

The background features decorative elements in the corners: top-left and bottom-left have overlapping teal and white geometric shapes with thin teal lines; top-right and bottom-right have teal shapes with a grid of small teal circles.

WALMART SALES & PERFORMANCE ANALYSIS

Shrishti Jain

CONTENTS

- About the Project
- About the Data
- Data Cleaning
- Data Exploration
- Data Visualization Dashboard

PROJECT OVERVIEW

This project analyzes a Walmart Sales Dataset . In particular, it will find answer to the following questions:

1. Find different payment methods, number of transactions, and total quantities sold.
2. Identify the highest rated category in each branch with average ratings.
3. Identify the busiest day for each branch based on the number of transactions.
4. Calculate the total number of items sold per payment method.
5. Determine the average, minimum, and maximum ratings of each category per city.
6. Calculate total profit for each category.
7. Determine the most common payment method for each branch.
8. Categorize sales into morning, afternoon, and evening shifts. Then find the number of invoices for each shift.
9. Identify the top 5 branches with the highest decrease in revenue from 2022 to 2023.

ABOUT THE DATA

- The dataset used in analysis contains 11 columns & 10,051 sales records from Walmart Stores. This dataset was downloaded from official Kaggle Website.
- Imported the data into SQL server and then further analysis was carried out.

DATA CLEANING

In this project, we mainly used SQL for both data cleaning and exploration.

Data cleaning included the following steps:

- Split original dataset into two separate tables: transactions and product for better relational structure.
- Checked for Duplicate values and removed them.
- Checked for NULL values and removed them.
- Checked the data type of each column and converted them to suitable data type.
- Reformat the unit_price column values (e.g.= \$63 → 63).
- Verified integrity of joins between transactions and product using invoice_id.

DATA EXPLORATION

- Added a new column name total_spending in transactions table for better analysis.
- SQL skills used for further analysis are:-
 - Joins
 - Common Table Expressions (CTEs)
 - Aggregate Functions
 - Data type Conversion
 - Relational operators
 - Window Functions
 - CASE statements

DATA EXPLORATION

I. Find different payment methods, number of transactions, and total quantities sold.

This query shows the number of transactions made and the total quantity of items sold for each payment method. It reveals that the majority of transactions and items sold were through the Credit Card method, with 4,256 transactions and 9,567 quantities sold, making it the most preferred payment option among customers.

Results Messages			
	payment_method	transaction_count	quantities_sold
1	Credit card	4256	9567
2	Ewallet	3881	8932
3	Cash	1832	4984

DATA EXPLORATION

2. Identify the highest rated category in each branch with average ratings.

For each branch, the category with the highest rating is selected. For example, Branch WALM001 shows Fashion Accessories & Electronic accessories as the top-rated category with maximum rating above 9, indicating strong customer satisfaction in that segment.

Results		Messages		
	Branch	category	max_rating	avg_rating
1	WALM001	Fashion accessories	9.5	6.36
2	WALM001	Electronic accessories	9.5	7.45
3	WALM002	Fashion accessories	9.7	5.96
4	WALM003	Electronic accessories	9.9	5.52
5	WALM004	Food and beverages	9.3	9.3
6	WALM005	Electronic accessories	9.9	6.8
7	WALM006	Home and lifestyle	9.8	5.88
8	WALM007	Food and beverages	9.6	7.55
9	WALM008	Home and lifestyle	9.7	6.63
10	WALM009	Sports and travel	9.6	9.6
11	WALM010	Fashion accessories	9.6	6.94
12	WALM011	Health and beauty	9.5	6.27
13	WALM012	Food and beverages	9.9	6.87
14	WALM013	Fashion accessories	9.8	6.21
15	WALM014	Fashion accessories	9	6.69
16	WALM014	Home and lifestyle	9	6.4
17	WALM015	Fashion accessories	9	6.99

DATA EXPLORATION

3. Identify the busiest day for each branch based on the number of transactions

By counting daily transactions and ranking them, the busiest day is identified. For instance, Branch WALM001 experiences the most footfall on Thursday, suggesting higher customer activity during weekends.

