Unveiling the Wave: Unraveling Sales Trends and Unlocking Market Basket Insights

Ming Ann, Teh[1], Tong Ming, Lim[2]*, Wan Yoke, Chin[3]

[1,3]*Faculty of Computing and Information Technology, Tunku Abdul Rahman University of Management & Technology; E-mail: limtm@tarc.edu.my*

[2]*Centre for Business Incubation and Entrepreneurial Ventures, Tunku Abdul Rahman University of Management & Technology*

Abstracts: This article presents sales trends and market basket analysis for Food and Beverages (F&B) items recommendation using linear regression time series and association rules. The purpose of this study is to help businesses to make better informed decisions by understanding their customers' purchasing behaviors and the sales trends of the F&B items. This includes the study of both cross- selling and up-selling patterns using market basket analysis. To improve the market basket analysis, food name clustering has been carried out to group those foods having similarity ratio higher or equal to 80 based on the Levenshtein distances between the food names. This study analyzes sales data from a mobile app platform that register the Malaysian F&B hawkers (or merchants) and app users (or buyers) who sell and purchase F&B respectively. This study is based on sales trends analysis and association rules of product recommendations by examining FP-Growth and the relationship between product items bought. This study provides valuable insights for local Malaysian hawkers who are looking to improve their F&B items recommendations and to make better data-driven decisions while planning their marketing activities. It helps to improve sales by either up or cross-sell their products more effectively.

Keywords: Sales Trend Analysis, Market Basket Analysis, Association Rules, Product Recommendation.

1. INTRODUCTION

The Food and Beverages (F&B) retail industry has been experiencing rapid changes and advancements technologically both from promotion and marketing perspective in recent years (Nimbly Technologies, 2022), making it more competitive than ever before. In order to stay ahead of the game, businesses need to constantly analyze and interpret their sales data to make informed decisions. Sales trend analysis (Lavendar Nguyen. 2022; Bizstats. 2019; Efti, S. 2023) and market basket analysis (Kadlaskar, A., 2021; Loshin, D. 2013; Coumar, N. 2020) are two powerful analyses that can provide insights on customer behavior and help businesses improve their sales and marketing strategies. Sales trend analysis involves examining sales data over a period of time to identify patterns and trends in customer behavior. This technique helps businesses understand the impact of various factors on their sales performance, such as seasonal changes, promotional activities, and changes in customer preferences.

Market basket analysis, on the other hand, is a technique that examines the relationships between different products purchased by customers. By analyzing customer transaction data, businesses can identify which products are frequently purchased together, and use this information to improve product placement, promotions, and cross-selling opportunities. In this paper, we will explore the use of sales trend analysis and market basket analysis to help businesses make data-driven decisions. We will use Python programming language, along with the Matplotlib library, to visualize sales trends and identify associations between different food items. For market basket analysis, techniques such as association rules and frequent itemsetmining are valuable tools for identifying patterns and relationships between products purchased by customers. This information can be used by businesses to make informed decisions, such as product recommendations and optimizing store layout. By utilizing these techniques, businesses can increase revenue, improve customer satisfaction, and reduce waste by managing inventory more efficiently. To allowtest and validation of the works developed, we had developed a Flask based web application that automates the analysis process and provides businesses with an interactive tool for decision- making. While there may be some limitations to our research, such as the limited types of data visualization techniques used, we believe that this study will provide valuable insights for businesses looking to improve their sales and marketing strategies.

2. LITERATURE REVIEWS

2.1. Sales analysis

Sales trend analysis is an important aspect of business decision-making that involves analyzing sales data over a period of time to identify trends and patterns. In recent years, there has been an increase in the use of data analysis techniques in various industries to gain insights into customer behavior and preferences. The following literature review discusses some of the relevant studies and techniques used in sales trend analysis.

One popular approach to sales trend analysis is time series analysis. Time series analysis is a statistical technique that involves analyzing data over time to identify trends and patterns. According to Neelam Tyagi(Tyagi, N, 2021), time series analysis has been used to analyze sales data in various industries, including retail, finance, and healthcare. Time series analysis can help businesses identify seasonal trends, cyclical patterns, and overall trends in sales data.

Another approach to sales trend analysis is machine learning. Machine learning involves training algorithms to recognize patterns in data and make predictions based on those patterns. According to Ostap Zabolotnyy (Yamasaki, D., 2017), machine learning has been used to analyze sales data in various industries, including e-commerce and retail. Machine learning algorithms can help businesses identify customer preferences, predict sales trends, and optimize pricing strategies.

In summary, sales trend analysis is an important aspect of business decision-making that involves analyzing sales data over a period of time to identify trends and patterns. Time series analysis, machine learning, and association rule mining are some of the relevant techniques that can be used in sales trend analysis. It neednot be dry words or uninteresting numbers.

