MARKET BASKET AND TIME SERIES ANALYSIS: CASE STUDY PT XYZ SALES DATA

Muhammad Wildan[1], Moh Thaha Rizieq Hentihu[2], Riyanto Jayadi[3]
Universitas Bina Nusantara, Indonesia[123]
Email: muhammad.wildan@binus.ac.id[1], moh.thaha@binus.ac.id[2],
riyanto.jayadi@binus.ac.id[3]

Abstract
PT. XYZ is a retail company that has been operating since 2015 in New York, which seeks to increase sales through attractive package bidding strategies and ensuring stock availability for the following month's highest-selling items. The project aims to analyze and forecast the items that will have the highest sales as well as provide package recommendations that can increase sales. The analysis was carried out using a machine learning model, with the Market Basket Analysis (MBA) method for package recommendations and the Auto Regressive Integrated Moving Average (ARIMA) for sales predictions. The results of the analysis show that the "USB-C Charging Cable" and "Bose SoundSport Headphones" have a strong association with the "Vareebadd Phone", which indicates the potential for increased sales if offered together. In addition, ARIMA's predictions suggest that the MacBook Pro Laptop will generate the highest revenue, with an average projection of around $788,067.45. Based on these findings, we recommend PT. XYZ is preparing a package offering for the Vareebadd Phone with USB-C Charging Cable and Bose SoundSport Headphones, as well as preparing stock for the MacBook Pro Laptop, to meet market demand and maximize revenue potential.

Keywords: Market Basket Analysis, Auto Regressive Integrated Moving Average, Machine learning

*Correspondence Author: Muhammad Wildan
Email: muhammad.wildan@binus.ac.id

INTRODUCTION

PT. XYZ, a retail entity founded in 2015 in New York, specializes in the sale of electronic products through physical and online platforms, including Amazon. In an effort to increase its sales figures, the company integrates package offerings and aims to maintain a sufficient supply of products with high demand every month (Datta et al., 2024; Wu et al., 2022). The novelty of this study lies in the application of two different machine learning methodologies, namely Market Basket Analysis (MBA) to recommend optimal product packages and Auto Regressive Integrated Moving Average (ARIMA) to predict best-selling items. This approach is expected not only to help PT. XYZ in proactively managing inventory, but it also provides deeper insights into market dynamics that can improve their sales and marketing strategies.

The package offer policy provides an opportunity for consumers to acquire some items at a lower price compared to individual purchases. Research by Chen et al. (2016) shows that the implementation of bundling strategies can significantly increase sales volume by utilizing existing product associations (D'Angelo & Minchin, 2021). This is not only attractive to discount seekers, but also to customers looking for complementary

product combinations (Wolle et al., 2019). In addition, the importance of effective inventory management to maintain the availability of popular products so as not to lose customers to competitors has also been discussed in research by Huang and Sarigöllü (2016) (Vermeulen et al., 2023).

To achieve this, an efficient inventory management system is indispensable, which can monitor stock levels in real-time and coordinate inventory across various sales channels, both online and offline. Identifying potential items for bundling requires an in-depth understanding of the relationships between various products (Priessner & Hampl, 2020; Volles et al., 2024). Association analysis plays an important role in this context, as it explores the relationships between items in a dataset. Techniques such as the A priori algorithm can be used to identify sets of items that often occur together in transactions above a certain threshold. By identifying these associations, researchers can strategically place related products to increase sales (Martinez et al., 2021).

Additionally, to understand patterns and predict future trends, predictive analytics is used, often utilizing machine learning techniques to improve accuracy and insights in data handling, pattern detection, and outcome prediction (Dubey et al., 2019; Gunasekaran et al., 2017). Previous research has shown that market analysis has been widely used by various companies to inform their promotional strategies based on identified product associations (Qisman et al., 2021). For example, Limitedbrands has shown how targeted promotional campaigns, such as "buy two, get three" or "buy products, get rewards", can greatly benefit from insights gained from market analysis by ensuring that the promoted items are paired appropriately (Kurnia et al., 2019).

