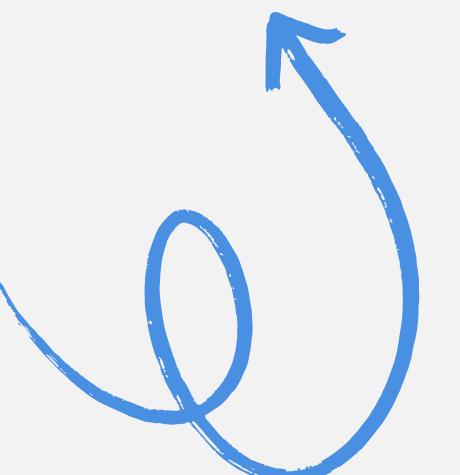


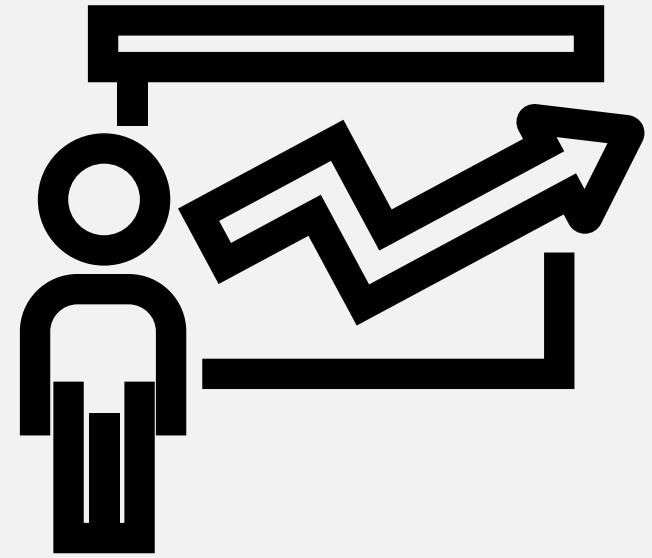
# Sales Performance Analysis

Vedang Mishra



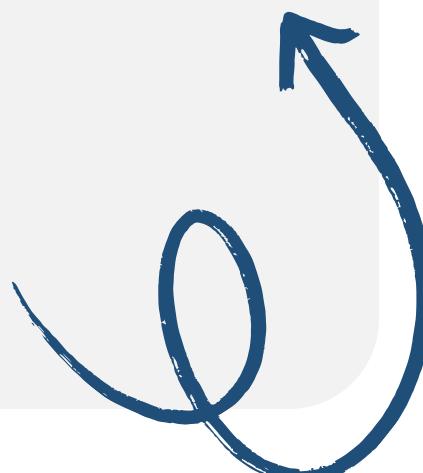
# Project Overview

- Analyzed 1000+ Walmart transactions
- Used SQL CTEs, Window Functions, Ranking, Aggregations
- Objective: Extract business insights across sales, customers & products



## PROJECT OBJECTIVES

- Identify key sales trends and profitability patterns across branches.
- Segment customers and detect anomalies for better retail decision-making.
- Provide actionable insights to improve operations, marketing, and customer targeting.

