SALES ANALYSIS USING RELATIONAL DATABASE MANAGEMENT SYSTEM

INTRODUCTION

MySQL is a robust rational database management system used for data analysis, organizing data into tables and ensuring integrity and scalability. SQL queries facilitate efficient data retrieval and manipulation, utilizing functions like SUM, AVG, COUNT, and clauses such as GROUP BY and WHERE for filtering and aggregating data. Advanced features, including joins, subqueries, and window functions, support complex analyses.

The object of this report is centered on analyzing sales performance, focusing on operational efficiency, and overall profitability of a car repair shop. The data used for this analysis were ingested from csv files into a MySQL database and processed using various SQL queries to extract valuable insights. And proposing value added recommendations.

METHDOLOGY

Data Cleaning and Preparation

Data Cleaning: Python codes was used to sort columns in the datasets, by corelating the expected data types, ensure all foreign keys in the datasets corresponded to valid primary keys in their respective tables. Also observed discrepancies in the invoice data and recalculated the columns to ensure accuracy.

Database Setup: Created necessary tables (customer, job, vehicle, part, invoices) with appropriate schemas using MySQL, ensuring columns created in each table are in alignment with the csv data

Data Ingestion: Imported cleaned datasets into the respective database tables. This involved using MySQL script to load the data from CSV files into the database, ensuring that the data was correctly mapped to the corresponding columns.

SQL Queries for Analysis
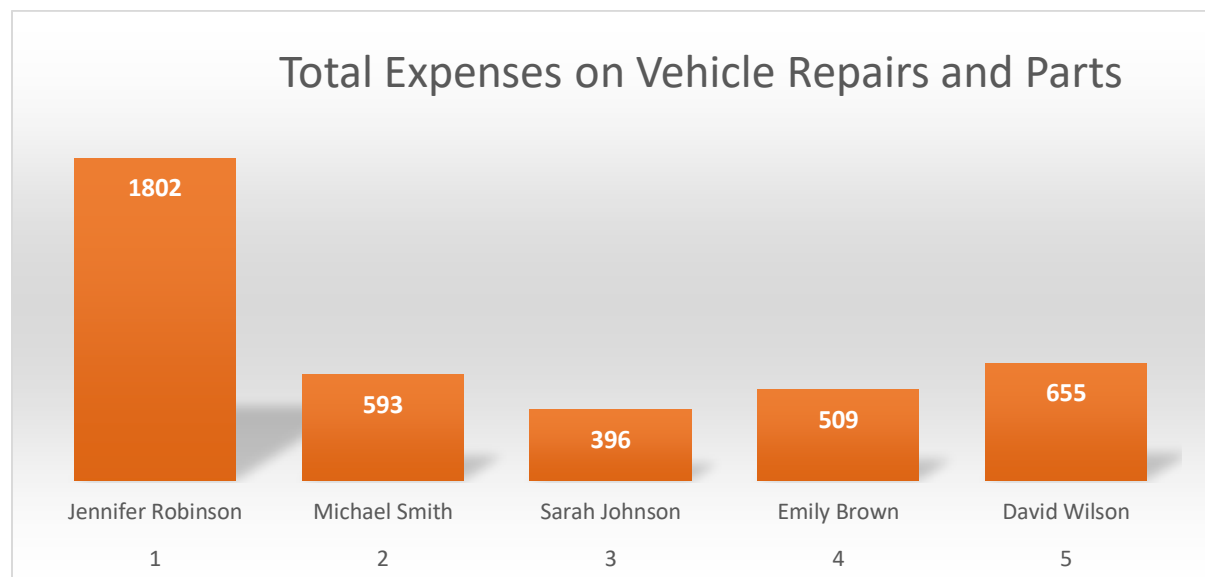
Customer Analysis

Top 5 customers who have spent the most on vehicle repairs and parts:

This involves aggregating the total spending for each customer by summing the labor costs, parts costs, and sales tax. The results shows that the highest total expenditure is 1802 by Jennifer Robinson. This comprehensive total is essential for accurately identifying the shop's most valuable customers and understanding their full financial contribution.

```
52
53        #Top 5 customer who have spent the most on vehicle repairs and parts
54  •    SELECT c.customer_id, c.Name, SUM(i.Total) AS TotalSpent
55        FROM customer c
56        JOIN invoices i ON c.customer_id = i.CustomerID
57        GROUP BY c.customer_id, c.Name
58        ORDER BY TotalSpent DESC
59        LIMIT 5;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

customer_id	Name	TotalSpent
1	Jennifer Robinson	1802
5	David Wilson	655
2	Michael Smith	593
4	Emily Brown	509
3	Sarah Johnson	396

Visualization

Total Expenses on Vehicle Repairs and Parts

	Jennifer Robinson	Michael Smith	Sarah Johnson	Emily Brown	David Wilson
Value	1802	593	396	509	655
	1	2	3	4	5

Average Spending of Customers on Repairs and Parts.

This involves calculating the total in the invoice table. This metric helps gauge the typical expenditure of customers and can be used to benchmark individual customer spending against the average. The average expenditure is said to be 791,0000

