**DEPARTMENT OF INFORMATION TECHNOLOGY**

**Course: Data Mining & Business Intelligence Lab (ITL601)**

**B.Tech. (Information Technology) – Semester VI**

**Academic Year: 2023-24 (Even Semester)**

**PRACTICAL 2**

**Aim:** Exploration of WEKA and Implementation of Data Mining tasks using WEKA

**Lab Objective:**

1. Identify sources of data for mining, design a Data Warehouse schema, and perform data

exploration.

2. Organize and prepare the data needed for data mining algorithms in terms of attributes and

class inputs, training, validating, and testing files.

**Theory:**

Data Mining:

Data mining is the process of discovering patterns, trends, correlations, or valuable insights from large sets of data. It involves using various techniques and algorithms to analyze and extract meaningful and previously unknown information from structured or unstructured data. The goal of data mining is to uncover hidden knowledge and make informed decisions based on the patterns and relationships identified within the data. This interdisciplinary field combines elements of statistics, machine learning, database management, and artificial intelligence to transform raw data into actionable knowledge for business, research, or decision-making purposes.

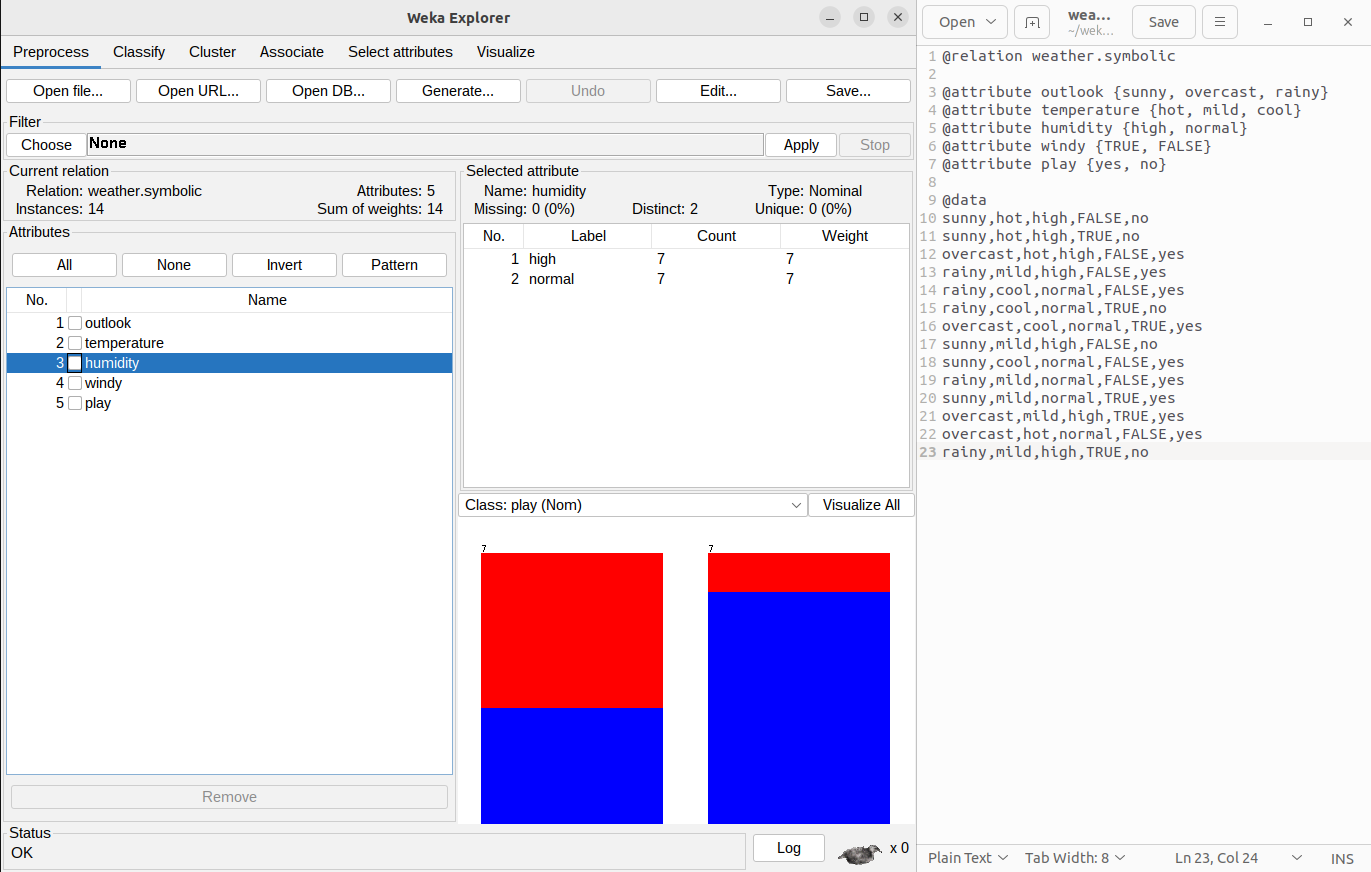
Weka Software:

