

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Programme | : | **B.Tech** | Semester | : | **Win Sem 21-22** |
| Course | : | **Web Mining** | Code | : | **CSE3024** |
| Faculty | : | **Dr.Bhuvaneswari A** | Slot | : | **L7+L8** |
| Date | : | **11-04-2022** | Marks | : | **10 Points** |

**NAME: RAHUL GARG  
REGISTER NUMBER: 19BCE1431**

**GOOGLE COLAB LINK:** [**https://colab.research.google.com/drive/1oq\_644Y9HqyCwlEt8glOEAyYUHZfJSFA?usp=sharing**](https://colab.research.google.com/drive/1oq_644Y9HqyCwlEt8glOEAyYUHZfJSFA?usp=sharing)

**Exercise13:Association Rule Mining and Analysis - Apriori Algorithm**

Market-Basket-Analysis ofGrocery Dataset collected from an online grocery billing data.

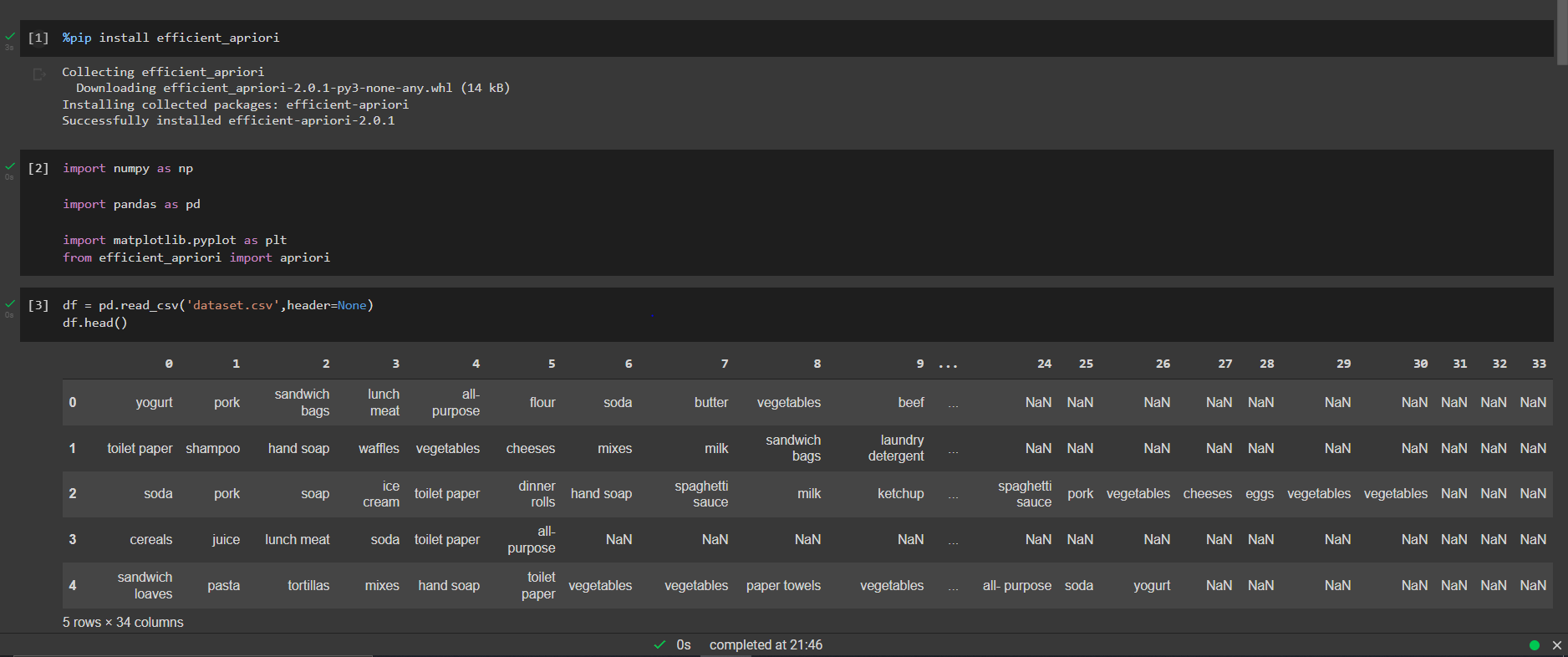
Market Basket Analysis is the analysis of past buying behavior of online customers to find out which are the products that are bought together by the customers. That means to find out the association between various products. If the retail's management can find this association, while placing the products in the shop, these associated products can be put together. Or, when seeing that a customer is buying a product, the salesman can offer the associated product to the customer.