To forecast inventory needs and sales trends, various advanced statistical and machine learning models are applied. These include the Fuzzy Logic (FL) approach to optimize inventory order quantities (Li et al., 2020), Artificial Neural Network (ANN) to predict stock market trends, and the Autoregressive Integrated Moving Average (ARIMA) method to forecast sales and stock prices (Aldino et al., 2021). In this study, we involved two different machine learning methodologies: Market Basket Analysis (MBA) to recommend optimal product packages and Auto Regressive Integrated Moving Average (ARIMA) to predict best-selling items, thus allowing PT. XYZ to proactively manage inventory in order to meet anticipated demand. By combining these two approaches, this research provides a strong foundation for PT. XYZ to improve their sales and inventory management strategies, as well as strengthen their competitive position in the electronics retail market.

RESEARCH METHODS

Data Understanding

This data comprises sales transaction data from January to December 2019, totaling 185,919 transactions. Among these transactions, 7,136 indicate more than 1 product being bought together (Monteserin & Armentano, 2018).

Table 1. Data Understanding Table

Attribute name	Data type	Description
OrderID	Numerical - Discrete	Unique identifier assigned to each individual order or purchase made by a customer.
Product	Categorical - Nominal	The name or description of the item(s) that were purchased in the order
Quantity	Numerical - Discrete	The number of units or items of a particular product that were purchased in the order.
Price	Numerical - Discrete	The cost or price of a single unit of the product
OrderDate	Categorical - Ordinal	The date and time when the order was placed by the customer
Purchase Address	Categorical – Nominal	The location where the order was delivered or the customer's billing address

Data Preparation

1) Merging data

The data we get is sales data per month for 1 year in 2019, therefore we first merge the data before continuing with the next data preparation process and also import data into RapidMiner.

Fig 1. Merge Data

2) Delete Empty Row

In the data we get, after several rows of data, there will usually be a row whose contents are empty before moving on to the next row, and this happens consistently across all sheets. Because of this, we delete this empty row so that the data remains sorted from above without being truncated.

Fig 2. Empty Row

3) Reformat Order Dat

In the data we got, there were inconsistent date data formats, some were automatically converted to date, but some were still strings. This is one of the factors that makes it difficult to enter existing data into rapidminer. Therefore, we need to change the date format to be consistent with the same date format.

Fig 3. Reformat Order Date

Modelling

a. Market Basket Analysis

The Market Basket Analysis (MBA) model has been made using rapidminer. In order to make the data usable in rapidminer, the data that has been preprocessed by merging and performing necessary data cleaning steps needs to be processed again using rapidminer. We first import our data into rapidminer and start the data manipulation using rapidminer operators.
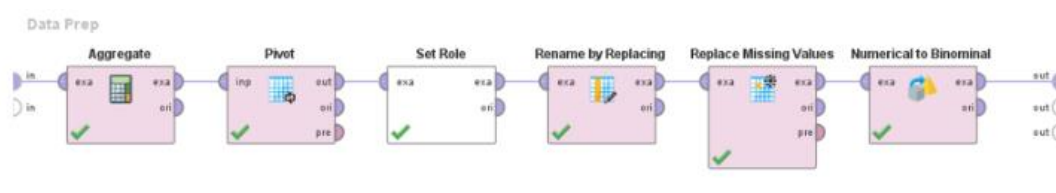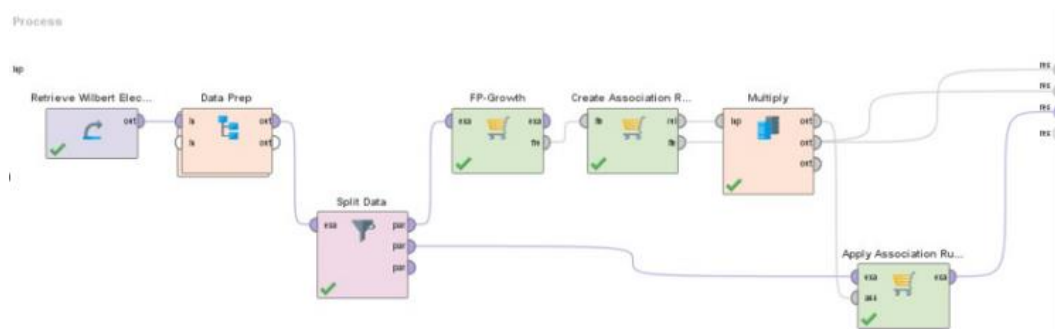

Fig 4. Market Basket Model

Our goal here is to make our data set readable by Rapidminer and to do that the operators we used are Aggregate, Pivot, Set Role, Rename by Replacing, Replacing Missing Values, and Numerical to Binomial. On the aggregation attributes we choose Quantity Ordered with sum functions, and we group it by Order ID and Product. and by using Pivot we group the attribute by Order ID and column grouping attribute using Product also setting the aggregation attributes by Quantity Ordered. We set the role of Order ID to id and simplify the name of each product using the Rename By Replacing operator. We replace the missing value with zero and change the numerical value to binomial. From this data preparation the output that we want is to place the Order ID on the left column and placing all the Product Name on the top row. Each column under the Product Name will be true if the product is purchased from the Order ID on the left column, otherwise it will be false. The output will be shown using the Apply Association Rules operator.

