Car Sales Performance & Customer Insights Analysis

Portfolio Project Report

(Tools Used: Excel • SQL • Power BI)

---

1. Introduction

This project analyses a mid-sized automotive dealership's performance using three integrated datasets: Cars, Customers, and Sales.
The goal is to understand the factors that drive sales performance, identify key customer segments, and determine which car models generate the highest revenue.

The analysis demonstrates the use of Excel for data cleaning, SQL for analytical querying, and Power BI for interactive visualisation, showcasing full end-to-end data analytics capability.

---

2. Objectives

The project aims to:

1. Clean and prepare the datasets for accurate analysis.
2. Explore sales performance trends and revenue drivers.
3. Profile customers based on demographics and purchase behaviour.
4. Evaluate inventory status and car characteristics affecting sales.
5. Build a Power BI dashboard that presents insights in a clear and actionable way.

---

3. Data Sources

The datasets were downloaded from Kaggle under Apache 2.0 Licence and include:

Cars.csv

- Car ID, brand, model, year
- Engine type, transmission
- Price, quantity in stock
- Status (available, reserved, sold)

Customers.csv

- Customer ID, name, gender
- Age, phone, email
- City

Sales.csv

- Sale ID, sale date
- Customer ID, car ID
- Quantity, sale price
- Salesperson, payment method

---

4. Data Cleaning Process

Data cleaning was performed in Excel and later verified in SQL and Power BI.

Steps Performed:

4.1 Removing Duplicates

All three datasets were checked and cleaned to ensure unique entries for cars, customers, and sales.

4.2 Correcting Data Types

- Sale price and price → converted to numeric
- Sale date → converted to Date format
- Customer age → converted to integer

4.3 Validation Checks

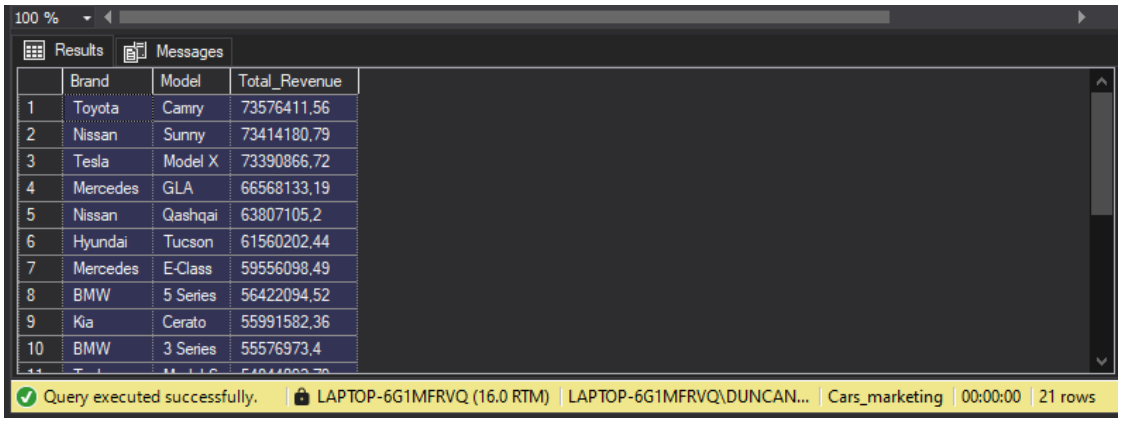
- Ensured every Car_ID in Sales exists in Cars table
- Ensured every Customer_ID matches Customers table

This ensured the dataset was clean, consistent, and analysis ready.

5. SQL Analysis

SQL was used to derive deeper insights.
Key queries included:

