Axon Sales Report

Abstract:

This document outlines a capstone project addressing Axon's sales data management challenges through a Business Intelligence (BI) solution using Microsoft PowerBI and SQL. It begins with data extraction and cleaning from a MySQL database, followed by PowerBI dashboard design for interactive data visualization and analysis. SQL is utilized for advanced analytics to derive actionable insights. Rigorous testing ensures solution reliability before deployment to the management team. The project emphasizes user-friendliness and comprehensive documentation for successful adoption. By leveraging BI tools, Axon gains the ability to make informed decisions, driving operational efficiency and business growth in the competitive classic car retail market.

Introduction:

In the realm of classic vehicle sales, Axon faces the intricate task of managing and deciphering sales data across a diverse range of offerings, including cars, airplanes, trains, buses, trucks, and bikes. This complexity underscores the critical need for a comprehensive Business Intelligence (BI) solution to illuminate sales insights and inform strategic decision-making. Hence, this capstone project embarks on implementing a robust BI framework, leveraging Microsoft PowerBI and SQL, tailored to Axon's multifaceted product portfolio.

Within this endeavor, the significance of PowerBI is paramount. Its versatile tools and intuitive interface empower Axon to transform raw data into actionable insights across various vehicle categories. By harnessing PowerBI's visualization capabilities, Axon gains real-time visibility into sales performance, customer preferences, and market trends across their entire product spectrum. This not only enhances operational efficiency but also fosters proactive decision-making, enabling Axon to adapt swiftly to market shifts and capitalize on emerging opportunities.

In essence, PowerBI serves as the linchpin of Axon's BI solution, facilitating comprehensive analysis and strategic agility across their diverse range of

offerings, ultimately positioning them as a dynamic force in the classic vehicle retail market.

Problem statement:

1. **Data Integration Challenge**: Axon struggles with integrating sales data from its MySQL database into PowerBI, hindering the creation of a comprehensive view of sales performance across various product categories.
2. **Data Quality and Preparation Hurdle**: The process of cleaning and transforming data for analysis poses a significant challenge, impacting the accuracy and reliability of insights derived from the data.
3. **Lack of Insightful Reporting**: The absence of interactive dashboards and reports in PowerBI inhibits the sales team and management from gaining actionable insights and making informed decisions based on sales data.
4. **Limited Analytical Capabilities**: Axon lacks the ability to perform advanced analytics using SQL, limiting its capacity to extract meaningful insights that could drive sales improvement initiatives and strategic decision-making.

Objective

To develop and deploy a Business Intelligence (BI) solution utilizing PowerBI and SQL, aimed at enhancing Axon's sales data management and analysis capabilities for informed decision-making and improved business performance.

Project Approach

1. **Database Creation**: Executed SQL script in MySQL Workbench to establish the "classic models" database, housing the sales data.
2. **Integration with PowerBI**: Utilized PowerBI's SQL connector to seamlessly integrate the "classic models" database, enabling direct access to the sales data for analysis.
3. **Data Cleaning in Power Query**: Leveraged Power Query within PowerBI to perform comprehensive data cleaning, ensuring accuracy and consistency for subsequent analysis.

4. **Data Modeling and DAX Calculations**: Employed data modeling techniques and DAX calculations within PowerBI to create measured and calculated columns, enriching the dataset with insightful metrics.
5. **Data Visualization**: Utilized PowerBI's robust visualization capabilities to create compelling dashboards and reports, facilitating intuitive exploration of sales data.
6. **SQL Insights**: Leveraged SQL queries to extract additional insights, complementing the analysis conducted within PowerBI and providing deeper understanding of the sales data.

Tools and Technologies

Microsoft PowerBI:

- PowerBI is a robust business intelligence (BI) tool known for its intuitive interface and powerful data visualization capabilities. It enables users to import, clean, analyze, and visualize data from various sources, facilitating informed decision-making.
- In this project, PowerBI serves as the primary BI tool, allowing for the creation of interactive dashboards and reports to analyze Axon's sales data effectively.

