**A**

***Project Report On***

**Implementation of Association Rule Mining**

****

Submitted for partial fulfillment of requirement for the award of degree

Of

***BACHELOR OF TECHNOLOGY***

(COMPUTER SCIENCE AND ENGINEERING)

Of

***Assam Down Town University***

***Session 2018-22***

***Submitted by : -***

Name : - Chandra Shekhar Yadav

Roll no : - 1814010

Registration no : - (ADTU/2018-22/BCS/010)

Name : - Nilratam Sarkar

Roll no : - 1814017

Registration no : - (ADTU/L/2018-22/BCS/017)

***Under The Guidence of : -***

Eirene Barua, Assistant Professor

Faculty of Engineering and Technology

Computer Science and Engineering

***Assam Down Town University***

Sankar MadhabPath Gandhi Nagar, Panikhaiti,

781026, Guwahati, Assam,India



**Assam Down Town University**

**Computer Science & Engineering**

**Faculty of Engineering & Technology**

Sankar Madhab Path , Gandhi Nagar,

Panikhaiti , 781026 , Guwahati , Assam , India

**CERTIFICATE**

This is to certify that A Project titled **“Implementation of Association Rule Mining”**  submitted by *Chandra Shekhar Yadav bearing Registration no : - ADTU/2018-22/BCS/010 & Roll no : - 1814010 and Nilratan Sarkar bearing Registration no : -ADTU/L/2018-22/BCS/017 & Roll no : - 1814017* , students of 7th sem , B.Tech C. S . E , carried under my guidance and supervision for the award of Degree Bachelor of Technology in Computer Science & Engineering of *Assam Down Town University* and the work is original and not a copy of any other project.

Date : -

( *Signature of Dean* ) ( *Signature of Supervisor )*

**DECLARATION**

We the undersigned solemnly declare that the report of the project work entitled “**Implementation of Association Rule Mining”**, is based on our own work carried out during the course of our study under the supervision of Eirene Barua .

We assert that the statements made and conclusions drawn are an outcome of the project work. We further declare that to the best of our knowledge and belief that the project report does not contain any part of any work which has been submitted for the award of any other degree in this University or any other University.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(Signature of the Candidate)

Name : -

Roll No .: -

Enrollment No.: -

**ACKNOWLEDGEMENT**

Place : -

Date : -

We would like to express our gratitude towards Mrs. Eirene Barua , Assistant Professor, Computer Science & Engineering, Faculty of Engineering & Technology , Assam Down Town University for her support in accomplishment of our project on

“**Implementation of Association Rule Mining” .**

I would like to extend my deep appreciation to my fellow group member, without his support and coordination we would not have been able to complete this project.

Name : -

Roll No.: -

Enrollment No.: -

Signature : -

**Contents : -**

List of Contents Page No.

1. Abstract……………...……………………………………………..6
2. Introduction……..………………………………………………….7
3. Aim and objective………………..………………………………...8
4. Platform………………………………………………………...…..9
5. Technologies used……………………………………………...…..9
6. Software requirement………………………………………...……10
7. Hardware requirement……………………………………………..10
8. Number of outputs corresponding to the given min\_support…..…11
9. EXPERIMENTS(SCREENSHOTS)………..……...……………...12-17
10. Plotted image of the output…………………..…………………...17-18
11. Conclusion…………………………………………………………19

***1 . ABSTRACT : -***

Task of extracting useful and interesting knowledge from large data is called data mining. It has many aspects like clustering, classification, association mining, outlier detection, regression etc. Among them association rule mining is one of the important aspect for data mining.

Best example of association rule mining is market-basket analysis. Applications of association rule mining are stock analysis, web log mining, medical diagnosis, customer market analysis bio-informatics etc.

The APriori Data Mining Algorithm is used to create association rules from sets of items. The algorithm finds patterns of items. Algorithm uses knowledge from previous iteration phase to produce frequent item-sets that are frequently associated together. A confidence measure is created for each rule generated from the frequent item-sets.

***2 . INTRODUCTION : -***

* We in this project will First import a data-set .
* Then convert the data-set into one-hot encoding array ,where each unique label is represented as a column in the new array.
* Then mine frequent item-set from it with given input for minimum support metric
* Then mine association rules using apriori algorithm with metric of support , confidence , lift , leverage .
* Then visualize the output for better understanding of the data.
* At the end we will save the data in a csv file.

***3 . AIM & OBJECTIVE :-***

* To find frequent item-set from a given data-set.
* Mine association rules from the generated frequent item-set .
* Sort the data on basis of metric concordance and lift with given minimum threshold .
* Visualize the mined association rules for better understanding of the data.
* Then store the data in a csv file .
* The generated data can be used for many things like market basket analysis,ux design, disease predication etc and many more.

***4 . PLATFORM : -***

* **Operating system : -** Microsoft windows.

***5 . Technologies used : -***

* **Python :-** Python is an interpreted high-level general-purpose programming language. Python is often used to build websites and software, automate tasks, and conduct data analysis,data visualization,AI and machine learning.
* **Libraries used : -**

1. **Numpy : -** NumPy is Python library used for working with arrays. It also has functions for working in domain of linear algebra, fourier transform, and matrices.
2. **Pandas : -**Pandas is a software library written for the Python programming language for data manipulation and analysis.
3. **Mlxtend : -**Mlxtend (machine learning extensions) is a Python library of useful tools for the day-to-day data science tasks.
4. **Matplotlib : -**Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python.

***6 . Software Requirements : -***

* **Anaconda :-** Anaconda is a distribution of the Python and R programming languages for scientific computing, that aims to simplify package management and deployment.

