Data Mining Project Midterm Report:

**Introduction**

This document describes how to use and run the python script for generating association rules using Brute-Force, Apriori and FP-Growth and also compare their performance.

**Tools and Libraries:**

To execute the program we need following libraries:

1. mlxtend
2. pandas
3. numpy
4. itertools

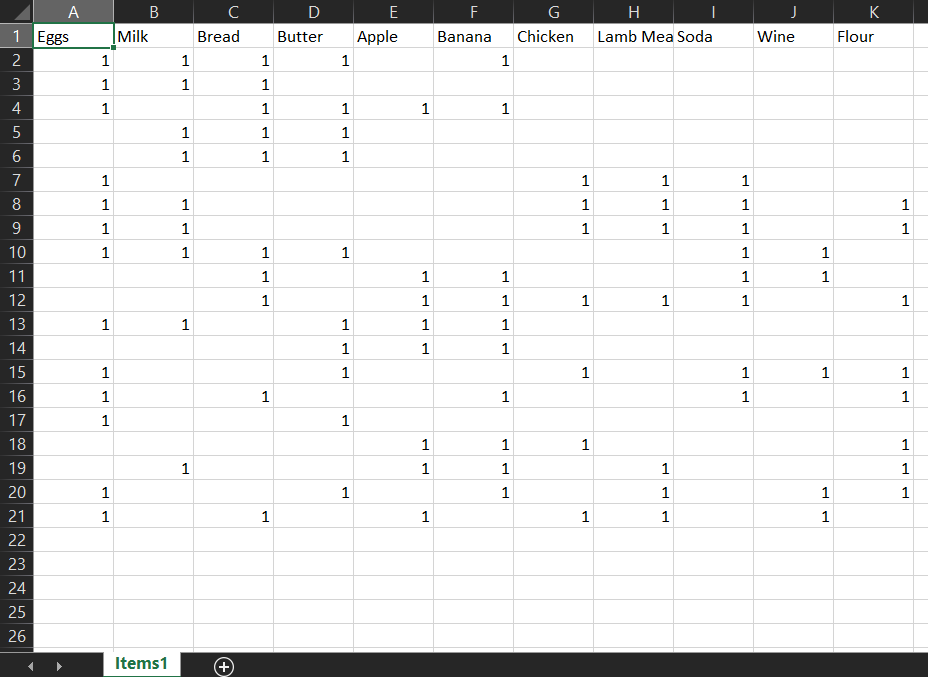
**Set-up:**

1. Clone the github repository from link : **add git\_link**
2. Go to project directory
3. Install python libraries by executing **pip install -r requirements.txt**
4. **Run command “python .\mid\_term\_project.py” to start the program**

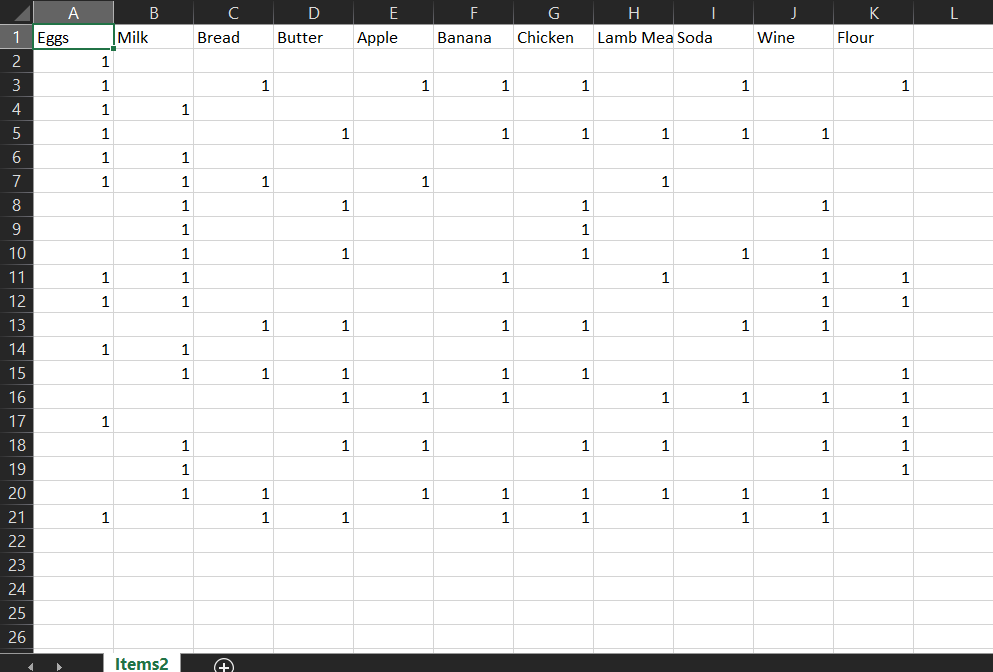
**Data:**

For testing the implementation I have created 5 datasets with 20 transactions and 10 unique items.

