

UNDERSTANDING WHOLESALE CLIENT BEHAVIOR AT A VIETNAMESE GARMENT OUTSOURCING FACTORY

Lu Thi Kim Phung^{1,*}, Nguyen Duy Dat²

¹*Faculty of Science, Dong Nai University, Vietnam*

²*CEO, Fide Tech Co., Ltd, Ho Chi Minh City, Vietnam*

*Email: luphungw@gmail.com

Received: 12 October 2025; Revised: 16 March 2026; Accepted: 9 April 2026

ABSTRACT

To increase sales, businesses need to understand customer behavior in both wholesale and retail markets. This includes considering crucial factors like product type, timing, and location. This study, based on data from a Vietnamese garment outsourcing factory, reveals distinct differences in consumer behavior between the country's wholesale and retail sectors. The findings allowed us to categorize customers into two groups: Key accounts, who contribute over 50% of sales value and all other customers. The research also examines the unique needs of each group and proposes effective strategies for engaging with them. Ultimately, this study introduces a new approach for those working in Vietnam's wholesale and retail industries.

Keywords: Consumer behavior, outsourcing company, Pearson correlation coefficient, wholesale buyer.

1. INTRODUCTION

This study proposes a novel method to address the needs of wholesale buyers in Vietnam's garment industry, offering practical criteria and solutions to support the sector's future growth. Vietnam, with a population of approximately 93 million, is a leading global garment exporter, consistently ranking among the top five worldwide. Its rapid development is demonstrated by export growth from \$2.4 billion in 2002 to \$27 billion in 2015 and \$36 billion in 2018.

Most Vietnamese garment firms operate as outsourcing manufacturers for global brands and face major challenges. These include heavy dependence on imports, with 70% of silk fabric and materials sourced from mainland China, and limited access to modern machinery. About 65% of equipment was acquired before 2000, 20% after 2005, and only 15% after 2015. As a result, companies are highly dependent on global market demand driven by fashion retailers and wholesale buyers.

Addressing these challenges requires a clear understanding of wholesale customer preferences to develop effective sales strategies. This study derives adaptive solutions from a private dataset provided by a garment outsourcing company in Ho Chi Minh City and compares the results with the UK Online Retail dataset.

2. PREVIOUS WORK

Past research has explored various approaches to understanding customer behavior and sales. For instance, Vahid [1] utilized Wi-Fi tracking and other data to forecast a coffee shop's revenue based on customer visit duration. Daqing Chen and colleagues [2] employed an RFM

(Recency, Frequency, Monetary) model and other features for customer classification. In contrast, our study will categorize customers based on their revenue contributions.

While other researchers [3, 4, 5, 6] have applied data mining techniques, this paper prioritizes research efficiency by directly addressing two key questions:

- What are the specific needs of different customer segments?
- How can a CEO adapt their strategies to meet these needs effectively?

It's important to note that much of the existing research on this topic relies on survey data rather than actual sales data. This often leads to a lack of reliable and detailed information, which is why we will not extensively refer to those studies in later sections.

3. PROPOSE METHOD

This study aims to identify emerging trends across two customer groups in a garment outsourcing company. We analyze key factors influencing purchasing decisions and examine relationships among customer features, with a focus on differences between Vietnamese wholesale buyers and international retail customers. The findings support more effective customer engagement strategies and revenue growth.

The proposed method follows a structured workflow: (1) Data Collection, (2) Data Pre-processing, (3) Customer Grouping and Characterization, (4) Behavior Comparison between key buyers and other customers, (5) Monthly Revenue Prediction, and (6) Strategy Recommendation for each customer group.

3.1. Data collection

The dataset for this study was collected from a garment outsourcing factory in Vietnam, spanning from October 1, 2017, to August 31, 2019. Although data collection began in May 2017, the initial months had limited billing information, so we focused on a more complete period. The dataset includes standard billing details such as: Receipt ID, Customer ID, Requirement, Amount, Price, Created Time.

