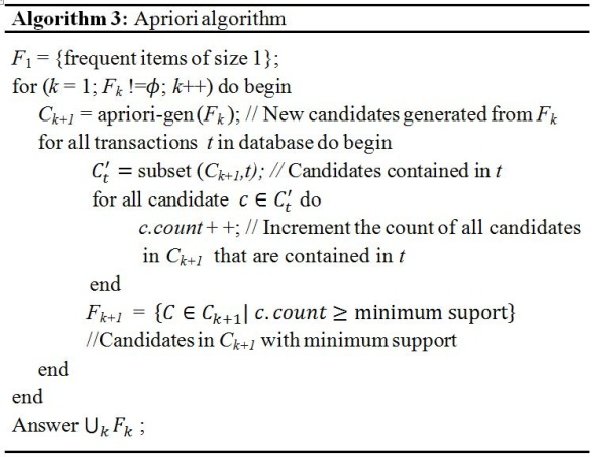
|  |  |
| --- | --- |
| Algorithm: Apriori Algorithm | |
| USN: 1MS18CS024 | NAME: Aravind S |
| USN: 1MS18CS029 | NAME: Arun J Kennedy |

**Description of the Algorithm: <<Write 2-3 Paragraphs about the Algorithm>>**

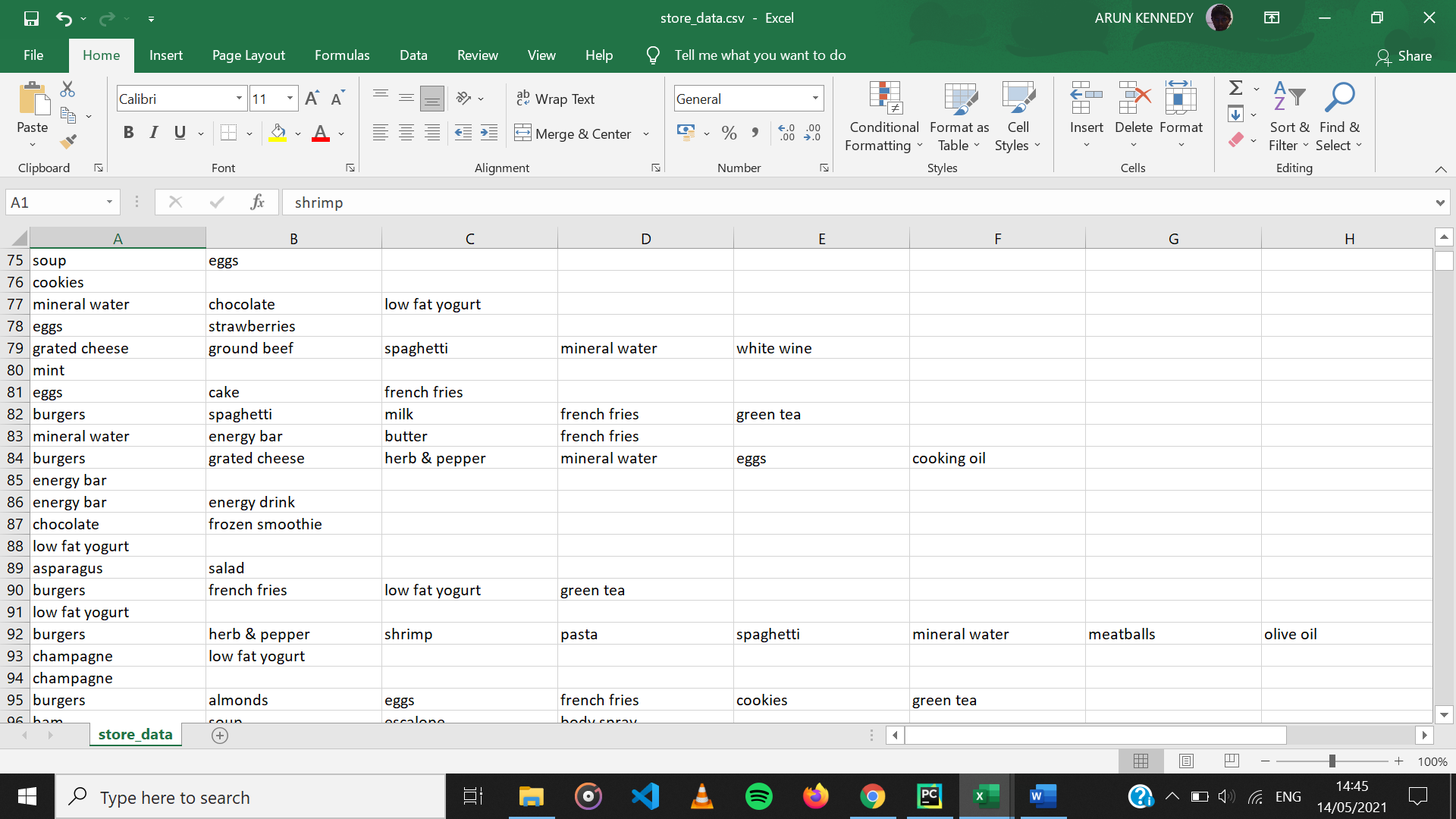
Apriori algorithm is for finding frequent itemsets in a dataset for boolean association rule. Name of the algorithm is Apriori because it uses prior knowledge of frequent itemset properties. We apply an iterative approach or level-wise search where k-frequent itemsets are used to find k+1 itemsets.

