Finolex Academy of Management and Technology, Ratnagiri

**Department of Information Technology**

| | | | |
|---|---|---|---|
| **Subject:** | **R Programming Lab. (ITL804)** | | |
| **Class:** | **BE IT / Semester – VIII (Rev-2016) / Academic year: 2019-20** | | |
| **Name of Student:** | **Kazi Jawwad A Rahim** | | |
| **Roll No:** | **28** | **Date of performance (DOP) :** | |
| **Assignment/Experiment No:** | **04** | **Date of checking (DOC) :** | |
| **Title:** Exploratory data analysis such as- Range, summary, mean, variance, median, standard deviation, histogram, boxplot, scatterplot | | | |
| **Marks:** | | **Teacher's Signature:** | |

**1. Aim**: To understand the exploratory data analysis and the methods required to do it in R.

**2. Prerequisites**:

1. Basics of R programming, various data structures, functions etc.

**3. Hardware Requirements**:

1. PC with minimum 2GB RAM

**4. Software Requirements:**

1. Windows / Linux OS.
2. R version 3.6 or higher

**5. Learning Objectives:**

1. To understand decision and loop control instructions.
2. To understand function definition and calling to it.

**6. Learning Objectives Applicable: LO 3. LO 4**

**7. Program Outcomes Applicable: PO 2, PO 3**

**8. Program Education Objectives Applicable: PEO 2, PEO 3**

**Range:**

```
a=c(1,5,7,8,9,6,4,9,5,8,9,6)
range(a)
```

**OUTPUT:**

```
[1] 1 9
```

**Summary:**

```
a=c(1,5,7,8,9,6,4,9,5,8,9,6)
summary(a)
```

**OUTPUT:**

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1.000   5.000   6.500   6.417   8.250   9.000
```

**Mean:**

```
a=c(1,5,7,8,9,6,4,9,5,8,9,6)
mean(a)
```

**OUTPUT:**

```
[1] 6.416667
```

**Mode:**

```
a=c(1,5,7,8,9,6,4,9,5,8,9,6)
table(a)
```

**OUTPUT:**

```
a
1 4 5 6 7 8 9
1 1 2 2 1 2 3
=> Mode=9
```

**Median:**

```
a=c(1,5,7,8,9,6,4,9,5,8,9,6)
median(a)
```

**OUTPUT:**

```
[1] 6.5
```

**Variance:**

```
a=c(1,5,7,8,9,6,4,9,5,8,9,6)
var(a)
```

**OUTPUT:**

```
[1] 5.901515
```

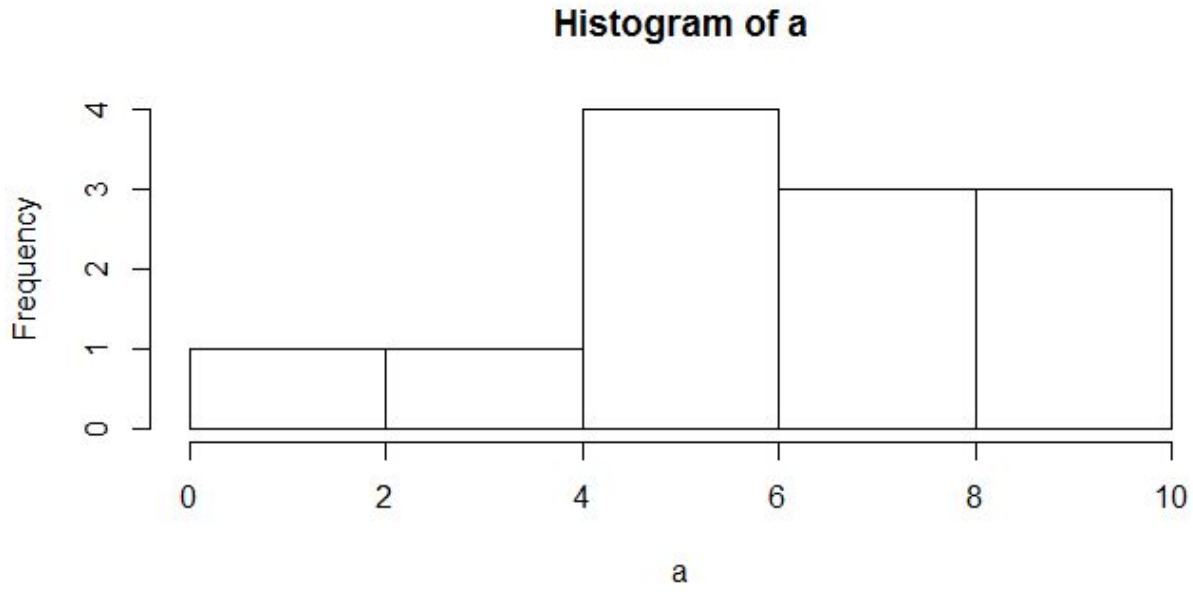
**Standard Deviation:**