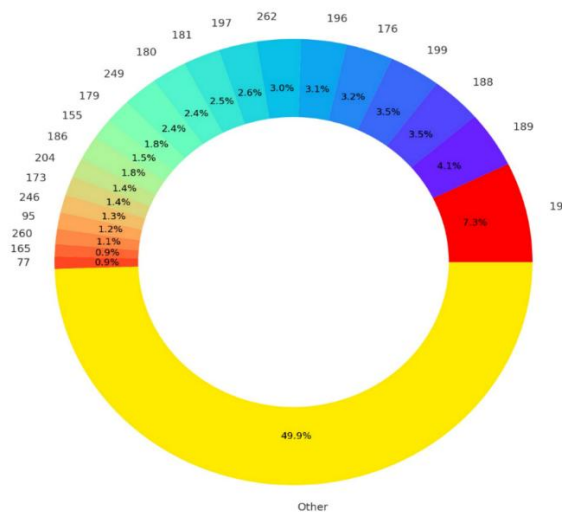


Fig. 1. Percentages of top buyers in group 1 (top) and the rest buyers in group 2 (bottom)

For this analysis, we exclusively used these relevant features. Other information, like customer names, locations, or job descriptions, was excluded as it wasn't pertinent to our research and was omitted to protect data privacy.

3.2. Data preprocessing

To uncover hidden relationships between our data and real-world customer behavior, we will combine existing features to generate new ones. This will help us understand how revenue, items, and time interact from various perspectives.

Since our dataset was meticulously compiled, all features are reliable. We will create essential new features to provide both a broad overview and detailed insights into specific customer groups and revenue fluctuations.

For instance, the "Requirement" feature is rich with details like fabric type, silk depth, color, and processing instructions. We will extract a new feature focusing solely on silk type, as identifying the best-selling fabric is crucial for the factory owner. This insight allows them to adjust material procurement more accurately based on product trends.

From the "Created Time" field, we can extract valuable information such as the week of the year, week within the dataset, whether it's a holiday or workday, month, and year. This data will enable us to analyze revenue and best-selling items over time, revealing the shopping trends of different customer segments. For example, we can explore the relationship between revenue and preferred materials or track changes in revenue across various silk types in recent months.

3.3. Customer grouping

To effectively analyze the varied behaviors of both high-value and lower-value customers, we've established a threshold of 50% of the total product value to divide our customers into two distinct groups:

- **Group 1: Full Outsourcing Clients** These are primarily clothing manufacturers who design their own garments and then contract our factory for the entire production process. Collectively, they contribute over 50% of the company's total revenue.
- **Partial Outsourcing Clients** This group consists mainly of small shop owners who produce nearly finished products and then engage our company for the final stages of the manufacturing process. They account for less than 50% of the company's profit.

Group 1 is comprised entirely of large business owners. In contrast, Group 2 encompasses numerous smaller shop owners and companies, many of whom employ fewer than ten workers.

Figure 1 illustrates Group 1's significant contribution to the factory's revenue. Our dataset includes 400 customers, with a total revenue of approximately \$2.78 million USD over two years. Notably, just 21 key buyers in Group 1 generate 50.1% of this revenue, while the remaining 379 customers in Group 2 contribute 49.9% of the sales value. In subsequent sections, we will delve deeper into the data to uncover the reasons behind these differing contributions.

3.4. Analyzing shopping behaviors in two groups

Our analysis begins with a global overview to identify which customer groups or features contribute most significantly to revenue. Following this, we will individually examine highly correlated factors to uncover any unusual patterns. Finally, we will provide logical explanations by connecting data evidence with additional insights.

Given the overview presented in Figure 1, we now turn to Table 1, which details the relationship between the shopping frequency and total order value of our top buyers. We used the entire dataset to identify the most frequent customers (based on the number of bills they generated) and their corresponding total purchase values. Customers are then ranked according to these past actions. Note that Table 1 lists customer IDs.

