Performance Assessment

WGU | D212

D212 Part 3

Narcisse, Laurie

2024

Contents

[**Part I: Research Question** 2](#_Toc157593363)

[A.  Describe the purpose of your data mining report by doing the following: 2](#_Toc157593364)

[1.  Propose one question relevant to a real-world organizational situation that you will answer using market basket analysis. 2](#_Toc157593365)

[2.  Define one goal of the data analysis. Ensure your goal is reasonable within the scope of the selected scenario and is represented in the available data. 2](#_Toc157593366)

[**Part II: Market Basket Justification** 2](#_Toc157593367)

[B.  Explain the reasons for using market basket analysis by doing the following: 2](#_Toc157593368)

[1.  Explain how market basket analyzes the selected data set. Include expected outcomes. 2](#_Toc157593369)

[2.  Provide one example of transactions in the data set. 2](#_Toc157593370)

[3.  Summarize one assumption of market basket analysis. 3](#_Toc157593371)

[**Part III: Data Preparation and Analysis** 3](#_Toc157593372)

[C.  Prepare and perform market basket analysis by doing the following: 3](#_Toc157593373)

[1.  Transform the data set to make it suitable for market basket analysis. Include a copy of the cleaned data set. 3](#_Toc157593374)

[2.  Execute the code used to generate association rules with the Apriori algorithm. Provide screenshots that demonstrate that the code is error free. 11](#_Toc157593375)

[3.  Provide values for the support, lift, and confidence of the association rules table. 13](#_Toc157593376)

[4.  Explain the top three relevant rules generated by the Apriori algorithm. Include a screenshot of the top three relevant rules. 13](#_Toc157593377)

[**Part IV: Data Summary and Implications** 17](#_Toc157593378)

[D.  Summarize your data analysis by doing the following: 17](#_Toc157593379)

[1.  Summarize the significance of support, lift, and confidence from the results of the analysis. 17](#_Toc157593380)

[2.  Discuss the practical significance of your findings from the analysis. 17](#_Toc157593381)

[3.  Recommend a course of action for the real-world organizational situation from part A1 based on the results from part D1. 18](#_Toc157593382)

[**Part V: Attachments** 19](#_Toc157593383)

[E.  Provide a Panopto video recording that includes the presenter and a vocalized demonstration showing all code used, the code being executed, and the results of all code used in the task. 19](#_Toc157593384)

[F.  Record all web sources you used to acquire data or segments of third-party code to support the application. Ensure the web sources are reliable. 19](#_Toc157593385)

# **Part I: Research Question**

## A.  Describe the purpose of your data mining report by doing the following:

### 1.  Propose **one** question relevant to a real-world organizational situation that you will answer using market basket analysis.

What items are frequently purchased together by customers in the telecommunications industry?

### 2.  Define **one** goal of the data analysis. Ensure your goal is reasonable within the scope of the selected scenario and is represented in the available data.

Identify associations between items in customer purchases to provide insights for creating targeted discount offers.

# **Part II: Market Basket Justification**

## B.  Explain the reasons for using market basket analysis by doing the following:

### 1.  Explain how market basket analyzes the selected data set. Include expected outcomes.

In my analysis using the Apriori algorithm for market basket analysis, I aim to uncover associations and patterns within my dataset. The Apriori algorithm works by initially identifying frequent individual items in the dataset. It then systematically explores larger itemsets, ensuring that they are also frequent based on a predefined support threshold. The process involves generating itemsets, checking their frequency in the dataset, and iteratively narrowing down to those that meet the support criteria.

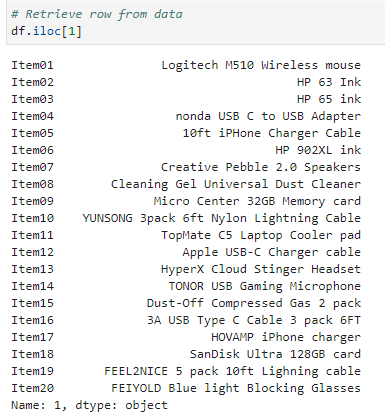
The output will reveal association rules, indicating the likelihood of certain items co-occurring. These rules consist of antecedents and consequents, where antecedents represent the items already in the basket, and consequents represent the items likely to be added to the basket. This information allows me to analyze why, what, and how customers purchase items.

