

**COMSATS UNIVERSITY ISLAMABAD, ABBOTTABAD**

Data science

Class Task# 01

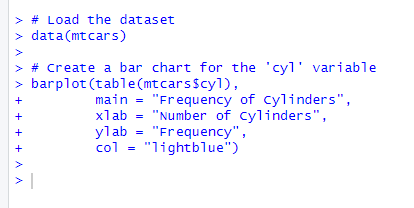
***Submitted by:***

Laiba binte tahir FA21-BSE-019

***Submitted to:***

Dr. Ghulam Mujtaba

**Task-01: Bar Chart for Categorical Variable**



A graph of a number of cylinders

Description automatically generated

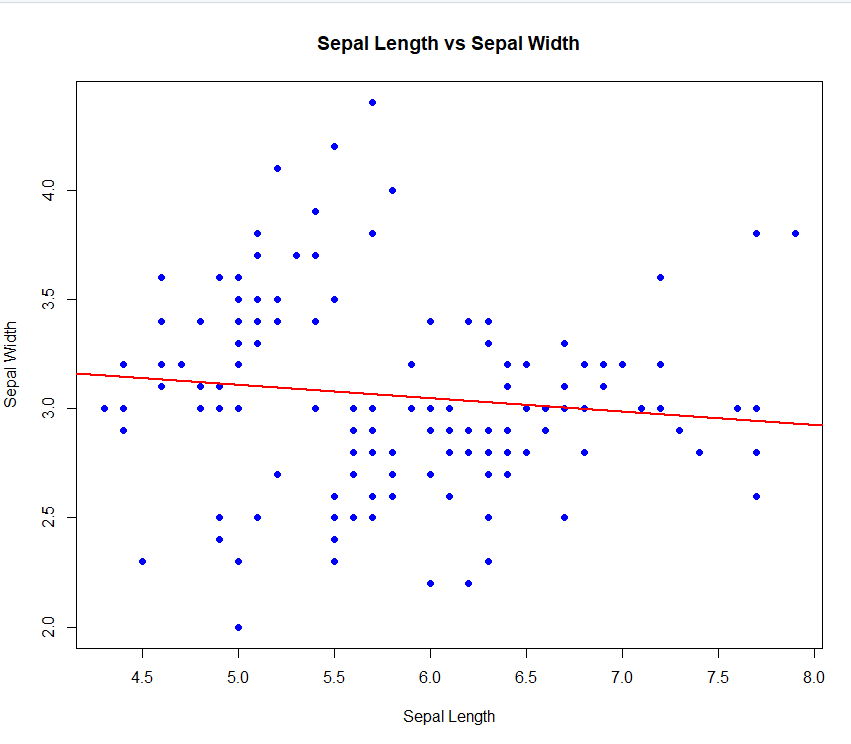
**Task-02: Histogram for Continuous Variable**A white background with blue text

