3 Branches, 3 Months, 1 Goal: Decoding Supermarket Sales Patterns

MARK ANTHONY A. BUNA

Data Analyst

CONTEXT

• Supermarkets are rapidly expanding in densely populated cities, leading to increased market competition. This dataset captures historical sales from a supermarket company, recorded across three branches over a three-month period. While limited in time range, the dataset is structured in a way that makes it suitable for applying predictive and exploratory data analytics techniques.

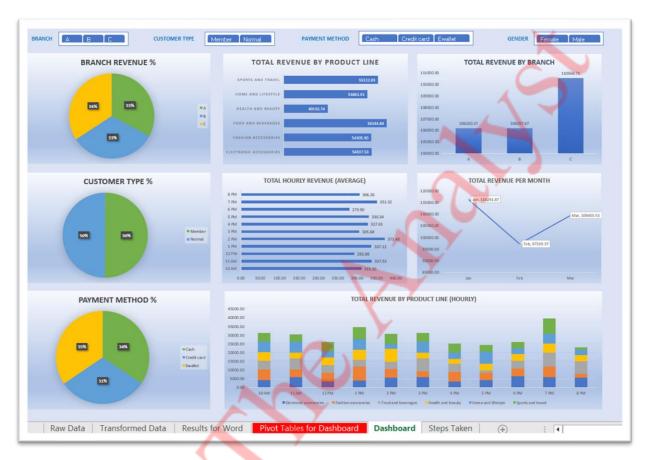
BUSINESS PROBLEM

• With supermarkets rapidly expanding in dense urban areas, competition is intensifying. This project aims to analyze short-term sales data from three branches to uncover performance patterns and product trends. Despite the limited timeframe, the goal is to support strategic decisions in inventory, marketing, and operations to enhance competitiveness and drive growth.

O LINK

• Kaggle: https://www.kaggle.com/datasets/aungpyaeap/supermarket-sales/data

DASHBOARD



(right click + open link)

Q DATA TO BE IDENTIFIED

1. SALES & REVENUE INSIGHTS

- 1.1 What is the total revenue, and how does it vary across different branches?
- 1.2 Which product line generates the highest revenue?
- 1.3 Which product line sells the most in terms of quantity?
- 1.4 Which branch is the most profitable, and why?
- 1.5 What are the peak sales hours and days of the week?

2. CUSTOMER BEHAVIOR ANALYSIS

- 2.1 Which gender spends more on average, and in which product line?
- 2.2 Which customer type (Member vs. Normal) contributes more to total revenue?
- 2.3 Is there a difference in purchasing behavior across different cities?

2.4 How does customer spending vary by payment method?

3. PRODUCT & INVENTORY ANALYSIS

- 3.1 Which product lines have the highest and lowest sales?
- 3.2 *Is there a correlation between unit price and quantity sold?*
- 3.3 Are some product lines consistently underperforming?
- 3.4 Which product lines have the highest gross margin?

4. PRICING & DISCOUNT ANALYSIS

- 4.1 How does the tax impact total sales and profitability?
- 4.2 Do higher-priced products sell less compared to lower-priced ones?
- 4.3 Which branch offers the highest average unit price?

5. CUSTOMER SATISFACTION (RATINGS ANALYSIS)

- 5.1 Which branch has the highest and lowest customer ratings?
- 5.2 Which product line has the best and worst ratings?
- 5.3 Is there a relationship between total sales and customer ratings?
- 5.4 Do high-spending customers leave better ratings?