Fig 5. Market Basket Model Extend

The model for Market Basket Analysis can be seen on the figure. The data that has been formatted is placed on the Data Prep operator. Here we apply FP-Growth and Association Rules in our model. By implementing FP-Growth our aim here is to construct an FP-Tree and mine frequent items. We also apply Association rules to find the important parameters on making decisions for bundling which are Support, Confidence, and lift. Support is the percentage of transactions that contain all of the items in an itemset. High support values indicate that the items in the itemset are often bought together, making them good candidates for bundling. Confidence is the probability that a transaction that contains the items on the left hand side of the rule also contains the item on the right hand side. Higher confidence values suggest a stronger likelihood that the presence of itemset A will be followed with the presence of itemset B.

Lift is the probability of all of the items in a rule occurring together divided by the product of the probabilities of the items on the left and right hand side occurring as if there was no association between them. If the lift is greater than 1, it implies a positive correlation between items A and B. This indicates that the presence of itemset A in a transaction increases the likelihood of itemset B being present. If the lift is equal to 1, it indicates that there's no association between the two itemsets. If the lift is less than 1, it implies that customers who buy itemset A are less likely to buy itemset B. After importing the data and doing the required data reprocessing, we divide the data using the Split Data operator into FP-Growth and Apply Association Rules operator. On our model the parameter on FP-Growth is set on default and connected to the Create Association Rules operator. We set the criterion in Create Association Rules operator to confidence and set the confidence to 0.001

to get more results. Finally, we use the Multiply operator to copy the format from Create Association Rules to output and into the Apply Association Rules operator. In order to get the result from FP-Growth we connect the item sets of the Create Association Rules operator to the output. Finally, by giving the Apply Association rules operator the required data we connect the example set into the output. This ends us with 3 types of output which are Frequent item sets (FP-Growth), Association Rules (Create Association Rules), and Example sets (Apply Association Rules).

b. ARIMA (auto Regressive Integrated Moving Average)

The ARIMA model is performed using Python. Before performing modelling, the dataset will require additional preparation by splitting data into training and testing sets (Kaur & Kang, 2016). The process of splitting the dataset will be based on the date and time marker on each transaction that happened. Below, is showed the code on how the data will be split into two different sets based on the date.

```python
# Converting 'Order Date' to datetime format
df['Order Date'] = pd.to_datetime(df['Order Date'])

# Grouping the data by 'Product' and 'Order Date', and summing up the 'Total Price'
data_grouped = df.groupby(['Product', df['Order Date'].dt.date]).sum()['Total Price'].reset_index()

# Converting 'Order Date' back to datetime format
data_grouped['Order Date'] = pd.to_datetime(data_grouped['Order Date'])

# Initializing empty dictionaries to hold the training and testing datasets for each product
train_data_dict = {}
test_data_dict = {}

# Looping through each unique product to create training and testing datasets
for product in data_grouped['Product'].unique():
    product_data = data_grouped[data_grouped['Product'] == product]

    # Filter data for training: from January 1st to November 30th
    train_data = product_data[(product_data['Order Date'] >= '2019-01-01') & (product_data['Order Date'] <= '2019-11-30')]

    # Filter data for testing: Dates other than January 1st to November 30th
    test_data = product_data[(product_data['Order Date'] < '2019-01-01') | (product_data['Order Date'] > '2019-11-30')]

    # Storing the training and testing data into dictionaries with the product name as the key
    train_data_dict[product] = train_data
    test_data_dict[product] = test_data

# Let's check the first product's training and testing datasets as a sample
first_product = list(train_data_dict.keys())[0]
train_data_dict[first_product].head(), test_data_dict[first_product].head()
```