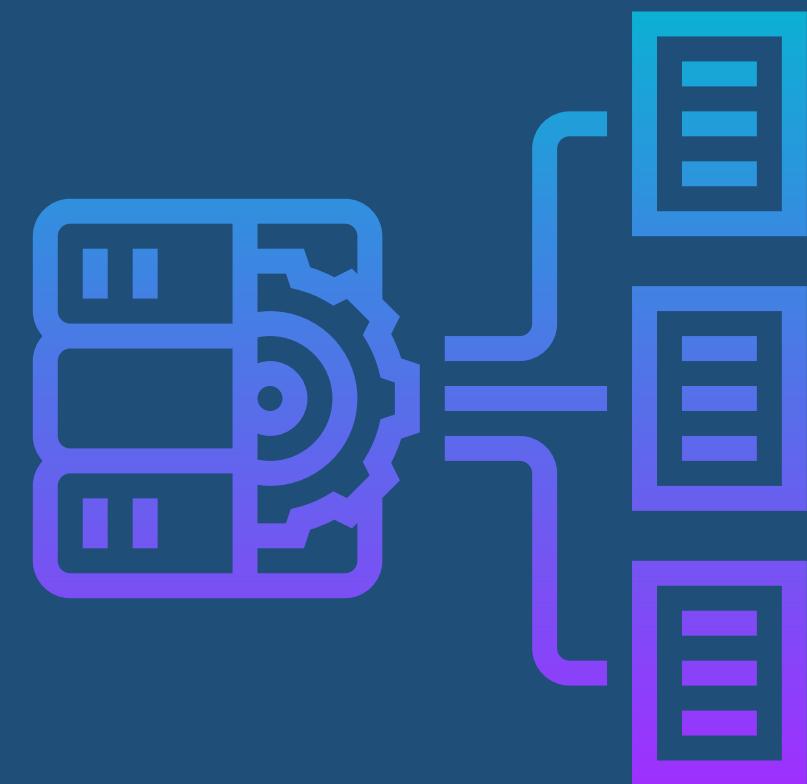


# Dataset Overview and Attributes

- Fields: Branch, Product Line, Total, Gross Income, Gender, Payment, Customer Type
- Contains Dates + Product + Revenue + Customer Behavior
- 3 Branches: A, B, C
- 17 attributes



Invoice ID	Date	Branch	City	Customer type	Product line	Total
450-42-3339	2019-01-01	A	Mandalay	Member	Health and beaut	888.405
699-88-1972	2019-01-01	A	Naypyitaw	Normal	Sports and travel	925.321
607-65-2441	2019-01-01	B	Yangon	Normal	Health and beaut	960.617
628-90-8624	2019-01-01	B	Yangon	Normal	Food and bevera	950.205
743-04-1105	2019-02-01	C	Mantoy	Member	Health and beaut	753.567
280-17-4359	2019-02-01	A	Mandalay	Food and travel	Food and bevera	952.157
766-85-7061	2019-02-01	B	Naypyitaw	Food and bevera-	Electronic access	905.205
267-62-7380	2019-02-01	B	Yangon	Electronic acces-	Electronic access	900.573
431-66-2305	2019-02-01	C	Yangon	Electronic access-	Electronic access	805.331
533-33-5337	2019-02-01	B	Naypyitaw	Food and tre	Electronic access	950.208
227-07-4446	2019-02-01	A	Yangon	Electronic acce-	Electronic access	860.651
817-69-8206	2019-03-01	C	Nayptaw	Electronic ac-	Electronic access	789.392
861-77-0145	2019-03-01	B	Naypyitaw	Electronic access-	Electronic access	850.665



# Task 1: Top Branch Analysis

```
-- Task 1: Identifying the Top Branch by Sales Growth Rate
with monthly_sales as (
    select Branch, DATE_FORMAT(STR_TO_DATE(Date, '%d-%m-%Y'), '%Y-%m') as month, SUM(Total) as total_sales
    from walmartsales
    group by Branch, month
),
growth_calc as (
    select Branch, month, total_sales - lag(total_sales) over (partition by Branch order by month)
    as growth_amount
    from monthly_sales
)
select Branch, month, growth_amount
from growth_calc
where growth_amount IS NOT NULL
order by growth_amount DESC
limit 1;
```

Result Grid			Filter Rows:	Export:	Wrap Cell Content:
	Branch	month	growth_amount		
▶ A	2019-03	7799.001000000004			



# TASK 2: Most Profitable Product Line per Branch

```
-- Task 2: Finding the Most Profitable Product Line for Each Branch
WITH profit_calc AS (
    SELECT Branch, `Product line`, round(SUM(`gross income`),2) as total_profit
    FROM walmartsales
    GROUP BY Branch, `Product line`
),
ranked AS (
    select *, rank() over (partition by Branch order by total_profit DESC) as rnk
    from profit_calc
)
select Branch, `Product line`, total_profit
from ranked
where rnk = 1;
```

Result Grid | Filter Rows:  | Export: | Wrap Cell Content

	Branch	Product line	total_profit
▶	A	Home and lifestyle	1067.49
	B	Sports and travel	951.82
	C	Food and beverages	1131.75



# Task 3: Customer Segmentation Insights



- Effective customer segmentation allows for tailored marketing strategies, enhancing engagement and sales.
- By analyzing purchasing behavior and demographics, we can identify distinct customer groups for targeted outreach.

```
-- Task 3: Analyzing Customer Segmentation Based on Spending
with customer_spending as (
    select `Customer ID`, round(SUM(Total),3) as total_spent
    from walmartsales
    group by `Customer ID`
),
percentile_split as (
    select `Customer ID`, total_spent, NTILE(4) over (ORDER BY total_spent) AS quartile
    from customer_spending
)
select `Customer ID`, total_spent,
case
when quartile = 4 then 'High'
when quartile = 3 then 'Medium'
else 'Low'
end as spending_category
-- ... --
```

