

Experiment 10: Illustrate the descriptive statistics using summary() in R

Aim: To generate the statistical measurements for a given dataset using summary() method in R.

A descriptive statistic is a summary statistic that quantitatively describes or summarizes features from a collection of information. They help describe, show or summarize data in a meaningful way such that, for example, patterns might emerge from the data. Descriptive statistics do not, however, allow us to make conclusions beyond the data we have analyzed or reach conclusions regarding any hypotheses we might have made, they are simply a way to describe our data.

Descriptive statistics are very important because if we simply presented our raw data it would be hard to visualize what the data was showing, especially if there was a lot of it. Descriptive statistics therefore enable us to present the data in a more meaningful way, which allows simpler interpretation of the data.

Experiment 10: Illustrate the descriptive statistics using summary() in R

Aim: To generate the statistical measurements for a given dataset using summary() method in R.

Typically, there are two general types of statistic that are used to describe data:

1. Measures of central tendency: these are ways of describing the central position of a frequency distribution for a group of data. **The most common measures of central tendency are the arithmetic mean, the median, and the mode.**

2. Measures of spread: these are ways of summarizing a group of data by describing how spread out the scores are. **Measures of spread include the range, quartiles and the interquartile range, variance and standard deviation.**

Experiment 10: Illustrate the descriptive statistics using summary() in R

Aim: To generate the statistical measurements for a given dataset using summary() method in R.

Inbuilt datasets in R can be accessed by using:

>data()

>data(iris)

#Load the dataset into current workspace

>print(iris)

#Display the contents of dataset

>str(iris)

structure of dataset

>head(iris)

#First 6 records can be displayed(default)

>head(iris,10)

#First 10 records can be displayed

>tail(iris)

#Last 6 records can be displayed

Experiment 10: Illustrate the descriptive statistics using summary() in R

Aim: To generate the statistical measurements for a given dataset using summary() method in R.

```
> mean(iris$Sepal.Length)  
[1] 5.843333
```

```
> median(iris$Sepal.Length)  
[1] 5.8
```

```
> quantile(iris$Sepal.Length)  
 0%   25%   50%   75%  100%  
4.3   5.1   5.8   6.4   7.9
```

Experiment 10: Illustrate the descriptive statistics using summary() in R

Aim: To generate the statistical measurements for a given dataset using summary() method in R.

First and third quartile:

```
>quantile(iris$Sepal.Length, 0.25) # first quartile  
25%  
5.1
```

```
>quantile(iris$Sepal.Length, 0.75) # third quartile  
75%  
6.4
```

Experiment 10: Illustrate the descriptive statistics using summary() in R

Aim: To generate the statistical measurements for a given dataset using summary() method in R.

```
>mean(iris$Sepal.Length)
```

```
[1] 5.843333
```

```
>median(iris$Sepal.Length)
```

```
[1] 5.8
```

```
>sd(iris$Sepal.Length)
```

```
[1] 0.8280661
```

Experiment 10: Illustrate the descriptive statistics using summary() in R

Aim: To generate the statistical measurements for a given dataset using summary() method in R.

```
>summary(iris)
```

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
Min. :4.300	Min. :2.000	Min. :1.000	Min. :0.100
1st Qu.:5.100	1st Qu.:2.800	1st Qu.:1.600	1st Qu.:0.300
Median :5.800	Median :3.000	Median :4.350	Median :1.300
Mean :5.843	Mean :3.057	Mean :3.758	Mean :1.199
3rd Qu.:6.400	3rd Qu.:3.300	3rd Qu.:5.100	3rd Qu.:1.800
Max. :7.900	Max. :4.400	Max. :6.900	Max. :2.500

Species

setosa :50
versicolor:50
virginica :50

Experiment 10: Illustrate the descriptive statistics using summary() in R

Aim: To generate the statistical measurements for a given dataset using summary() method in R.

```
>summary(iris[1:50,])
```

```
>summary(iris[51:100,])
```

```
>summary(iris[101:150,])
```