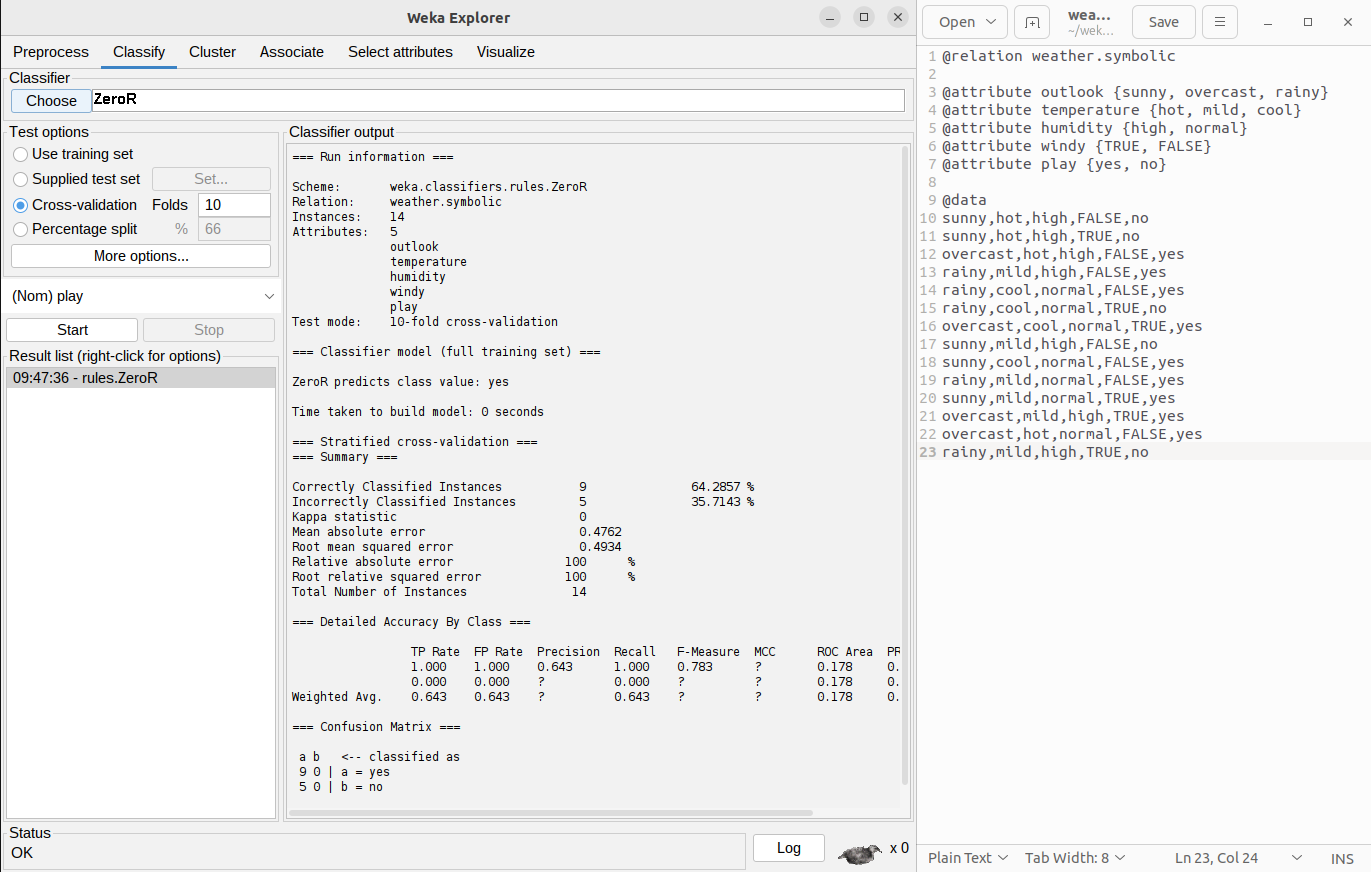
Weka (Waikato Environment for Knowledge Analysis) is a popular open-source suite of machine learning and data mining tools. Developed at the University of Waikato in New Zealand, Weka provides a comprehensive collection of algorithms and tools for data preprocessing, classification, regression, clustering, association rules, and visualization.

