Coffee Sales Analysis Report

Using SQL & Power BI

A comprehensive data-driven approach to optimizing vending machine performance, revenue, and customer satisfaction through advanced analytics.

$2,935.72	3,898	$0.75	10
Total Revenue (USD)	Total Orders	Avg Order Value	Coffee Varieties

PostgreSQL   Power BI   Valentina Studio   USD Currency   Mar 2024 - Mar 2025

Prepared by: Data Analyst

Project Overview

This report consolidates, cleans, and analyzes two years of coffee vending machine sales data (2024 & 2025) using SQL to uncover business performance trends, customer purchasing habits, and product-level insights.

Field	Details
Domain	Retail / Vending Machine Sales
Tools	PostgreSQL, Valentina Studio, Power BI
Data Period	March 2024 – March 2025 (ongoing)
Currency	US Dollar (USD)
Tables	index_1 (2024), index_2 (2025), coffee_sales (merged)
Total Records	3,898 transactions
Status	Completed

Dataset Structure

Column	Data Type	Description
date	DATE	Date of transaction
datetime	TIMESTAMP	Full timestamp of purchase
cash_type	VARCHAR	Payment method: cash or card
card	VARCHAR	Anonymous card number (2024 only)
money	NUMERIC	Transaction amount in US Dollars (USD)
coffee_name	VARCHAR	Name of coffee product purchased

Objectives

1	Combine 2024 & 2025 datasets into a unified PostgreSQL database
2	Analyze total revenue, growth trends, and monthly performance
3	Identify top-performing coffee products by revenue and order count
4	Understand customer payment preferences (cash vs. card)
5	Visualize key findings through an interactive Power BI dashboard

Data Preparation & Cleaning

Step 1 — Combine Tables Using UNION ALL

Both yearly tables share the same schema. They were merged into a single **coffee_sales** table using UNION ALL.

```sql
CREATE TABLE coffee_sales AS

SELECT date, datetime, cash_type, money::NUMERIC, coffee_name FROM index_1

UNION ALL

SELECT date, datetime, cash_type, money::NUMERIC, coffee_name FROM index_2;
```

Step 2 — Convert Money Column to NUMERIC

The **money** column was stored as text. ALTER TABLE was used to cast it to NUMERIC before the merge.

```sql
ALTER TABLE index_1 ALTER COLUMN money TYPE NUMERIC USING money::NUMERIC;

ALTER TABLE index_2 ALTER COLUMN money TYPE NUMERIC USING money::NUMERIC;
```

Resulting Unified Table — Sample Records

date	datetime	cash_type	money (USD)	coffee_name
2024-03-01	2024-03-01 10:15:50	card	$0.93	Latte
2024-03-01	2024-03-01 12:19:22	card	$0.93	Hot Chocolate
2024-03-01	2024-03-01 13:46:33	card	$0.69	Americano
2024-03-01	2024-03-01 15:39:47	card	$0.81	Americano with Milk
2024-03-02	2024-03-02 10:30:35	cash	$0.96	Latte
2024-03-02	2024-03-02 10:41:41	card	$0.81	Americano with Milk

Data Quality Notes

Issue / Observation	Resolution / Action
money column stored as TEXT	Cast to NUMERIC using ::NUMERIC operator before merge
card column absent in index_2 (2025)	Excluded from merged table; only present in 2024 data
No missing values in key fields	Verified with COUNT(*) vs COUNT(column) checks
Date format consistent across tables	Used TO_DATE() for month extraction in time-series queries

Revenue Analysis

KEY METRICS

Overall Revenue Summary

$2,935.72	3,898	$0.75	$0.36
Total Revenue (USD)	Total Orders	Avg Order Value	Min Order

```sql
SELECT COUNT(*) AS total_orders, SUM(money) AS total_revenue,

MIN(money) AS min_order, MAX(money) AS max_order,

ROUND(AVG(money),2) AS avg_order_value

FROM coffee_sales;
```

Year-over-Year Revenue Comparison

Sales data for 2025 only covers January–March (partial year). The difference in totals reflects data coverage gap, not an actual business decline.