	Customer ID	total_spent	spending_category
▶	1	22634.545	Medium
	2	23392.278	High
	3	23402.263	High
	4	17656.716	Low
	5	19632.039	Low
	6	20693.956	Low
	7	20628.09	Low
	8	26634.342	High
	9	19661.597	Low
	10	20723.934	Low
	11	21398.822	Medium
	12	21720.647	Medium
	13	21063.662	Low
	14	21049.403	Low

# Task 4: Sales Anomaly Detection Insights

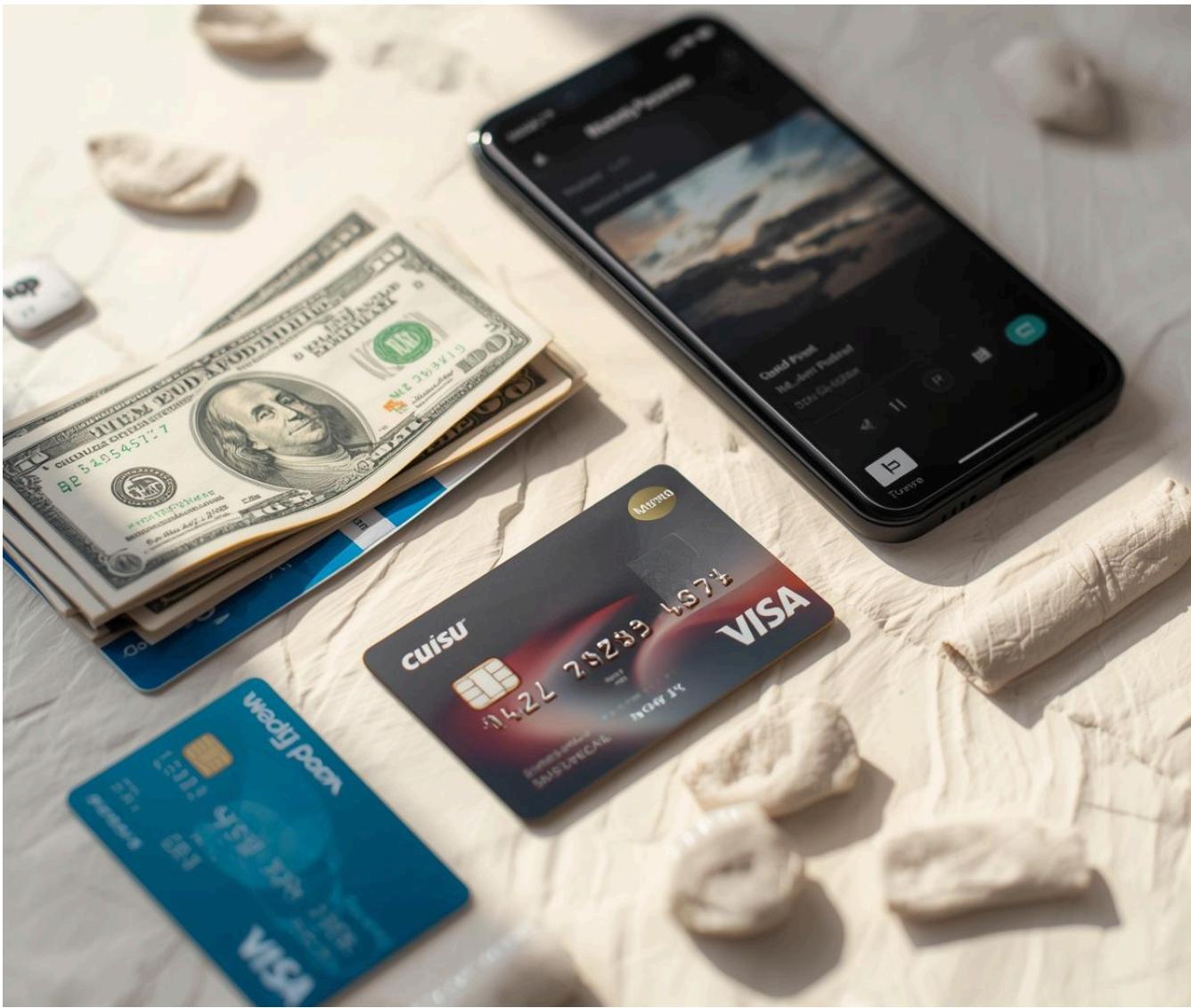
```
-- task 4 Detecting Anomalies in Sales Transactions
WITH stats AS (
    SELECT `Product line`, round(AVG(Total),2) AS avg_sale, STDDEV(Total) AS std_sale
    FROM walmartsales
    GROUP BY `Product line`
)
select w.`Invoice ID`, w.`Product line`, w.Total, s.avg_sale, s.std_sale,
case
when w.Total > s.avg_sale + 2 * s.std_sale then 'High Anomaly'
when w.Total < s.avg_sale - 2 * s.std_sale then 'Low Anomaly'
end as anomaly_flag
FROM walmartsales w
JOIN stats s USING (`Product line`) WHERE w.Total > s.avg_sale + 2*s.std_sale OR w.Total < s.avg_sale - 2*s.std_sale
```

