Sales Analysis Using SQL

Agenda

In this project, we conducted a comprehensive analysis of Walmart's sales dataset using SQL to extract actionable business insights. The key goal was to support operational decisions across branches, products, timeframes, and customer behaviour.

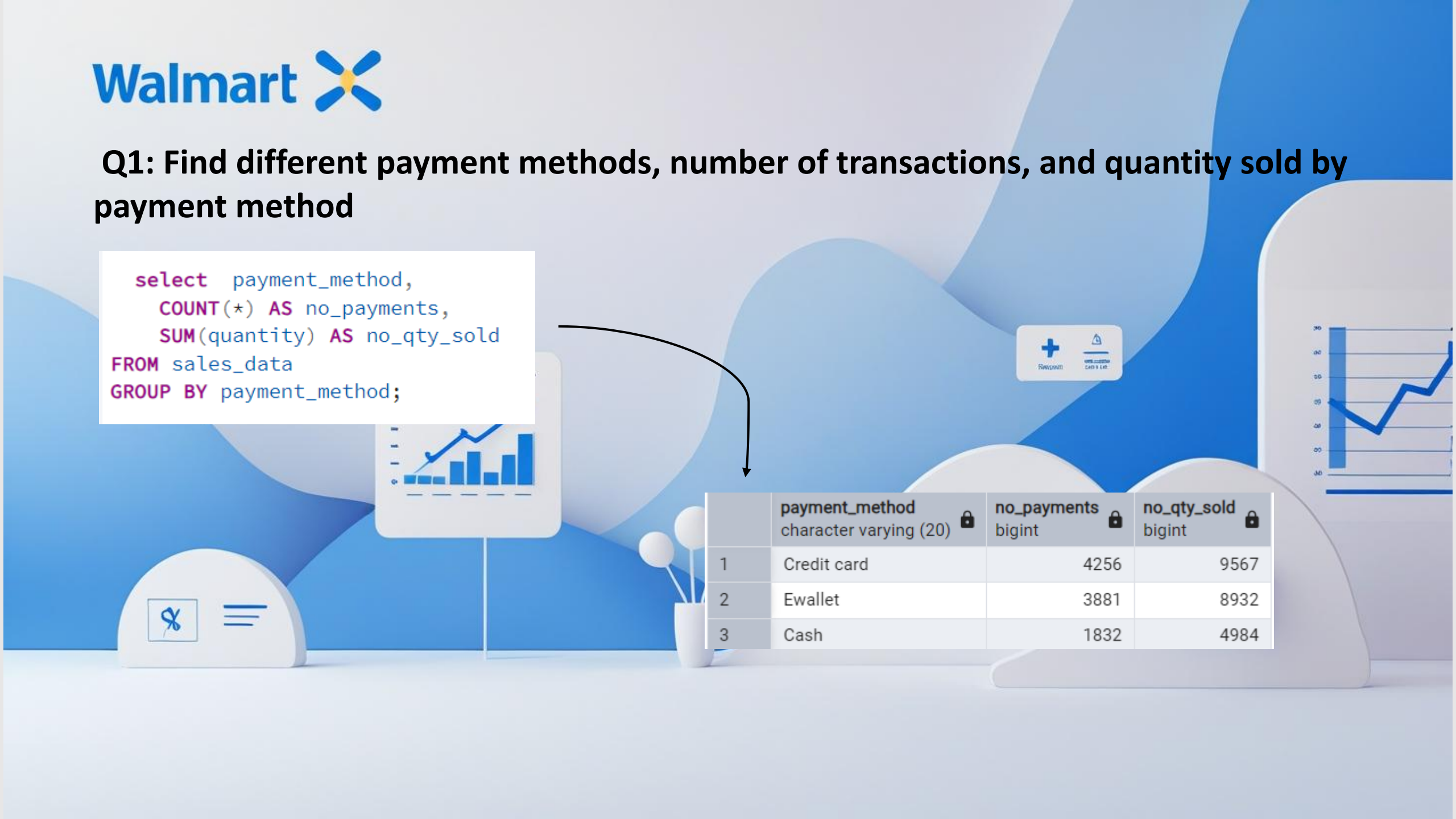
Key Business Questions Addressed:

- What are the most popular **payment methods** and how do they influence sales quantity?
- Which **product categories** are most highly rated in each branch?
- What is the **busiest day** for each branch, based on transaction count?
- How do **sales ratings** vary by **city and category**?
- What is the **most common payment method** in each branch?
- How are sales distributed across **shifts** (morning, afternoon, evening)?
- Which branches experienced the highest **year-over-year revenue decline**?
- What is the **total profit** generated per product category?

