DAY-2 LAB EXPERIMENTS

R PROGRAMMING

EXPERIMENT 1:

The following values are the number of pencils available in the different boxes. Create a vector and find out the mean, median and mode values of set of pencils in the given data.

Box1	Box2	Box3	Box4	Box5	Box6	Box7	Box8	Box9	Box 10
9	25	23	12	11	6	7	8	9	10

AIM:

To find the mean,median,mode of the given data set

SOFTWARE REQUIRES:

R SOFTWARE

PROGRAM:

```
pencil_counts <- c(9, 25, 23, 12, 11, 6, 7, 8, 9, 10)

mean_value <- mean(pencil_counts)

median_value <- median(pencil_counts)

mode_value <- names(table(pencil_counts))[table(pencil_counts) == max(table(pencil_counts))]

# Print results

cat("Mean:", mean_value, "\n")

cat("Median:", median_value, "\n")

cat("Mode:", mode_value, "\n")
```

OUTPUT:

```
> # Create a vector with the number of pencils in each box
> pencil_counts <- c(9, 25, 23, 12, 11, 6, 7, 8, 9, 10)
>
> # Calculate mean
> mean_value <- mean(pencil_counts)
>
> # Calculate median
> median_value <- median(pencil_counts)
>
> # Calculate mode (there can be multiple modes)
> mode_value <- names(table(pencil_counts))[table(pencil_counts) == max(table(pencil_counts))]
>
> # Print results
> cat("Mean:", mean_value, "\n")
Mean: 12
> cat("Median:", median_value, "\n")
Median: 9.5
> cat("Mode:", mode_value, "\n")
Mode: 9
>
```

EXPERIMENT 2:

the following table would be plotted as (x,y) points, with the first column being the x values as number of mobile phones sold and the second column being the y values as money. To use the scatter plot for how many mobile phones sold.

x :4 1 5 7 10 2 50 25 90 36

y :12 5 13 19 31 7 153 72 275 110

AIM:

To draw the scatter plot for the mobiles phones sold in the given data

SOFTWARE REQUIRES:

R SOFTWARE

PROGRAM:

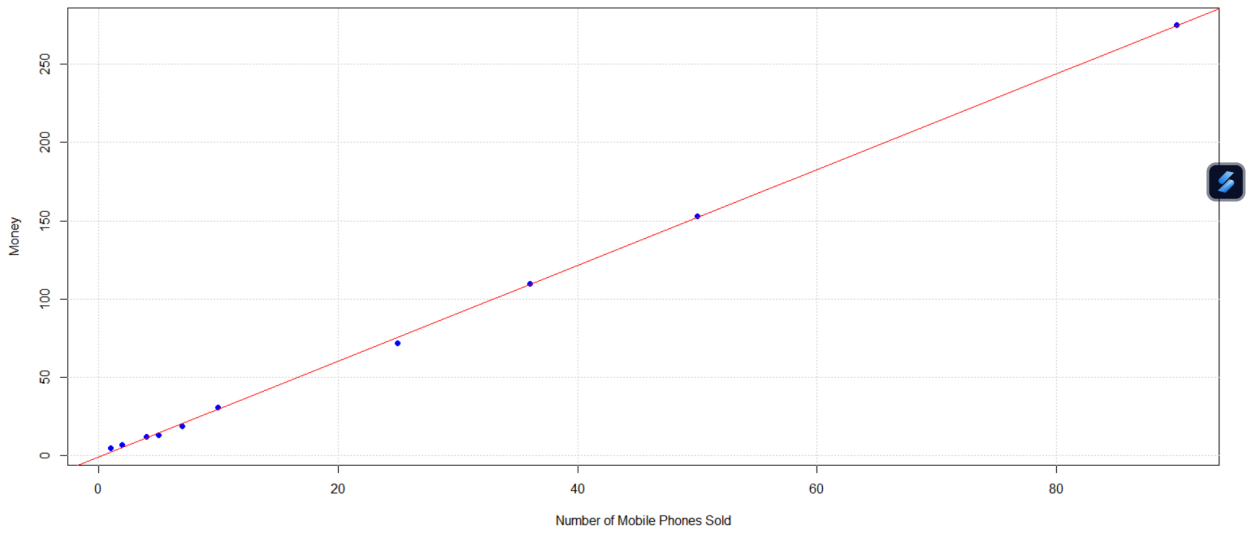
```
x <- c(4, 1, 5, 7, 10, 2, 50, 25, 90, 36)

y <- c(12, 5, 13, 19, 31, 7, 153, 72, 275, 110)

plot(x, y, main = "Scatter Plot of Mobile Phones Sold",

xlab = "Number of Mobile Phones Sold", ylab = "Money",
```

col = "blue", pch = 19)

grid()

abline(lm(y ~ x), col = "red")

OUTPUT:



EXPERIMENT 3:

. Implement of the R script using marks scored by a student in his model exam has been sorted as follows: 55, 60, 71, 63, 55, 65, 50, 55,58,59,61,63,65,67,71,72,75. Partition them into three bins by each of the following methods. Plot the data points using histogram.