Table 1: Papers reviewed on Sales analysis

Author	Objective	Techniques	Limitations
Shrapnel, W. S., &Butcher, B. E. (2020)	To analyze trends in national volume sales of sugar-sweetened beverages (SSBs)	Linear Regression	Since the datasets only included a portion of the total volume sales of water-based beverages, some imputation was required to increase the values and account for all convenience volume sales as well as an estimate of foodservice, vending, and dining expenditures.
Romano, S.(2020)	To characterize the impact of the COVID-19 outbreak on outpatient medicines' sales and shortages.	Min-max normalization, time-trend analysis	The results for April are provisional estimates interms of data gathering because monthly pharmaceutical sales are verified in the second half of the following month. However, because very little sales were deleted throughout the validation process, these estimates can be used for analysis without running the risk of significant data differences.
Hoefler, R. et al, (2023)	To investigate the trends in antidepressants sales in Brazil	Join point regression, chi-square trend	The information excludes both the therapeutic indications for antidepressants prescribed in hospitals, government agencies, and compounding pharmacies.
Gupta, A. (2015)	To predict the future sales and growth of athletic shoes market	Trend analysis	Factors like culture, economy and taste willinfluence the sales of athletic shoes also.
amasaki, D.et al, (2017)	to evaluate the utility of the national database (NDB) based on health insurance claims data for antimicrobial use (AMU) surveillance in medical institutions in Japan.	Trend analysis, Pearson's correlation analysis	Although this should account for a minor proportion, the database we utilised in this investigation did not contain antimicrobialsprescribed in the dental clinic and department of the hospital. A follow-up investigation with more details is necessary.

Wisesa, O,(2020)	To analyze Business to Business (B2B) Telecommunication services using Machine Learning Techniques	Machine Learning, Decision Tree, Random Forest	The most recent design had a Mean Squared Error for GBT Prediction of 0.18. To enhance the performance of the MSE, it was suggested to create the GBT algorithm.
Ostapenko, R. et al,(2020)	To identify the major trends in production and sales of the organic agricultural products in Ukraine.	Statistics, Random sampling	Because conventional farms were chosen at random as opposed to organic farms, this study's analysis of conventional farms is constrained. Propensity score matching and bootstrapping can be used in future investigations to achieve more precise inference.
Choi, T-M (2014)	To achieve real time sales forecasting for fast fashion operations in the future	3F algorithm	Fast fashion has a very short product life cycle and many operational decisions in the supply chain that must be made on a tight timeframe.
Zeitouny, S, (2021)	To mapping global trends in vaccine sales before and during the first wave of the COVID-19 pandemic	Time series analysis	The study's main flaw is that none of the specific nations for which data were available fit the description of low-income nations.
Massaro, A., (n.d.)	To predict the sales based on the effect of weather	Regressio nanalysis	Lack of model used, in the future, Data Mining model such as decision tree can be used to compare their performance.

2.2. Market basket analysis

Market basket analysis (MBA) determines consumer purchasing patterns by identifying association between the many goods customers place in their "shopping baskets." By learning which products consumers regularly purchase together, retailers and marketers can use the finding of this kind of relationship to generate marketing strategies. For instance, how likely is it that consumers who are purchasing milk will also purchase bread or butter during the same trip to the retails store (Kadlaskar, A., 2021)? By assisting shops with targeted marketing and ledge space planning, this information may enhance sales. In addition, retailers may utilize a customer's affection for a bundle of products as a sign that they may be more likely to buy them all at once. This makes it possible to present products for cross-selling or may imply that buyers could be likely to purchase additional products when specific products are groupedtogether (Loshin, D. 2013).

MBA has been widely used in various industries such as retail, healthcare, and finance. In retail, MBA hasbeen applied to analyze customer purchases and recommend products based on their shopping habits. In healthcare, MBA has been used to identify potential drug interactions and recommend treatment plans for patients (Rao, A. B, 2021). In finance, MBA has been applied to detect fraudulent activities and to make investment decisions based on market trends (Coumar, N. 2020).

Table 2: Papers reviewed on Market Basket Analysis

Author	Objectives	Techniques	Limitations/ Future works
Gupta, S., & Mamtora, R. (2014)	To survey about the existing data mining algorithm for market basketanalysis	Association Rule Mining, Apriori Algorithms	It frequently examines the database. Each time, more options will be generated during the scanning process. This adds to the work required to search the database. Therefore, a large volume of data servicesmust be stored in a database. As a result, there isn't enough memory to hold those extra pieces of data.
N, I. (2018).	To discover a pattern in consumer purchasing behaviour that the business owner mayuse to inform their plans	Association Rule, Frequent Itemset Mining	More data should be used in future studies.This would improve the results' accuracy.
Musalem, A. (2018)	To outline a method for identifying relationships between product categories that canbe utilised to dividea retailer's company into smaller categories.	Market Basket Analysis	The empirical results are unique to the shop under study. Additionally, while the suggested methodology detects cross-category interrelations, it does not account for their underlying causes (such as marketing or non-marketing interrelations).
Kaur, M., & Kang, S. (2016)	To explores association rulemining, a data mining approach, and presents a new algorithm that could be useful for analyzing consumer behavior and boosting sales.	Data Mining, Association rule	The manual threshold values for score have a significant impact on the results, hence automating the threshold values is necessary for improved outlier detection. Additionally, this strategy focuses exclusively on Market Basket Data; however, it might be expanded to other areas.
Kumar, B. (2017)	To build prediction model using market basket analysis	Association rule mining, Predictive Modelling, Market Basket Analysis	The creation of intelligent prediction models to produce association rules that can be incorporated into recommendation systems in order to make them more functionally operational will be the focus of future work. The performance of the recommendation system can be improved by using more superior and efficient rule mining approaches.
Solnet, D., (2016)	To find attractive product and services of hotels in order to increase sales	Market Basket Analysis	In future, market basket analysis should give more attention in tourism and hospitality
Apolinario, K.(2016).	To analyse retail transactional data using MBA, using the findings as a prescriptive model for the optimal design of the sell floor and for directing the creation of relevant in-store marketing campaigns.	Market Basket Analysis, Sell floor Layout Optimization	Future analyses could take into account loyalty or reward cards, which give a newperspective and enable the creation of customised promotional plans for each client.

