D212: Data Mining II - Task 3

Market Basket Analysis of WGU Telecom Transactional History

Andrew Mecchi

Masters of Data Analytics, Western Governors University

The following report covers Task 3 of the Performance Assessment for D212 Data Mining II. This document is categorized by the questions defined in the rubric.

Part I: Research Question

A1. Propose one question relevant to a real-world organizational situation that you will answer using market basket analysis.

Retaining profitable customers in the telecommunications industry is difficult knowing churn rates can be as high as 25% as described by the WGU Telecom data dictionary. Worseso, obtaining new clientele can generate costs upwards of ten times more than the price to retain subscribers. To counteract potential customer churn, WGU Telecom executives want to offer discounts to patrons on purchased products of interest, but an examination of purchase habits has not been performed. To assist with the discovery of customer purchase habits, the following research question is proposed:

Can market basket analysis be used to identify what items telecom subscribers frequently purchase together?

A2. Define one goal of the data analysis.

The main goal of the market basket analysis is to identify products that customers frequently purchase together. If the analysis is successful and can adequately classify concurrent items purchased, WGU Telecom can use the results to offer discounts on commonly purchased products as a campaign to retain existing customers.

Part II: Market Basket Justification

B1. Explain how the market basket analyzes the dataset with expected outcomes.

Market basket analysis is a systematic examination of customer purchase histories whereby rules and associations are defined by an algorithm and performance metrics determine the statistical significance of items frequently purchased together. Identifying products frequently purchased together allows businesses to feature these items and offer discounts or promotions to bolster sales. To perform a market basket analysis, a dataset of consumer transaction history composed of individual transactions is required.

A customer transaction history is a dataset composed of individual purchases and an itemized list of products for each transaction. Each transaction is represented by a row in the dataset and the column(s) contain the information of purchased items. For example, using a grocery store's transaction dataset, a transaction may include the purchase of milk, cereal, bread, and butter. This purchase would occupy a single row in the dataset while the products bought are either classified in an individual column separated by a delimiter, or, there are multiple item columns and each product is represented by a value in a column. The market basket method takes the data and analyzes each transaction and creates association rules. Association rules are relationship probabilities, but do not imply causal relationships, only co-occurrences; or, statements that connect an {antecedent} item to a {consequent} item (Sivek, 2020). The list of possible combinations (rules) translates to, “if an antecedent item is purchased then a consequent item is also purchased.” Using the grocery store example, a simplified association rule would be written as {milk}→{cereal}, or, if a customer purchases milk, then they also buy cereal.

Market basket analysis takes an entire dataset of transactions and classifies each transaction into itemsets, or items that appear together in a given purchase. These itemsets are analyzed by an algorithm which computes performance metrics to create a set of association rules. These association rules are generated from common metrics relative to the identified itemsets; including, support and confidence, which expresses the likelihood of one item being purchased given the purchase of another item (Kadlaskar, 2023). Ultimately, the results identify rules relative to their strength of association between purchased products and eliminate the rules with low significance. The products with the strongest associations are presented to businesses as items of interest from which promotional sale campaigns can be designed.

Market basket analysis will be implemented with the WGU Telecom transaction data and frequent itemsets analyzed. The frequent itemsets output from the algorithm will be limited to those with a defined minimum support value, or, the frequency of which an item appears in a given purchase relative to all transactions. Following the creation of frequent itemsets, association rules are generated using additional performance metrics to filter the results, retaining only relationships with the strongest associations (lift is used in this analysis). From the refined rules, additional performance filters are applied to identify the products most often purchased together. It is my expectation that the market basket analysis will be able to identify items frequently purchased together with statistical significance. The items with the strongest relationships classified by this analysis will be presented to WGU Telecom executives as products to target for promotional sales and marketing campaigns.

B2. Provide one example of a transaction from the dataset.