Results				Messages			
	Branch	day_name	no_of_trans				
1	WALM001	Thursday	16				
2	WALM002	Thursday	15				
3	WALM003	Tuesday	33				
4	WALM004	Sunday	14				
5	WALM005	Wednesday	19				
6	WALM006	Thursday	15				
7	WALM007	Sunday	12				
8	WALM007	Friday	12				
9	WALM008	Tuesday	17				
10	WALM009	Sunday	42				
11	WALM010	Wednesday	12				
12	WALM011	Tuesday	18				
13	WALM012	Sunday	20				
14	WALM013	Monday	13				
15	WALM014	Sunday	12				
16	WALM015	Friday	15				
17	WALM016	Tuesday	16				

DATA EXPLORATION

4. Calculate the total number of items sold per payment method.

By summing up the quantity of items per method, we gain insights into customer preferences. For example, **Credit Card** transactions account for the **highest** number of items sold ie. **9,567** items, indicating it's a popular payment option among customers.

Results Messages		
	payment_method	total_items_sold_per_payment_method
1	Credit card	9567
2	Ewallet	8932
3	Cash	4984

DATA EXPLORATION

5. Determine the average, minimum, and maximum ratings of each category per city.

This helps in evaluating customer satisfaction regionally. For instance, in **Abilene**, the **Health and Beauty** category has consistently high average rating of **9.7**, indicating strong customer preference in that city.

Results		Messages			
	City	category	avg_rating	min_rating	max_rating
1	Abilene	Health and beauty	9.7	9.7	9.7
2	Abilene	Electronic accessories	7.97	7.1	8.8
3	Abilene	Food and beverages	6.95	6	8.9
4	Abilene	Fashion accessories	6.24	4	9
5	Abilene	Home and lifestyle	6.1	4	9
6	Alamo	Health and beauty	7.95	7.7	8.2
7	Alamo	Sports and travel	7.3	5	10
8	Alamo	Fashion accessories	6.87	3	9
9	Alamo	Home and lifestyle	6.3	3	9
10	Alamo	Food and beverages	5.2	5.2	5.2
11	Alice	Food and beverages	7.67	5	9.2
12	Alice	Electronic accessories	7.3	7.3	7.3
13	Alice	Sports and travel	6.93	6.5	7.9
14	Alice	Home and lifestyle	6.04	4	9
15	Alice	Fashion accessories	5.93	3	9
16	Allen	Fashion accessories	6.59	3	9
17	Allen	Electronic accessories	6.4	6.4	6.4

DATA EXPLORATION

6. Calculate total profit for each category.

This query calculates the total profit for each product category using the formula (unit price × quantity × profit margin). It helps determine the most profitable categories. From the results, **Fashion accessories** emerges as the top contributor to overall profit with the profit of 1,92,314.89, followed by Home and Lifestyle.

Results Messages		
	category	total_profit
1	Fashion accessories	192314.8932
2	Home and lifestyle	192213.6381
3	Electronic accessories	30772.4895
4	Food and beverages	21552.8622
5	Sports and travel	20613.8082
6	Health and beauty	18671.7345

DATA EXPLORATION

7. Determine the most common payment method for each branch.

This query identifies the most commonly used payment method in each branch by counting the number of transactions per method. It reveals customer preferences in different regions. For instance, **Ewallet** was the most preferred method in Branch **WALM001**, while **Credit Card** dominated in Branch **WALM003**.

Results Messages			
	Branch	payment_method	preferred_payment_method
1	WALM001	Ewallet	45
2	WALM002	Ewallet	37
3	WALM003	Credit card	115
4	WALM004	Ewallet	44
5	WALM005	Ewallet	56
6	WALM006	Ewallet	50
7	WALM007	Ewallet	52
8	WALM008	Ewallet	39
9	WALM009	Credit card	139
10	WALM010	Ewallet	47
11	WALM011	Ewallet	39
12	WALM012	Ewallet	52
13	WALM013	Ewallet	44
14	WALM014	Ewallet	28
15	WALM015	Ewallet	57
16	WALM016	Ewallet	46
17	WALM017	Ewallet	46

DATA EXPLORATION

8. Categorize sales into morning, afternoon, and evening shifts. Then find the number of invoices for each shift

This query categorizes transactions into three shifts: **Morning** (00:00–11:59), **Afternoon** (12:00–16:59), and **Evening** (17:00–23:59), based on the time of purchase. It shows the number of invoices generated in each shift. The analysis reveals that the **Evening** shift had the highest number of transactions, indicating peak shopping hours.