MySQL:

- MySQL is a popular relational database management system (RDBMS) that utilizes SQL (Structured Query Language) for managing and manipulating data. It provides a reliable platform for storing and retrieving structured data efficiently.
- For this project, MySQL is utilized for data storage and management. SQL queries are employed to extract, transform, and analyze sales data, enabling advanced analytics and insights to improve Axon's sales strategies.

Database Description:

The MySQL sample database schema comprises 8 tables:

1. **Customers**: Contains customer data including name, address, and order details.

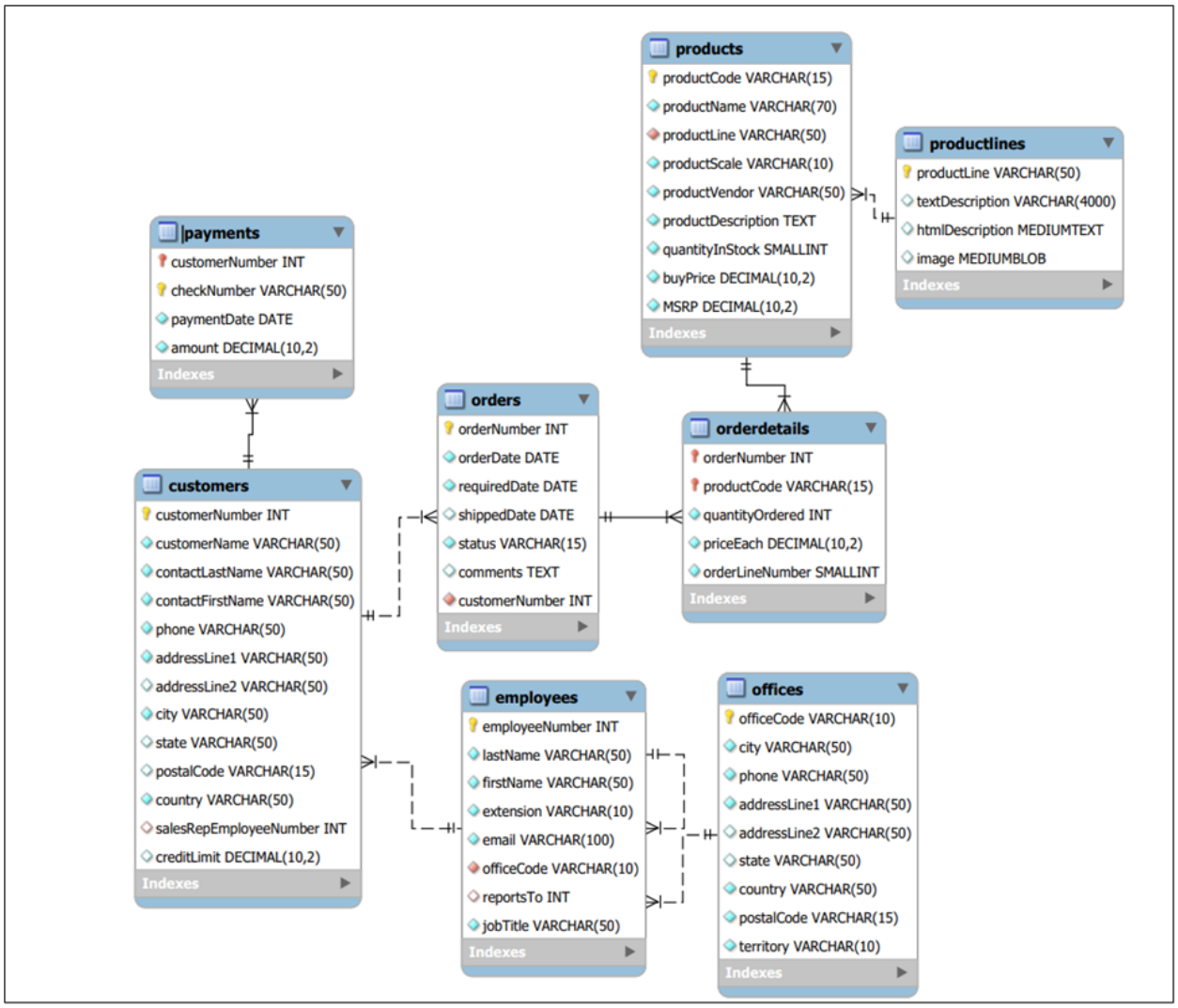
2. **Products**: Lists various vehicles such as classic cars, trains, motorbikes, buses, trucks, and airplanes, along with attributes like name, cost price, MRSP, vendors, and stock quantity.
3. **ProductLines**: Stores categories of product lines.
4. **Orders**: Records sales orders placed by customers.
5. **OrderDetails**: Stores line items for each sales order, including quantity ordered, delivery date requested by customers, shipment date by the company, and price of each item.
6. **Payments**: Tracks payments made by customers based on their accounts.
7. **Employees**: Contains employee information and organizational structure, specifying reporting relationships.
8. **Offices**: Stores data related to sales offices.