Metric	2024 (Full Year)	2025 (Jan–Mar Only)	Note
Total Revenue (USD)	$2,770.36	$165.36	Partial year 2025
Growth Difference		-$2,604.72	Not comparable — partial data
Months Covered	10 months (Mar–Dec)	3 months (Jan–Mar)	2025 data still growing

* The large negative growth difference is because 2025 data covers only 3 months vs 10 months in 2024. A fair comparison requires the same time window.

```sql
WITH totals AS (

SELECT '2024' AS year, SUM(money) AS sales FROM index_1 UNION ALL

SELECT '2025', SUM(money) FROM index_2

)

SELECT MAX(CASE WHEN year='2024' THEN sales END) AS sales_2024,

MAX(CASE WHEN year='2025' THEN sales END) AS sales_2025,

MAX(CASE WHEN year='2025' THEN sales END) -

MAX(CASE WHEN year='2024' THEN sales END) AS growth_difference

FROM totals;
```

Product Performance Analysis

TOP SELLERS

Top 10 Coffee Types by Total Revenue (Combined)

Latte is the clear revenue leader, followed by Americano with Milk and Cappuccino. These three products together account for the majority of total revenue.

Rank	Coffee Name	Total Revenue (USD)	Revenue Share
1	Latte	$687.80	~23.4%
2	Americano with Milk	$606.46	~20.7%
3	Cappuccino	$444.34	~15.1%
4	Americano	$370.49	~12.6%
5	Hot Chocolate	$244.14	~8.3%
6	Cocoa	$208.28	~7.1%
7	Cortado	$180.84	~6.2%
8	Espresso	$76.49	~2.6%
9	Americano with milk	$26.40	~0.9%
10	Irish whiskey	$12.60	~0.4%

Most Popular by Order Count

Rank	Coffee Name	Total Orders	Avg Revenue/Order (USD)
1	Americano with Milk	824	~$0.74
2	Latte	806	~$0.85
3	Americano	593	~$0.63
4	Cappuccino	517	~$0.86
5	Cortado	292	~$0.62
6	Hot Chocolate	282	~$0.87
7	Cocoa	243	~$0.86
8	Espresso	152	~$0.50

```
SELECT coffee_name, SUM(money) AS revenue

FROM coffee_sales GROUP BY coffee_name ORDER BY revenue DESC;
```

Monthly Trends & Payment Analysis   TIME SERIES

Monthly Revenue — 2024 vs 2025 (USD)

Month	Sales 2024 (USD)	Sales 2025 (USD)	YoY Diff	Growth %
2024-03	$169.20	$0	-$169.20	-100%
2024-04	$161.29	$0	-$161.29	-100%
2024-05	$217.52	$0	-$217.52	-100%
2024-06	$186.21	$0	-$186.21	-100%
2024-07	$165.98	$0	-$165.98	-100%
2024-08	$182.73	$0	-$182.73	-100%
2024-09	$239.73	$0	-$239.73	-100%
2024-10	$333.39	$0	-$333.39	-100%
2024-11	$206.17	$0	-$206.17	-100%
2024-12	$197.70	$0	-$197.70	-100%
2025-01	$153.57	$0	-$153.57	-100%
2025-02	$317.17	$86.14	-$231.03	-72.84%
2025-03	$239.67	$79.22	-$160.45	-66.95%

* -100% for 2024-only months occurs because 2025 data does not exist for those months. Valid YoY comparison only available for Feb and March.

Payment Method Analysis

Card payments represent 95.66% of all transactions — confirming near-cashless customer behavior.

Payment Type	Count 2024	Count 2025	Total	Share
Card	3,547	182	3,729	95.66%
Cash	89	80	169	4.34%
TOTAL	3,636	262	3,898	100%