To improve the efficiency of level-wise generation of frequent itemsets, an important property is used called Apriori property which helps by reducing the search space.

Apriori assumes that “All subsets of a frequent itemset must be frequent (Apriori propertry). If an itemset is infrequent, all its supersets will be infrequent.”

**Algorithm Pseudocode:**

**Data set Used: (Attach Screen shot of the few rows and also the Kaggle/Dataset link)**



Dataset link: <https://drive.google.com/file/d/1y5DYn0dGoSbC22xowBq2d4po6h1JxcTQ/view?usp=sharing>

**Challenges faced during the implementation of the program:**

No particular challenges were faced to retrieve the information.

Python contains a library called apriori in which the algorithm can be readily implemented.

**Advantages & Disadvantages of the Algorithm:**

***Advantages:***

* Uses large itemset property
* Easily parallelized
* Easy to implement
* Withdraws rescanning the database.
* Decreases the size of the candidate itemsets and stimulate both the meeting and the pruning process.

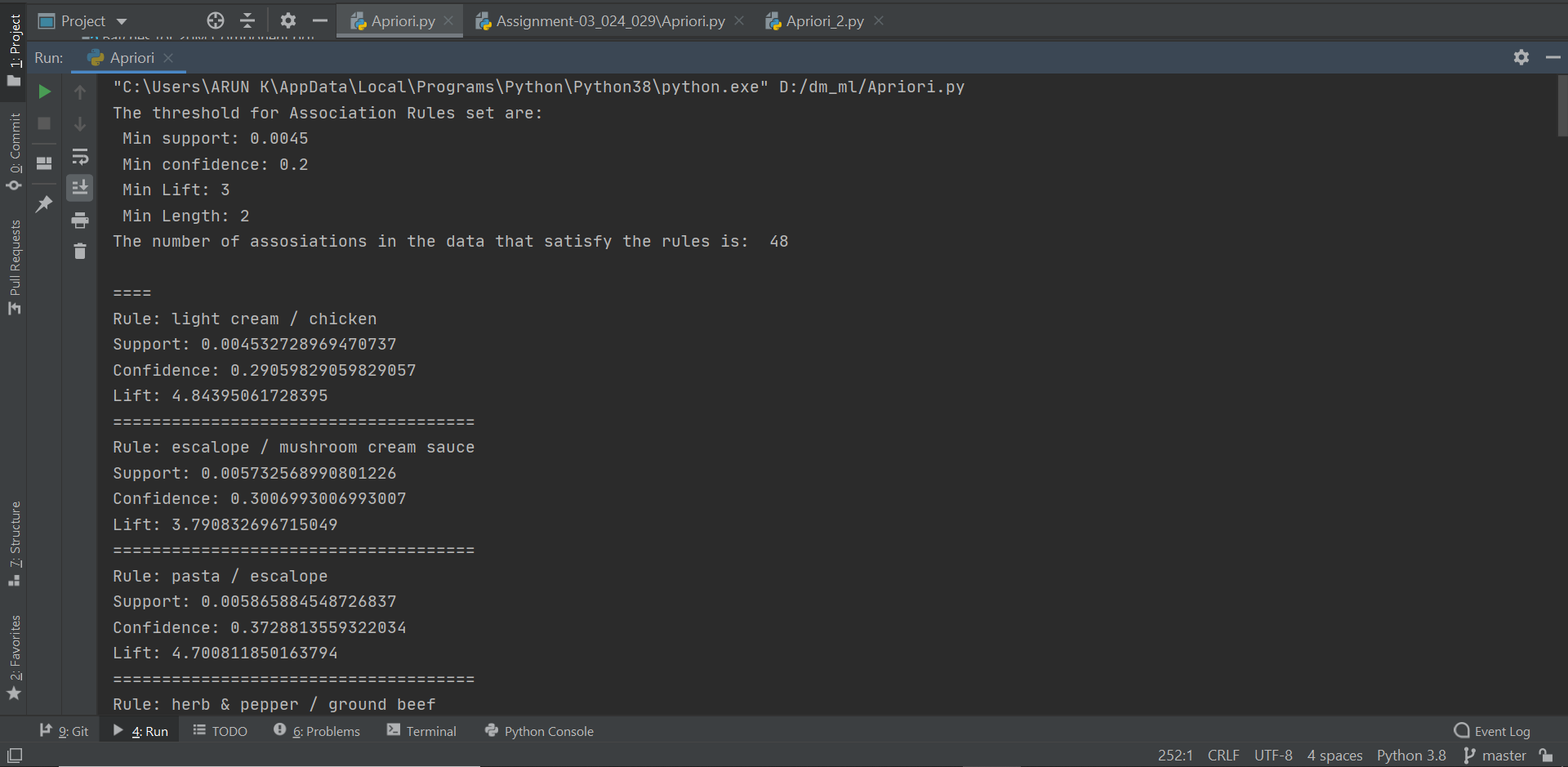
***Disadvantages:***

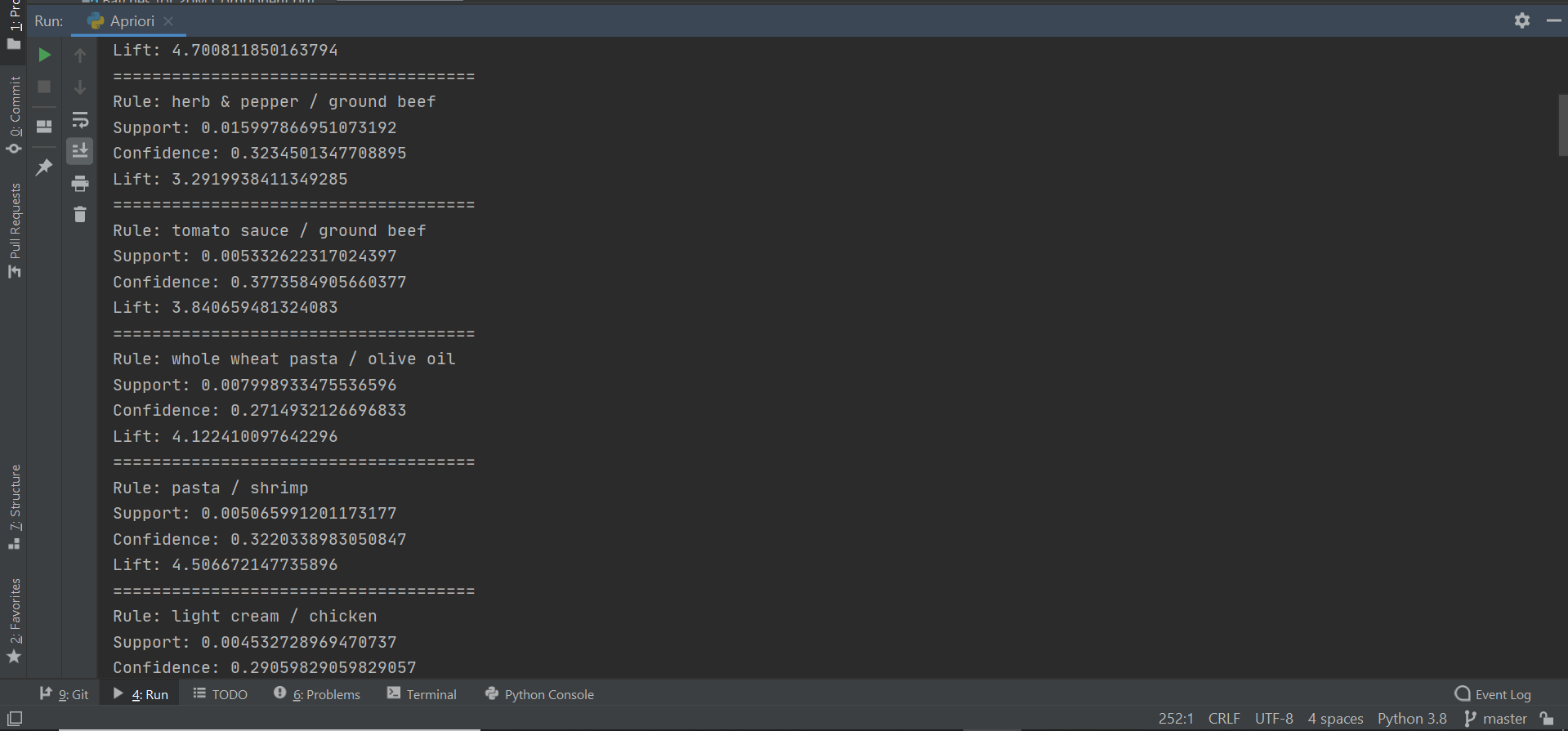
* Assumes transaction database is memory resident.
* Requires many database scans
* The algorithm gets dismissed when various itemsets cannot be prolonged further.
* The performance time is more as consumed in generating contestants every time, it also needs more exploration space and computational cost is too expensive.