These tables collectively house typical business data crucial for sales analysis and decision-making processes within the organization.

Project Implementation

Data extraction:

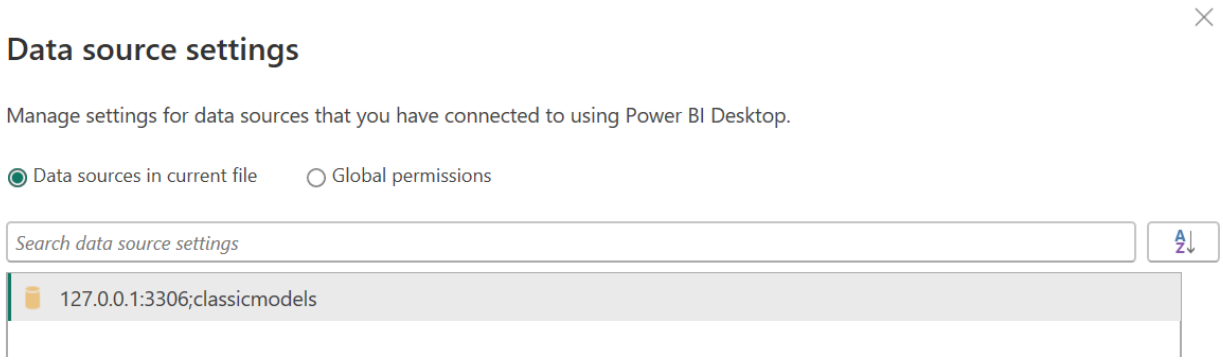
Database is creating by run sql script of the data. As given in data description, database has eight tables customers, office, orders, orderdetails, employees, payments, products and productlines. Order table is acting as a fact table. Orders table acts as a centre for all other table. Here schema is snow flake scheme. Order details and customers are directly connected to the orders table, so they are acting as dimension table. Remaining tables like payments, employees, offices, products and productlines are subdimension tables thus forming the snowflake schema. Ordernumber, customer number, employee number, product code, office code are primary key for customers, employees, products, offices table respectively. Totally there are 23 employees, 122 customers, 110 products and 326 orders



Entity relationship diagram (ERD) for the Axon sales database

Loading Data into PowerBI Using Direct Query

PowerBI offers versatile data connectivity options, including importing, direct query, and live connection, facilitating seamless integration with various data sources such as files, databases, and web services. In this project, I opted for the direct query method to connect PowerBI directly to the SQL database using a connector. This approach enables real-time access to the data stored in the SQL database, ensuring that PowerBI always reflects the latest information without the need for data duplication or storage within the PowerBI environment. This title encapsulates the approach used to load data into PowerBI, emphasizing the utilization of direct query for dynamic data retrieval and analysis.



Data Cleaning: Ensuring Data Integrity and Quality

Data cleaning is a critical step in the data analysis process, aimed at enhancing data integrity and reliability. In this project, I meticulously examined the dataset using Power Query within PowerBI to ensure its quality and accuracy. The analysis revealed that the dataset exhibited no errors, missing values, or inconsistencies. Additionally, thorough scrutiny confirmed that the column names were appropriate, and the data types assigned to each column were correct.