6. PAYMENT & TRANSACTION ANALYSIS

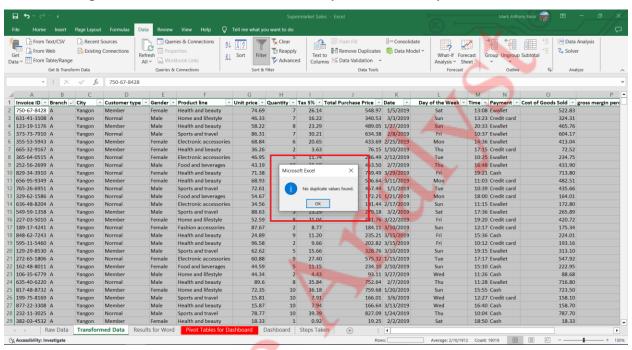
- 6.1 Which payment method is most commonly used?
- 6.2 Are there differences in spending based on the payment method? Do customers using a specific payment method buy more expensive products?

7. TREND & TIME-BASED ANALYSIS

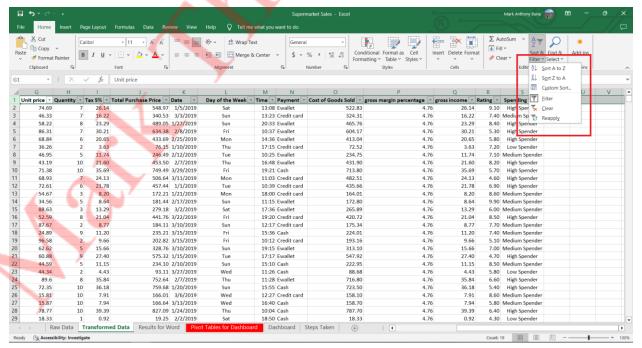
- 7.1 Are there seasonal trends in supermarket sales by branch?
- 7.2 *Is revenue increasing or decreasing over time?*

OBJECT OF THE PARATION

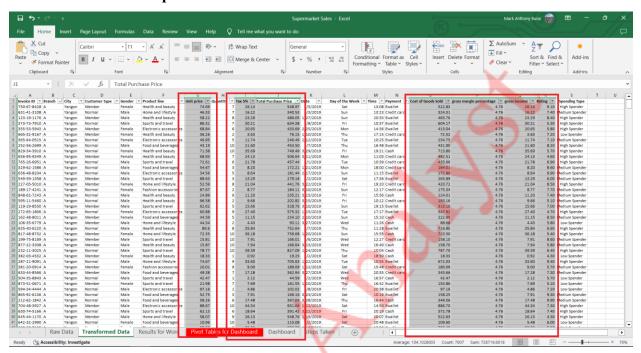
1. **Removed duplicate entries** to ensure data accuracy and consistency.



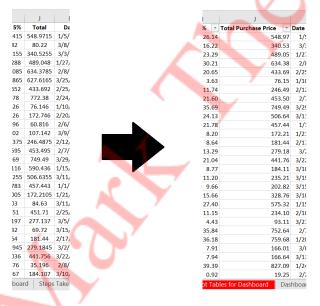
2. **Applied filters** to easily identify any typos or missing values.



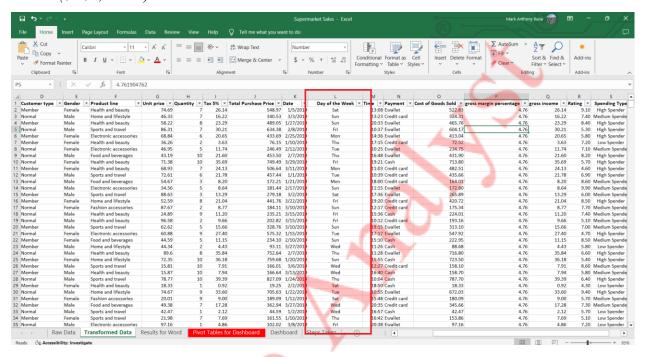
3. Standardized decimal places to two for a more uniform and readable format.