Fig 6. Arima Model

In this process, the training dataset is all the transactions of every product that happened from January 1st, 2019, until November 30th, 2019. For the testing dataset, it will be filled with all transactions of every product from December 1st, 2019, until December 31st, 2019.

```
# Function to train and test ARIMA model with arbitrary best-practice parameters (5,1,2)
def train_test_arima(product_data_train, product_data_test, order=(5,1,2)):
    """
    Trains an ARIMA model on the training dataset and tests it on the testing dataset.

    Parameters:
    - product_data_train: DataFrame containing the training data
    - product_data_test: DataFrame containing the testing data
    - order: Tuple containing the (p, d, q) parameters for the ARIMA model

    Returns:
    - A plot comparing the training, testing, and predicted data
    """
    # Preparing the data
    train_series = product_data_train['Total Price']
    test_series = product_data_test['Total Price']

    # Creating the ARIMA model and fitting it to the training data
    model = ARIMA(train_series, order=order)
    model_fit = model.fit()

    # Making predictions
    start = len(train_series)
    end = start + len(test_series) - 1
    predictions = model_fit.predict(start=start, end=end, typ='levels')
```

Fig 7. Arima Model Extend

After split the datasets we will be using method ARIMA that we can just import from statmodels package (statsmodels.tsa.arima.model). The ARIMA will have 3 important parameters: p is the number of autoregressive terms, d is the number of nonseasonal differences needed for stationarity, and q is the number of lagged forecast errors in the prediction equation. For the initial training, we will use p = 5, d = 1, and q = 2 as they are known for the best number for the ARIMA setup. Then we will do model.fit() to fit the training dataset into the model with set parameters. Then, we can do testing by using model_fit.predict into the testing dataset so that we are able to know how the model is performed.

RESULTS AND DISCUSSION

Market Basket Analysis

	Size	Support ↓	Item 1	Item 2	Item 3
No. of Sets: 115	1	0.10596	Wired Headphones		
Total Max. Size: 3	1	0.08672	Apple Airpods Headphones		
	1	0.07417	Bose SoundSport Headpho...		
Min. Size: 1	1	0.04222	27in FHD Monitor		
Max. Size: 3	1	0.03786	iPhone		
Contains Item:	1	0.03480	34in Ultrawide Monitor		
	1	0.03472	27in 4K Gaming Monitor		
Update View	1	0.03070	Google Phone		
	1	0.02685	Flatscreen TV		
	1	0.02655	Macbook Pro Laptop		
	1	0.02309	20in Monitor		
	1	0.02273	ThinkPad Laptop		
	1	0.01146	Vareebadd Phone		
	2	0.00558	USB-C Charging Cable	Google Phone	
	2	0.00553	Lightning Charging Cable	iPhone	

Fig 8. Market Analysis Result

The figure shows the result of the data that shows the products and products bought together with the highest support by using the FP-Growth operator in Rapidminer. From this data, it can be seen that the product most purchased is Wired Headphones with the support of 0.10596 and most purchased products together with confidence of 0.00558, which consist of USB-C Charging Cable and Google Phone. Based on this information, we can see that the item bought together is connected in terms of being used together. This will help us in making the decision on which products should be bundled together.

No.	Premises	Conclusion	Support	Confide... ↓	Lift
203	Lightning Charging Cable, Wired Headphones	iPhone	0.00034	0.50549	13.3509
202	USB-C Charging Cable, Wired Headphones	Google Phone	0.00049	0.42038	13.6916
201	Lightning Charging Cable, Apple Airpods Headphones	iPhone	0.00025	0.40964	10.8192
200	USB-C Charging Cable, Bose SoundSport Headphones	Google Phone	0.00017	0.29114	9.48225
199	Wired Headphones, Bose SoundSport Headphones	Google Phone	0.00010	0.26415	8.60326

No.	Premises	Conclusion	Support	Confidence	Lift ↓
193	USB-C Charging Cable, Bose SoundSport Headphones	Vareebadd Phone	0.00010	0.17722	15.47055
202	USB-C Charging Cable, Wired Headphones	Google Phone	0.00049	0.42038	13.69163
203	Lightning Charging Cable, Wired Headphones	iPhone	0.00034	0.50549	13.35096
191	USB-C Charging Cable, Wired Headphones	Vareebadd Phone	0.00017	0.14650	12.78889
201	Lightning Charging Cable, Apple Airpods Headphones	iPhone	0.00025	0.40964	10.81924

Fig 9. Association rules

The result we got from the Create Association Rules operator shows the relation of each possible item combination. This result also shows all required parameters to make the important decision for the bundling of the product which is Support, Confidence, and Lift. We sort the result table to show the highest possible Support, Confidence, and Lift to decide the best possible items to bundle depending on the purpose of the bundle. Initial inspection of the dataset via the Create Association Rules operator shows recurring item sets, with the standout being the "Google Phone" and "USB-C Charging Cable" pairing, boasting the highest support value of 0.00558. This finding indicated a strong tendency for these items to be purchased together, prompting consideration of a bundling strategy to enhance sales.