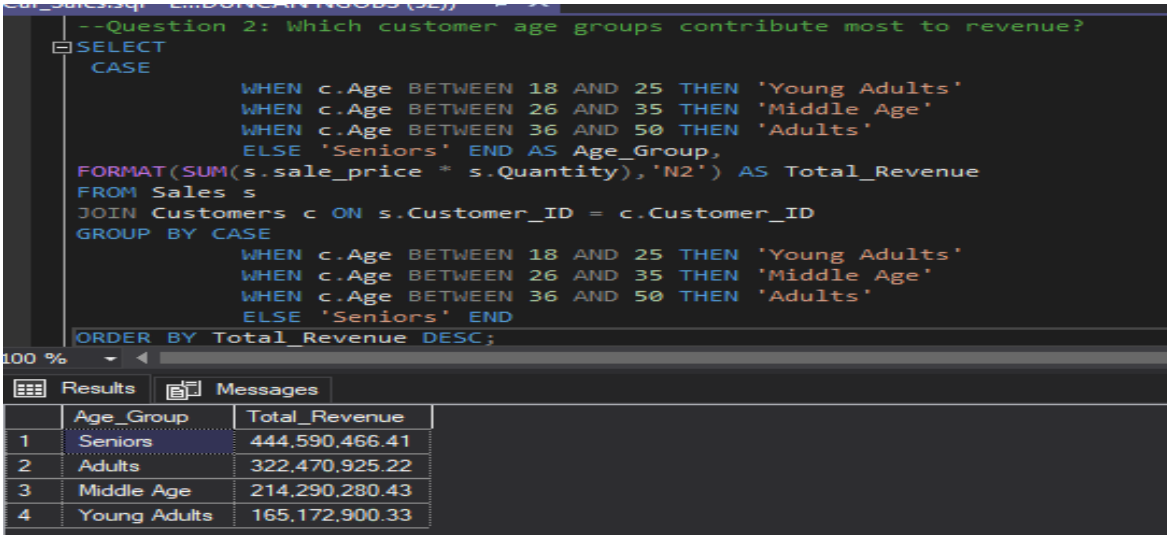
Which car models have generated the highest total revenue?



	Brand	Model	Total_Revenue
1	Toyota	Camry	73576411,56
2	Nissan	Sunny	73414180,79
3	Tesla	Model X	73390866,72
4	Mercedes	GLA	66568133,19
5	Nissan	Qashqai	63807105,2
6	Hyundai	Tucson	61560202,44
7	Mercedes	E-Class	59556098,49
8	BMW	5 Series	56422094,52
9	Kia	Cerato	55991582,36
10	BMW	3 Series	55576973,4

Query executed successfully. | LAPTOP-6G1MFRVQ (16.0 RTM) | LAPTOP-6G1MFRVQ\DUNCAN... | Cars_marketing | 00:00:00 | 21 rows

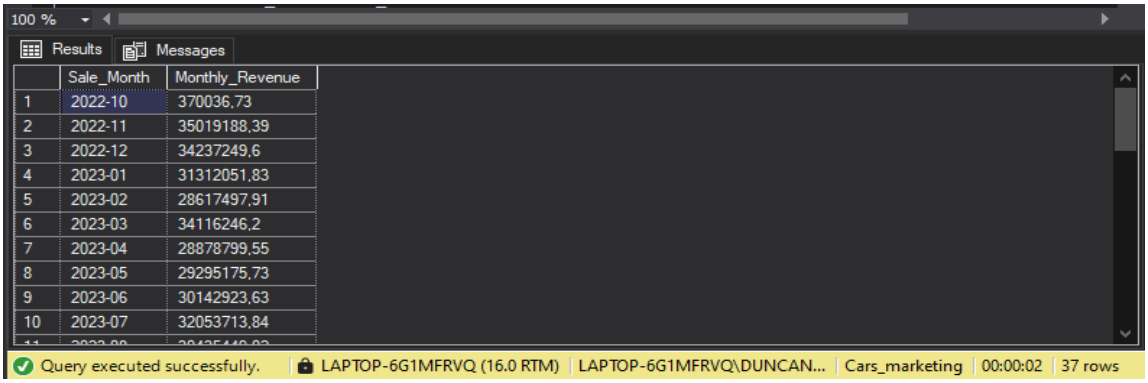
Which customer age groups contribute most to revenue?



```
--Question 2: Which customer age groups contribute most to revenue?
SELECT
 CASE
            WHEN c.Age BETWEEN 18 AND 25 THEN 'Young Adults'
            WHEN c.Age BETWEEN 26 AND 35 THEN 'Middle Age'
            WHEN c.Age BETWEEN 36 AND 50 THEN 'Adults'
            ELSE 'Seniors' END AS Age_Group,
FORMAT(SUM(s.sale_price * s.Quantity),'N2') AS Total_Revenue
FROM Sales s
JOIN Customers c ON s.Customer_ID = c.Customer_ID
GROUP BY CASE
            WHEN c.Age BETWEEN 18 AND 25 THEN 'Young Adults'
            WHEN c.Age BETWEEN 26 AND 35 THEN 'Middle Age'
            WHEN c.Age BETWEEN 36 AND 50 THEN 'Adults'
            ELSE 'Seniors' END
ORDER BY Total_Revenue DESC;
```

