

Fundamentals of Machine Learning - Assignment 1

Raghu Manusanipalli

2023-09-09

#Importing the data

```
data = read.csv("C:/Users/manus/OneDrive/Desktop/Data.csv")
View(data)
summary(data)
```

```
##      Country      League      Club      Player.Names
## Length:660      Length:660      Length:660      Length:660
## Class :character Class :character Class :character Class :character
## Mode  :character Mode  :character Mode  :character Mode  :character
##
##
##
## Matches_Played Substitution      Mins      Goals
## Min.   : 2.00    Min.   : 0.000    Min.   : 264    Min.   : 2.00
## 1st Qu.:14.00    1st Qu.: 0.000    1st Qu.:1364    1st Qu.: 8.00
## Median :24.00    Median : 2.000    Median :2246    Median :11.00
## Mean   :22.37    Mean   : 3.224    Mean   :2071    Mean   :11.78
## 3rd Qu.:31.00    3rd Qu.: 5.000    3rd Qu.:2822    3rd Qu.:14.00
## Max.   :38.00    Max.   :26.000    Max.   :4177    Max.   :37.00
##      xG      xG.Per.Avg.Match      Shots      OnTarget
## Min.   : 0.710    Min.   :0.0700    Min.   : 5.00    Min.   : 2.00
## 1st Qu.: 6.100    1st Qu.:0.3400    1st Qu.: 37.75    1st Qu.:17.00
## Median : 9.285    Median :0.4350    Median : 62.00    Median :26.00
## Mean   :10.090    Mean   :0.4762    Mean   : 64.18    Mean   :28.37
## 3rd Qu.:13.252    3rd Qu.:0.5700    3rd Qu.: 86.00    3rd Qu.:37.00
## Max.   :32.540    Max.   :1.3500    Max.   :208.00    Max.   :102.00
## Shots.Per.Avg.Match On.Target.Per.Avg.Match      Year
## Min.   :0.800      Min.   :0.240      Min.   :2016
## 1st Qu.:2.335      1st Qu.:0.980      1st Qu.:2017
## Median :2.845      Median :1.250      Median :2019
## Mean   :2.948      Mean   :1.316      Mean   :2018
## 3rd Qu.:3.382      3rd Qu.:1.540      3rd Qu.:2019
## Max.   :7.200      Max.   :3.630      Max.   :2020
```

#discriptive statistics

```
mean(data$Goals)
```

```
## [1] 11.78485
```

```
sd(data$Matches_Played)
```

```
## [1] 9.754658
```

```
str(data$Shots)
```

```
## int [1:660] 48 88 120 117 50 162 69 105 78 64 ...
```

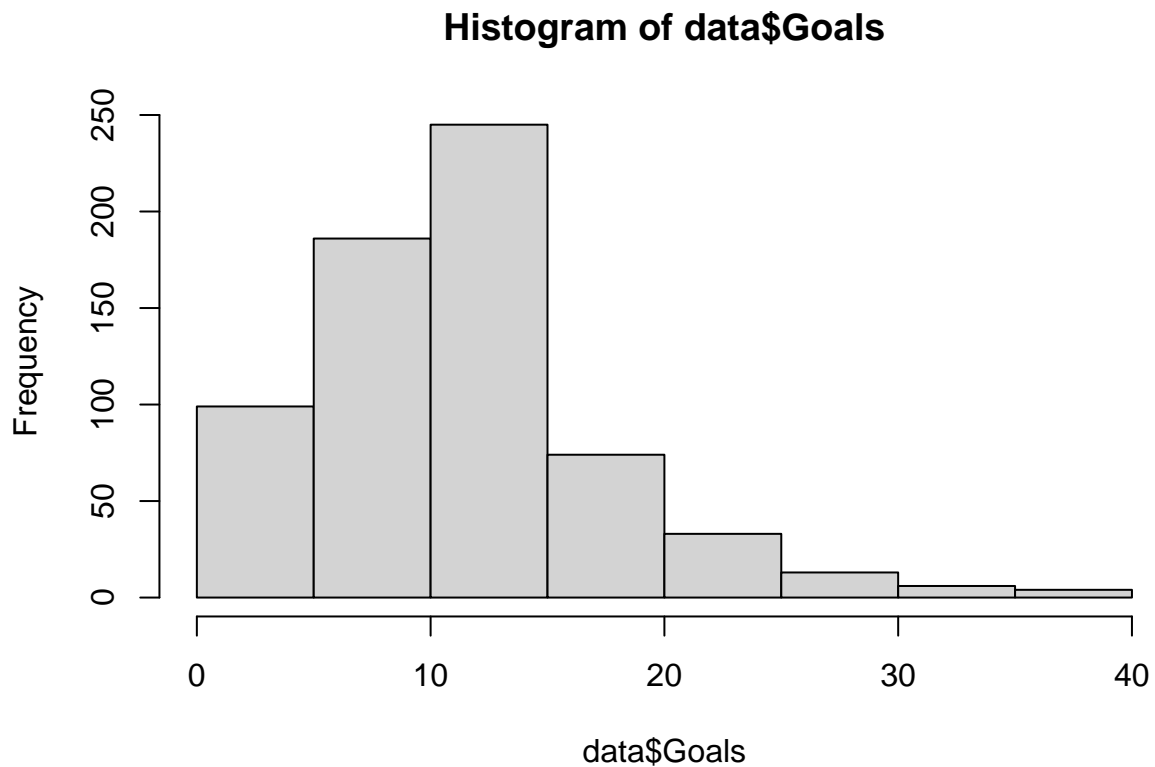
```
#data transformation of variable
```

```
MatchesPlayedVariance = sd(data$Matches_Played)* sd(data$Matches_Played)  
MatchesPlayedVariance
```

```
## [1] 95.15334
```

```
#histogram
```

```
hist(data$Goals)
```



```
#scatter plot
```

```
x <- data$Matches_Played  
y <- data$Goals
```

```
plot(x,y, main = "Matches Played and Goals", xlab = "MatchesPlayed", ylab = "Goals")
```

Matches Played and Goals

