## [1] 0.1298
```

```
#Spread
quantile(stats_2019$ptsPerMinute)
```

```
##     0%    25%    50%    75%   100%
## 0.1898 0.3405 0.4118 0.5000 0.9810
```

```
x <- seq(0, size, 1)
y <- dnorm(stats_2019$ptsPerMinute, mean(stats_2019$ptsPerMinute), sd(stats_2019$ptsPerMinute))
```

```r
#Graph
ggplot(stats_2019) +
  geom_line(aes(x = x, y = ptsPerMinute), color = "red") +
  xlab("Player") +
  ylab("Points Per Minute") +
  ggtitle("Visualizing Points Per Minute Distribution")
```



Assits Per Game Distribution

```r
stats_2019 <- stat_2019[order(stat_2019$astPerGame),]

#Mean
mean(stats_2019$astPerGame)
```

```
## [1] 2.256
```

```r
#Standard Deviation
sd(stats_2019$astPerGame)
```

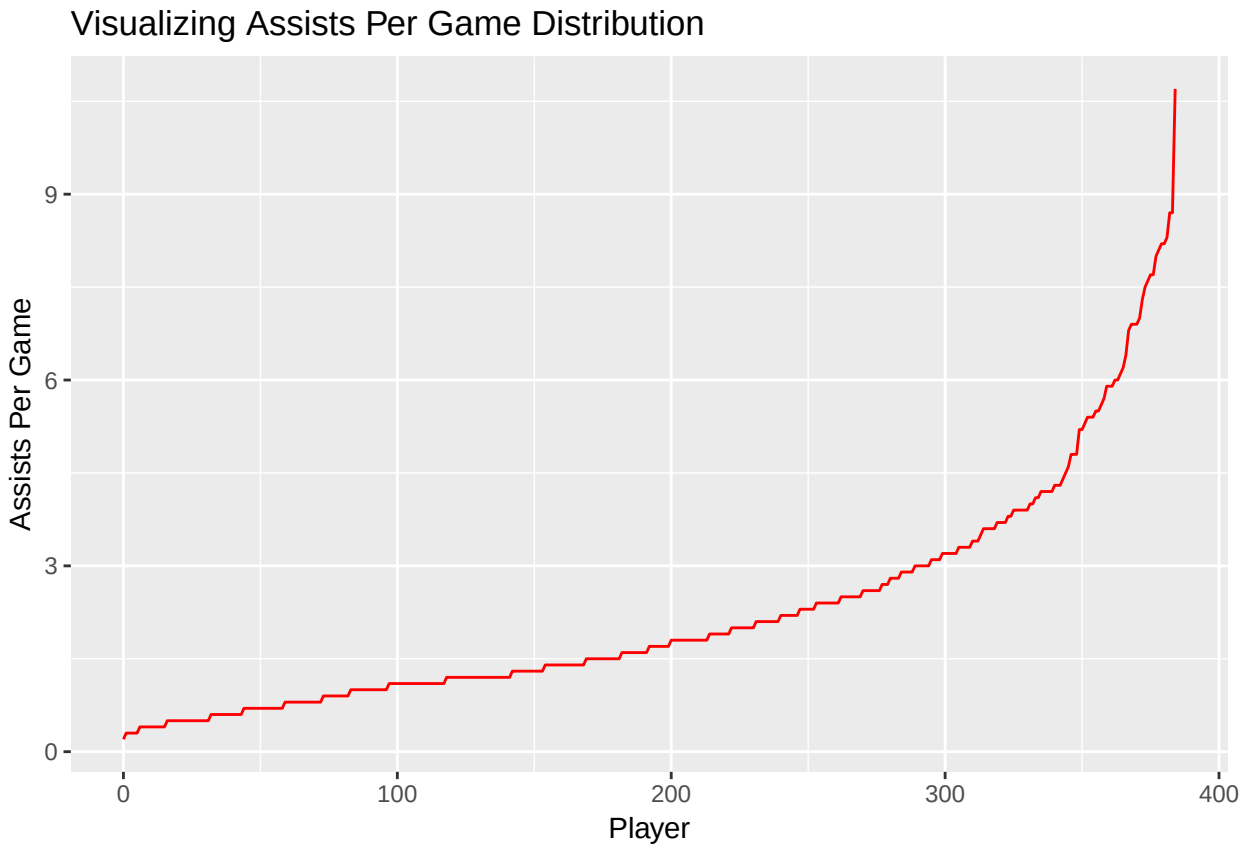
```
## [1] 1.826
```

```r
#Spread
quantile(stats_2019$astPerGame)
```

```
##   0%  25%  50%  75% 100%
##  0.2  1.0  1.7  2.9 10.7
```

```r
x <- seq(0, size, 1)
y <- dnorm(stats_2019$astPerGame, mean(stats_2019$astPerGame), sd(stats_2019$astPerGame))

#Graph
ggplot(stats_2019) +
  geom_line(aes(x = x, y = astPerGame), color = "red") +
  xlab("Player") +
  ylab("Assists Per Game") +
  ggtitle("Visualizing Assists Per Game Distribution")
```



Assists Per Minute Distribution

```r
stats_2019 <- stat_2019[order(stat_2019$astPerMinute),]