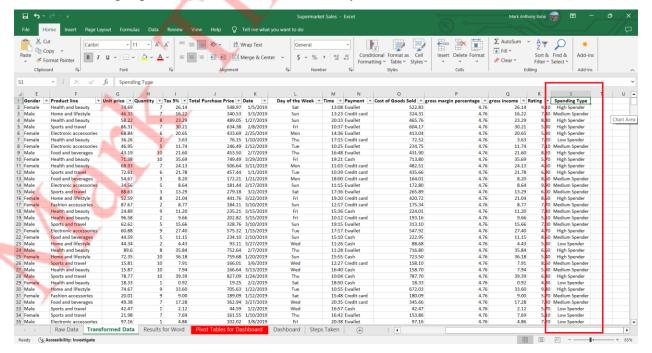
4. Renamed the "Total" column to "Total Purchase Price" for better clarity.



5. Added a new column to extract the day of the week from each date using the formula =TEXT(\$K\$2, "ddd").



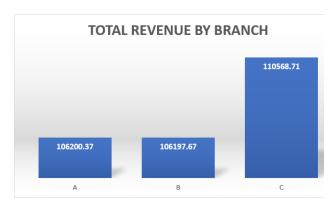
6. Created a "Spending Type" column by segmenting data into three percentiles (low, medium, and high spenders) to facilitate further analysis.



(2) INTERPRETATION AND REPORTING

1. SALES & REVENUE INSIGHTS

1.1 What is the total revenue, and how does it vary across different branches?



Branch Sum of Total Purchase Price (\$)

A 106200.37

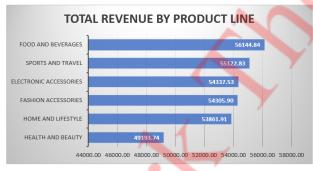
B 106197.67

C 110568.71

Grand Total 322966.75

The total revenue of the three branches is \$322,966.75 and as the details show, there's a little difference between branch A and B while the branch C garnered the highest number of sales.

1.2 Which product line generates the highest revenue?



Product Line Sum of Total Purchase Price (\$)

Health and beauty 49193.74

Home and lifestyle 53861.91

Fashion accessories 54305.90

Electronic accessories 54337.53

Sports and travel 55122.83

Food and beverages 56144.84

Grand Total 322966.75

The product line that generates the highest revenue is <u>Food and Beverages</u> having \$56,144.83 in total revenue.

1.3 Which product line sells the most in terms of quantity?



Product Line 🚭	Sum of Quantity
Electronic accessories	971
Food and beverages	952
Sports and travel	920
Home and lifestyle	911
Fashion accessories	902
Health and beauty	854
Grand Total	5510

The product line that sells the most in terms of quantity is **Electronics Accessories** which sold a quantity of **971** in total across the 3 branches.

1.4 Which branch is the most profitable?



Branch Sum of 1	Total Purchase Price (\$)	Sum of Cost of Goods Sold (\$)	Sum of gross in
Α	106200.37	101143.21	5057.16
В	106197.67	101140.64	5057.03
С	110568.71	105303.53	5265.18
Grand Total	322966.75	307587.38	15379.3

The most profitable branch is **Branch C**, which has a total gross income of \$5,265.18. It is also worth noting that Branch C has the highest total purchase amount among the three branches, totaling \$110,568.71.

1.5 What are the peak sales hours and days of the week?

		PEA	AK SALES E	BY HOUR A	ND DAY				
Sum	of Total Purchase Price	Day of the Week 🔻							
	Purchase Time	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Grand Total (\$)
+	10 AM	4067.34	3738.42	4585.48	3177.33	6885.23	5671.69	3295.99	<i>*</i> 31421.48
+	11 AM	4736.77	2872.63	5223.88	5004.96	4402.89	2649.29	5486.91	30377.33
+	12 PM	4728.20	4726.10	3739.07	3238.33	2564.09	2159.24	4910.85	26065.88
+	1 PM	5159.55	3758.63	3962.53	6734.34	2594.22	6824.31	5689.64	34723.23
+	2 PM	4676.12	2004.35	4953.17	3212.57	5474.05	5166.27	5341.88	30828.40
+	3 PM	2681.86	5141.35	7020.21	4855.94	4313.09	3696.06	3471.02	31179.51
+	4 PM	2241.23	4690.99	3831.44	3229.62	4327.19	3092.34	3813.52	25226.32
+	5 PM	4169.34	3179.26	2975.53	3101.27	4056.59	1893.41	5069.81	24445.22
+	6 PM	3637.27	2806.09	2417.58	4233.77	3021.57	3175.29	6738.75	26030.34
+	7 PM	5963.64	2598.62	9198.67	4160.17	3185.09	5475.93	9117.38	39699.51
+	8 PM	2396.56	2382.63	3574.69	2782.84	4525.24	4122.50	3185.07	22969.53
	Grand Total (\$)	44457.89	37899.08	51482.25	43731.14	45349.25	43926.34	56120.81	322966.75