Figure 1 shows an example of an individual purchase from the WGU Telecom transactional dataset. This individual transaction includes the purchase of three items, Apple Lightning to Digital AV Adapter, TP-Link AC1750 Smart WiFi Router, and an Apple Pencil. As evidenced by the purchase, upwards of twenty items could be included in an individual purchase with NaN values entered for rows where no additional items were bought. The entire dataset contains 7501 purchases with a maximum of 20 items in a single purchase (Fig. 2)

Figure 1: Example of Individual Transaction

```
[59] # For question B2 on Task 3, print an example of a dataset transaction
     print(df.iloc[1])

     Item01    Apple Lightning to Digital AV Adapter
     Item02          TP-Link AC1750 Smart WiFi Router
     Item03                              Apple Pencil
     Item04                                       NaN
     Item05                                       NaN
     Item06                                       NaN
     Item07                                       NaN
     Item08                                       NaN
     Item09                                       NaN
     Item10                                       NaN
     Item11                                       NaN
     Item12                                       NaN
     Item13                                       NaN
     Item14                                       NaN
     Item15                                       NaN
     Item16                                       NaN
     Item17                                       NaN
     Item18                                       NaN
     Item19                                       NaN
     Item20                                       NaN
     Name: 1, dtype: object
```

Figure 2: Transaction Count and Max Items Purchased

```
[62] # Identify the number of transactions in the dataset
     print("Number of individual transactions: " + str(df.shape[0]))
     print("Number of max items per transaction: " + str(df.shape[1]))

     Number of individual transactions: 7501
     Number of max items per transaction 20
```

B3. Summarize one assumption of market basket analysis.

When employing market basket analysis on a dataset, it is important to understand the assumptions behind the method. The underlying assumption in market basket analysis is the joint occurrence of two or more products imply that these products are complements in purchase, therefore, the purchase of one will lead to purchase of others (Hua, 2015). The items purchased together are classified as itemsets and the apriori algorithm computes frequent itemsets, or, items commonly purchased together given a minimum frequency threshold. This assumption of market basket analysis purports that products in frequent itemsets are complements to each other and are classified as the association rules. Based on the strength of relationship of a given association rule leads to the identification of products purchased together.

Part III: Data Preparation and Analysis

C1. Transform the dataset to make it suitable for analysis. Attach copy of clean dataset.

1) Import market basket analysis transactional data

```
#Upload CSV file
df =
pd.read_csv(r'C:\Users\andrew\Desktop\WGU_MSDA\D212_Data_Mining_II\Data
Sets\churn\teleco_market_basket.csv')
```

2) View the first 10 rows of data. Identify multiple rows missing values.

```
# View first 10 rows of dataframe
# dataframe appears to be full of NaN rows
df.head(10)
```

3) Following identification of rows with missing values, all rows with NaNs are dropped. Rows within transactional data are individual purchases, therefore, for a purchase to be entered, a minimum of one item is necessary for purchase. Empty rows are dropped if there is no value for "Item01," otherwise, the minimum value to qualify a transaction. Index is reset after the drop and confirmation of dropped rows is visualized.

```
# Entire rows missing values
# Clean dataframe and remove empty rows
df = df[df['Item01'].notna()]

# Reset index after dropping empty rows
df.reset_index(drop = True, inplace = True)

# View dataframe after rows dropped
df.head(10)
```

4) Identify the number of transactions and maximum number of purchases within a transaction

```
# Identify the number of transactions in the dataset
print("Number of individual transactions: " + str(df.shape[0]))
print("Number of max items per transaction: " + str(df.shape[1]))
```

5) Create a list of lists in preparation of encoding transaction list

```
# Transform transactional data into list of lists
# (Kamara, 2022)
com_item_list = []

for row in range (len(df)):
```

```
  com_item_list.append([str(df.values[row, value])
  for value in range(len(df.columns))])
```

6) Encode list of lists using Transaction Encoder, applying fit and transform methods.

```
# Instantiate encoder object
encoder = TransactionEncoder()

# Fit and transform encoder
encoded = encoder.fit(com_item_list).transform(com_item_list)
```