Table 1. Customer ranking by purchase frequency and total revenue

Rank	Frequency Ranking	Total Revenue Ranking
1	191	191
2	145	189
3	262	188
4	84	199
5	197	176
6	176	196
7	203	262
8	260	197
9	66	181
10	188	180
11	148	249
12	204	179
13	179	155
14	172	186
15	77	204
16	64	173
17	199	246
18	116	95
19	189	260
20	150	165
21	95	77

Table 1 highlights a significant disparity between a customer's total revenue contribution and their shopping frequency. While some customers appear on both lists, others do not. This can be attributed to the nature of their businesses: small shop owners typically order small batches, often fewer than 100 units per order. In contrast, larger clothing companies or distribution agents operate on a much bigger scale, supplying multiple shops and even setting new fashion trends. Consequently, their orders range from 300 to a minimum of 6,000 units.

Consider customer IDs 191 and 145 as specific examples. Customer 191 is a standout, leading both in buying frequency and total money spent. Customer 145, though not part of Group 1 (our high-value customers), ranks second in the total number of bills placed. To put this in perspective, customer 191 generated over 900 receipts, while customer 145 made approximately 540 requests.

With additional context, we know that customer 191 owns a medium-sized clothing company, accounting for a substantial 7.3% of total sales value. This company frequently orders clothes made from sesame and spandex fabrics in various sizes. Their material preference likely stems from their customer base being predominantly in Southern Vietnam, a hot, tropical region. These fabrics are ideal for sportswear and everyday items due to their excellent sweat permeability. Conversely, customer 180 frequently requests collar processing, indicating they lack the necessary machinery or workforce to complete their products in-house.

To understand how revenue changes across our two customer groups, we can look at the quarterly revenue bar chart in Figure 2. This chart reveals clear trends for both groups. Since

we only have data for July and August of 2019, we've multiplied the third-quarter 2019 data by 1.5 to estimate a full quarter.

Both groups generally show a downtrend in revenue during the first quarter of each year.

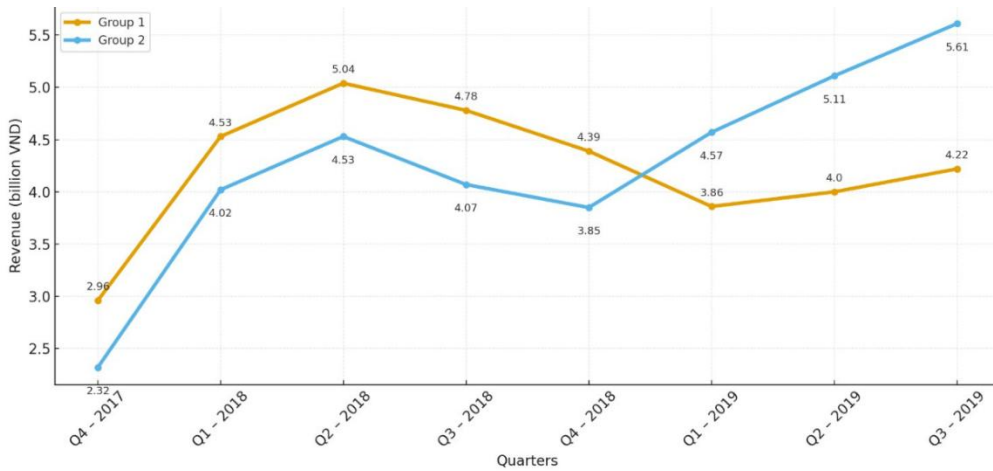


Fig. 2. Compare revenue from two groups in recent quarters

However, overall sales tend to increase in subsequent years, especially within Group 2. Specifically, Group 1's total contribution began at 2.9 billion VND, surged by 40% in 2018, and then slightly decreased in 2019. In contrast, Group 2's revenue started at 2.3 billion VND and significantly jumped to 5.6 billion VND. For the first five quarters, Group 1's sales consistently outpaced Group 2's by approximately 500 to 700 million VND. However, in the last three quarters, Group 2's total sales saw a noticeable increase, starting at over half a billion VND higher and ultimately reaching a 1.4 billion VND difference in total sales value.