Using advanced SQL techniques such as window functions, case statements, and subqueries, we were able to identify patterns in customer behaviour, operational performance, and sales strategy.
This data-driven approach equips Walmart with actionable insights for optimizing product offerings, staffing, marketing strategies, and overall retail performance.

Q1: Find different payment methods, number of transactions, and quantity sold by payment method

```
  select  payment_method,
    COUNT(*) AS no_payments,
    SUM(quantity) AS no_qty_sold
FROM sales_data
GROUP BY payment_method;
```



	payment_method character varying (20)	no_payments bigint	no_qty_sold bigint
1	Credit card	4256	9567
2	Ewallet	3881	8932
3	Cash	1832	4984

Q2: Highest-Rated Category in Each Branch ---- Display the branch, category, and avg rating

```sql
SELECT branch, category, avgr, rank
FROM (
  SELECT branch, category, AVG(rating) AS avgr,
         RANK() OVER(PARTITION BY branch ORDER BY AVG(rating) DESC) AS rank
  FROM sales_data
  GROUP BY 1, 2
) AS ranked
WHERE rank = 1
```

	branch character varying (10)	category character varying (50)	avgr numeric	rank bigint
1	WALM001	Electronic accessories	7.4500000000000000	1
2	WALM002	Food and beverages	8.2500000000000000	1
3	WALM003	Sports and travel	7.5000000000000000	1
4	WALM004	Food and beverages	9.3000000000000000	1
5	WALM005	Health and beauty	8.3666666666666667	1
6	WALM006	Fashion accessories	6.7970588235294118	1
7	WALM007	Food and beverages	7.5500000000000000	1
8	WALM008	Food and beverages	7.4000000000000000	1
9	WALM009	Sports and travel	9.6000000000000000	1
10	WALM010	Electronic accessories	9.0000000000000000	1
11	WALM011	Food and beverages	7.0000000000000000	1
12	WALM012	Health and beauty	7.4500000000000000	1
13	WALM013	Health and beauty	7.6000000000000000	1
14	WALM014	Electronic accessories	6.8333333333333333	1
15	WALM015	Home and lifestyle	6.2230769230769231	1
16	WALM016	Sports and travel	9.1000000000000000	1
17	WALM017	Electronic accessories	7.0000000000000000	1
18	WALM018	Electronic accessories	8.7500000000000000	1
19	WALM019	Electronic accessories	8.4000000000000000	1
20	WALM020	Food and beverages	8.3333333333333333	1
21	WALM021	Sports and travel	7.2000000000000000	1

Q3: Calculate the total quantity of items sold per payment method

```sql
SELECT payment_method, SUM(quantity) as total_quantity
FROM sales_data
GROUP BY payment_method;
```

	payment_method character varying (20)	total_quantity bigint
1	Credit card	9567
2	Ewallet	8932
3	Cash	4984

Q4: Determine the average, minimum, and maximum rating of categories for each city

```
SELECT city, category, AVG(rating), MIN(rating), MAX(rating)
FROM sales_data
GROUP BY 1, 2;
```

	city character varying (50)	category character varying (50)	avg numeric	min numeric	max numeric
1	Little Elm	Fashion accessories	6.1181818181818182	4.0	9.6
2	Mesquite	Sports and travel	7.8000000000000000	7.8	7.8
3	Canyon	Health and beauty	6.9000000000000000	5.8	8.9
4	McKinney	Home and lifestyle	5.9270270270270270	3.0	9.0
5	Brownwood	Food and beverages	7.8000000000000000	6.4	9.2
6	Flower Mound	Health and beauty	7.9500000000000000	6.4	9.5
7	Edinburg	Fashion accessories	6.7307692307692308	3.0	9.0
8	Pharr	Health and beauty	9.2000000000000000	9.2	9.2
9	Mineral Wells	Fashion accessories	6.1866666666666667	3.0	9.0
10	Edinburg	Health and beauty	8.6000000000000000	8.6	8.6
11	Lufkin	Fashion accessories	5.9750000000000000	3.0	9.9
12	La Porte	Fashion accessories	6.5870967741935484	3.0	9.1
13	San Angelo	Sports and travel	9.3000000000000000	9.3	9.3
14	Seguin	Electronic accessories	5.0615384615384615	3.0	7.0
15	Kerrville	Health and beauty	8.3000000000000000	7.9	8.7
16	Kerrville	Home and lifestyle	6.2125000000000000	4.0	9.0
17	Rowlett	Electronic accessories	5.9666666666666667	3.0	9.5
18	North Richland Hills	Food and beverages	5.8666666666666667	5.0	6.3
19	Haltom City	Electronic accessories	8.0333333333333333	4.5	9.9
20	Flower Mound	Home and lifestyle	6.2589743589743590	4.0	9.6
21	Big Spring	Sports and travel	7.3400000000000000	6.8	8.0