(a) equal-frequency (equi-depth) partitioning (b) equal-width partitioning

AIM:

To draw the histogram graph for the given points

SOFTWARE REQUIRES:

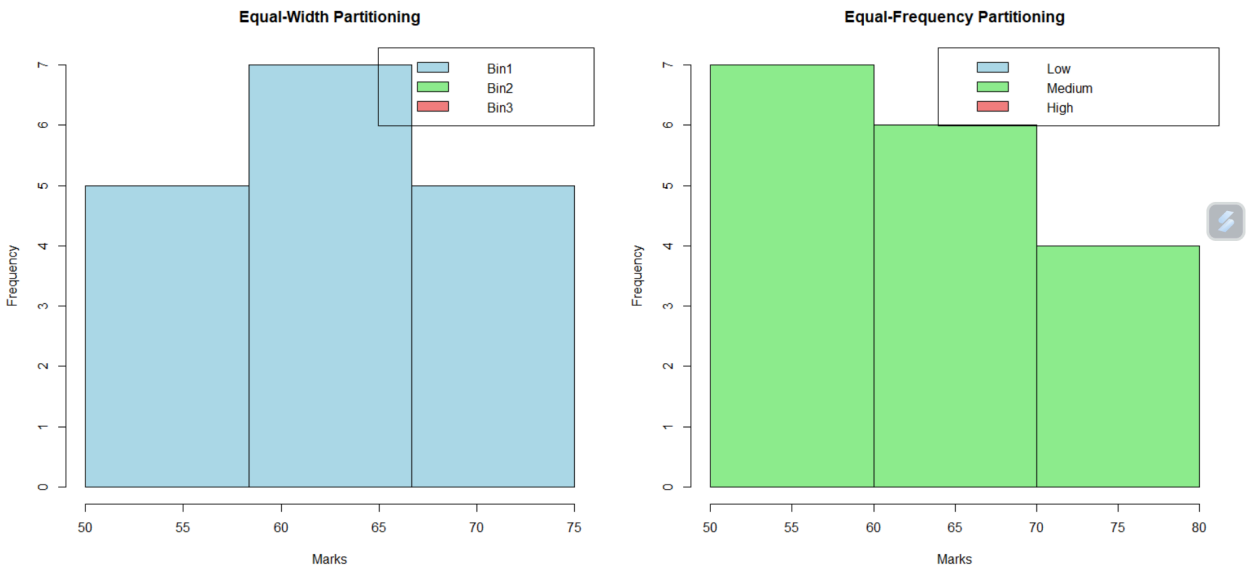
R SOFTWARE

PROGRAM:

marks <- c(55, 60, 71, 63, 55, 65, 50, 55, 58, 59, 61, 63, 65, 67, 71, 72, 75)

```
sorted_marks <- sort(marks)

# Equal-frequency partitioning (equi-depth)

bins_ef <- cut(sorted_marks, breaks = 3, labels = c("Low", "Medium", "High"), include.lowest = TRUE)

# Equal-width partitioning

min_mark <- min(sorted_marks)

max_mark <- max(sorted_marks)

width <- (max_mark - min_mark) / 3

breaks_ew <- seq(min_mark, max_mark, by = width)

bins_ew <- cut(sorted_marks, breaks = breaks_ew, labels = c("Bin1", "Bin2", "Bin3"), include.lowest = TRUE)

# Plot histograms

par(mfrow = c(1, 2))  # Set up the layout for subplots

hist(sorted_marks, breaks = breaks_ew, main = "Equal-Width Partitioning", xlab = "Marks", ylab = "Frequency", col = "lightblue")

legend("topright", legend = c("Bin1", "Bin2", "Bin3"), fill = c("lightblue", "lightgreen", "lightcoral"))

hist(sorted_marks, breaks = 3, main = "Equal-Frequency Partitioning", xlab = "Marks", ylab = "Frequency", col = "lightgreen")

legend("topright", legend = c("Low", "Medium", "High"), fill = c("lightblue", "lightgreen", "lightcoral"))
```

OUTPUT:



EXPERIMENT 4:

Suppose that the speed car is mentioned in different driving style.

Regular 78.3 81.8 82 74.2 83.4 84.5 82.9 77.5 80.9 70.6 Speed

Calculate the Inter quantile and standard deviation of the given data.

AIM:

T o find the interquartile and standard deviation of the given data

SOFTWARE REQUIRES:

R SOFTWARE

PROGRAM:

```
speed <- c(78.3, 81.8, 82, 74.2, 83.4, 84.5, 82.9, 77.5, 80.9, 70.6)

iqr <- IQR(speed)

std_dev <- sd(speed)

cat("Interquartile Range (IQR):", iqr, "\n")

cat("Standard Deviation:", std_dev, "\n")
```

OUTPUT:

```
> # Speed data for regular driving style
> speed <- c(78.3, 81.8, 82, 74.2, 83.4, 84.5, 82.9, 77.5, 80.9, 70.6)
>
> # Calculate interquartile range (IQR)
> iqr <- IQR(speed)
>
> # Calculate standard deviation
> std_dev <- sd(speed)
>
> # Print results
> cat("Interquartile Range (IQR):", iqr, "\n")
Interquartile Range (IQR): 4.975
> cat("Standard Deviation:", std_dev, "\n")
Standard Deviation: 4.445835
>
```

EXPERIMENT 5:

Suppose that the data for analysis includes the attribute age. The age values for the data tuples are (in increasing order) 13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70.

Can you find (roughly) the first quartile (Q1) and the third quartile (Q3) of the data?

AIM:

To find the quartiles of the given data

SOFTWARE REQUIRES:

R SOFTWARE

PROGRAM:

ages <- c(13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70)

q1 <- quantile(ages, probs = 0.25, type = 1)

q3 <- quantile(ages, probs = 0.75, type = 1)

cat("First Quartile (Q1):", q1, "\n")

cat("Third Quartile (Q3):", q3, "\n")

OUTPUT:

```
> # Age data
> ages <- c(13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70)
>
> # Calculate first quartile (Q1)
> q1 <- quantile(ages, probs = 0.25, type = 1)
>
> # Calculate third quartile (Q3)
> q3 <- quantile(ages, probs = 0.75, type = 1)
>
> # Print results
> cat("First Quartile (Q1):", q1, "\n")
First Quartile (Q1): 20
> cat("Third Quartile (Q3):", q3, "\n")
Third Quartile (Q3): 35
>
```