Key Insights & Recommendations

ACTION ITEMS

Key Findings Summary

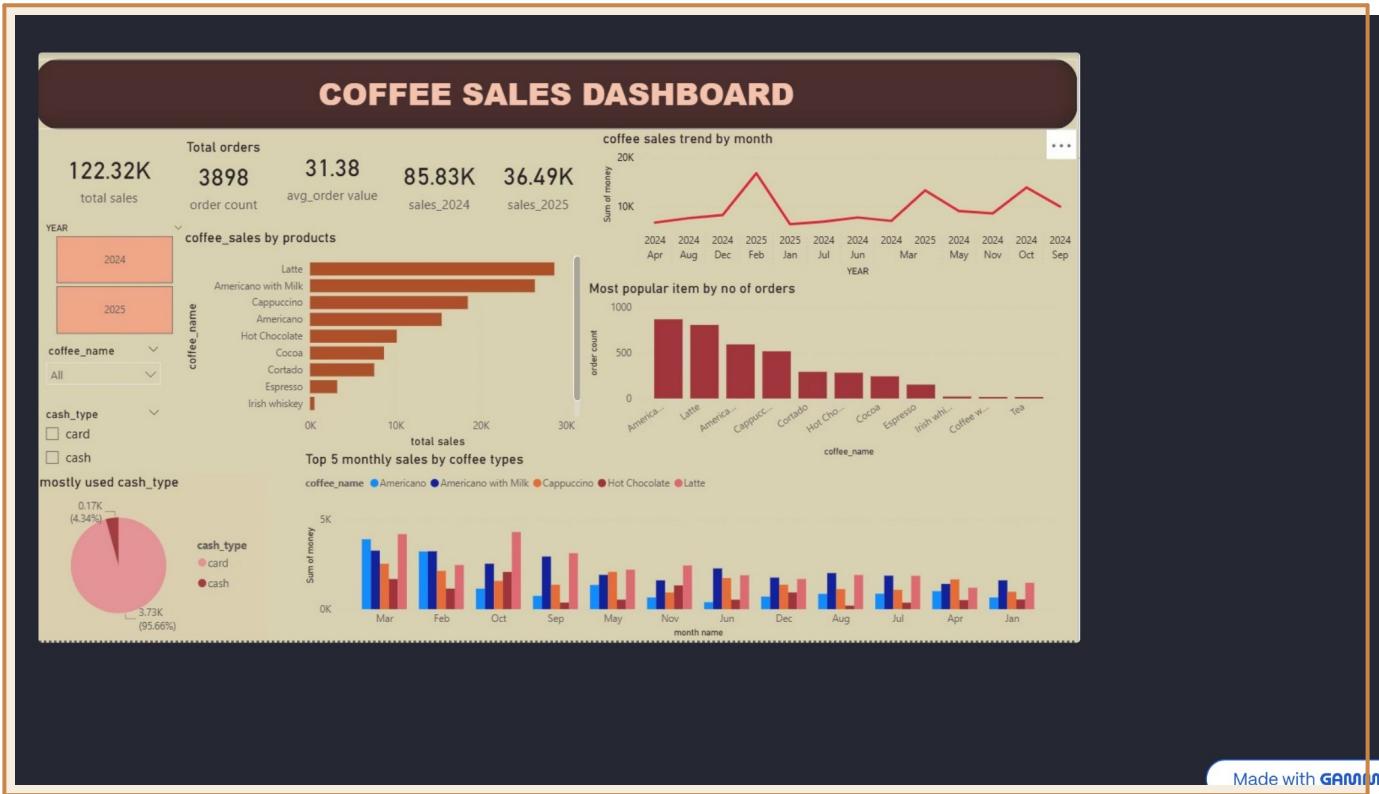| Category	Finding
Revenue	Total revenue of $2,935.72 USD from 3,898 orders over ~13 months of data
Top Product	Latte leads revenue ($687.80); Americano with Milk leads order count (824 orders)
Peak Month	October 2024 was the highest revenue month at $333.39 USD
Payment	95.66% of transactions are card-based — nearly cashless operation
2025 Trend	Feb–Mar 2025 shows continued activity; full year comparison pending
Product Breadth	18 distinct products; top 4 account for ~72% of all revenue

Recommendations

1. Stock Optimization	Prioritize Latte, Americano with Milk, and Cappuccino. These 3 SKUs account for >50% of revenue. Ensure consistent availability.
2. Menu Rationalization	Remove low-volume specialty items (Irish whiskey, 21 orders; $12.60 total) to free machine space for higher-demand products.
3. Contactless Payments	With 95.66% card usage, ensure card reader reliability and consider adding QR/digital payment support.
4. Replicate October Peak	October 2024 was the strongest month ($333.39). Investigate what drove this and replicate conditions in 2025.
5. Continue Monitoring	Re-run monthly as 2025 data grows for a true YoY comparison. Use same SQL queries on refreshed data.
6. Price Testing	Avg order value is $0.75. Americano has the lowest avg (~$0.63). Test modest price increase to lift the blended average.

Power BI Dashboard

The interactive Power BI dashboard below visualizes all key findings — revenue trends, product performance, payment preferences, and monthly comparisons. Built using the unified **coffee_sales** PostgreSQL table exported to Power BI Desktop.