Adding to this, the analysis shows the item set with the highest confidence with a notable 0.50549, connecting "Lightning Charging Cable," "Wired Headphones," and "iPhone." This high confidence underscored a robust association, implying that customers purchasing the latter item also consistently opted for the former, reinforcing cross-selling and promotional opportunities. Furthermore, the highest lift value of 15.47055 shows a connection for "USB-C Charging Cable" and "Bose SoundSport Headphones" combination in pair with the "Vareebadd Phone". With this, the idea of lift goes beyond just showing connections. It suggests that when you buy certain items together, it makes it much more likely that you'll also end up getting the "Vareebadd Phone". Discovering the highest support, confidence, or lift values for itemsets unveils significant purchase patterns. High support signifies frequently co-purchased items, guiding inventory decisions. High confidence highlights strong item associations, informing cross-selling approaches. A high lift value indicates meaningful item connections beyond chance, aiding bundling and

promotional strategies. Ultimately, these insights empower businesses to optimize inventory, enhance cross-selling, and strategically bundle items for improved customer experience and revenue growth.

Based on our analysis of the association rules, we have identified several promising opportunities for bundling products. The itemset with the highest support, consisting of the "Google Phone" and "USB-C Charging Cable," presents an ideal pairing that could be marketed together as a bundle, potentially boosting sales for both items. Additionally, we observed an intriguing relationship involving the "USB-C Charging Cable," "Bose SoundSport Headphones," and "Vareebadd Phone." Specifically, when customers purchase the charging cable and the headphones, they are likely to also buy the Vareebadd Phone. This insight suggests a strategic opportunity for Willbert Electronics to enhance product placement by positioning these items in close proximity to one another or creating a bundled offer that includes all three products.

The impact of such bundling strategies on business performance could be significant. By leveraging the identified associations, Willbert Electronics can increase the average transaction value and improve customer satisfaction through convenience, as customers appreciate the ease of purchasing complementary products together. Furthermore, these bundles can attract price-sensitive consumers who are looking for deals, thereby expanding the customer base. From a stock management perspective, accurately predicting the demand for bundled items can lead to more effective inventory management, ensuring that popular products are readily available and reducing the risk of stockouts. This proactive approach to inventory forecasting not only enhances sales but also strengthens customer loyalty, as satisfied customers are more likely to return for future purchases.

Regression Analysis

From the modelling, we can predict all of the product selling performance within the month of December. Here we will be using a green line as a marker for the prediction result for the next 30 days after November 2019. In this analysis, we will only show the top 5 products with the highest selling dollar based in the month of December. Below are showed 5 plots showing the blue line which was the data used to train the model, the orange line that is used for testing the model, and the green line is the predicted revenue.
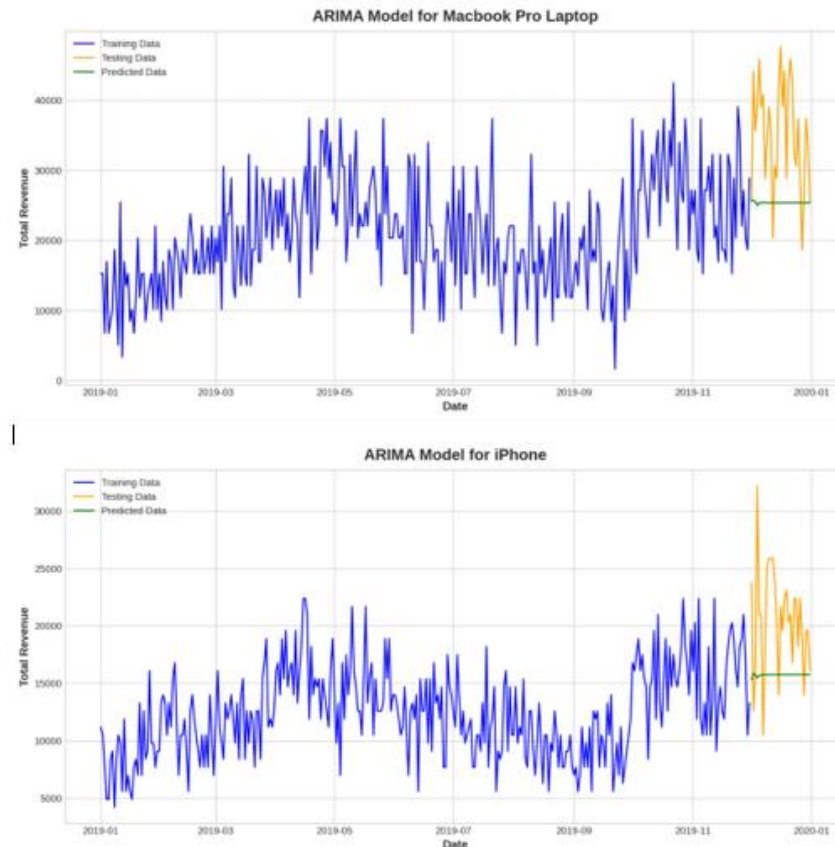
Fig 10. Arima Result