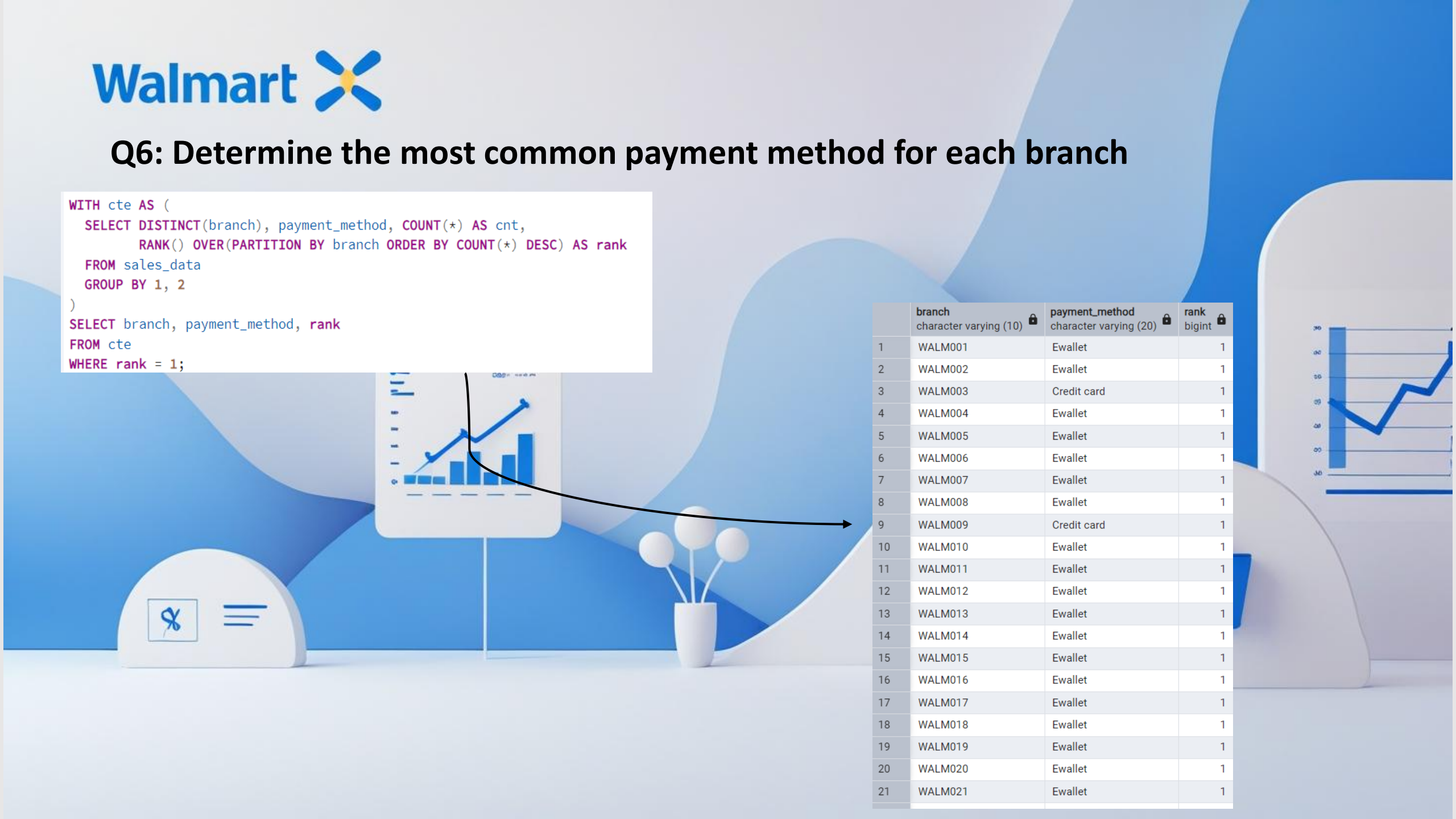
Q5: Calculate the total profit for each category

```sql
SELECT category,
       SUM(total_sales),
       SUM(total_sales * profit_margin)
FROM sales_data
GROUP BY 1;
```