	Age_Group	Total_Revenue
1	Seniors	444,590,466.41
2	Adults	322,470,925.22
3	Middle Age	214,290,280.43
4	Young Adults	165,172,900.33

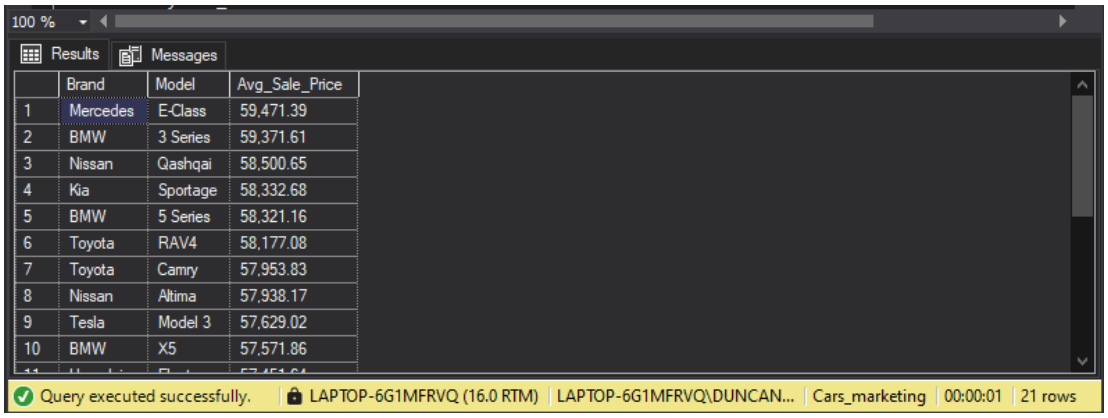
How has revenue changed month by month?



	Sale_Month	Monthly_Revenue
1	2022-10	370036,73
2	2022-11	35019188,39
3	2022-12	34237249,6
4	2023-01	31312051,83
5	2023-02	28617497,91
6	2023-03	34116246,2
7	2023-04	28878799,55
8	2023-05	29295175,73
9	2023-06	30142923,63
10	2023-07	32053713,84

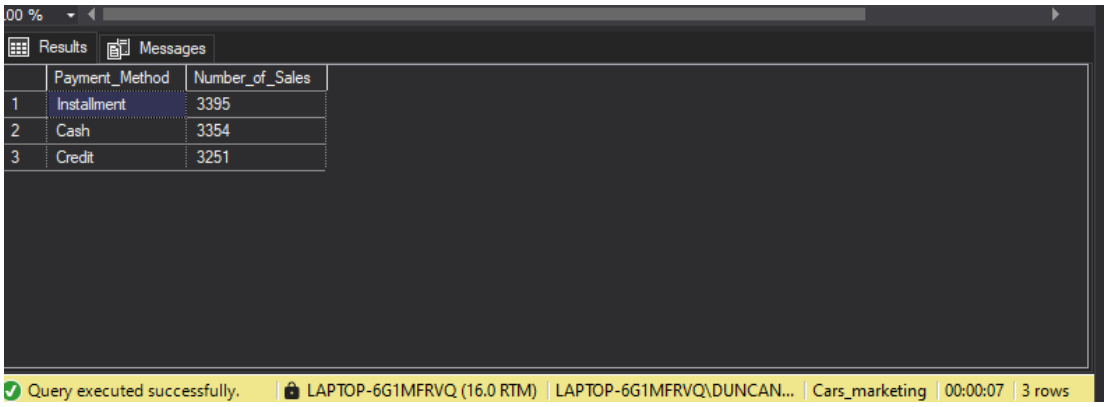
Query executed successfully. | LAPTOP-6G1MFRVQ (16.0 RTM) | LAPTOP-6G1MFRVQ\DUNCAN... | Cars_marketing | 00:00:02 | 37 rows

What is the average sale price per car model?