#Mean
mean(stats_2019$astPerMinute)
```

```
## [1] 0.09515
```

```r
#Standard Deviation
sd(stats_2019$astPerMinute)
```

```
## [1] 0.05584
```

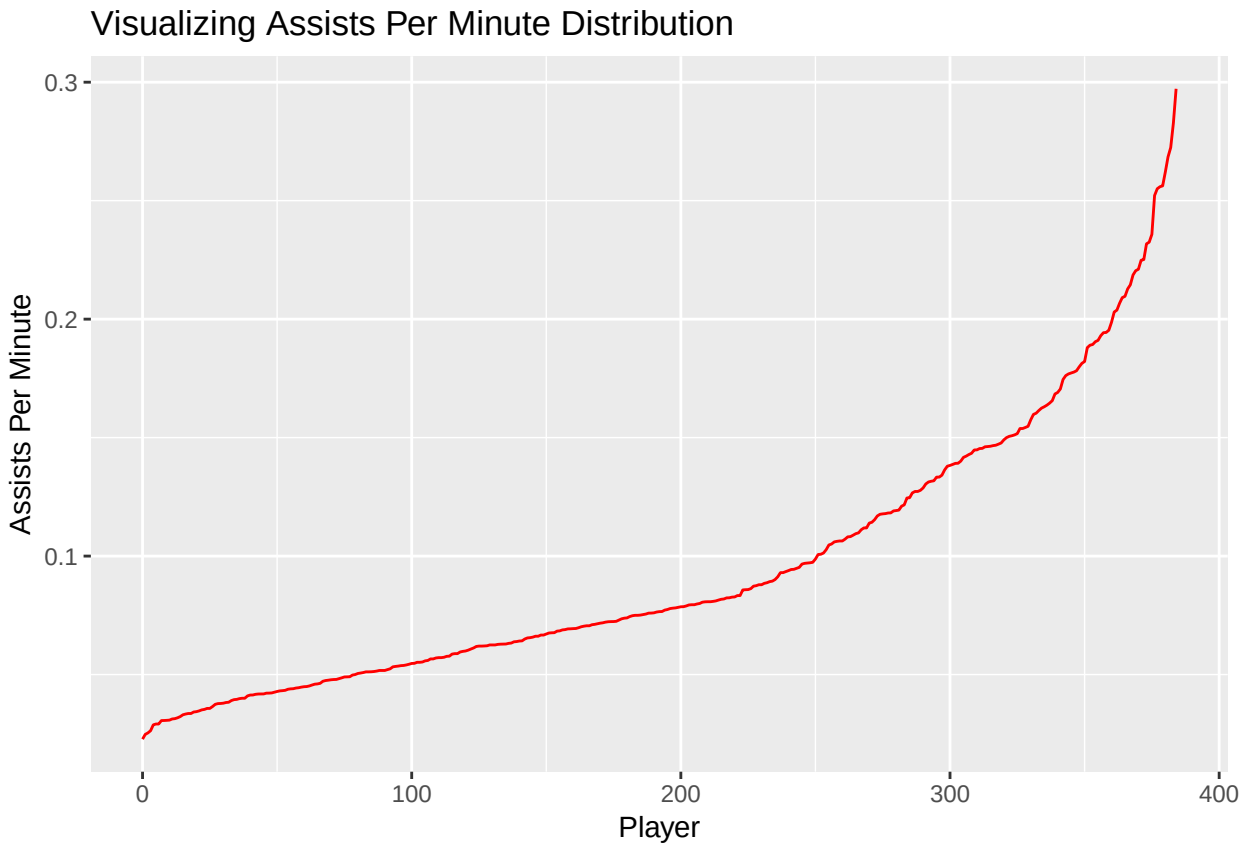
```r
#Spread
quantile(stats_2019$astPerMinute)
```

```
##      0%     25%     50%     75%    100%
```

```
## 0.02273 0.05376 0.07656 0.12727 0.29722
```

```
x <- seq(0, size, 1)
y <- dnorm(stats_2019$astPerMinute, mean(stats_2019$astPerMinute), sd(stats_2019$astPerMinute))

#Graph
ggplot(stats_2019) +
  geom_line(aes(x = x, y = astPerMinute), color = "red") +
  xlab("Player") +
  ylab("Assists Per Minute") +
  ggtitle("Visualizing Assists Per Minute Distribution")
```



Total Rebounds Per Game Distribution

```
stats_2019 <- stat_2019[order(stat_2019$trbPerGame),]

#Mean
mean(stats_2019$trbPerGame)
```

```
## [1] 4.109
```

```
#Standard Deviation
sd(stats_2019$trbPerGame)
```

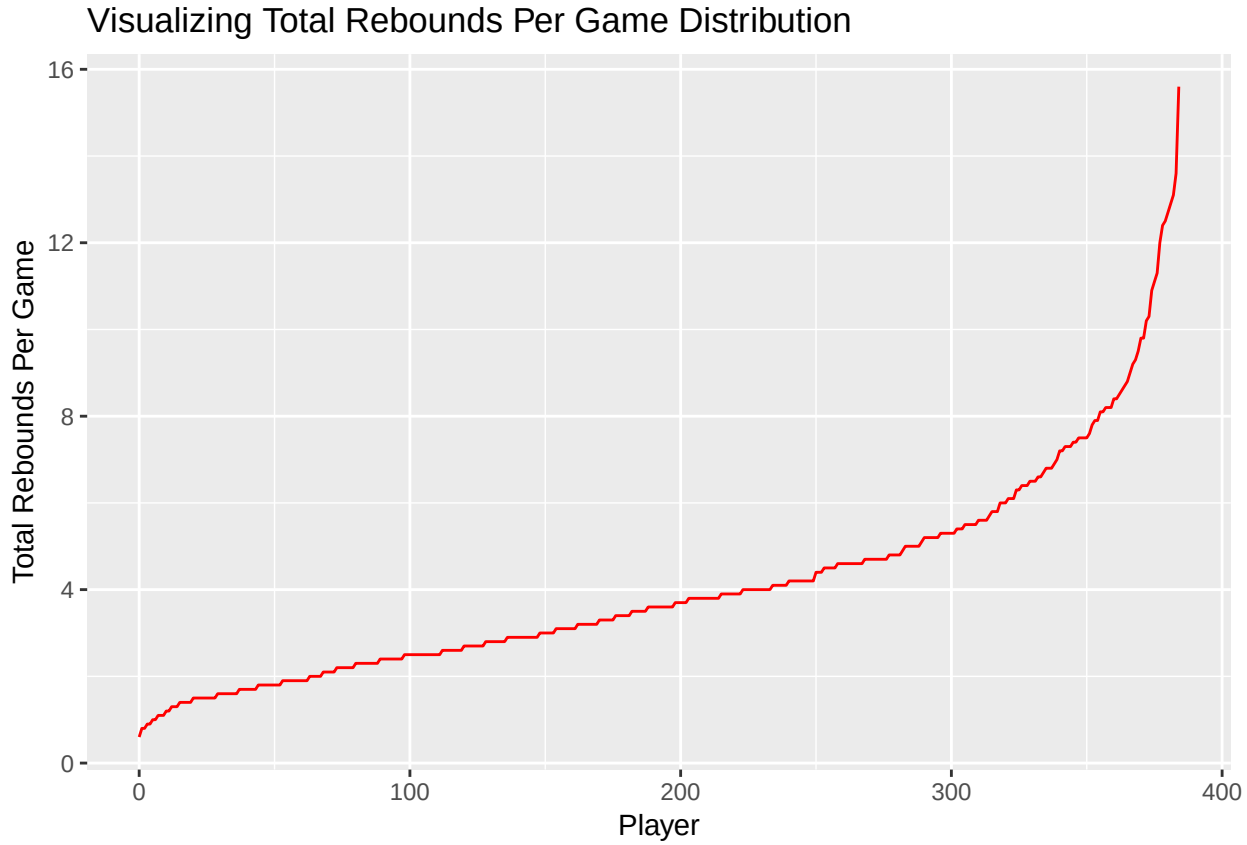
```
## [1] 2.447
```