The plots above are showed to be the top 5 highest products that create revenue the most. In details, the table below showing the predicted revenue in the month of December for the top 5 products, by total amount, and by the daily average

Table 2. Highest Product

Product	Total Predicted Revenue	Daily Average Predicted Revenue
Macbook Pro Laptop	788,067.45	25,421.53
iPhone	488,233.72	15,749.47
ThinkPad Laptop	363,486.46	11,725.37
Google Phone	305,502.62	9,854.92
27in 4K Gaming Monitor	255,158.02	8,230.90

The ARIMA model evaluated by using Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). Both metrics used for standard measurement for regression analysis. The RMSE can tell how many units the model is wrong on average. While MAPE tells how wrong the forecasts are percentage- wise, for example if the MAPE value 0.02 it means that the forecast are 98% accurate.

The overall error metrics for the ARIMA models across all products are as follows:
a. Overall Root Mean Square Error (RMSE): 4049.61
b. Overall Mean Absolute Percentage Error (MAPE): 0.288 or 28.8

The RMSE of 4049.61 indicates that, on average, the model's predictions are around 4049.61 units away from the actual values. The MAPE of 28.8 suggests that the model's predictions are off by about 28.8%. Based on the predictions result above, related to the

stocks, it is advisable for the owner to bundle product that includes the one of the or both form the 5 products that dominate the revenue the most.

CONCLUSION

Exploring the realm of Market Basket Analysis (MBA) for bundling in an electronics store environment presents several promising avenues for future development. Considering the dynamic nature of electronics and rapid technological advancements, integrating product lifecycle data could unveil insights into optimal bundling strategies that align with product trends and obsolescence. Additionally, incorporating external factors such as customer reviews, ratings, and online discussions related to electronic products could enrich the analysis, while integrating sentiment analysis would provide a deeper understanding of consumer preferences and the effectiveness of the suggested bundling strategy. Currently, the ARIMA model produces a flat prediction line, indicating it forecasts the same value for all future data points; to address this, transforming data to stabilize variance and decomposing seasonal components could enhance accuracy. Feature engineering techniques, such as using lagged variables and incorporating external variables as exogenous factors, may also yield better predictions. The implications of this research extend beyond sales optimization, as refining bundling strategies and enhancing inventory forecasting can improve customer satisfaction and loyalty, ultimately driving long-term profitability. Furthermore, integrating customer sentiment analysis can lead to more targeted marketing efforts, aligning offerings with consumer desires, while future studies might explore the impact of promotional campaigns on bundling effectiveness, contributing valuable insights to the broader field of retail analytics.