Expected Outcomes: Discovering item associations can help the telecommunications company understand which products customers typically buy together, enabling the creation of targeted discounts to retain customers.

### 2.  Provide **one** example of transactions in the data set.

Here are a few transactions from the dataset (Transaction # does not correlate with ipynb, also it is not complete transaction):

* Transaction 1: Logitech M510 Wireless mouse, HP 63 Ink, HP 65 ink
* Transaction 2: Apple Lightning to Digital AV Adapter, TP-Link AC1750 Smart WiFi Router, Apple Pencil
* Transaction 3: Cat8 Ethernet Cable, HP 65 ink



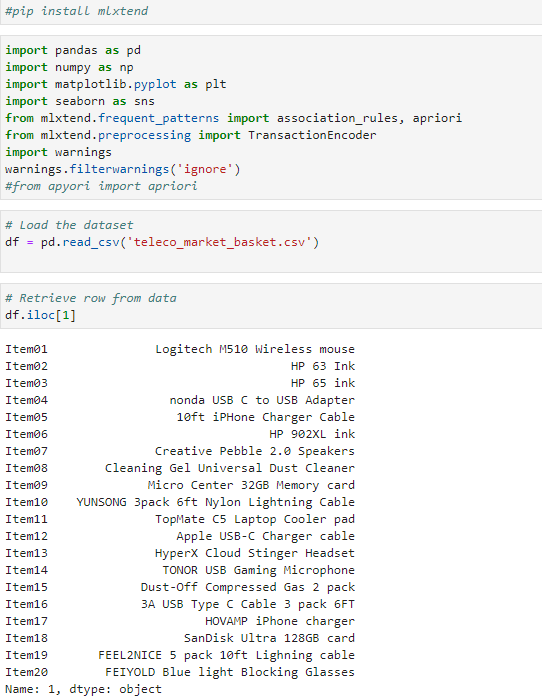
### 3.  Summarize **one** assumption of market basket analysis.

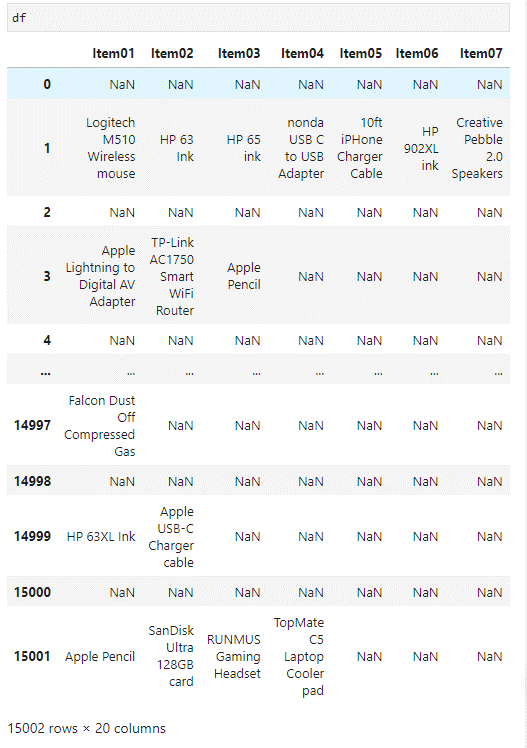
One assumption is that items purchased together in the past are likely to be purchased together in the future. It assumes that there is a meaningful association between items based on historical data.