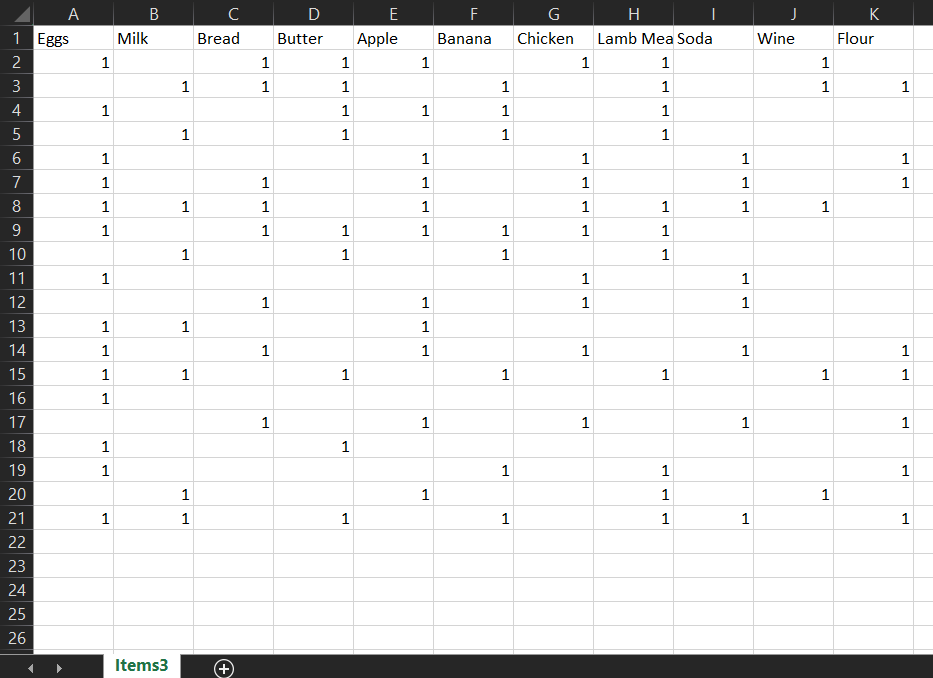
**Items1.csv**

****

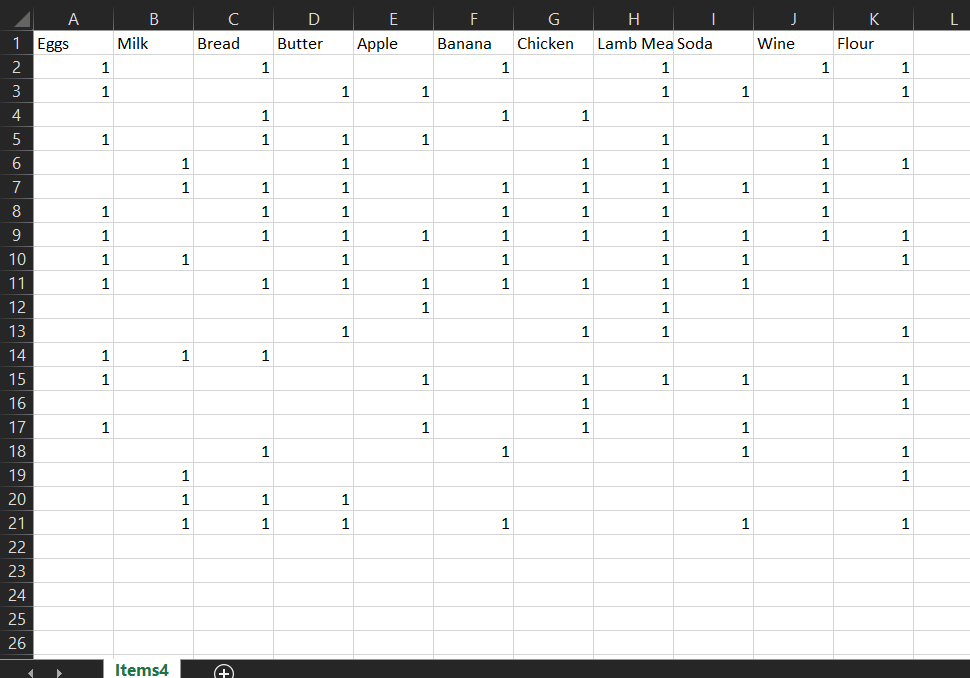
**Items2.csv**

****

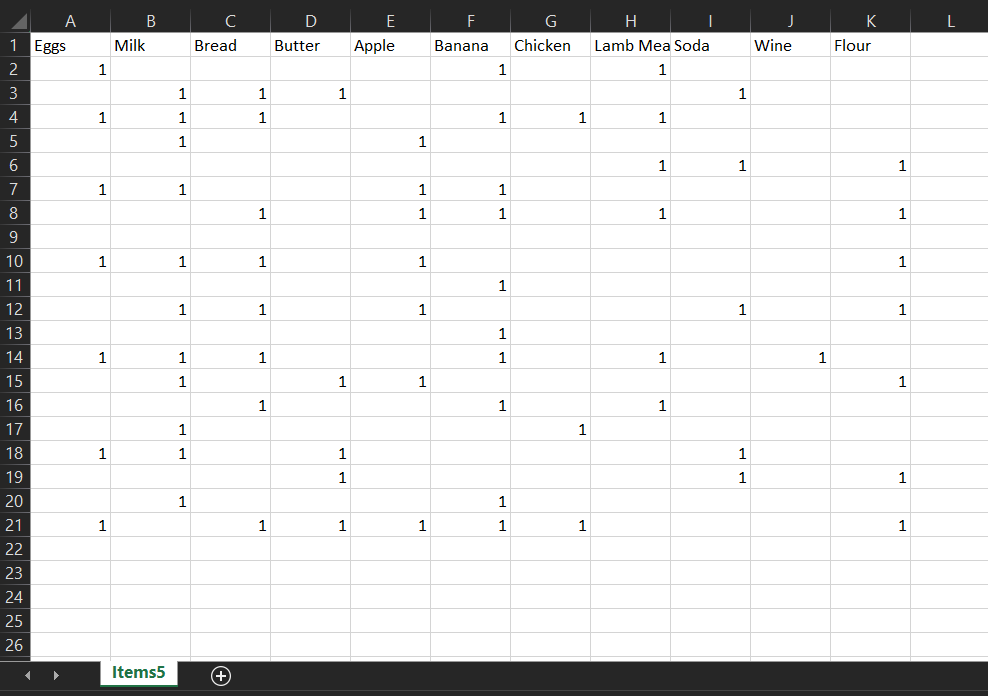
**Items3.csv**

****

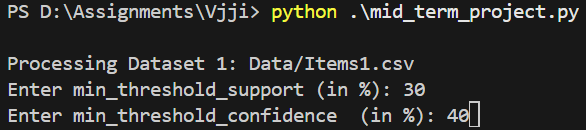
**Items4.csv**

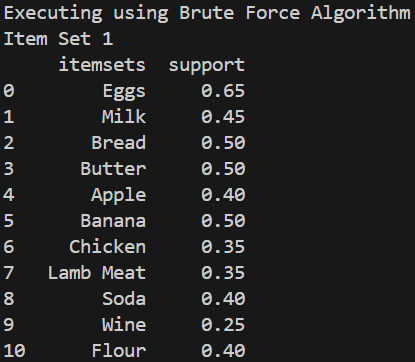