	Invoice ID	Product line	Total	avg_sale	std_sale	anomaly_flag
▶	450-42-3339	Health and beauty	888.405	323.64	236.7561479276635	High Anomaly
	699-88-1972	Health and beauty	832.944	323.64	236.7561479276635	High Anomaly
	607-65-2441	Health and beauty	860.475	323.64	236.7561479276635	High Anomaly
	628-90-8624	Health and beauty	867.09	323.64	236.7561479276635	High Anomaly
	743-04-1105	Health and beauty	918.729	323.64	236.7561479276635	High Anomaly
	280-17-4359	Health and beauty	950.25	323.64	236.7561479276635	High Anomaly
	766-85-7061	Health and beauty	922.635	323.64	236.7561479276635	High Anomaly
	267-62-7380	Electronic accessories	864.57	319.63	245.22119001027605	High Anomaly
	431-66-2305	Electronic accessories	833.9625	319.63	245.22119001027605	High Anomaly
	457-94-0464	Electronic accessories	830.3715	319.63	245.22119001027605	High Anomaly
	533-33-5337	Electronic accessories	833.595	319.63	245.22119001027605	High Anomaly
	227-07-4446	Electronic accessories	820.365	319.63	245.22119001027605	High Anomaly
	817-69-8206	Electronic accessories	942.4485	319.63	245.22119001027605	High Anomaly
	638-60-7125	Electronic accessories	836.304	319.63	245.22119001027605	High Anomaly
	861-77-0145	Electronic accessories	860.685	319.63	245.22119001027605	High Anomaly
	704-48-3927	Electronic accessories	931.035	319.63	245.22119001027605	High Anomaly
	303-96-2227	Home and lifestyle	1022.49	336.64	253.80379730921942	High Anomaly
	744-16-7898	Home and lifestyle	1022.385	336.64	253.80379730921942	High Anomaly



I detected unusually high or low sales transactions using mean and standard deviation.  
If a sale exceeded the mean by 2 standard deviations, it was labeled a High Anomaly.

# Task 5:Most Popular Payment Methods by City



This analysis seeks to uncover the most frequently utilized payment methods across various cities, highlighting trends in consumer preferences and adapting to demographic behaviors seen in shopping patterns.

```
-- Task 5: Most Popular Payment Method by City
with payment_count as (
    select City, Payment, COUNT(*) as txn_count
    from walmartsales
    group by City, Payment
),
ranked as (
    select * , RANK() OVER (partition by City order by txn_count DESC) as rnk
    FROM payment_count
)
select City, Payment as most_used_payment_method, txn_count
from ranked where rnk = 1;
```

Result Grid			
	City	most_used_payment_method	txn_count
▶	Mandalay	Ewallet	113
	Naypyitaw	Cash	124
	Yangon	Ewallet	126

# Task 6: Monthly Sales Distribution by Gender

-- Task 6 Monthly Sales Distribution by Gender

```
select DATE_FORMAT(STR_TO_DATE(Date, '%d-%m-%Y'), '%Y-%m') as month, Gender as total_sales from walmarthsales group by month, Gender order by month, Gender;
```

Result Grid | Filter Rows: \_\_\_\_\_

	month	Gender	total_sales
▶	2019-01	Female	59138.98
	2019-01	Male	57152.89
	2019-02	Female	56335.56
	2019-02	Male	40883.82
	2019-03	Female	52408.39
	2019-03	Male	57047.12

This analysis focuses on the sales distribution by gender for each month, revealing critical insights into purchasing behavior and enabling targeted marketing strategies for both male and female customers.



# Task 7: Best Product Line by Customer Type

```
-- Task 7 Best Product Line by Customer Type
with sales_data as (
    select `Customer type`, `Product line`, round(SUM(Total),3) as total_sales
    from walmartsales
    group by `Customer type`, `Product line`
),
ranked as (
    select *, rank() over (partition by `Customer type` order by total_sales desc) as rnk
    from sales_data
)
select `Customer type`, `Product line`, total_sales
from ranked
where rnk = 1;
```

Result Grid | Filter Rows: \_\_\_\_\_ | Export: \_\_\_\_\_ | Wrap Cell Content: \_\_\_\_\_

	Customer type	Product line	total_sales
▶	Member	Food and beverages	31357.62
	Normal	Electronic accessories	29839.037

Result 21 ×

Output: \_\_\_\_\_

\_\_\_\_\_



This analysis identifies which product lines are favored by various customer types, providing insights that can enhance targeted marketing strategies and improve overall sales performance across different demographics.