Dashboard Components

Visual Component	Description
KPI Cards	Total Sales ($2,935.72 USD), Order Count (3,898), Avg Order Value ($0.75), Sales by Year
Sales Trend Line Chart	Monthly revenue trend showing Oct 2024 peak and 2025 partial-year data
Products Bar Chart	Horizontal bars ranking all coffee types by total revenue — Latte #1
Popular Items Bar Chart	Order count by product — Americano with Milk leads with 824 orders
Monthly by Coffee Type	Grouped bar chart of top 5 products across all months
Cash Type Donut	95.66% card vs 4.34% cash — near-cashless customer behavior
Year & Product Slicers	Interactive filters by 2024/2025, specific products, or payment types

SQL Query Reference

All SQL queries used in this analysis are documented below for reproducibility.

Q1 — Combine Tables

```
CREATE TABLE coffee_sales AS

SELECT date, datetime, cash_type, money::NUMERIC, coffee_name FROM index_1

UNION ALL

SELECT date, datetime, cash_type, money::NUMERIC, coffee_name FROM index_2;
```

Q2 — Total Sales by Year

```
SELECT SUM(money) AS total_sales_2024 FROM public.index_1;

SELECT SUM(money) AS total_sales_2025 FROM public.index_2;
```

Q3 — YoY Growth

```
WITH totals AS (

SELECT '2024' AS year, SUM(money) AS sales FROM index_1 UNION ALL

SELECT '2025', SUM(money) FROM index_2

)

SELECT MAX(CASE WHEN year='2024' THEN sales END) AS sales_2024,

MAX(CASE WHEN year='2025' THEN sales END) AS sales_2025,

MAX(CASE WHEN year='2025' THEN sales END) -

MAX(CASE WHEN year='2024' THEN sales END) AS growth_difference

FROM totals;
```

Q4 — Top 10 by Revenue

```
SELECT coffee_name, SUM(money) AS revenue

FROM coffee_sales

GROUP BY coffee_name

ORDER BY revenue DESC LIMIT 10;
```

Q5 — Revenue 2024 vs 2025

```
SELECT COALESCE(a.coffee_name, b.coffee_name) AS coffee,

COALESCE(a.total_2024, 0) AS sales_2024,

COALESCE(b.total_2025, 0) AS sales_2025

FROM (SELECT coffee_name, SUM(money::NUMERIC) AS total_2024 FROM index_1 GROUP BY coffee_name) a

FULL JOIN (SELECT coffee_name, SUM(money::NUMERIC) AS total_2025 FROM index_2 GROUP BY coffee_name) b

ON a.coffee_name = b.coffee_name ORDER BY sales_2024 DESC LIMIT 10;
```

Q6 — Payment Methods

```
SELECT COALESCE(a.cash_type, b.cash_type) AS cash_type,

COALESCE(a.total_2024, 0) AS count_2024,
```

```
COALESCE(b.total_2025, 0) AS count_2025

FROM (SELECT cash_type, COUNT(*) AS total_2024 FROM index_1 GROUP BY cash_type) a

FULL JOIN (SELECT cash_type, COUNT(*) AS total_2025 FROM index_2 GROUP BY cash_type) b

ON a.cash_type = b.cash_type;
```

Q7 — Monthly Comparison

```
SELECT COALESCE(a.month, b.month) AS month,

COALESCE(a.total_2024, 0) AS sales_2024,

COALESCE(b.total_2025, 0) AS sales_2025

FROM (SELECT TO_CHAR(TO_DATE(date,'YYYY-MM-DD'),'YYYY-MM') AS month,

SUM(money::NUMERIC) AS total_2024 FROM index_1 GROUP BY month) a

FULL JOIN (SELECT TO_CHAR(TO_DATE(date,'YYYY-MM-DD'),'YYYY-MM') AS month,

SUM(money::NUMERIC) AS total_2025 FROM index_2 GROUP BY month) b

ON a.month = b.month ORDER BY month;
```

Q8 — Daily Sales

```
SELECT date, COUNT(*) AS total_orders, SUM(money) AS total_revenue

FROM coffee_sales GROUP BY date ORDER BY date;
```

Q9 — Most Popular Item

```
SELECT coffee_name, COUNT(*) AS total_orders

FROM coffee_sales GROUP BY coffee_name ORDER BY total_orders DESC;
```

Q10 — Revenue Summary

```
SELECT COUNT(*) AS total_orders, SUM(money) AS total_revenue,

MIN(money) AS min_order, MAX(money) AS max_order, AVG(money) AS avg_order

FROM coffee_sales;
```