**Final Year B.Tech. (CSE) – I : [ 2022-23 ]**

**5CS462 : PE5 Data Mining Lab**

**PRN/ Roll No: 2019BTECS00113**

**Full name: Sanket Rajaram Mote**

**Batch: B8**

**Assignment No. 2**

**Title: Data Analysis Tool (GUI)**

**Objective:**

To calculate Chi- Square

2. To Calculate Pearson coefficient and covariance

3. Implement Min-max Normalization

4. To Implement Z-Score normalisation

**Introduction & Theory:**

**1.Chi-Square Test :**

A chi-square test is used to help determine if observed results are in line with expected results, and to rule out that observations are due to chance

**2. Correlation coefficient :**

It measures the extent to which the value of one variable changes

with another.

**3.Min-max normalization :** Min Max is a data normalization

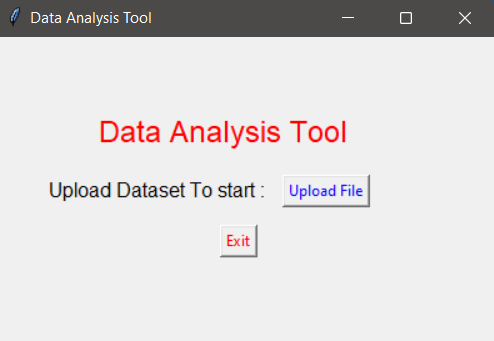
technique like Z score, decimal scaling, and normalization with standard deviation.

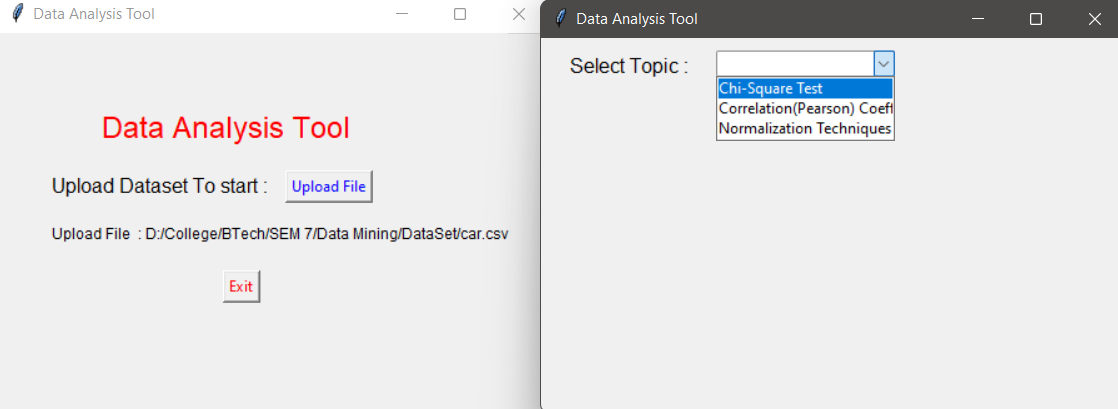
**4. Z-Score normalization :**

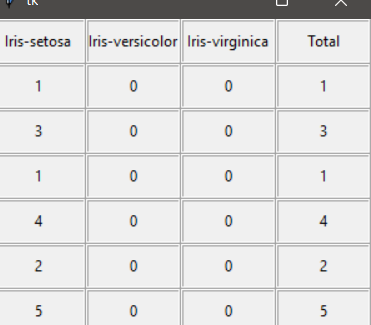
Z-Score value is to understand how far the data point is from the

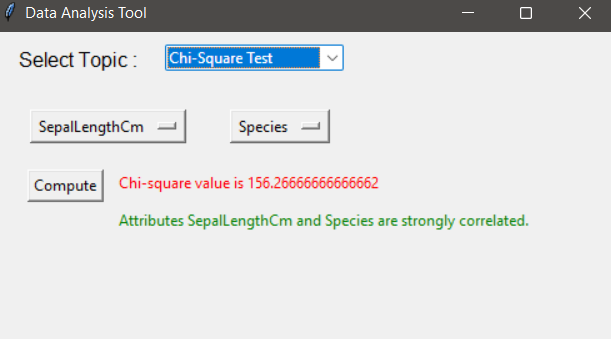
mean. Technically, it measures the standard deviations below or above the mean.

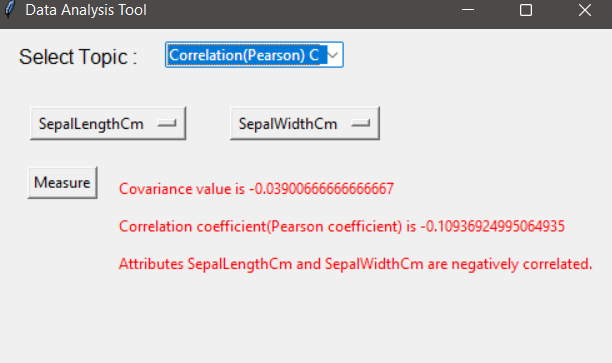
**Screen Shots :**

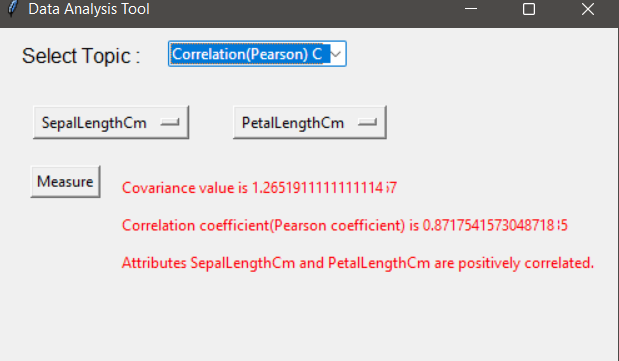
****

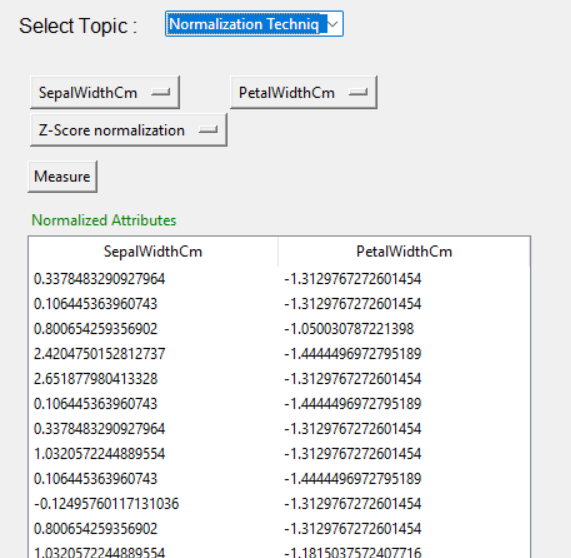
****

****

****

****

****

****

**Conclusion:**

Successfully implemented chi square test , Correlation coefficient ,

Covariance, Different Type of Normalisation using Gui

**References:**

[**https://www.geeksforgeeks.org/data-normalization-in-data-mining/**](https://www.geeksforgeeks.org/data-normalization-in-data-mining/)

[**https://www.kaggle.com/code/dehaozhang/chi-squared-test-on-attrition/notebook**](https://www.kaggle.com/code/dehaozhang/chi-squared-test-on-attrition/notebook)

**Dataset:**

[**https://archive.ics.uci.edu/ml/datasets/Iris**](https://archive.ics.uci.edu/ml/datasets/Iris)

[**https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29**](https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29)