

# ASSIGNMENT 1

Krishna Krupa

2023-09-09

#Downloading a Dataset

<https://www.kaggle.com/datasets/adityaab1407/employee-productivity-and-satisfaction-hr-data>

#Importing the Dataset into R

```
employee = read.csv("hr_dashboard_data.csv", header = TRUE, sep = ",")  
head(employee)
```

```
##           Name Age Gender Projects_Completed Productivity  
## 1 Douglas Lindsey 25 Male 11 57  
## 2 Anthony Roberson 59 Female 19 55  
## 3 Thomas Miller 30 Male 8 87  
## 4 Joshua Lewis 26 Female 1 53  
## 5 Stephanie Bailey 43 Male 14 3  
## 6 Jonathan King 24 Male 5 63  
## Satisfaction_Rate_Per Feedback_Score Department Position  
Joining_Date  
## 1 25 4.7 Marketing Analyst  
Jan-20  
## 2 76 2.8 IT Manager  
Jan-99  
## 3 10 2.4 IT Analyst  
Jan-17  
## 4 4 1.4 Marketing Intern  
Jan-22  
## 5 9 4.5 IT Team Lead  
Jan-05  
## 6 33 4.2 Sales Junior Developer  
Jan-21  
##           Salary  
## 1 63,596.00  
## 2 1,12,540.00  
## 3 66,292.00  
## 4 38,303.00  
## 5 1,01,133.00  
## 6 48,740.00
```

#Descriptive statistics of Quantitative variable

Mean, Median, Min, Max and Standard deviation of Dataset

```
summary(employee$Age)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    22.00  26.00   32.00   34.65  41.00   60.00
```

```
summary(employee$Projects_Completed)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.00   6.00   11.00   11.46  17.00   25.00
```

```
summary(employee$Productivity)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.00  23.00   45.00   46.76  70.00   98.00
```

```
summary(employee$Satisfaction_Rate_Per)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.00  25.75   50.50   49.94  75.25  100.00
```

```
summary(employee$Feedback_Score)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      1.000   1.900   2.800   2.883   3.900   4.900
```

#Descriptive statistics of categorical variable

1.Number of male and female in the Dataset

```
table(employee$Gender)
```

```
##
## Female    Male
##     100     100
```

2.Proportion of each department in the Dataset

```
count= table(employee$Department)
prop.table(count)
```

```
##
## Finance      HR      IT Marketing    Sales
##    0.205    0.160    0.190    0.210    0.235
```

#Transformation of Quantitative Variable

converting the feedback values into percentage

```
result= (employee$Feedback_Score/5)*100
print(result)
```

```
##      [1] 94 56 48 28 90 84 46 56 46 22 24 50 36 62 36 36 74 78 44 88 40 48 76
##      24 94
##      [26] 26 56 78 24 84 98 56 30 26 28 82 54 94 56 26 54 80 86 24 34 60 30 90
```

```

94 46
## [51] 54 64 42 64 36 44 56 68 62 40 38 30 82 28 78 40 42 78 72 80 62 88 30
72 30
## [76] 78 48 74 76 72 38 86 40 62 74 38 70 42 34 84 64 46 66 22 30 52 98 60
26 82
## [101] 82 82 66 76 38 78 86 80 48 68 70 92 28 86 62 48 56 22 38 72 72 60 96
58 46
## [126] 88 40 68 76 86 82 50 24 24 78 96 48 56 78 72 40 74 44 48 64 78 32 22
70 22
## [151] 24 38 72 26 24 98 54 92 38 54 48 34 46 98 80 50 52 42 62 78 74 58 24
94 52
## [176] 52 20 32 90 62 78 32 46 88 86 40 22 84 34 84 74 34 26 94 68 70 56 32
68 40

```

#Plotting the variables

Bar plot for number of departments in the Dataset

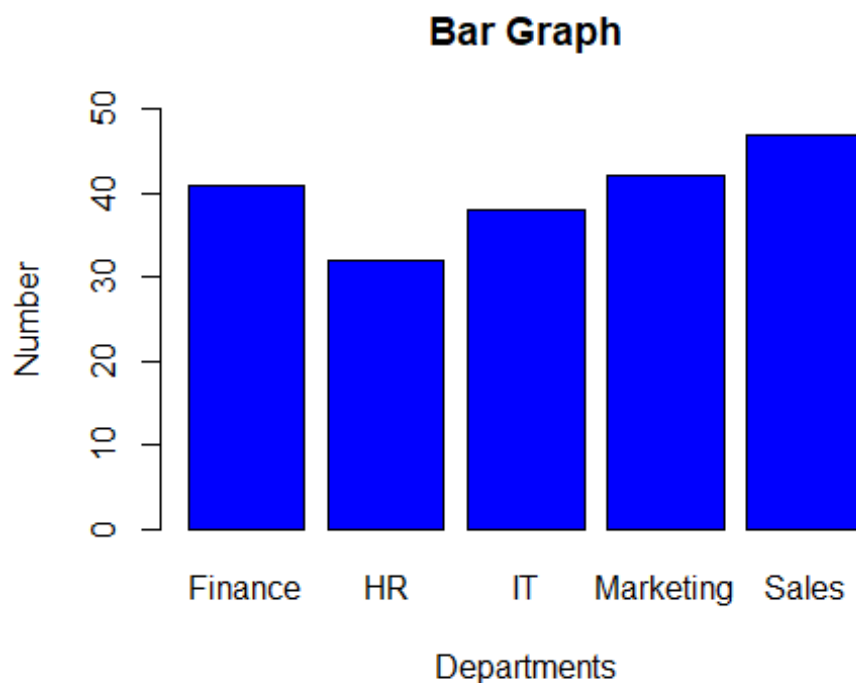
```

x=table(employee$Department)
print(x)

##
## Finance      HR      IT Marketing      Sales
##      41      32      38      42      47

barplot.default(x,names.arg = c('Finance','HR','IT', 'Marketing', 'Sales' ),
                xlab= 'Departments', ylab = 'Number', ylim = c(0,50),main =
'Bar Graph',col = 'blue' )

```



Scatter plot for age and the number of projects completed in the Dataset

```
Age= employee$Age  
projects= employee$Projects_Completed  
plot(x = Age, y = projects, xlab='Age', ylab= 'No.of projects', xlim =  
c(0,100), ylim = c(0,50), main = 'scatterplot')
```

