Name: Allan Rodrigues

Class: TE IT A

Roll no: 59

Pid:191104

St. Francis Institute of Technology, Mumbai-400 103

**Department Of Information Technology**

A.Y. 2021-2022

Class: TE-ITA/B, Semester: VI

Subject: **Business Intelligence Lab**

**Experiment – 3: Data Preprocessing in WEKA Tool**

1. **Aim:** Data Preprocessing in WEKA Tool.
2. **Objectives:** After study of this experiment, the students will be able to
   * Understand and knew how data is preprocessed in Weka.
3. **Outcomes:**

After study of this experiment, the students will be able to

**CO2:** Organize and prepare the data needed for data mining using pre preprocessing

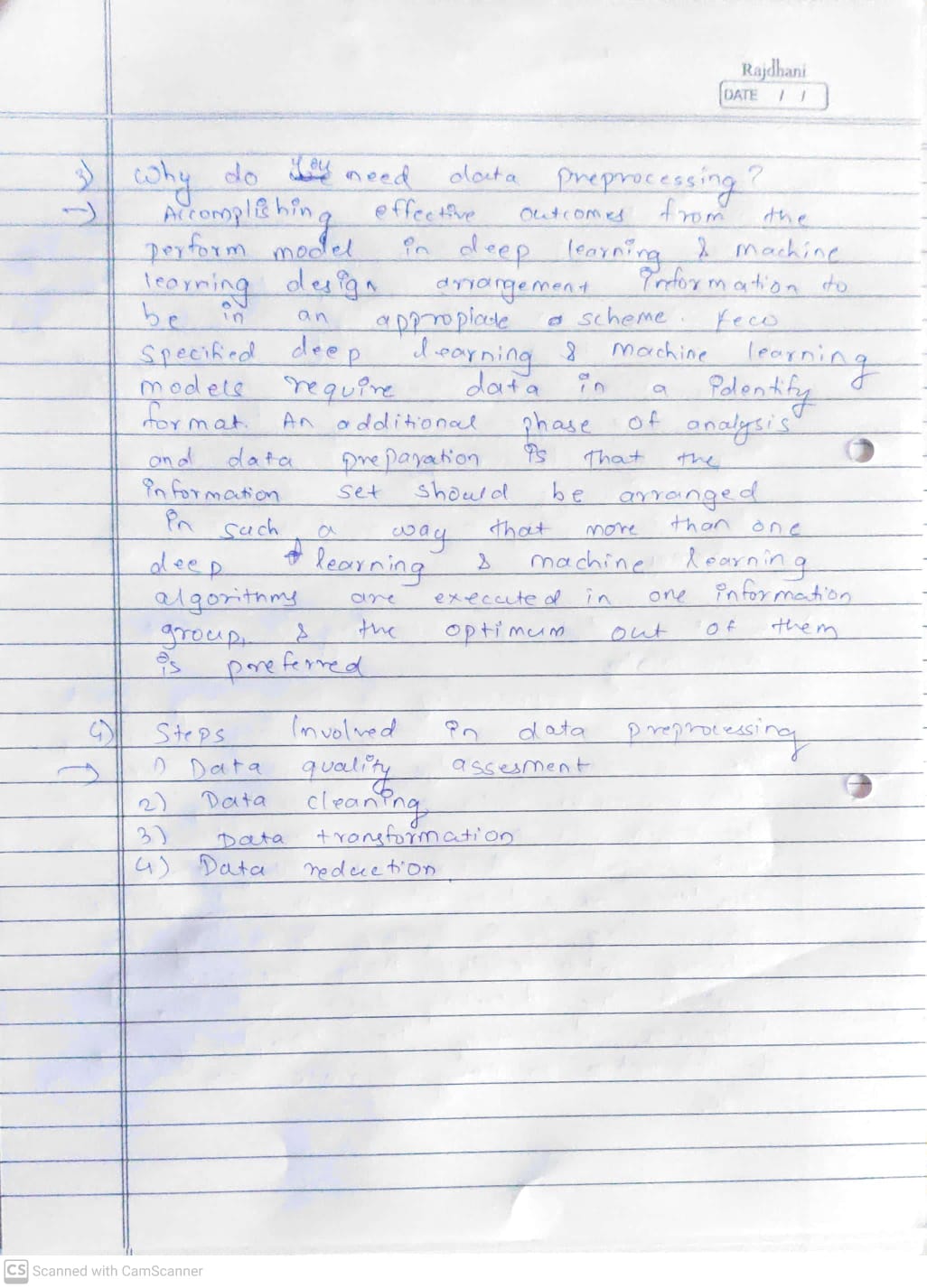
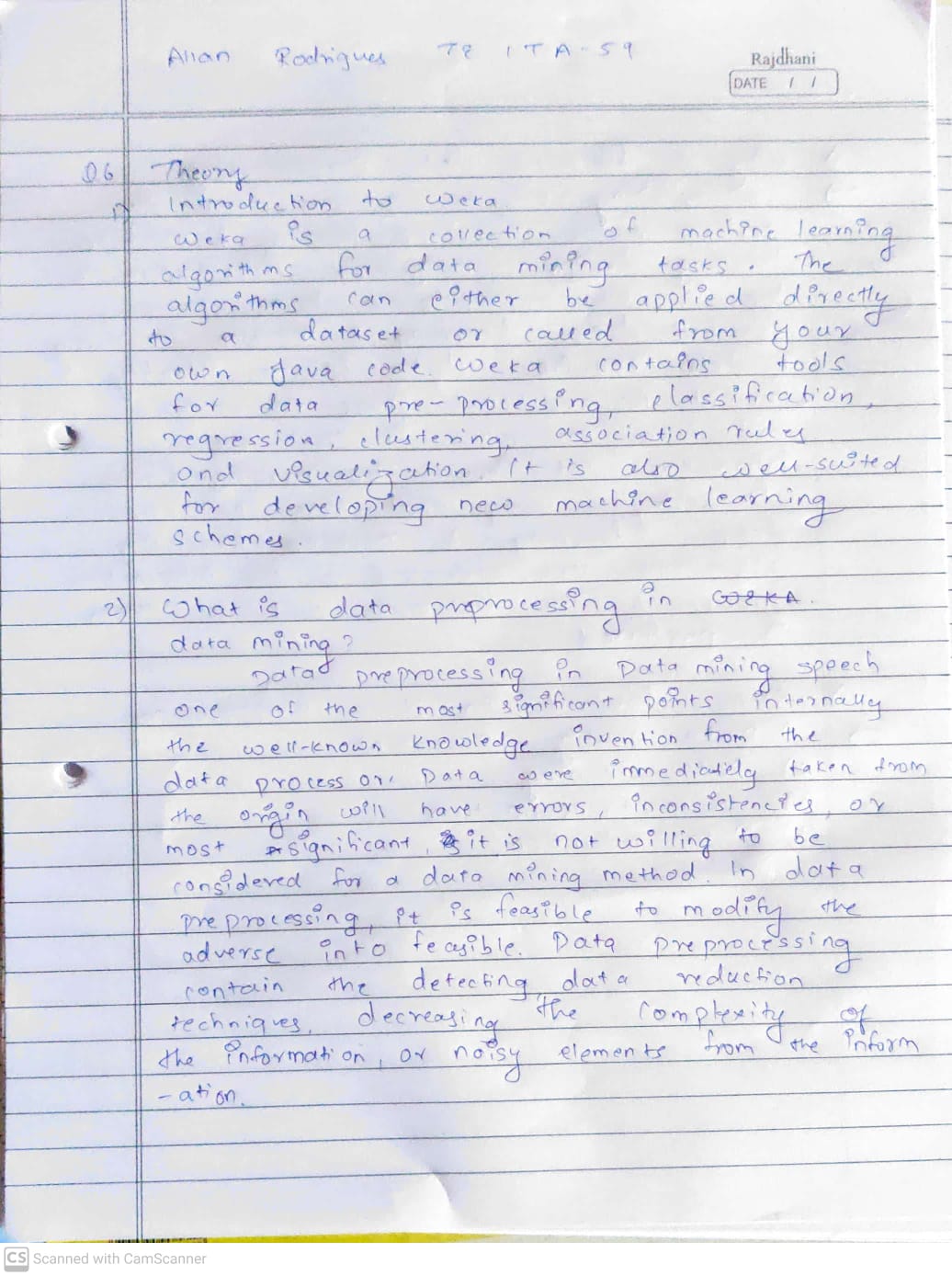
techniques.

**CO3:** Perform exploratory analysis of the data to be used for mining

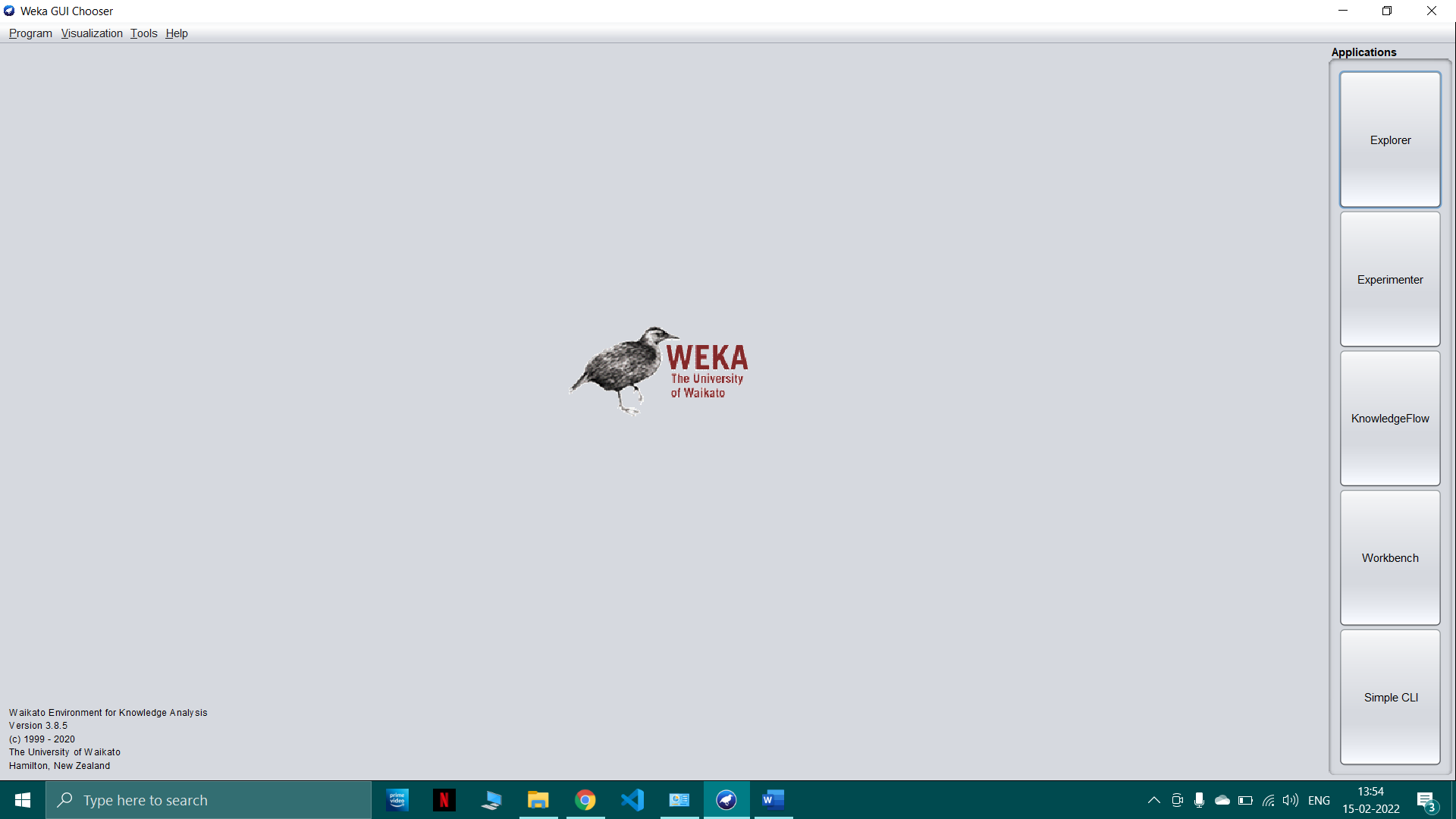
1. **Prerequisite:** Introduction to Steps in data preprocessing.
2. **Requirements:** Personal Computer, Windows XP operating system/Windows 7, Internet

Connection, Microsoft Word, WEKA tool.