2.3. Critical Remarks

Based on the journals reviewed above, it is found that sales trend analysis is a valuable tool for businesses to track and analyze their sales performance over time. However, it is important to note that the accuracy and effectiveness of sales trend analysis can be influenced by various factors, such as the quality of the datacollected, the period analyzed, and the methodology used. Therefore, businesses should be cautious when interpreting the results of sales trend analysis and ensure that they consider all relevant factors before making any decisions

based on the analysis. Additionally, incorporating multiple techniques and data sources can help provide a more comprehensive and accurate understanding of sales trends.

It can be observed that market basket analysis is a widely used technique for identifying purchasing patterns and associations between products in retail settings. However, while many algorithms have been developed for this purpose, there is no one-size-fits-all solution, and the choice of algorithm depends on various factors such as the size of the dataset, the complexity of the associations, and the desired level of accuracy.

Additionally, it is important to note that market basket analysis is not a panacea for all retail-related problems. While it can provide insights into customer behavior and product associations, it cannot explain why certain associations exist or predict future trends. Moreover, the accuracy and relevance of the insights obtained from market basket analysis depend on the quality and completeness of the underlying data. Therefore, it is crucial to ensure that the data used for analysis is relevant, accurate, and up to date.

Overall, while market basket analysis can be an additional valuable tool for businesses looking to optimize their sales and marketing strategies, it should be used in conjunction with other analytical techniques and should not be solely relied upon for decision-making (Kadlaskar, A., 2021; Loshin, D. 2013; Coumar, N. 2020).

3. PROBLEMS AND OBJECTIVES

The following are problems identified based on observations with from the industry partners and past research investigated:

1. Our study had found that inconsistent F&B product ID is serious and it is technically difficult to analyze their sales trends and customers' purchase behavior.

2. It is also found that F&B hawkers are serious about their customers' patterns but unable to analyze correctly.

3.1. Objectives:

1. To investigate using the Levenshtein Distance (Rosette. 2017; Deep, M. 2022; Nijhuis, M., 2022) based on F&B product name due to inconsistency in F&B product IDs used by F&B hawkers.

2. To study using clustering (Rosette. 2017; Deep, M. 2022; Nijhuis, M., 2022) and market basket analysis (Musalem, A., 2018; Kaur, M., & Kang, S., 2016; Kumar, B., 2017) on the similarity ratio by time series for customer behavior analysis.

3. To analyze transaction data using market basket analysis to identify trends in customers' purchasing patterns and use the insights to optimize menu offerings and pricing in order to up-selling and cross-selling strategies.

4. METHODOLOGY

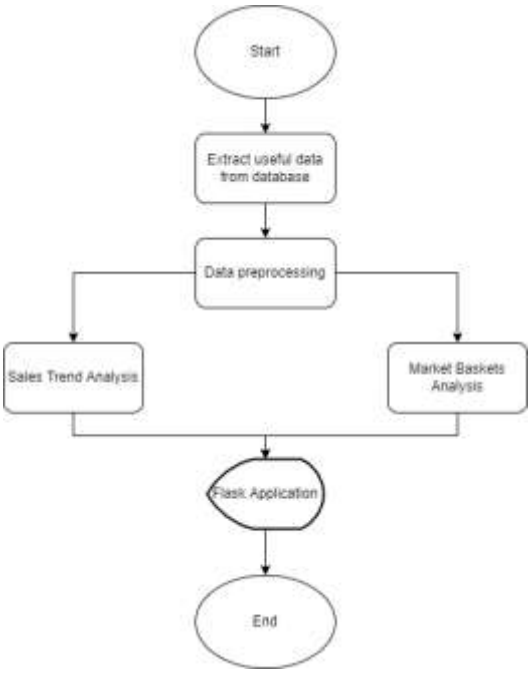

Figure 1: Flow chart of the system

Data Extraction: The first step is to extract useful data from the database that is valuable for sales trend analysis and market basket analysis. This involves identifying the appropriate database and selecting the relevant data that will be used for the analysis. The data set used in this research has 20496 rows transactional records.

Data Preprocessing: Once the data is extracted, the next step is to perform necessary data preprocessing to clean the data, remove garbage data that could affect the analysis result, and improve the analysis performance. This step could involve techniques such as data cleaning, data transformation, and data normalization.

Sales Trend Analysis: With the clean data, the sales trend analysis can be performed by grouping the data by day, week, and month to enable users to view and understand the trends and patterns in sales over specific periods. This analysis will help to identify the overall sales trends and patterns in the business.

Market Basket Analysis: The next step is to use association rules to identify the relationship between the items purchased by customers. This analysis will help to identify the product items that are commonly purchased together and can be used to generate recommended product items.

Flask-based web Application: The final step is to build a Flask based web application to enable users to view all the analysis results and interact with them. This step will involve data visualization techniques to make the results more comprehensive and user-friendly.

Overall, the methodology provides a clear and logical sequence of steps for extracting and analyzing data from a database and presenting the results in a user-friendly dashboard. The analysis will provide valuable insights to help the business make informed decisions to improve sales and customer satisfaction.

5. DATA PREPROCESSING

This section explains data preprocessing works taken place prior to the modelling.

Preparing data for analysis by cleaning, formatting, and transformation is known as data preparation. It

ensures that the data are of good quality and are in a format that is appropriate for analysis, making it a crucial stage in the data analysis process.

The data Is sourced from the TapFood company database, which is used to store information about merchants, customer data, and transaction data on their online food ordering platform. TapFood provides a platform for people to order food online.