The data also highlights the most popular items sold for each group. For Group 1 (full outsourcing), sesame fabric is the top seller. Within this, white color accounts for a substantial 66% of total sesame silk value, followed by black at 14%, and sesame rubber fabric at 7%. Other types of sesame fabric, such as hardened sesame, gray, blue, yellow, and red, make up the rest. Group 2 (partial outsourcing) has a much wider range of requirements, five times more diverse than Group 1's clothing manufacturing needs. Most of their requests involve hardening shirt collars using sesame fabric, which makes up about 38.5% of their outsourcing value.

For a more detailed view, Figure 3, showing sales per month, helps us predict future business conditions. The data indicates that the partial outsourcing group (Group 2) acquired more new customers in 2019 than in 2018. Furthermore, they brought in enough new orders to offset the revenue decline from Group 1, thus balancing the factory's total income.

Monthly sales generally start slowly, but a clear trend emerges from the third month onwards. February consistently shows the lowest sales due to the Vietnamese Lunar New Year (Tet). January and March are typically the best-selling months, especially the three weeks immediately before and after the New Year holidays.

In Vietnam, there's a strong tradition where workers often spend their end-of-year bonuses, typically an extra month's salary, in the month leading up to the Lunar New Year, particularly in the final week. This consumer spending surge means garment factories need to prepare a month in advance to ensure products reach stores punctually.



Fig. 3. Compare revenue from two groups in recent months.

The weeks immediately following the Lunar New Year holiday are also crucial for sales. During this period, consumers often purchase gifts for friends and colleagues, making items like clothes, shoes, and silk products highly popular due to their reasonable prices.

However, a subtle difference in purchasing behavior emerges while mid-range clothing sells exceptionally well, requests for outsourced products are typically delayed by a week before experiencing a sharp rise two weeks later. This delay is likely because consumers frequently update their fashion wardrobes, whereas silk items are replaced less often throughout the year.

Another peak time for selling silk products is the week leading up to Vietnam's Independence Day on September 2nd, which also marks the start of the new school year. This confluence creates a significant demand for silk shoes, clothes, and bags for students. This explains the sudden surge in outsourcing orders from small business owners about two weeks before the new school year begins.

In the subsequent months, revenue generally experiences a slight decline until the end of December. Overall, the data indicates that 2019's monthly revenue was more stable compared to 2018. Analyzing the times when most bills are created reveals our customers' preferred ordering hours. A stark contrast emerges when comparing the shopping habits of UK online retail customers with those of Vietnamese wholesale customers.

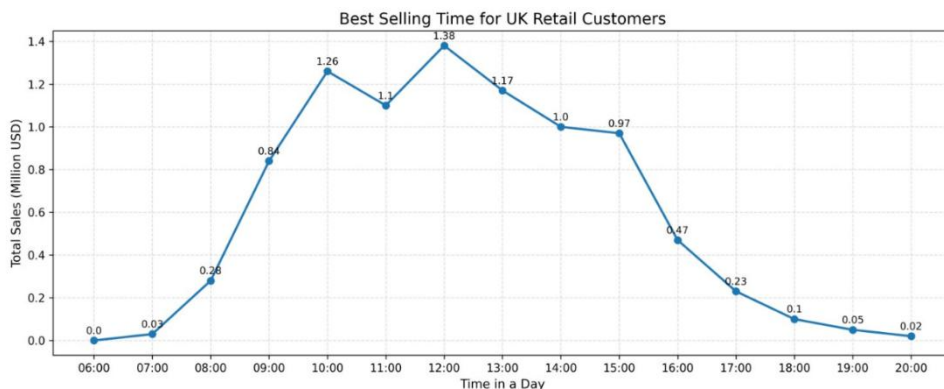


Fig. 4. Bestselling time for UK online retail customers.

Figure 4 indicates that retail customers in the UK prefer to make purchases during the daytime, particularly around noon (10 AM to 2 PM). Conversely, Figure 5 shows that Vietnamese wholesale buyers tend to place orders during the beginning of the morning shift (9 AM to 10 AM) and the closing hours of the afternoon shift (3 PM to 5 PM).

