

Assignment 1

Akhil Kornala

2023-09-23

```
knitr::opts_chunk$set(echo = TRUE, comment=NA)
```

1. For this assignment I used the Best Batsman of IPL data, which included both quantitative and qualitative variables. Dataset Source: <https://www.kaggle.com/datasets/yaranathakur/ipl-all-time-best-batsman>

```
library(readr)
Batsman <- read_csv("C:/Users/LENOVO/Desktop/Courses/FML/archive/All-Time-Best-Batsman.csv")
```

```
Rows: 135 Columns: 15
-- Column specification -----
Delimiter: ","
chr  (3): Player, Team, HS
dbl (12): POS, Mat, Inns, NO, Runs, Avg, BF, SR, 100, 50, 4s, 6s

i Use 'spec()' to retrieve the full column specification for this data.
i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
View(Batsman)
```

2. Descriptive Statistics for the variables “Player” and “Runs Scored”

```
summary(Batsman[,c('Player', 'Runs')])
```

Player	Runs
Length:135	Min. : 503.0
Class :character	1st Qu.: 826.5
Mode :character	Median :1405.0
	Mean :1923.7
	3rd Qu.:2655.5
	Max. :7263.0

3. Transforming the “Runs” variable into Square Root

```
square_runs <- (Batsman$Runs)^(1/2)
square_runs
```

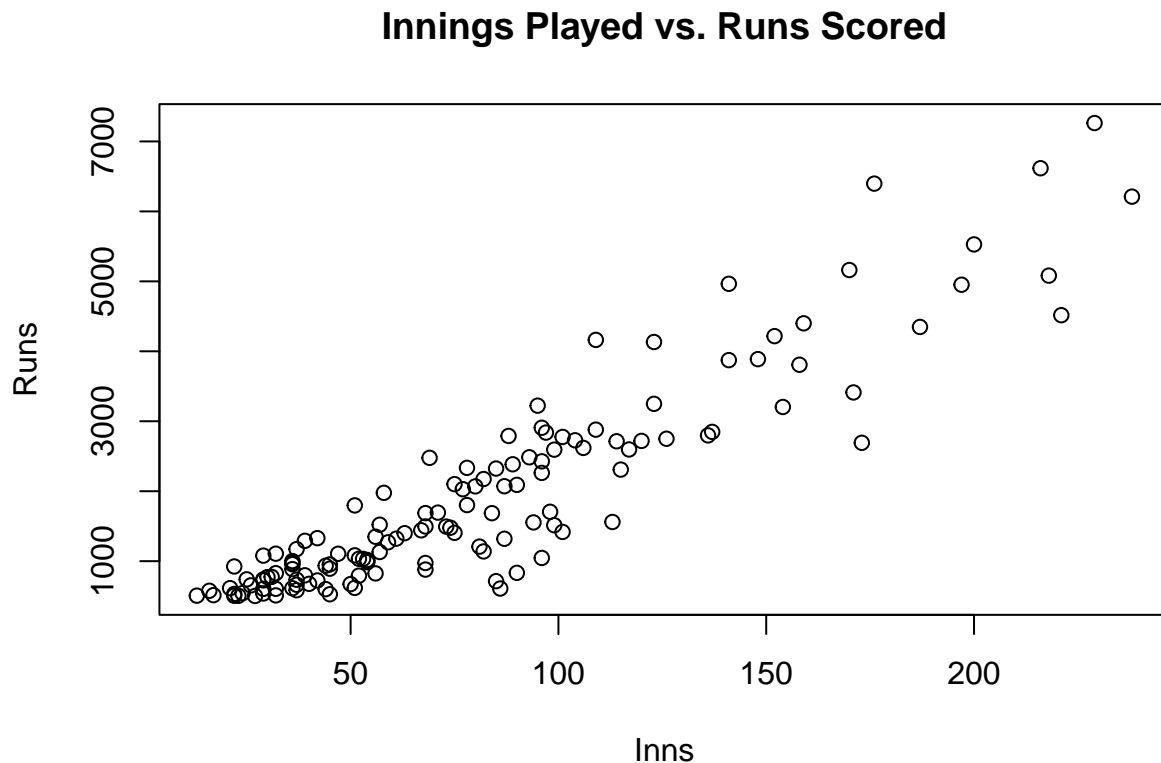
```

[1] 85.22324 81.34494 79.98125 78.80990 74.35052 71.84706 71.28815 70.46276
[9] 70.37045 67.20119 66.33250 65.93937 64.93843 64.52131 64.28841 62.35383
[17] 62.24147 61.70899 58.41233 57.00000 56.77147 56.60389 53.91660 53.66563
[25] 53.36666 53.27288 52.89612 52.82045 52.68776 52.44044 52.23026 52.14403
[33] 52.09607 51.88449 51.17617 50.96077 50.93133 49.84977 49.76947 49.26459
[41] 48.83646 48.31149 48.20788 48.05206 47.56049 46.62617 45.83667 45.72745
[49] 45.50824 45.48626 45.04442 44.46347 42.44997 42.39104 41.30375 41.15823
[57] 41.07311 41.06093 39.49684 39.42081 39.00000 38.91015 38.67816 38.65230
[65] 38.44477 37.96051 37.65634 37.48333 37.41657 36.72874 36.45545 36.35932
[73] 36.33180 35.93049 35.63706 34.75629 34.23449 33.74907 33.63034 33.27161
[81] 33.25658 32.90897 32.84814 32.34192 32.15587 32.12476 31.89044 31.63858
[89] 31.62278 31.38471 31.20897 31.20897 30.91925 30.88689 30.57777 30.39737
[97] 29.88311 29.76575 29.66479 28.86174 28.77499 28.72281 28.24889 28.16026
[105] 27.83882 27.78489 27.20294 27.18455 27.11088 26.92582 26.90725 26.72078
[113] 26.00000 25.92296 25.74879 25.57342 24.85961 24.77902 24.75884 24.67793
[121] 24.59675 24.57641 24.49490 24.18677 24.02082 23.30236 23.19483 23.04344
[129] 22.95648 22.67157 22.60531 22.51666 22.49444 22.47221 22.42766

```

4. Plotting the variables “Innings Played” and “Runs Scored” against each other

```
plot(Batsman$Inns, Batsman$Runs, main= "Innings Played vs. Runs Scored", xlab='Inns', ylab='Runs')
```



5. Scatter Plot for the variables “Innings Playes” and “Runs Scored”

```
library(ggplot2)
ggplot(data=Batsman, aes(x= Inns,y=Runs))+ geom_point(color = "red", size = 2) +
  labs(title = "Scatter Plot of Innings Played vs. Runs Scored")
```