	category character varying (50)	sum numeric	sum numeric
1	Fashion accessories	489480.90	192314.8932
2	Electronic accessories	78175.03	30772.4895
3	Health and beauty	46851.18	18671.7345
4	Food and beverages	53471.28	21552.8622
5	Sports and travel	52497.93	20613.8082
6	Home and lifestyle	489250.06	192213.6381

Q6: Determine the most common payment method for each branch

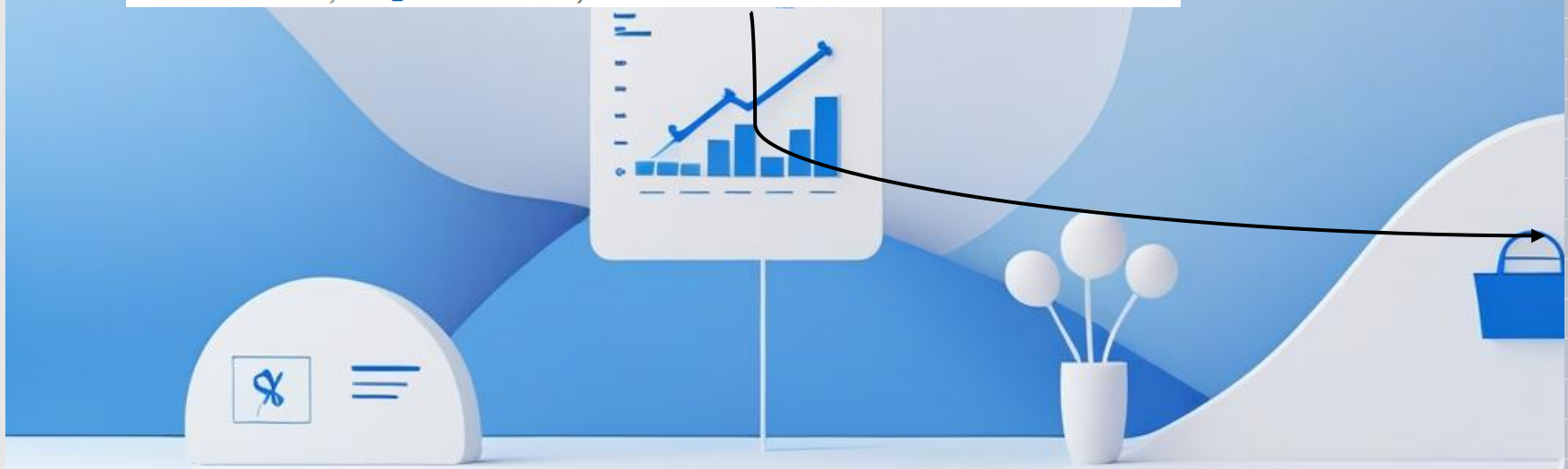
```sql
WITH cte AS (
  SELECT DISTINCT(branch), payment_method, COUNT(*) AS cnt,
         RANK() OVER(PARTITION BY branch ORDER BY COUNT(*) DESC) AS rank
  FROM sales_data
  GROUP BY 1, 2
)
SELECT branch, payment_method, rank
FROM cte
WHERE rank = 1;
```



	branch character varying (10)	payment_method character varying (20)	rank bigint
1	WALM001	Ewallet	1
2	WALM002	Ewallet	1
3	WALM003	Credit card	1
4	WALM004	Ewallet	1
5	WALM005	Ewallet	1
6	WALM006	Ewallet	1
7	WALM007	Ewallet	1
8	WALM008	Ewallet	1
9	WALM009	Credit card	1
10	WALM010	Ewallet	1
11	WALM011	Ewallet	1
12	WALM012	Ewallet	1
13	WALM013	Ewallet	1
14	WALM014	Ewallet	1
15	WALM015	Ewallet	1
16	WALM016	Ewallet	1
17	WALM017	Ewallet	1
18	WALM018	Ewallet	1
19	WALM019	Ewallet	1
20	WALM020	Ewallet	1
21	WALM021	Ewallet	1