The size of the data Is specified as containing 20,496 rows of transaction data.The fields in the data include:

orderId: This field represents the unique identifier of each order. orderDateTime: It captures the date and time when the order was placed.

totalAmount: This field indicates the total amount of the order, which is typically the sum of individualfood items' prices.

Id: This field represents the unique identifier of the customer who placed the order.Name: This field contains the name of the food item that was ordered.

Quantity: It specifies the quantity or number of units of the food item in the order.Amount: This field indicates the unit price of the food item in the order.

merchantId: It represents the unique identifier of the merchant from whom the order was placed. Name.1: This field contains the name of the merchant.

These fields provide essential information about the orders, including customer details, food items, quantities, prices, and merchant information.

5.1 Remove Duplicate Record

Removing duplicate data before conducting analysis is crucial for ensuring the accuracy, integrity, and efficiency of the analysis. Duplicate data can lead to inaccurate results, distorting the findings and compromising the reliability of the analysis. By removing duplicates, the dataset becomes streamlined andmore efficient, allowing for faster and more effective analysis. Additionally, duplicate data can introduce bias into the analysis, potentially skewing the results and creating a misleading representation of the true trends and patterns in the data. It also helps prevent overrepresentation of certain factors or attributes, ensuring a more balanced and accurate analysis. Overall, eliminating duplicate data is essential for maintaining data quality, enhancing the validity of the analysis, and making informed decisions based on reliable insights.

5.2. Near Name Matching

In this paper, we address the challenge of different food items having varying names by implementing a near name matching approach. This issue is prevalent in the food industry, where the same food can be referred to differently based on regional, cultural, or establishment-specific factors. Without near name matching, accurately analyzing "market baskets" and providing relevant recommendations based on the data would be challenging.

To overcome this challenge, we employ the fuzzywuzzy library, which utilizes the Levenshtein distance algorithm (Rosette. 2017; Deep, M. 2022; Nijhuis, M., 2022) to calculate similarity ratios between food names. A similarity ratio of 80 or higher is considered acceptable for grouping similar food items together. By standardizing the food names and grouping them based on this similarity threshold, we improve the performance of the market basket analysis. This approach enables a more comprehensive analysis of customer buying patterns, leading to more accurate and actionable insights for optimizing sales strategies and making product recommendations.

By employing the fuzzywuzzy algorithm (Rosette. 2017; Deep, M. 2022; Nijhuis, M., 2022) and implementing a standardized naming approach, our study enhances the accuracy and effectiveness of market basket analysis. This methodology enables us to overcome the challenges posed by variations in food names, ultimately contributing to improved decision-making processes and more successful sales strategies in the food industry.

5.3. Data Transformation

In order to perform market basket analysis effectively, certain data transformations are necessary. One common transformation involves encoding the data into a binary format, where a value of 0 represents "no purchase" and a value of 1 represents "purchased."

The reason for this transformation is to convert the transaction data into a format that is suitable for association rule mining, which is a fundamental technique used in market basket analysis. Association rule mining aims to discover relationships or associations between items based on their co-occurrence in transactions.

By encoding the data into 0s and 1s, we create a binary representation of the presence or absence of each item in a transaction. This binary format allows us to identify item sets or combinations of items that occur frequently together in transactions.

The transformation simplifies the data and allows for efficient analysis. It helps in identifying frequent itemsets, such as sets of items that are frequently purchased together. These frequent item sets can then be used to generate association rules that provide insights into customer behavior, preferences, and potential cross-selling opportunities.

Additionally, the binary representation facilitates the application of various algorithms and techniques used in market basket analysis, such as the Apriori algorithm. This algorithm relies on binary item sets to efficiently mine frequent item sets and generate association rules.

Overall, the transformation of data into a binary format by encoding it into 0s and 1s is essential for conducting market basket analysis effectively. It enables the identification of frequent item sets and the generation of meaningful association rules, leading to valuable insights for businesses to optimize their sales strategies and improve customer satisfaction.

5.4. Filter order that only have one items purchased

In market basket analysis, it is common practice to remove order records that consist of only one item purchased. This step is crucial to ensure the reliability and accuracy of the analysis results. Several reasonsjustify the exclusion of these single-item transactions.

Firstly, single-item transactions do not provide any meaningful associations as they lack the presence of multiple items. The goal of market basket analysis is to identify relationships and associations between items based on their co-occurrence in transactions. Including single-item transactions can introduce noise and make it challenging to uncover significant patterns or associations among items.

Secondly, single-item transactions may not possess sufficient statistical significance to draw reliable conclusions. The limited data points in these transactions make it difficult to establish statistically significant relationships or associations. By removing these transactions, the analysis can focus ontransactions with multiple items that have a higher likelihood of revealing significant patterns andassociations.

Furthermore, from a business perspective, single-item transactions may not provide substantial insights into customer behavior or inform effective sales and marketing strategies. Analyzing these transactions might not yield actionable insights for cross-selling or bundling strategies, which are commonly derived from transactions involving multiple items. By excluding single-item transactions, the analysis can concentrate on transactions that

are more likely to provide meaningful insights and drive impactful decision-making.

Lastly, removing single-item transactions improves resource efficiency. Analyzing these transactions can be computationally inefficient and add unnecessary complexity to the analysis. By filtering out these transactions, the dataset becomes more streamlined and easier to process, resulting in faster and more efficient analysis.

In summary, the removal of order records consisting of only one item purchased is a crucial step in market basket analysis. By excluding these transactions, the analysis can focus on transactions with multiple items, increasing the potential for discovering meaningful associations and patterns. This filtering process enhances the quality and relevance of the analysis results, enabling businesses to make informed decisions and optimize their sales strategies while improving customer satisfaction.