```
#Spread
quantile(stats_2019$trbPerGame)
```

```
##   0%  25%  50%  75% 100%
##  0.6  2.4  3.6  5.0 15.6
```

```
x <- seq(0, size, 1)
y <- dnorm(stats_2019$trbPerGame, mean(stats_2019$trbPerGame), sd(stats_2019$trbPerGame))

#Graph
ggplot(stats_2019) +
  geom_line(aes(x = x, y = trbPerGame), color = "red") +
  xlab("Player") +
  ylab("Total Rebounds Per Game") +
  ggtitle("Visualizing Total Rebounds Per Game Distribution")
```



Total Rebounds Per Minute Distribution

```
stats_2019 <- stat_2019[order(stat_2019$trbPerMinute),]

#Mean
mean(stats_2019$trbPerMinute)
```

```
## [1] 0.1857
```

```
#Standard Deviation
sd(stats_2019$trbPerMinute)
```

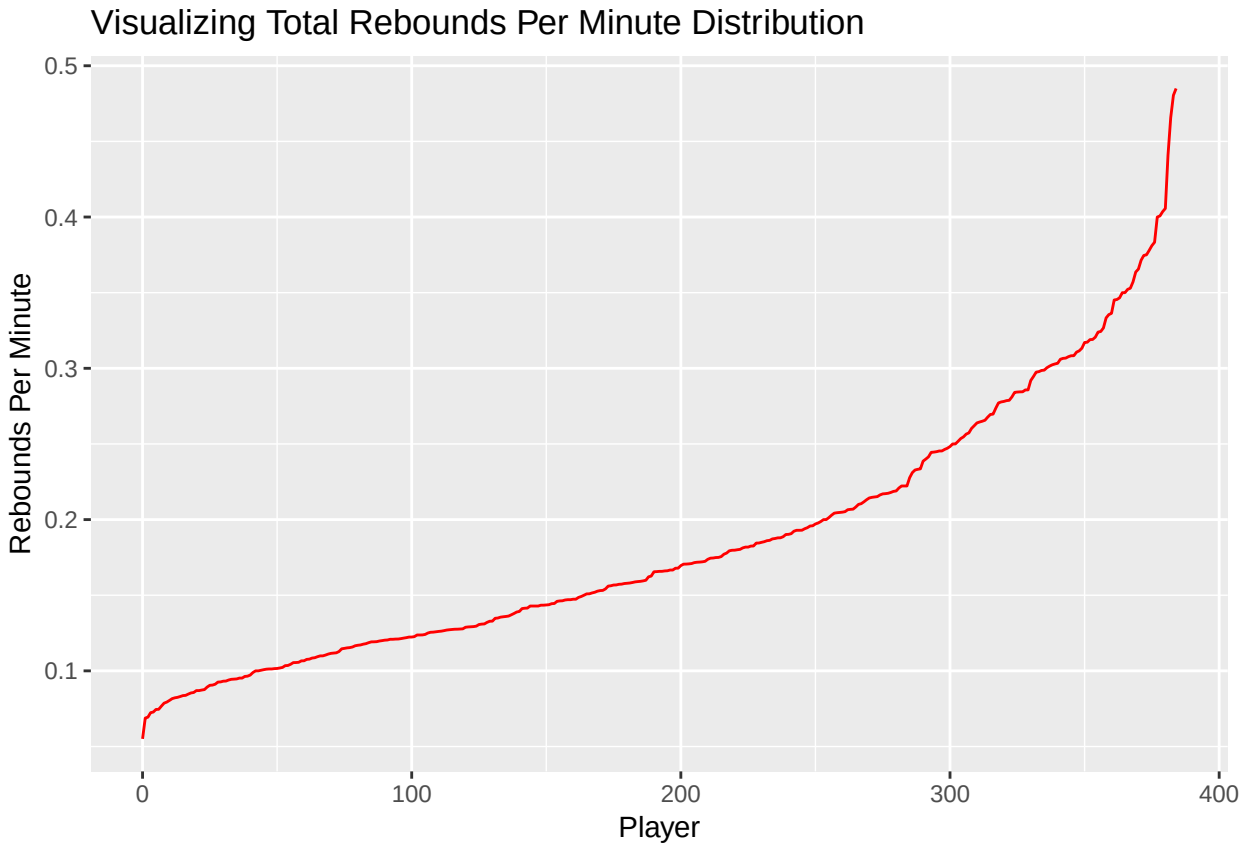
```
## [1] 0.08422
```

```
#Spread
quantile(stats_2019$trbPerMinute)
```

```
##      0%     25%     50%     75%    100%
## 0.05505 0.12132 0.16578 0.23323 0.48498
```

```r
x <- seq(0, size, 1)
y <- dnorm(stats_2019$trbPerMinute, mean(stats_2019$trbPerMinute), sd(stats_2019$trbPerMinute))

#Graph
ggplot(stats_2019) +
  geom_line(aes(x = x, y = trbPerMinute), color = "red") +
  xlab("Player") +
  ylab("Rebounds Per Minute") +
  ggtitle("Visualizing Total Rebounds Per Minute Distribution")
```



Offensive Rebounds Per Game Distribution

```r
stats_2019 <- stat_2019[order(stat_2019$orbPerGame),]