****

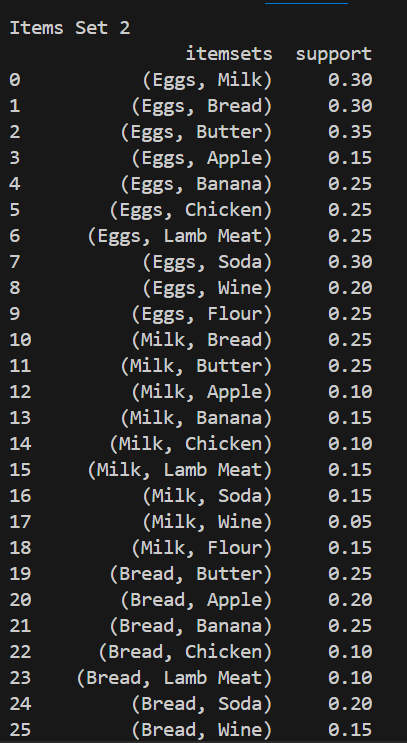
**Items5.csv**

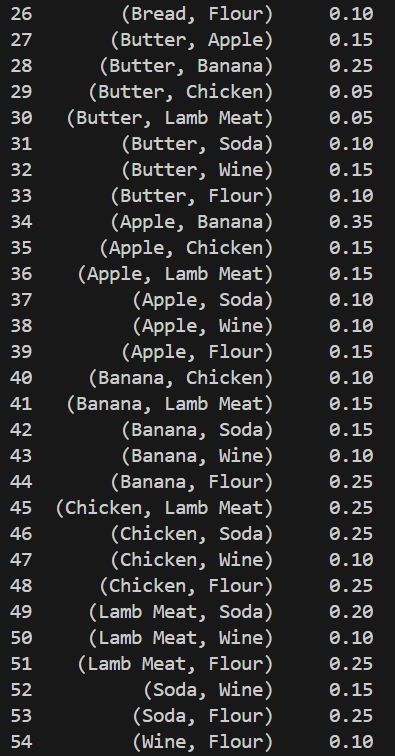
****

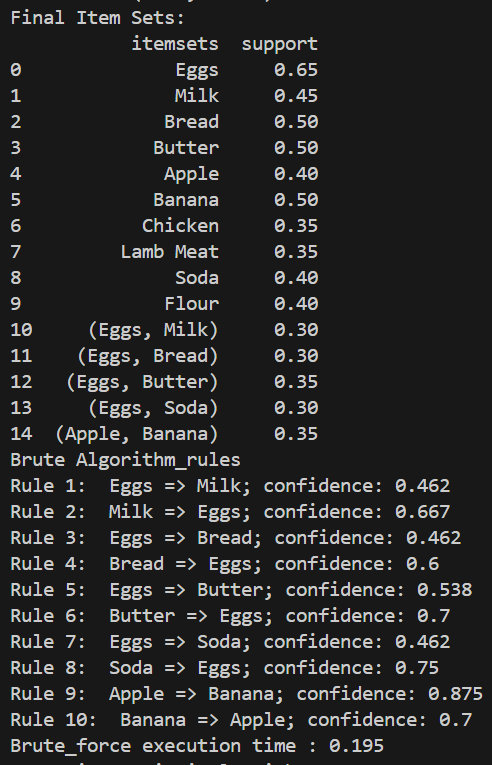
**Working:**

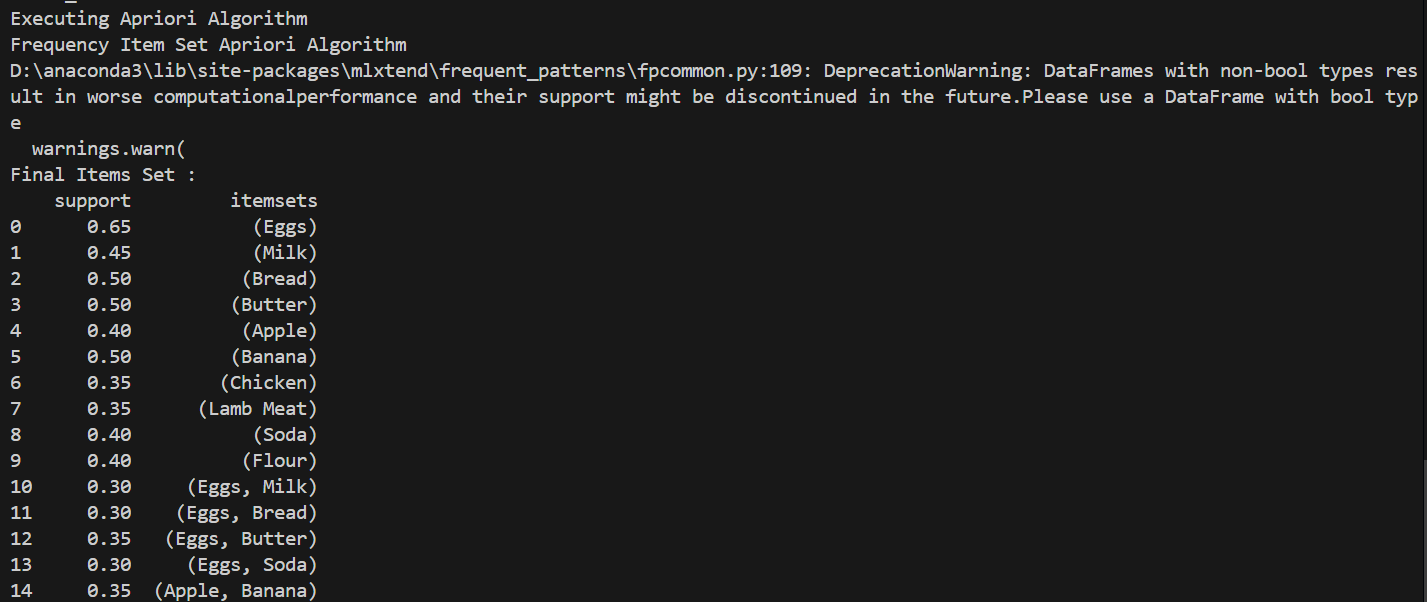
****

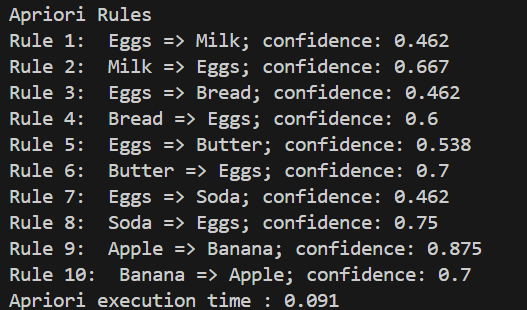
****

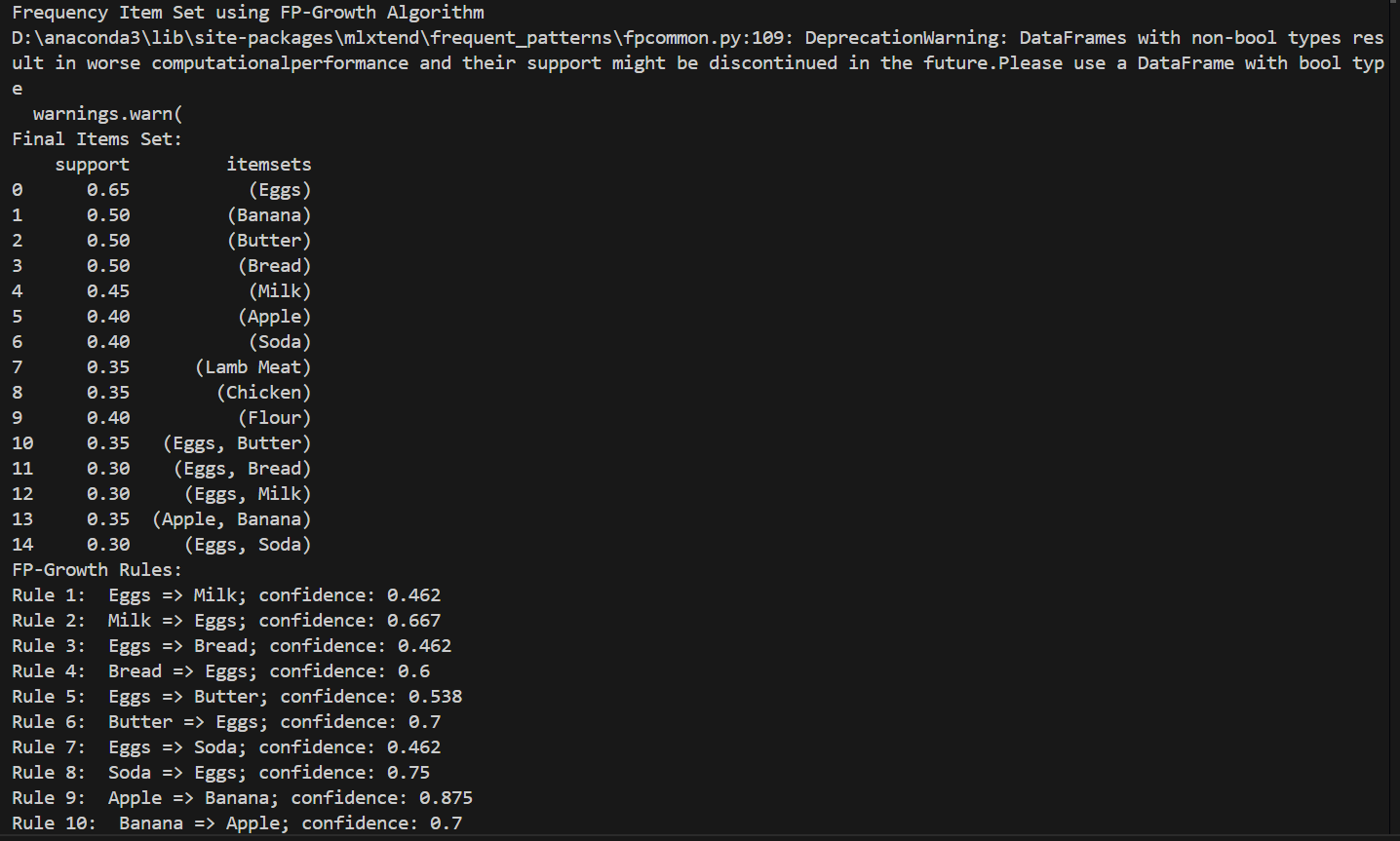