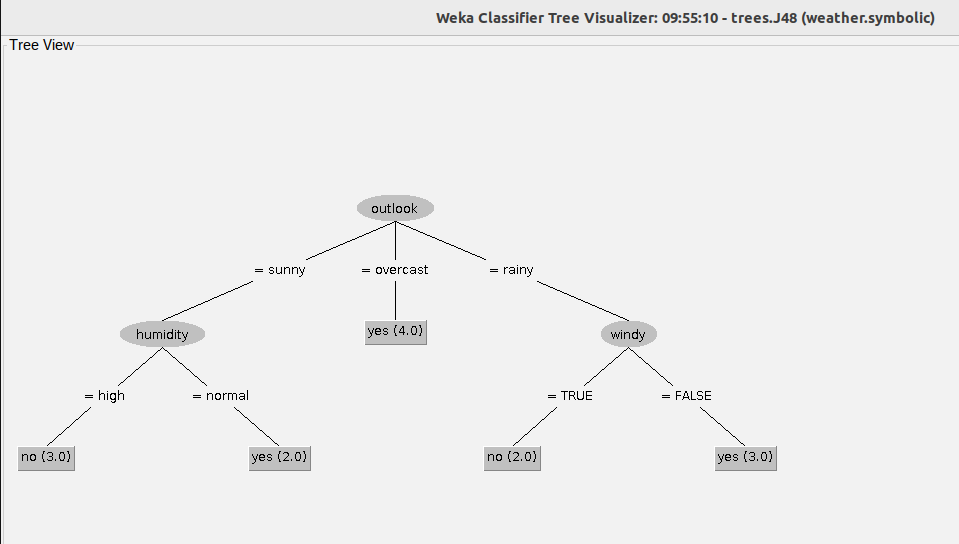
Weka offers a user-friendly graphical interface that simplifies the process of data analysis and model building. It supports various file formats and provides a wide range of data preprocessing techniques, such as attribute selection, data cleaning, and transformation. Weka's extensive library of machine learning algorithms allows users to build and evaluate models for different tasks, including classification, regression, clustering, and more.

**Schema Designs / <Code with Output>:**

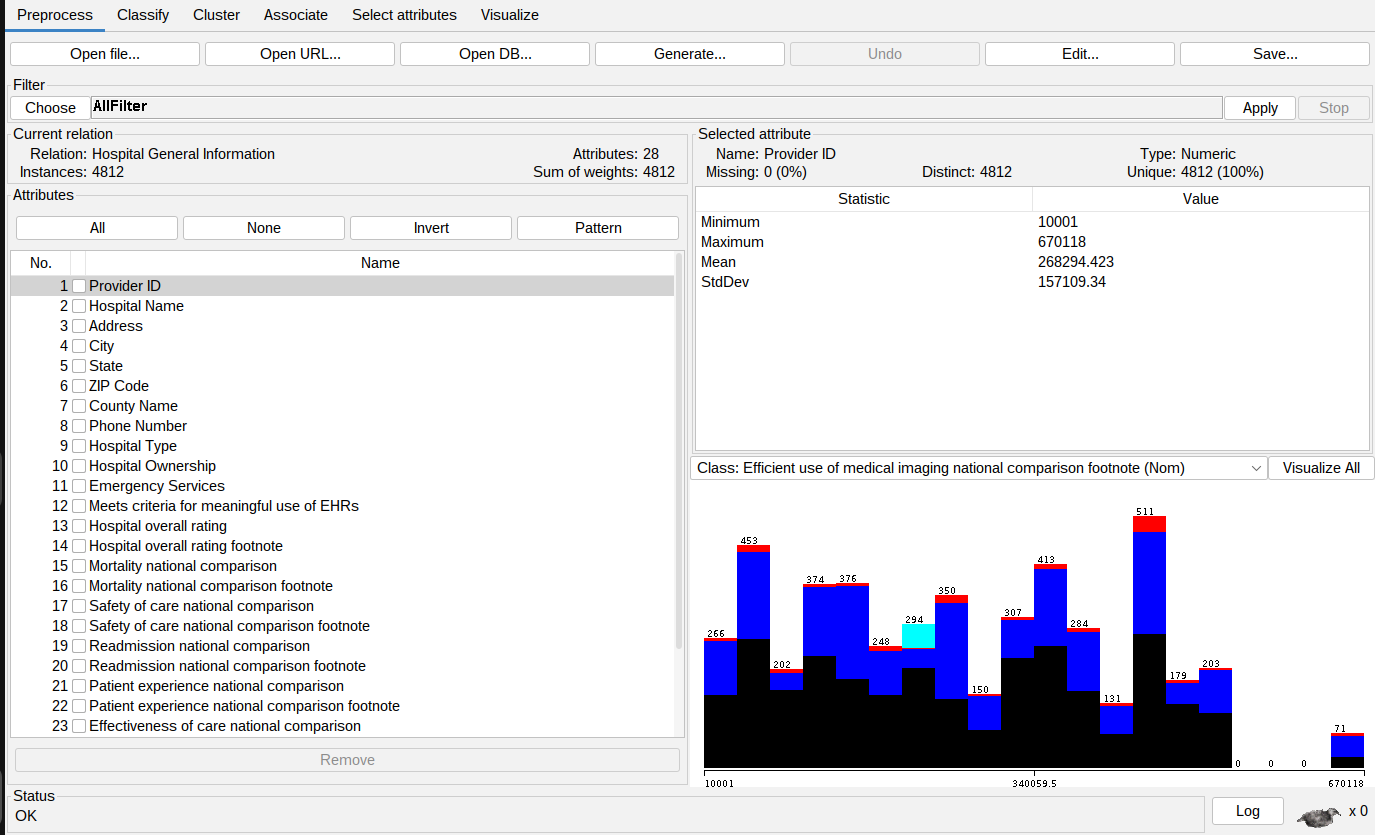
* **Data Exploration in WEKA Using Inbuilt Dataset:**