Results			Messages		
	no_of_transactions		shift		
1	4273		Evening		
2	3609		Afternoon		
3	2087		Morning		

DATA EXPLORATION

9. Identify the top 5 branches with the highest decrease in revenue from 2022 to 2023.

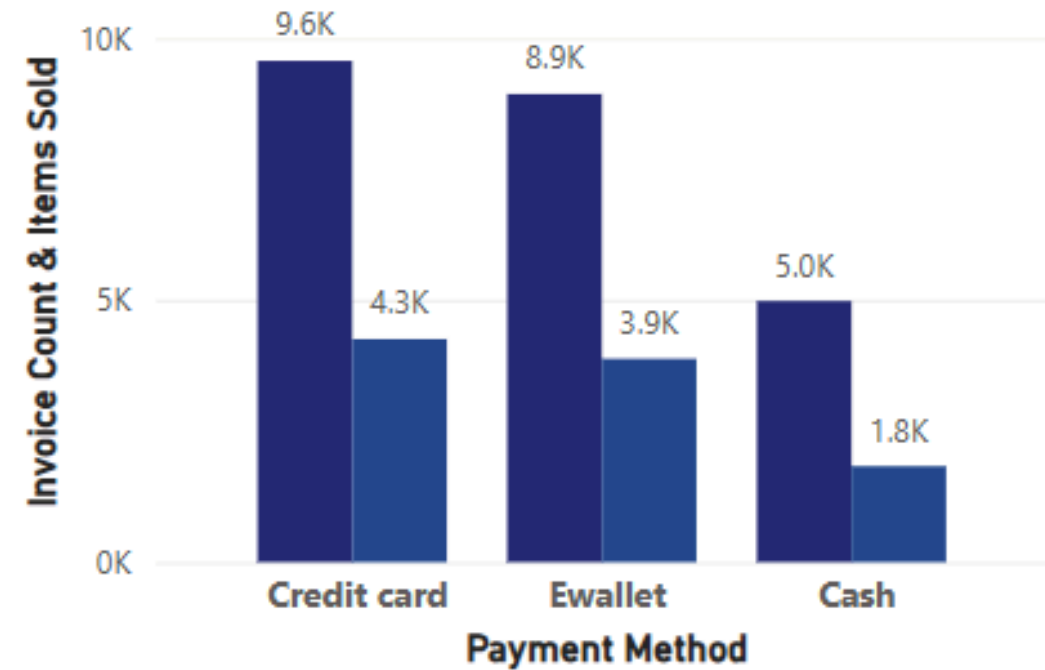
It calculates the percentage decrease and lists the top 5 branches with the most significant drop. This helps in identifying underperforming branches that may need attention or strategic intervention.

Results		Messages		
	Branch	revn_2022	revn_2023	dec_ratio
1	WALM045	1731	647	62.62
2	WALM047	2581	1069	58.58
3	WALM098	2446	1030	57.89
4	WALM033	2099	931	55.65
5	WALM081	1723	850	50.67