****

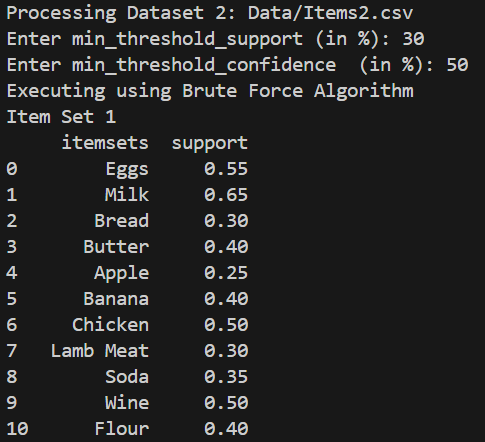
****

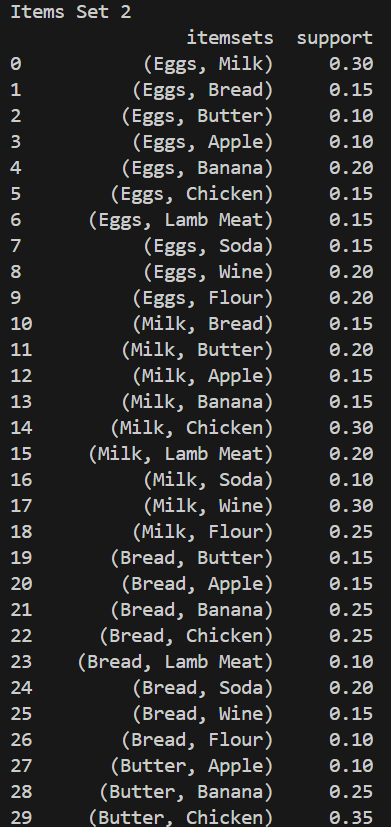
****

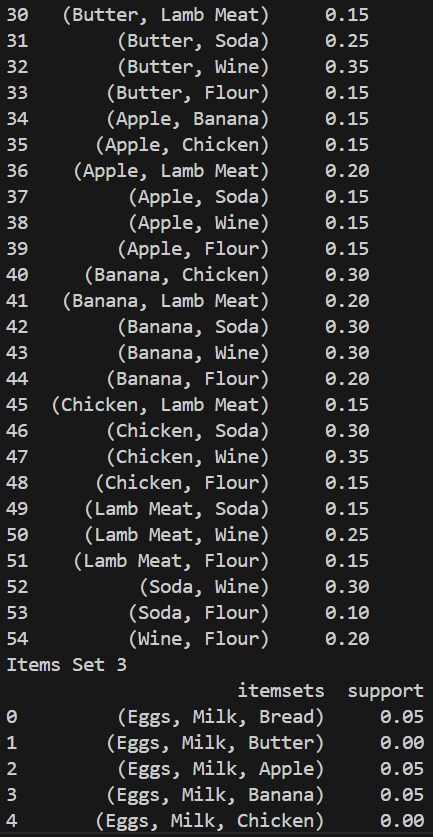
****

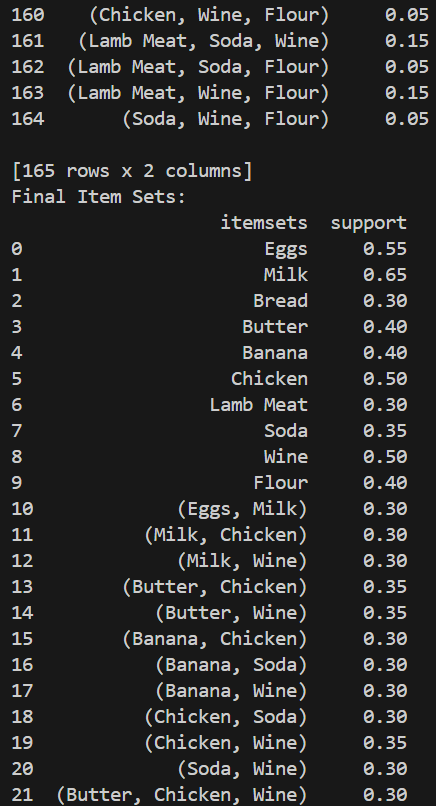
****

****

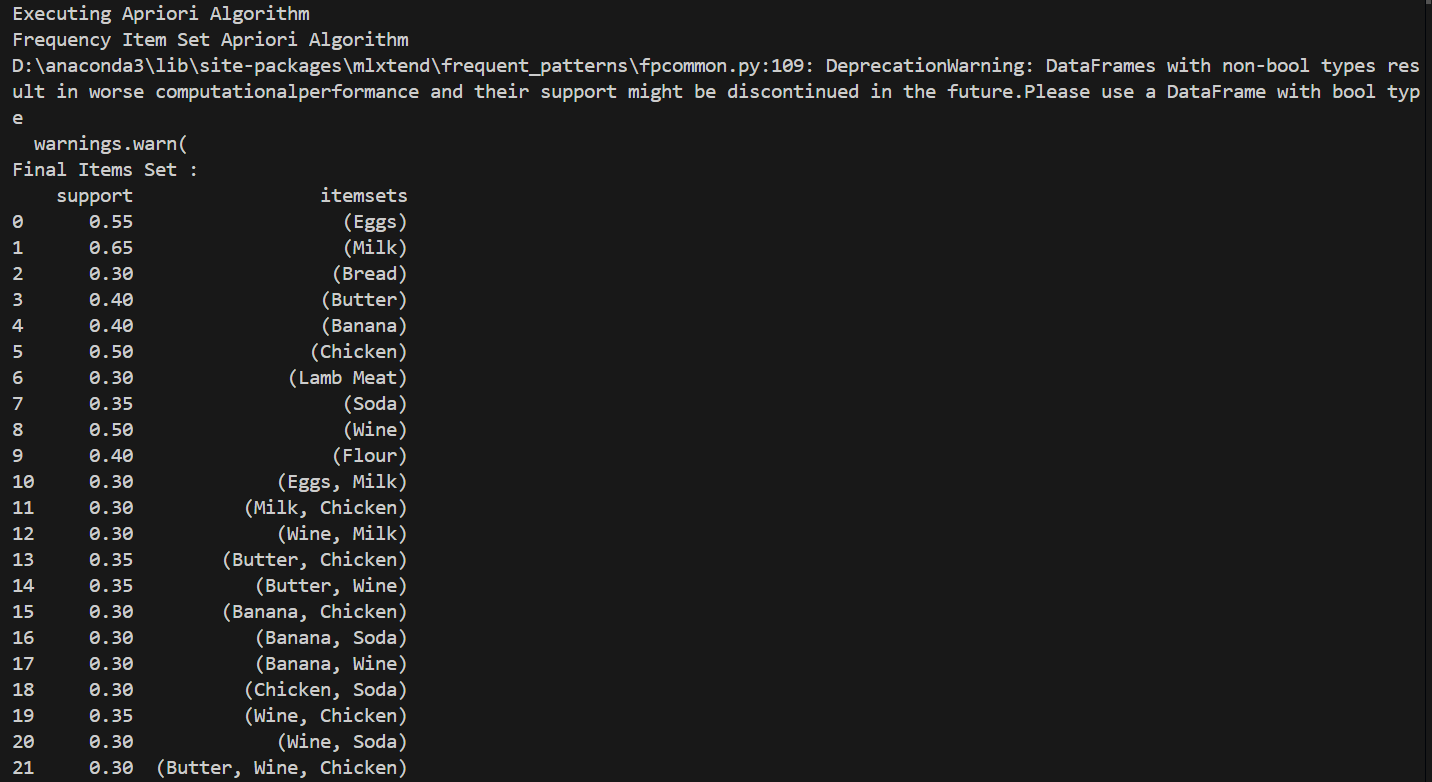




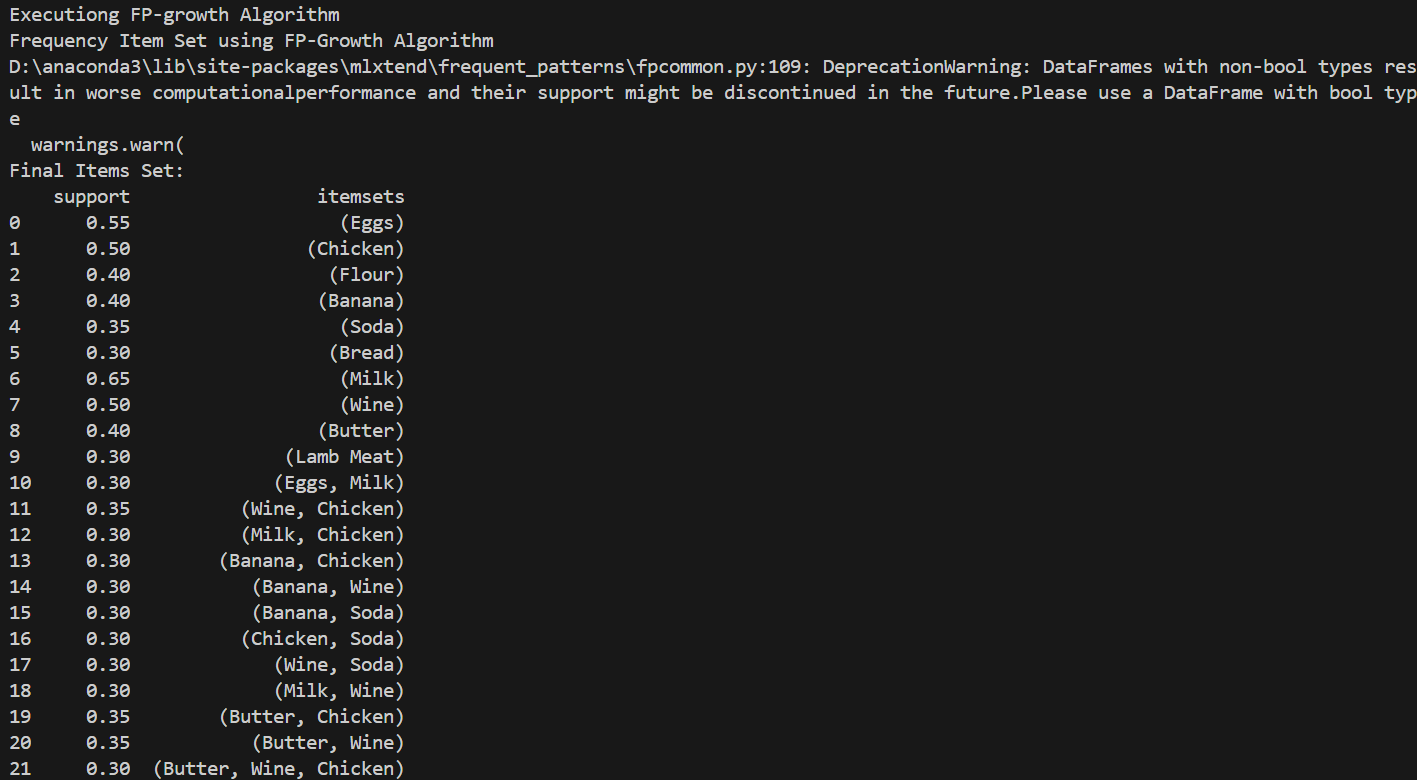