	Brand	Model	Avg_Sale_Price
1	Mercedes	E-Class	59,471.39
2	BMW	3 Series	59,371.61
3	Nissan	Qashqai	58,500.65
4	Kia	Sportage	58,332.68
5	BMW	5 Series	58,321.16
6	Toyota	RAV4	58,177.08
7	Toyota	Camry	57,953.83
8	Nissan	Altima	57,938.17
9	Tesla	Model 3	57,629.02
10	BMW	X5	57,571.86

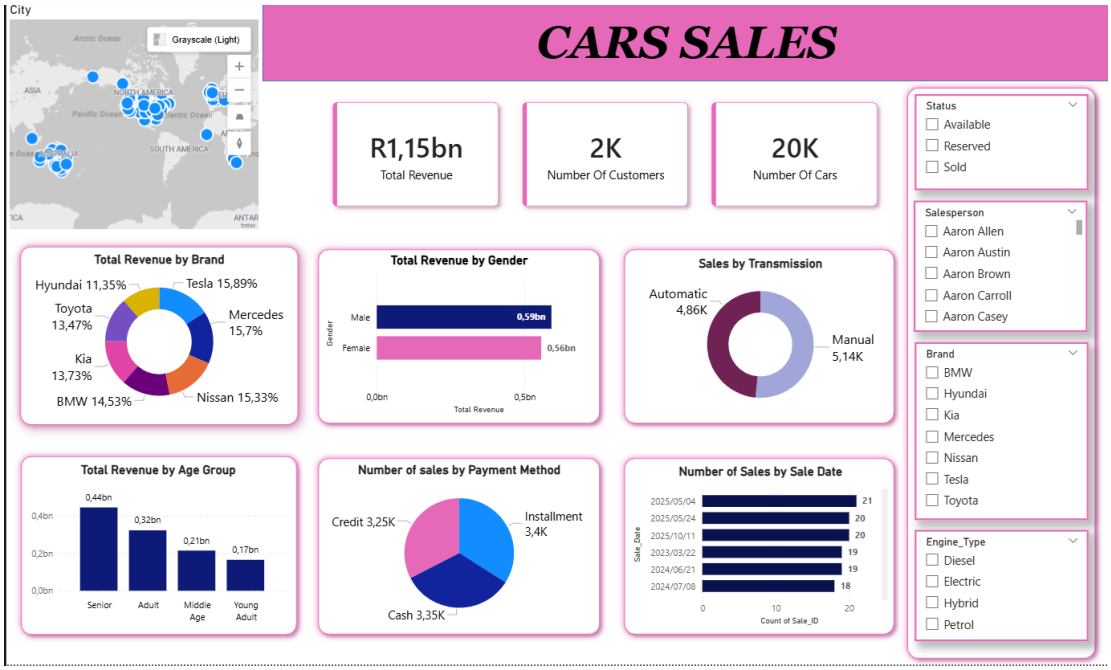
Which payment method do customers mostly prefer?



	Payment_Method	Number_of_Sales
1	Installment	3395
2	Cash	3354
3	Credit	3251

6. Power BI Dashboard

A full interactive dashboard was developed:



6.1 Sales Performance

- Revenue by salesperson
- Units sold by brand
- Payment method breakdown
- Sales by engine type

6.2 Customer Insights

- Revenue by age group
- Map of revenue by city
- Gender spending analysis
- Slicers for demographic segmentation

6.3 Inventory & Car Insights

- Cars in stock by brand
- Slow-moving inventory

---

7. Key Findings (Insights)

1. Tesla and Mercedes generate the highest revenue, driven by strong sales of Model X and GLA respectively.
2. Sales peak between July and October, showing seasonal demand trends.
3. Customers over age 51 are the largest buying segment, contributing the most revenue.
4. Payment methods are dominated by Instalment, suggesting financing is important for customers.
5. Several models show high stock but low sales, indicating potential over-ordering or poor demand.
6. Certain salespeople consistently outperform others, contributing disproportionately to revenue.

---

8. Tools Used

- **Excel:** Data cleaning, labelling, validation, and preparation
- **SQL:** Analytical queries, data exploration, relational integrity checks
- **Power BI:** Dashboard building, modelling, DAX measures, storytelling

---

9. Conclusion

This project demonstrates the full data analysis workflow from cleaning raw data to generating insights and building a professional dashboard.
It highlights my ability to:

- Clean and transform complex datasets
- Write efficient SQL queries
- Build interactive Power BI dashboards
- Communicate insights clearly and visually
- Combine business understanding with technical skill

This showcases end-to-end analytical capabilities suitable for roles in data analytics, business intelligence, and data science.