```
a=c(1,5,7,8,9,6,4,9,5,8,9,6)
sqrt(var(a))
```

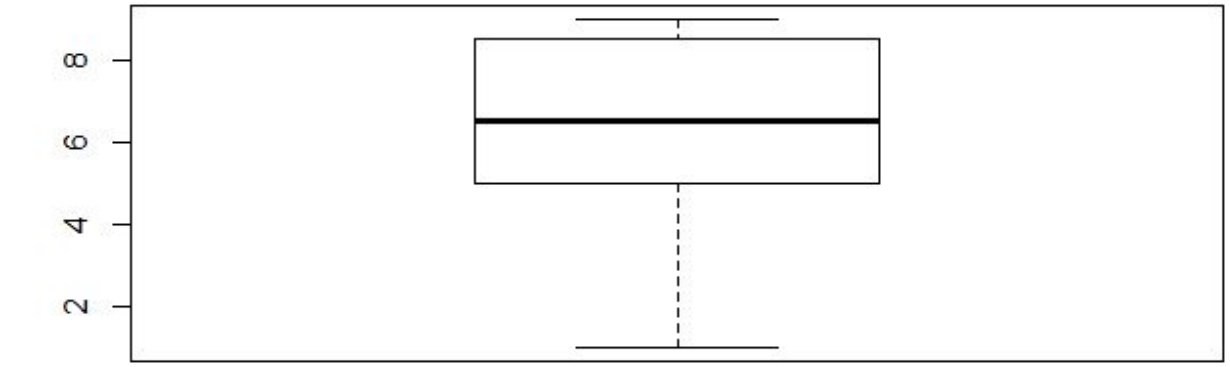
**OUTPUT:**

```
[1] 2.429303
```

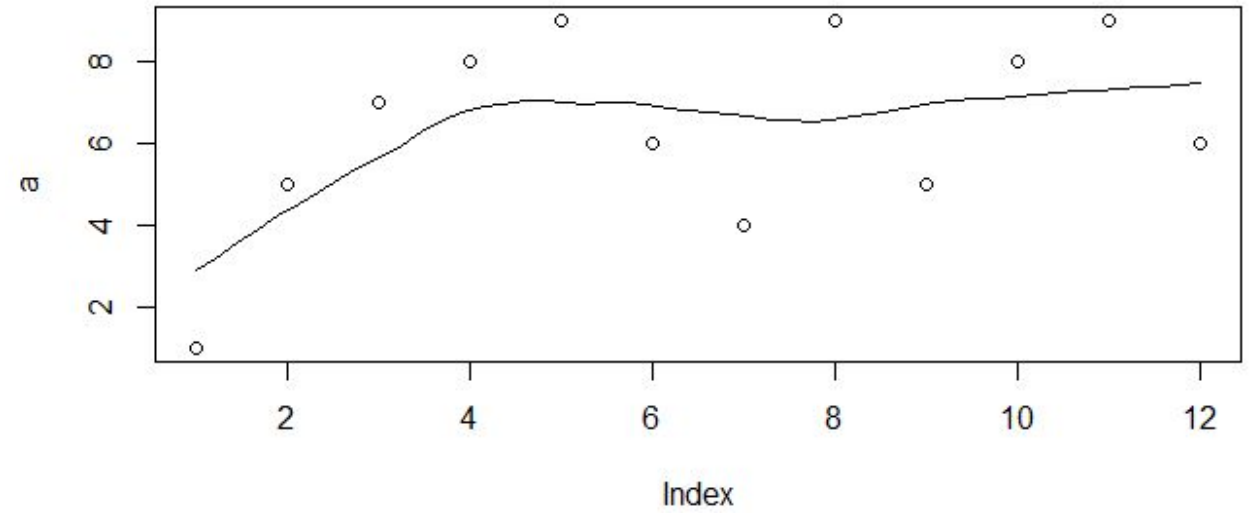
**Histogram:**




**Boxplot:**




**Scatterplot:**




## Learning Outcomes Achieved:

1. We understood decision and loop control instructions.
2. We understood the function definition and it's calling.

## Conclusion:

We have successfully demonstrated the data analysis such as- Range, summary, mean, variance, median, standard deviation, histogram, boxplot, scatterplot.

## 13. Experiment/Assignment Evaluation

| **Experiment/Assignment Evaluation:** | | | | |
|---|---|---|---|---|
| **Sr. No.** | **Parameters** | | **Marks obtained** | **Out of** |
| **1** | Technical Understanding (Assessment may be done based on Q & A **or** any other relevant method.) Teacher should mention the other method used - | | | 6 |
| **2** | Neatness/presentation | | | 2 |
| **3** | Punctuality | | | 2 |
| **Date of performance (DOP)** | | **Total marks obtained** | | **10** |
| **Date of checking (DOC)** | | **Signature of teacher** | | |

# References:


1. URL: https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf ( Online Resources)
2. R Cookbook Paperback – 2011 by Teetor Paul O Reilly Publications
3. Beginning R: The Statistical Programming Language by Dr. Mark Gardener, Wiley Publications
4. R Programming For Dummies by Joris Meys Andrie de Vries, Wiley Publications


# Viva Questions

1. What is exploratory data analysis ?
2. What is summary of the data ?
3. What is importance of median of the data collection ?
4. What is histogram? Why is it important in data?
5. What information does the box plot provides?
6. List various R library functions used in exploratory data analysis.