Furthermore, beyond the absence of errors and missing values, data cleaning often involves several other important tasks, including:

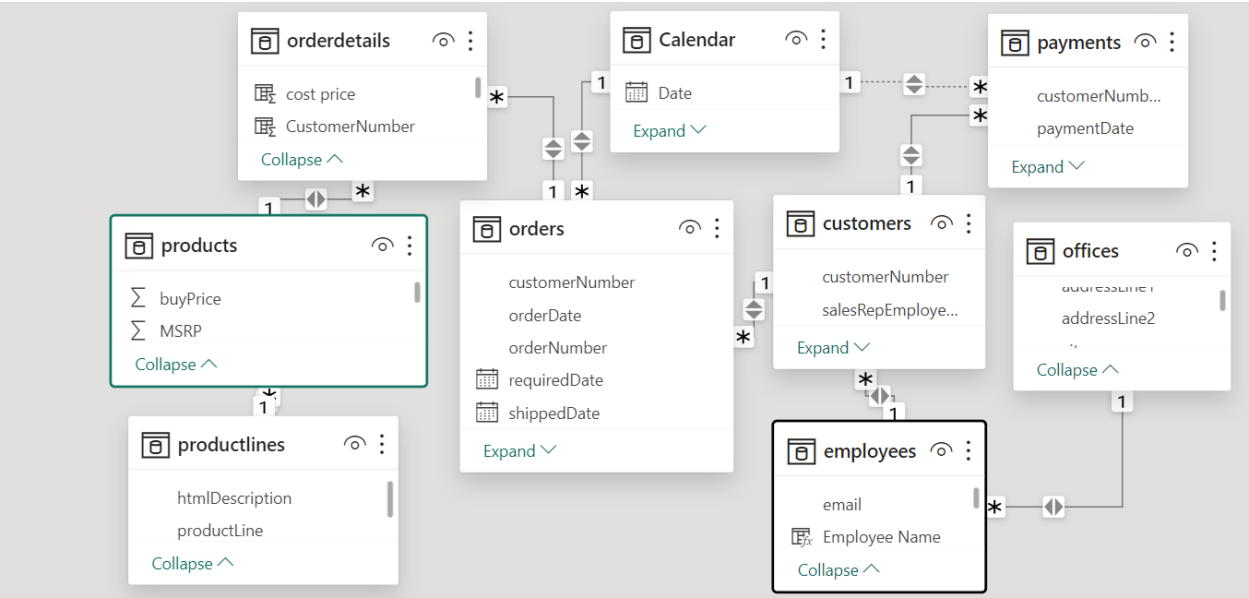
1. **Handling Duplicates**: Identifying and removing duplicate records to prevent data redundancy and maintain data accuracy.
2. **Standardizing Data Formats**: Ensuring uniformity in data formats and units across the dataset to facilitate accurate analysis and interpretation.
3. **Addressing Outliers**: Detecting and addressing outliers or anomalous data points that may skew analysis results or misrepresent trends.
4. **Dealing with Inconsistencies**: Resolving any inconsistencies or discrepancies in the data, such as formatting inconsistencies or contradictory information.
5. **Imputing Missing Values**: Imputing missing values using appropriate techniques, such as mean imputation or interpolation, to preserve the integrity of the dataset while mitigating data loss.

Although the initial analysis indicated that the dataset was clean and error-free, it is essential to remain vigilant and address any potential data quality issues that

may arise during subsequent stages of analysis. By prioritizing data cleanliness and quality assurance, we ensure the reliability and accuracy of our insights and conclusions derived from the data.

Data modeling:

Data modeling in Power BI involves structuring and organizing data to facilitate analysis and visualization. It includes defining relationships between tables, specifying cardinality (one-to-one, one-to-many, or many-to-many), and determining the direction of relations (single or both). In our dataset, all relationships are one-to-many, reflecting the snowflake model's hierarchical structure. By setting relation direction as both, we ensure efficient data flow during visualization, enabling seamless exploration of insights across interconnected tables. This approach optimizes the analysis process, enhancing our ability to derive actionable insights and make informed decisions based on the data.