***7 . Hardware Requirements : -***

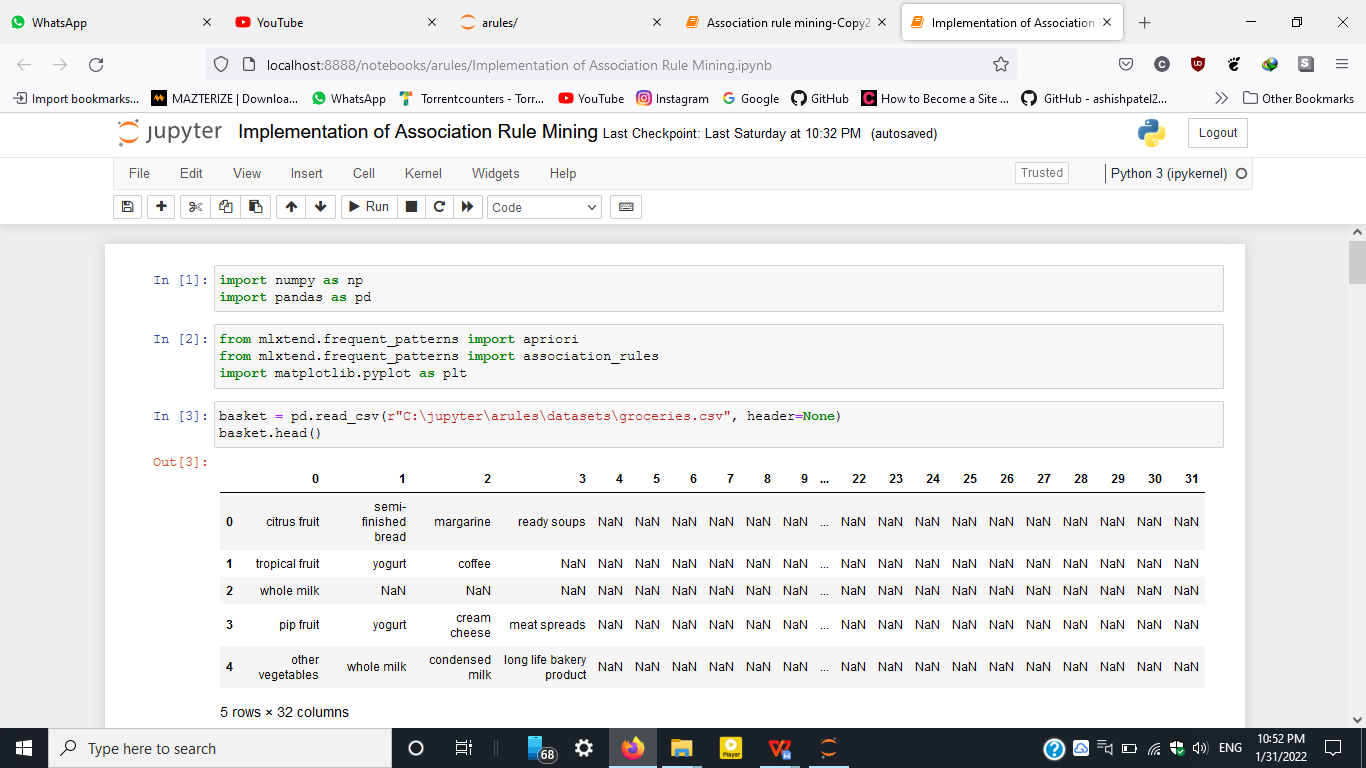
* Intel core i3, i5,i7 processor or equivalent or higher.
* 1GB Ram or Higher.
* 20 GB HDD or Higher.
* 1.5GB free space.