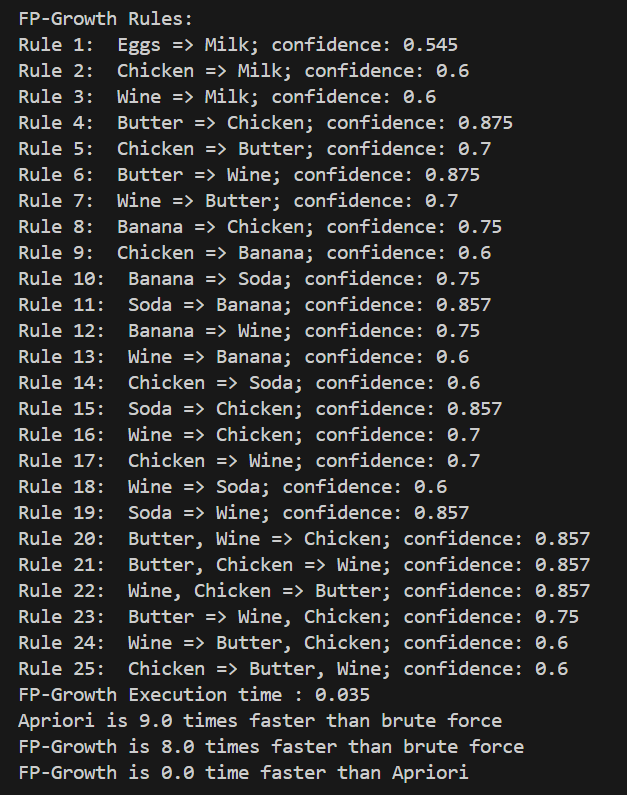


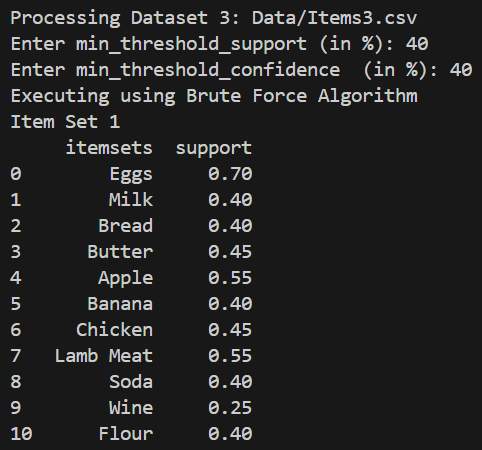


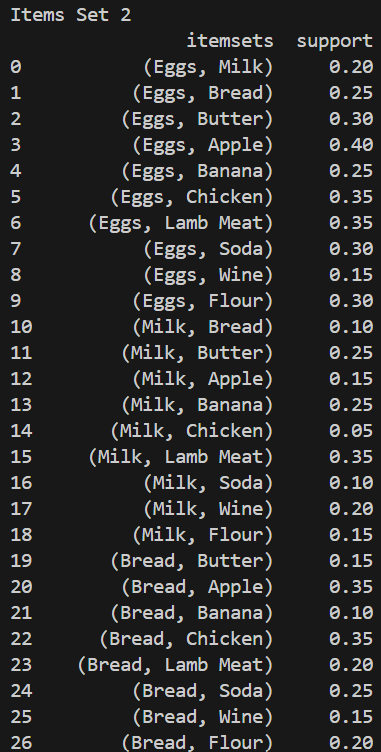


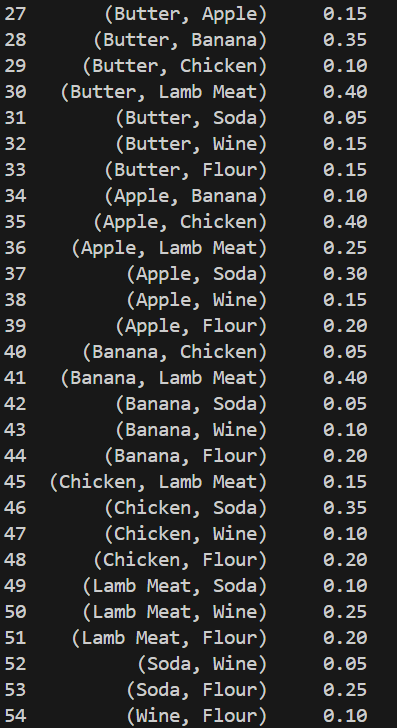


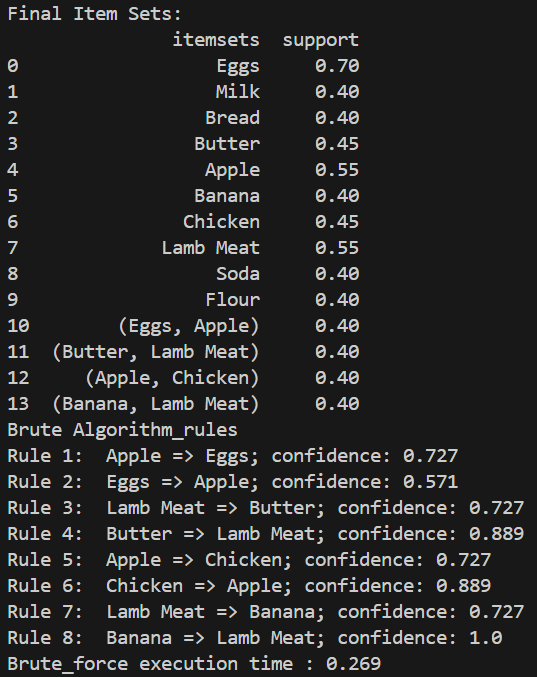


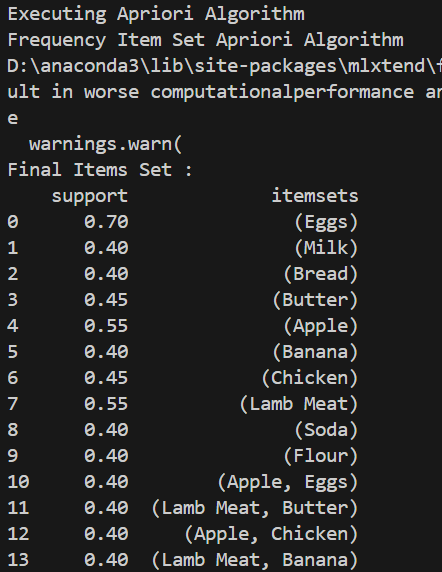


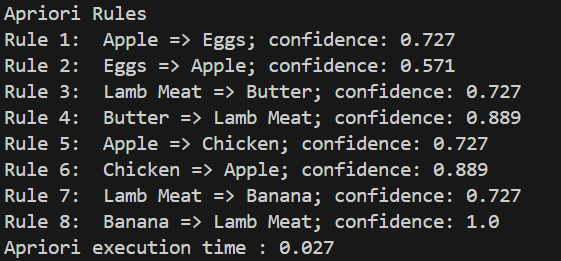


