**D212 Performance Assessment Task 3**

Renzo Espinoza

College of Information Technology, Western Governors University

D212: Data Mining III

Dr. Keiona Middleton

February 13, 2023

**A1:Proposal of Question**

A question relevant to a real-world organizational situation that I will answer by using market basket analysis is: can I find which items customers usually buy together?

**A2:Defined Goal**

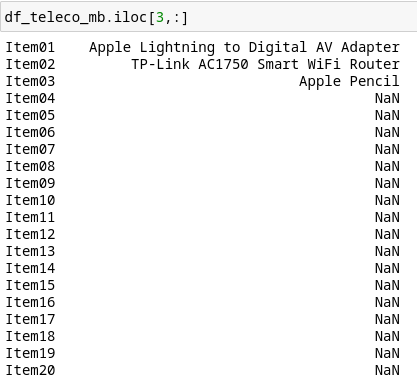
The goal of my analysis is to identify items of interest that the telecommunications company can put on discount to attempt to retain customers. The dataset provided to me contains the purchase history of around 7,500 customers and I believe this analysis is well within the scope of the scenario and the data.

**B1:Explanation of Market Basket**

According to Kadlaskar (2021), market basket analysis analyzes datasets containing past customer transactions by finding associations between the items a customer placed in their shopping cart when completing a transaction. I expect this analysis to result in finding common combinations of items that the telecommunications company offers, so they can then offer discounts on those popular items.

**B2:Transaction Example**

Here is an example of a transaction from the dataset:

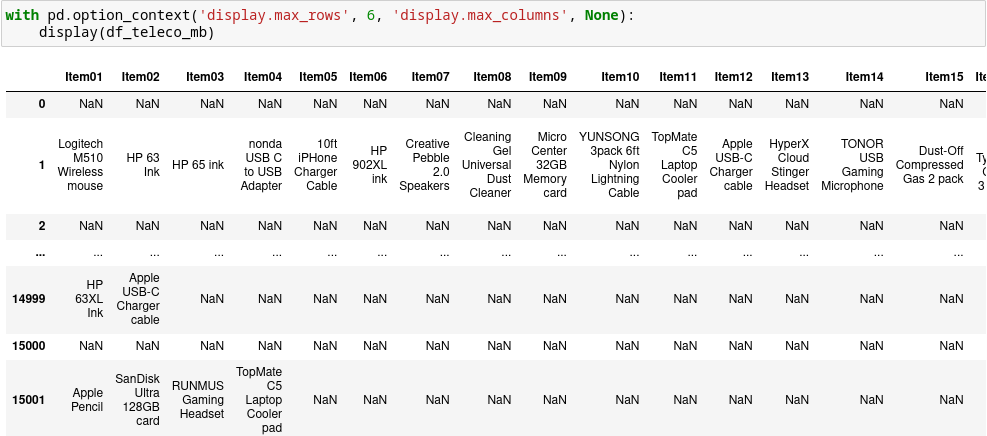


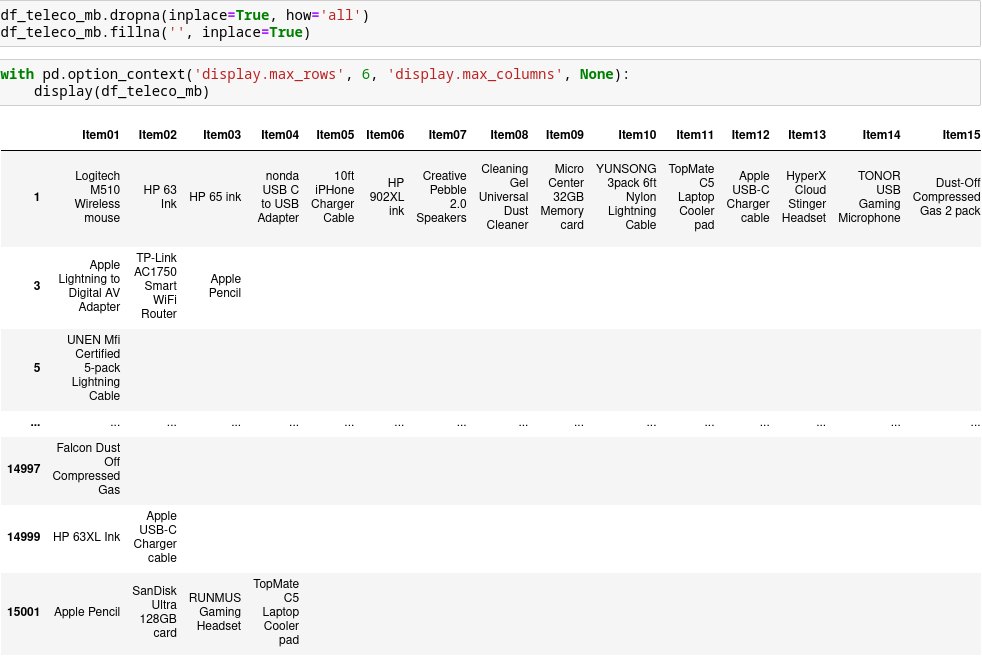
The transactions in this dataset contain twenty items at max, the transaction shown above contains only three items. The first and third items added to their cart are directly related as they are both Apple products. The market basket analysis that I will perform should uncover whether this type of association is frequent in the dataset.