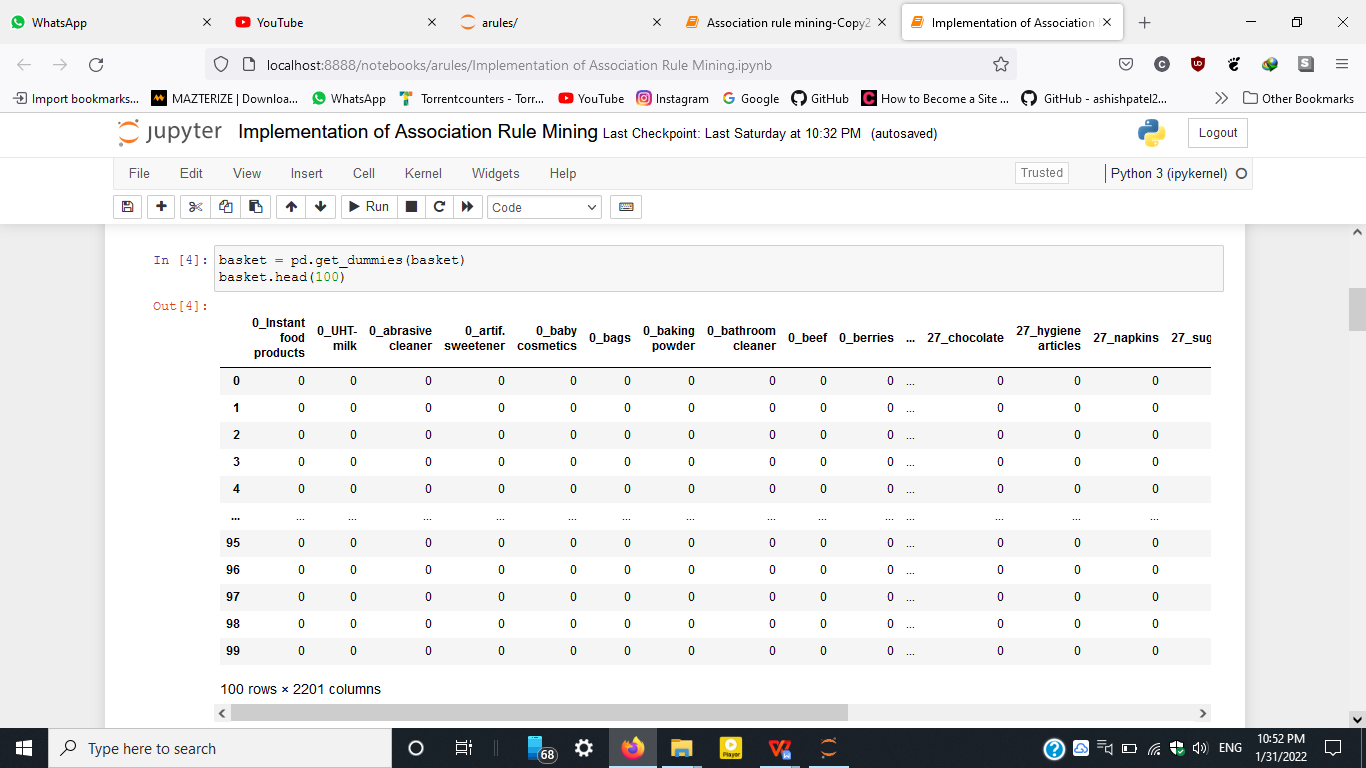
**8.**

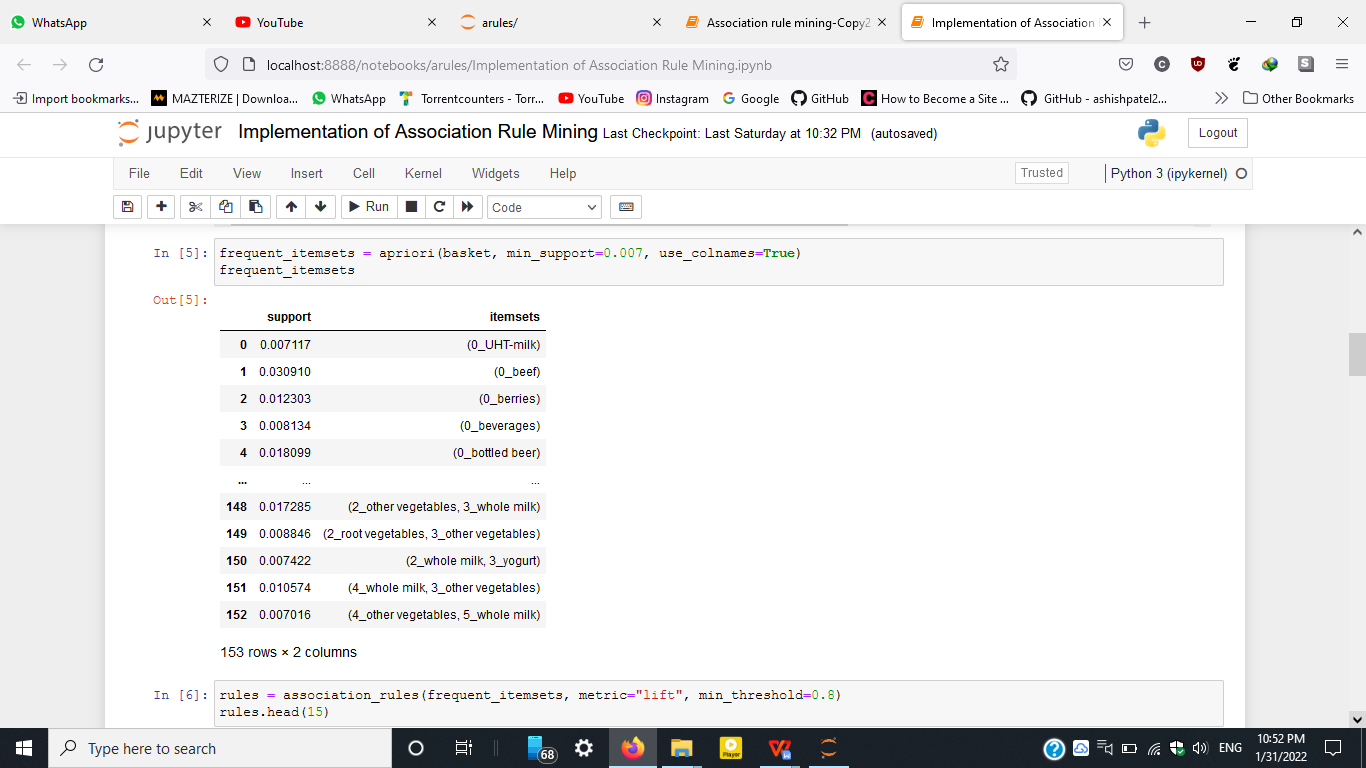
***Number of outputs corresponding to the given min\_support : -***

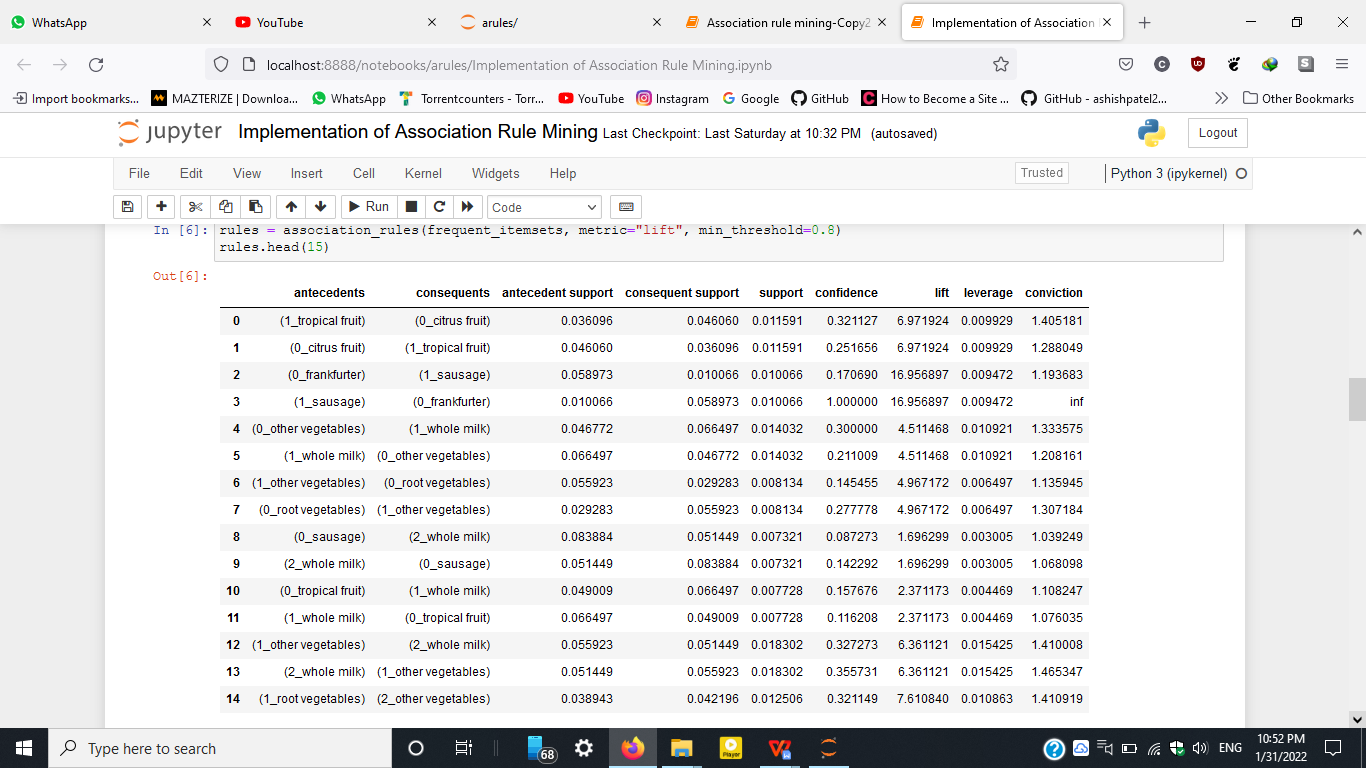
|  |  |  |
| --- | --- | --- |
| Serial no. | min\_support | Number Of Outputs |
| 01 | 0.0838 | 1 |
| 02 | 0.0755 | 1 |
| 03 | 0.0729 | 2 |
| 04 | 0.0675 | 2 |
| 05 | 0.0664 | 3 |
| 06 | 0.0600 | 3 |
| 07 | 0.0589 | 4 |
| 08 | 0.0560 | 4 |
| 09 | 0.0559 | 5 |
| 10 | 0.0525 | 5 |
| 11 | 0.0514 | 6 |
| 12 | 0.0500 | 6 |
| 13 | 0.0490 | 7 |
| 14 | 0.0475 | 7 |
| 15 | 0.0467 | 8 |
| 16 | 0.0462 | 8 |
| 17 | 0.0460 | 9 |
| 18 | 0.0425 | 9 |
| 19 | 0.0421 | 10 |
| 20 | 0.0395 | 10 |
| 21 | 0.0391 | 11 |
| 22 | 0.0390 | 11 |
| 23 | 0.0389 | 12 |
| 24 | 0.0385 | 12 |
| 25 | 0.0384 | 13 |
| 26 | 0.0370 | 13 |
| 27 | 0.0360 | 14 |
| 28 | 0.0345 | 14 |
| 29 | 0.0341 | 15 |
| 30 | 0.0321 | 15 |