7) Recast encoded array into a new, “clean” dataframe

```
# Recast encoded array as a dataframe
df_clean = pd.DataFrame(encoded, columns = encoder.columns_)
```

8) Explore the newly created “clean” dataframe. Recast the encoded array into a dataframe will classify each transaction into a row, and each column a specific item. Exploring columns will indicate if all columns are indeed purchasable items.

```
# Confirm no empty columns in clean df
for col in df_clean.columns:
  print(col)
```

9) Results of column exploration revealed a column labeled as NaN, or, not a number, also, not an item. Proceed to drop NaN column and view shape of df.

```
# Found a NaN column, drop empty column
df_clean = df_clean.drop(columns = ['nan'], axis = 1)

# Confirm drop of NaN column, should now have 119 columns
df_clean.shape
```

10) Run another loop to view columns of clean dataframe to confirm removal of NaN column

```
# Perform column search to confirm removal of NaN column
for col in df_clean.columns:
  print(col)
```

11) All columns refer to purchasable items, view the first three rows of cleaned dataframe.

```
# Preview first 3 rows of cleaned dataframe
df_clean.head(3)
```

12) Export cleaned dataframe

```
# Export encoded dataframe of itemized list of rx's to csv file
```

```
df_clean.to_csv(r'C:\\Users\\andrew\\Desktop\\WGU_MSDA\\D212_Data_Mining_I
I\\PA\\Task_3\\df_clean.csv')
```

13) Cleaned copy of dataframe.

SEE ATTACHED: df_clean.csv

C2. Execute the code used to generate association rules with the Apriori algorithm.

The generation of association rules is a process that involves running two functions. First, the apriori algorithm is applied to the cleaned dataset and identifies frequent itemsets, or, items frequently purchased together. To limit the frequent itemset permutations, a minimum support threshold of 0.02 was implemented, therefore, limiting results to products that appear in 2% of all transactions. The association rules are determined by inputting the frequent itemsets into the association_rules function and applying another metric value filter. The apriori algorithm limited the data to 103 frequent itemsets (Fig. 3) and after applying a minimum lift value of 1.0 with the association_rules function, the number of associations was reduced to 94 (Fig. 4).

1) Compute and view frequent itemsets using apriori function setting minimum support of 0.02.

```
# Identify frequent itemsets
fq_item_sets = apriori(df_clean, min_support = 0.02, use_colnames = True)

fq_item_sets
```

Figure 3: Functionality of Apriori Algorithm

```
[125] # Identify frequent item sets
      fq_item_sets = apriori(df_clean, min_support = 0.02, use_colnames = True)

      fq_item_sets
```

	support	itemsets
0	0.050527	(10ft iPHone Charger Cable 2 Pack)
1	0.042528	(3A USB Type C Cable 3 pack 6FT)
2	0.029463	(Anker 2-in-1 USB Card Reader)
3	0.068391	(Anker USB C to HDMI Adapter)
4	0.087188	(Apple Lightning to Digital AV Adapter)
...	...	...
98	0.023730	(Screen Mom Screen Cleaner kit, USB 2.0 Printe...
99	0.035462	(Screen Mom Screen Cleaner kit, VIVO Dual LCD ...
100	0.020131	(Stylus Pen for iPad, USB 2.0 Printer cable)
101	0.025197	(Stylus Pen for iPad, VIVO Dual LCD Monitor De...
102	0.027596	(USB 2.0 Printer cable, VIVO Dual LCD Monitor ...

103 rows × 2 columns

2) Generate association rules applying minimum lift value of 1 as values above one are indicative that the items are not purchased together by random chance.

```
# Define association rules, metric will be lift
rules = association_rules(fq_item_sets, metric = 'lift', min_threshold =
1.5)

# Identify number of associations
print("Number of Associations: " + str(rules.shape[0]))
```

Figure 4: Functionality of Association Rules

```
# Define association rules, metric will be lift
rules = association_rules(fq_item_sets, metric = 'lift', min_threshold = 1)

# Identify number of associations
print("Number of Associations: " + str(rules.shape[0]))
```

```
Number of Associations: 94
```

C3. Provide values for the support, lift, and confidence of the association rules table.