6. SALES TRENDS AND MARKET BASKET ANALYSIS ALGORITHM CONSIDERATIONS AND JUSTIFICATIONS

The following subsections present the algorithms that have adopted with justifications in term of its strengths and applicability.

6.1. Association rules for market basket analysis

In this research, the Association Rules (Kadlaskar, A., 2021; Loshin, D. 2013) (Rao, A. B, 2021) used are as follow:

(support (A -> B) = (number of transactions containing both A and B) / (total number of transactions)

(confidence (A -> B) = (number of transactions containing both A and B) / (number of transactions containing A)

(lift (A -> B) = (support (A -> B) / (support(A) * support(B))

where A and B are two items, and A -> B is the association rule that indicates that if a customer purchases item A, they are also likely to purchase item B.

In the context of Market Basket Analysis, support is a measure of the frequency with which a given item or set of items appears in a transactional dataset. It represents the proportion of transactions that contain the item or itemset and is used to identify the most frequently occurring itemset, which are then used to generate association rules.

Another important metric in Market Basket Analysis is confidence, which measures the likelihood that the rule A -> B holds true given that A has been purchased.

Finally, lift is a metric that measures the strength of the association between A and B compared to what would be expected by chance. A lift value greater than 1 indicates a positive association, while a value less than 1 indicates a negative association.

Association rules are employed in market basket analysis to determine relationships between various products in a dataset. It is frequently required to preprocess the data and set minimum confidence thresholds (as known as strong rules), minimum support threshold for the strength of the associations being identified in order to provide usable association rules (Kadlaskar, A., 2021). Following that, the resulting rules are often assessed using variables like support, confidence, and lift. The Apriori algorithm and the FP-growth algorithm are two popular algorithms employed for this purpose (Kadlaskar, A., 2021).

6.2. Algorithm used in Market Basket Analysis

6.2.1. Frequent Pattern Growth Algorithm

Frequent pattern growth algorithm (Javapoint, n.d.; Software Testing Help, 2023; Gupta, A., 2019) is also known as FP growth algorithm. This algorithm is similar to the Apriori algorithm, but it is more efficient for large datasets. It works by creating a prefix tree structure (called an FP-tree or frequent pattern) from the data, which allows it to quickly identify frequent item sets. The frequent pattern can be generated without the necessity for candidate generation. The relationship between the item sets is maintained by thefrequent pattern tree structure. Frequent Pattern Growth Algorithms are as below.

A tree structure called a frequent pattern tree is created using the earliest itemset of the data. The FP tree's main objective is to mine the most prevalent patterns. A component of that itemset is represented by each node in the FP tree. The lower nodes indicate the data itemset, whereas the root node represents the null value. While building the tree, the relationship between these nodes and the lower nodes, or between itemset,is maintained.

7. IMPLEMENTATION OF SALES TREND ANALYSIS

Time aggregation and visualization are essential steps in analyzing sales trends. Time aggregation involves grouping sales data into meaningful time intervals, such as days, weeks, or months, to gain a higher-level view of the sales patterns over time. This aggregation allows for a more concise and manageable representation of the data, making it easier to identify trends and patterns.

Once the sales data is aggregated, visualization techniques come into play. Visualizing the aggregated data through line charts, bar charts, or other suitable plots provides a visual representation of the sales trends. These visualizations offer a clear and intuitive way to observe changes and patterns in sales over time. Byexamining the plotted data, it becomes easier to identify seasonality, growth trends, or any notable fluctuations in sales volume or revenue.

Visualizations also help in understanding the impact of certain events or marketing campaigns on sales. By overlaying additional information, such as promotional periods or advertising campaigns, on the sales trends, it becomes possible to evaluate their effectiveness and their influence on sales performance.

In summary, time aggregation and visualization are crucial steps in analyzing sales trends. Time aggregation enables a higher-level view of sales data, while visualizations provide a clear and accessible means to identify patterns, trends, and the impact of various factors on sales performance. Together, these techniques empower businesses to make data-driven decisions and develop strategies to optimize sales and enhance overall performance.

8. IMPLEMENTATION OF MARKET BASKET ANALYSIS FOR UP- AND CROSS-SELLING

Once the data has been encoded into a binary format, the next step in market basket analysis is to apply association rule mining techniques such as Apriori or FP-Growth. These algorithms analyze the co- occurrence patterns of items in transactions to uncover meaningful associations or relationships.

First, the support for each item or itemset is calculated. Support refers to the proportion of transactions in which a specific item or itemset appears. It indicates the frequency or popularity of an item or itemset in the dataset. The support value helps in determining which associations are significant and worth considering.

To find the support, the dataset is scanned, and the number of transactions containing a particular item or itemset is counted. The support is calculated as the ratio of the count of transactions containing the item oritemset to the total number of transactions.

Once the support values are obtained, the market basket analysis can be performed. The association rule mining algorithms, such as Apriori or FP-Growth, use these support values to generate frequent itemset and association rules. The support threshold is set to 0.01, indicating that only itemset or rules with a support value equal to or higher than 0.01 (1%) will be considered significant.

Apriori algorithm generates frequent itemset by iteratively pruning and combining itemset based on their support values. It starts with individual items and progressively builds larger itemset, eliminating those thatfall below the support threshold.

FP-Growth algorithm, on the other hand, constructs a compact data structure called the FP-tree to representthe dataset. It then mines frequent itemset by recursively exploring the tree and finding patterns that meet the support threshold.

