	Bansilal Ramnath Agarwal Charitable Trust's Vishwakarma Institute of Information Technology **Department of Artificial Intelligence and Data Science**

Name: Siddhesh Dilip Khairnar		
Class: SY	Division: B	Roll No: 272028
Semester: IV	Academic Year: 2022-2023	
Subject Name & Code: ES22201AD: Probability and Statistics		
Title of Assignment: Variance, standard deviation, quartiles, inter quartiles in R		
Date of Performance: 03-04-2023	Date of Submission: 10-04-2023	

ASSIGNMENT NO. 6

Background Information:

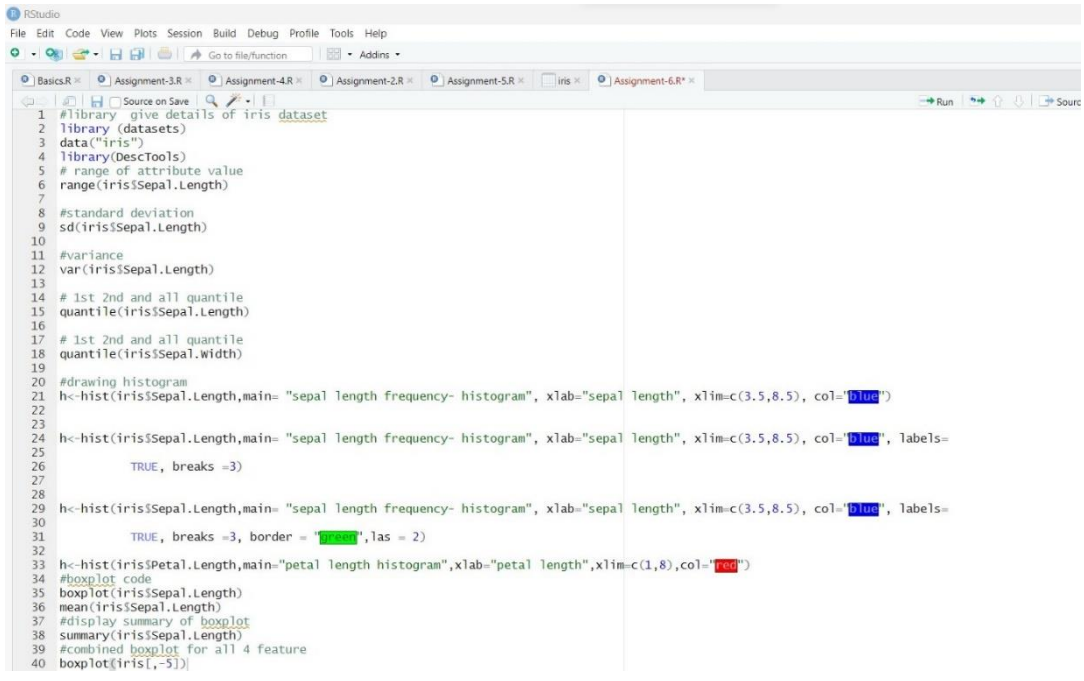
1. Variance: Variance is a measure of how spread out a dataset is, calculated as the average of the squared differences from the mean. A higher variance indicates a wider range of values in the dataset. In R, you can calculate variance using the `var () ` function.

2. Standard deviation: Standard deviation is a measure of how spread out a dataset is, calculated as the square root of the variance. A higher standard deviation indicates a wider range of values in the dataset. In R, you can calculate standard deviation using the `Sd () ` function.

3. Quartiles: Quartiles are values that divide a dataset into four equal parts. The first quartile (Q1) is the value below which 25% of the data falls, the second quartile (Q2) is the median of the data, and the third quartile (Q3) is the value below which 75% of the data falls. In R, you can calculate quartiles using the `quantile () ` function.

4. Interquartile range (IQR): The interquartile range (IQR) is the range between the first and third quartiles and represents the middle 50% of the data. In R, you can calculate the IQR by subtracting the first quartile from the third quartile.