# Task8: Repeat Customers Within 30 Days

```
-- Task 8 Identifying Repeat Customers
WITH purchases AS (
    SELECT `Customer ID`, STR_TO_DATE(Date, '%d-%m-%Y') AS purchase_date, LAG(STR_TO_DATE(DATE,
        PARTITION BY `Customer ID` ORDER BY STR_TO_DATE(Date, '%d-%m-%Y')) ) AS prev_purchase
    FROM walmartsales
),
repeat_events AS (
    SELECT `Customer ID`
    FROM purchases
    WHERE prev_purchase IS NOT NULL AND DATEDIFF(purchase_date, prev_purchase) <= 30
)
SELECT `Customer ID`, COUNT(*) AS repeat_count_within_30_days
FROM repeat_events
GROUP BY `Customer ID`
ORDER BY repeat_count_within_30_days DESC;
```

Result Grid		
	Customer ID	repeat_count_within_30_days
▶	1	66
	2	66
	3	66
	4	66
	5	66
	6	66
	7	66
	8	66
	9	66
	10	66
	11	65
	12	65
	13	65
	14	65
	15	65



This analysis identifies the number of customers who made repeat purchases within a 30-day window, providing insights into customer loyalty and shopping patterns at Walmart stores during the specified period.

# Task 9 :Top 5 Customers by Sales Volume

-- Task 9 Finding Top 5 Customers by Sales Volume

```
SELECT `Customer ID`, round(SUM(Total),2) AS total_spent  
FROM walmartsales  
GROUP BY `Customer ID`  
ORDER BY total_spent DESC  
LIMIT 5;
```

Result Grid		
	Customer ID	total_spent
▶	8	26634.34
	3	23402.26
	2	23392.28
	15	22674.46
	1	22534.55

This analysis identifies the top five customers who contributed the most significant sales volume during the specified period, offering valuable insights into customer purchasing behaviors and preferences for targeted marketing strategies.



# Task10: Sales Trend Analysis by Day

```
-- Task 10 Sales Trend by Day of the Week  
SELECT DAYNAME(STR_TO_DATE(Date, '%d-%m-%Y')) AS day_of_week, round(SUM(Total),3) AS total_sales  
FROM walmartsales  
GROUP BY day_of_week  
ORDER BY total_sales asc;
```

Result Grid | Filter Rows:

day_of_week	total_sales
Monday	37899.078
Wednesday	43731.135
Friday	43926.341
Sunday	44457.892
Thursday	45349.248
Tuesday	51482.246
Saturday	56120.809

This section examines weekly sales trends for Walmart, highlighting variations in consumer behavior across different days. Understanding these trends can guide inventory management and staffing decisions effectively.



# FINAL INSIGHTS & BUSINESS CONCLUSION

5-Step: A consolidated view of Walmart's sales, customer behavior, product performance, payment trends, and weekly transaction patterns.

Sales Trends & Anomalies	Product Line Insights	Payment & City Trends	Branch Performance	Customer Behavior
Saturday records the highest sales volume weekly.	Food & Beverages emerged as the most profitable category overall.	E-Wallet is the most popular method in Yangon & Mandalay.	Branch A achieved the highest sales growth in March 2019.	High-spending customers contribute the largest revenue share.
Monday shows the lowest sales—opportunity for weekday offers.	Home & Lifestyle leads profit in Branch A.	Cash dominates in Naypyitaw.	Branch performance varies month-to-month, with A leading overall.	Medium spenders form the majority of the customer base.
High anomalies represent bulk or special purchases.	Electronics Accessories perform well but with lower profit margins.	Useful for designing city-specific promo partnerships.	Stable but moderate growth observed in Branch C.	Several customers repurchase within 30 days—indicating loyalty.
<ul style="list-style-type: none"><li>Low anomalies signal discounts or unusually low transactions.</li></ul>	Product-level insights support inventory and demand planning.	Payment preferences vary significantly across cities	<ul style="list-style-type: none"><li>Growth insights help identify operational optimization needs.</li></ul>	Customer segmentation guides targeted marketing strategies.



# Thank You

## Queries & Discussion

Feel free to connect for questions, clarifications, or deeper insights into any SQL operation, workflow, or business analysis covered in this project.

Presented by: Vedang Mishra  
[vedangm64@gmail.com](mailto:vedangm64@gmail.com)  
For Your Attention and Support