#Mean
mean(stats_2019$orbPerGame)
```

```
## [1] 0.9351
```

```r
#Standard Deviation
sd(stats_2019$orbPerGame)
```

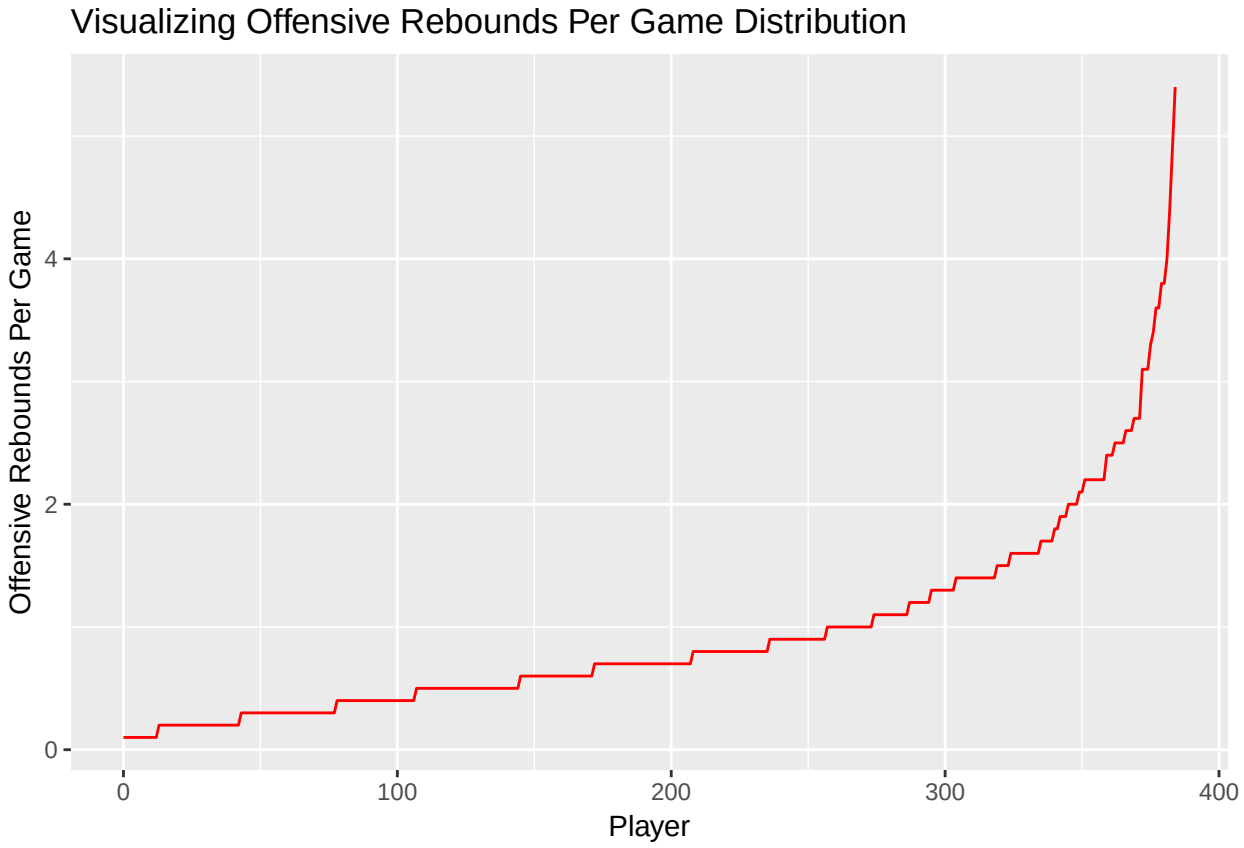
```
## [1] 0.8018
```

```r
#Spread
quantile(stats_2019$orbPerGame)
```

```
##   0%  25%  50%  75% 100%
##  0.1  0.4  0.7  1.2  5.4
```

```r
x <- seq(0, size, 1)
y <- dnorm(stats_2019$orbPerGame, mean(stats_2019$orbPerGame), sd(stats_2019$orbPerGame))

#Graph
ggplot(stats_2019) +
  geom_line(aes(x = x, y = orbPerGame), color = "red") +
  xlab("Player") +
  ylab("Offensive Rebounds Per Game") +
  ggtitle("Visualizing Offensive Rebounds Per Game Distribution")
```



Offensive Rebounds Per Minute Distribution

```r
stats_2019 <- stat_2019[order(stat_2019$orbPerMinute),]

#Mean
mean(stats_2019$orbPerMinute)
```

```
## [1] 0.04352
```

```r
#Standard Deviation
sd(stats_2019$orbPerMinute)
```

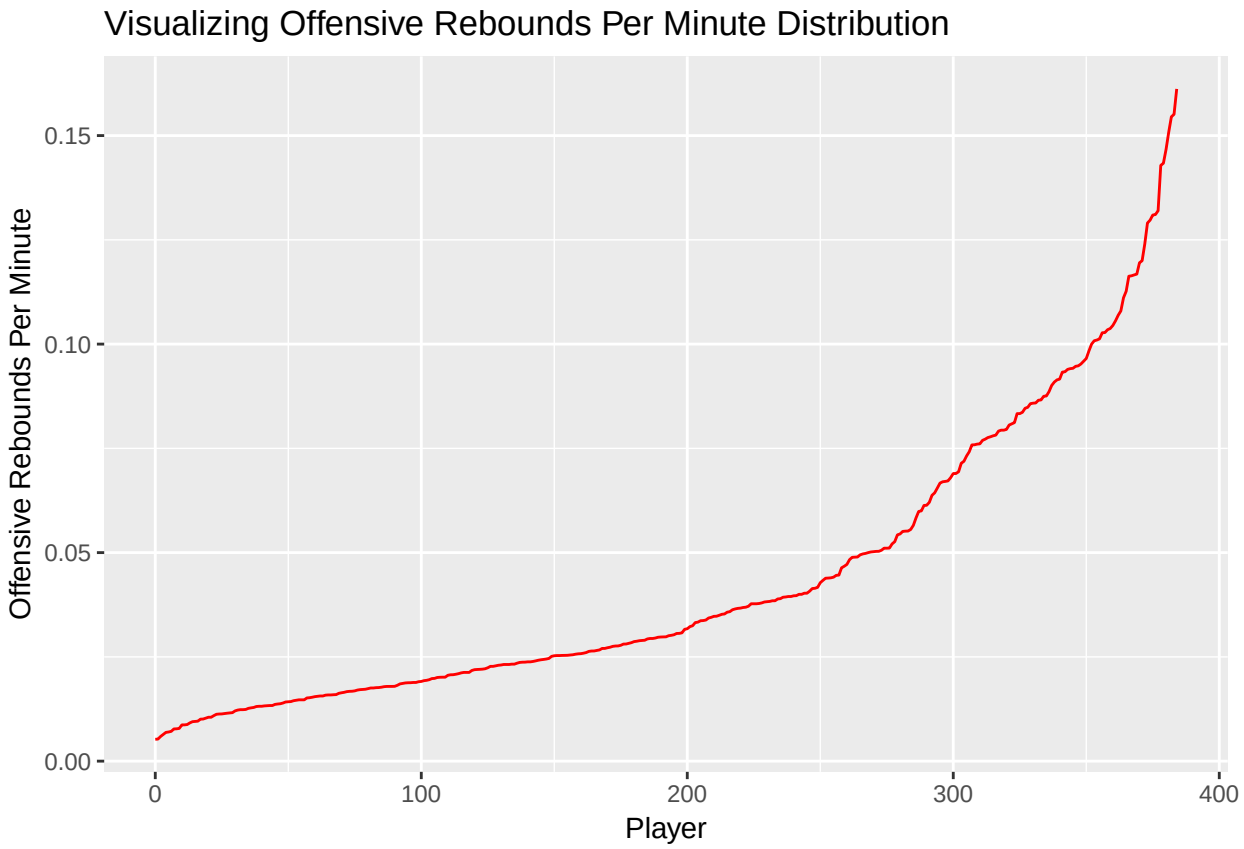
```
## [1] 0.03358
```