By applying these association rule mining techniques, market basket analysis reveals associations between items, helping businesses understand customer behavior and uncovering potential cross-sellingopportunities. The generated frequent itemset and association rules provide insights into which items are frequently purchased together, enabling businesses to optimize their product placement, marketing strategies, and recommendations to improve customer satisfaction and increase sales.

9. RESULTS AND ANALYSIS

9.1 Data visualization of Sales Trend Analysis in Flask based Web Application

The data visualization for the sales trend analysis module was implemented using Python's Matplotlib library. The module allows the user to plot graphs based on the sales data by day, week, or month.

The user can select the food item from the food name table or enter the food name in the text box to plot the sales trend graph. The graph is plotted using a line plot with the x-axis representing the date and the y- axis representing the sales amount.

Overall, the data visualization module enhances the user experience by providing an interactive and visually appealing way to analyze sales trends. The dashboard of the sales trend module is shown in figure 2.

Figure 2: Sales trend for “Fried Noodles” daily

9.2. Flask-based Web Application for Product Items Recommendation

Read the pickle file and list the merchants in the drop box, after users pick a merchant, display the recommended product items in two tables up selling and cross selling. If the second item is more expensivethan the first item, categorize it to the up-selling table.

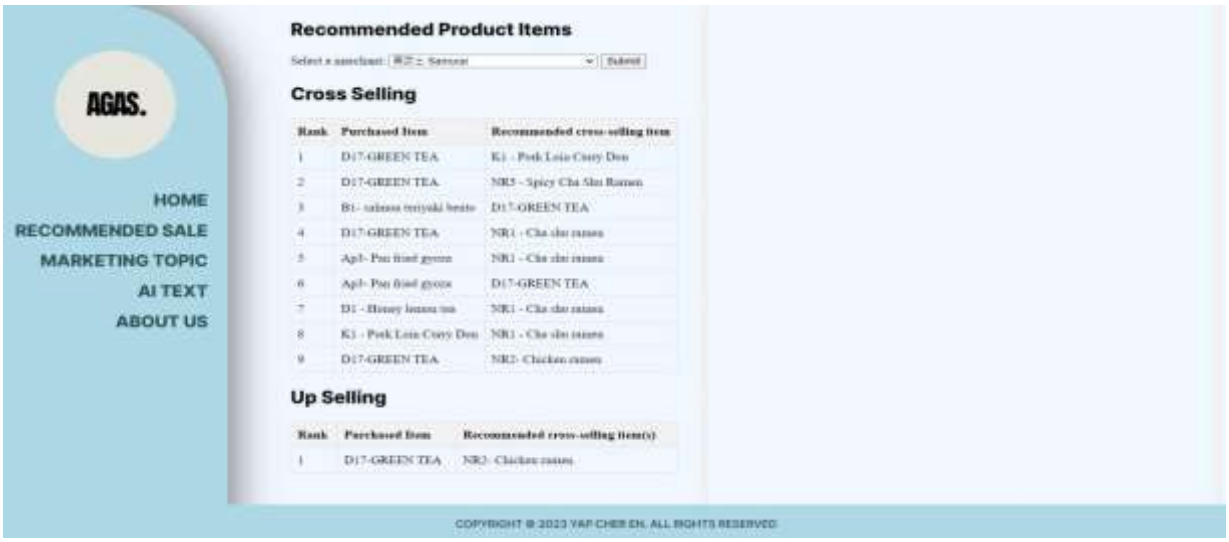

Figure 3: UI of the Recommended Product Items

	name_x	name_y	COUNT
795	D17-GREEN TEA	NR1 - Cha shu ramen	108
1162	NR1 - Cha shu ramen	D17-GREEN TEA	108
1163	NR1 - Cha shu ramen	D17-Green tea hot	83
842	D17-Green tea hot	NR1 - Cha shu ramen	83
717	D1 - Honey lemon tea	NR1 - Cha shu ramen	63

Figure 4: Top 5 sales food combination of merchant "黑武士Samurai"

Figure 4 provides compelling evidence that the items recommended to be sold together are indeed good recommendations. Figure 2 showcases the top 5 food items frequently purchased together based on real-life data. Remarkably, these top 5 food combinations are mostly "green tea" and "Cha shu ramen" food types and align perfectly with the recommendations generated by our market basket analysis algorithm.

This alignment between the recommended items and the actual sales data suggests a strong association and validates the effectiveness of our recommendation system. The fact that the top sales foods from real-life data match the recommended combinations demonstrates that our algorithm is capturing meaningful patterns and accurately identifying items that customers tend to purchase together.

By leveraging market basket analysis, we are able to uncover hidden associations and uncover synergies between different food items. The fact that our recommendations align with real-life data affirms that customers are responding positively to the suggested combinations. This not only provides confidence in the reliability of our recommendations but also highlights the potential for increased sales and customer satisfaction by strategically promoting these recommended food combinations.

Overall, Figure 4 serves as compelling evidence that our market basket analysis is capable of identifying and recommending food items that customers are likely to purchase together. This insight can empower businesses to optimize their product offerings, adjust pricing strategies, and devise targeted marketing campaigns that cater to customer preferences and boost overall sales performance.

DISCUSSION AND CONCLUSION

In conclusion, this study aimed to address the issue of F&B product ID inconsistency by investigating the use of Levenshtein Distance based on F&B product names. By employing this approach, we aimed to achieve more accurate and consistent analysis in understanding customer behavior.

Additionally, we explored the application of clustering and market basket analysis on the similarity ratio obtained from time series data. This allowed us to gain deeper insights into customer behavior patterns and uncover associations among items in customer transactions. Through this analysis, we were able to identifytrends and patterns that shed light on customer preferences, enabling us to make data-driven decisions.