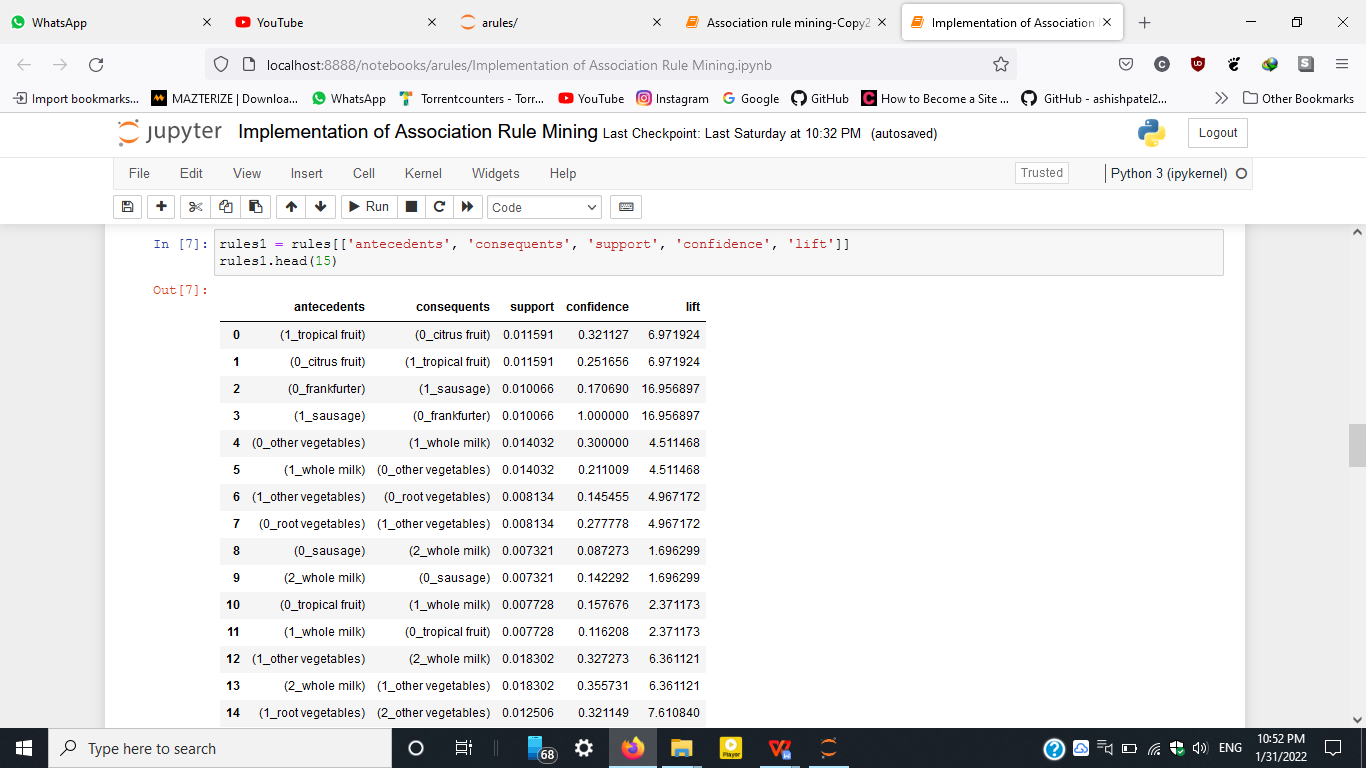
***9 . EXPERIMENTS(SCREENSHOTS) : -***

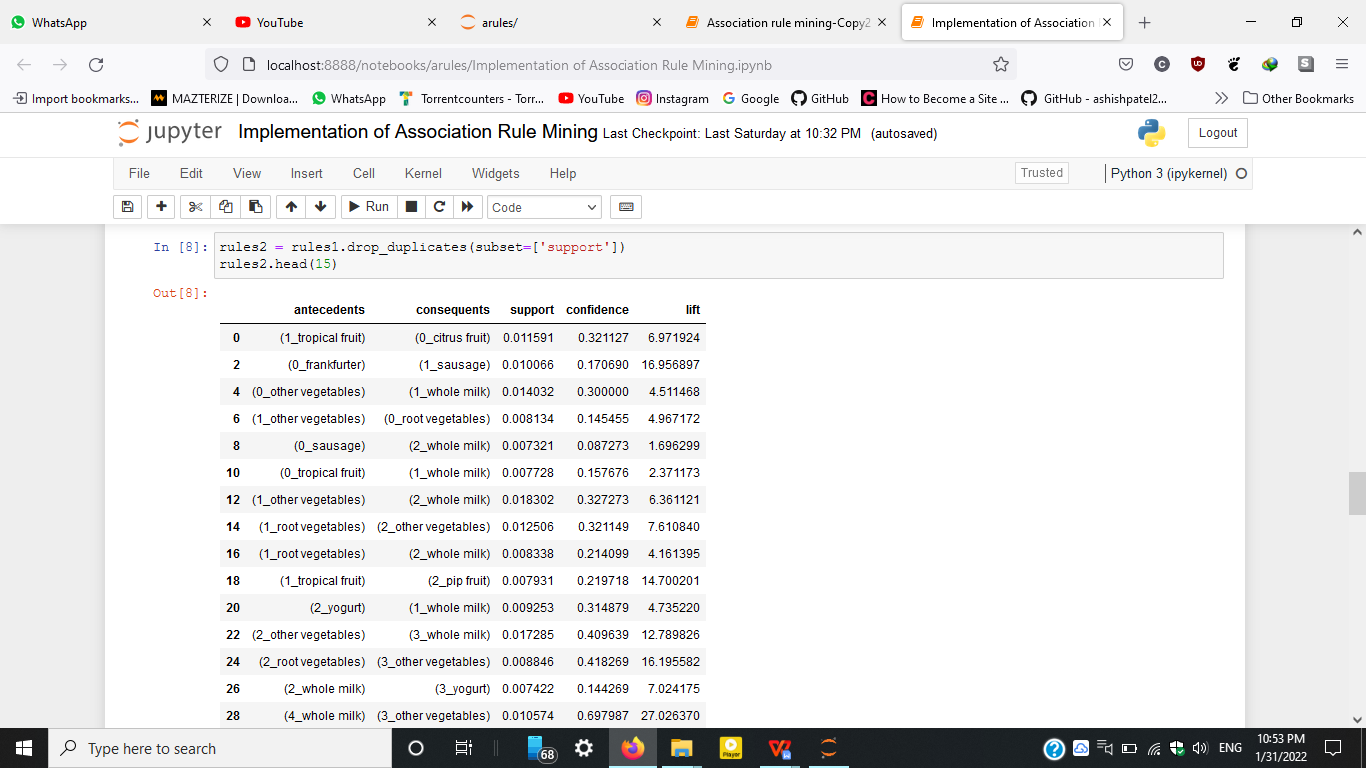
**

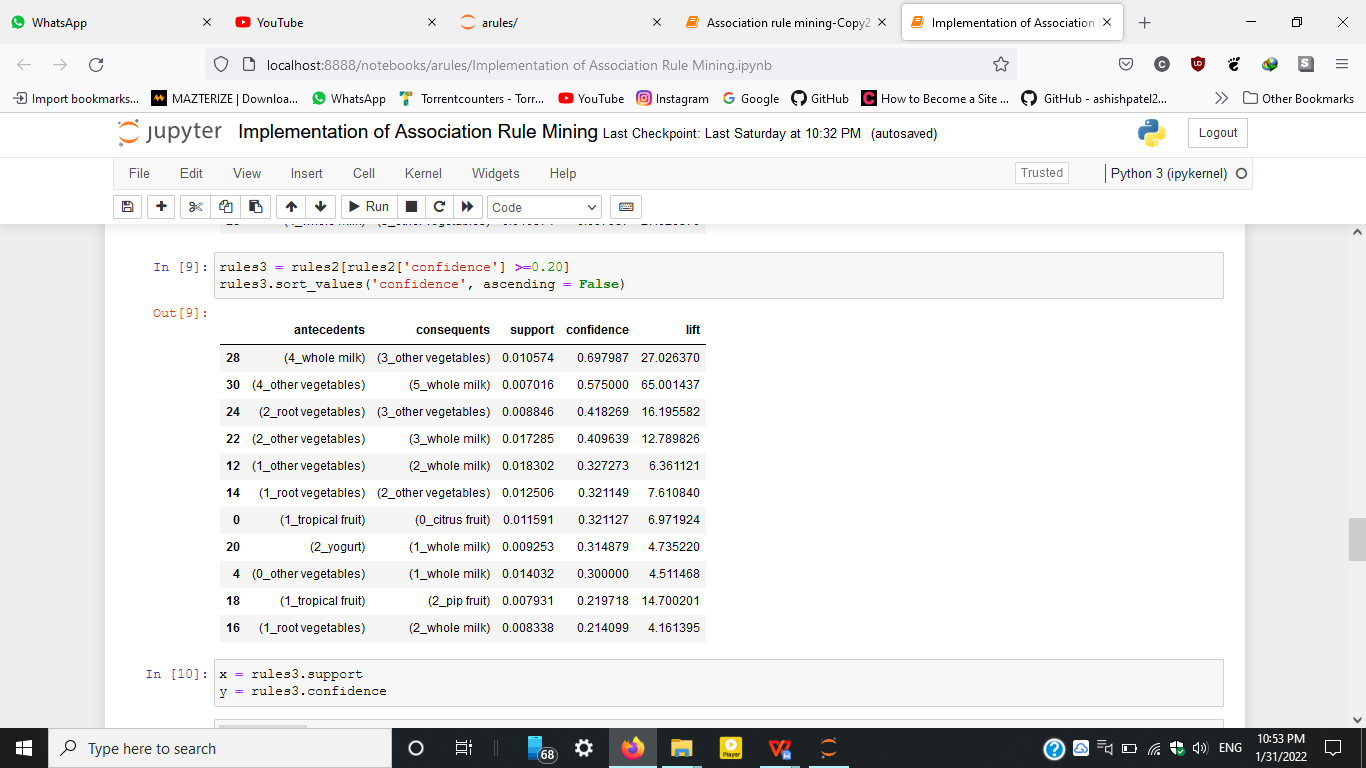
****

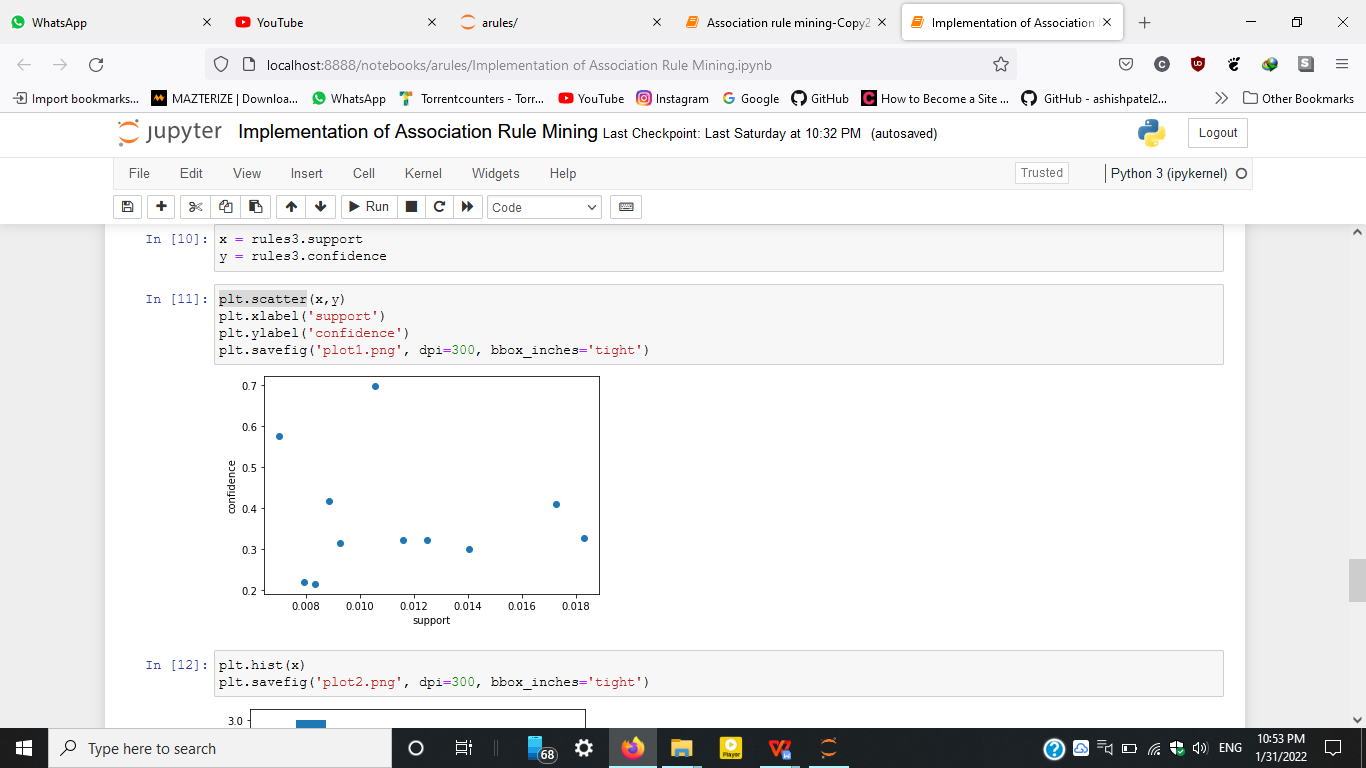
