# Association Rules

**Instructions:**

Please share your answers filled in-line in the word document. Submit code separately wherever applicable.

Please ensure you update all the details:

**Name: Biswajeet Padhi Batch ID:** 280921

**Topic: Association Rules**

**Grading Guidelines:**

**1. An assignment submission is considered complete only when correct and executable code(s) are submitted along with the documentation explaining the method and results. Failing to submit either of those will be considered an invalid submission and will not be considered for evaluation.**

**2. Assignments submitted after the deadline will affect your grades.**

**Grading:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Ans** | **Date** |  |  | **Ans** | **Date** |
| Correct | On time | A | 100 |  |  |
| 80% & above | On time | B | 85 | Correct | Late |
| 50% & above | On time | C | 75 | 80% & above | Late |
| 50% & below | On time | D | 65 | 50% & above | Late |
|  |  | E | 55 | 50% & below |  |
| Copied/No Submission |  | F | 45 |  |  |

* **Grade A: (>= 90):** When all assignments are submitted on or before the given deadline.
* **Grade B: (>= 80 and < 90):** 
  + When assignments are submitted on time but less than 80% of problems are completed.

(OR)

* + All assignments are submitted after the deadline.
* **Grade C: (>= 70 and < 80):** 
  + When assignments are submitted on time but less than 50% of the problems are completed.

(OR)

* + Less than 80% of problems in the assignments are submitted after the deadline.
* **Grade D: (>= 60 and < 70):**
  + Assignments submitted after the deadline and with 50% or less problems.
* **Grade E: (>= 50 and < 60):** 
  + Less than 30% of problems in the assignments are submitted after the deadline.

(OR)

* + Less than 30% of problems in the assignments are submitted before the deadline.
* **Grade F: (< 50):** No submission (or) malpractice.

Hints:

1. Business Problem
   1. **What is the business objective?**
   2. **Are there any constraints?**
2. Work on each feature of the dataset to create a data dictionary as displayed in the below image**:**



1. Data Pre-processing
   1. Data Cleaning, Feature Engineering, etc.
2. Model Building

4.1 Application of Apriori Algorithm

* 1. Build most frequent item sets and plot the rules
  2. Work on Codes

5.Deployment

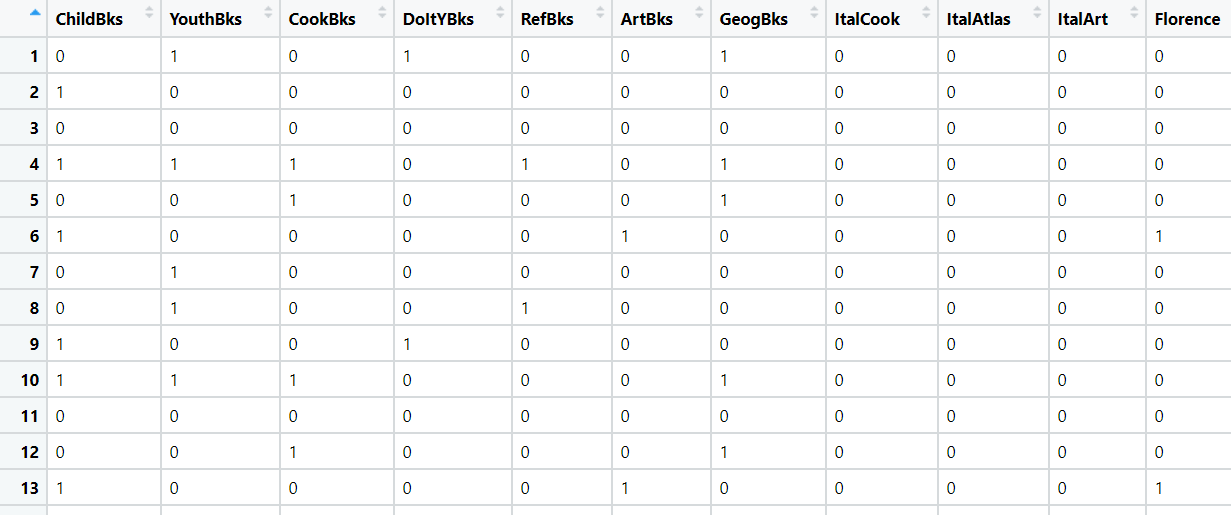
5.1 Deploy solutions

6. Write about the benefits/impact of the solution - in what way does the business (client) benefit from the solution provided?

**Problem Statement: -**

Kitabi Duniya, a famous book store in India, which was established before Independence, the growth of the company was incremental year by year, but due to online selling of books and wide spread Internet access its annual growth started to collapse, seeing sharp downfalls, you as a Data Scientist help this heritage book store gain its popularity back and increase footfall of customers and provide ways the business can improve exponentially, apply Association RuleAlgorithm, explain the rules, and visualize the graphs for clear understanding of solution.