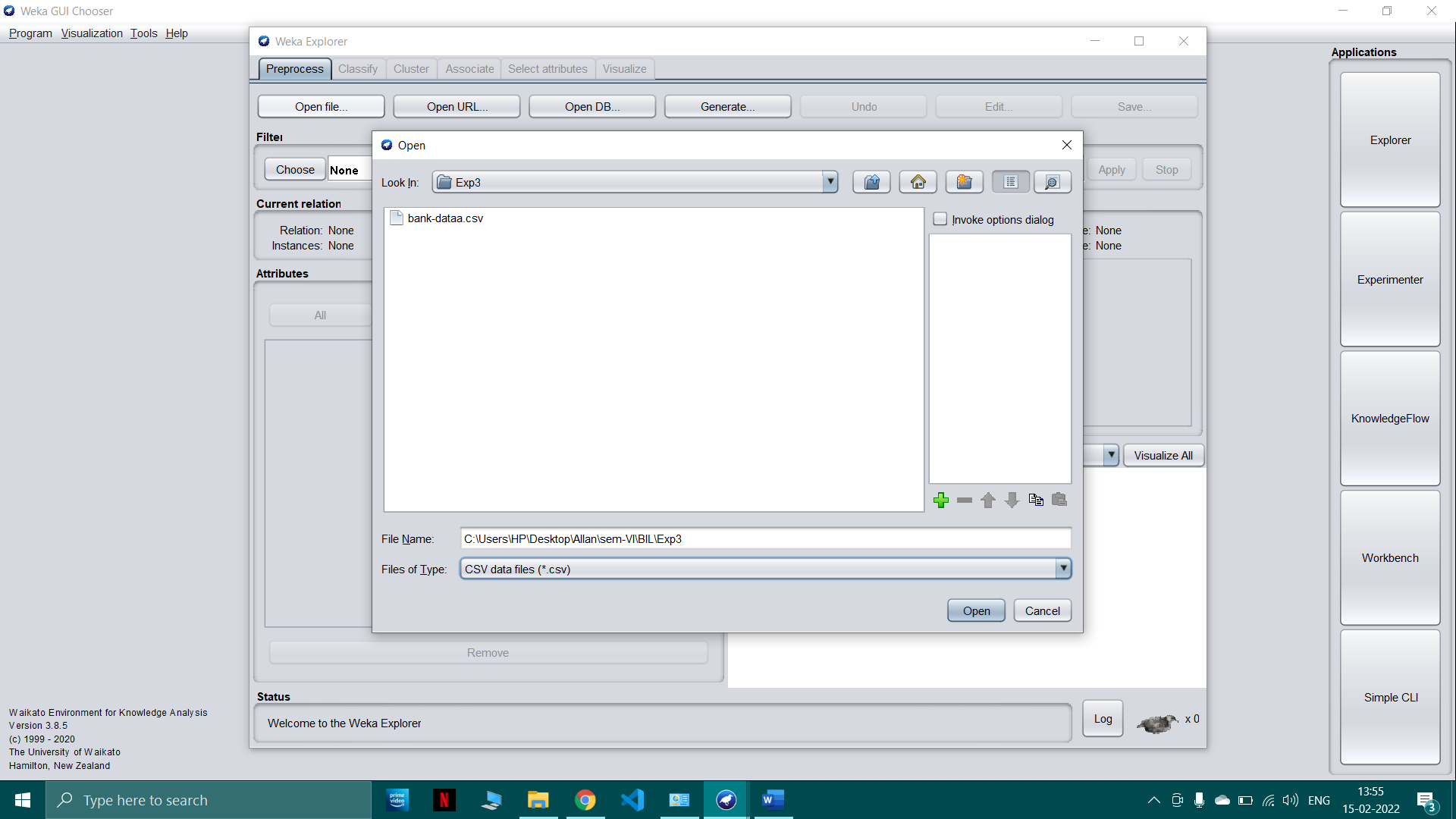
1. **Theory:**
2. Introduction to Weka.
3. What is Data Preprocessing in data Mining?
4. Why do you need Preprocessing?
5. Steps involved in Data Preprocessing.

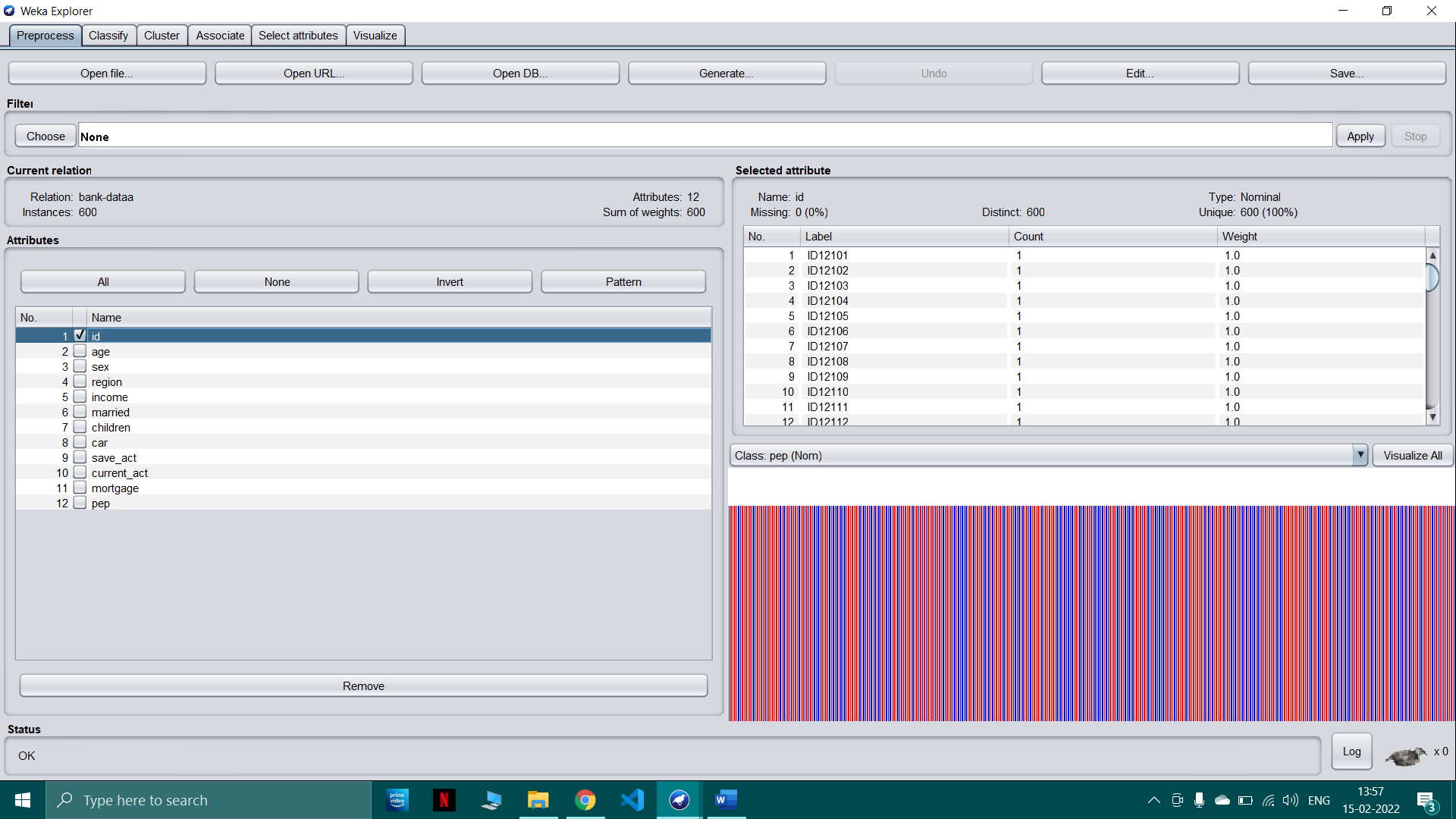


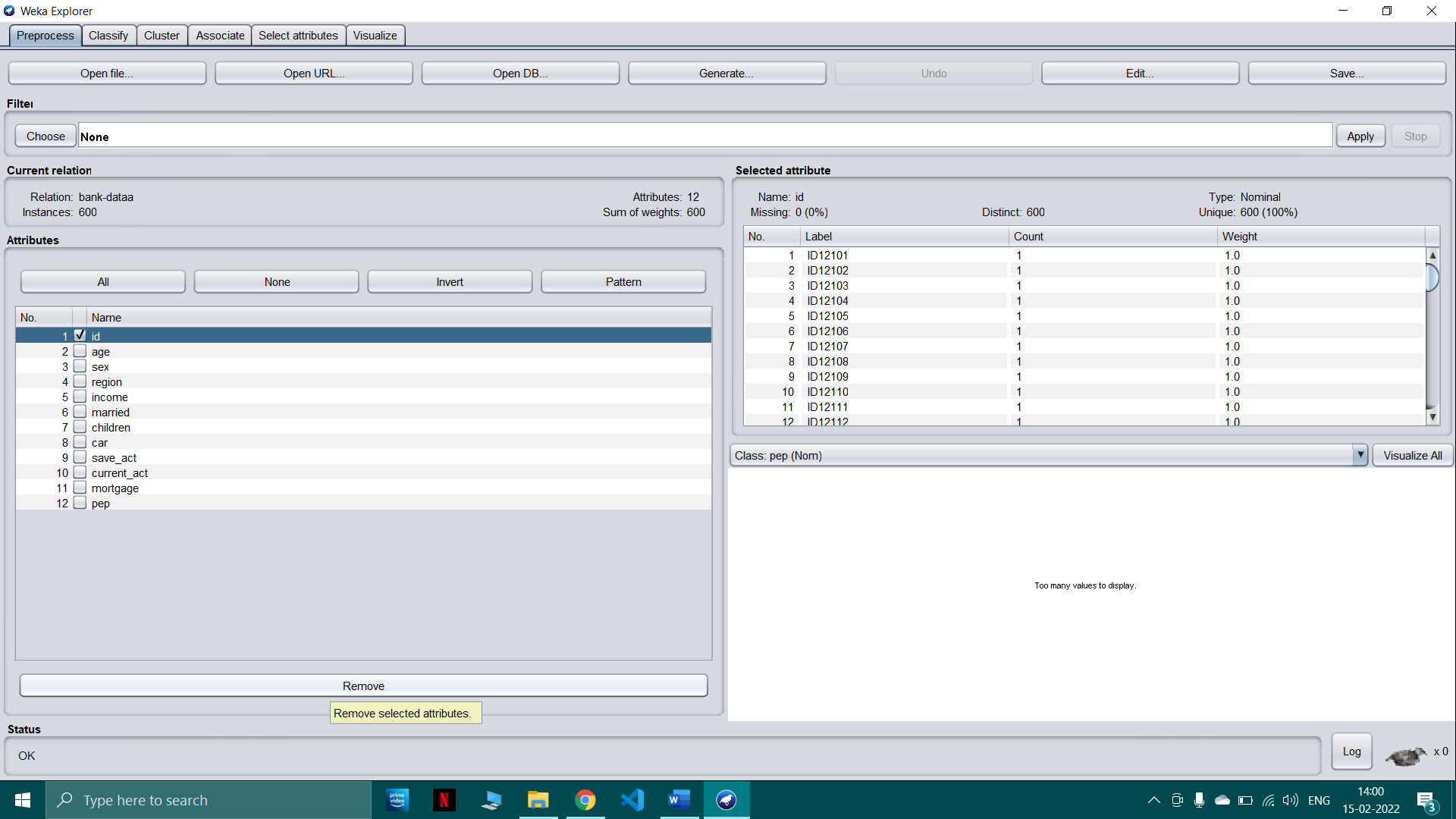
1. **Laboratory Exercise:** Implementation of Data Preprocessing Process in WEKA and Take Printout of Implementation along with coding and snapshot.