****

****

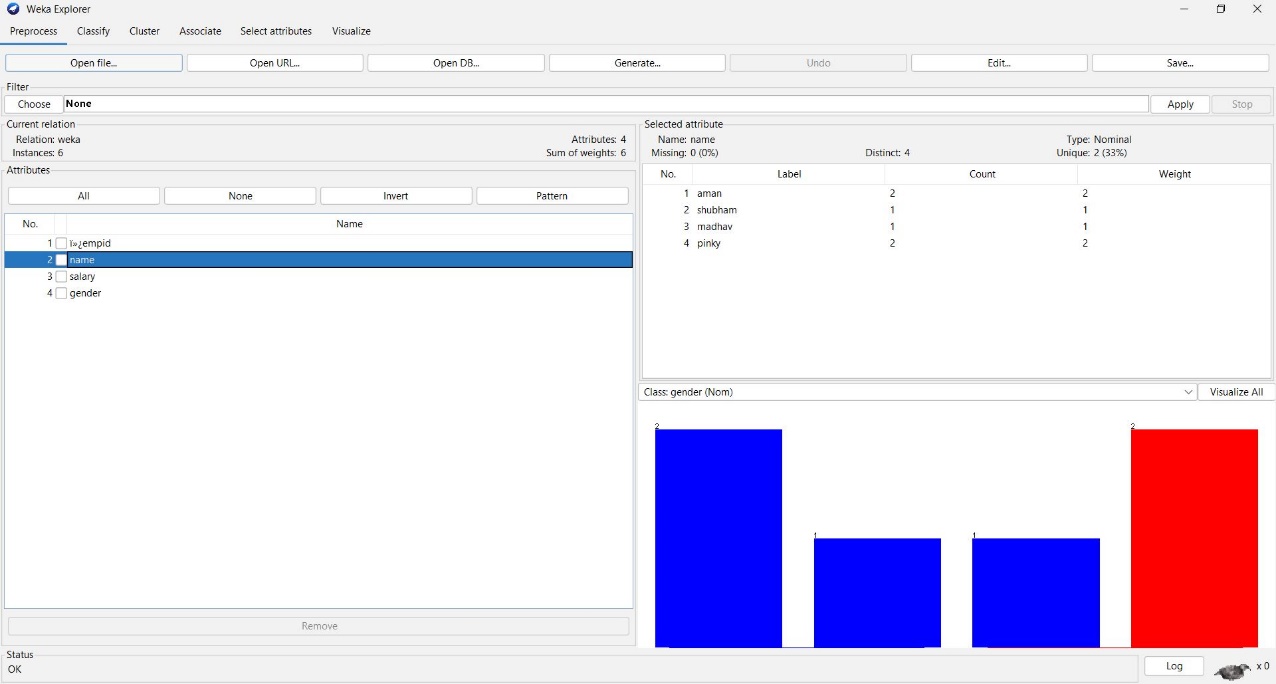
****

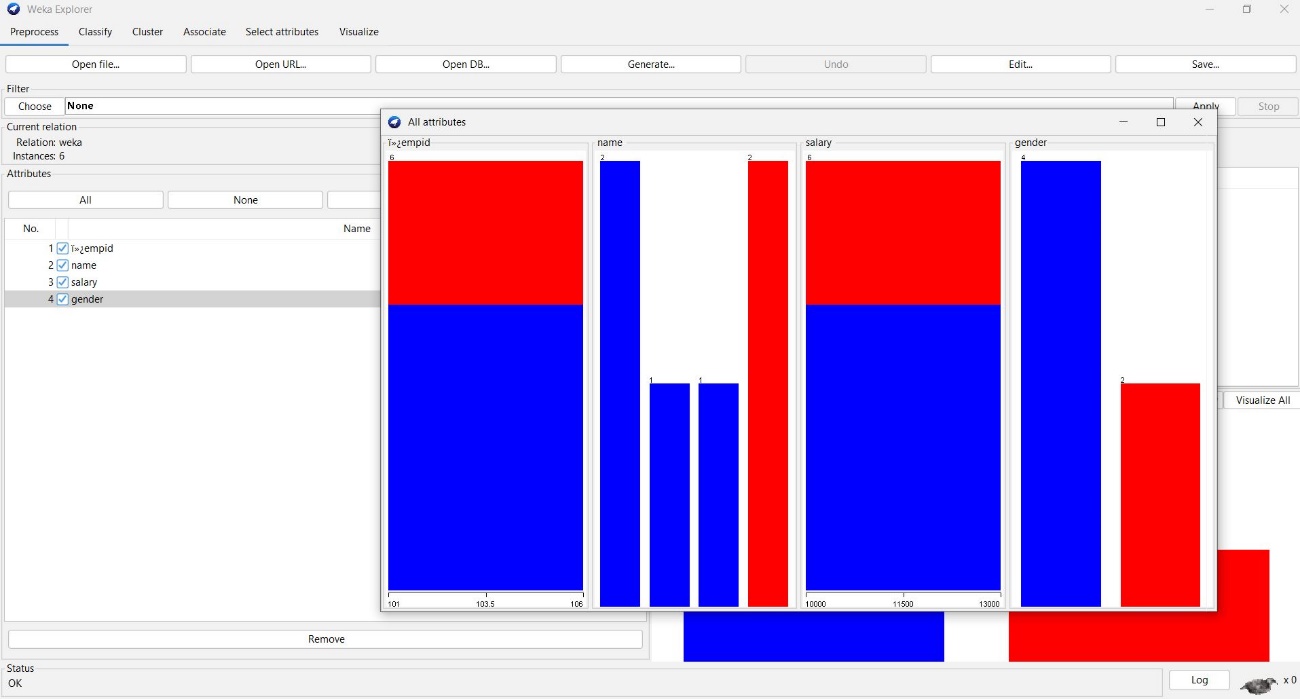
* **Data Exploration in WEKA Using Dataset Imported