# **Part III: Data Preparation and Analysis**

## C.  Prepare and perform market basket analysis by doing the following:

### 1.  Transform the data set to make it suitable for market basket analysis. Include a copy of the cleaned data set.



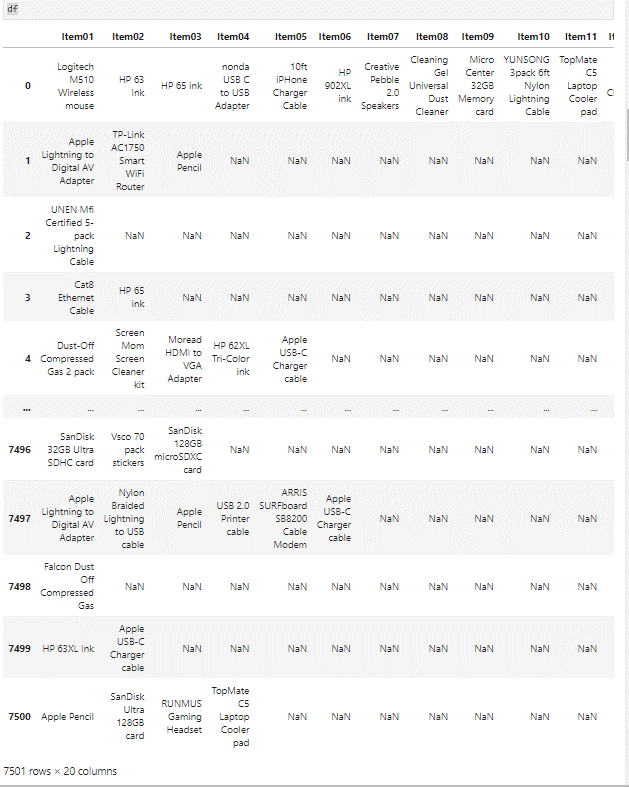


A screenshot of a computer

Description automatically generated

A white rectangular object with a white rectangle

Description automatically generated



A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

### 2.  Execute the code used to generate association rules with the Apriori algorithm. Provide screenshots that demonstrate that the code is error free.

A screenshot of a computer program

Description automatically generated

A screenshot of a computer

Description automatically generated

### 3.  Provide values for the support, lift, and confidence of the association rules table.

A screenshot of a computer

Description automatically generated

### 4.  Explain the top **three** relevant rules generated by the Apriori algorithm. Include a screenshot of the top three relevant rules.

A screenshot of a computer

Description automatically generated

 A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

In the first rule, when customers purchase the SanDisk Ultra 64GB card, there's a 39.89% likelihood that they will also buy the VIVO Dual LCD Monitor Desk mount. This association is relatively strong, as indicated by a lift of 2.29, meaning the purchase of the card is around 2.29 times more likely when the monitor mount is also bought. Additionally, the conviction of 1.37 suggests that customers are 37.40% more likely to buy the SanDisk Ultra 64GB card without the monitor mount.

Moving to the second rule involving FEIYOLD Blue light Blocking Glasses and VIVO Dual LCD Monitor Desk mount, there's a 34.82% confidence level, indicating a moderate association. The lift of 1.99 suggests that the purchase of the glasses is approximately twice as likely when the monitor mount is also in the transaction. The conviction of 1.27 reveals a slight inclination for customers to buy the glasses independently.

Finally, the third rule involves the purchase of a 10ft iPhone Charger Cable 2 Pack alongside the Dust-Off Compressed Gas 2 pack. This rule has a high confidence level of 45.65%, and the lift of 1.91 indicates that the charger cable is almost twice as likely to be bought when the compressed gas pack is also purchased. The conviction of 1.40 suggests a substantial increase in the likelihood of buying the charger cable independently.

These rules, derived from the Apriori algorithm, provide valuable insights into customer purchase patterns, aiding businesses in strategic product placement and marketing decisions.

# **Part IV: Data Summary and Implications**

## D.  Summarize your data analysis by doing the following:

### 1.  Summarize the significance of support, lift, and confidence from the results of the analysis.

Defining:

* Support: The proportion of transactions that contain both antecedent and consequent. Higher support values indicate that the rule is applicable to a larger portion of the dataset. It helps identify items that frequently co-occur, highlighting popular combinations.
* Lift: The ratio of the observed support to the expected support. Lift > 1 implies a positive correlation between antecedent and consequent.
* Confidence: The probability of the occurrence of the consequent given the antecedent. Higher confidence indicates a stronger association.