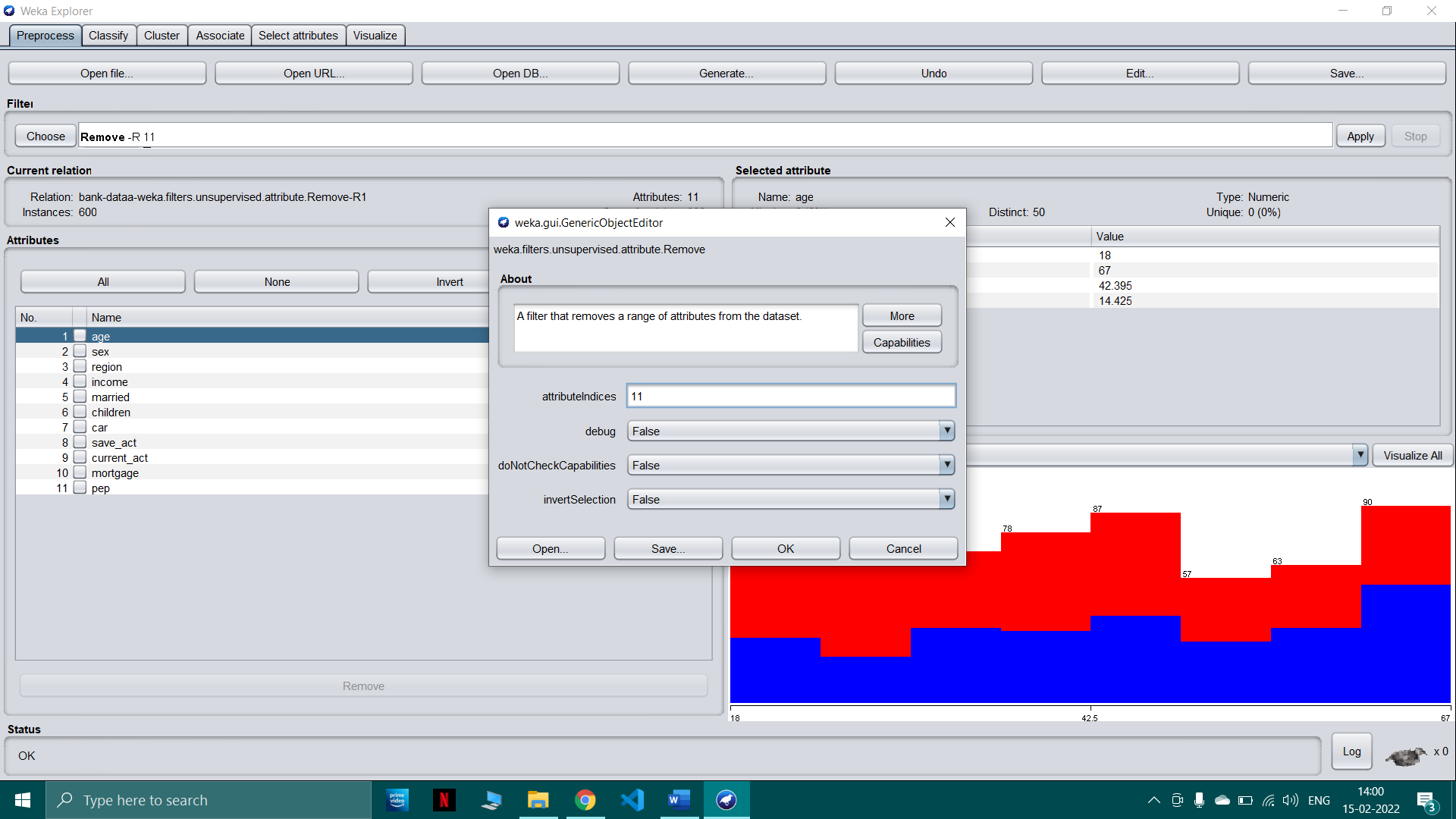


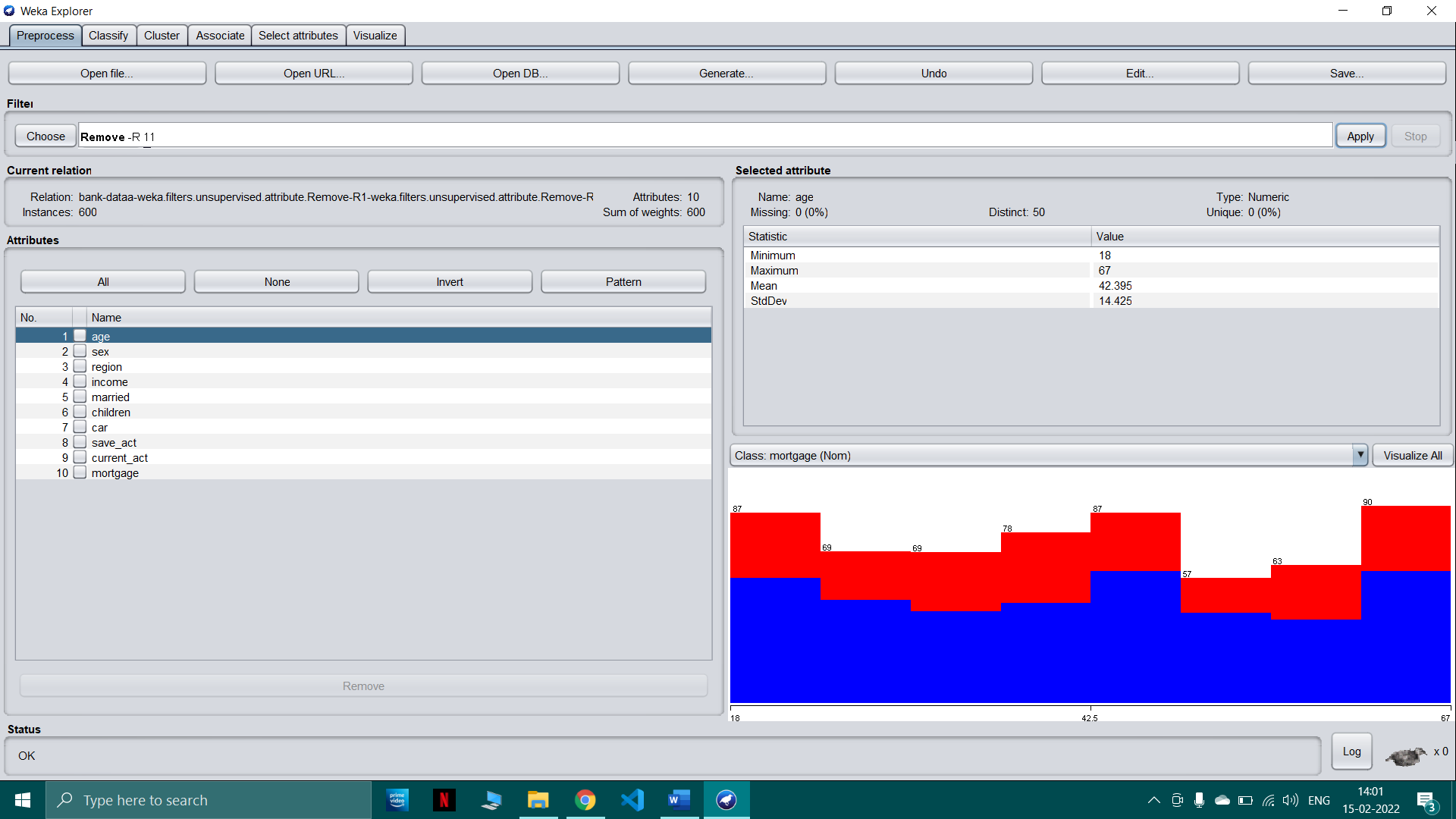


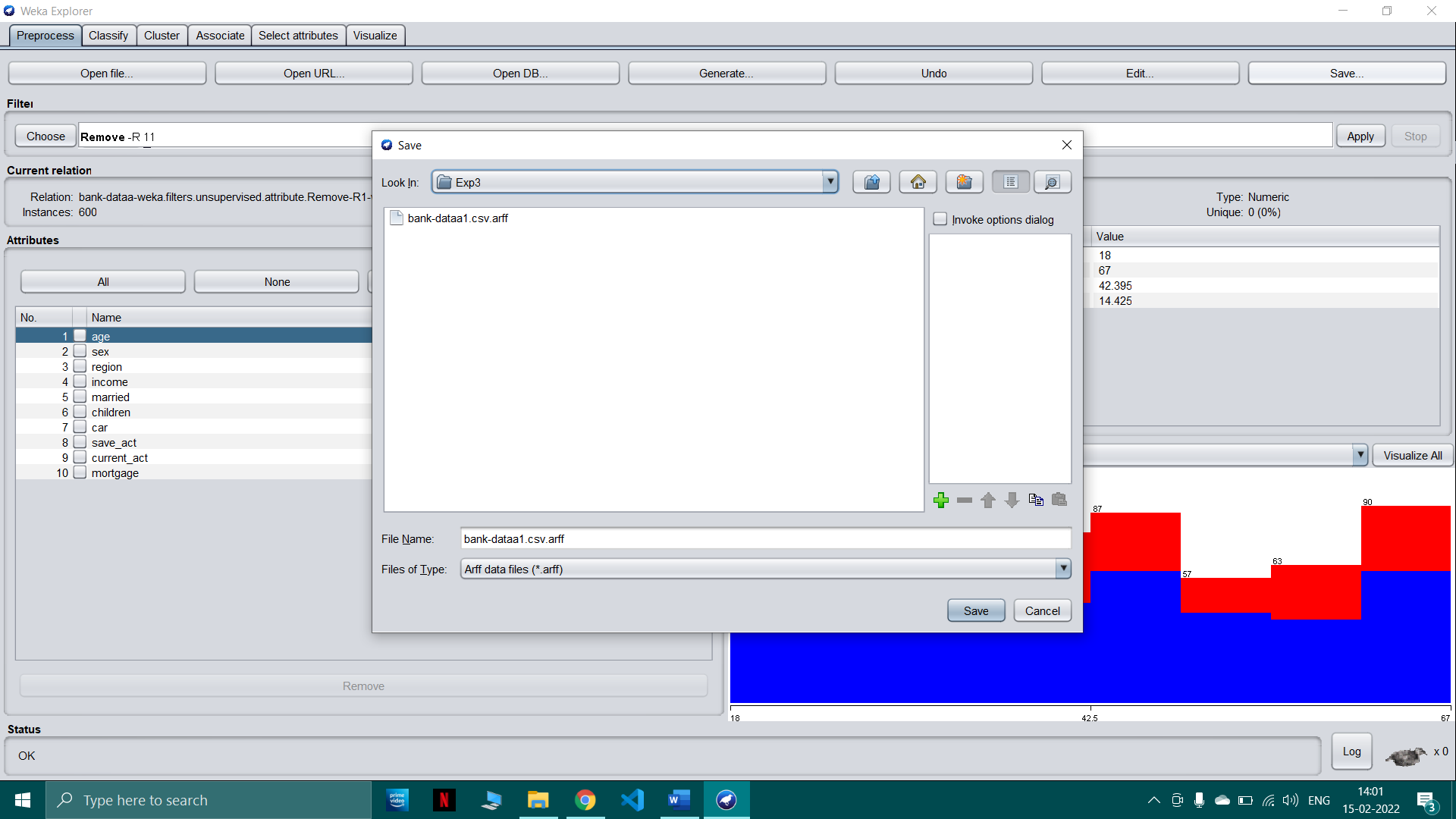


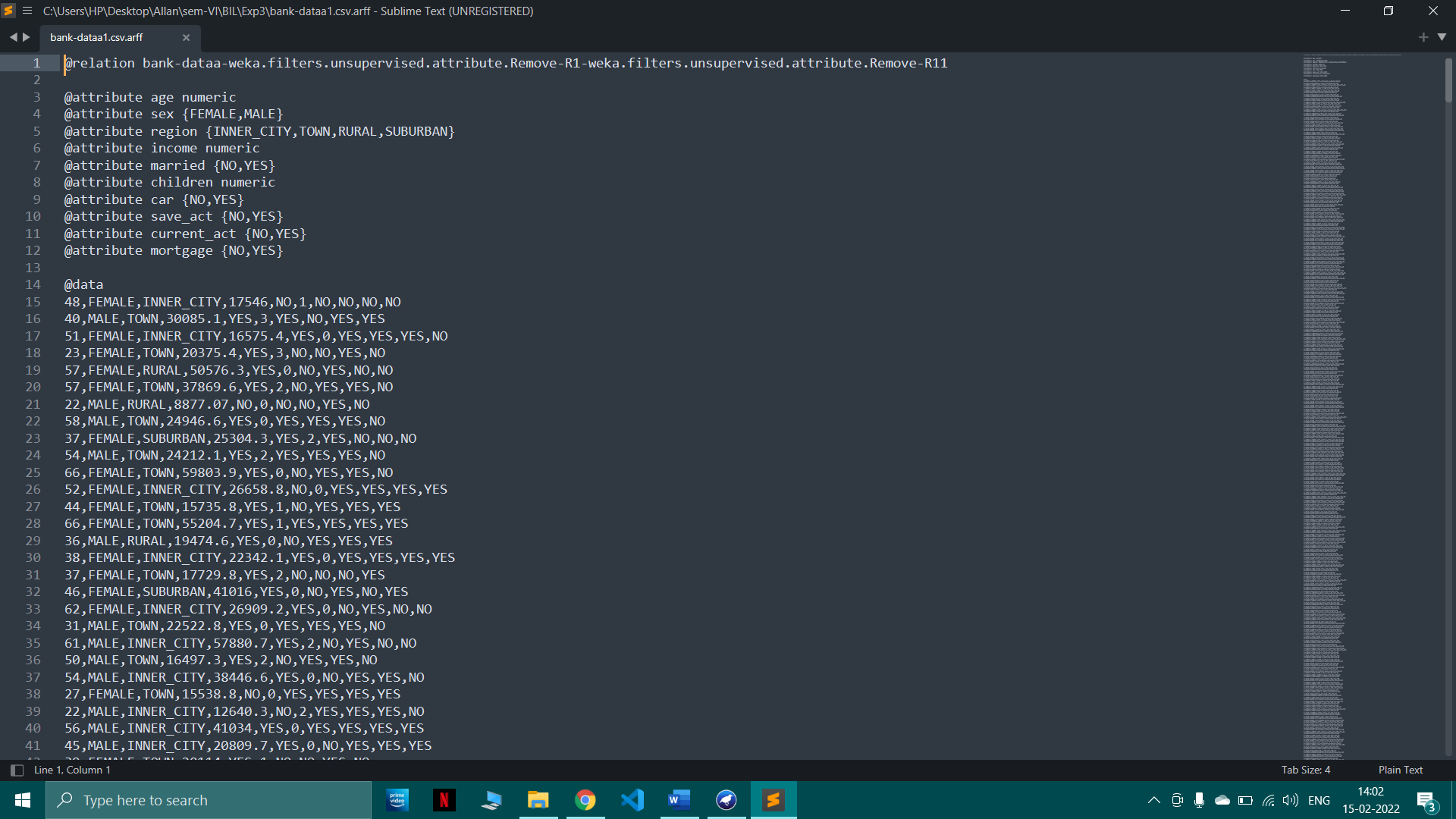




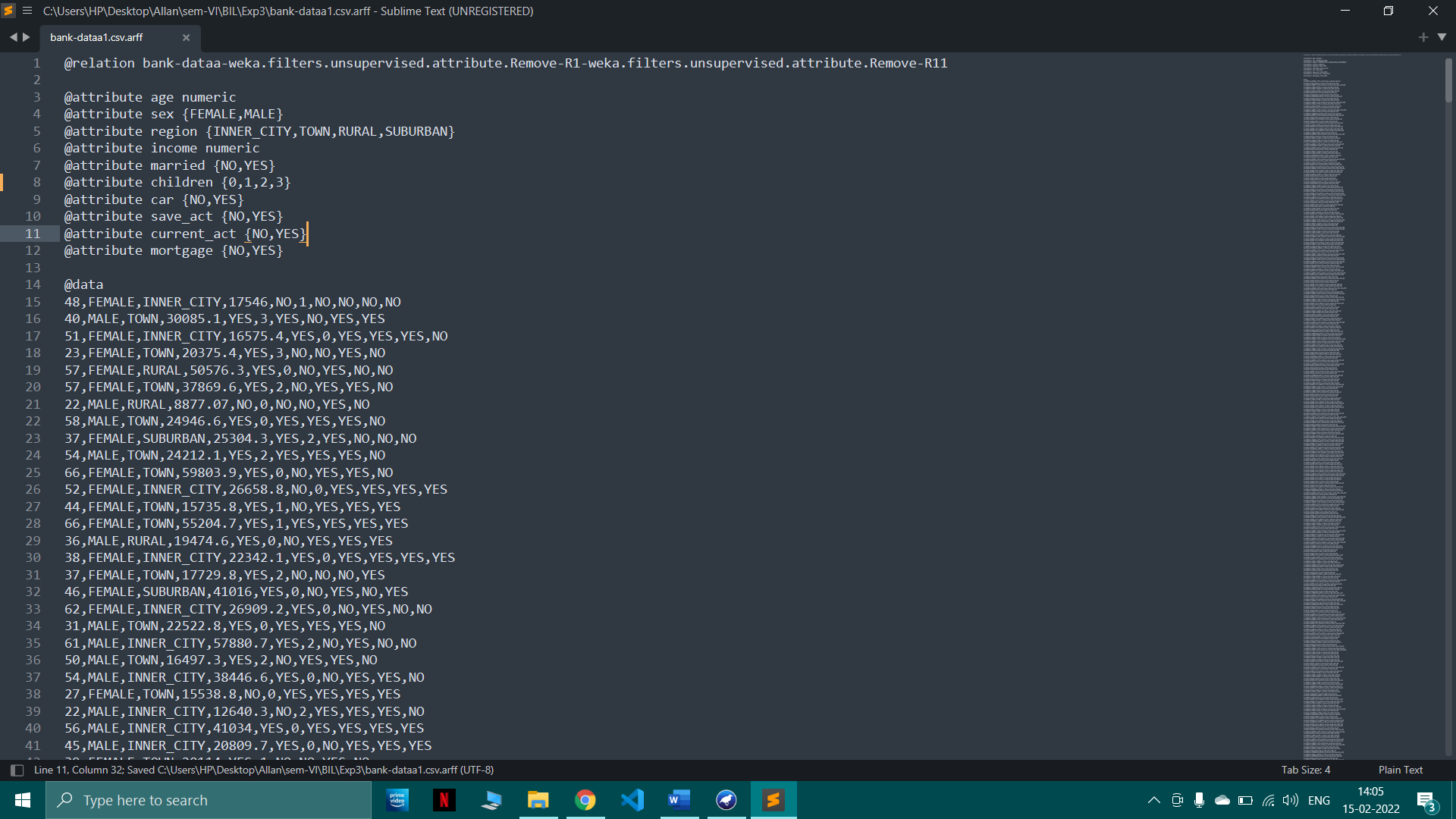


