**Your Name: Purnabhishek Sripathi**

**Grading Policy: In order to get credit for each screenshot,** you **MUST** have the repository named as ***BAN540\_YourLastName*** with**GroceryStoreList\_YourLastName** and**ProblemSolving\_Association\_YourLastName** in it. The repository and data files **MUST** be clearly shown in your screenshot as my examples below. you will get ZERO for that particular screenshot which doesn’t show them clearly.

**Screenshot #1 (6 points)**

A screenshot of a computer

Description automatically generated

**Screenshot #2 (6 points)**

A screenshot of a computer

Description automatically generated

1. **Question:** Based on the list of rules shown in the previous step, please identify the rule with the largest **confidence** value, and type in your answer here **(3 points)**.

**Answer**:

* The rule with the greatest confidence value among the provided rules is [hummus] --> [peroni] (confidence: 0.826). This criterion sticks out in the study because it suggests that there is a good chance that consumers who buy hummus will also probably buy Peroni.
* A strong and consistent pattern in consumer purchasing behavior is indicated by such a high degree of confidence, and this provides insightful information for targeted marketing and product placement tactics in the grocery store environment.

1. **Question:** Interpret the rule that you’ve identified for the above Question. Specifically, please describe what this rule tells us about the relationship between the antecedent item set (i.e., Premises) and the consequent item set (i.e., Conclusion). Please type in your answer here **(3 points)**.

**Answer:**

We can learn a lot about customer behavior from the association rule ([hummus] --> [peroni], confidence: 0.826). Customers who purchase hummus are significantly more likely to purchase Peroni, according to this guideline. More specifically, in around 82.6% of hummus-related purchases, Peroni is also purchased. There may be a complementary or matching quality between these two products, as indicated by this relationship, which highlights a trend in consumer preferences. When developing targeted marketing campaigns and improving product placement, an understanding of these patterns can be quite helpful.

**Screenshot #3 (6 points)**

A screenshot of a computer

Description automatically generated

1. **Question**: What insights can the grocery store obtain about the purchasers of peroni? Please type in your answer here **(6 points)**.

**Answer**:

A grocery store can obtain useful information for strategic planning from the Peroni purchase data:

* Purchases of Peroni and Hummus are strongly correlated; approximately 82.6% of consumers purchase both ([hummus] --> [peroni], confidence: 0.826). This shows that there may be room for both product sales to be increased through cross-merchandising and focused marketing.
* Likewise, around 80.8% of soda buyers also frequently buy Peroni ([soda] --> [peroni], confidence: 0.808). To capitalize on this trend, the business may use this information to inform the creation of packaged offers or exclusive discounts.
* Furthermore, Peroni is significantly associated with both crackers (75% confidence) and risotto (66.6% confidence), suggesting that both pairings are well-liked by consumers. This information can be used by the store to better arrange products and run promotions, which could enhance consumer satisfaction and increase sales.

**Screenshot #4 (8 points)**

*<Please add your screenshot #4 here>*

A screenshot of a computer

Description automatically generated

1. **Question:** How many rules satisfy this criterion (**min support value = 0.30**) **(3 points)**? By increasing the min support value from 0.25 to 0.30, do you get more or less rules in RapidMiner **(3 points)**?

**Answer:**

We saw a decrease in the number of qualifying rules when we raised the minimum support value in our RapidMiner study from 0.25 to 0.30; only 2 rules met the higher criteria. This result is a well-known example of how support values affect association rule mining. A more discerning rule set that concentrates on the most common item relationships is produced by increasing the support value. Although this method offers a more condensed and possibly more applicable set of guidelines, it also narrows the range of correlations found. One of the most important factors in data analytics is striking a balance between the number and quality of rules.

A screenshot of a computer

Description automatically generated

1. **Question:** How many rules satisfy this criterion now (**min support value = 0.20**) **(3 points)**? By decreasing the support value from 0.25 to 0.20, do you get rules in RapidMiner **(3 points)**?

**Answer:**

* Setting the minimum support value to 0.20 in our RapidMiner data analysis produced a significant change. A total of forty rules were found that satisfy this requirement. This rise in rules over the 0.25 support value scenario highlights a fundamental idea in data mining: when the support threshold falls, more item combinations are included, resulting in a bigger collection of rules.
* Although it might include fewer common and possibly less significant relationships, this growth is indicative of capturing a wider range of associations. To properly comprehend and use association rules in data-driven decision-making, it is imperative to comprehend this dynamic.

**Now you are done with Problem-Solving Assignment 2, please make sure to fill in your name on the first page of this document, and rename this document to reflect your own last name before submitting it to Canvas.**