# 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: MD SABIULLAH Batch ID : PDS 07052024**

**Topic: Association Rules**

# 

**Guidelines:**

**1. An assignment submission is considered complete only when the correct and executable code(s) is 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 a correct submission.**

**2. Ensure that you submit your assignments correctly and in full. Resubmission is not allowed.**

**3. Post the submission you can evaluate your work by referring to the keys provided. (will be available only post the submission).**

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**:**



Data Dictionary :

| Column Name | Description | Data Type | Range | Min Value | Max Value |
| --- | --- | --- | --- | --- | --- |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ChildBks | Children Books Purchased | int64 | 0 - 1 | 0 | 1 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| YouthBks | Youth Books Purchased | int64 | 0 - 1 | 0 | 1 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| CookBks | Cookbooks Purchased | int64 | 0 - 1 | 0 | 1 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| DoItYBks | Do It Yourself Books Purchased | int64 | 0 - 1 | 0 | 1 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| RefBks | Reference Books Purchased | int64 | 0 - 1 | 0 | 1 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ArtBks | Art Books Purchased | int64 | 0 - 1 | 0 | 1 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| GeogBks | Geography Books Purchased | int64 | 0 - 1 | 0 | 1 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ItalCook | Italian Cookbooks Purchased | int64 | 0 - 1 | 0 | 1 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ItalAtlas | Italian Atlases Purchased | int64 | 0 - 1 | 0 | 1 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ItalArt | Italian Art Books Purchased | int64 | 0 - 1 | 0 | 1 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Florence | Books About Florence Purchased | int64 | 0 - 1 | 0 | 1 |

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

4.1 Application of Apriori Algorithm

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

5. Deployment

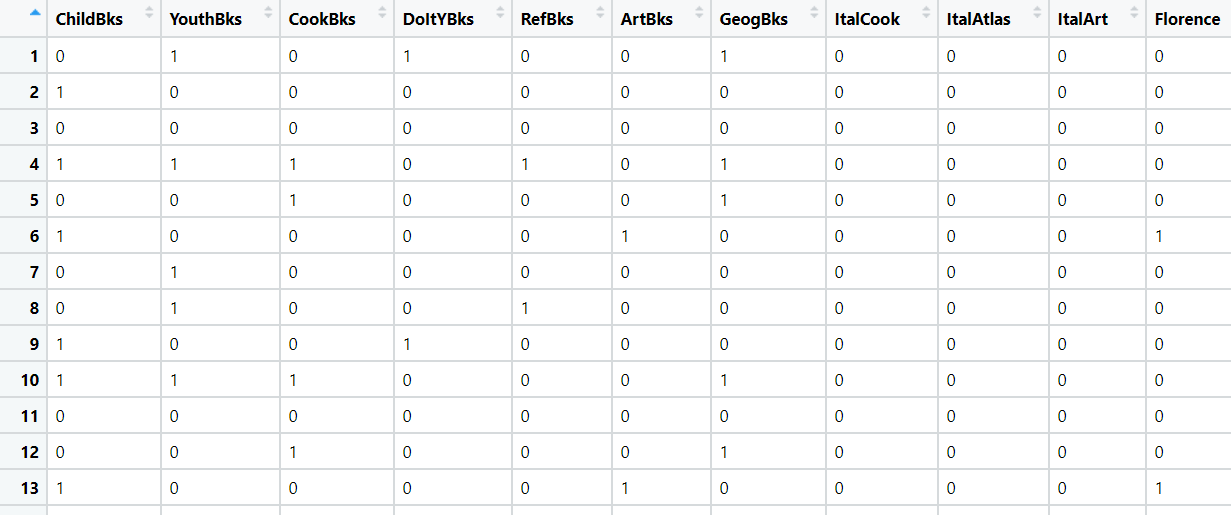
5.1 Deploy solutions

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

**Problem Statement: -**

Q1. Kitabi Duniya, a famous bookstore in India, was established before Independence, the growth of the company was incremental year by year, but due to the online selling of books and widespread Internet access, its annual growth started to collapse. As a Data Scientist, you must help this heritage bookstore gain its popularity back and increase the footfall of customers and provide ways to improve the business exponentially to an expected value at a 25% improvement of the current rate. Apply the pattern mining techniques (Association Rules Algorithm) to identify ways to improve sales. Explain the rules (patterns) identified, and visually represent the rules in graphs for a clear understanding of the solution.

**1.) Data: Books.csv**



**Business Objective:**

**Business Constraints:**

**Success Criteria:**

**Business Success Criteria:**

**ML Success Criteria:**

**Economic Success Criteria:**

**Questions to Ignite your Thinking process:**

Q1. Which library/package is used for the Association rules algorithm?

Q2. Which functions are used in the Association rules algorithm?

Q3. What is the keyword used to import any package to the Python session’s memory?

Q4. What type of data is usually worked in Association rules?

Q5. Association rules are also named as

Q6. What is the IF part called in an Association rule?

Q7. What is the THEN part called in an Association rule?

Q8. In which sector is the Association rules algorithm mainly used?

Q9. What do slotting fees mean?

Q10. How is Support calculated in the Association rules algorithm?

Q11. What is the drawback of Support in Association rules algorithm?

Q12. To remove the infrequent items from data which algorithm is used?

Q13. The drawback of Support is captured by

Q14. How is confidence calculated in the Association rules algorithm?

Q15. What is the drawback of Confidence in Association rules algorithm?

Q16. How is the Lift ratio calculated?

Q17. What is the threshold value of the Lift ratio of a rule to declare it as a good Association rule?