Figures 5 & 6 show the long-form and refined view of the association rules table. Figure 5 provides all metrics while Figure 6 limits the view of the table to focus on the values for support, lift, and confidence.

Figure 5: View of Association Rules Table

```
[133] # View association rules table
      rules
```

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction	zhangs_metric
0	(10ft iPHone Charger Cable 2 Pack)	(Dust-Off Compressed Gas 2 pack)	0.050527	0.238368	0.023064	0.456464	1.914955	0.011020	1.401255	0.503221
1	(Dust-Off Compressed Gas 2 pack)	(10ft iPHone Charger Cable 2 Pack)	0.238368	0.050527	0.023064	0.096756	1.914955	0.011020	1.051182	0.627330
2	(Anker USB C to HDMI Adapter)	(Dust-Off Compressed Gas 2 pack)	0.068391	0.238368	0.024397	0.356725	1.496530	0.008095	1.183991	0.356144
3	(Dust-Off Compressed Gas 2 pack)	(Anker USB C to HDMI Adapter)	0.238368	0.068391	0.024397	0.102349	1.496530	0.008095	1.037830	0.435627
4	(Anker USB C to HDMI Adapter)	(VIVO Dual LCD Monitor Desk mount)	0.068391	0.174110	0.020931	0.306043	1.757755	0.009023	1.190117	0.462740
...	...	...	...	...	...	...	...	...	...	...
89	(VIVO Dual LCD Monitor Desk mount)	(Screen Mom Screen Cleaner kit)	0.174110	0.129583	0.035462	0.203675	1.571779	0.012900	1.093043	0.440468
90	(Stylus Pen for iPad)	(USB 2.0 Printer cable)	0.095054	0.170911	0.020131	0.211781	1.239135	0.003885	1.051852	0.213256
91	(USB 2.0 Printer cable)	(Stylus Pen for iPad)	0.170911	0.095054	0.020131	0.117785	1.239135	0.003885	1.025766	0.232768
92	(Stylus Pen for iPad)	(VIVO Dual LCD Monitor Desk mount)	0.095054	0.174110	0.025197	0.265077	1.522468	0.008647	1.123778	0.379218
93	(VIVO Dual LCD Monitor Desk mount)	(Stylus Pen for iPad)	0.174110	0.095054	0.025197	0.144717	1.522468	0.008647	1.058066	0.415518

94 rows × 10 columns

Figure 6: Association Rules Table - Support, Lift, Confidence Metrics

```
# PA - C3 -  provide values for the support, lift, and confidence of the association rules table
rules[['antecedents', 'consequents', 'support', 'lift', 'confidence']]
```

	antecedents	consequents	support	lift	confidence
0	(10ft iPhone Charger Cable 2 Pack)	(Dust-Off Compressed Gas 2 pack)	0.023064	1.914955	0.456464
1	(Dust-Off Compressed Gas 2 pack)	(10ft iPhone Charger Cable 2 Pack)	0.023064	1.914955	0.096756
2	(Anker USB C to HDMI Adapter)	(Dust-Off Compressed Gas 2 pack)	0.024397	1.496530	0.356725
3	(Dust-Off Compressed Gas 2 pack)	(Anker USB C to HDMI Adapter)	0.024397	1.496530	0.102349
4	(Anker USB C to HDMI Adapter)	(VIVO Dual LCD Monitor Desk mount)	0.020931	1.757755	0.306043
...	...	...	...	...	...
89	(VIVO Dual LCD Monitor Desk mount)	(Screen Mom Screen Cleaner kit)	0.035462	1.571779	0.203675
90	(Stylus Pen for iPad)	(USB 2.0 Printer cable)	0.020131	1.239135	0.211781
91	(USB 2.0 Printer cable)	(Stylus Pen for iPad)	0.020131	1.239135	0.117785
92	(Stylus Pen for iPad)	(VIVO Dual LCD Monitor Desk mount)	0.025197	1.522468	0.265077
93	(VIVO Dual LCD Monitor Desk mount)	(Stylus Pen for iPad)	0.025197	1.522468	0.144717

94 rows × 5 columns

C4. Identify the top three rules generated by the Apriori algorithm. Include a screenshot of the top rules along with their summaries.

