**ITA0448 – STATISTICS WITH R PROGRAMMING FOR VECTORIZED EXPRESSIONS**

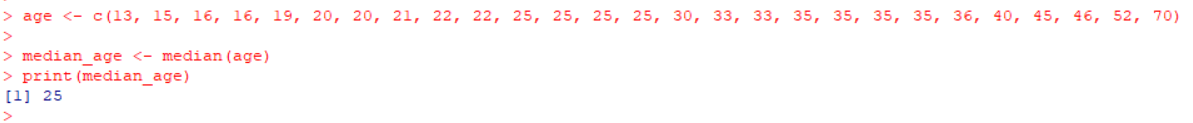
**SUBMITTED BY: DEVADARSHINI J REG NO: 192124092**

**DATE: 23/03/2023**

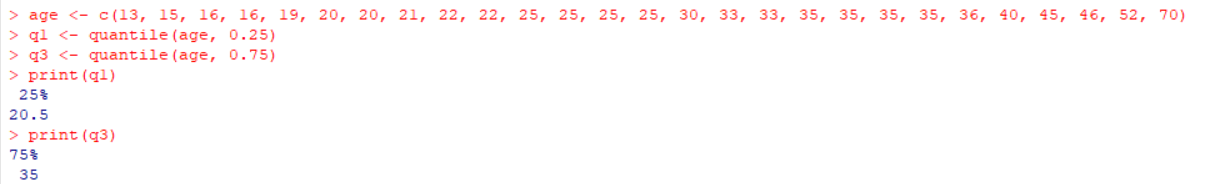
1. **Suppose that the data for analysis includes the attribute age. The age values for the data**

**tuples are (in increasing order) 13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33,**

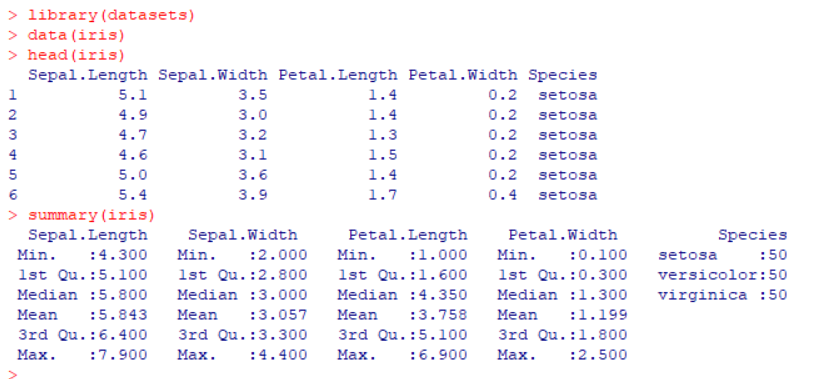
**33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70. What is the median?**

****

1. **Suppose that the data for analysis includes the attribute age. The age values for the data tuples are (in increasing order) 13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70. Can you find (roughly) the first quartile (Q1) and the third quartile (Q3) of the data?**

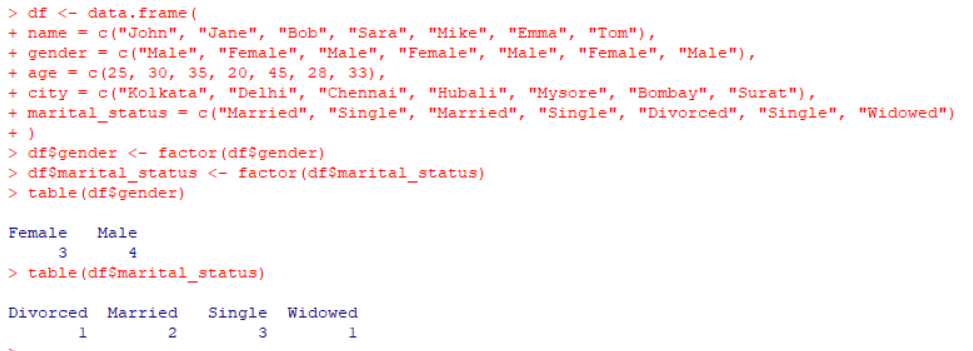
****

1. **Load iris Dataset which is inbuilt in R .explore the dataset in terms of dimension and summary statistics**

****

1. **Find the categorical column data and convert that to factor form, also find the number of**

**rows for each factors in dataset.**

****

1. **Find mean of numeric data in dataset based on Species group. and plot Bar chart (use ggplot ) to interpret same**