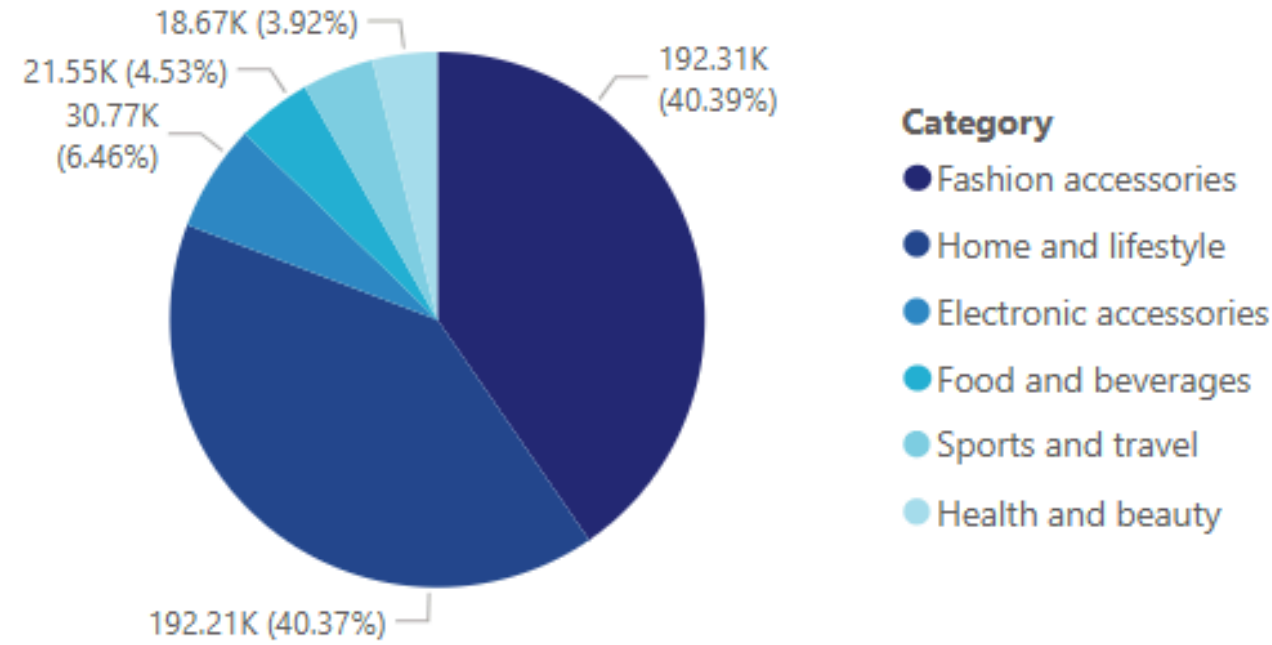
POWER BI DASHBOARD

Payment Method vs Sales & Transactions

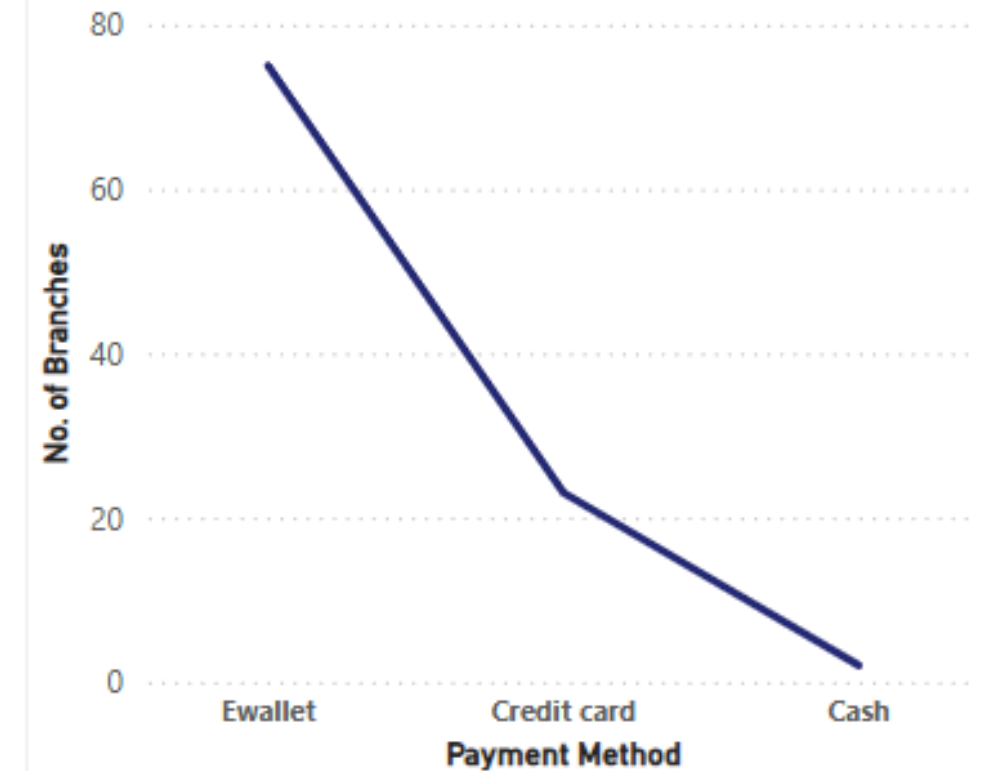
● Sum of Quantity Sold ● Sum of Transactions