As per data, peak sales by hour happen around $\frac{7pm (7:00 - 7:49pm)}{7:49pm}$ and while by day, it occurs during <u>Saturday</u>.

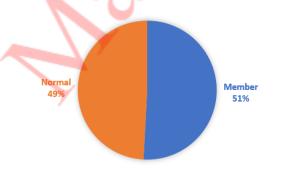
2. CUSTOMER BEHAVIOR ANALYSIS

2.1 Which gender spends more on average, and in which product line?

	GENDER SPENDING TRENDS BY PRODUCT LINE						
Sum of Total Purchase Price	Sum of Total Purchase Price Column Labels 🔻 💮						
Row Labels	Electronic accessories	Fashion accessories	Food and beverages	Health and beauty	Home and lifestyle	Sports and travel	Grand Total (\$)
Female	27102.02	30437.40	33170.92	18560.99	30036.88	28574.72	167882.93
Male	27235.51	23868.50	22973.93	30632.75	23825.04	26548.11	155083.82
Grand Total (\$)	54337.53	54305.90	56144.84	49193.74	53861.91	55122.83	322966.75

Female spends more on average than the male gender and the product line that they spent with the highest is **Food and Beverages** with a total amount **of \$33,170.92**.

2.2 Which customer type (Member vs. Normal) contributes more to total revenue?



REVENUE BY CUSTOMER TYPE

The customer type that contributes more to total revenue is <u>Member</u> type of customers having purchased a total amount of \$164,223.44.

REVENUE BY CUSTOMER TYPE					
Customer Type Total Purchase Price (\$)					
Member	164223.44				
Normal	158743.31				
Grand Total	322966.75				

2.3 Is there a difference in purchasing behavior across different cities?

PURCHA	PURCHASING BEHAVIOR DIFFERENCES ACROSS CITIES						
	City -						
Branch 🔻	Mandalay	Naypyitaw	Yangon	Total Revenue (\$)			
Α			106200.37	106200.37			
В	106197.67			106197.67			
С		110568.71		110568.71			
Total Revenue (\$)	106197.67	110568.71	106200.37	322966.75			

<u>There is not much difference among the three branches</u> in terms of total revenue, however, it can also be noted that branch C which is located in Naypyitaw gathered the most amount in total revenue of \$110,568.71.

2.3 What is the average number of items per transaction for each product line?



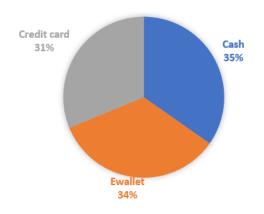
AVG QUANTITY SOLD PER PRODUCT LINE Product Line - Average of Quantity Electronic accessories 5.71 Home and lifestyle 5.69 Health and beauty 5.62 Sports and travel 5.54 5.47 Food and beverages Fashion accessories 5.07 **Grand Total** 5.51

Electronics and accessories had the highest average number of items per transaction, resulting in an average quantity of 5.71, followed by Home and Lifestyle, Health and Beauty, Sports and Travel, Food and Beverages, and lastly, Fashion Accessories.

Overall, customers purchase approximately 5.51~6 items per transaction on average.