```
#Spread
quantile(stats_2019$orbPerMinute)
```

```
##       0%      25%      50%      75%     100%
## 0.005236 0.018809 0.029810 0.060000 0.161194
```

```
x <- seq(0, size, 1)
y <- dnorm(stats_2019$orbPerMinute, mean(stats_2019$orbPerMinute), sd(stats_2019$orbPerMinute))

#Graph
ggplot(stats_2019) +
  geom_line(aes(x = x, y = orbPerMinute), color = "red") +
  xlab("Player") +
  ylab("Offensive Rebounds Per Minute") +
  ggtitle("Visualizing Offensive Rebounds Per Minute Distribution")
```



Defensive Rebounds Per Game Distribution

```
stats_2019 <- stat_2019[order(stat_2019$drbPerGame),]

#Mean
mean(stats_2019$drbPerGame)
```

```
## [1] 3.175
```

```r
#Standard Deviation
sd(stats_2019$drbPerGame)
```

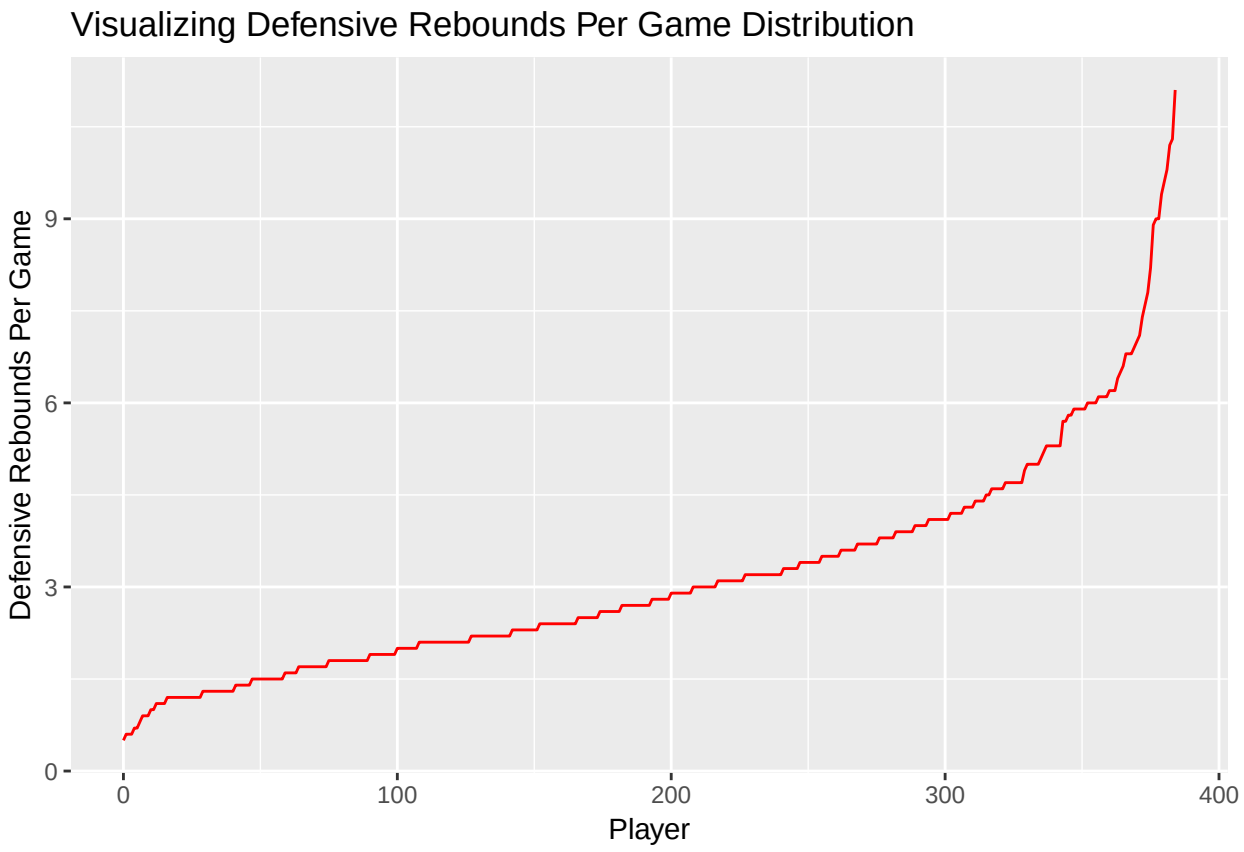
```
## [1] 1.807
```

```r
#Spread
quantile(stats_2019$drbPerGame)
```

```
##   0%  25%  50%  75% 100%
##  0.5  1.9  2.7  3.9 11.1
```

```r
x <- seq(0, size, 1)
y <- dnorm(stats_2019$drbPerGame, mean(stats_2019$drbPerGame), sd(stats_2019$drbPerGame))