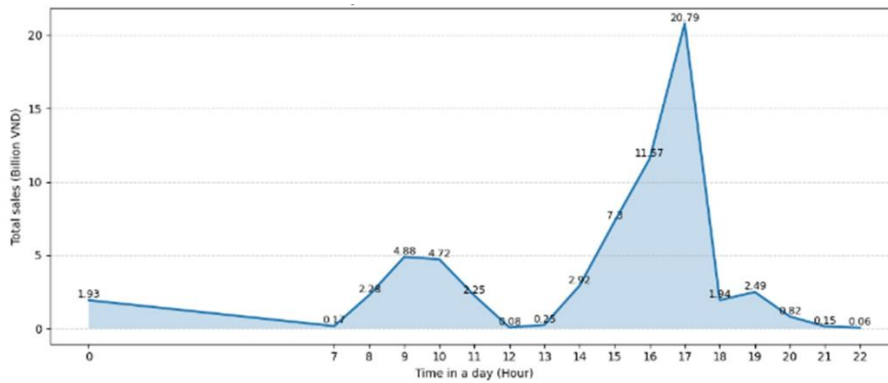


Fig. 5. Bestselling time for Vietnamese wholesales buyers

This difference likely stems from their respective social positions. Retail buyers are often employees or staff, who might use company computers for personal shopping during work hours. In contrast, wholesale consumers are typically high-ranking managers or business owners. They adhere to structured schedules and meticulously gather information and calculate needs before placing orders. Thus, they might make purchases in the early morning when they have free time, or during the last hour of their workday after receiving requests from their own customers throughout the day.

Finally, we will examine revenue across weekdays to identify the bestselling day of the week. Figure 6 provides a detailed view of how revenue changes on weekdays. Generally, total income from Monday to Saturday remains consistent, with the exception of Sunday. In Vietnam, end-consumers often visit shops and supermarkets on Sundays. This means shop owners are busy selling their existing stock and typically don't place new orders with the factory on Sundays.

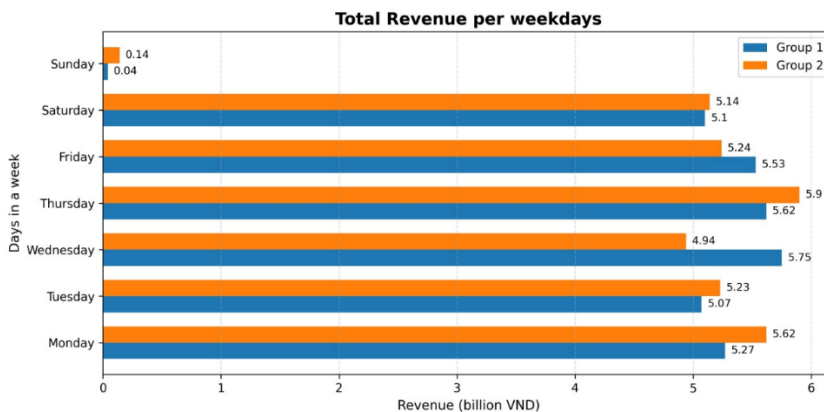


Fig. 6. Revenue in days of a week

The bar chart clearly shows a difference in peak activity between the two customer groups. For Group 1, Wednesday is their busiest ordering day, which also translates into our highest revenue day from them. In contrast, Mondays and Thursdays are most important for Group 2. This difference likely stems from the order quantities involved. Group 2 customers typically order smaller batches, often 100 units or less. This allows them to sell out quickly and then re-order by Thursday to ensure stock for Friday and Saturday sales. In Vietnam, the weekend is the prime shopping period, not Monday. If a business has around 1,000 items to sell over the weekend, there are often leftover items that need to be sold on Monday and Tuesday.

Additionally, orders placed by Group 2 on Wednesday would likely arrive by Friday, ready for weekend sales and beyond.

3.5. Correlation between features to revenue

Based on our previous analysis, we've identified purchasing trends, top-selling products, seasonal variations, and evolving customer needs. Now, we will delve deeper to intensely analyze the relationship between revenue and other features. After pinpointing the most strongly correlated factors, we will examine them closely to capture even the slightest differences between our two customer groups. To achieve this, this study will utilize the Pearson correlation coefficient for the entire population.