Regarding my results to my analysis, looking at the various association rules, certain patterns stand out, particularly in the context of the antecedents and consequents involved. One notable association involves the purchase of a 10ft iPhone Charger Cable 2 Pack alongside the Dust-Off Compressed Gas 2 pack. The support value of 0.023064 indicates a moderate occurrence of this combination, and the confidence of 0.456464 highlights a significant likelihood of customers buying the charger cable when purchasing the compressed gas pack. The lift of 1.914955 suggests that these items are almost twice as likely to be bought together compared to their individual probabilities, indicating a potentially strategic pairing for marketing or product bundling.

Another interesting association revolves around the FEIYOLD Blue light Blocking Glasses and the Dust-Off Compressed Gas 2 pack. With a confidence of 0.419028 and a lift of 1.757904, it appears that customers who purchase the glasses also show a propensity for buying the compressed gas pack. This association could be leveraged for targeted promotions or product recommendations.

Moreover, the association between the SanDisk Ultra 64GB card and the Dust-Off Compressed Gas 2 pack showcases a different dynamic. The confidence of 0.416554 and lift of 1.747522 suggest that there's a substantial likelihood of customers acquiring both items together. This association could be explored further for potential cross-selling strategies.

Considering these insights, businesses can strategically position these items in proximity or run targeted promotions to capitalize on these observed associations, ultimately enhancing the customer shopping experience and potentially increasing sales.

Also in Part C4, I discussed the significance in the top three relevant rules from the analysis. Please refer to the section for further information.

### 2.  Discuss the practical significance of your findings from the analysis.

Discovering these associations through the analysis is like uncovering a treasure trove of actionable insights for businesses. The link between the 10ft iPhone Charger Cable 2 Pack and Dust-Off Compressed Gas 2 pack, for instance, suggests an opportunity to strategically position these items together. As a business, I would consider placing them in close proximity in physical stores or featuring them in joint promotions online to encourage customers to see the value in purchasing both items.

The analysis also reveals a chance to fine-tune marketing strategies. Recognizing that customers who buy FEIYOLD Blue light Blocking Glasses might also be interested in the Dust-Off Compressed Gas 2 pack, I'd design targeted campaigns to spotlight these associations. This personalized approach could lead to increased cross-selling and a more engaging shopping experience.

Furthermore, the high lift values associated with certain pairs, like the 10ft iPhone Charger Cable 2 Pack and Dust-Off Compressed Gas 2 pack, signal an opportunity for bundle offerings and discounts. Implementing these strategies could entice customers with added value and potentially boost sales.

Optimizing inventory management is another practical application. If certain items are frequently bought together, I'd consider stocking them in close proximity, making it easier for customers to find related products and potentially increasing sales.

Overall, the findings empower me to make informed, data-driven decisions. From strategic product placement to targeted marketing and inventory management, these insights provide a roadmap for enhancing the customer experience and maximizing business outcomes. It's like having a compass guiding me toward customer-centric strategies in a competitive landscape.

### 3.  Recommend a course of action for the real-world organizational situation from part A1 based on the results from part D1.

After analyzing the association rules in the telecommunications industry, I see a clear opportunity to enhance our customer experience and boost sales. Specific product pairs, like devices or accessories, have emerged as frequently purchased together. My first course of action would be to strategically bundle these items, creating comprehensive packages that offer added value to our customers.

Implementing online recommendations based on these associations is another crucial step. By guiding customers toward related products during their online journey, we can increase cross-selling opportunities and make their shopping experience more personalized and intuitive. This digital approach aligns with modern consumer behavior and expectations.