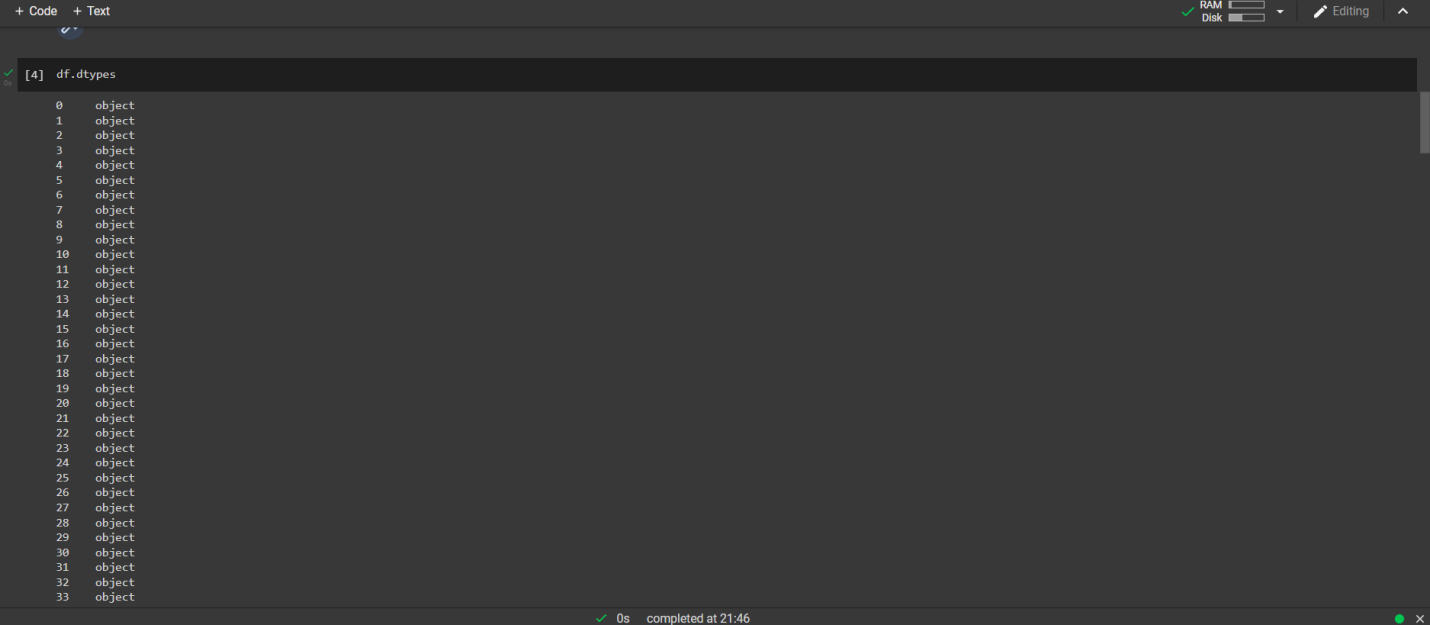
We find this association by Association Rule learning which is a machine-learning rule based approach that generates relationship between variables in a dataset. It has major application in retail industry including e-commerce.