Leveraging DAX Formulas for Comprehensive Analysis

In the data analysis process, DAX (Data Analysis Expressions) formulas played a pivotal role. Measures were meticulously crafted to extract essential metrics such as total revenue, total orders, and customer count, providing invaluable insights into sales performance.

👉 cost price = RELATED(products[buyPrice])

For calculated columns, DAX functions including RELATED and fundamental mathematical operations were employed to accurately compute profit margins and other pertinent indicators. Furthermore, the creation of a calculated table for the calendar was imperative, ensuring streamlined visualization by harmonizing disparate dates within the orders and payments data.

👉 Calendar = CALENDAR(min(orders[orderDate]),
max(payments[paymentDate]))

Noteworthy is the utilization of the 'USE relationship' function to establish a secondary relationship between the calendar and payments table, enabling precise calculation of total payments.

👉 Total Payment = CALCULATE(sum(payments[amount]),
USERELATIONSHIP(payments[paymentDate], 'Calendar'[Date]))

Data visualization:

Harnessing the capabilities of Power BI, I transformed raw data into actionable insights through engaging dashboards and reports. Using a diverse range of visualization tools such as Donut charts, treemaps, line graphs, and gauge meters, I provided comprehensive views of categorical data contributions, sales trends over time, and comparative analyses of total sales and payment settlements. Additionally, I utilized tables, metrics, and bar charts to highlight top-performing

customers, products, and employees. Geographic maps allowed for a global perspective on customer distribution, complemented by slicers and filters for dynamic data exploration. With Power BI's dynamic and interactive visualizations, I facilitated effective communication of findings, trend identification, and data-driven decision-making, driving business success.

Key Insight from PowerBI dashboard:

1. **Financial Overview:** Axon's total revenue stands at $8.9 million, with a profit of $3.59 million, representing an impressive overall profit margin of 40%. Notably, sales exhibit a significant peak in the fourth quarter of each year, particularly in November and December.
2. **Product Performance:** Cars emerge as the top-selling product category, contributing to over 50% of total revenue and profit. Among these, the Ferrari stands out as the most popular car.
3. **Global Revenue Drivers:** The USA, France, and Spain emerge as the top countries in terms of revenue and profit generation for Axon, indicating lucrative markets for further expansion.
4. **Customer Analysis:** Axon identifies 24 inactive customers, with 10 located in Germany. Addressing the needs and expectations of these customers could potentially revitalize sales in this segment.
5. **Workforce Composition:** With a total of 23 employees, the majority (17) are sales representatives, underscoring the emphasis on sales-driven strategies within the company.
6. **Customer Insights:** Euro+ Shopping Channel stands out as the top customer from Spain among Axon's 122 customers.
7. **Seasonal Sales Trends:** Sales consistently peak in the fourth quarter, particularly during November and December. Leveraging marketing efforts during these months could further enhance sales performance.

These insights provide a comprehensive understanding of Axon's sales dynamics, highlighting areas of strength and opportunities for improvement to inform strategic decision-making and drive sustainable growth.

MySql analysis:

MySQL is a potent tool for project file analysis, offering robust SQL capabilities. By importing and structuring data, users can extract valuable insights through querying and advanced analytics. I utilized MySQL to delve into and analyze sales data from Axon, conducting thorough examination and deriving insights to inform decision-making.

▼ Q1: What is the average time it takes for customers to make payments after placing an order?

```
CREATE TEMPORARY TABLE payment_time AS (
WITH cte AS (
SELECT o.customernumber, o.orderdate,
SUM(od.priceeach * od.quantityordered) AS total_price
FROM orders o
JOIN orderdetails od USING (ordernumber)
WHERE status = 'shipped'
GROUP BY o.ordernumber
)
SELECT c.customernumber, p.paymentdate, c.orderdate,
DATEDIFF(p.paymentdate, c.orderdate) AS
timetaken_to_pay, c.total_price
FROM cte c
JOIN payments p
ON (c.customernumber = p.customernumber)
AND (c.total_price = p.amount)
);
```

This code generates a temporary table containing the payment settlement time for each order. It enables us to calculate the average payment date for individual customers and also for all customers collectively.

```
SELECT customernumber, AVG(timetaken_to_pay) AS
average_payment_settlement_time
FROM payment_time
GROUP BY customernumber
ORDER BY average_payment_settlement_time ASC;
```

This code calculates the average time taken by each customer.