Given a pair of two random variables (X, Y), the formula for ρ is:

$$\rho_{X,Y} = \frac{cov(X,Y)}{\delta_X \delta_Y} \tag{1}$$

Where:

- $cov(X, Y)$ is the covariance between X and Y
- δ_X, δ_Y are the standard deviations of X, Y

We will start by identifying overall correlations which are presented in Table 2. After that, we will focus on strongly related factors and examine them in more detail, as shown in Table 3.

Table 2. Correlation between revenue and other features in our dataset.

	Price	Quantity	Day in week	Holiday	Month	Hour
Revenue	0.313	0.778	-0.009	-0.009	-0.015	-0.048

Overall, our factory's earnings are primarily driven by the total quantity of products sold, rather than their individual prices. Interestingly, our revenue shows a very weak correlation with various time periods, including weekdays, months, years, or even hours. Therefore, we will intensely focus on the relationship between revenue, price, and sales volume within our two distinct customer groups.

Table 3. Correlation between revenue versus price and quantity in two groups.

		Price	Quantity
G1	Revenue	0.202	0.82
G2	Revenue	0.266	0.73

In the key account group (Group 1), the volume of sales has a much greater impact on the factory's revenue than the price. This is likely because these top buyers are large business owners whose profits depend more on the quantity of goods they move rather than unit price.

Conversely, for the remaining buyers (Group 2), while sales quantity is still important, price holds a more significant influence. Since customers in this group are typically small shop owners with order volumes usually under 100 units, the price per item becomes a crucial factor for their purchasing decisions.

3.6. Monthly revenue prediction

This study aimed to predict the factory's profit for August 2019. To select an appropriate learning algorithm, we considered the data's characteristics. Given that we had 23 months of data and monthly incomes showed stable fluctuations, we chose Gaussian Regression with an RBF kernel as our training model. Figure 7 illustrates the results.

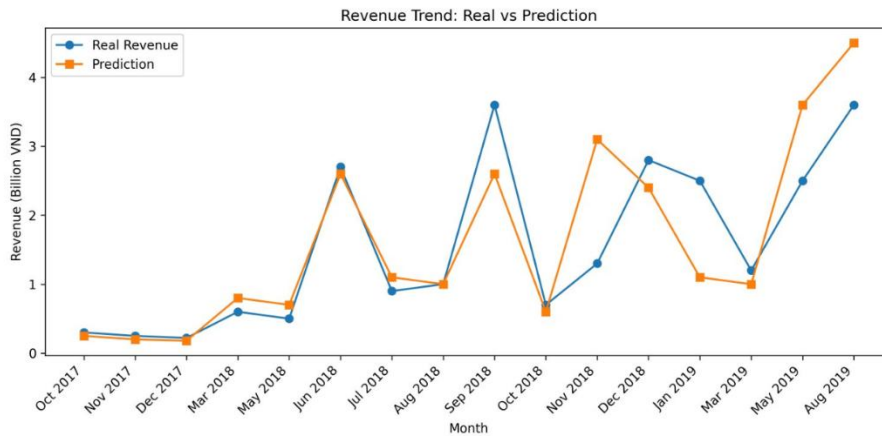


Fig. 7. Monthly revenue prediction using Gaussian Regression model

After 500 training iterations to estimate the error, our profit prediction for August 2019 was remarkably close to the actual figures, with only a 2.3% difference. Specifically, the predicted value was 3.17 billion VND, while the true income was 3.24 billion VND. Gaussian Regression was selected because it offers reliable predictions even with less data, unlike models like Long Short-Term Memory (LSTM), which typically require thousands of samples for strong predictions.

In the future, we plan to expand this model to make more detailed predictions, including will-sold quantities of specific items, high-probability best-selling products, and weekly income, which tends to fluctuate more than monthly profit.

3.7. Suggesting solution

Based on the data we've analyzed, here are some suggestions for the factory owner:

First, it's crucial to understand why we have so few fully outsourced customers (Group 1) and why our revenue, customer count, and product quantity are so heavily dependent on them. We also need to develop strategies to retain these valuable clients, as the cost of acquiring new customers significantly outweighs the cost of keeping existing loyal ones.