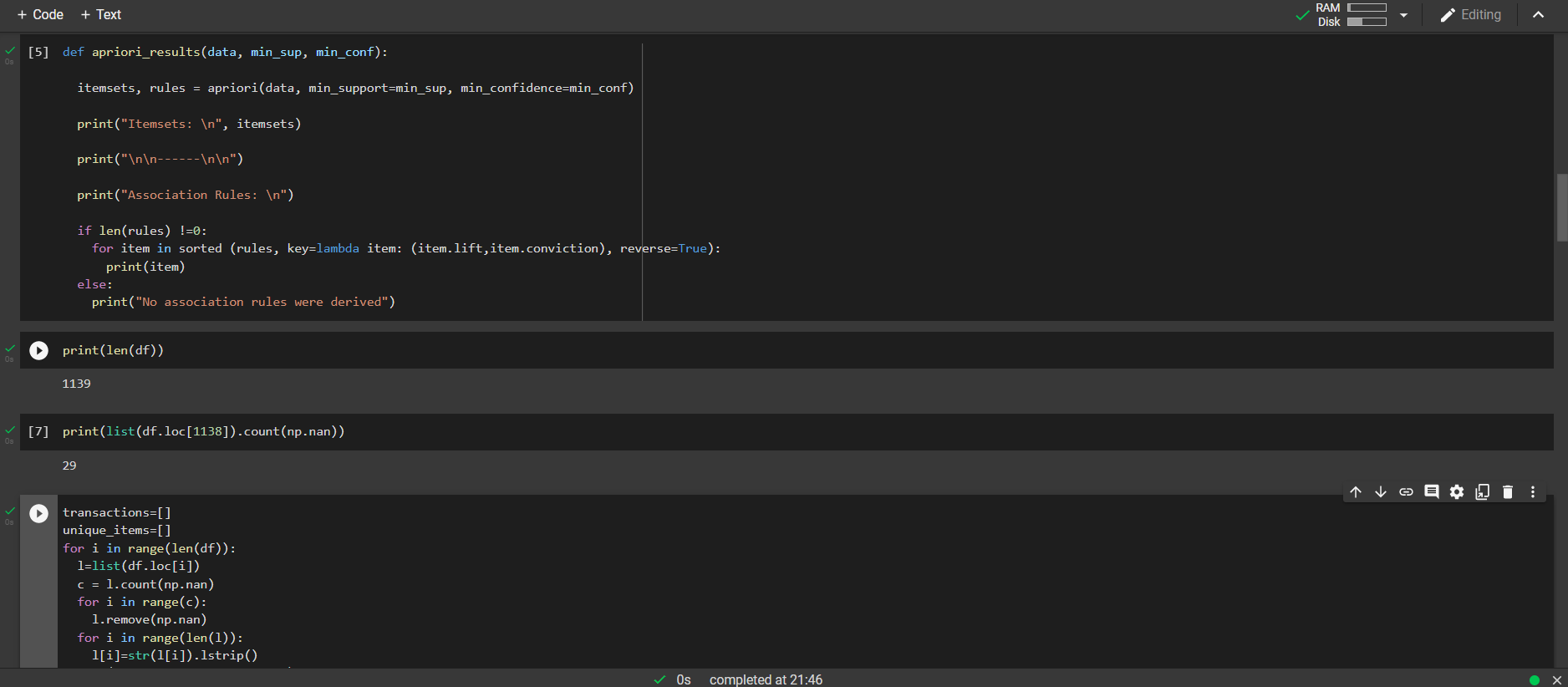
Problem Statement

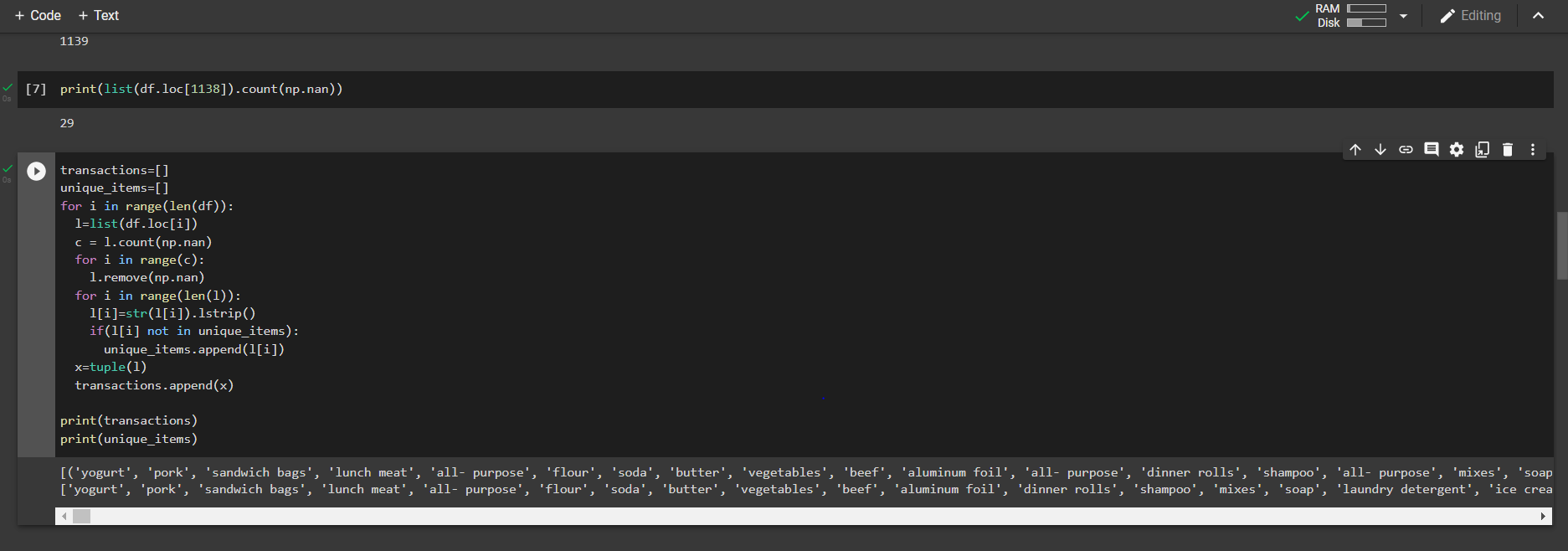
To determine the association between various products in the basket by analysing the customer purchase pattern of multiple items.