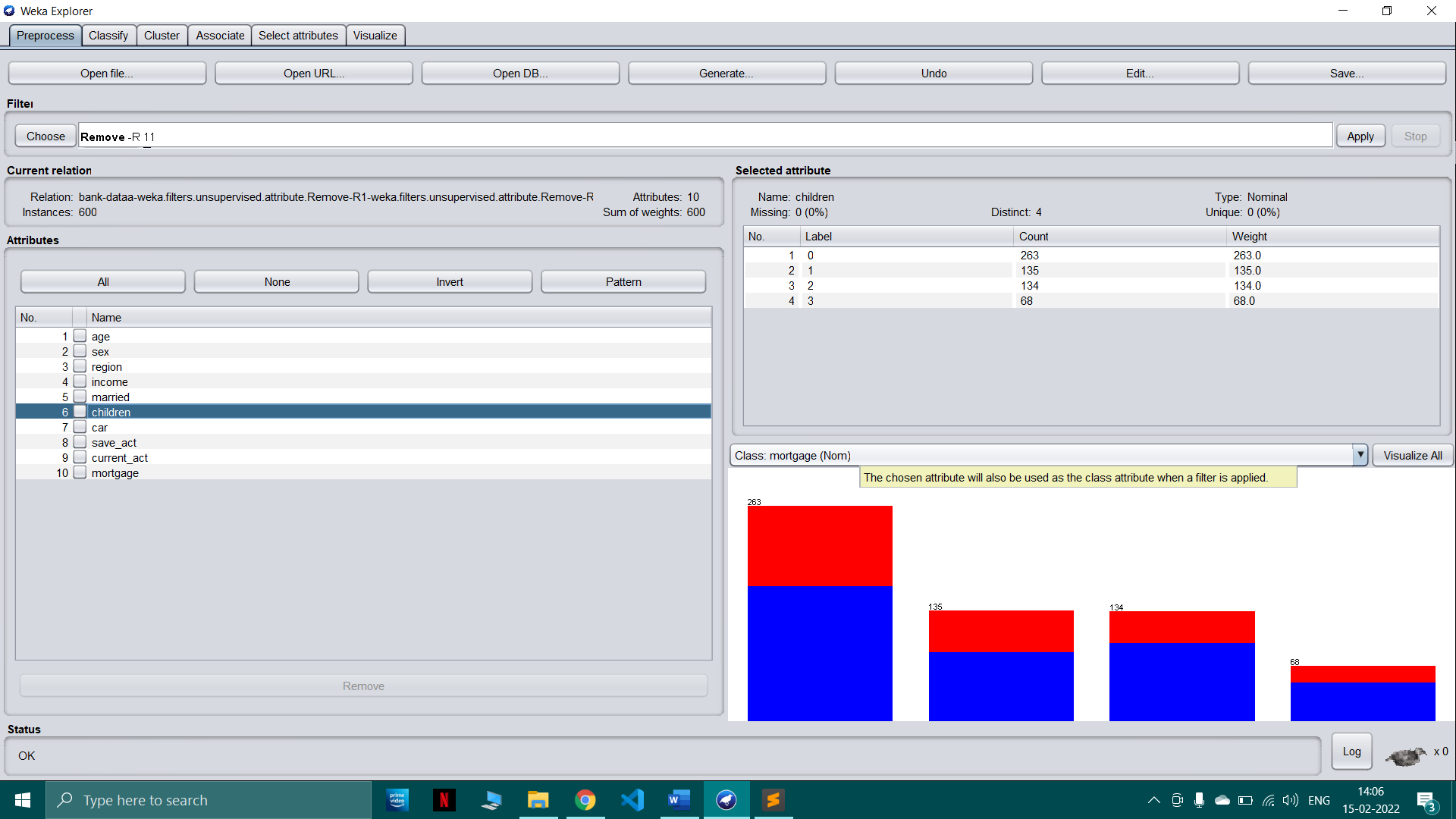


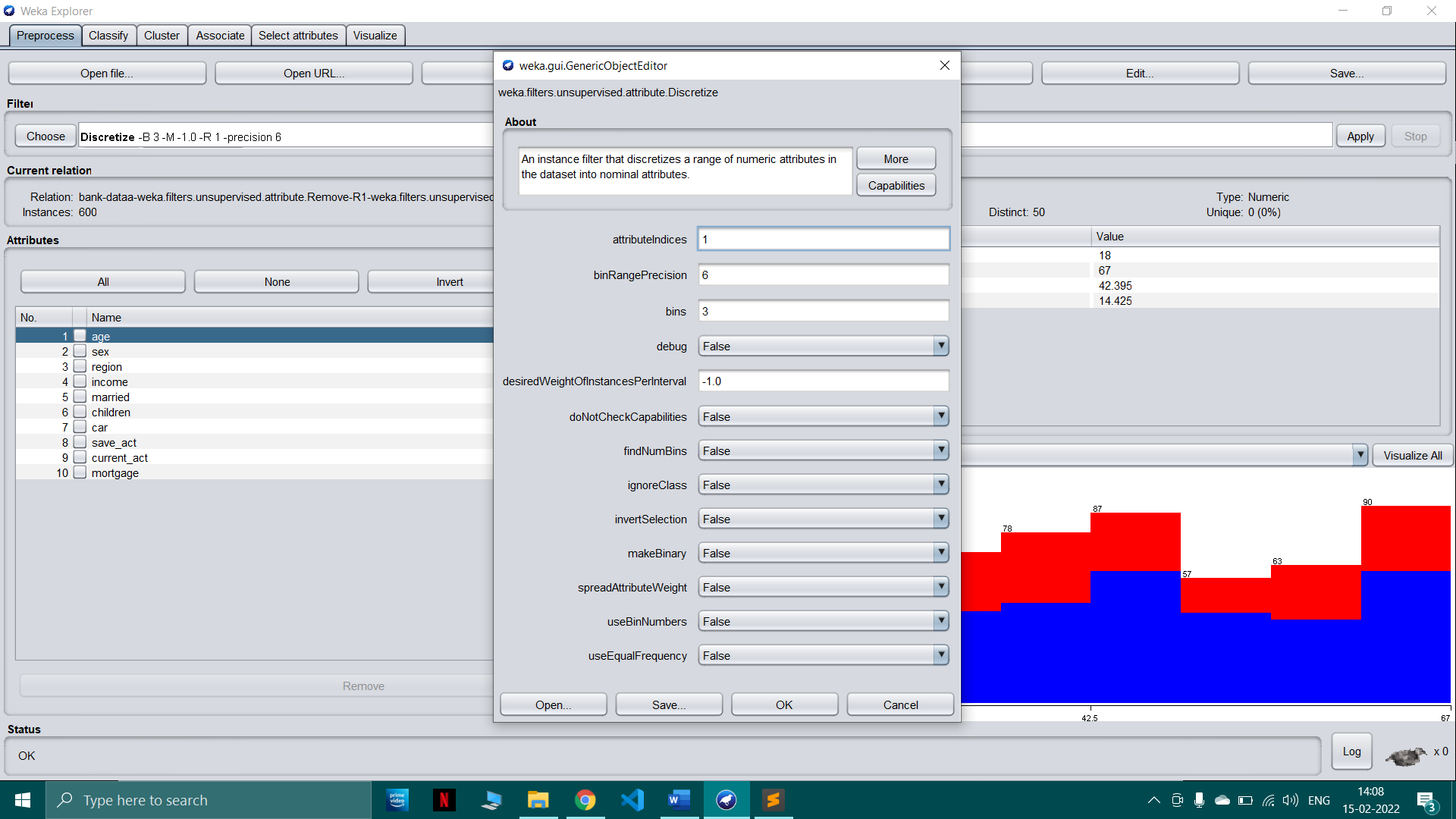


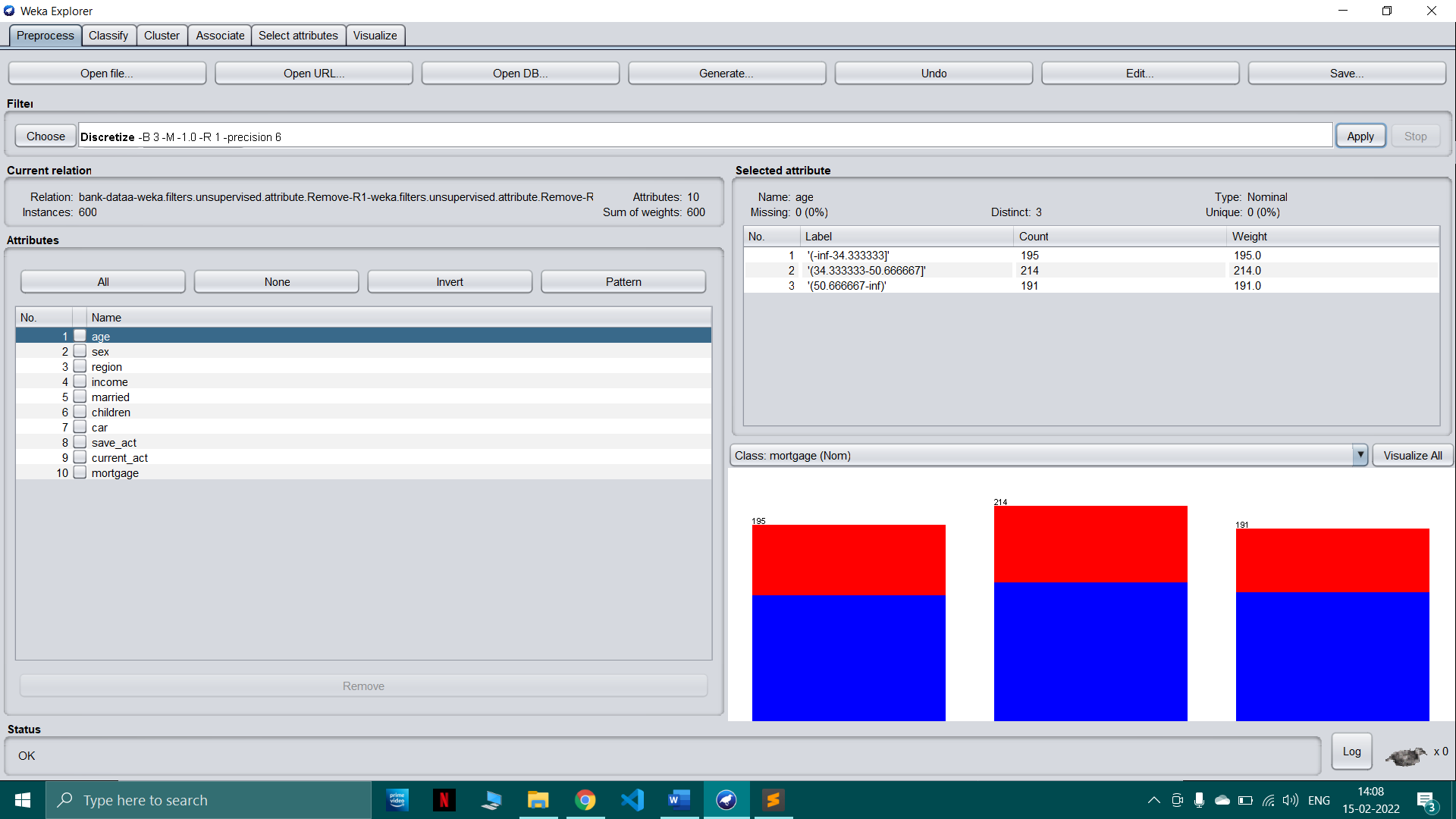


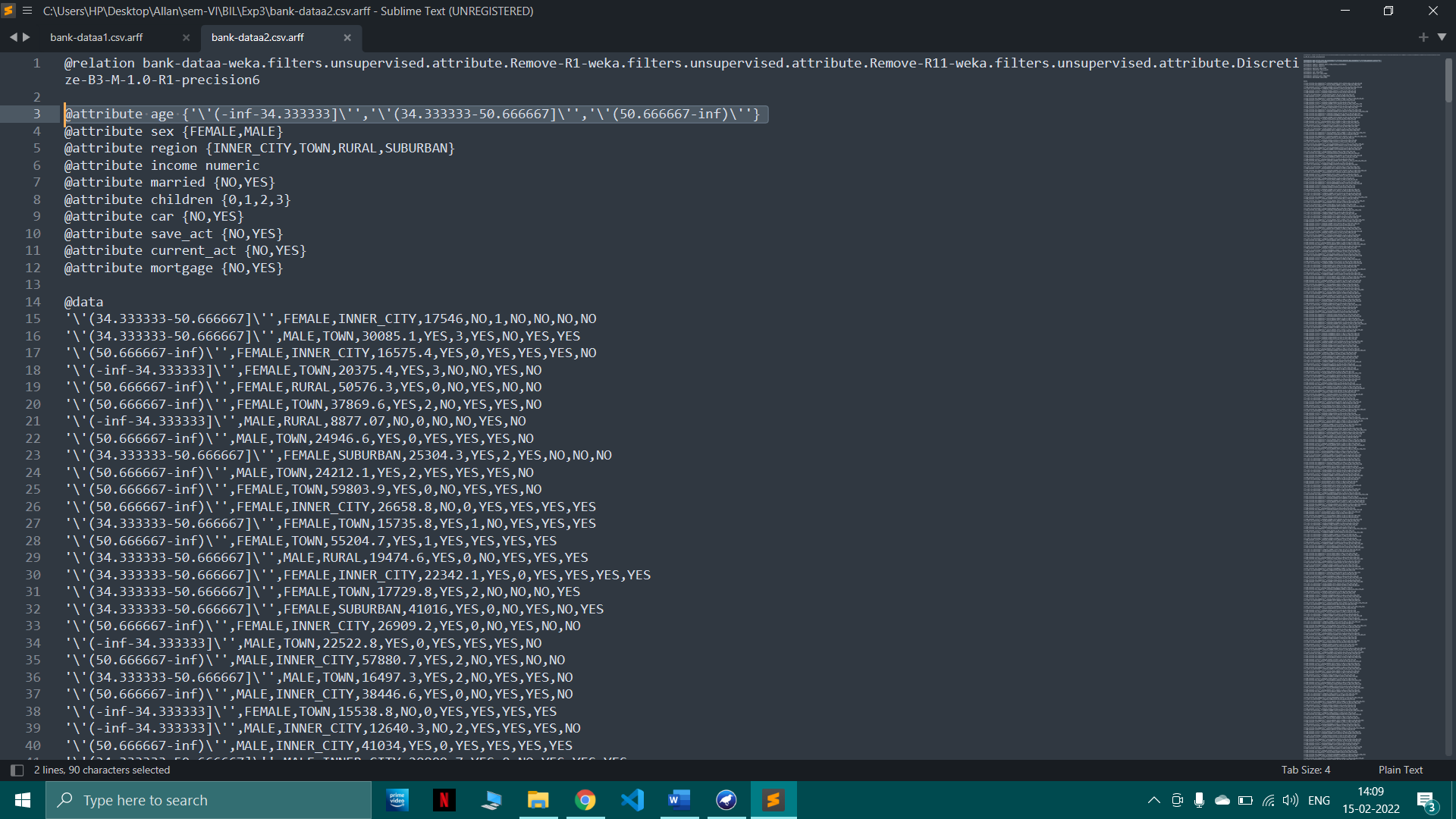
Discretization