****

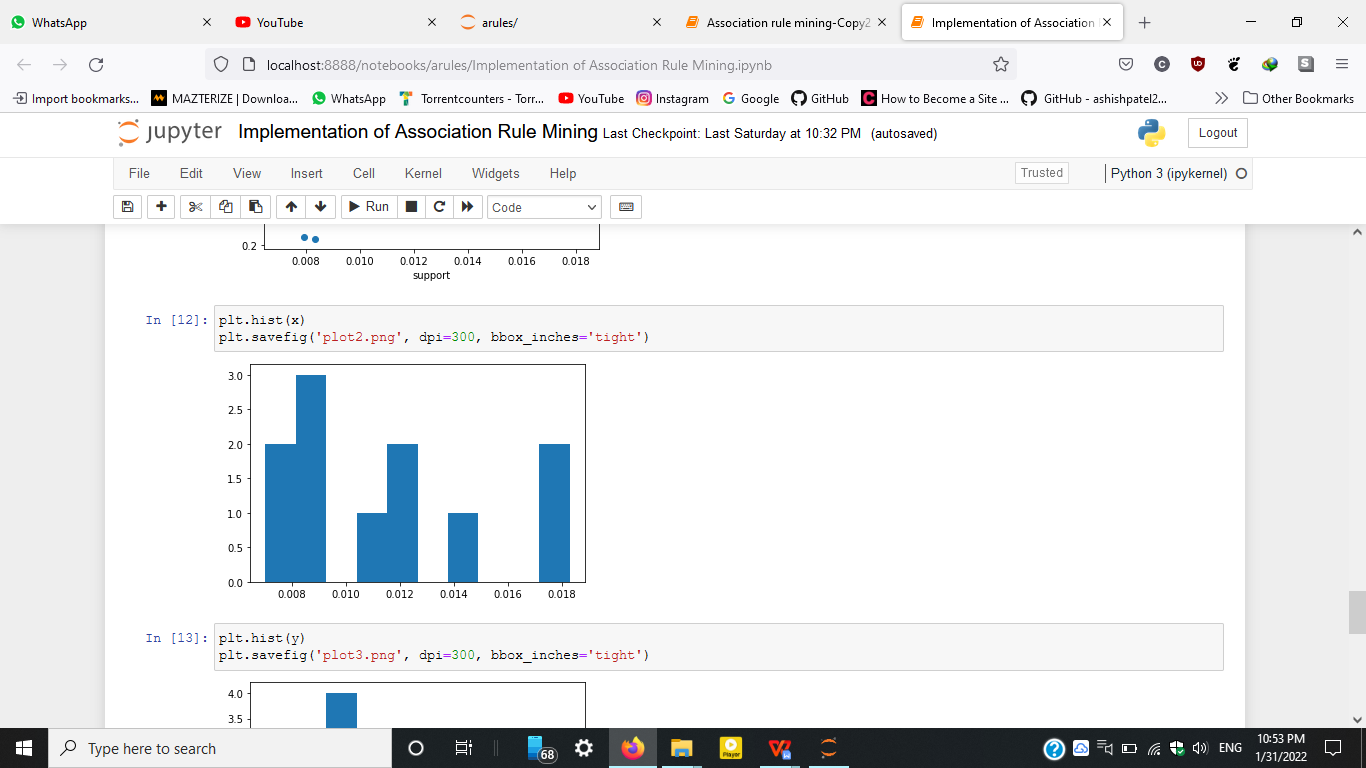
****

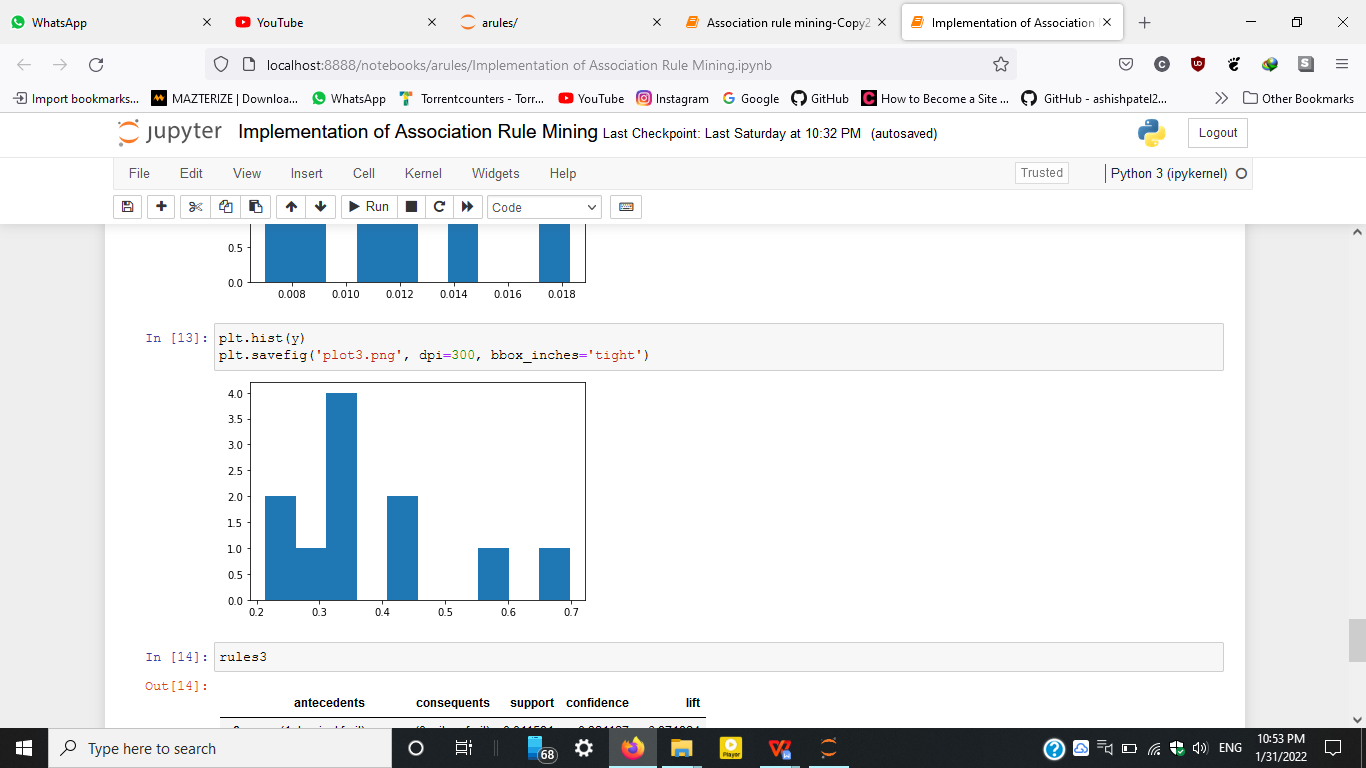
****

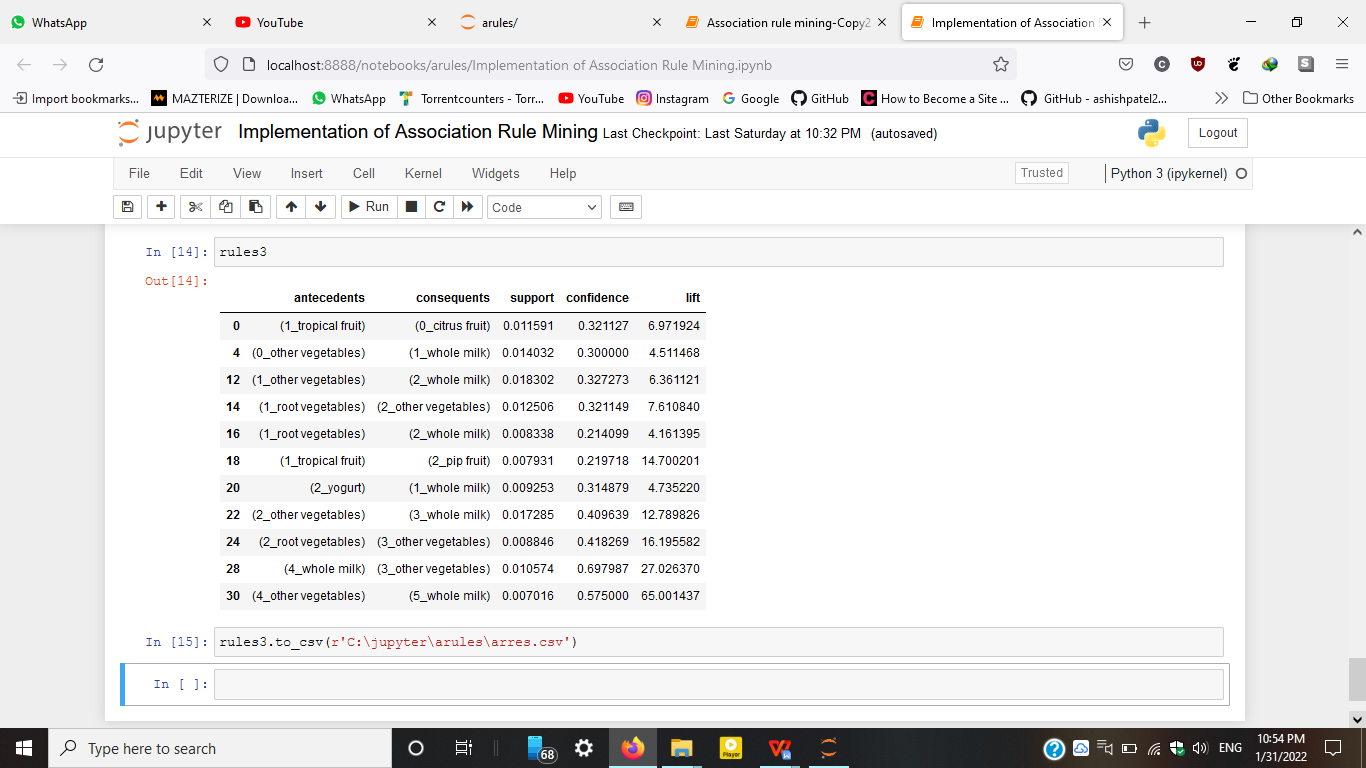
****

****

****

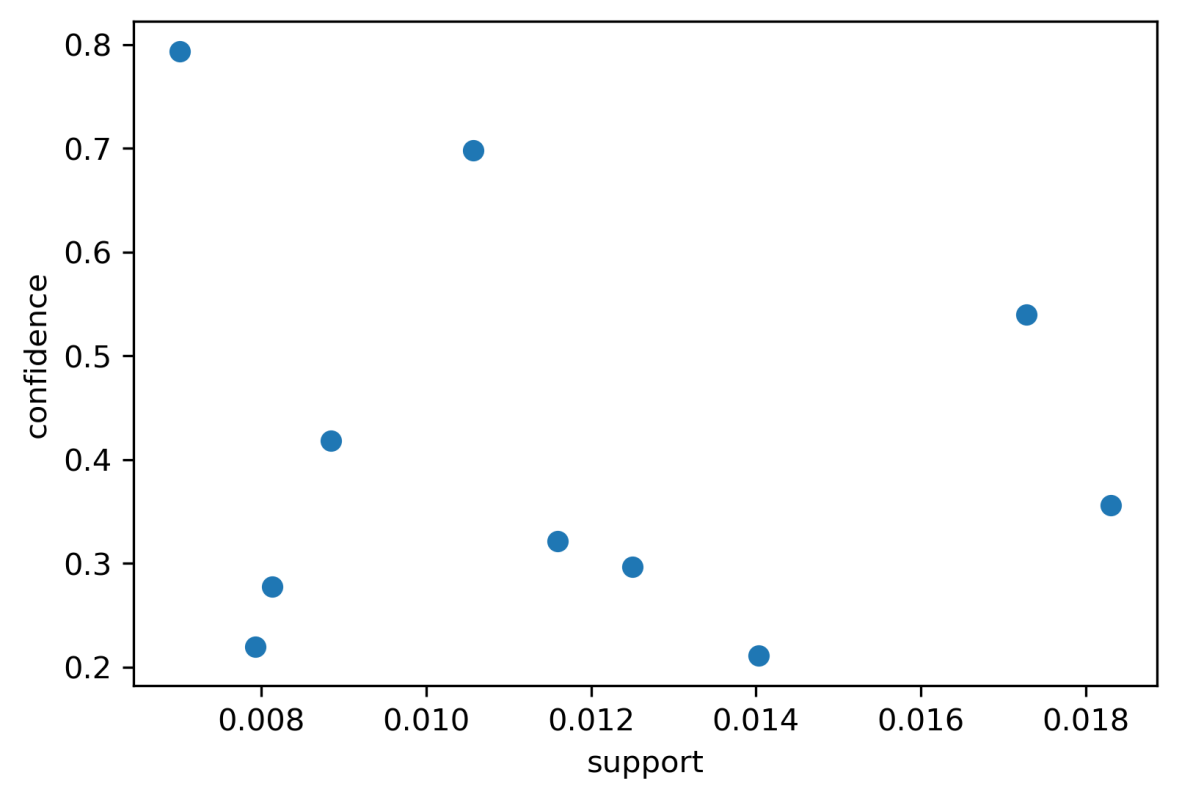