#Graph
ggplot(stats_2019) +
  geom_line(aes(x = x, y = drbPerGame), color = "red") +
  xlab("Player") +
  ylab("Defensive Rebounds Per Game") +
  ggtitle("Visualizing Defensive Rebounds Per Game Distribution")
```



Defensive Rebounds Per Minute Distribution

```r
stats_2019 <- stat_2019[order(stat_2019$drbPerMinute),]

#Mean
mean(stats_2019$drbPerMinute)
```

```
## [1] 0.1422
```

```
#Standard Deviation
sd(stats_2019$drbPerMinute)
```

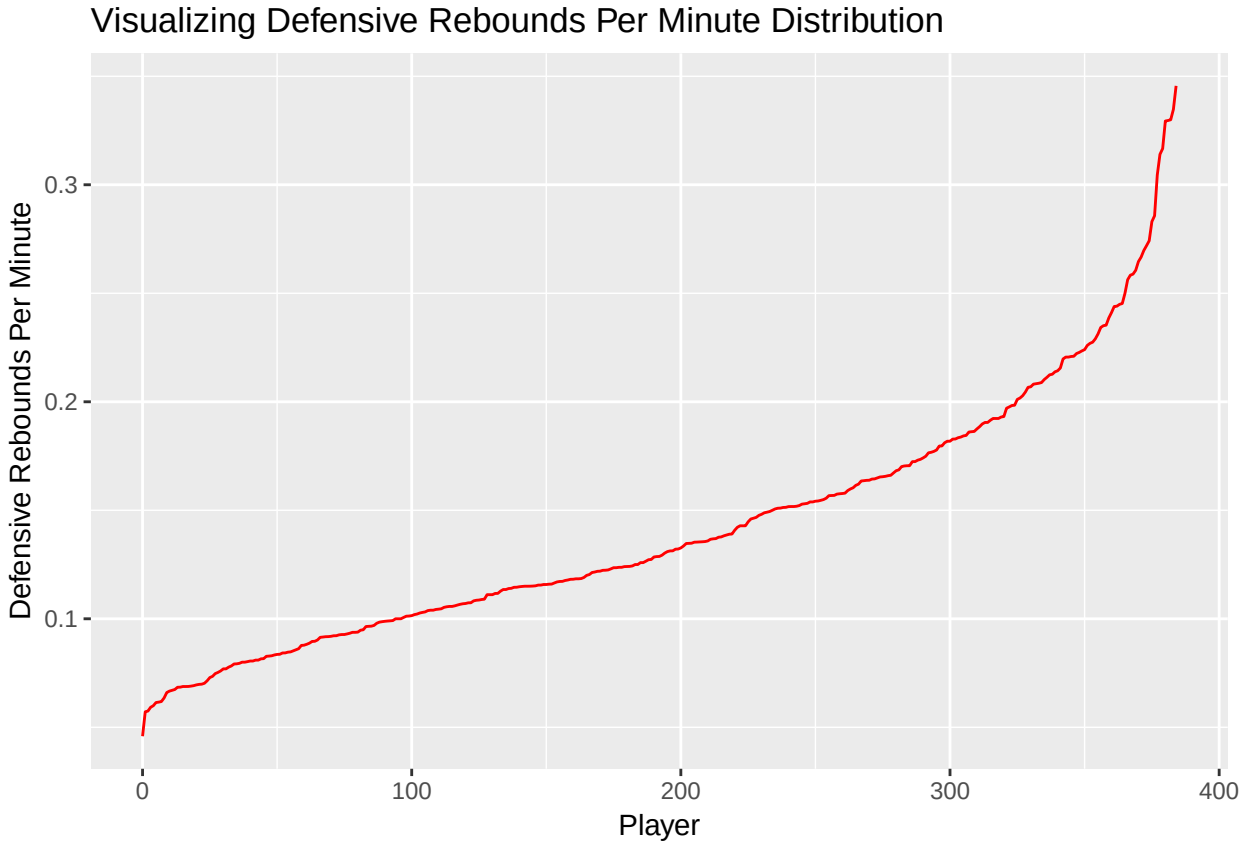
```
## [1] 0.05709
```

```
#Spread
quantile(stats_2019$drbPerMinute)
```

```
##      0%     25%     50%     75%    100%
## 0.04587 0.10000 0.12871 0.17308 0.34559
```

```
x <- seq(0, size, 1)
y <- dnorm(stats_2019$drbPerMinute, mean(stats_2019$drbPerMinute), sd(stats_2019$drbPerMinute))

#Graph
ggplot(stats_2019) +
  geom_line(aes(x = x, y = drbPerMinute), color = "red") +
  xlab("Player") +
  ylab("Defensive Rebounds Per Minute") +
  ggtitle("Visualizing Defensive Rebounds Per Minute Distribution")
```



Steals Per Game Distribution

```
stats_2019 <- stat_2019[order(stat_2019$stlPerGame),]

