5CS462 : PE5 Data Mining Lab

PRN/ Roll No: 2019BTECS00113

Full name: Sanket Rajaram Mote

Batch: B8

Assignment No. 1

Title: Data Analysis Tool (GUI)

Objective:

1. To Calculate and show the measures of central tendency for uploaded data : mean ,median , mode , midrange , variance and standard deviation

2. To Calculate and show the dispersion of data : range , quartiles , interquartile range, five-number summary

3. To Graphical display of above calculated statistical description of data (provide the facility - UI form to choose different attributes from uploaded data set) : a. Quantile plot b. Quantile-quantile (q-q) plot c. Histogram d. Scatter plot e. Boxplot

Introduction & Theory:

Measures of Central Tendency :
1. Mean : add up the values in the data set and then divide by the number of values that you added.
2. Median : list the values of the data set in numerical order and identify which value appears in the middle of the list. To find the mode, identify which value in the data set occurs most often.
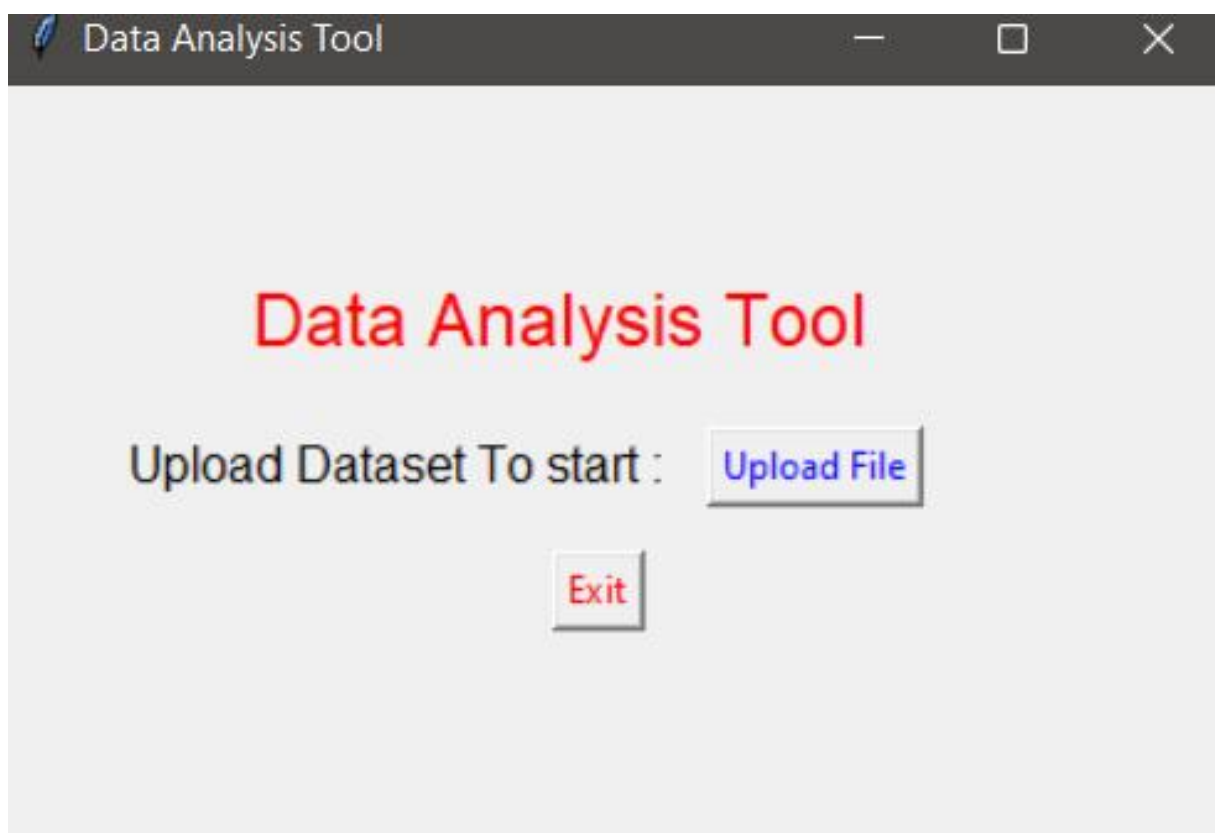
3. Mode : To identify which value in the data set occurs most often

4. Midrange : the number that is exactly halfway between the minimum and maximum numbers in a set of data.

5. Variance : It is a measure of dispersion that takes into account the spread of all data points in a data set.

6. standard deviation : It Is providing a measure of variability in a sample or population.