Regarding our labor force, we should adjust working hours. In the mornings, more staff should be allocated to processing tasks from the previous day. Additionally, we should have more personnel available to take orders from customers during the final hour of the afternoon shift.

Wednesdays and Thursdays are the two most important days of the week for our business. Specifically, on Wednesdays, we should ensure more order receivers with deep knowledge of fabric types and the entire garment manufacturing process are available. On Thursdays, more staff proficient in common garment processing should be on standby to handle requests from Group 2 customers.

Finally, we could consider a slight price increase for our key account group. Since these large buyers distribute products to end-users at double or even triple our price, a small adjustment on our end is unlikely to deter them. Conversely, prices have a strong impact on our smaller business owners (Group 2), who outnumber Group 1 customers by a significant 19:1 ratio. Maintaining a stable, competitive price for this group is more likely to foster loyalty and attract new clients.

From a national perspective, as Vietnam continues to be a major garment exporter, we should consider developing our own products and focusing on direct exports to maximize income and proactively reach end-users.

4. CONCLUSIONS

This paper unveiled intriguing insights into the behavior of two distinct groups of wholesale buyers in Vietnam. We drew comparisons not only between these two wholesale segments but also with retail consumer behaviors.

Our key findings highlight significant differences in shopping times between wholesale and retail buyers, identify the bestselling items for each group, and pinpoint the most active days of the week for each. Crucially, we also analyzed the correlation between revenue and both price and quantity.

We found that wholesale customers tend to be more predictable, largely adhering to their work schedules and being influenced by major national holidays and events, such as the Lunar New Year and the start of the school year.

Ultimately, we provided practical suggestions and made necessary predictions using AI technology to help adapt to the needs of different customer groups. These strategies aim to enhance business efficiency, streamline company operations, and maximize profit margins in the future.

REFERENCES

- [1] G. V. Golderzahi and H.-K. Pao, "Understanding Customers and Their Grouping via WiFi Sensing for Business Revenue Forecasting," in *Proc. MLDM 2018, Part II*, New York, NY, USA, Jul. 2018, pp. 56–71. doi: https://doi.org/10.1007/978-3-319-96133-0_5.
- [2] D. Chen, S. L. Sain, and K. Guo, "Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining," *Journal of Database Marketing & Customer Strategy Management*, vol. 19, pp. 197–208, Aug. 2012. doi: <https://doi.org/10.1057/dbm.2012.17>.
- [3] R. Chakraborty, J. Lee, S. Bagchi-Sen, S. Upadhyaya, and H. R. Rao, "Online shopping intention in the context of data breach in online retail stores: An examination of older and younger adults," *Decision Support Systems*, vol. 83, pp. 47–56, Mar. 2016. doi: <https://doi.org/10.1016/j.dss.2015.12.007>.
- [4] R. A. El-Deen Ahmeda, M. E. Shehaba, S. Morsya, and N. Mekawiea, "Performance study of classification algorithms for consumer online shopping attitudes and behavior using data mining," in *Proc. 5th Communication Systems and Network Technologies (CSNT)*, Gwalior, Oct. 2015, pp. 1344–1349. doi: <https://doi.org/10.1109/CSNT.2015.50>.
- [5] B. M. Ramageri, "Role of data mining in retail sector," *International Journal on Computer Science and Engineering*, vol. 19, Aug. 2013. [Online]. Available: https://www.researchgate.net/publication/354653024_ROLE_OF_DATA_MINING_IN_RETAIL_SECTOR
- [6] P. S. Raju, V. R. Bai, and G. K. Chaitanya, "Data mining: Techniques for Enhancing Customer Relationship Management in Banking and Retail Industries," *Journal of Database Marketing & Customer Strategy Management*, vol. 2, no. 1, Jan. 2014. [Online]. Available: <https://www.rroj.com/open-access/data-mining-techniques-for-enhancing-customerrelationship-management-in-banking-and-retailindustries.pdf>