

EXERCISE: 3

AIM:

To perform Univariate, Bivariate and Multivariate analysis on Iris Dataset

PROCEDURE:

Step 1: Start R studio

Step 2: Install Pre-requisite packages

Step 3: Load required libraries

Step 4: Load Iris Dataset – preloaded in R

Step 5: Compute Mean, Median, Mode, Standard Deviation, Variance, Quantile

Step 6: Visualize using Histogram and Barplot

Step 7: Normalize numeric data in the dataset using Z-score.

Step 8: View Z-score

CODE:

```
# install pre-requisite packages #
install.packages("stats")
install.packages("dplyr")
# Load required libraries #
library(stats)
library(dplyr)
# Iris Data is preloaded in R #
# Loading data on to my data object #
mydata = iris
str(mydata)
```

UNIVARIATE ANALYSIS

1. Data Summarization

```
SL = mydata$Sepal.Length
# Central Tendencies #
# Mean #
mean(SL)
# Median #
```

```
median(SL)
# Mode #
Mode(SL)
# Measures of variability #
# Calculate Standard Deviation #
sd(SL)
# Calculate Variance - Square of SD #
var(SL)
# Calculate IQR #
quantile(SL)
```

BIVARIATE ANALYSIS

```
# 2. Data Visualization #
str(mydata)
Species = mydata$Species
# Numeric Data - histogram #
hist(SL)
# Categorical Data - Barplot #
table(Species)
barplot(table(Species))
```

MULTIVARIATE ANALYSIS

```
# 3. Data Normalization #
str(mydata)
# Select all numeric data in the data set #
mydata_numeric = select(mydata,c(1,2,3,4))
# Normalize dataset (Z-score) - All variables on the same scale #
zscore_mydata = scale(mydata_numeric)
# Check all the variables converted to Z Scores #
View(zscore_mydata)
```

RESULT

Thus, Univariate, Bivariate and Multivariate analysis has been carried out successfully.