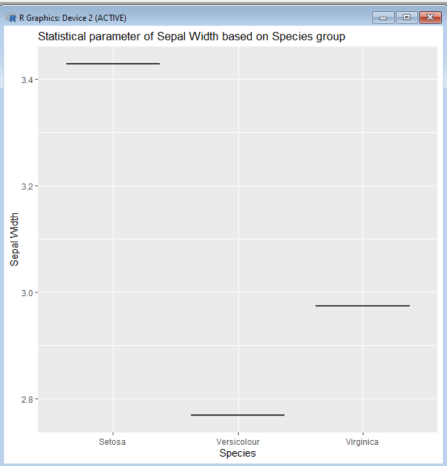
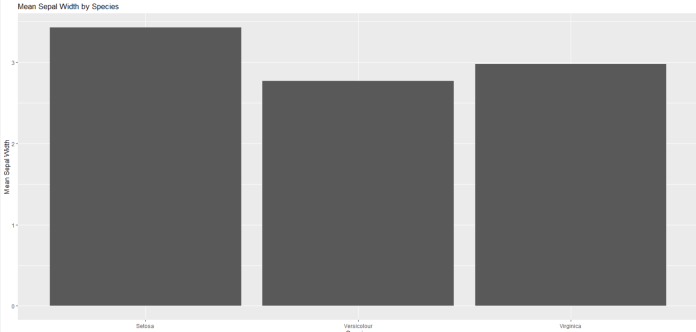
**Species Sepal Length Sepal Width Petal Length Petal Width**

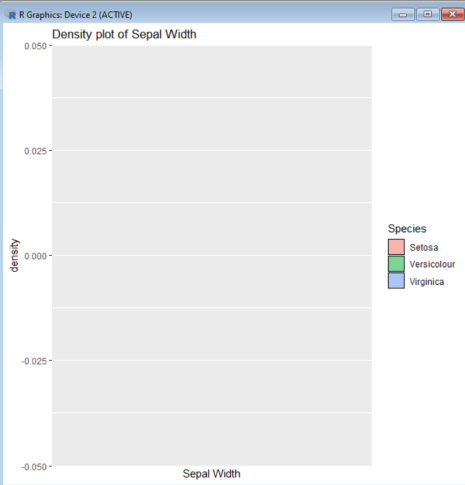
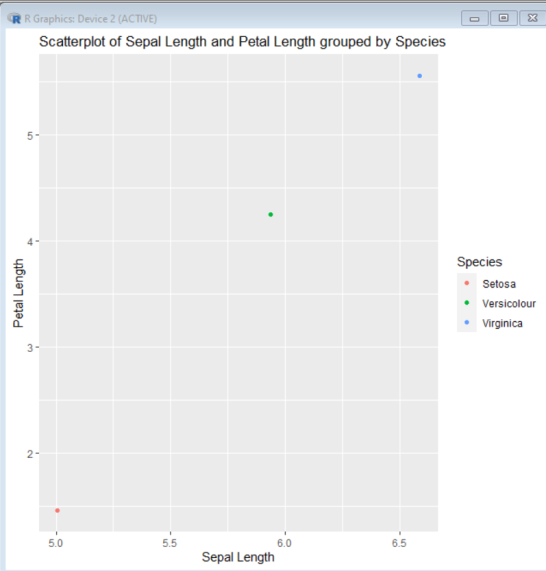
**Setosa 5.006 3.428 1.462 0.246**

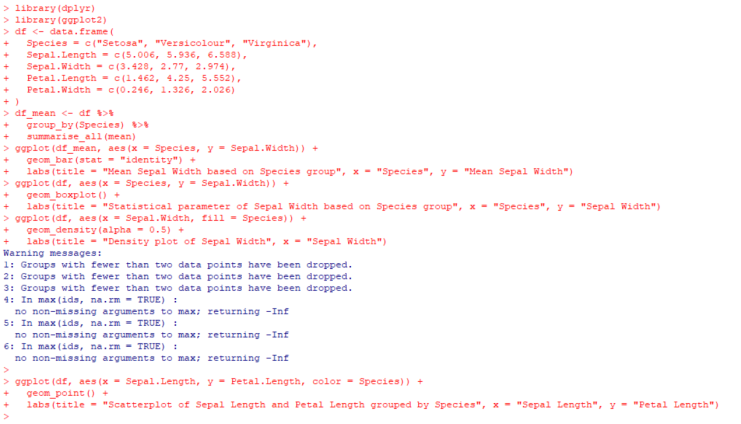
**Versicolour 5.936 2.770 4.250 1.326**

**Virginica 6.588 2.974 5.552 2.026**

1. **Draw a suitable plot which summaries statistical parameter of Sepal.Width based on Species group**
2. **Draw a suitable plot to find the skewness of the data for Sepal.Width and print the comment about skewness.**
3. **Draw ggplot2 scatterplot showing the variables Sepal.Length and Petal.Length grouped by the three-level factor “Species”.**

****

** **

****