#Mean
mean(stats_2019$stlPerGame)
```

```
## [1] 0.7083
```

```
#Standard Deviation
sd(stats_2019$stlPerGame)
```

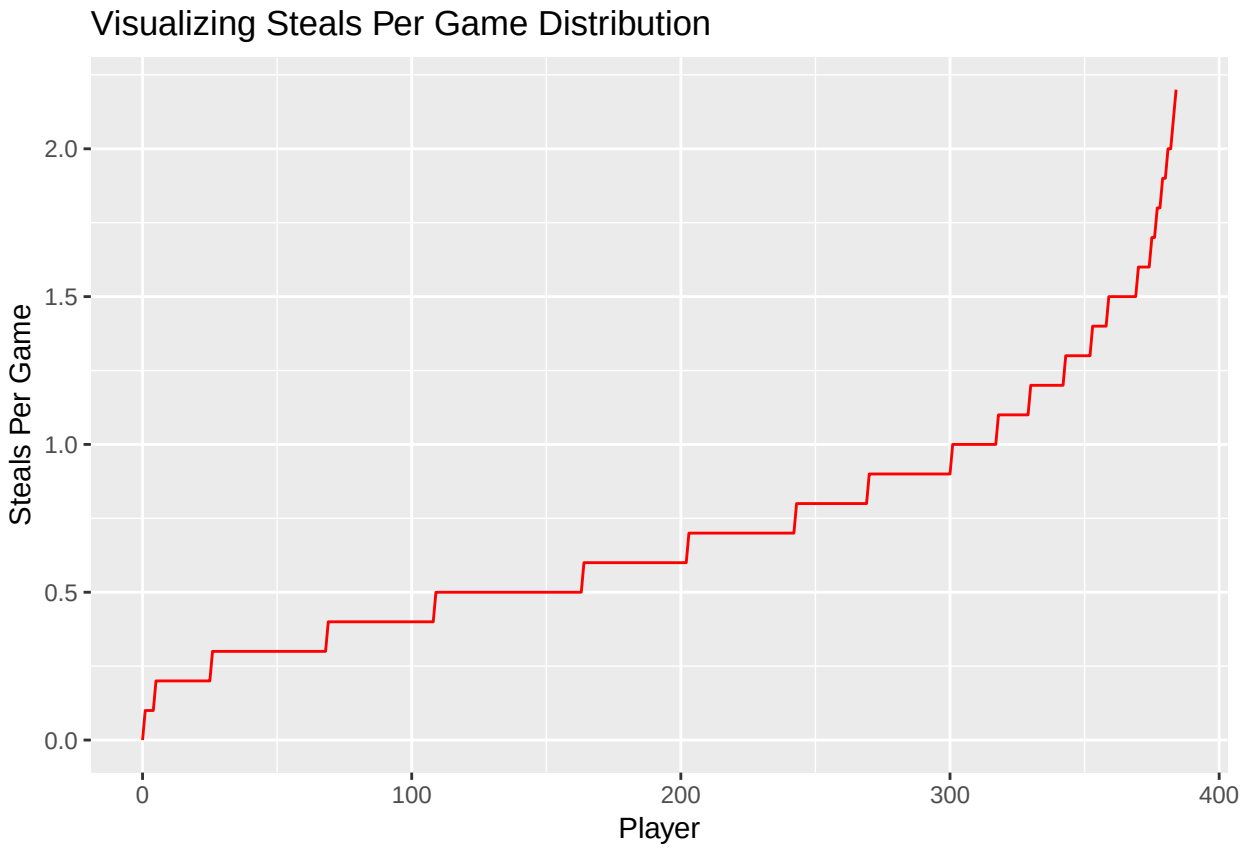
```
## [1] 0.3984
```

```
#Spread
quantile(stats_2019$stlPerGame)
```

```
##   0%  25%  50%  75% 100%
##  0.0  0.4  0.6  0.9  2.2
```

```
x <- seq(0, size, 1)
y <- dnorm(stats_2019$stlPerGame, mean(stats_2019$stlPerGame), sd(stats_2019$stlPerGame))

#Graph
ggplot(stats_2019) +
  geom_line(aes(x = x, y = stlPerGame), color = "red") +
  xlab("Player") +
  ylab("Steals Per Game") +
  ggtitle("Visualizing Steals Per Game Distribution")
```



Steals Per Minute Distribution

```
stats_2019 <- stat_2019[order(stat_2019$stlPerMinute),]
```

```r
#Mean
mean(stats_2019$stlPerMinute)
```

```
## [1] 0.03125
```

```r
#Standard Deviation
sd(stats_2019$stlPerMinute)
```

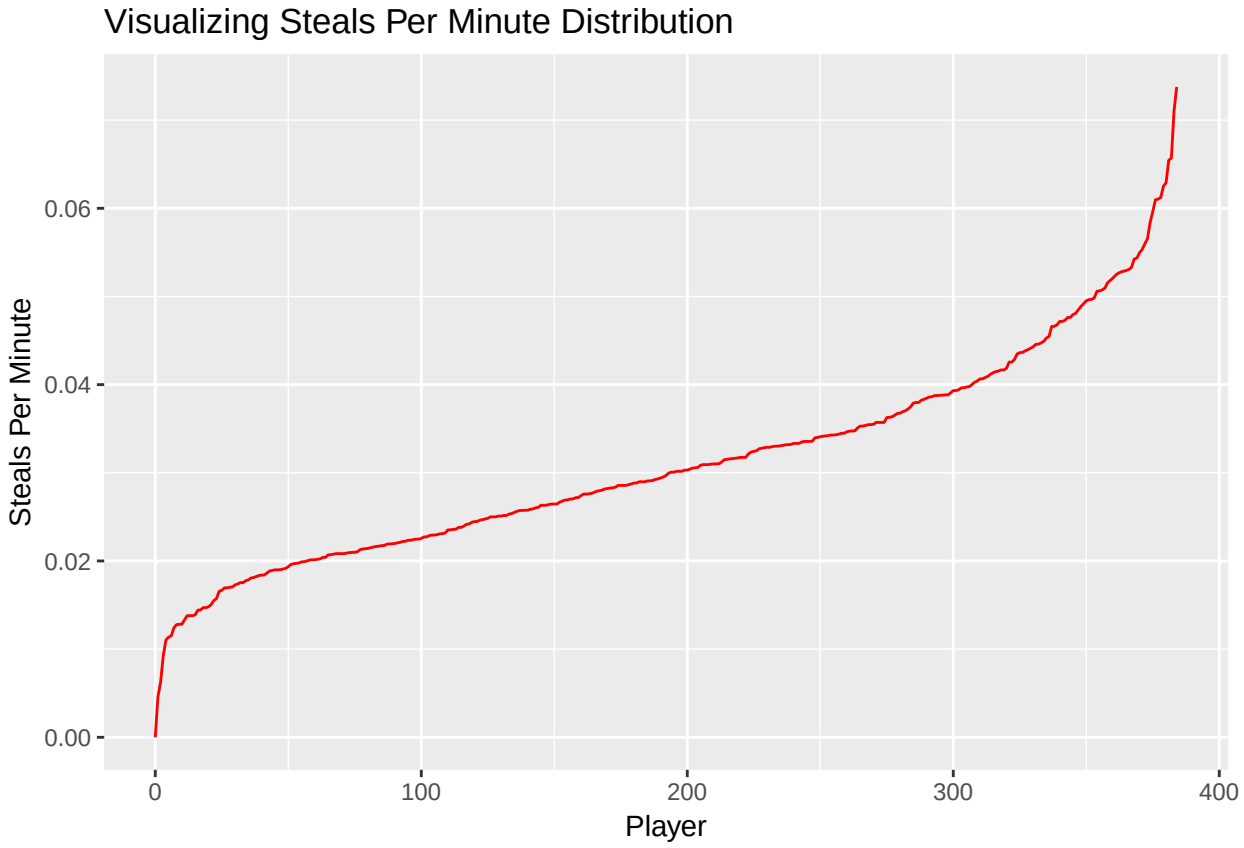
```
## [1] 0.01184
```