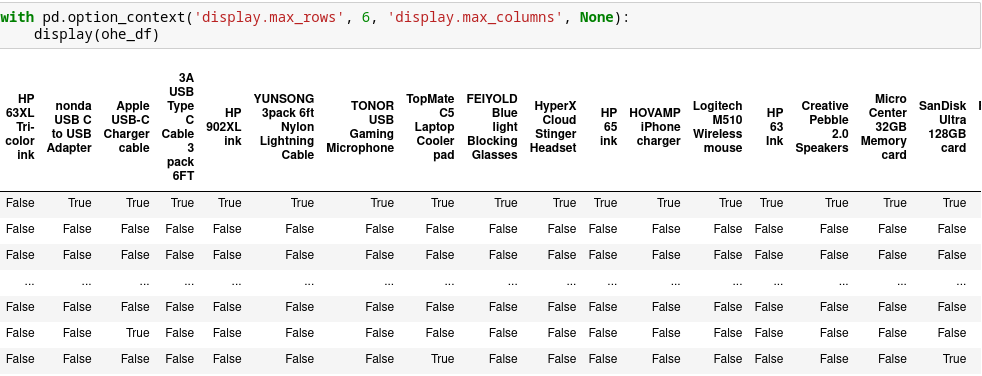
**B3:Market Basket Assumption**

The algorithm that I will use in my market basket analysis is the Apriori algorithm. The key assumption of this algorithm is that the subset of frequent sets are frequent as well. With this principle, the algorithm ruses it to remove sets that are infrequent and keep sets that do appear frequently.

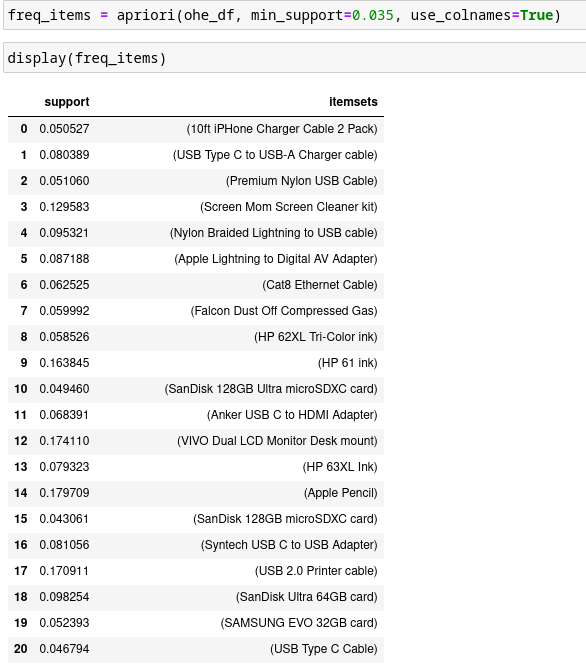
**C1:Transforming the Dataset**

To prepare the data for market basket analysis, I first needed to remove missing data. The dataset contains empty rows in between each transaction row, so I used the ‘dropna’ method from the Pandas library to remove rows entirely consisting of nulls.



The next step was to reformat the data so it can be used by the ‘MLxtend’ library’s Apriori function. The function requires the data to contain only one’s and zero’s or boolean values. So I performed one-hot encoding to transform the dataset given to me.