Optimizing in-store product placement is equally important. Placing frequently purchased items in close proximity and creating dedicated displays for complementary products can drive impulse purchases and improve overall customer satisfaction. This strategy aims to seamlessly integrate the online and offline shopping experiences.

Crafting tailored marketing campaigns is the next step. Promotions and advertisements should emphasize the convenience and benefits of purchasing specific items together. By reaching out to our audience through various channels, both digital and traditional, we can effectively communicate these offerings and attract more customers.

Continuous monitoring of the implemented strategies is crucial for success. By analyzing sales data, customer feedback, and engagement metrics, we can assess the impact and make data-driven refinements. This iterative process ensures that our approach remains dynamic and responsive to changing customer behavior.

Lastly, fostering collaboration with suppliers can unlock additional opportunities. Exploring joint promotions, exclusive bundles, and special offers in partnership with suppliers aligns with the identified associations and strengthens our overall value proposition. By working closely with suppliers, we can create mutually beneficial initiatives that resonate with our customer base.

In summary, these actionable steps derived from the association rule analysis pave the way for a customer-centric and strategically aligned approach in the telecommunications industry. By capitalizing on identified product associations, we can not only meet but exceed customer expectations, driving sustainable growth for our organization.

# **Part V: Attachments**

## E.  Provide a Panopto video recording that includes the presenter and a vocalized demonstration showing all code used, the code being executed, and the results of all code used in the task.

Please see attached.

## F.  Record all web sources you used to acquire data or segments of third-party code to support the application. Ensure the web sources are reliable.

DataCamp. (n.d.). *Market Basket Analysis in Python course*. <https://www.datacamp.com/courses/market-basket-analysis-in-python>

Hyunjoonbok. (n.d.). *Python-Projects/GeneralML/Cohort\_Basket\_Analysis/Market Basket Analysis 101 with Real Example - Association rules, Lift, Confidence, Support.ipynb at master · hyunjoonbok/Python-Projects*. GitHub. <https://github.com/hyunjoonbok/Python-Projects/blob/master/GeneralML/Cohort_Basket_Analysis/Market%20Basket%20Analysis%20101%20with%20Real%20Example%20-%20Association%20rules%2C%20Lift%2C%20Confidence%2C%20Support.ipynb>

Kadlaskar, A. (2023, November 16). *Market Basket Analysis: A Comprehensive guide for businesses*. Analytics Vidhya. <https://www.analyticsvidhya.com/blog/2021/10/a-comprehensive-guide-on-market-basket-analysis/>

Learning Champion. (2020, November 17). *Market Basket Analysis || Support || Lift || Confidence || Data Science for Beginners* [Video]. YouTube. <https://www.youtube.com/watch?v=xXE5k0XT35Y>

Oduwaiye, I. (2023, December 14). Market Basket Analysis with Python and Apriori - DevOps.dev. *Medium*. <https://blog.devops.dev/market-basket-analysis-with-python-and-apriori-algorithm-82cef11fb950>

Online, S. (2023, October 2). *Market basket Analysis in Python - Shiksha online*. Shiksha.com. <https://www.shiksha.com/online-courses/articles/market-basket-analysis-in-python/>

Roshan, B. (2021, December 24). 🍎Market Basket Analysis🍞- Association Rule Mining with visualizations. *Medium*. <https://medium.com/analytics-vidhya/market-basket-analysis-association-rule-mining-with-visualizations-cda24d537019>

Selvaraj, N. (2023, April 24). *How to perform market basket analysis in Python*. 365 Data Science. <https://365datascience.com/tutorials/python-tutorials/market-basket-analysis/>

Unfold Data Science. (2020, March 13). *Market Basket Analysis In Python|How to implement market basket analysis in Python|apriori algorithm* [Video]. YouTube. <https://www.youtube.com/watch?v=4QIWJVVWJdQ>

Yugagrawal. (2019, February 19). *Market basket Analysis (Apriori) in Python*. Kaggle. <https://www.kaggle.com/code/yugagrawal95/market-basket-analysis-apriori-in-python>