Program and Output:



RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

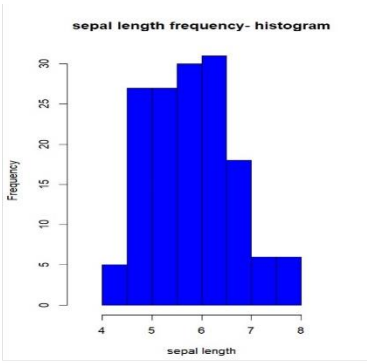
Go to file/function | Addins

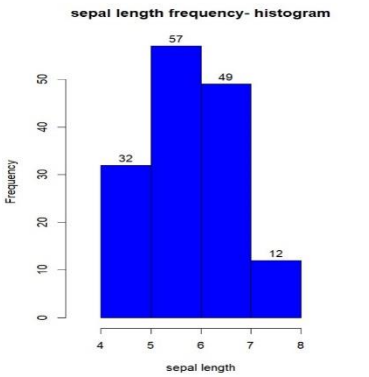
Source

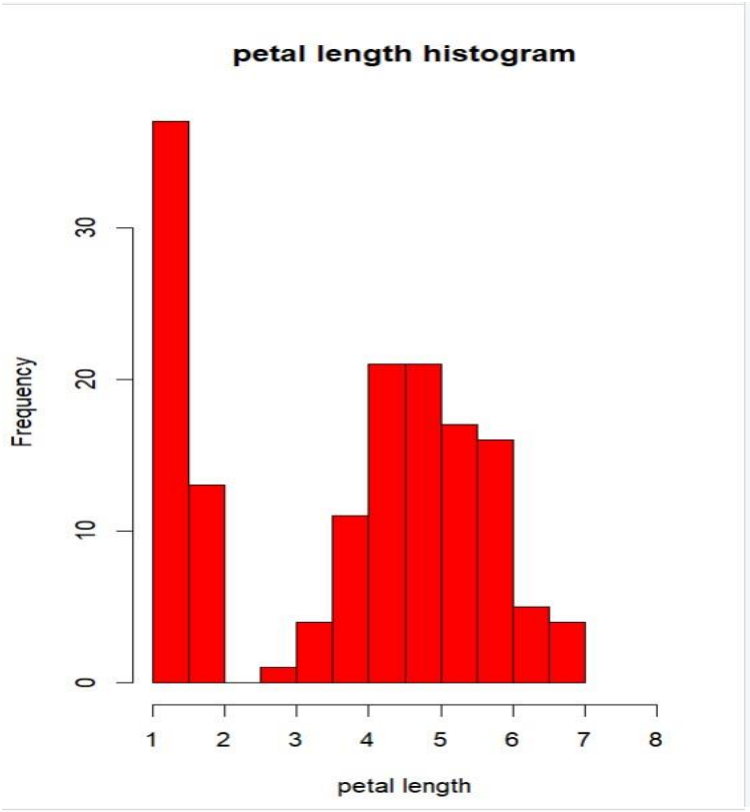
Console | Terminal | Background Jobs

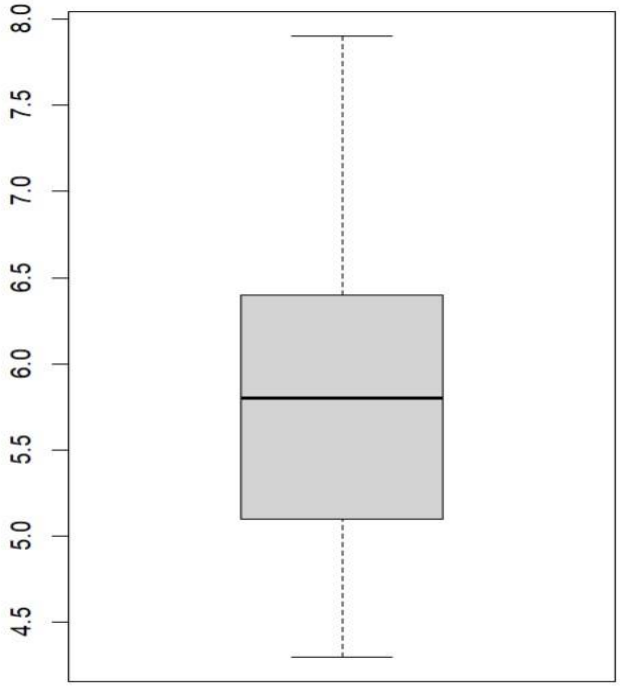
R 4.2.2 · ~/

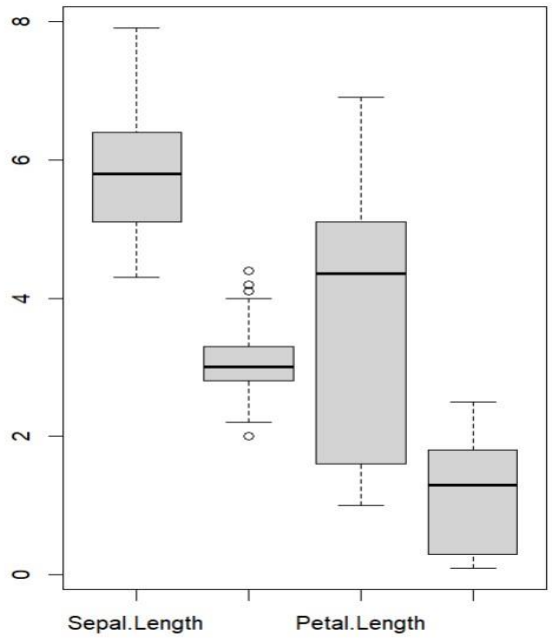
```
> #library  give details of iris dataset
> library (datasets)
> data("iris")
> library(DescTools)
> # range of attribute value
> range(iris$Sepal.Length)
[1] 4.3 7.9
> #standard deviation
> sd(iris$Sepal.Length)
[1] 0.8280661
> #variance
> var(iris$Sepal.Length)
[1] 0.6856935
> # 1st 2nd and all quantile
> quantile(iris$Sepal.Length)
  0%  25%  50%  75% 100%
 4.3  5.1  5.8  6.4  7.9
> # 1st 2nd and all quantile
> quantile(iris$Sepal.Width)
  0%  25%  50%  75% 100%
 2.0  2.8  3.0  3.3  4.4
>
```






petal length histogram
Frequency
0
10
20
30
1
2
3
4
5
6
7
8
petal length


4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0




```
> #drawing histogram
> h<-hist(iris$Sepal.Length,main= "sepal length frequency- histogram", xlab="sepal length", xlim=c(3.5,8.5), col="blue")
> h<-hist(iris$Sepal.Length,main= "sepal length frequency- histogram", xlab="sepal length", xlim=c(3.5,8.5), col="blue", labels=
+
+           TRUE, breaks =3)
> h<-hist(iris$Sepal.Length,main= "sepal length frequency- histogram", xlab="sepal length", xlim=c(3.5,8.5), col="blue", labels=
+
+           TRUE, breaks =3, border = "green",las = 2)
> h<-hist(iris$Petal.Length,main="petal length histogram",xlab="petal length",xlim=c(1,8),col="red")
> #boxplot code
> boxplot(iris$Sepal.Length)
> mean(iris$Sepal.Length)
[1] 5.843333
> #display summary of boxplot
> summary(iris$Sepal.Length)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  4.300   5.100   5.800   5.843   6.400   7.900
> #combined boxplot for all 4 feature
> boxplot(iris[,-5])
```

Conclusion: Hence, in this assignment we've learned and implemented various variability measure in R Studio