**BASIC OUTPUT:**





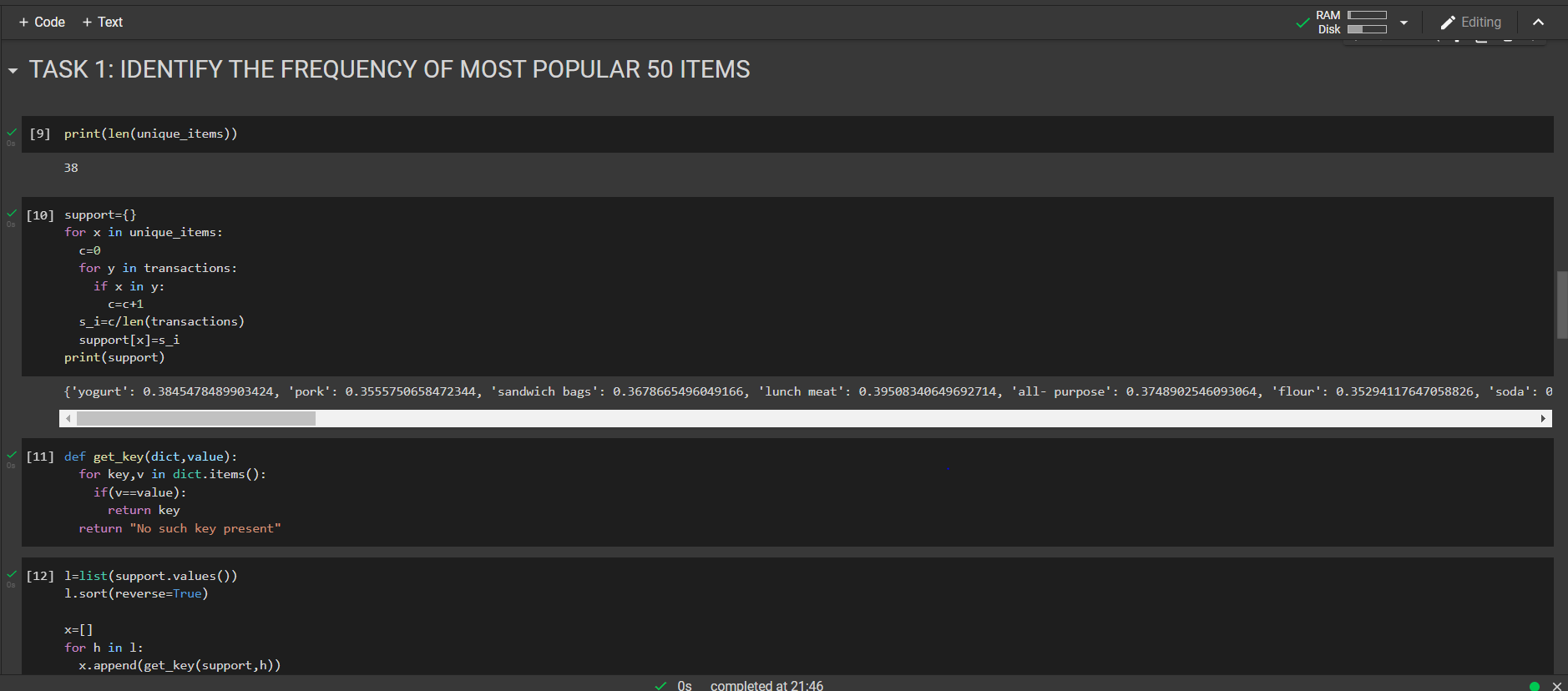


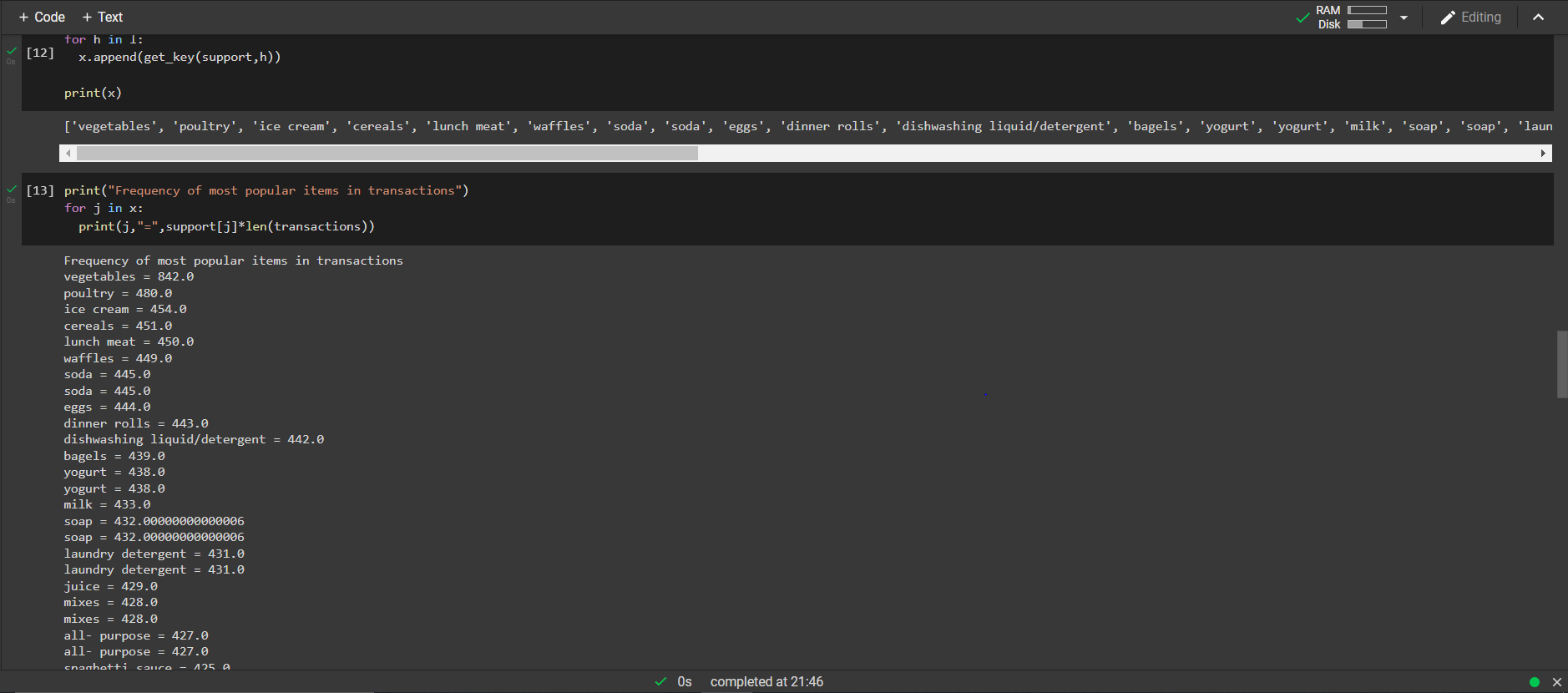


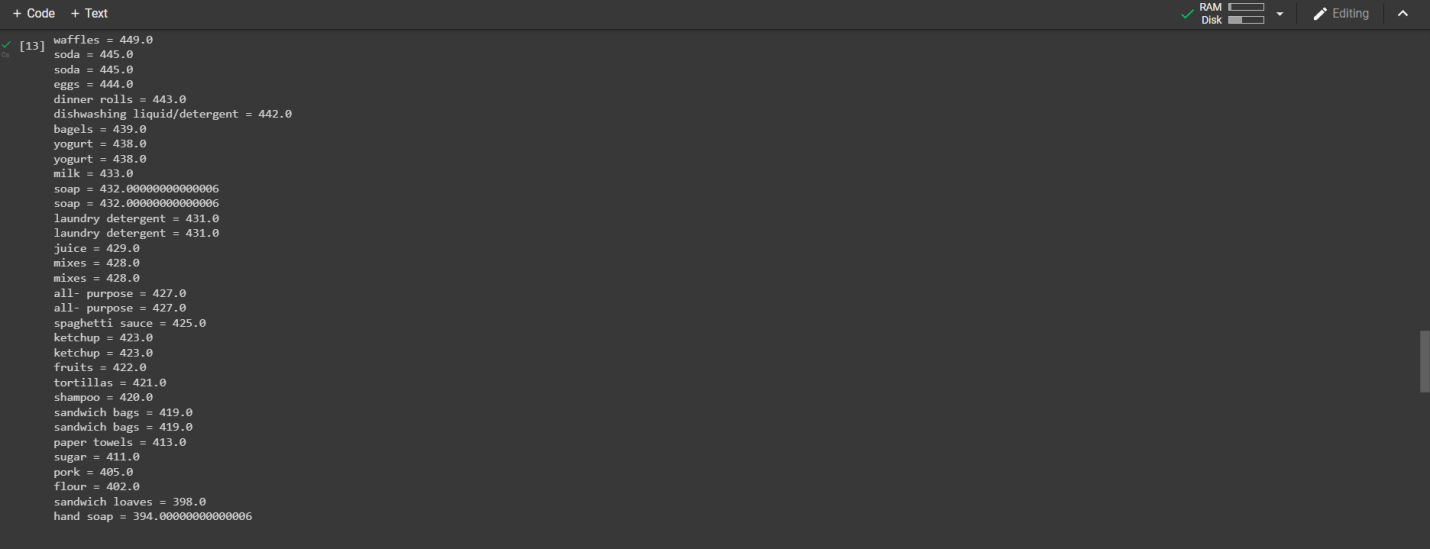
**Task 1: Identify the frequency of most popular 50 items**

Input is dataset file named 'dataset.csv'.

**OUTPUT:**







Task 2: Run aprior algorithm using Python for the following support and confidence

|  |
| --- |
| Case 1 (minimum support=0.15 and minimum confidence=0.6 (60%))  Identify the Itemset and derive association rules.  **OUTPUT:** |
| Case 2 (minimum support=0.3 and minimum confidence=0.7)  Identify the Itemset and derive association rules.  Expected Output:  As per our rules with Min. Confidence of 70%,  **OUTPUT:** |
| Case 3 (minimum support=0.4 and minimum confidence=0.85)  Identify the Itemset and derive association rules.  **OUTPUT:** |

SAMPLE CODE FOR REFERENCE