from Public Repositories(Kaggle) :**

****

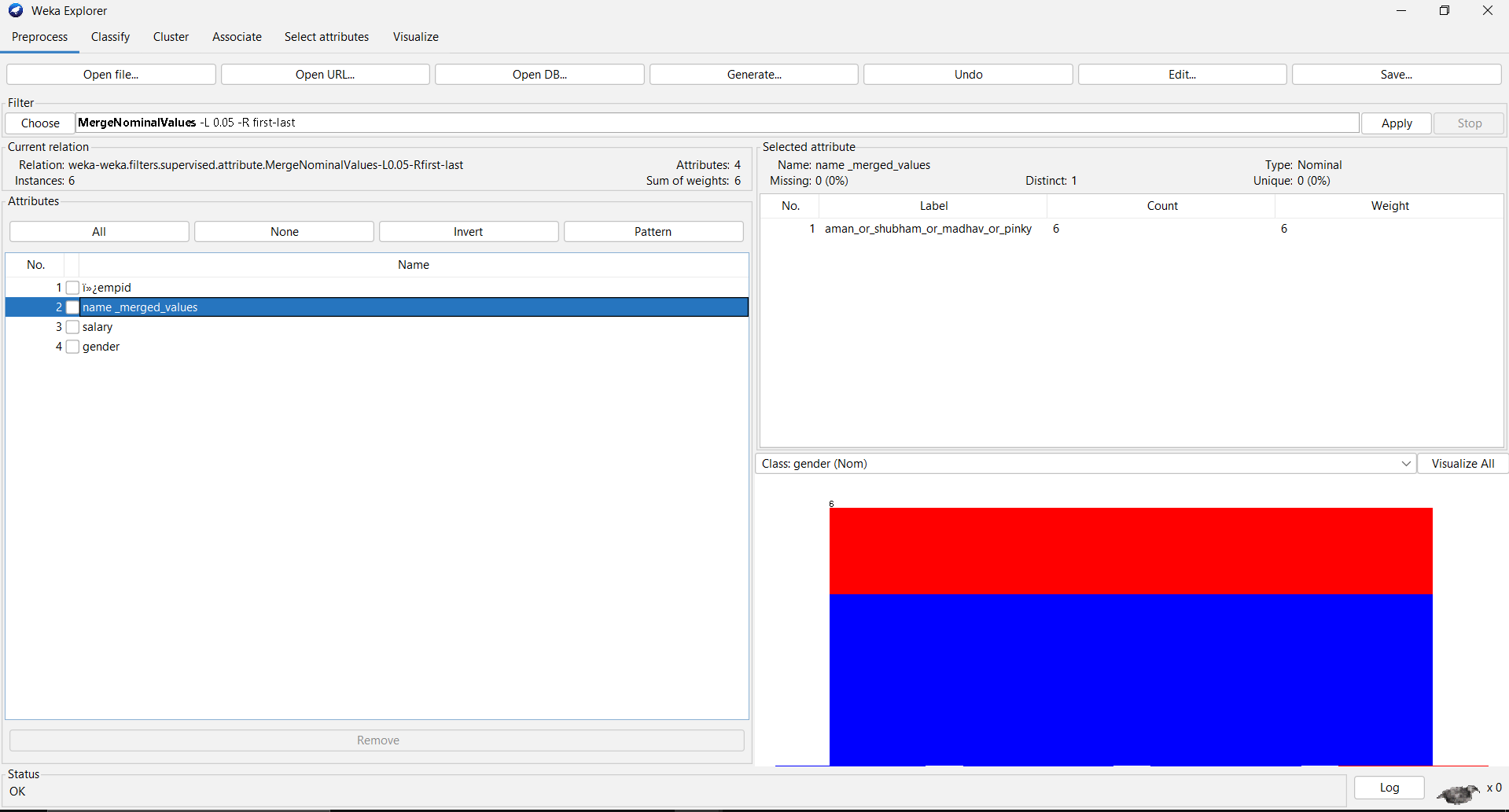
* **Data Exploration in WEKA Using Self-Created Dataset:**