BIBLIOGRAPHY

Aldino, A. A., Pratiwi, E. D., Setiawansyah, Sintaro, S., & Putra, A. D. (2021). Comparison of Market Basket Analysis to Determine Consumer Purchasing Patterns Using Fp-Growth and Apriori Algorithm. *2021 International Conference on Computer Science, Information Technology, and Electrical Engineering, ICOMITEE 2021.* https://doi.org/10.1109/ICOMITEE53461.2021.9650317

D'Angelo, D., & Minchin, R. E. (2021). Project bundling in transportation construction. *Proceedings of International Structural Engineering and Construction*, *8*(1). https://doi.org/10.14455/ISEC.2021.8(1).CON-09

Datta, A., Sarkar, B., Dey, B. K., Sangal, I., Yang, L., Fan, S.-K. S., Sardar, S. K., & Thangavelu, L. (2024). The impact of sales effort on a dual-channel dynamical system under a price-sensitive stochastic demand. *Journal of Retailing and Consumer Services*, *76*, 103561. https://doi.org/10.1016/j.jretconser.2023.103561

Dubey, R., Gunasekaran, A., Childe, S. J., Papadopoulos, T., Luo, Z., Wamba, S. F., & Roubaud, D. (2019). Can big data and predictive analytics improve social and environmental sustainability? *Technological Forecasting and Social Change*, *144*, 534–545. https://doi.org/10.1016/j.techfore.2017.06.020

Gunasekaran, A., Papadopoulos, T., Dubey, R., Wamba, S. F., Childe, S. J., Hazen, B., & Akter, S. (2017). Big data and predictive analytics for supply chain and organizational performance. *Journal of Business Research*, *70*, 308–317. https://doi.org/10.1016/j.jbusres.2016.08.004

Kaur, M., & Kang, S. (2016). Market Basket Analysis: Identify the Changing Trends of Market Data Using Association Rule Mining. *Procedia Computer Science*, *85*.

https://doi.org/10.1016/j.procs.2016.05.180

Kurnia, Y., Isharianto, Y., Giap, Y. C., Hermawan, A., & Riki. (2019). Study of application of data mining market basket analysis for knowing sales pattern (association of items) at the O! Fish restaurant using apriori algorithm. *Journal of Physics: Conference Series*, *1175*(1). https://doi.org/10.1088/1742-6596/1175/1/012047

Li, H., Wu, Y. J., & Chen, Y. (2020). Time is money: Dynamic-model-based time series data-mining for correlation analysis of commodity sales. *Journal of Computational and Applied Mathematics*, *370*. https://doi.org/10.1016/j.cam.2019.112659

Martinez, M., Escobar, B., García-Díaz, M. E., & Pinto-Roa, D. P. (2021). Market basket analysis with association rules in the retail sector using Orange. Case study: Appliances sales company. *CLEI Eletronic Journal (CLEIej)*, *24*(2). https://doi.org/10.19153/cleiej.24.2.12

Monteserin, A., & Armentano, M. G. (2018). Influence-based approach to market basket analysis. *Information Systems*, *78*. https://doi.org/10.1016/j.is.2018.01.008

Priessner, A., & Hampl, N. (2020). Can product bundling increase the joint adoption of electric vehicles, solar panels and battery storage? Explorative evidence from a choice-based conjoint study in Austria. *Ecological Economics*, *167*, 106381. https://doi.org/10.1016/j.ecolecon.2019.106381

Qisman, M., Rosadi, R., & Abdullah, A. S. (2021). Market basket analysis using apriori algorithm to find consumer patterns in buying goods through transaction data (case study of Mizan computer retail stores). *Journal of Physics: Conference Series*, *1722*(1). https://doi.org/10.1088/1742-6596/1722/1/012020

Vermeulen, H., Meyer, F., & Schönfeldt, H. C. (2023). A basic healthy food basket approach to evaluate the affordability of healthy eating in South Africa and Kenya. *Frontiers in Sustainable Food Systems*, *7*. https://doi.org/10.3389/fsufs.2023.1181683

Volles, B. K., Ribbers, D., Van Kerckhove, A., & Geuens, M. (2024). Beyond bundles: Choosing product bundles increases shopping basket size. *Journal of Retailing and Consumer Services*, *81*, 104035. https://doi.org/10.1016/j.jretconser.2024.104035

Wolle, M. M., Stadig, S., & Conklin, S. D. (2019). Market basket survey of arsenic species in the top ten most consumed seafoods in the United States. *Journal of Agricultural and Food Chemistry*, *67*(29). https://doi.org/10.1021/acs.jafc.9b02314

Wu, X., Zha, Y., & Yu, Y. (2022). Asymmetric retailers' sales effort competition in the presence of a manufacturer's help. *Transportation Research Part E: Logistics and Transportation Review*, *159*, 102625. https://doi.org/10.1016/j.tre.2022.102625