2.4 How does customer spending vary by payment method?

SPENDING VARIATION BY PAYMENT METHOD



Paying by <u>Cash</u> comes first with a total amount of \$112,206.57, followed by <u>Ewallet</u> worthing \$109,993.11 and lastly paying through <u>credit card</u> with \$100,767.07.

CUSTOMER SPENDING VARIATION BY PAYMENT METHOD			
Payment Method	Total Purchase Price (\$)		
Cash	112206.57		
Ewallet	109993.11		
Credit card	100767.07		
Grand Total	322966.75		

3. PRODUCT & INVENTORY ANALYSIS

3.1 Which product lines have the highest and lowest sales?



The product line that has the highest sales is **Food and Beverages** with \$56,144.84 while the lowest is **Health and Beauty** with \$49,193.74.

HIGHEST AND LOWEST SALES BY PRODUCT LINE					
Product Line 🚽	Total Purchase Price (\$)				
Food and beverages	56144.84	highest			
Sports and travel 55122.83					
Electronic accessories	54337.53				
Fashion accessories	54305.90				
Home and lifestyle 53861.91					
Health and beauty 49193.74 lowest					
Grand Total	322966.75				

3.2 Is there a correlation between unit price and quantity sold?

CORRELATION 0.01078



A correlation of <u>0.0108</u> indicates an **extremely weak positive correlation** between the two variables, in this case, **unit price** and **quantity sold**.

This means:

- There is almost no relationship between unit price and quantity sold, as the correlation is very close to 0.
- The tiny positive value suggests that, if anything, as unit price increases, quantity sold might increase very slightly, but this relationship is almost negligible.

In practical terms, this result suggests that unit price and quantity sold are virtually unrelated in this dataset.

3.3 Are some product lines consistently underperforming?

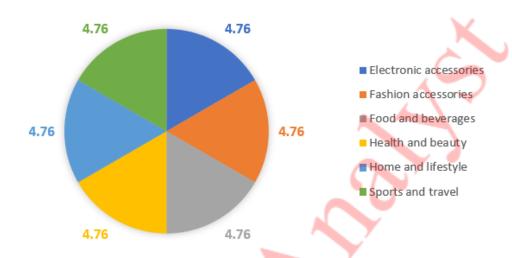


Average of Total Purchase Price	Month 🔻			
Product Line	Jan	Feb	Mar	Total Average (\$)
Electronic accessories	348.73	321.54	292.63	319.63 declining
Fashion accessories	302.27	316.83	295.39	305.09
Food and beverages	349.47	322.59	295.96	322.67 declining
Health and beauty	334.35	317.44	319.44	323.64
Home and lifestyle	347.37	327.22	332.27	336.64
Sports and travel	309.53	321.15	370.68	332.07
Total Average (\$)	330.37	320.86	317.26	322.97

Based on the 3-month sales result from the dataset, it can be observed that product lines that are declining in sales performance are <u>Electronics Accessories and Food and beverages</u>. Both product lines have been declining over time.

3.4 Which product lines have the highest gross margin?

GROSS MARGIN PER PRODUCT LINE %



GROSS MARGIN BY PRODUCT LINE						
Product Line	Average of gross margin percentage (%)	Average of gross income (\$)				
Electronic accessories	4.76	15.22				
Fashion accessories	4.76	14.53				
Food and beverages	4.76	15.37				
Health and beauty	4.76	15.41				
Home and lifestyle	4.76	16.03				
Sports and travel	4.76	15.81				
Grand Total	4.76	15.38				

It can be noted that all three branches use the same gross margin percentage of 4.76% with <u>Home and Lifestyle</u> having the highest average gross income of \$16.03.

