**D212 PA**

Jon Isak

D212: Data Mining II

Contents

[**D212 Task 3 3**](#_3xni6yjcaxld)

[**Part I: Research Question 3**](#_on428cwqsz8k)

[**Part II: Market Basket Justification 3**](#_mytzr5pv706l)

[**Part III: Data Preparation and Analysis 4**](#_103541cwhku6)

[**Part IV: Data Summary and Implications 6**](#_gbpyfqak8j0a)

[**Part V: Attachments 7**](#_ixwt1kkbz9xk)

## 

## D212 Task 3

## Part I: Research Question

A1. This data analysis will look into the question, What are the most common item combinations purchased by customers?

A2. The goal of the data analysis is to use market basket analysis on the dataset to uncover relationships between the various products that customers buy. By understanding these associations, the company can improve marketing and create bundles that cater to customer preferences, thus improving customer satisfaction and potentially increasing revenue.

## Part II: Market Basket Justification

B1. The market basket analysis will take a look at the items purchased by a customer and find associations between items purchased together. The market basket analysis uses association rules to identify items frequently bought together and association rules. It should make discoveries in metrics like support, lift and confidence that could show the Logitech M510 Wireless mouse and HP 63 Ink are often bought together and customers that buy the mouse are more likely to purchase the ink (Kadlaskar, n.d.).

B2. One example of transactions in the data set is one customer purchased item01: ‘Logitech M510 Wireless mouse’ and Item02: ‘HP 63 Ink’.

B3. One of the core assumptions of market basket analysis is that customer purchasing behavior exhibits a pattern of association, where the purchase of one item increases the probability of purchasing another. (Indeed, 2022).

## Part III: Data Preparation and Analysis

C1. The data is attached.

A screenshot of a computer program

Description automatically generated

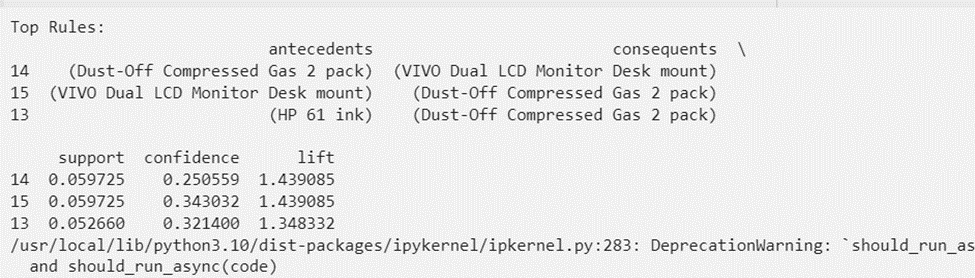
C2. The code used is attached in the screenshot, and the results of the code are below the code.

C3. The values for the support, lift, and confidence can be found in the screenshot.

A screenshot of a computer program

Description automatically generated

C4. The Apriori algorithm extracted three significant rules from the dataset. The rules were sorted based on the lift in descending order in order to get the top 3 rules. Both Rules 1 and 2 exhibit a relationship between Dust-Off Compressed Gas 2 pack and VIVO Dual LCD Monitor Desk mount, signifying a reciprocal association between these items. The support value of approximately 0.05 indicates the frequency of occurrence for these item combinations. Specifically, Rule 1 shows that when customers purchase Dust-Off Compressed Gas 2 pack, there is a 25% probability of them also buying VIVO Dual LCD Monitor Desk mount, while Rule 2 suggests a higher likelihood (34%) of purchasing Dust-Off Compressed Gas 2 pack after buying VIVO Dual LCD Monitor Desk mount. The lift value, greater than 1, signifies a positive association between these products. Rule 3 indicates an association between HP 61 ink and Dust-Off Compressed Gas 2 pack. The support value of approximately 0.05 highlights the frequency of these items occurring together. Customers purchasing HP 61 ink show a 32% probability of also purchasing Dust-Off Compressed Gas 2 pack. The lift value greater than 1.34 implies a stronger association between HP 61 ink and Dust-Off Compressed Gas 2 pack.



## Part IV: Data Summary and Implications

D1. The support metric denotes the frequency of items purchased together, ranging from approximately 0.05 to 0.6 in our dataset. This range reflects varying degrees of association, where lower support values signify less frequent item combinations, while higher values indicate more frequent co-occurrences. Confidence, spanning from around 0.214 to 0.343 in our analysis, reveals the conditional probability of customers purchasing the consequent item given the antecedent item's presence. These values showcase the strength of association between items, demonstrating the likelihood of customers buying the consequent item after buying the antecedent. Lift measures the strength of association between antecedent and consequent items, accounting for their individual support. The lift values, ranging between approximately 1.189 and 1.439, signify the departure from independence between item pairs. Lift values exceeding 1 indicate a positive association, indicating that these item combinations are more likely to be purchased together than if they were bought independently, further reinforcing the strength of their relationship.

D2. The practical significance of the findings from the analysis comes from the support, lift, and confidence metrics. The analysis helps the company understand customer behavior. The association between the Dust-Off Compressed Gas 2 pack and VIVO Dual LCD Monitor Desk mount and the association between HP 61 ink and Dust-Off Compressed Gas 2 pack gas gives insight on customer preferences and interests. This information can be used to improve marketing strategies and increase sales. The product relationships with high confidence can be used to market products more effectively to customers through cross-selling.

D3. Based on insights from the analysis, the company can use the information to improve their marketing strategies. The company can create product bundles with the Dust-Off Compressed Gas 2 pack and VIVO Dual LCD Monitor Desk mount to increase sales and provide convenience to customers. The company can also use the high confidence relationships to promote the product through cross-selling through product recommendations. The company can continue to track transactions and customer behavior to continue discovering associations. With continuous monitoring, the company can also create email marketing campaigns to promote products based on previous purchases to increase repeat purchases.

## Part V: Attachments

E. https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=d9cc5568-d3ac-483d-b522-b0d80004429a

F.

Rasbt. (n.d.). *TransactionEncoder - mlxtend*. GitHub Pages. Retrieved December 13, 2023, from https://rasbt.github.io/mlxtend/user\_guide/preprocessing/TransactionEncoder/

G.

Indeed. (2022, October 12). *FAQ: What Is Market Basket Analysis? (Types Plus Examples)*. Indeed. Retrieved December 12, 2023, from https://sg.indeed.com/career-advice/career-development/market-basket-analysis

Kadlaskar, A. (n.d.). *Market Basket Analysis: A Comprehensive Guide for Businesses*. Analytics Vidhya. Retrieved December 13, 2023, from https://www.analyticsvidhya.com/blog/2021/10/a-comprehensive-guide-on-market-basket-analysis/