To identify the top three association rules, a scatterplot of the rules' confidence and support metrics (lift represented by size attribute) was used to help assist with refining metric filters (Fig. 7). The addition of the scatterplot to aid in filtering association rules was based on the work of Bayardo-Agarwal, who suggested the most interesting rules exist on the border of a confidence vs. support scatter plot (Paul, 2021). Therefore, the support, lift, and confidence metrics were tuned until a statistically significant top-three presented itself. The final metrics applied to identify the top-three association rules were; support = 0.025, lift = 1.75, and confidence = 0.33 (Fig. 8).

Figure 7: Scatterplot of Confidence vs. Support

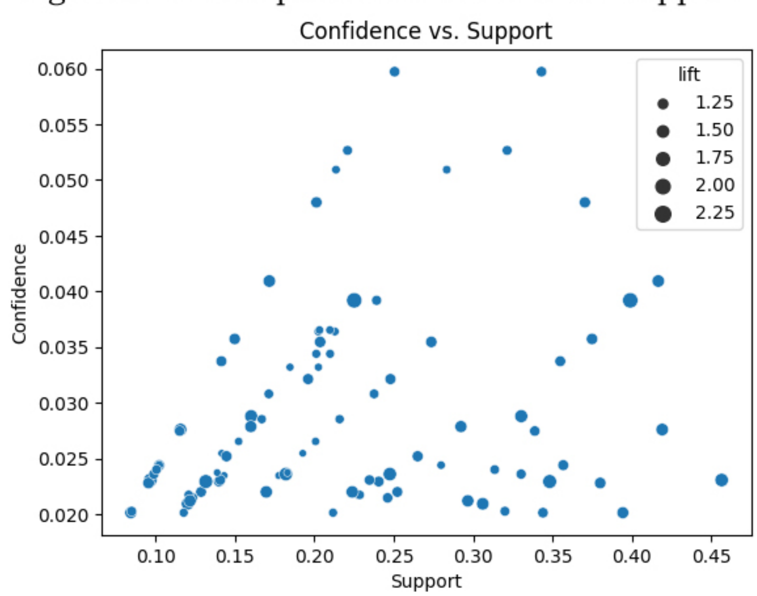

Figure 8: Top Three Association Rules and Filtered View of Results

```
[145] # Look for top rules adjusting metrics for support, lift, and confidence
      top_3_rules = rules[(rules['support'] >= 0.025) & (rules['lift'] >= 1.75) & (rules['confidence'] >= 0.33)].sort_values(by = ['lift'], ascending = False)

      top_3_rules[['antecedents', 'consequents', 'support', 'lift', 'confidence']]
```

	antecedents	consequents	support	lift	confidence
85	(SanDisk Ultra 64GB card)	(VIVO Dual LCD Monitor Desk mount)	0.039195	2.291162	0.398915
7	(Apple Lightning to Digital AV Adapter)	(Apple Pencil)	0.028796	1.837830	0.330275
36	(FEIYOLD Blue light Blocking Glasses)	(Dust-Off Compressed Gas 2 pack)	0.027596	1.757904	0.419028

Part IV: Data Summary and Implications

D1. Summarize the significance of support, lift, and confidence from the results of the analysis.