Output

customernumber	average_payment_Settlement_time
357	-167.5000
201	-141.5000
385	4.6667
177	5.5000
455	8.5000
486	9.0000
256	9.5000
456	10.0000
320	10.3333
382	11.0000
362	11.0000
347	11.0000
447	11.3333
475	11.5000

The average payment time for most customers ranges between 4.6 days (minimum) and 35 days (maximum). Interestingly, two customers have negative payment times, indicating they paid in advance for their orders.

▼ Q2: Which customers have placed orders for the greatest variety of unique products?

```
WITH cte AS (
SELECT ordernumber, productcode, productname,
productline, customerNumber, quantityordered
FROM orders o
JOIN orderdetails od USING (ordernumber)
JOIN products USING (productcode)
WHERE status = 'shipped'
)
SELECT customerNumber,
COUNT(DISTINCT ordernumber) AS no_of_orders,
COUNT(DISTINCT productline) AS no_of_productlines,
COUNT(DISTINCT productcode) AS no_of_products,
SUM(quantityordered) AS total_quantity
FROM cte
GROUP BY customerNumber
ORDER BY no_of_productlines DESC, no_of_products DESC,
total_quantity DESC;
```

Output:

	customerNumber	no_of_orders	no_of_productline	no_of_products	total_quantity
▶	141	22	7	87	7544
	353	5	7	37	1433
	398	4	7	32	1150
	124	16	6	77	6291
	496	3	6	39	1359
	386	3	6	39	1280
	148	5	6	37	1524
	320	3	6	34	1140
	282	3	6	31	1069

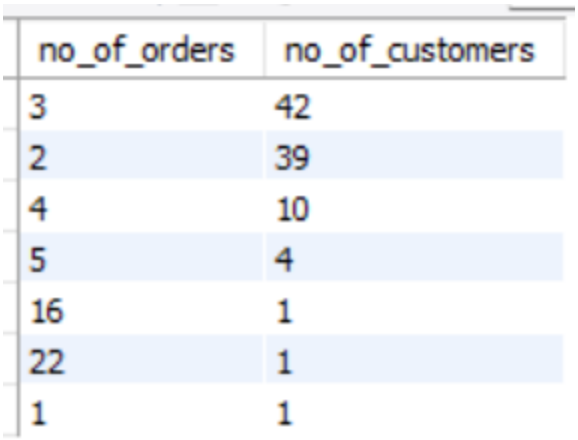
Customer number 141, known as Euro+ Shopping Channel from Spain, stands out as a significant contributor to sales analysis. They have demonstrated exceptional engagement by purchasing a diverse range of 87 unique products across all 7 product lines. This customer's extensive buying behavior showcases a deep interest in various offerings. Remarkably, despite the total unique product count being 110, this customer has managed to explore a significant portion of the available inventory. Furthermore, their substantial order quantity of 7544 units solidifies their position as one of the most valuable customers. This detailed analysis underscores the importance of customer segmentation and understanding individual purchasing patterns in driving sales growth.

▼ Q3: What is the ordering frequency for each customer?

```
SELECT no_of_orders, COUNT(customernumber) AS
no_of_customers
FROM (
SELECT customernumber, COUNT(ordernumber) AS
no_of_orders
FROM orders
WHERE status = 'shipped'
GROUP BY customernumber
) AS ordertable
```

```
GROUP BY no_of_orders
ORDER BY no_of_customers DESC;
```