The results of the market basket analysis provided valuable insights into customer buying patterns, which, in turn, informed key business strategies. By optimizing menu offerings based on the identified trends, adjusting pricing according to customer preferences, and developing tailored marketing strategies, businesses can enhance customer satisfaction and drive sales growth.

Overall, this study demonstrated the effectiveness of leveraging Levenshtein Distance, clustering, and market basket analysis in understanding customer behavior and optimizing business strategies. The insights gained from this analysis can guide decision-making processes, leading to improved operational efficiency, increased customer engagement, and ultimately, business success in the competitive F&B industry.

The development of a Flask app to automate the analysis process and provide real-time results makes this tool more accessible to businesses of all sizes. However, there are limitations to the type of datavisualization that can be achieved with the current technology.

However, our study has some limitations, such as the limited type of data visualization. There may be other visualization techniques that could provide additional insights into the sales trends and customer behavior. Additionally, the association rules generated from the market basket analysis may not necessarily indicate causality, as other factors such as marketing promotions or seasonal events may also influence customer purchasing behavior.

Despite these limitations, our study provides practical implications for businesses looking to make data- driven decisions. By leveraging sales trend analysis and market basket analysis, businesses can better understand customer behavior and make informed decisions about product placement and marketing promotions. This can ultimately lead to increased sales and customer satisfaction.

Future research should focus on improving the accuracy and speed of the association rules algorithm to provide even better recommendations. Additionally, further exploration into other data visualization techniques can lead to more insightful and the inclusion of other factors that may influence customer purchasing behavior, such as marketing promotions and seasonal events.

Overall, the use of sales trend analysis and market basket analysis can greatly benefit businesses in understanding consumer behavior and making informed decisions to improve their bottom line.

REFERENCES

[1] 7 Digital Transformation Trends in Food and Beverage Industry. (2022, July 4). Hellonimbly.com. https://hellonimbly.com/7-digital-transformation-trends-in-retail-and-fb-in-sea/

[2] Apolinario, K. (2016). Layout Optimization and Promotional Strategies Design in a Retail Store based on a Market Basket Analysis. https://www.researchgate.net/profile/Andres-Abad-2/publication/304092404_Layout_Optimization_and_Promotional_Strategies_Design_in_a_Retai

[3] l_Store_based_on_a_Market_Basket_Analysis/links/57661ebc08aeb4b9980711bb/Layout-Optimization-and-Promotional-Strategies-Design-in-a-Retail-Store-based-on-a-Market-Basket-Analysis.pdf

[4] Coumar, N. (2020, July 5). Market Basket Analysis — Understand you customer and invest effectively. Medium.

https://medium.com/@nandacoumar/market-basket-analysis-understand-you-customer-and-invest-effectively-c134acf46a46

[5] Deep, M. (2022, March 27). Surprisingly Effective Way To Name Matching In Python. Medium. https://towardsdatascience.com/surprisingly-effective-way-to-name-matching-in-python-1a67328e670e

[6] Efti, S. (2020, January 28). How to perform a sales analysis (step-by-step with methods & metrics). The Close Sales Blog. https://blog.close.com/sales-analysis/

[7] FP Growth Algorithm in Data Mining - Javatpoint. (n.d.). Www.javatpoint.com. https://www.javatpoint.com/fp-growth-algorithm-in-data-mining

[8] Frequent Pattern (FP) Growth Algorithm In Data Mining. (2023). Www.softwaretestinghelp.com. https://www.softwaretestinghelp.com/fp-growth-algorithm-data-mining/

[9] Gupta, A. (2015). Predictive Modeling of Athletic Footwear Sales using Trend Analysis. Journal of Textile and Apparel, Technology and Management, 9(2). https://jtatm.textiles.ncsu.edu/index.php/JTATM/article/view/7535/3602

[10] Gupta, A. (2019, July 21). ML | Frequent Pattern Growth Algorithm. GeeksforGeeks. https://www.geeksforgeeks.org/ml-frequent-pattern-growth-algorithm/

[11] Gupta, S., & Mamtora, R. (2014). A Survey on Association Rule Mining in Market Basket Analysis. International Journal of Information and Computation Technology, 4(4), 409–414. http://ripublication.com/irph/ijict_spl/ijictv4n4spl_11.pdf

[14] Sadiku, K. M. ., Mjaku, G. ., & Qarri , A. . (2023). Consumer Behavior in Decision Making - What Are the Factors?. International Journal of Membrane Science and Technology, 10(3), 1-6. https://doi.org/10.15379/ijmst.vi.1257

[15] Hoefler, R., Tiguman, G. M. B., Galvão, T. F., Ribeiro-Vaz, I., & Silva, M. T. (2023). Trends in sales of antidepressants in Brazil from 2014 to 2020: A time trend analysis with joinpoint regression. Journal of Affective Disorders, 323, 213–218. https://doi.org/10.1016/j.jad.2022.11.069

[16] Kadlaskar, A. (2021, October 2). Market Basket Analysis: A Comprehensive Guide for Businesses. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2021/10/a-comprehensive-guide-on-market-basket-analysis/#h2_4

[17] Kaur, M., & Kang, S. (2016). Market Basket Analysis: Identify the Changing Trends of Market Data Using Association Rule Mining. Procedia Computer Science, 85, 78–85. https://doi.org/10.1016/j.procs.2016.05.180

[18] Kumar, B. (2017). Building Prediction Model using Market Basket Analysis. https://www.researchgate.net/publication/315049886_Building_Prediction_Model_using_Market_Basket_Analysis

[19] Jam, F. A., Rauf, A. S., Husnain, I., Bilal, H. Z., Yasir, A., & Mashood, M. (2014). Identify factors affecting the management of political behavior among bank staff. African Journal of Business Management, 5(23), 9896-9904.