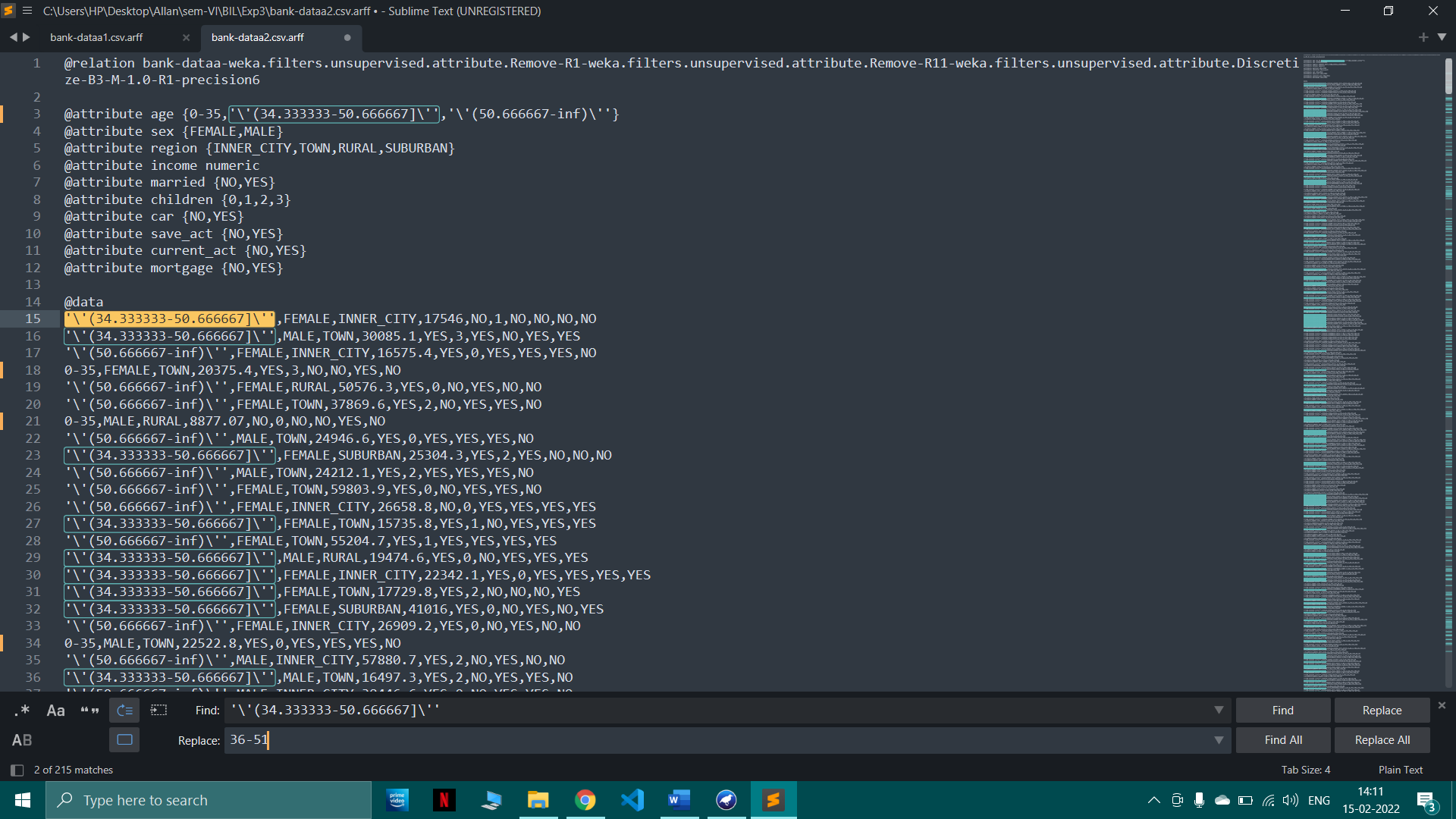


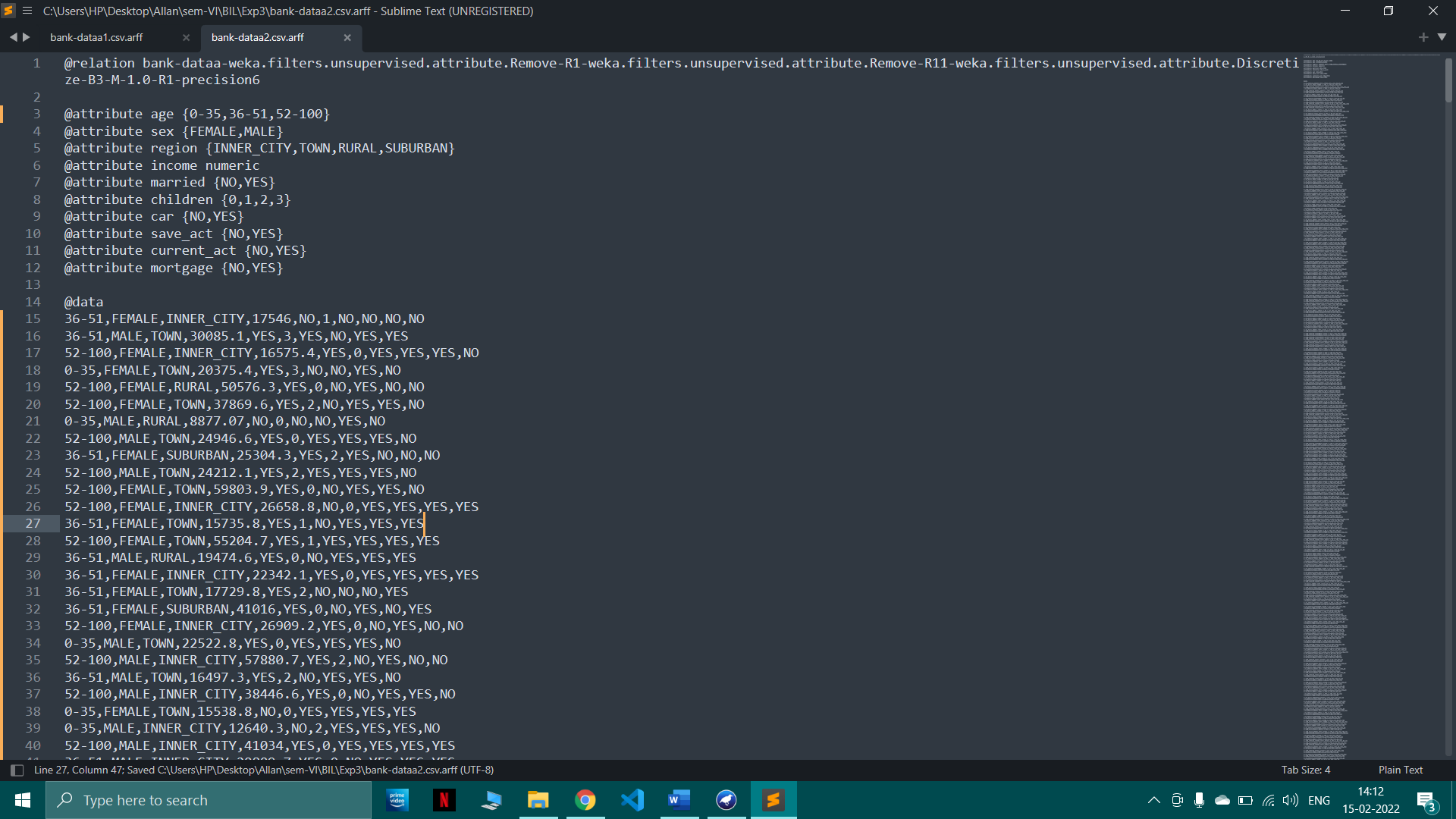


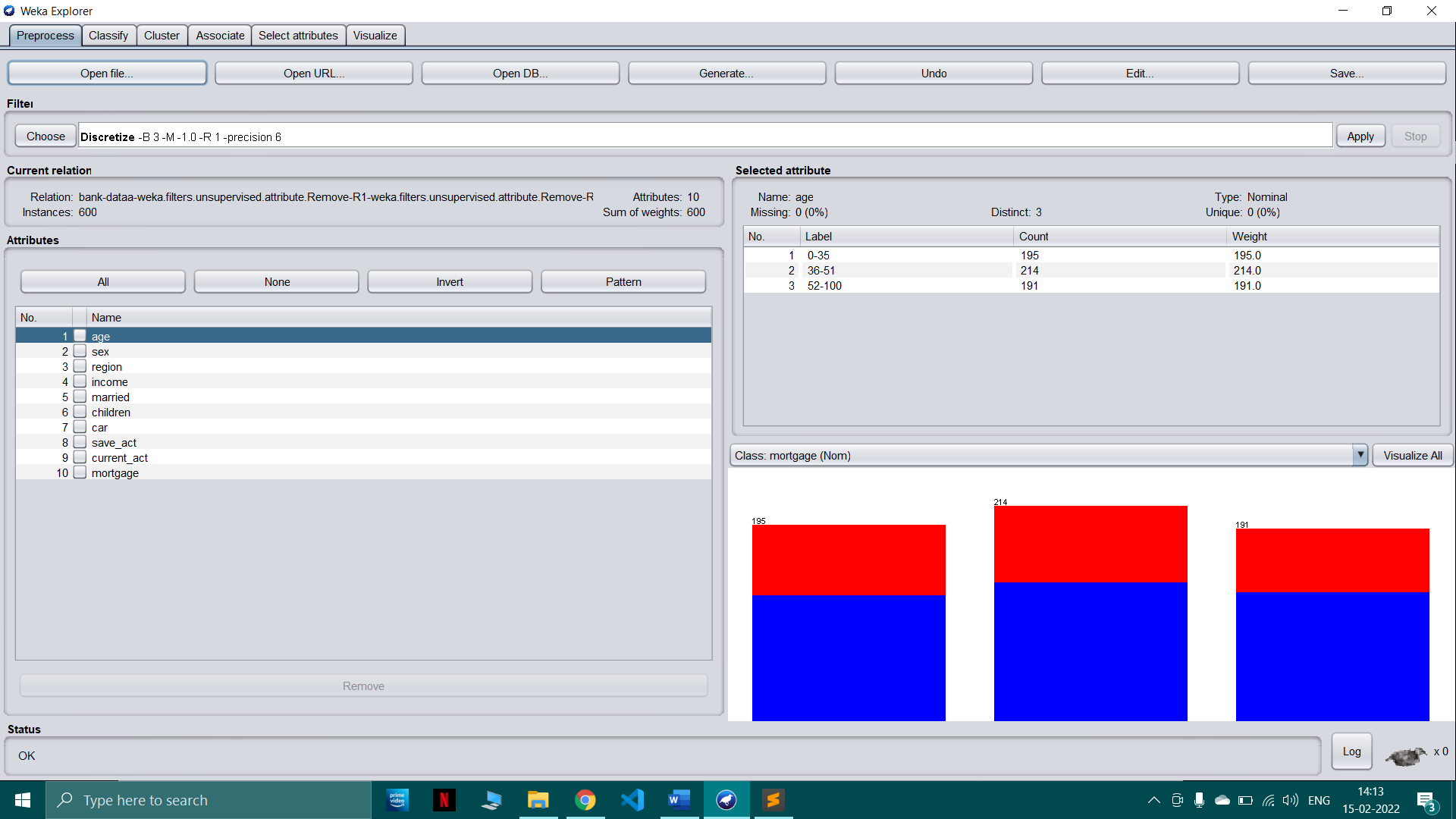




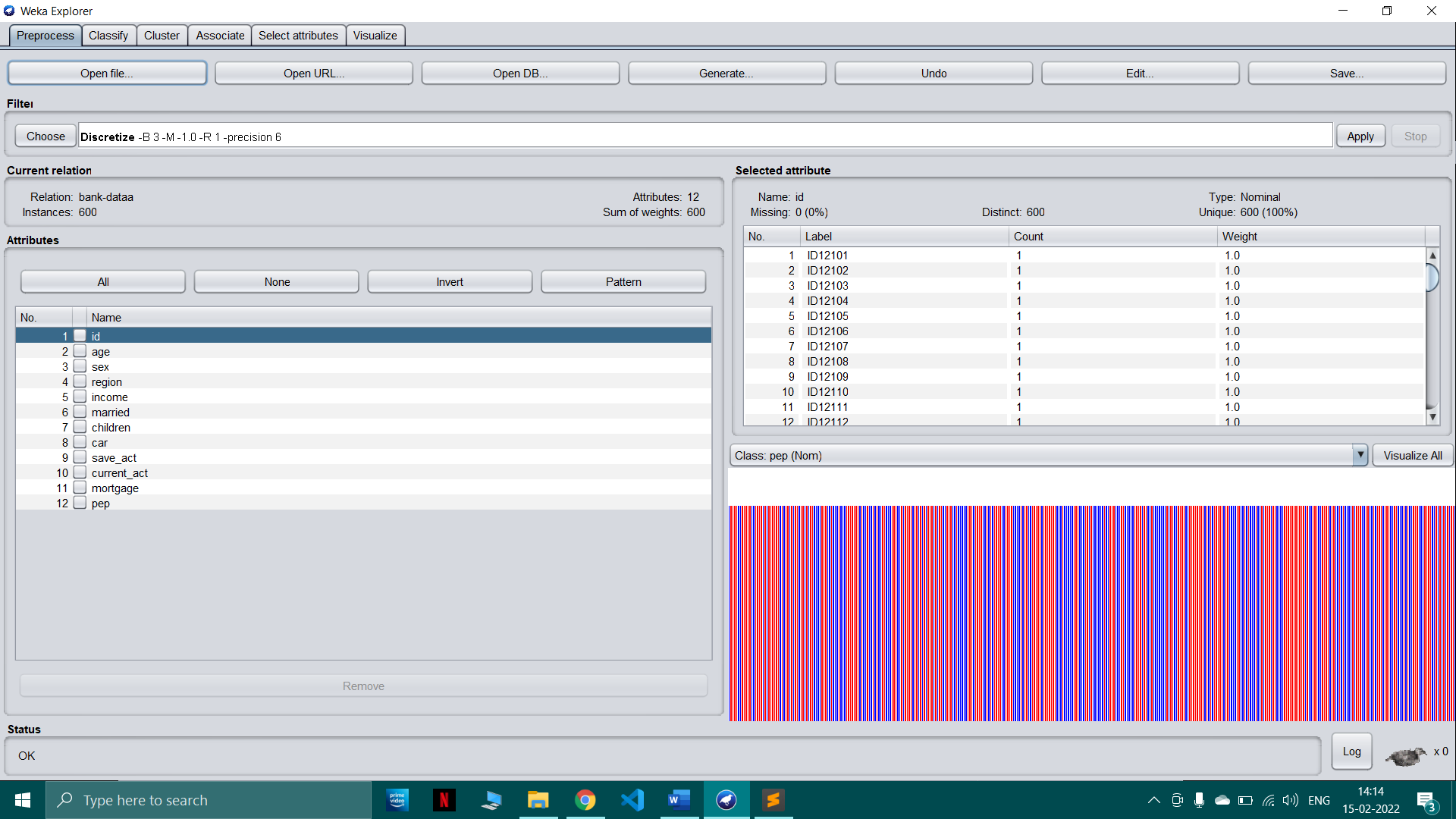


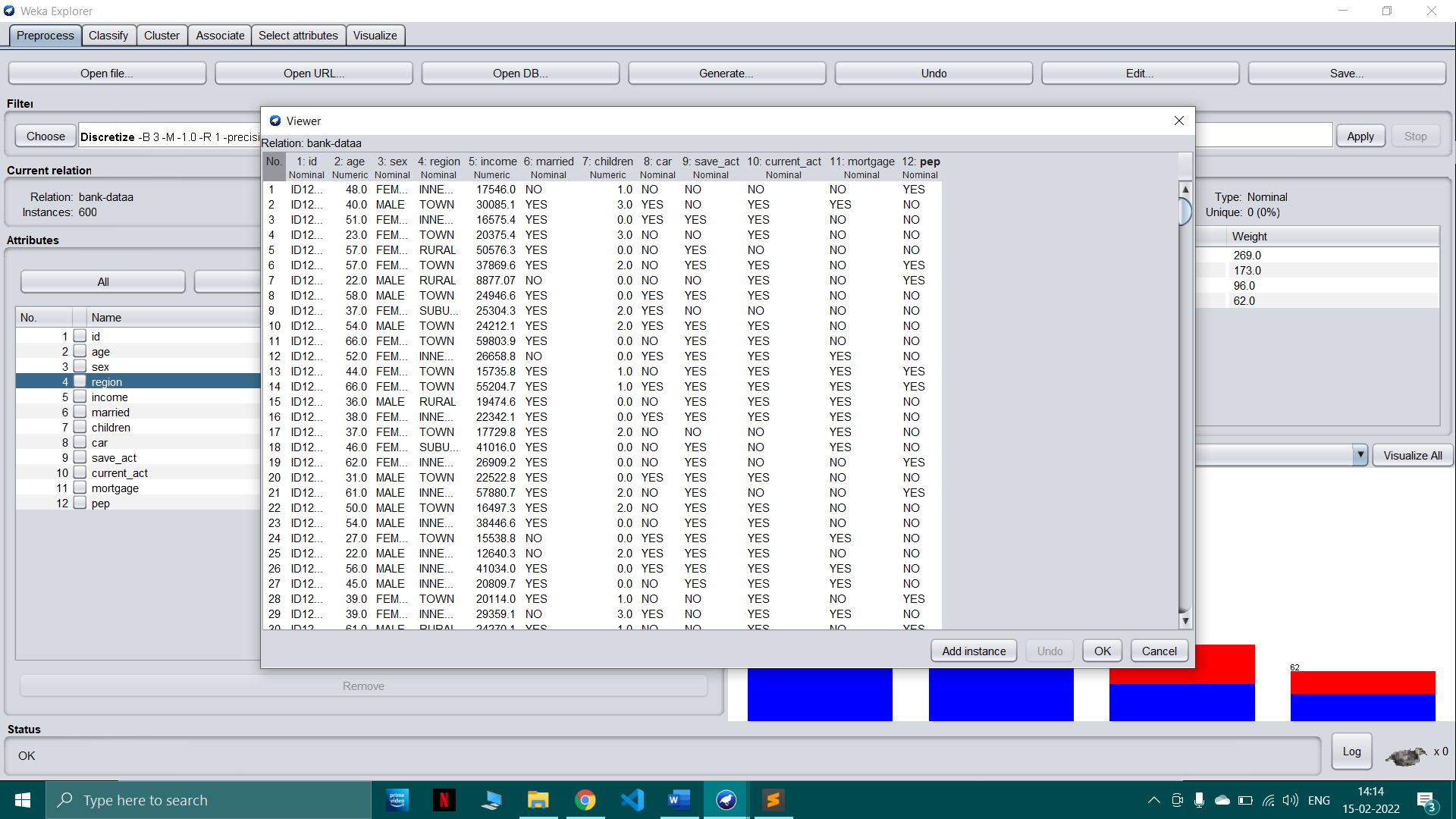


