Student Name : ..... Roll No. .....

Experiment No. : ..... Date : 26/2/24

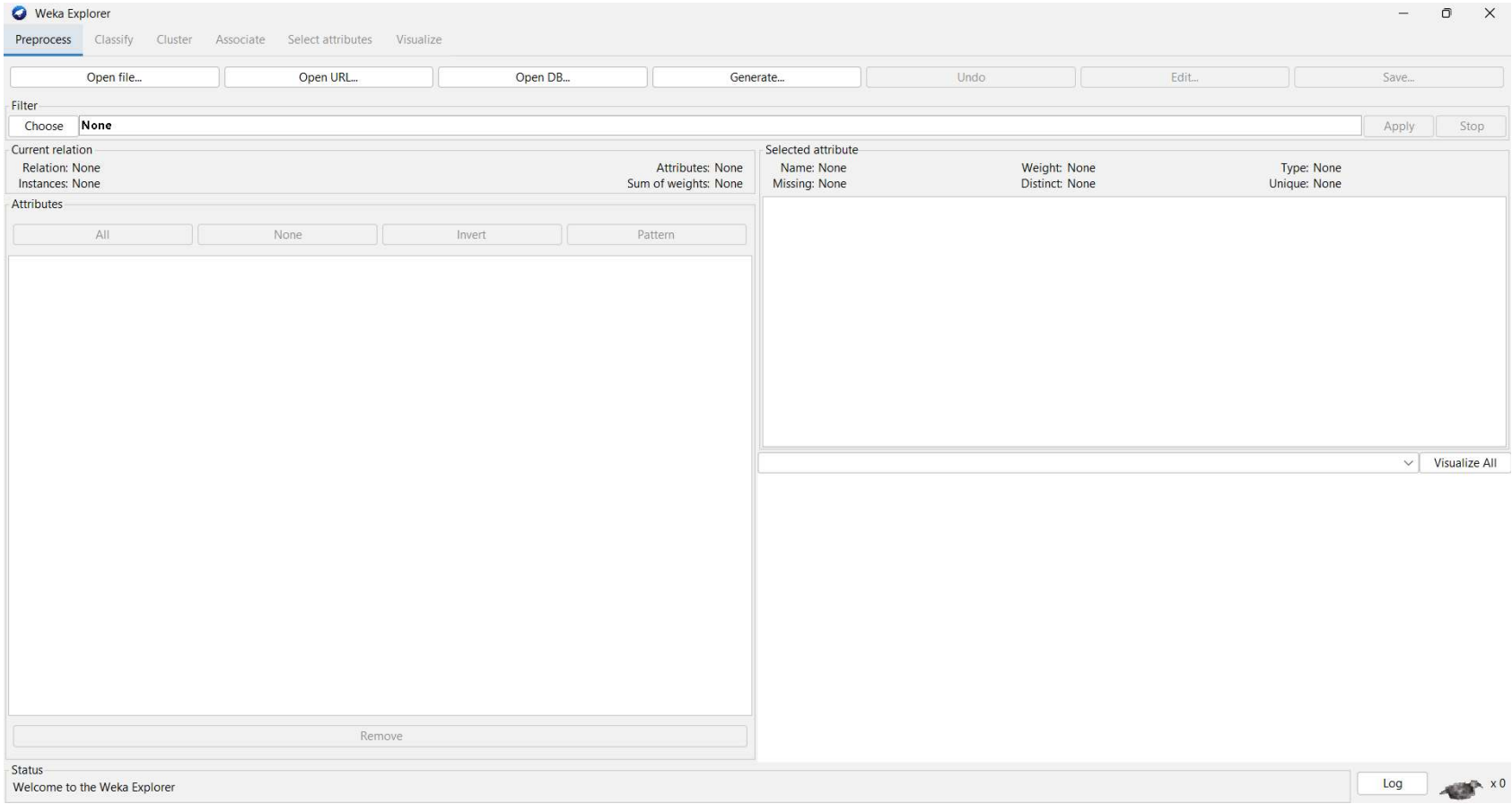
Experiment :- 2

- **Objective :-** Implementation of data preprocessing using weka explorer.
- **Tool used :-** Weka Software
- **Dataset used :-** IRIS dataset (.CSV)
- **Theory :-** Data preprocessing and data cleaning are crucial steps in any data analysis or machine learning project they involve transfering raw data into a clean, consistent and usable format that can be effectively analyzed.

(1) **Detect outliers →** An outliers in a data point that differ significantly from other observation in a dataset. In weka, outlier can be detected using the Inter quantile range (filter / weka unsupervised I) Click on apply to apply the filter. two column would appear as outlier and extreme also, these would be indicated differently using a different colour.

(2) **Detect & remove Null Values →** a Null Value is represents the absence of a value or a missing value in a database or programming content in weka, go to

EXPERIMENT 2



Student Name : ..... Roll No.: .....

Experiment No. : ..... Date : .....