4. PRICING & DISCOUNT ANALYSIS

4.1 How does the tax impact total sales and profitability?

IMP	ACT OF TAX ON	TOTAL SALES AND PROFITABI	LITY
Row Labels	Sum of Tax 5%	Sum of Total Purchase Price	Sum of gross income
Electronic accessories	2587.50	54337.53	2587.50
Fashion accessories	2586.00	54305.90	2586.00
Food and beverages	2673.56	56144.84	2673.56
Health and beauty	2342.56	49193.74	2342.56
Home and lifestyle	2564.85	53861.91	2564.85
Sports and travel	2624.90	55122.83	2624.90
Grand Total	15379.37	322966.75	15379.37

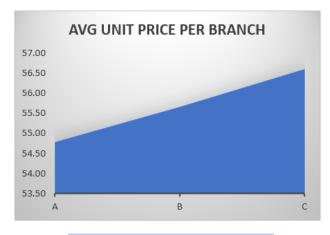
The store's pricing model is designed so that the 5% tax applied to each item also represents the gross income per item. This structure indicates that the store's profitability is entirely reliant on sales volume rather than expanding profit margins, as the tax essentially covers the entire profit margin.

4.2 Do higher-priced products sell less compared to lower-priced ones?



The correlation between product price and quantity sold is <u>0.0108</u>, indicating a <u>negligible linear</u> relationship. This suggests that, based on the data analyzed, <u>price does not significantly impact sales volume.</u>

4.3 Which branch offers the highest average unit price?



The branch that offers the highest average unit price is <u>Branch C</u> which has \$56.61 per item.

AVERAGE UNIT	AVERAGE UNIT PRICE PER BRANCH			
Branch				
Α	54.78			
В	55.66			
С	56.61			
Grand Total	55.67			

5. CUSTOMER SATISFACTION (RATINGS ANALYSIS)

5.1 Which branch has the highest and lowest customer ratings?

lowest

highest



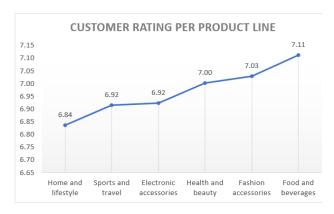
Grand Total

7.03

6.97

The branch the has the **highest rating** is **Branch C**, while the **lowest** is **Branch B**.

5.2 Which product line has the best and worst ratings?



The product line that has the **best rating** is **Food and Beverages** having a rating of 7.11/10 while the **worst** is **Home and Lifestyle** with 6.84/10.

CUSTOMER RATING F		
Product Line 🖵	Average of Rating	
Home and lifestyle	6.84	worst rating
Sports and travel	6.92	
Electronic accessories	6.92	
Health and beauty	7.00	
Fashion accessories	7.03	
Food and beverages	7.11	best rating
Grand Total	6.97	

5.3 Is there a relationship between total sales and customer ratings?

CORRELATION -0.036441705



The correlation between total sales and customer ratings is **-0.036**, indicating a very weak negative relationship. This suggests that there **is little to no meaningful relationship between total sales and customer ratings** in this dataset.

RELA	TIONSHIP BE	TWEEN TOTA	L SALES AN	D CUSTOME	R RATINGS	- REGRESSIC	ON ANALYSI	S
Regression S	Statistics							
Multiple R	0.036441705							
R Square	0.001327998							
Adjusted R Square	0.000327325							
Standard Error	1.718299005							
Observations	1000							
ANOVA								
	df	SS	MS	F	Significance F		4	
Regression	1	3.91834363	3.91834363	1.327104259	0.24959691		/ /	
Residual	998	2946.646366	2.952551469			<u></u>	_	
Total	999	2950.56471						
							1	
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	7.054960946	0.089730264	78.6240965	0	6.878879314	7.231042578	6.878879314	7.231042578
X Variable 1	-0.000254704	0.000221097	-1.152000112	0.24959691	-0.000688573	0.000179165	-0.000688573	0.000179165

The **R-squared** value is **0.0013**, indicating that the independent variable (total sales) explains only 0.13% of the variation in the dependent variable (customer ratings), which suggests a <u>very weak</u> relationship between the two variables.