****

****

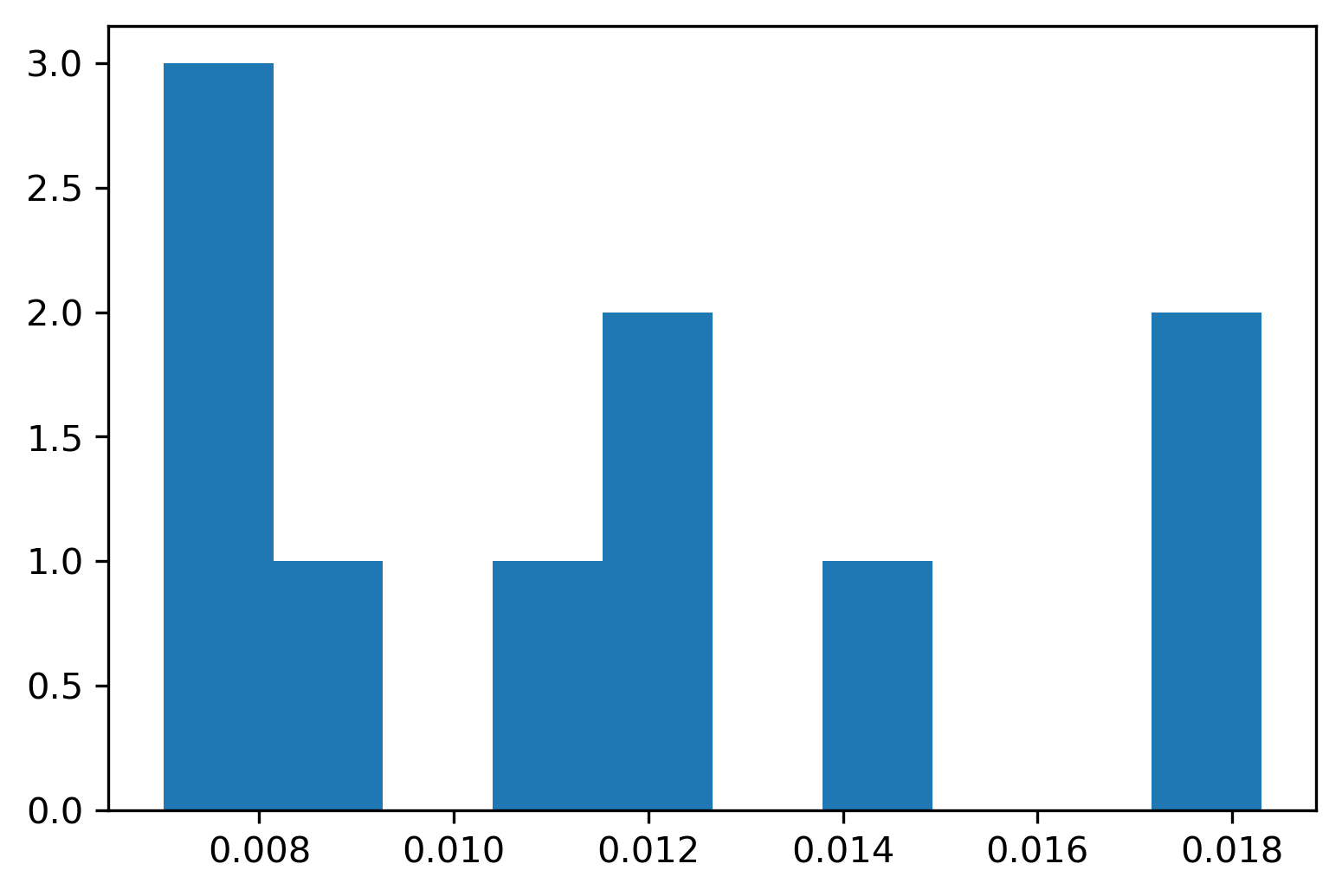
****

***10. Plotted image of the output : -***

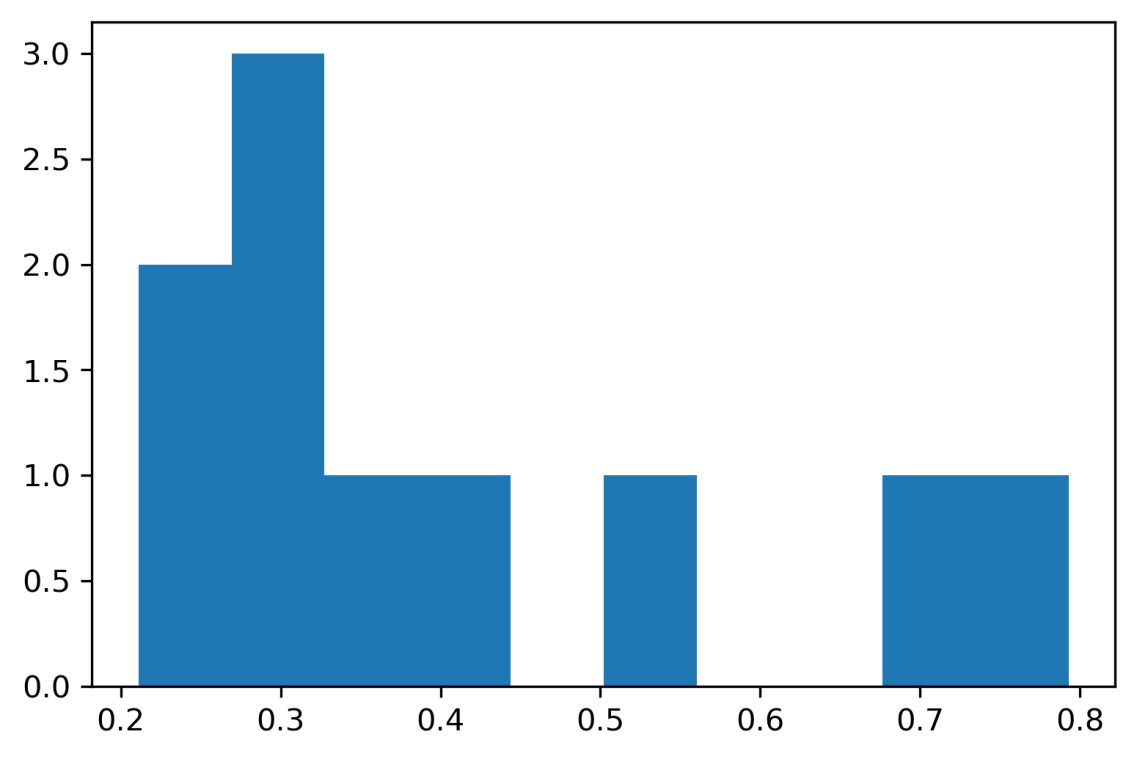
1. Each value in the data set is represented by a dot : -

****

1. Horizontal axis represents support & vertical axis represents number of rules: -

****

1. Horizontal axis represents confidence & vertical axis represents number of rules:-

****

***11 . Conclusion : -***

The project on **“Implementation of Association Rule Mining”** greatly helped us to develop our knowledge on Python .Association rule mining can be greatly useful for businesses willing to increase their sales whether it's a small retail store or e-commerce website .