Output:



no_of_orders	no_of_customers
3	42
2	39
4	10
5	4
16	1
22	1
1	1

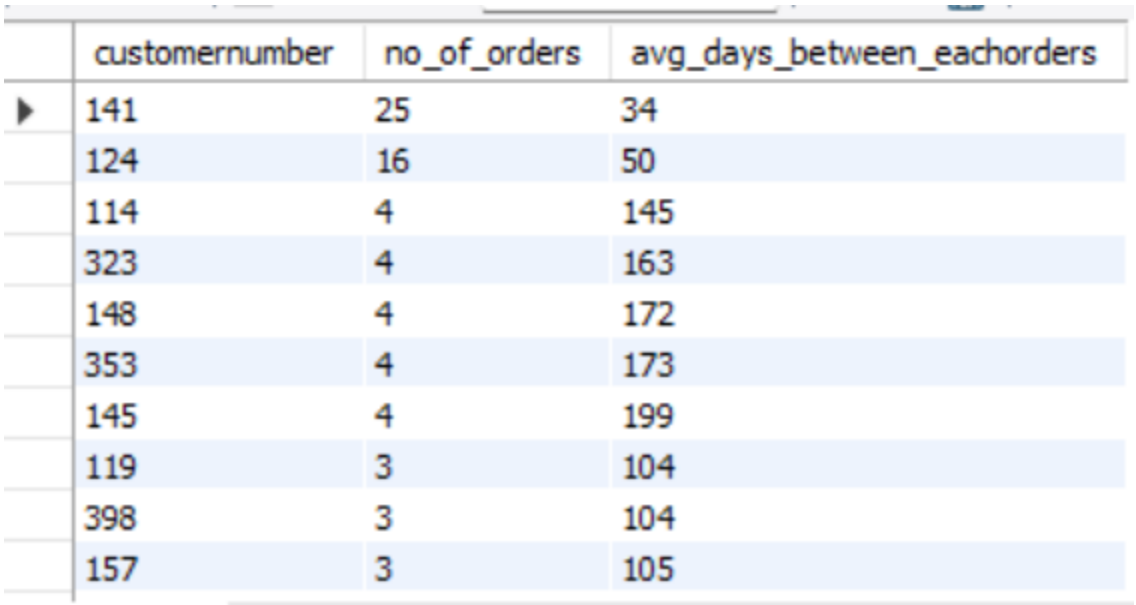
Among the 98 active customers recorded in Axon sales data from June 2003 to May 2005, notable patterns emerge: 42 customers placed 3 shipped orders, while 39 placed 2. This consistent engagement signals a loyal customer base and underscores the potential for tailored marketing initiatives to further cultivate customer loyalty and drive sales growth.

▼ Q4: what is the average number of days between each order for these customers?

```
WITH cte AS (
   SELECT customernumber, ordernumber, orderdate,
        LAG(orderdate) OVER (PARTITION BY
customernumber ORDER BY orderdate) AS lagdate
   FROM orders
),
cte1 AS (
```

```
    SELECT customernumber, ordernumber, orderdate,
lagdate,
        DATEDIFF(orderdate, lagdate) AS timediff
    FROM cte
)
SELECT customernumber, COUNT(ordernumber) AS
no_of_orders,
      AVG(timediff) AS avg_days_between_each_orders
FROM cte1
WHERE timediff IS NOT NULL
GROUP BY customernumber
ORDER BY no_of_orders DESC,
avg_days_between_each_orders;
```

Output:



customernumber	no_of_orders	avg_days_between_eachorders
141	25	34
124	16	50
114	4	145
323	4	163
148	4	172
353	4	173
145	4	199
119	3	104
398	3	104
157	3	105

Customer numbers 141 and 124 stand out for their significant order volume, with 25 and 16 orders respectively. On average, they place orders

approximately every 34 and 50 days, showcasing their consistent engagement and purchasing frequency.

Summary:

The implementation of a Business Intelligence (BI) solution utilizing PowerBI and SQL has provided Axon with actionable insights into their classic vehicle sales data. Analysis revealed consistent customer engagement, with notable customers placing multiple orders over time. These insights inform targeted marketing strategies and customer retention efforts. PowerBI's visualization tools facilitated dynamic exploration of sales performance, while SQL queries provided deeper insights. Overall, the BI solution enhances Axon's strategic decision-making capabilities, positioning them to capitalize on market opportunities and drive sustainable growth in the classic vehicle retail market.