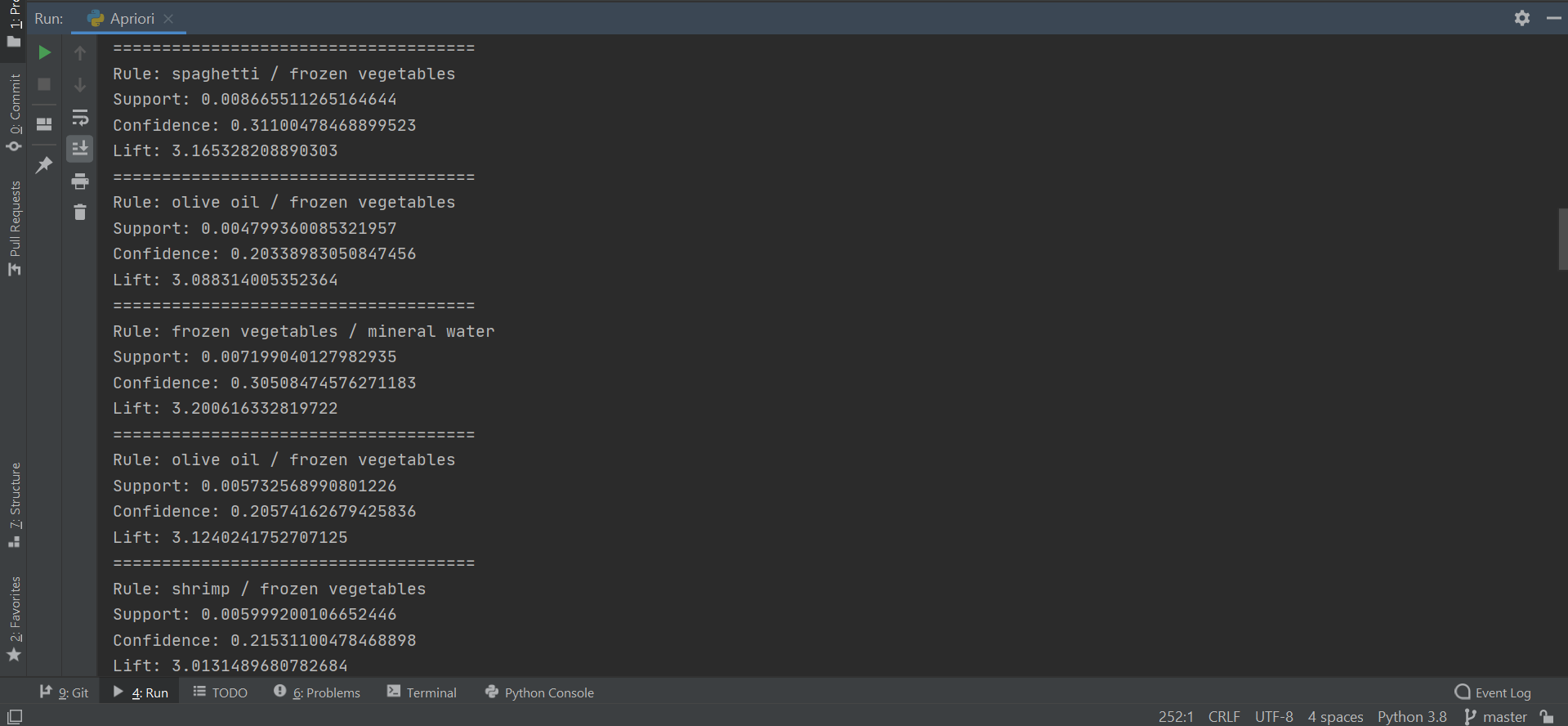
**Applications of the Algorithm:**

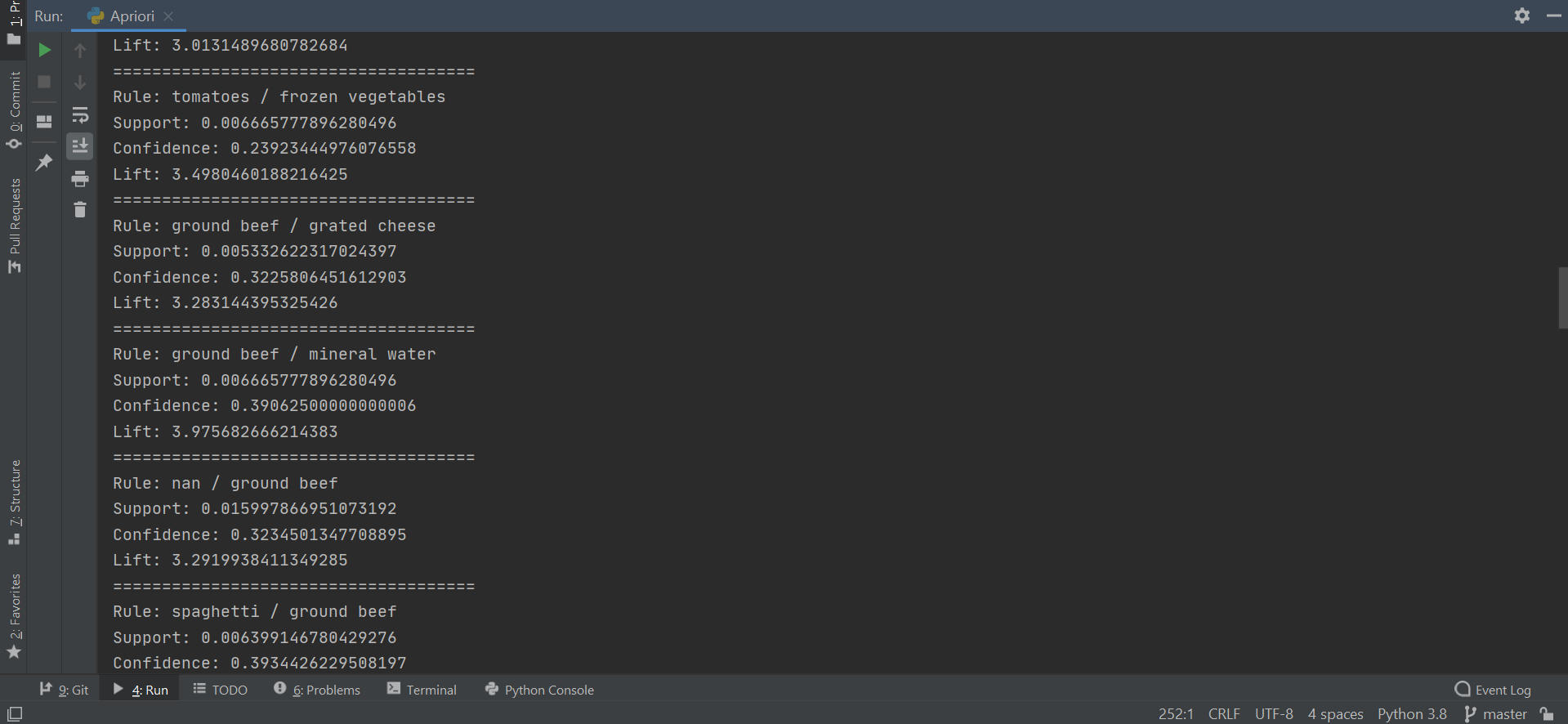
1. In Education Field: Extracting association rules in data mining of admitted students through characteristics and specialties.
2. In the Medical field: For example Analysis of the patient’s database.
3. In Forestry: Analysis of probability and intensity of forest fire with the forest fire data.
4. Apriori is used by many companies like Amazon in the Recommender System and by Google for the auto-complete feature.

**Output: (Screen shots)**









**References:**

* <https://stackabuse.com/association-rule-mining-via-apriori-algorithm-in-python/>
* <https://www.softwaretestinghelp.com/apriori-algorithm/#:~:text=Applications%20Of%20Apriori%20Algorithm,-Some%20fields%20where&text=In%20the%20Medical%20field%3A%20For,for%20the%20auto%2Dcomplete%20feature>.
* <https://www.geeksforgeeks.org/apriori-algorithm/>
* <https://www.slideshare.net/INSOFE/apriori-algorithm-36054672>
* <https://brainly.in/question/2071086>