[20] Li, S. (2017, September 25). A Gentle Introduction on Market Basket Analysis — Association Rules. Medium; Towards Data Science. https://towardsdatascience.com/a-gentle-introduction-on-market-basket-analysis-association-rules-fa4b986a40ce

[21] Loshin, D. (2013). Market Basket Analysis - an overview | ScienceDirect Topics. Www.sciencedirect.com. https://www.sciencedirect.com/topics/computer-science/market-basket-analysis

[22] Massaro, A., Barbuzzi, D., Vitti, V., Galiano, A., Aruci, M., & Pirlo, G. (n.d.). Predictive sales analysis according to the effect of weather. https://ceur-ws.org/Vol-1746/paper-09.pdf

[23] Musalem, A. (2018). Market basket analysis insights to support category management. https://www.emerald.com/insight/content/doi/10.1108/EJM-06-2017-0367/full/html

[24] N, I. (2018). Market Basket Analysis of Customer Buying Patterns at Corm Café. https://www.researchgate.net/profile/Norulhidayah-Isa/publication/336148574_Market_Basket_Analysis_of_Customer_Buying_Patterns_at_Corm_Cafe/links/5d92f82e92851c33e94b4737/Market-Basket-Analysis-of-Customer-Buying-Patterns-at-Corm-Cafe.pdf

[25] Nijhuis, M. (2022, March 3). Company Name Matching. DNB—Data Science Hub. https://medium.com/dnb-data-science-hub/company-name-matching-6a6330710334

[26] Ostapenko, R., Herasymenko, Y., Nitsenko, V., Koliadenko, S., Balezentis, T., & Streimikiene, D. (2020). Analysis of Production and Sales of Organic Products in Ukrainian Agricultural Enterprises. Sustainability, 12(8), 3416. https://doi.org/10.3390/su12083416

[27] Rao, A. B., Kiran, J. S., & G, P. (2021). Application of market–basket analysis on healthcare. International Journal of System Assurance Engineering and Management. https://doi.org/10.1007/s13198-021-01298-2

[28] Raschka, S. (n.d.). mlxtend: Machine Learning Library Extensions. PyPI. https://pypi.org/project/mlxtend/

[29] Romano, S. (2020). Time-trend analysis of medicine sales and shortages during COVID-19 outbreak: Data from community pharmacies. Research in Social and Administrative Pharmacy. https://doi.org/10.1016/j.sapharm.2020.05.024

[30] Rosette. (2017, December 12). An Overview of Fuzzy Name Matching Techniques. Rosette Text Analytics. https://www.rosette.com/blog/overview-fuzzy-name-matching-techniques/

[31] Sajidali. (2020). Market Basket Analysis. Kaggle.com. https://www.kaggle.com/code/sajidcse/market-basket-analysis#Apriori-Vs-FP-Growth

[32] Sales Trend Analysis with Examples and Illustrations. (2019). Bizstats.ai. https://www.bizstats.ai/blog/2019/01/11/sales-trend-analysis-with-examples-and-illustrations/

[33] Shrapnel, W. S., & Butcher, B. E. (2020). Sales of Sugar-Sweetened Beverages in Australia: A Trend Analysis from 1997 to 2018. Nutrients, 12(4), 1016. https://doi.org/10.3390/nu12041016

[34] Solnet, D., Boztug, Y., & Dolnicar, S. (2016). An untapped gold mine? Exploring the potential of market basket analysis to grow hotel revenue. International Journal of Hospitality Management, 56, 119–125. https://doi.org/10.1016/j.ijhm.2016.04.013

[35] Tarasenko, O. (2022, January 20). Sales Trend Analysis | Revenue Grid. Revenue Grid: Revenue Operations and Intelligence Platform.

https://revenuegrid.com/blog/sales-trend-analysis/

[36] Tyagi, N. (2021). 5 Applications of Time Series Analysis | Analytics Steps. Www.analyticssteps.com. https://www.analyticssteps.com/blogs/5-applications-time-series-analysis

[37] Wisesa, O., Andriansyah, A., & Ibrahim Khalaf, O. (2020). Prediction Analysis for Business To Business (B2B) Sales of Telecommunication Services using Machine Learning Techniques. MajlesiJournal of Electrical Engineering, 14(4), 145–153. https://doi.org/10.29252/mjee.14.4.145

[38] Yamasaki, D., Tanabe, M., Muraki, Y., Kato, G., Ohmagari, N., & Yagi, T. (2017). The first report of Japanese antimicrobial use measured by national database based on health insurance claims data (2011–2013): comparison with sales data, and trend analysis stratified by antimicrobial category and age group. Infection, 46(2), 207–214. https://doi.org/10.1007/s15010-017-1097-x

[39] Zabolotnyy, O. (n.d.). The key benefits of Machine Learning in Sales. First Bridge. https://firstbridge.io/blog/artificial-intelligence/the-key-benefits-of-machine-learning-in-sales

[40] Zeitouny, S., Suda, K. J., Mitsantisuk, K., Law, M. R., & Tadrous, M. (2021). Mapping global trends in vaccine sales before and during the first wave of the COVID-19 pandemic: a cross-sectional time- series analysis. BMJ Global Health, 6(12), e006874. https://doi.org/10.1136/bmjgh-2021-006874

DOI: https://doi.org/10.15379/ijmst.v10i3.1851