Description automatically generated

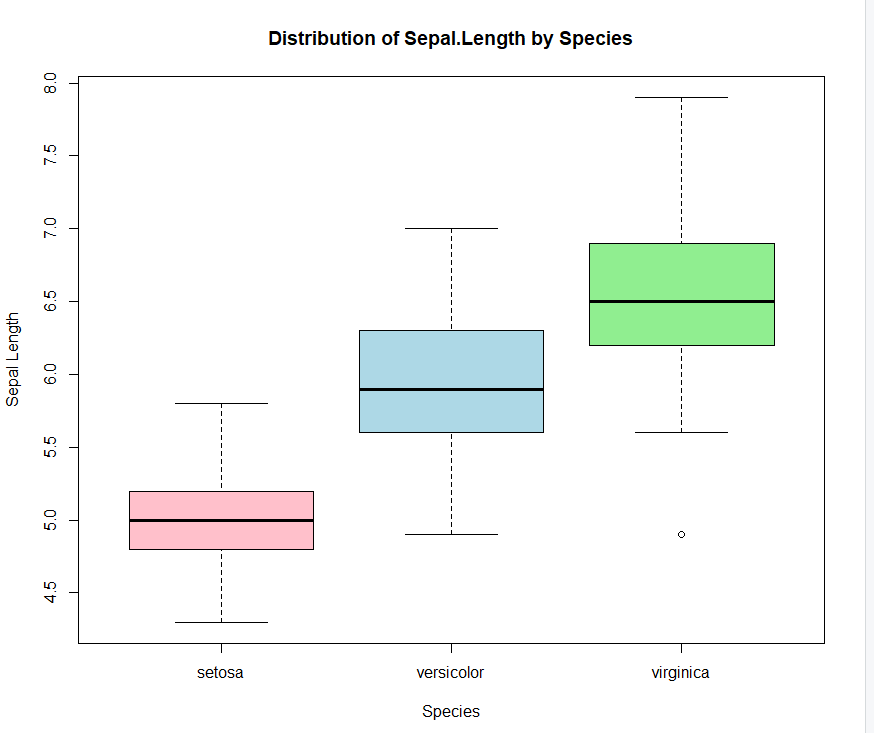
A graph of a number of green bars

Description automatically generated

**Task 2: Scatter Plot with Regression Line (Iris Dataset)**

****

**Task 3: Boxplot of Sepal.Length by Species**



boxplot(Sepal.Length ~ Species, data = iris,

main = "Distribution of Sepal.Length by Species",

xlab = "Species",

ylab = "Sepal Length",

col = c("pink", "lightblue", "lightgreen"))

**Task 4: Line Chart for mpg Dataset**

data(mtcars)

plot(mtcars$hp, mtcars$disp,

type = "l",

main = "Displacement vs Horsepower",

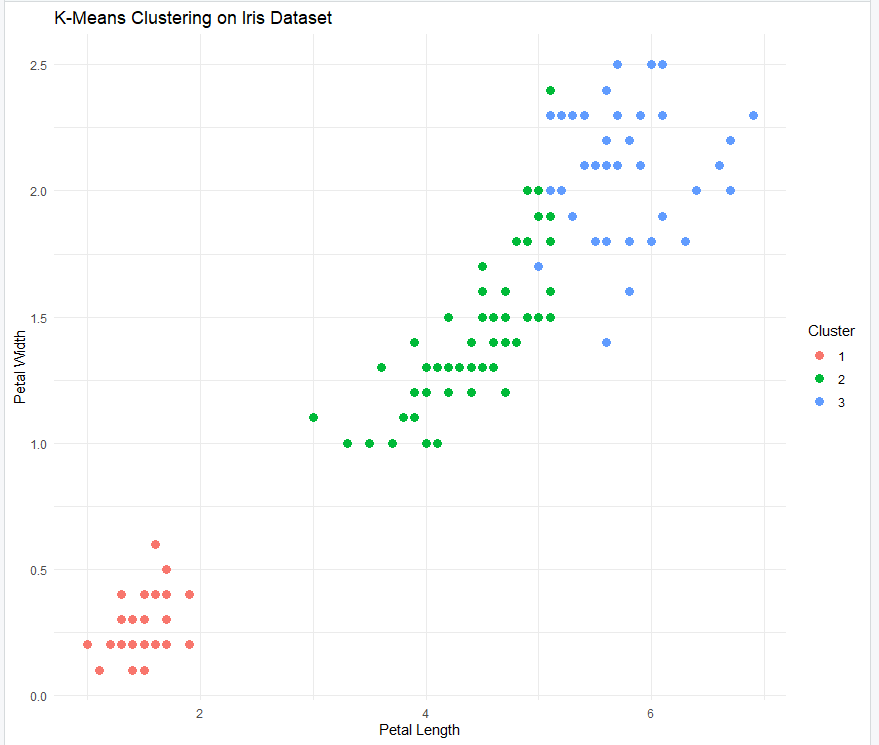
xlab = "Horsepower",

ylab = "Displacement",

col = "purple",

lwd = 2)

**Task: K Mean clustering**

****

iris\_features <- iris[, 1:4]

set.seed(123) # For reproducibility

kmeans\_result <- kmeans(iris\_features, centers = 3)

iris$Cluster <- as.factor(kmeans\_result$cluster)

library(ggplot2)

ggplot(iris, aes(x = Petal.Length, y = Petal.Width, color = Cluster)) +

geom\_point(size = 3) +

labs(title = "K-Means Clustering on Iris Dataset",

x = "Petal Length",

y = "Petal Width") +

theme\_minimal()