Ex1: STATISTICS USING R

PROBLEM 1:

Given the following data for Age, Diastolic pressure, and Systolic pressure:

- Age: 35, 65, 49, 30, 20, 40, 90, 54, 78, 45
- Systolic pressure: 122, 120, 120, 115, 130, 131, 118, 122, 120, 115
- Diastolic pressure: 83, 79, 78, 72, 90, 90, 82, 80, 82, 75

1. Create vectors for Age, Diastolic pressure, and Systolic pressure.
2. Calculate the following statistics for each variable:
    - Number of samples
    - Minimum value
    - Maximum value
    - Median
    - Mean
    - Variance
    - Standard deviation
3. Present the results in a table format.

PROGRAM:

```
age <- c(35, 65, 49, 30, 20, 40, 90, 54, 78, 45)
systolic_pressure <- c(122, 120, 120, 115, 130, 131, 118, 122, 120, 115)
diastolic_pressure <- c(83, 79, 78, 72, 90, 90, 82, 80, 82, 75)
calculate_stats <- function(data) {
  n <- length(data)
  min_val <- min(data)
  max_val <- max(data)
  median_val <- median(data)
  mean_val <- mean(data)
  variance_val <- var(data)
  sd_val <- sd(data)

  return(c(n, min_val, max_val, median_val, mean_val, variance_val, sd_val))
}
age_stats <- calculate_stats(age)
systolic_pressure_stats <- calculate_stats(systolic_pressure)
diastolic_pressure_stats <- calculate_stats(diastolic_pressure)
stats_table <- data.frame(
```

```
  Statistic = c("Number of samples", "Minimum value", "Maximum value", "Median", "Mean",
"Variance", "Standard deviation"),
  Age = age_stats,
  Systolic_Pressure = systolic_pressure_stats,
  Diastolic_Pressure = diastolic_pressure_stats
)

print(stats_table)
```

OUTPUT:

```
           Statistic       Age Systolic_Pressure Diastolic_Pressure
1  Number of samples  10.00000         10.000000          10.000000
2      Minimum value  20.00000        115.000000          72.000000
3      Maximum value  90.00000        131.000000          90.000000
4             Median  47.00000        120.000000          81.000000
5               Mean  50.60000        121.300000          81.100000
6           Variance 474.71111         29.566667          33.211111
7 Standard deviation  21.78787          5.437524           5.762908
```

PROBLEM 2:

1. Read the Iris dataset.
2. View the dataset and display the first 6 and last 6 rows.
3. Obtain the statistical summary of the dataset.
4. Calculate the mean petal width of the Iris versicolor species.
5. Create a frequency distribution of petal width using a histogram.
6. Calculate the mean, median, variance, and standard deviation of sepal length for each species (setosa, versicolor, virginica).
7. Display the calculated statistics in a table format.

PROGRAM:

```
demo<- read.csv('Iris.csv')

head(demo)

tail(demo)

summary(demo)

mean(demo $Petal.Width)

hist(iris$Petal.Width,
    main = "Histogram of Petal Width",
    xlab = "Petal Width",
```

```
    col = "lightblue",
    border = "black")
versi_petal<- mean(versi_data$Petal.Width)
versi_petal<- median(versi_data$Petal.Width)
versi_petal<- sd(versi_data$Petal.Width)
versi_petal<- var(versi_data$Petal.Width)

setosa_petal<-mean(setosa_data$Petal.Width)
setosa_petal<-median(setosa_data$Petal.Width)
setosa_petal<-sd(setosa_data$Petal.Width)
setosa_petal<-var(setosa_data$Petal.Width)

vir_petal<-mean(vir_data$Petal.Width)
vir_petal<-median(vir_data$Petal.Width)
vir_petal<-sd(vir_data$Petal.Width)
vir_petal<-var(vir_data$Petal.Width)

results <- data.frame(
  Statistic = c(
    "mean", "median",
    "Variance", "Standard deviation"
  ),
  versi= c( versi_petal

  ),
  setosa = c(
  setosa_petal
  ),
  virginica  = c(vir_petal)
)
print(results)
```

```
           Statistic     versi     setosa virginica
1               mean 0.1977527 0.01110612 0.2746501
2             median 0.1977527 0.01110612 0.2746501
3           Variance 0.1977527 0.01110612 0.2746501
4 Standard deviation 0.1977527 0.01110612 0.2746501
>
```