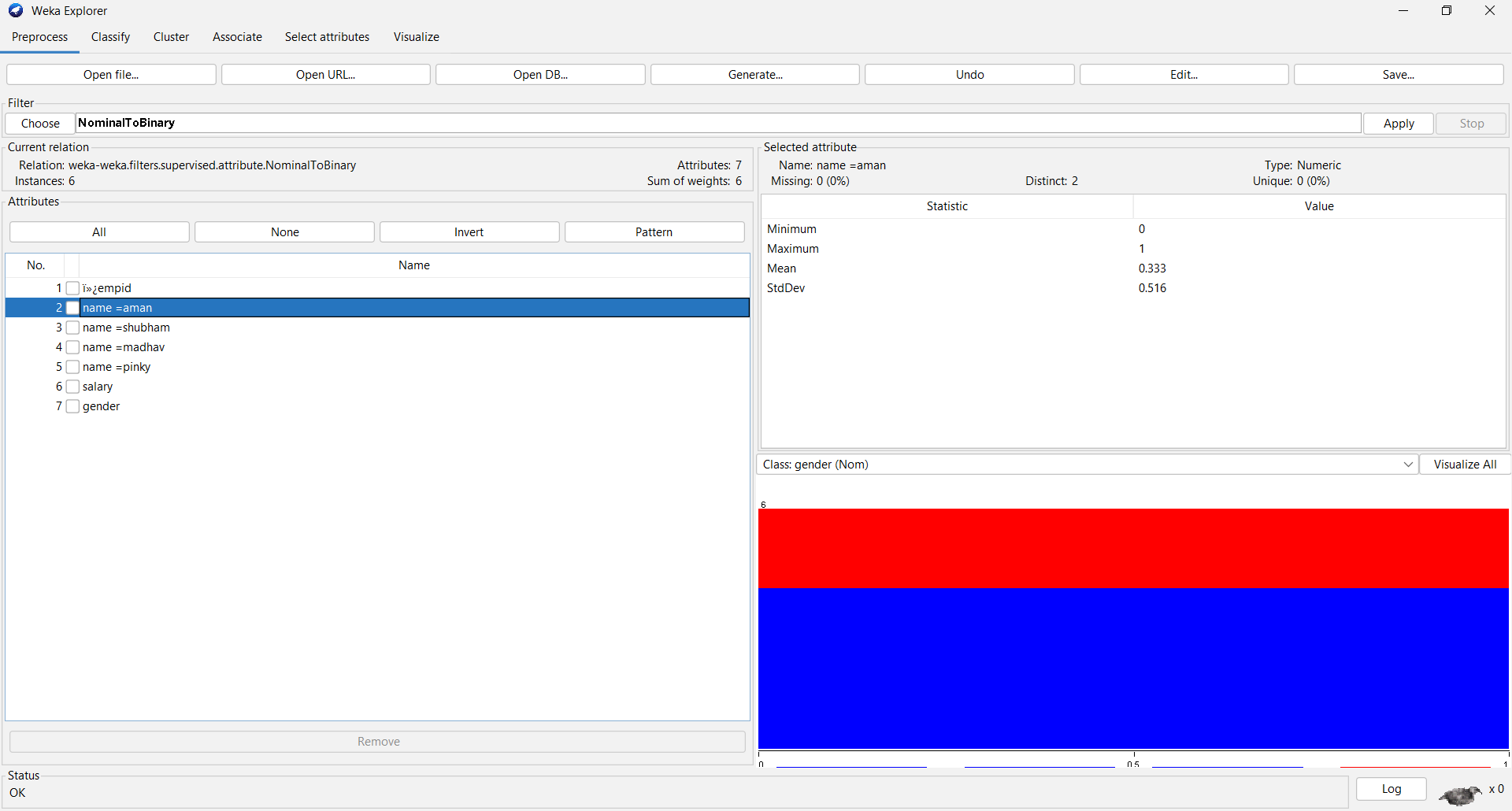
(The data set is about employee data in which there is empid(employee id ),name,salary,gender .