```r
#Spread
quantile(stats_2019$stlPerMinute)
```

```
##      0%     25%     50%     75%    100%
## 0.00000 0.02235 0.02970 0.03824 0.07377
```

```r
x <- seq(0, size, 1)
y <- dnorm(stats_2019$stlPerMinute, mean(stats_2019$stlPerMinute), sd(stats_2019$stlPerMinute))

#Graph
ggplot(stats_2019) +
  geom_line(aes(x = x, y = stlPerMinute), color = "red") +
  xlab("Player") +
  ylab("Steals Per Minute") +
  ggtitle("Visualizing Steals Per Minute Distribution")
```



Blocks Per Game Distribution

```
stats_2019 <- stat_2019[order(stat_2019$blkPerGame),]

#Mean
mean(stats_2019$blkPerGame)
```

```
## [1] 0.4629
```

```
#Standard Deviation
sd(stats_2019$blkPerGame)
```

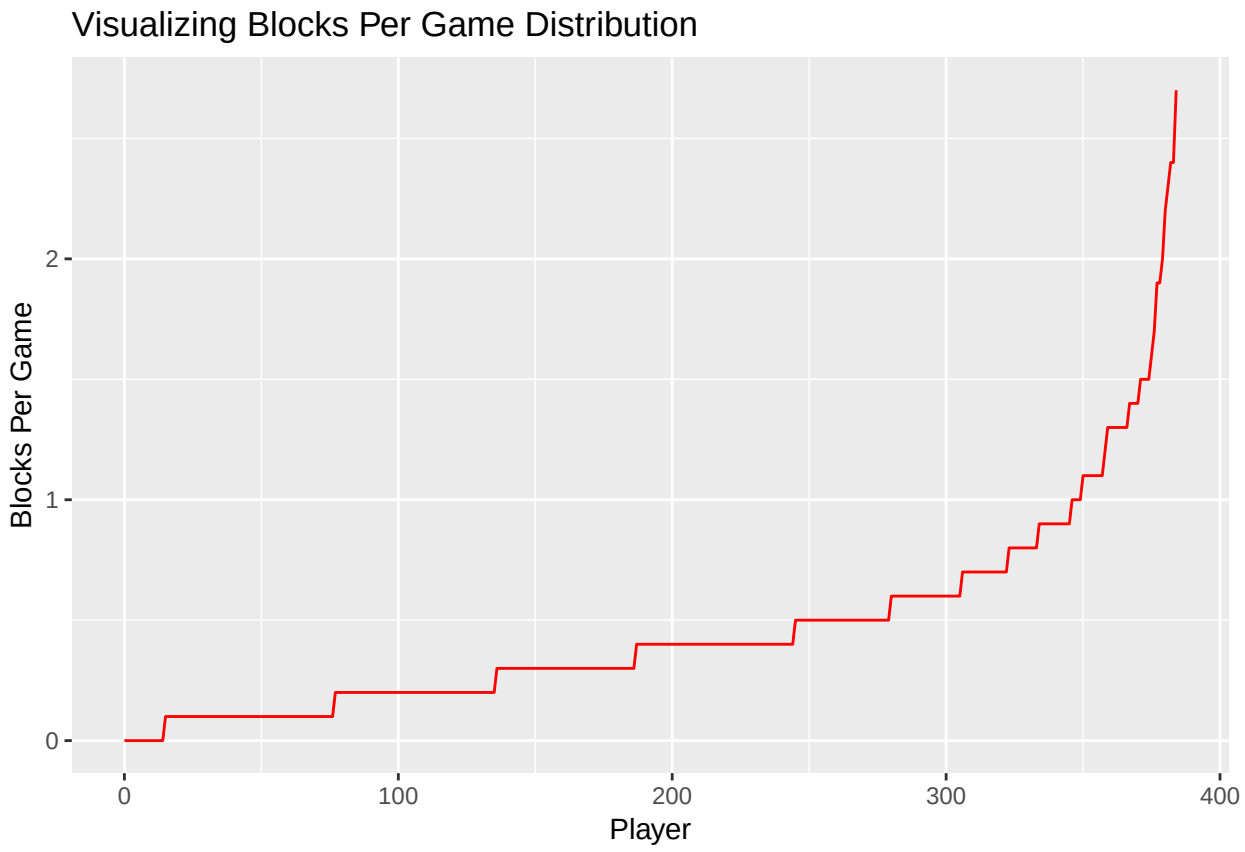
```
## [1] 0.4228
```

```
#Spread
quantile(stats_2019$blkPerGame)
```

```
##   0%  25%  50%  75% 100%
##  0.0  0.2  0.4  0.6  2.7
```

```
x <- seq(0, size, 1)
y <- dnorm(stats_2019$blkPerGame, mean(stats_2019$blkPerGame), sd(stats_2019$blkPerGame))

#Graph
ggplot(stats_2019) +
  geom_line(aes(x = x, y = blkPerGame), color = "red") +
  xlab("Player") +
  ylab("Blocks Per Game") +
  ggtitle("Visualizing Blocks Per Game Distribution")
```



Blocks Per Minute Distribution

```
stats_2019 <- stat_2019[order(stat_2019$blkPerMinute),]

#Mean
mean(stats_2019$blkPerMinute)
```

```
## [1] 0.02155
```

```
#Standard Deviation
sd(stats_2019$blkPerMinute)
```

```
## [1] 0.01915
```

```
#Spread
quantile(stats_2019$blkPerMinute)
```

```
##       0%      25%      50%      75%     100%
## 0.000000 0.009434 0.016129 0.027778 0.147727
```

```
x <- seq(0, size, 1)
y <- dnorm(stats_2019$blkPerMinute, mean(stats_2019$blkPerMinute), sd(stats_2019$blkPerMinute))

#Graph
ggplot(stats_2019) +
  geom_line(aes(x = x, y = blkPerMinute), color = "red") +
  xlab("Player") +
  ylab("Steals Per Minute") +
  ggtitle("Visualizing Blocks Per Minute Distribution")
```