The top-three rules have been identified based on the significance of their performance metrics; support, lift, and confidence. However, to understand what these association rules imply, one must recognize how to interpret the metrics. Support is the calculated average of an itemset relative to a dataset and is representative of the frequency with which an itemset appears among all transactions. Confidence measures the conditional probability of a consequent given the occurrence of the antecedent, in other words, out of the transactions that contain item A, how many also contain item B (Zhang, n.d.). Lift informs the relationship dependency of items and is the ratio of the probability of item B's occurrence given item A is present and the probability of item B's occurrence without knowing about the presence of item A (Kimnaruk, 2022). The value of lift is a directional relationship; if the value is greater than one, the presence of item A in the transaction causes a higher probability of B occurring in the same transaction. If the lift value is less than one, the opposite is true; the purchase of item A would decrease the probability of a purchase of item B, while a lift value equal to one suggests there is no causal relationship between items. Table 1 summarizes the top-three rules and their corresponding values of support, lift, and confidence.

Table 1: Top Three Association Rules and Performance Metrics (values rounded, 2-decimal places)

Rule Number	Antecedent	Consequent	Support	Lift	Confidence
A1	SanDisk Ultra 64GB card	VIVO Dual LCD Monitor Desk mount	0.04	2.29	0.40
A2	Apple Lightning to Digital AV Adapter	Apple Pencil	0.03	1.84	0.33
A3	FEIYOLD Blue Light Blocking Glasses	Dust-Off Compressed Gas 2 pack	0.03	1.76	0.42

The highest support of the top rules was A1 with a support value of 0.04, or 4% of all WGU Telecom transaction data included the purchase of a SanDisk Ultra 64GB card and a VIVO Dual LCD Monitor Desk mount. For rules A2 and A3, a purchase involving those {antecedent} → {consequent} itemsets appeared in roughly, 3% of all transactions. When assessing the value of lift for each of the top rules, all itemsets have a lift value that exceeds one, therefore, the presence of each rule's antecedent "lifts" the likelihood of also seeing the corresponding consequent purchased. For example, the lift for rule A2 (1.84) can be interpreted as; if a customer is purchasing an Apple Lightning to Digital AV Adapter (antecedent) there is an increased probability that the customer will also purchase an Apple pencil. Augmenting the interpretability of an itemset's causal relationships (lift) is confidence. The lift values of the top rules all indicate a positive association of the antecedents increasing the possibility of seeing the consequent and the confidence metric helps gauge that influence. Confidence adds meaning to lift by explaining the likelihood of the consequent being purchased given the presence of the antecedent, and the higher the confidence, the stronger the likelihood. Rule A3 scored the highest confidence value of 0.42, meaning, when a customer purchases Blue Light Blocking Glasses, there is an increased probability that the customer will also purchase a Dust-Off Compressed Gas 2-pack. The same holds true for rules A1 and A3 which had confidence scores of 0.40 and 0.33 respectively. The results of the market basket analysis are greatly informative, will assist executives in the design of promotional sales on products customers frequently purchase, and can be created with certainty given the statistical significance of the top three rules' support, lift, and confidence metrics.

D2. Discuss the practical significance of the findings from the analysis.

The market basket analysis of the WGU Telecom transactional dataset has proven practical in identifying the top rules that will assist executives with promotional campaigns. The goal of the analysis was to identify products frequently purchased together and the results identified three rules with great statistical significance. The results are informative of customer purchasing habits and will be presented to WGU Telecom executives. With the knowledge gained from the analysis, executives now have six products they can market to incentivize customers to purchase with the hope of retaining subscribers through promotional campaigns and reducing overall customer churn.

D3. Recommend a course of action for the real-world organizational situation from part A1 based on your results from part D1.

Provided the results yielded three rules with statistical significance relative to performance metrics; support, lift, and confidence, I will make suggestions pertaining to each rule; A1, A2, and A3 separately.