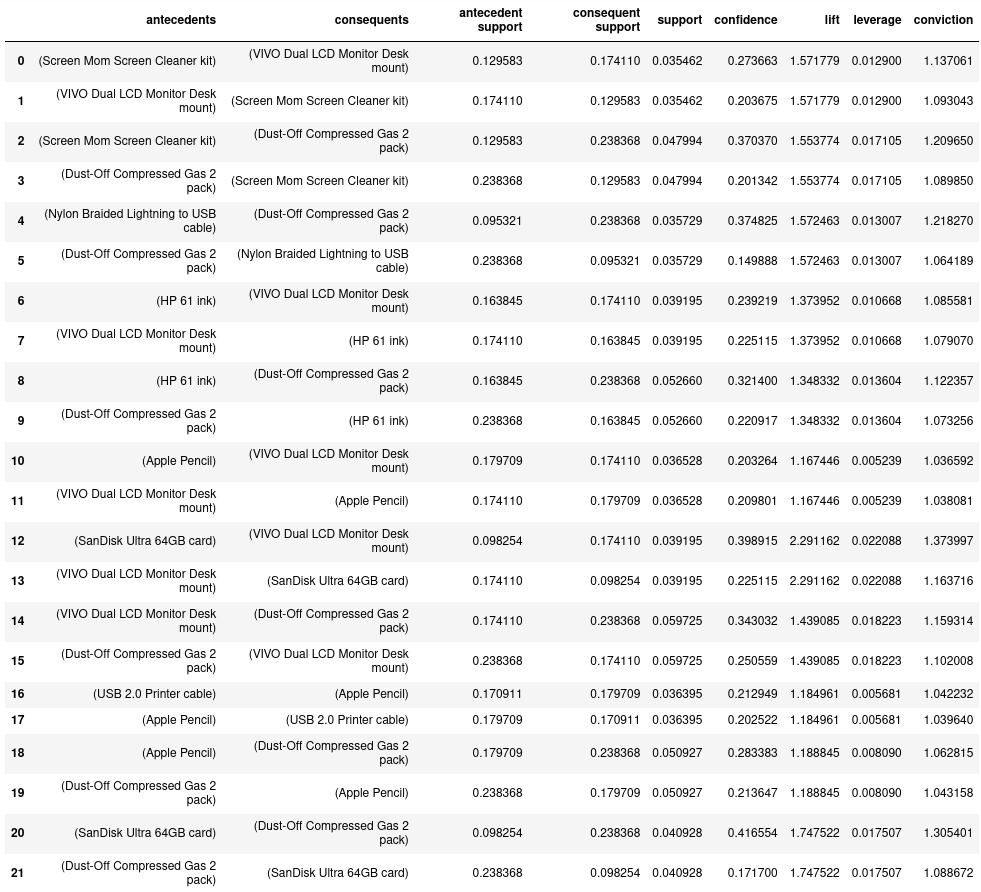
**C2:Code Execution**

****



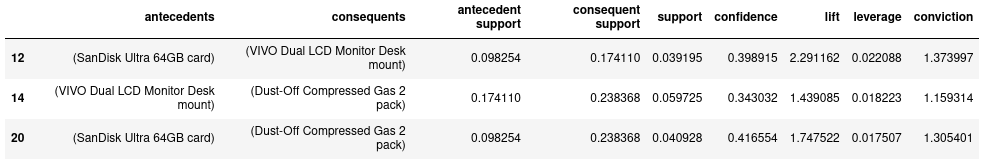
**C3:Association Rules Table**





**C4:Top Three Rules**

The top three rules that I have identified are rows 12, 14, and 20. Item set 12 has the highest lift, leverage, and conviction of all association rules. Row 14 has the highest support out of all the generated association rules. Finally, the 20th association rule has the highest confidence.

**D1:Significance of Support, Lift, and Confidence Summary**

The top three rules I selected have fairly high values for support, lift, and confidence, especially considering the datatset has 119 unique items and 7501 transactions. The high confidence and lift values show that the association between intra-set items are strong, and it is likely that the antecedent items leads to the consequent item. The high support means that the item sets appear frequently in transactions.

**D2:Practical Significance of Findings**

I believe my findings are practically significant, the lift values of the association rules are exceptionally high, so offering a discount on the antecedent items should theoretically result in the consequent items being bought more often. A limitation of my analysis is that my support threshold was fairly high when applying the Apriori alogrithm, resulting in my top three item sets consisting of cheap accessory items that are more likely to appear in the carts of shoppers than pricier items. For future analyses, I would use a threshold amount or metric and request prices to be included in the dataset.

**D3:Course of Action**

My recommended course of action would be to offer discounts on items identified in the market basket analysis. I would also recommend adding the prices of each item in a transaction, to help prioritize more expensive items.

**E:Panopto Recording**

**F & G: Sources**

Kadlaskar, A. (2021, October 2). *Market basket analysis: Guide on market basket analysis*. Analytics Vidhya. Retrieved February 7, 2023, from https://www.analyticsvidhya.com/blog/2021/10/a-comprehensive-guide-on-market-basket-analysis/