The **coefficient** for total sales is **-0.00025**, suggesting that for every unit increase in total sales, customer ratings decrease slightly. However, the **p-value** for this coefficient is **0.25**, which is greater than **the 0.05 threshold**, meaning that this relationship is **not statistically significant** and the result could be due to random chance.

In summary, based on these results, there is no meaningful or statistically significant relationship between total sales and customer ratings.

5.4 Do high-spending customers leave better ratings?

Percentile	Values	Spending Type
33rd percentile	< 163.36	Low Spender
66th Percentile	163.37 - 382.80	Medium Spender
100th percentile	> 382.80	High Spender

IMPACT OF HIGH SPENDING ON CUSTOMER RATINGS				
Spending Type	Number of Spender	Rating Average	Average of Total Purchase Price (\$)	
High Spender	340	6.93	615.03	
Medium Spender	330	7.06	256.46	
Low Spender	330	6.93	88.56	
Grand Total	1000	6.97	322.97	
Y		CORRELATION	-0.196856012	

The correlation between total spending and customer ratings is **-0.197**, indicating a weak negative relationship. This suggests that there is a slight inverse trend, where <u>higher spending may be weakly associated with lower ratings</u>, although the relationship is not strong enough to draw <u>definitive conclusions</u>

6. PAYMENT & TRANSACTION ANALYSIS

6.1 Which payment method is most commonly used?

CUSTOMER'S PAYMENT METHOD			
Payment Method -	Number of Customer		
Ewallet	345		
Cash	344		
Credit card	311		
Grand Total	1000		

The most commonly used payment method is the <u>Ewallet</u>, which has 345 customers.

6.2 Are there differences in spending based on the payment method? Do customers using a specific payment method buy more expensive products?



CUSTOMER'S PAYMENT METHOD			
Payment Metho	Number of Customer	Sum of Total Purchase Price	
Ewallet	345	109993.11	
Cash	344	112206.57	
Credit card	311	100767.07	
Grand Total	1000	322966.75	
170	CORRELATION	0.97818108	

The correlation between payment method and total purchase price is 0.978, indicating a very strong positive relationship. This suggests that a customer's total purchase price is closely linked to the payment method they choose. Specifically, customers using Cash are expected to have the highest total purchase price, followed by those using Ewallet, with Credit card users generally making the lowest total purchases.

7. TREND & TIME-BASED ANALYSIS

7.1 Are there seasonal trends in supermarket sales by branch?



Overview

Based on the sales data from January to March, there is a clear seasonal trend across all branches. January recorded strong sales, but February experienced a significant decline, followed by a recovery in March.

Key Findings

- January Sales: Sales were at their highest, indicating strong consumer demand at the beginning of the year.
- February Dip: There was a noticeable drop in sales across all branches, suggesting a potential seasonal slowdown.
- March Recovery: Sales increased again, indicating a return of consumer demand.

Recommendations

To mitigate the sales decline in February, the following strategies can be implemented:

- 1. **Promotional Campaigns:** Introduce discounts, bundled offers, or loyalty rewards to attract more customers.
- 2. **Seasonal Marketing Strategies:** Align promotions with events such as Valentine's Day or other seasonal trends.
- 3. **Product Mix Optimization:** Identify high-performing products from January and March and promote them more in February.
- 4. **Customer Engagement:** Enhance marketing efforts through targeted ads, email promotions, and in-store activations.

By implementing these strategies, we can work towards minimizing the sales dip in February and ensuring a more stable revenue flow throughout the first quarter.

7.2 Is revenue increasing or decreasing over time?

Revenue Trend Analysis (January - March)

From the recorded sales data, January had the highest revenue, but there was a notable decline in February across all branches. However, in March, sales rebounded, indicating a recovery.

- January → February: Revenue decreased significantly.
- February → March: Revenue increased again.