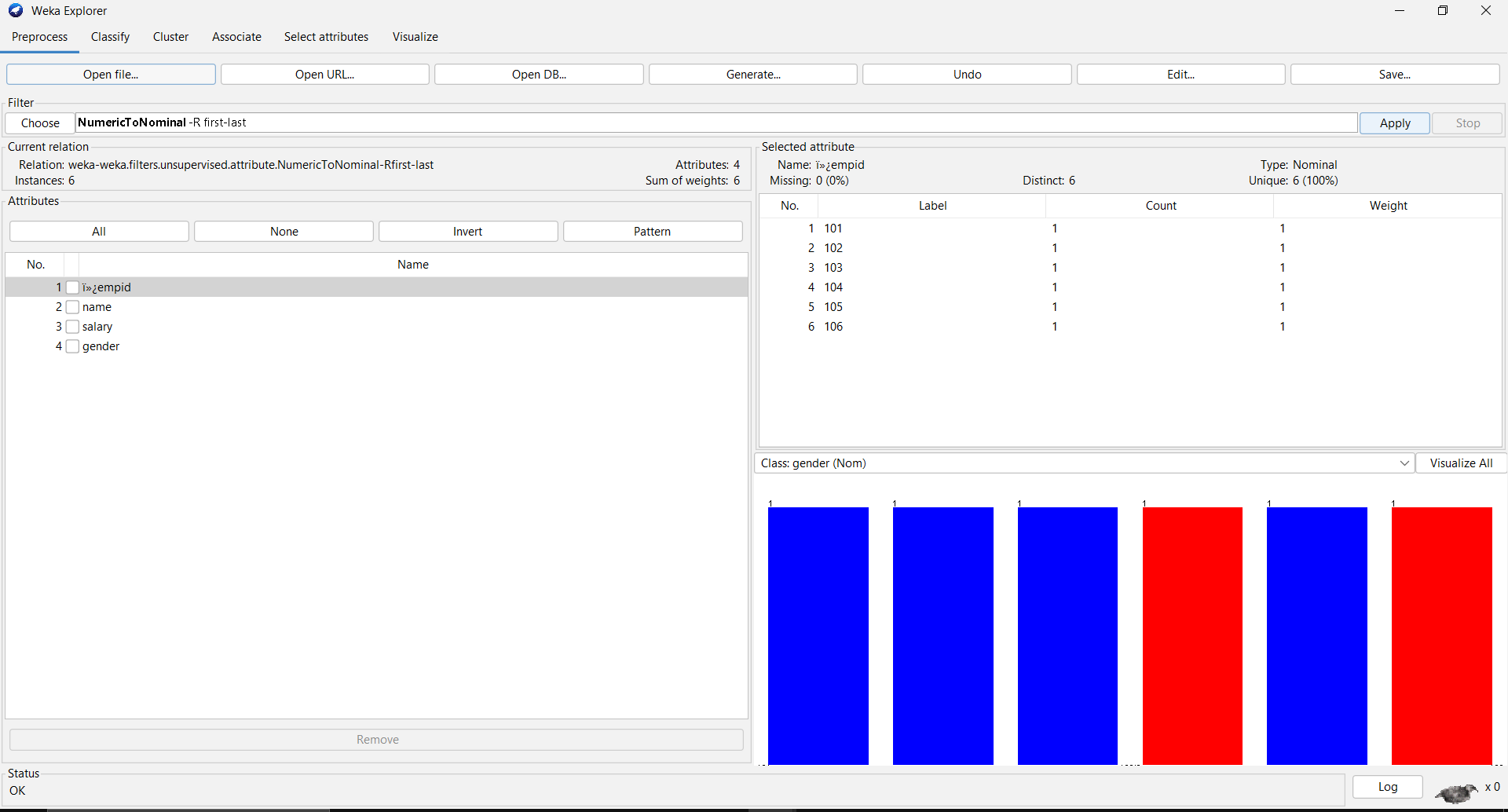


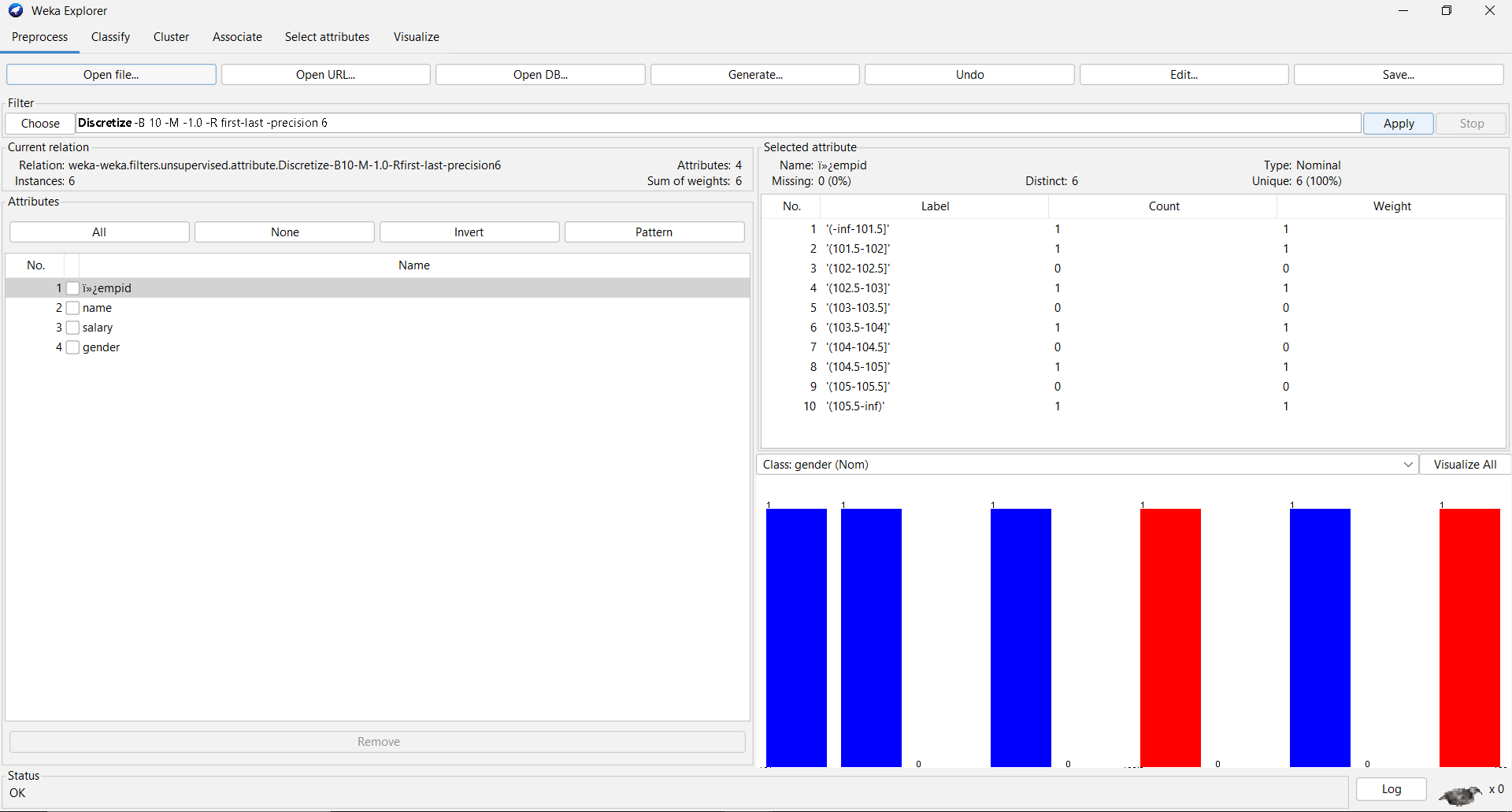


**B) Apply filters**









**Conclusion:**    
In conclusion, Weka serves as an accessible and versatile tool for data mining tasks. Its user-friendly interface, diverse set of built-in tools, and extensive documentation make it suitable for both beginners and experienced data scientists. Weka facilitates seamless exploration, preprocessing, and implementation of machine learning algorithms, making it a valuable asset for various applications in the field of data mining.

**Lab Outcome:** Ability to identify sources of data for mining and design a Data Warehouse schema.

**Submitted Details -**

**Name of Student:** **Parth Malviya**

**Roll No.: 22**

**Date of Performance: 16/01/2024**

**Date of Submission: 28/01/2024**