Remove with values in the filter tab (filter / unsupervised / attribute / remove with values). Configure it to remove instances with missing values in the desired attributes.

(3) Convert / transform data types → In weka, select the Numeric transform (filters / unsupervised / attribute / numeric transform) and select the instance you want to apply the transformation from.

(4) fill substitute of Null values :- In weka, we can use replace missing values (filter / unsupervised / attribute / replace missing values) click on apply button and the missing values will be replaced.

~~(5) Result~~

• Result :- We have successfully implemented all the various operations successfully.

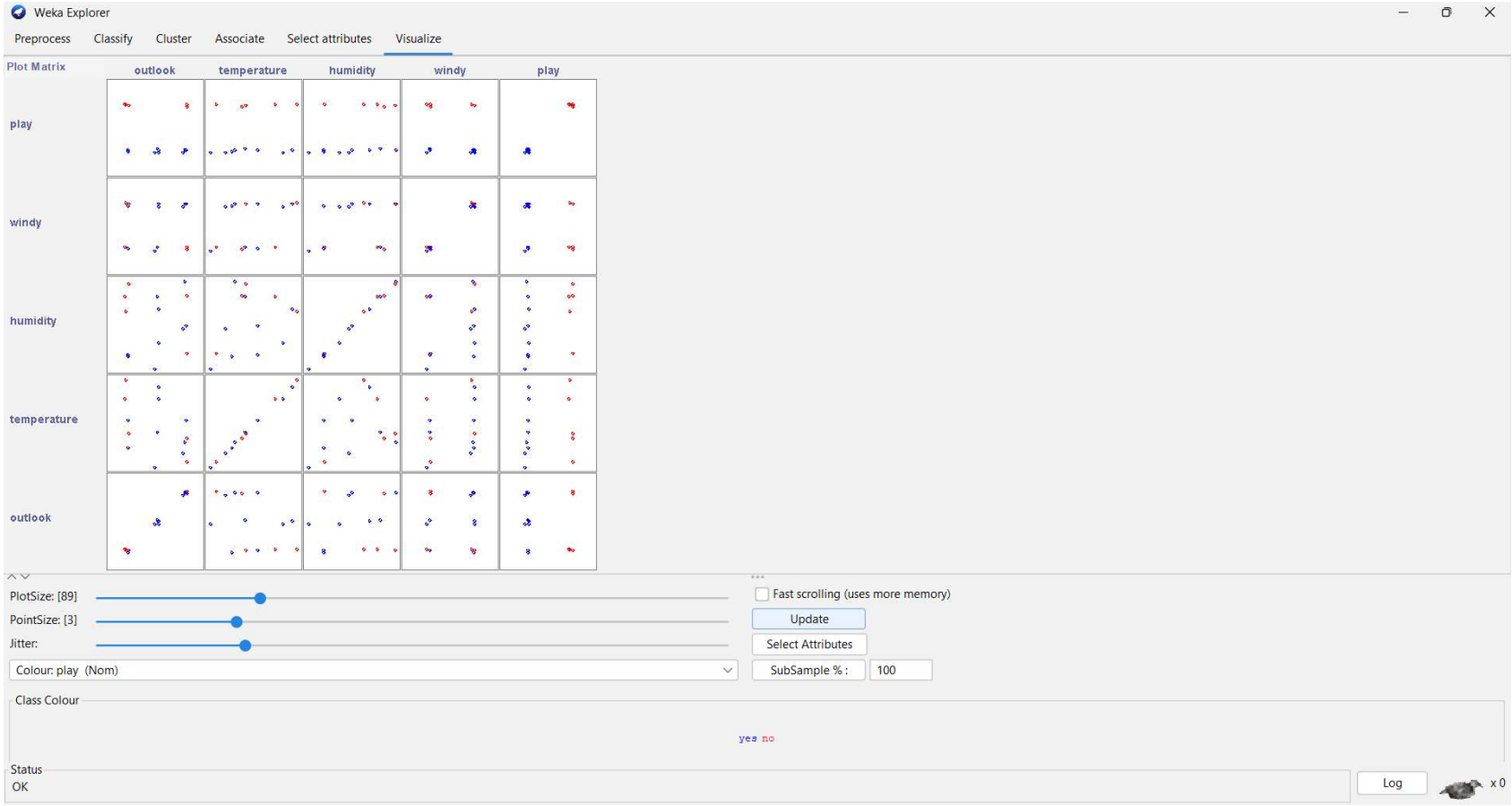
• Viva Question :

Q1 What is meant by data preprocessing ?

→ data preprocessing is the process of cleaning, transforming and organizing raw data into a suitable format for analysis for machine learning.

Page No. .....

EXPERIMENT 2



ent No. : ..... Date : .....

Q2 State the methods / processes of data preprocessing that can be implemented by weka?

→ Methods / process of data preprocessing that can be implemented using WEKA.

1. Handling missing values
2. data transformation
3. discretization
4. feature selection
5. data sampling
6. attribute construction
7. removing unwanted attribute