Q7: Categorize sales into Morning, Afternoon, and Evening shifts

```sql
SELECT branch,
       CASE
         WHEN EXTRACT(HOUR FROM time::time) < 12 THEN 'Morning'
         WHEN EXTRACT(HOUR FROM time::time) BETWEEN 12 AND 17 THEN 'Afternoon'
         ELSE 'Evening'
       END AS shift,
       COUNT(*) AS num_invoices
FROM sales_data
GROUP BY branch, shift
ORDER BY branch, num_invoices DESC;
```



	branch character varying (10)	shift text	num_invoices bigint
1	WALM001	Afternoon	36
2	WALM001	Evening	30
3	WALM001	Morning	8
4	WALM002	Afternoon	29
5	WALM002	Evening	21
6	WALM002	Morning	15
7	WALM003	Afternoon	95
8	WALM003	Morning	50
9	WALM003	Evening	41
10	WALM004	Afternoon	27
11	WALM004	Evening	24
12	WALM004	Morning	9
13	WALM005	Evening	35
14	WALM005	Afternoon	34
15	WALM005	Morning	15
16	WALM006	Afternoon	33
17	WALM006	Evening	31
18	WALM006	Morning	7
19	WALM007	Evening	33
20	WALM007	Afternoon	27
21	WALM007	Morning	10

Q8: Identify the 5 branches with the highest revenue decrease ratio from last year to current year (e.g., 2022 to 2023)

```sql
WITH revenue_2022 AS (
  SELECT branch, SUM(total_sales) AS revenue
  FROM sales_data
  WHERE EXTRACT(YEAR FROM date::date) = 2022
  GROUP BY branch
),
revenue_2023 AS (
  SELECT branch, SUM(total_sales) AS revenue
  FROM sales_data
  WHERE EXTRACT(YEAR FROM date::date) = 2023
  GROUP BY branch
)
SELECT r2.branch,
       r2.revenue AS rn2,
       r3.revenue AS rn3,
       ROUND((r2.revenue - r3.revenue)::numeric / r2.revenue::numeric * 100) AS decreasing_revenue
FROM revenue_2022 AS r2
JOIN revenue_2023 AS r3 ON r2.branch = r3.branch
WHERE r2.revenue > r3.revenue
ORDER BY decreasing_revenue DESC
LIMIT 5;
```

	branch character varying (10)	rn2 numeric	rn3 numeric	decreasing_revenue numeric
1	WALM045	1731.00	647.00	63
2	WALM047	2581.00	1069.00	59
3	WALM098	2446.00	1030.00	58
4	WALM033	2099.00	931.00	56
5	WALM081	1723.00	850.00	51

Walmart

Insight & Recommendation

1. Promote Ewallet Payment Method Insight:

-Ewallet is the most popular payment method across branches .

Recommendation :

-Encourage more customers to use Ewallet by offering exclusive discounts or reward points.

-Streamline and speed up Ewallet transactions to enhance customer satisfaction.

2. Optimize Shift Scheduling Based on Invoice Volume Insight:

-Some branches (e.g., WALM003) see much higher invoice volumes in the afternoon compared to morning or evening.

Recommendation :

-Analyse invoice trends per branch and adjust staffing accordingly.

-Allocate more employees during peak hours to reduce wait times and improve service.

3. Invest in High-Performing Categories Insight:

-Fashion Accessories and Home and Lifestyle generate the highest total sales (≈489K each).

Recommendation :

-Prioritize inventory and promotions for these top-performing categories.

-Explore customer preferences within these categories to introduce new but related products.

Insight & Recommendation

4. Improve Performance in Low-Scoring Cities Insight:
-Some cities like Seguin and Kerrville have the lowest average ratings (around 5.05 and 5.96).
Recommendation :
-Conduct surveys or feedback sessions in these locations to identify customer pain points.
-Improve product variety, service quality, or store ambiance in these cities.

5**. Targeted Category Improvement by Branch Insight:**
-Category performance varies by branch; for instance, “Food and Beverages” is highly rated in some branches and low in others.
Recommendation :
-Use localized marketing strategies to push underperforming categories where needed.
-Consider retraining staff or revamping layout in branches where specific categories underperform.
6. Align Inventory with Sales & Rating Data Insight:
-While some categories have high ratings, their sales may not match (e.g., “Sports and Travel” has high avg ratings but lower total revenue).
Recommendation :
-Increase visibility of high-rated but under-selling categories with better placement or bundling deals. Investigate why high-quality categories aren't converting into sales — pricing? visibility? availability?_