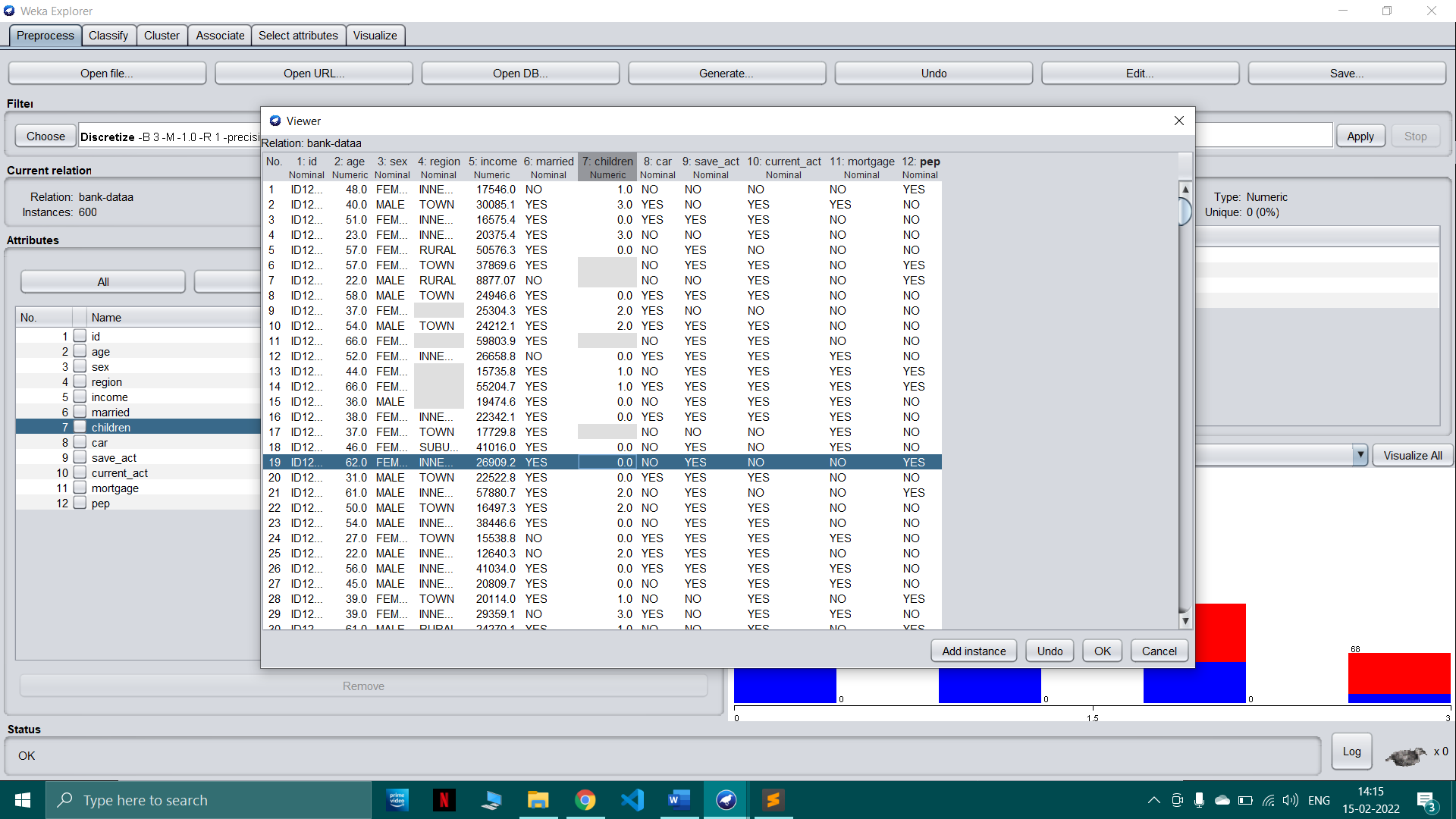


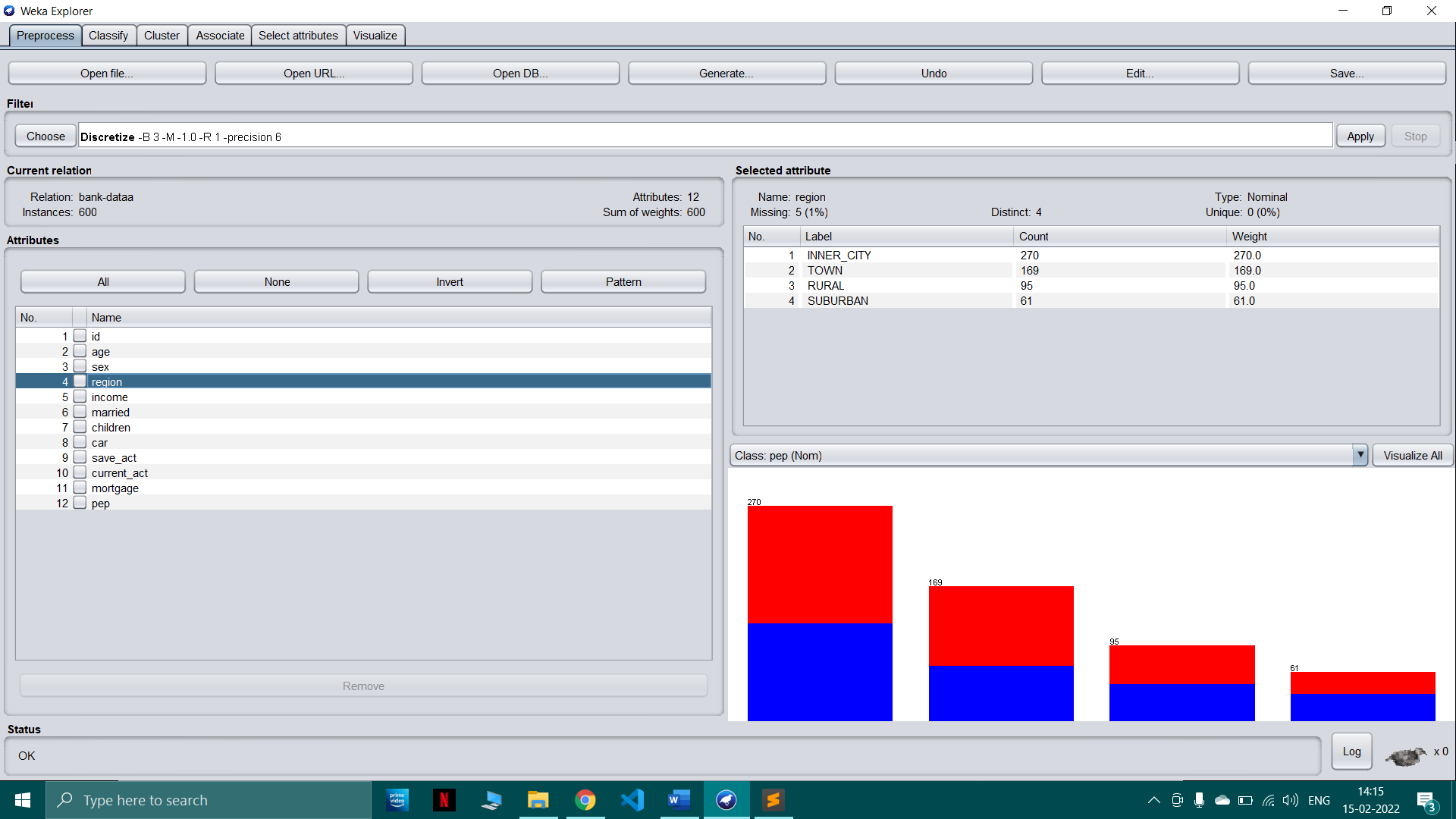


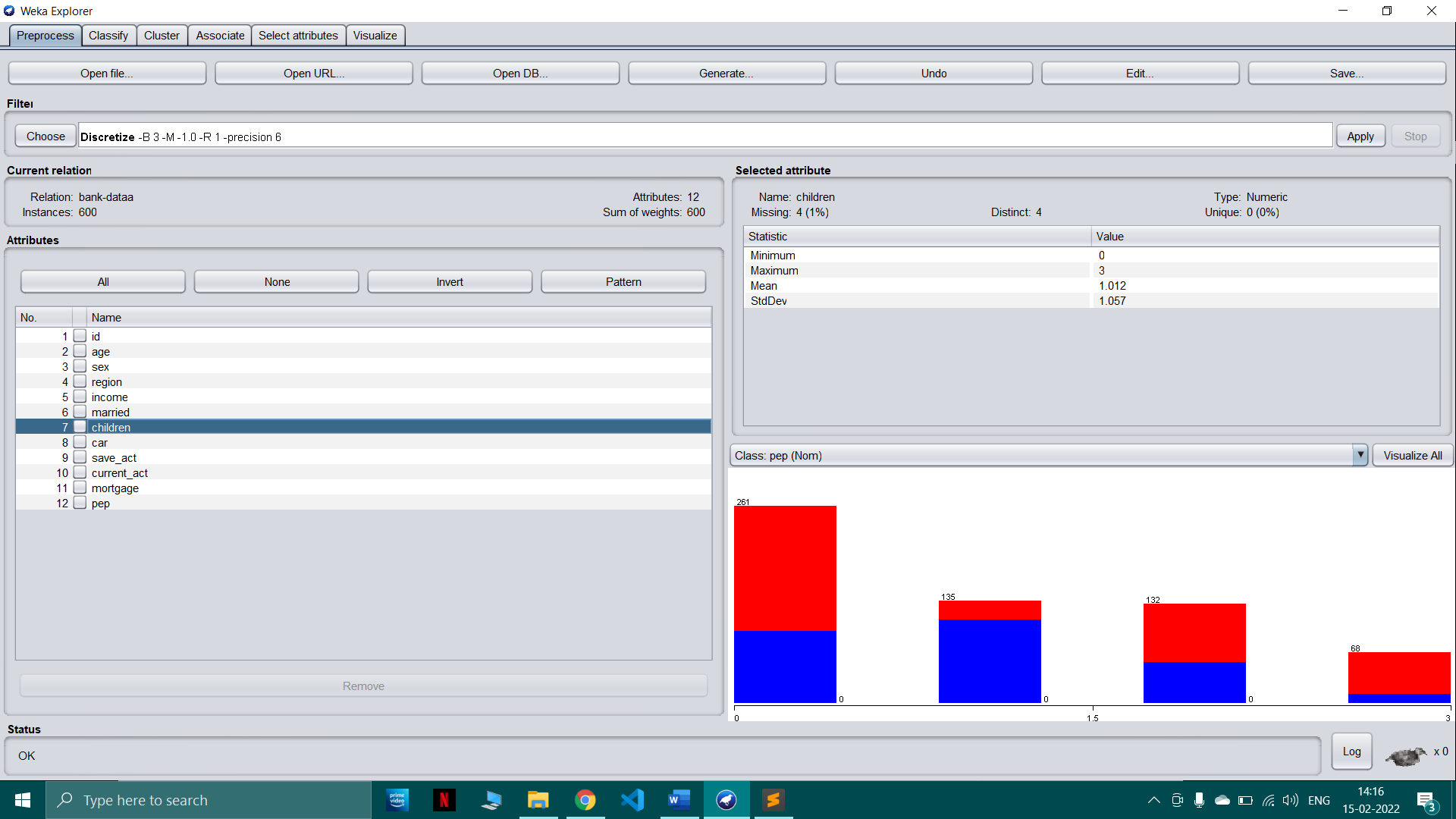
Missing values



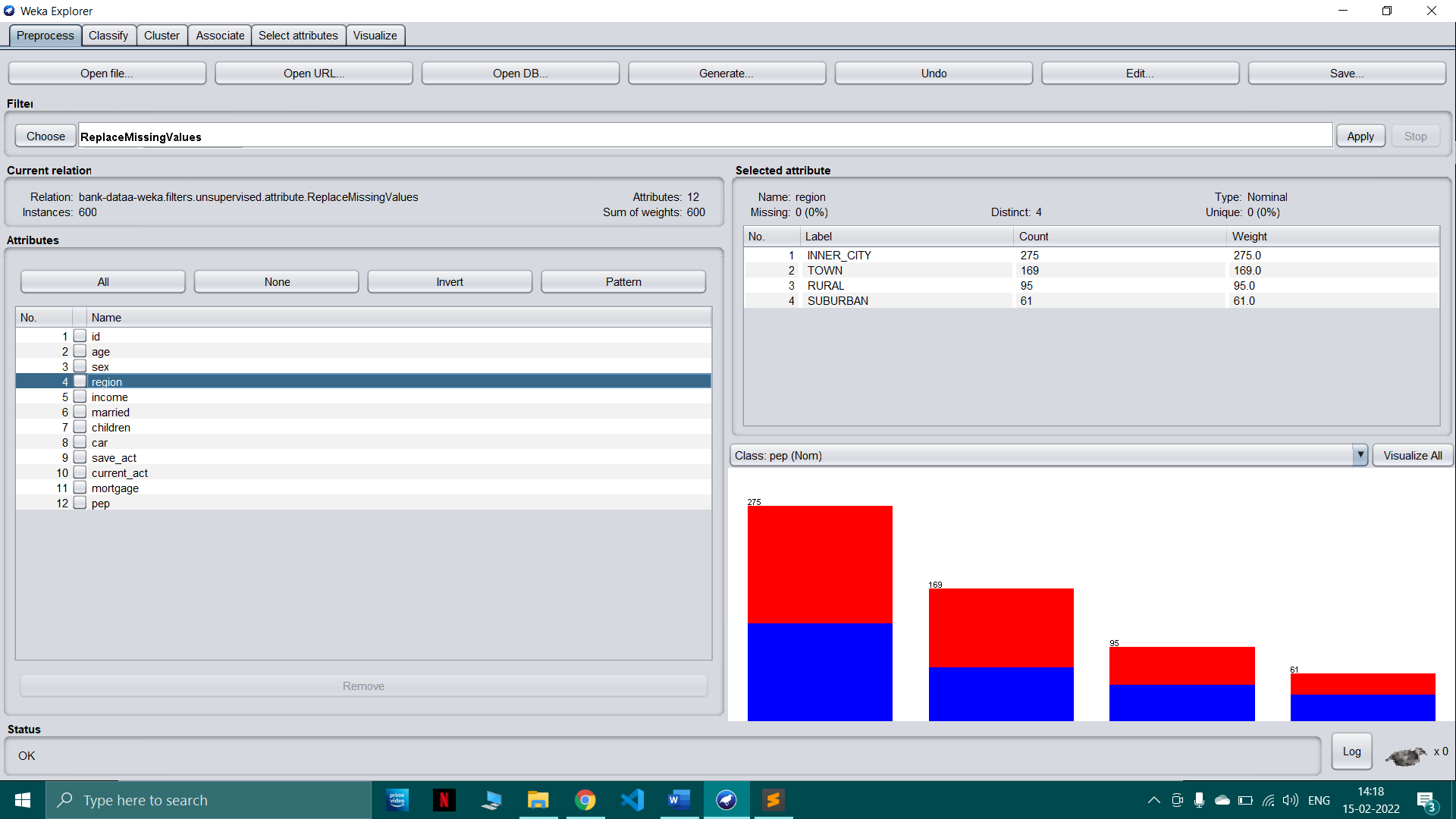