Procedure / Experiment:
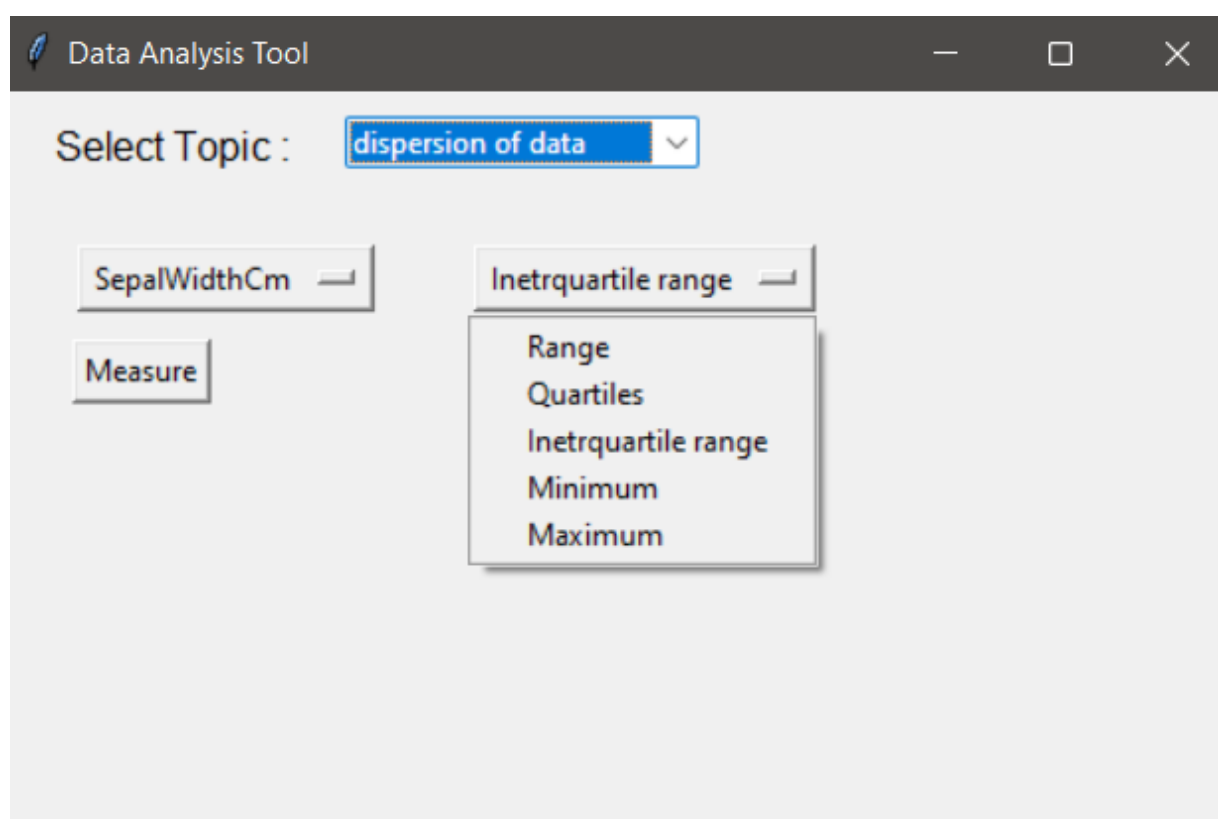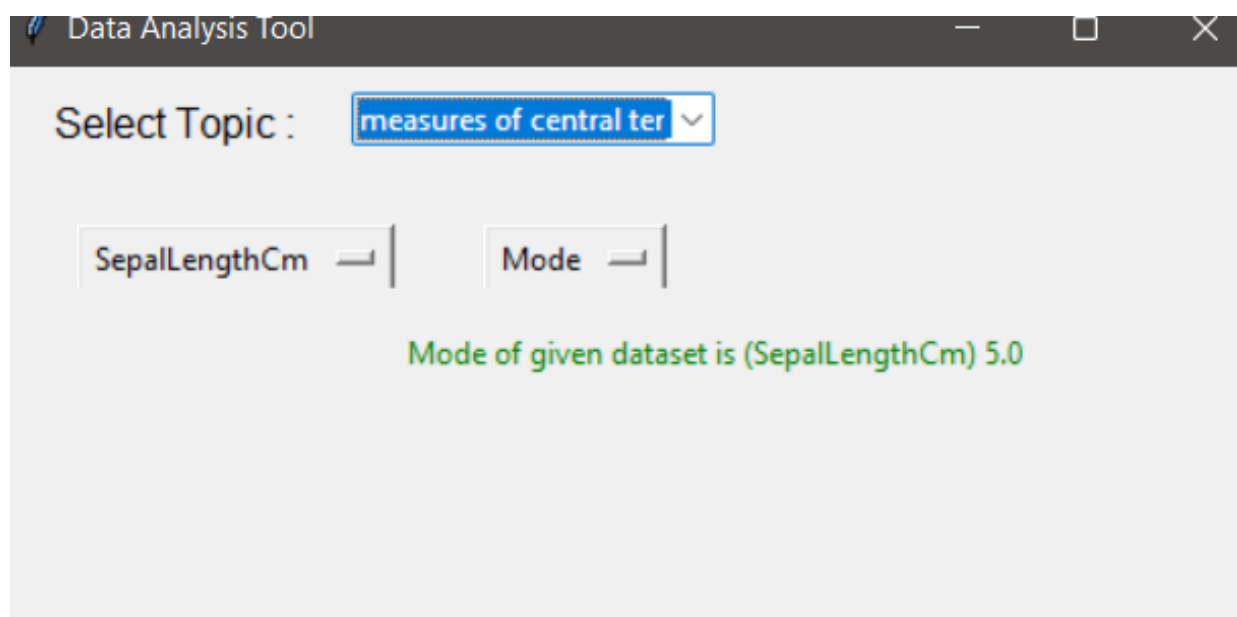
Data Analysis Tool
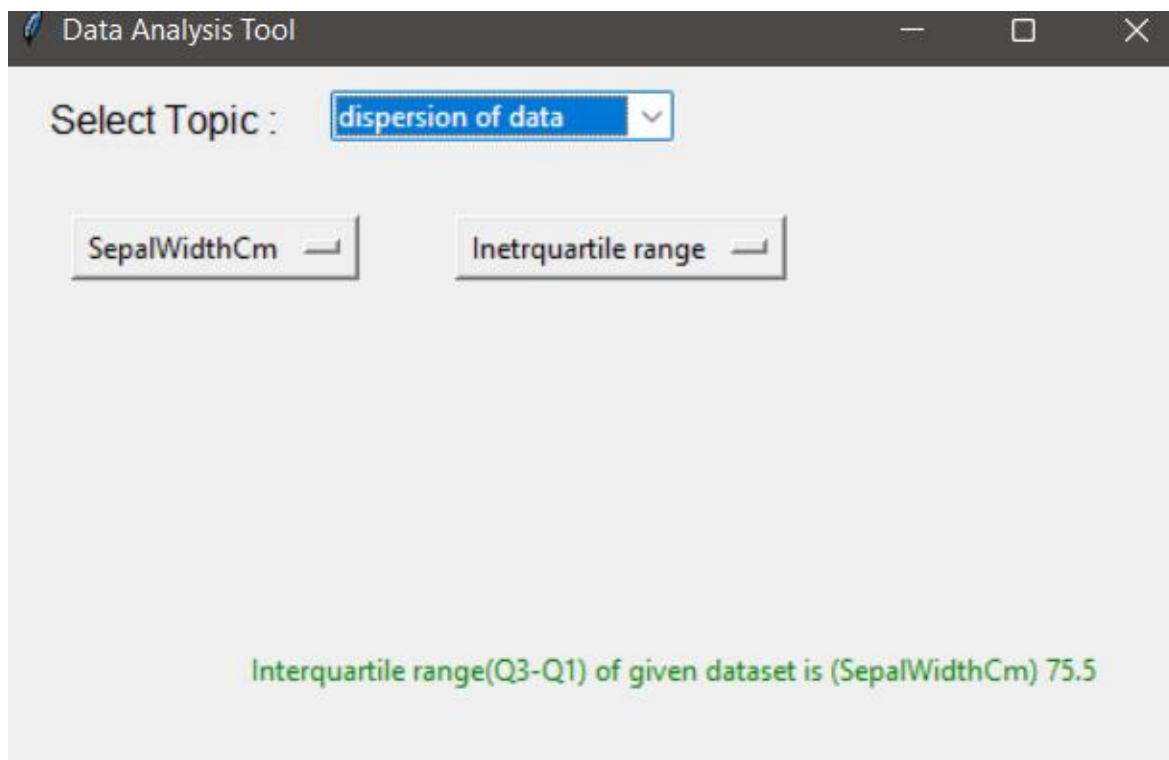
Select Topic :

Data Display
measures of central tender
dispersion of data
statistical description

tk

Id	epalLengthCr	SepalWidthCn	etalLengthCn	PetalWidthCn	Species
1	5.1	3.5	1.4	0.2	Iris-setosa
2	4.9	3.0	1.4	0.2	Iris-setosa
3	4.7	3.2	1.3	0.2	Iris-setosa
4	4.6	3.1	1.5	0.2	Iris-setosa
5	5.0	3.6	1.4	0.2	Iris-setosa
6	5.4	3.9	1.7	0.4	Iris-setosa

Data Analysis Tool

Select Topic : measures of central ter

SepalLengthCm   Mode

Mode of given dataset is (SepalLengthCm) 5.0

Data Analysis Tool

Select Topic : dispersion of data

SepalWidthCm   Inetrquartile range

Measure

Range
Quartiles
Inetrquartile range
Minimum
Maximum

Conclusion:

Successfully Calculated and show the measures of central tendency for uploaded data : mean ,median , mode , midrange , variance and standard deviation and displayed using GUI

References:

https://www.analyticsvidhya.com/blog/2021/04/dispersion-of-data-range-iqr-variance-standard-deviation/

Dataset:

https://archive.ics.uci.edu/ml/datasets/Iris