**1.) Books.csv**

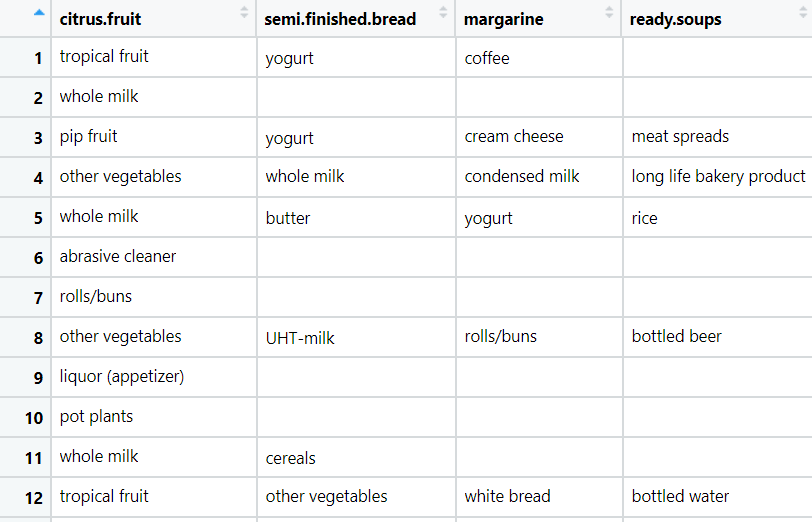


**Problem Statement: -**

The Departmental Store, has gathered the data of the products it sells on a Daily basis.

Using Association Rules concepts, provide the insights on the rules and the plots.

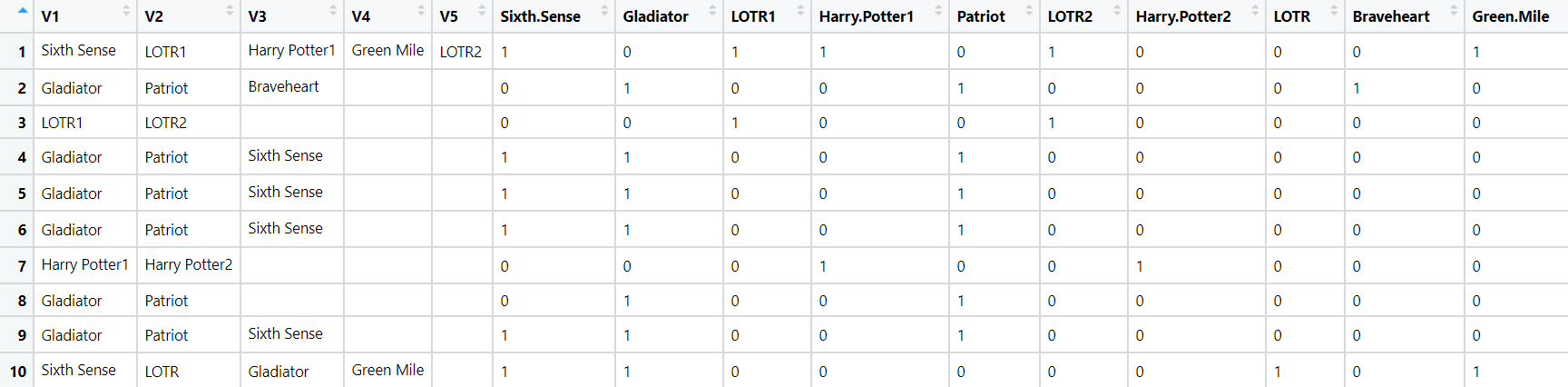
**2.) Groceries.csv**



**Problem Statement: -**

A film distribution company wants to target audience based on their likes and dislikes, you as a Chief Data Scientist Analyze the data and come up with different rules of movie list so that the business objective is achieved.

**3.) my\_movies.csv**



**Problem Statement: -**

A Mobile Phone manufacturing company wants to launch its three brand new phone into the market, but before going with its traditional marketing approach this time it want to analyze the data of its previous model sales in different regions and you have been hired as an Data Scientist to help them out, use the Association rules concept and provide your insights to the company’s marketing team to improve its sales.

**4.) myphonedata.csv**



**Problem Statement: -**

A retail store in India, has its transaction data, and it would like to know the buying pattern of the

consumers in its locality, you have been assigned this task to provide the manager with rules

on how the placement of products needs to be there in shelves so that it can improve the buying

patterns of consumes and increase customer footfall.

**5.) transaction\_retail.csv**