Rule A1: {SanDisk Ultra 64GB card} → {VIVO Dual LCD Monitor Desk mount}

The purchases of this itemset was represented in approximately 4% of all transactions with strong associations enhanced by the lift (2.29) and confidence (0.40) values. Therefore, I would recommend a promotional campaign that involves the purchase of the SanDisk Ultra 64 GB memory cards. Part of the campaign could also include a reduced price point for the memory card and a potential greater discount if the customer also adds the monitor desk mount to the same purchase. However, the focus on the memory card sales is rooted in that a customer will most likely (only) need one LCD monitor mount for their computers. Whereas, memory cards can run out of storage space, become lost or misplaced, or even break; therefore, the customer may need to purchase additional memory cards as needed or as replacements. Offering a reduced price on memory cards may incentivize consumers to make additional purchases knowing they will save money when the memory cards are added to their carts.

Rule A2: {Apple Lightning to Digital AV Adapter} → {Apple Pencil}

The support for purchases of the itemset of rule A2 was represented in roughly 3% of total transactions and registered strong associations with a lift value of 1.84 and confidence score of 0.33. Given both items are Apple products, I would strongly suggest marketing these Apple products together as well additional ancillary Apple products. It is possible that an AV adapter and Apple pencil are one-time purchases, therefore, marketing additional Apple products may influence the customer to add more items to their cart if promotional sales are incentivized.

Rule A3: FEIYOLD Blue Light Blocking Glasses} → {Dust-Off Compressed Gas 2 pack}

The purchases of an itemset involving Blue Light Blocking Glasses and Compressed Gas cleaner appeared in approximately 3% of all transactions and the significance of their association was reinforced with a lift value of 1.76 and confidence of 0.42. Here, I would suggest a cross-promotional campaign that targets blue light blocking glasses and compressed air cleaner, where the purchase of the glasses will offer a reduced price if the purchase also includes the compressed gas. While the need for blue light glasses is limited to an individual (pending a replacement) it is most likely a one-time purchase. However, the demand for compressed air is persistent, in that, a customer may be in-need of the product at any time given it has a limited use (runs out of gas). I would strongly suggest executives consider promoting the Dust-Off Compressed Gas 2-pack at a slightly reduced price if a customer adds it to any existing order. Additionally, I believe the compressed air should be advertised as, “add to cart,” when a customer proceeds to checkout as the need for a compressed air cleaner is potentially secondary to an overall purchase, and may boost sales with an incentivized impulse-buy price point.

Part V: Attachments

E. Provide a Panopto video recording that includes a demonstration of the functionality of the code used for the analysis and a summary of the programming environment.

SEE ATTACHED → Link: Panopto Presentation

Copy and Paste:
https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=d9dbc73a-2128-4b19-b74a-b067015223cd

CODE SOURCES

Kamara, K. (2022, September 1) Market Basket Analysis using Python. [Webinar]. Faculty of Masters of Data Analytics, Western Governors University. https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=dbe89ddb-e92f-4d40-a87a-af030178abf1.

REFERENCES

Hua, Sarah (2015, September 1). Market Basket Analysis. Hua's Analysis. https://sarahtianhua.wordpress.com/portfolio/market-basket-analysis/.

Kadlaskar, A. (2023, April 26). Market Basket Analysis: A Comprehensive Guide for Businesses. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2021/10/a-comprehensive-guide-on-market-basket-analysis/#h-how-does-market-basket-analysis-work.

Kimnaruk, Y. (2022, September 18).What are market basket analysis and the apriori algorithm. Medium. https://medium.com/mlearning-ai/what-are-market-basket-analysis-and-the-apriori-algorithm-fe0e8e6e34d.

Paul, A.C. (2021, August 13). A Conceptual Introduction into Association Rule Mining — Part 2. Medium. https://medium.com/delvify/a-conceptual-introduction-into-association-rule-mining-part-2-96c73c4ce87b.

Sivek, S.C. (2020, November 16). Market Basket Analysis 101: Key Concepts. Towards Data Science. https://towardsdatascience.com/market-basket-analysis-101-key-concepts-1ddc6876cd00#.

Zhang, L. (n.d.). Understanding Support, Confidence, Lift for Market Basket (Affinity) Analysis. The Data School. Retrieved August 22, 2023). https://www.thedataschool.co.uk/liu-zhang/understanding-lift-for-market-basket-analysis/.