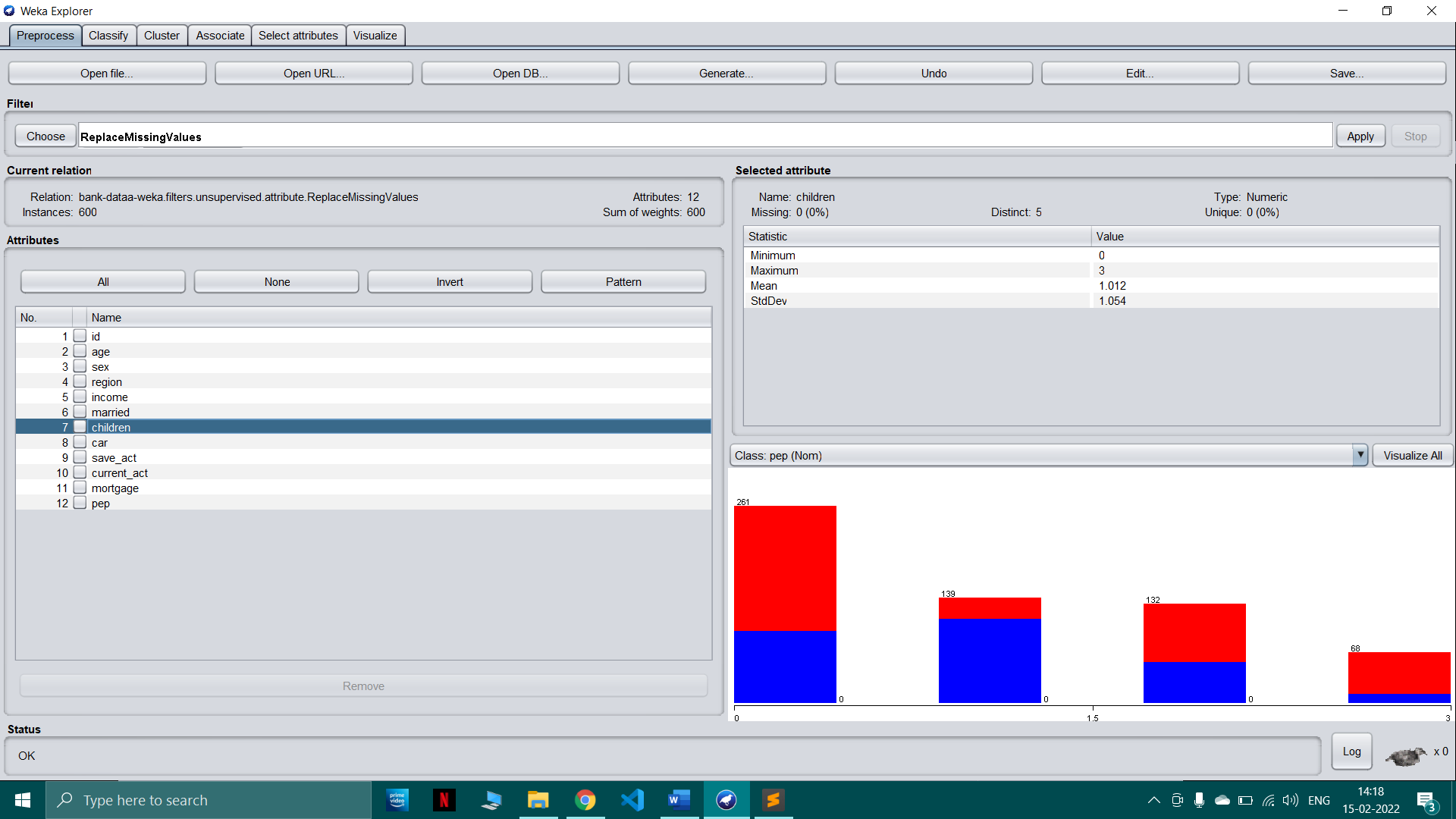


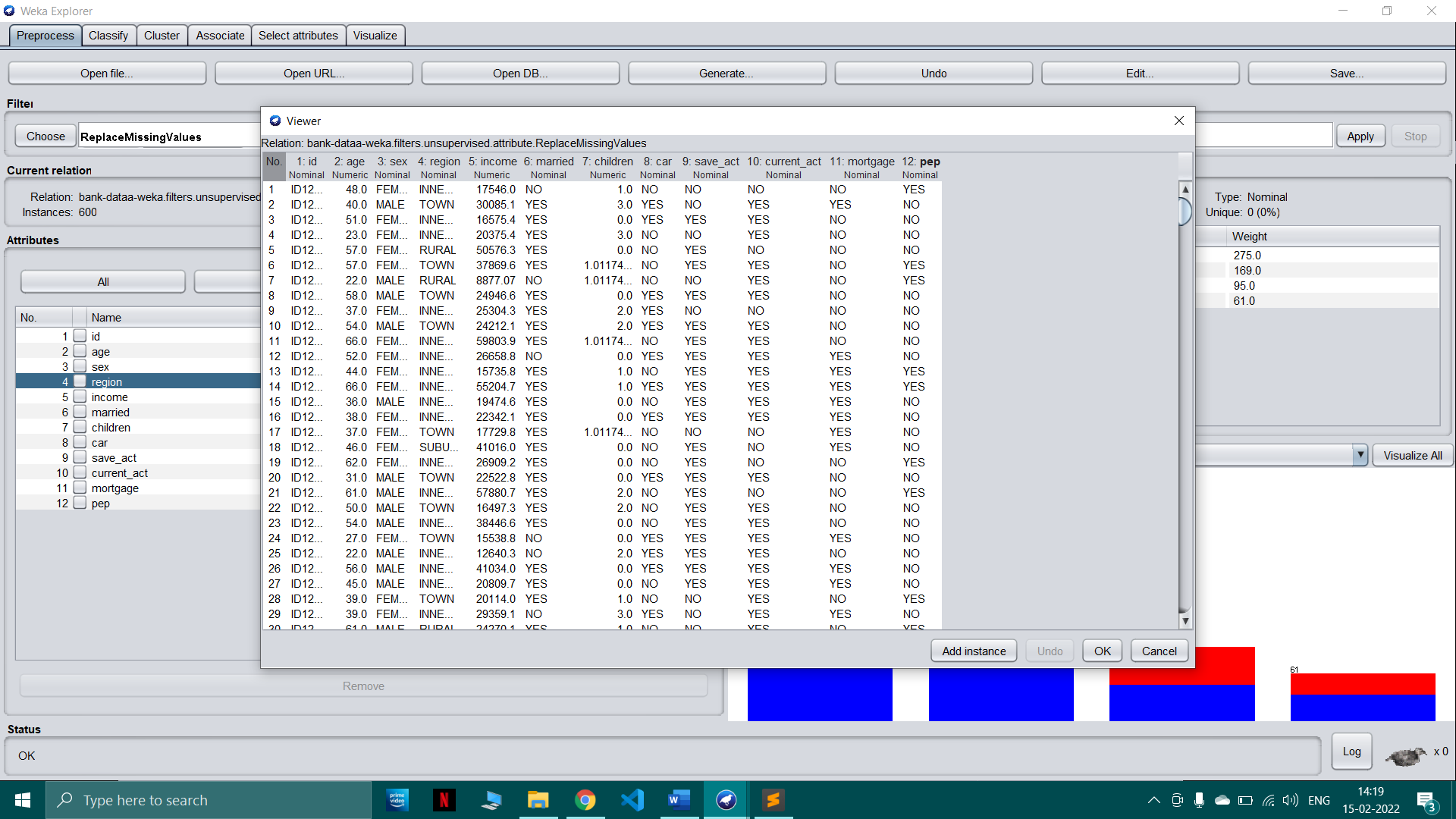












1. **Post-Experiments Exercise**

**A Questions:**

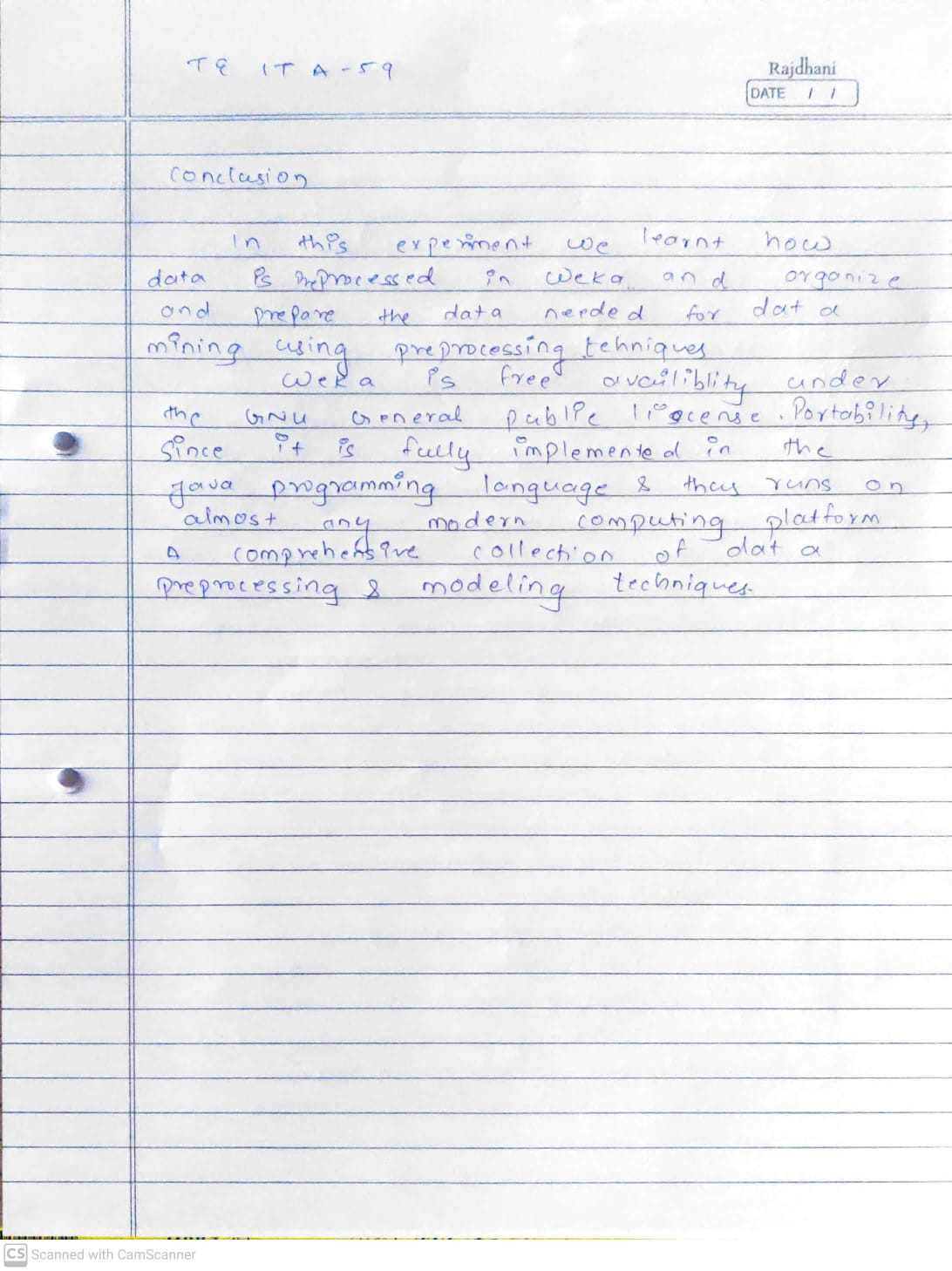
In form of MCQ type test

**B Conclusion:**

1. Summary of Experiment

2. Importance of Experiment

3. Application of Experiment



1. **Reference:** Data Mining: Concept & Techniques, 3rd Edition, Jiawei Han, Micheline Kamber, Jian Pei, Elsevier.