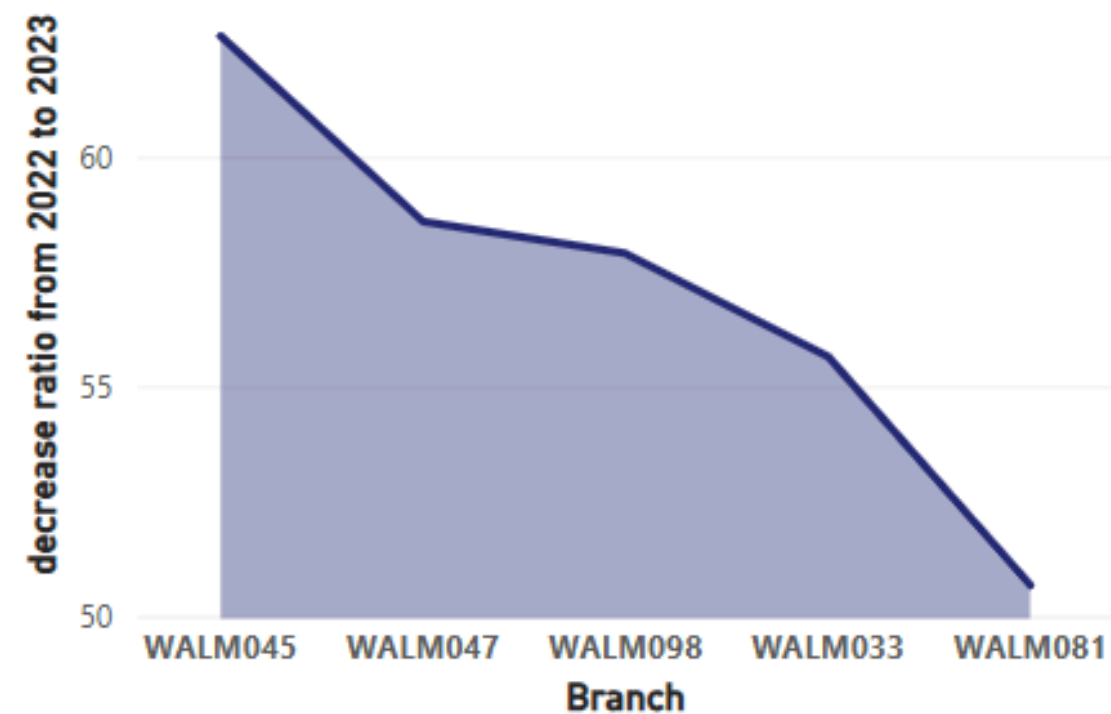
Category vs Total Profit



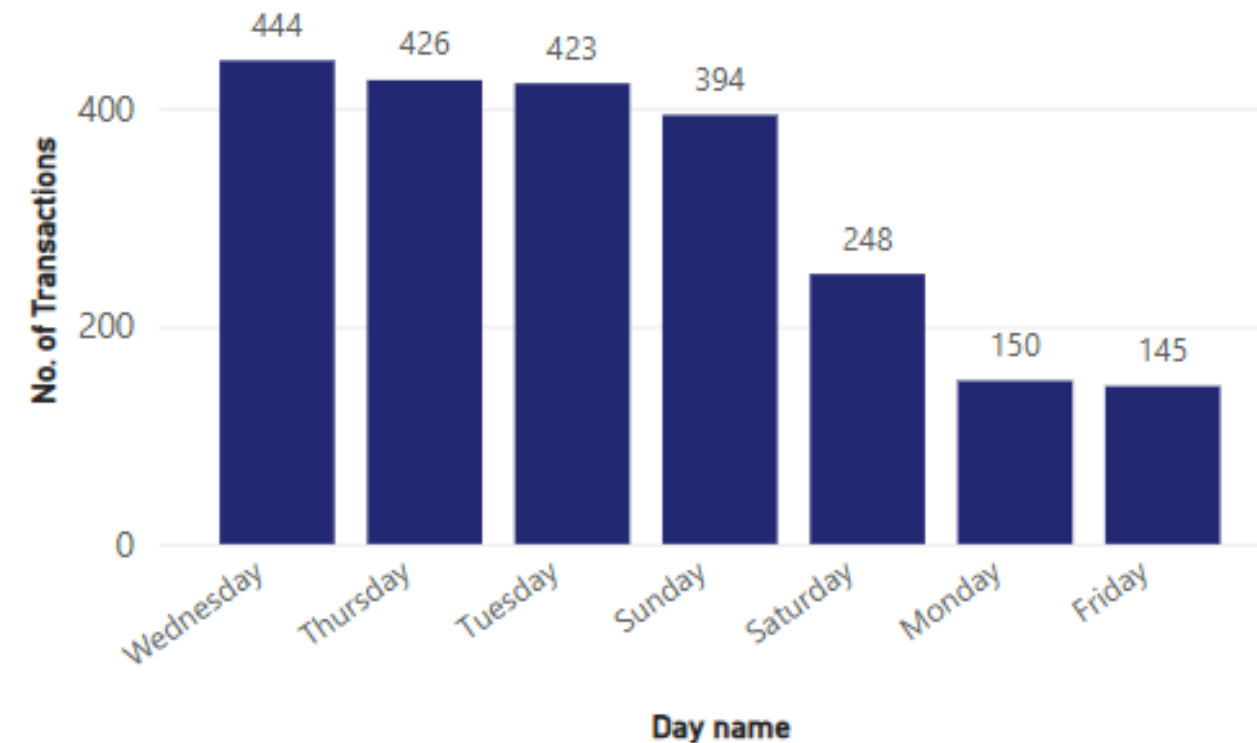
Preferred Payment Method