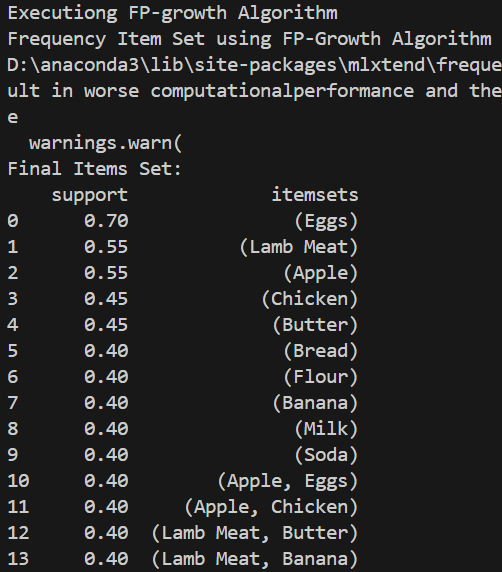


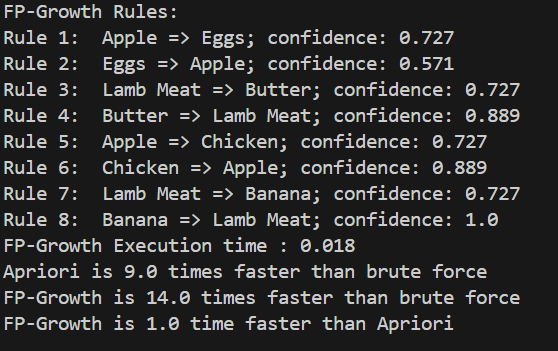


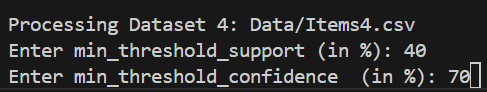


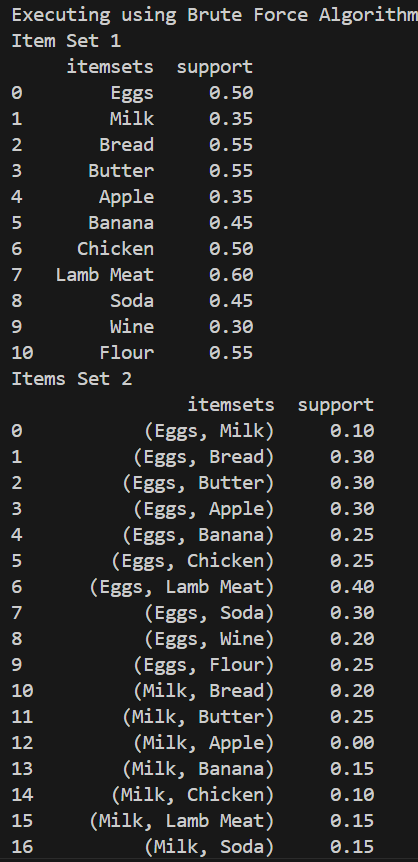


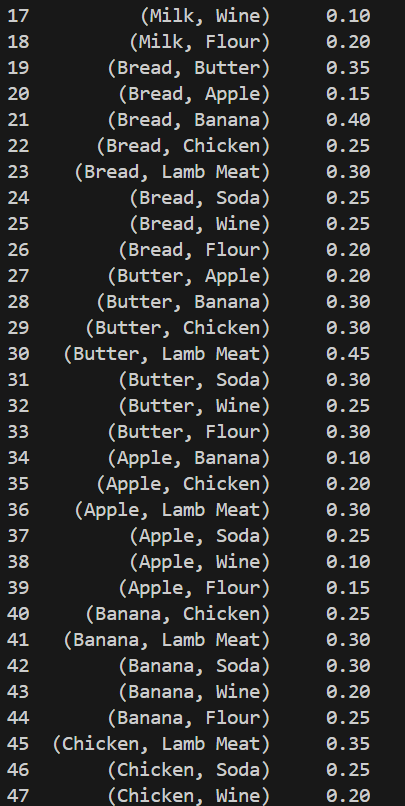


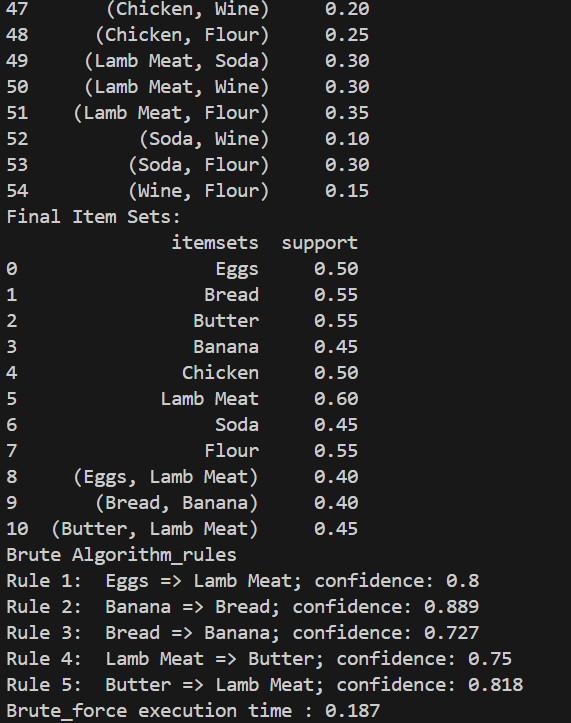


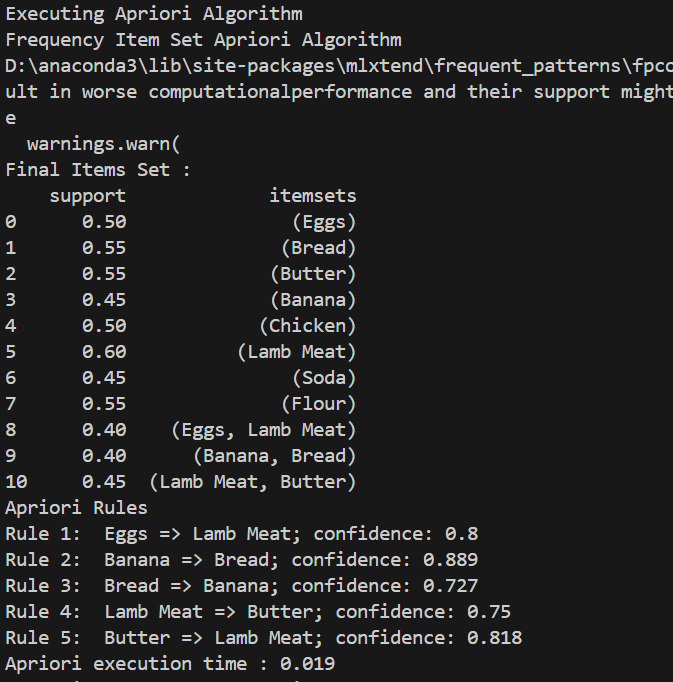