Since the data covers only three months, it is difficult to determine a long-term trend. However, the short-term pattern suggests fluctuations rather than a consistent increase or decrease. To establish a clearer trend, it would be beneficial to analyze sales over a longer period, ideally a year or more.

SUMMARY OF FINDINGS

1. SALES & REVENUE INSIGHTS

• The total revenue across the three branches is \$322,966.75, with Branch C generating the highest sales at \$110,568.71. There isn't much variation in total revenue among the branches. The Food and Beverages product line leads in revenue, bringing in \$56,144.83, while Electronics Accessories is the top-selling product by quantity, with 971 units sold across all branches. Peak sales occur around 7:00–7:49 PM, and Saturdays see the highest sales by day.

2. CUSTOMER BEHAVIOR ANALYSIS

• Females tend to spend more on average than males, with their highest expenditure on the Food and Beverages product line, totaling \$33,170.92. Member-type customers contribute significantly to total revenue, having spent \$164,223.44 in total. On average, customers purchase around 5.51~6 items per transaction. Cash payments lead with \$112,206.57, followed by Ewallets at \$109,993.11, and credit card payments at \$100,767.07.

3. PRODUCT & INVENTORY ANALYSIS

• Food and Beverages leads sales at \$56,144.84, while Health and Beauty trails at \$49,193.74. The correlation of 0.0108 shows a very weak link between unit price and quantity sold. Over the last three months, Electronics Accessories and Food and Beverages have declined in sales. All three branches share a 4.76% gross margin, with Home and Lifestyle achieving the highest average gross income of \$16.03.

4. PRICING & DISCOUNT ANALYSIS

• The store's profitability depends mainly on sales volume, as taxes cover most of the profit margin. The correlation between product price and quantity sold is 0.0108, indicating a negligible relationship, suggesting that price doesn't significantly impact sales volume. Branch C offers the highest average unit price at \$56.61 per item.

5. CUSTOMER SATISFACTION (RATINGS ANALYSIS)

• Branch C leads in ratings, while Branch B ranks lowest. Food and Beverages scores the highest with 7.11/10, and Home and Lifestyle the lowest at 6.84/10. The weak negative correlations between total sales (-0.036) and spending (-0.197) with customer ratings suggest no strong connection, though a slight inverse trend is observed with higher spending.

6. PAYMENT & TRANSACTION ANALYSIS

• The Ewallet is the most commonly used payment method, with 345 customers. The strong correlation of 0.978 between payment method and total purchase price shows that payment method is closely linked to purchase amounts. Cash users typically have the highest total purchases, followed by Ewallet users, while credit card users make the lowest purchases.

7. TREND & TIME-BASED ANALYSIS

• The sales data from January to March shows a seasonal trend, with strong sales in January, a decline in February, and a recovery in March. The short-term pattern indicates fluctuations, making it difficult to identify a long-term trend. Analyzing sales over a longer period would provide a clearer picture.

RECOMMENDATION

To sustain revenue and improve profitability, the business should:

- 1. Monitor product trends more closely, particularly the decline in Electronics Accessories and Food and Beverages, and adjust marketing or bundling strategies accordingly.
- 2. Capitalize on peak hours (7:00–7:49 PM) and weekends, optimizing staffing, inventory, and promotions during these periods.
- 3. Segment marketing efforts by gender and membership status, since female customers and members contribute more significantly to revenue.
- 4. Investigate product performance beyond price, as pricing appears to have minimal impact on sales volume; instead, focus on product placement, bundling, or loyalty perks.
- 5. Improve satisfaction for underperforming categories like Home and Lifestyle by gathering direct feedback and enhancing the product experience.
- 6. Extend sales tracking beyond the current 3-month window to detect long-term trends and forecast seasonal performance with greater accuracy.

REFLECTION & NEXT STEPS

• This project improved my ability to analyze retail data in Excel and draw insights from customer behavior, sales patterns, and product trends. Next, I plan to extend the dataset and explore advanced Excel tools like pivot charts and forecasting for deeper analysis.