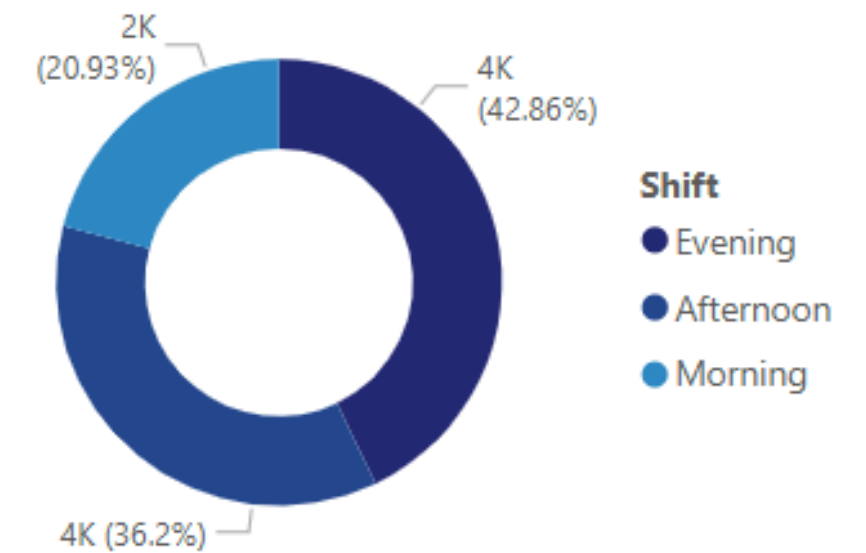
Branch vs Decrease ratio



Weekday vs No. of Transactions



Shift vs No. of Transactions





THANK YOU

By Shrishti Jain