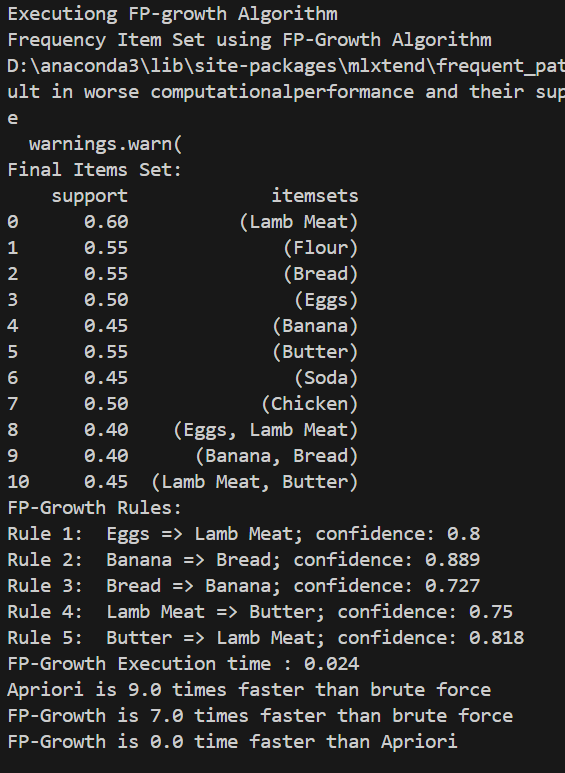




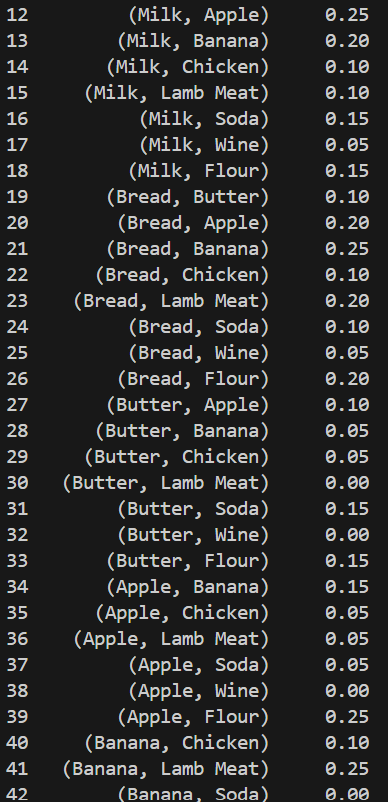


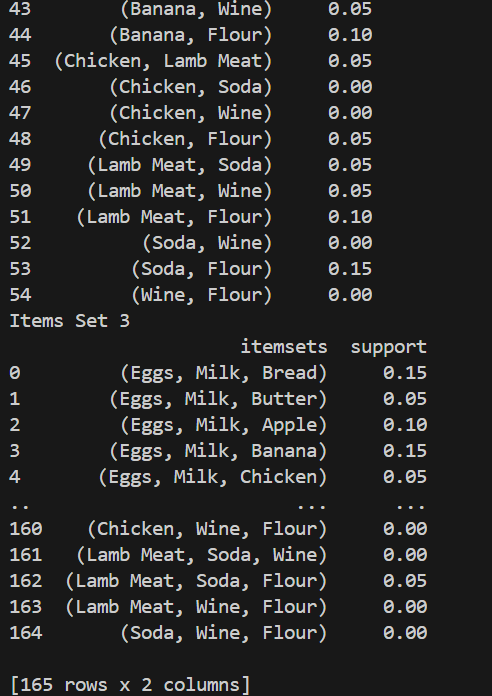


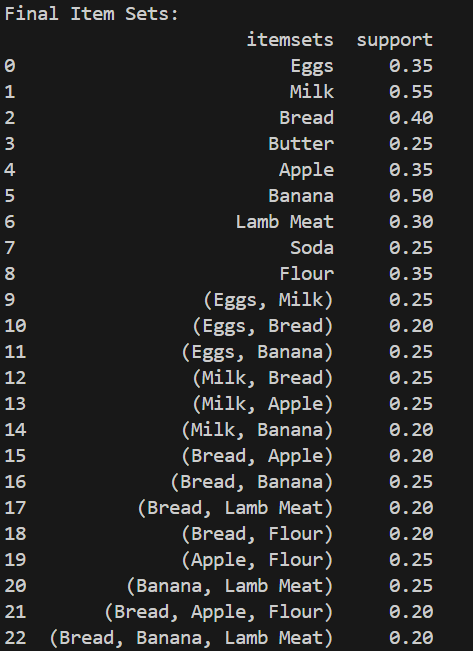


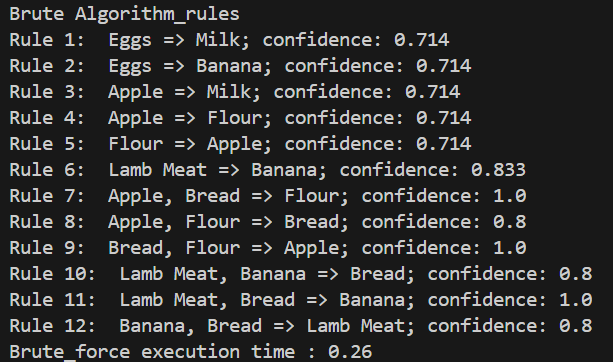




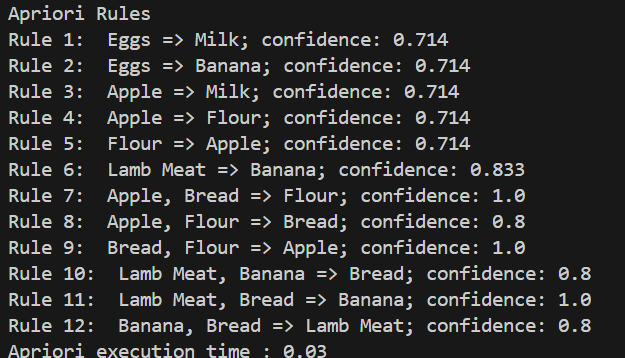


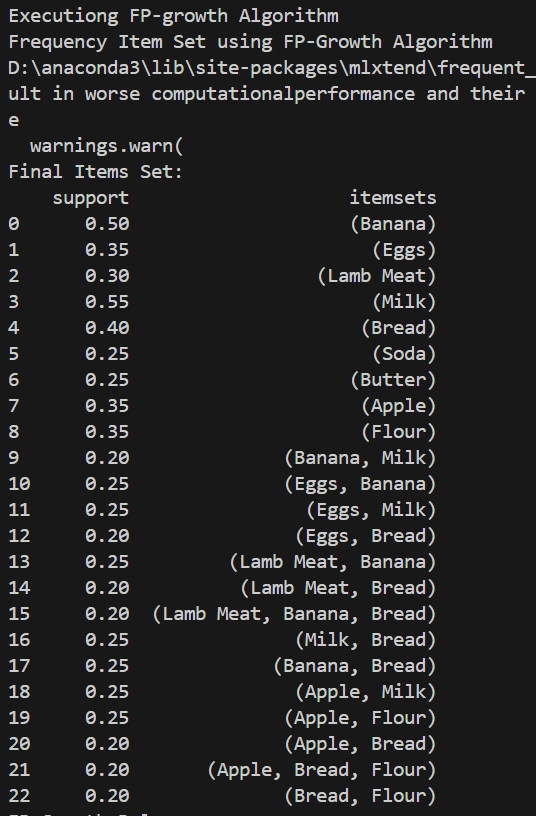


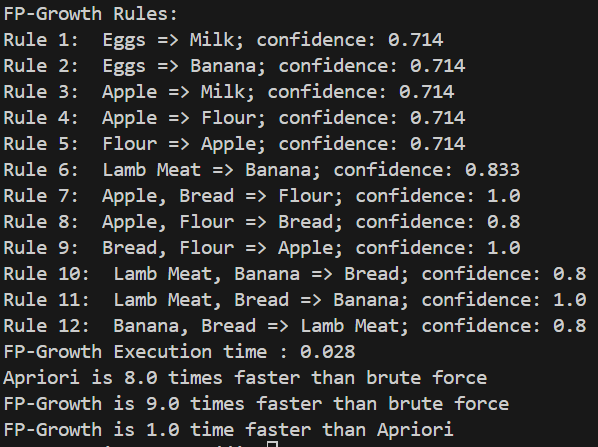












**Conclusion:**

* Brute force is the slowest the algorithm for associate rule mining
* All three algorithms produce same frequency items set
* All three algorithms produce same rules
* FP-Growth is the fastest algorithm for association rule mining.