```
59      LIMIT 5;
60
61      #Determine the average spending of customers on repairs and parts
62 •    SELECT AVG(i.Total) AS Average_Spending
63      FROM invoices i;
64
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Average_Spending
791.0000

Frequency of customers Visits and identify any Pattern

The frequency of customer visits is calculated by counting the number of invoices associated with each customer. This analysis highlights the most frequent visitors to the shop. it was found that each customer visited the car repair shop only once, the pattern noticed is that there is a lack of repeat visits from customers. Each customer has exactly one invoice, indicating they have only visited the shop once during the data collection period.

```
65      #Frequency of customer visits
66 •    SELECT c.customer_id, c.Name, COUNT(i.InvoiceID) AS Visit_Count
67      FROM customer c
68      JOIN invoices i ON c.customer_id = i.CustomerID
69      GROUP BY c.customer_id, c.Name
70      ORDER BY Visit_Count DESC;
71
72      #Average milage of vehicle serviced
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

customer_id	Name	Visit_Count
1	Jennifer Robinson	1
2	Michael Smith	1
3	Sarah Johnson	1
4	Emily Brown	1
5	David Wilson	1

Vehicle Analysis

Average Mileage of Vehicle Serviced

The average mileage of vehicles serviced at the shop is calculated. This provides insights into the typical wear and tear of vehicles brought in for repairs, which can inform maintenance recommendations and service offerings.

```
72      #Average milage of vehicle serviced
73 ●    SELECT AVG(v.Mileage) AS Average_Mileage
74      FROM vehicle v;
75
76      #The most common vehicle makes and models brought in for service
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Average_Mileage
33299.0000

Most Common Vehicle Makes and Models Brought in For Service

The most common makes and models of vehicles serviced were identified by counting the occurrences of each make and model in the vehicle data. It is shown that all vehicle makes and models were brought in once, which limit understanding the shop's expertise and reputation with specific vehicle brands.

```
76      #The most common vehicle makes and models brought in for service
77 ●    SELECT v.Make, v.Model, COUNT(*) AS Count
78      FROM vehicle v
79      GROUP BY v.Make, v.Model
80      ORDER BY Count DESC;
81
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Make	Model	Count
BMW	X5	1
Toyota	Corolla	1
Honda	Civic	1
Ford	Escape	1
Chevrolet	Malibu	1

Distribution of Vehicle Ages and Identifying Any Trends in Service Based on Vehicle Age

The distribution of vehicle ages was analyzed by calculating the age of each vehicle based on its year of manufacture and grouping them into age brackets. This helps in identifying trends in service requirements based on vehicle age, such as the types of services commonly needed for older over newer vehicles. From the result shown, it is observed that each make and model combination in the output has a count of 1, indicating that for each specific age group, there is only one vehicle of each make and model. This suggests that each make and model in the dataset is unique for the respective age group shown

```
#distribution of vehicle ages and identify any trends in service requirements based on vehicle age:
SELECT YEAR(CURDATE()) - v.Year AS VehicleAge, COUNT(*) AS Count
FROM vehicle v
GROUP BY VehicleAge, v.Make, v.Model
ORDER BY VehicleAge;
```

VehicleAge	Make	Model	Count
12	BMW	X5	1
9	Toyota	Corolla	1
8	Chevrolet	Malibu	1
6	Honda	Civic	1
4	Ford	Escape	1

Job Performance Analysis

Most Common Types of Job Performed and Their Frequency

The most common types of jobs performed were identified by counting the occurrences of each job description. This analysis provides insights into the primary services offered by the shop and helps in resource planning and inventory management.

```
87
88     #the most common types of jobs performed and their frequency
89  •   SELECT j.Description, COUNT(*) AS Count
90     FROM job j
91     GROUP BY j.Description
92     ORDER BY Count DESC;
93
```

esult Grid | ▦ | ↻ Filter Rows: [   ] | Export: ▦ | Wrap Cell Content: ↕A

Description	Count
Diagnose front wheel vibration	1
Replace front CV Axel	1
Balance tires	1
Oil change	1
Replace brake pads	1

Total Revenue Generated from Each Type of Job

The total revenue generated from each type of job was calculated by summing the amounts associated with each job description. This helps in identifying the most profitable services and can guide decisions on service offerings and pricing strategies. From the result show, it can be observed that Replace front CV Axel generated the highest revenue with a total revenue of $438

```
94     #the total revenue generated from each type of job
95  •   SELECT j.Description,
96         GROUP_CONCAT(j.JobID) AS JobIDs,
97         SUM(j.Amount) AS TotalRevenue
98     FROM job j
99     GROUP BY j.Description
L00    ORDER BY TotalRevenue DESC;
```
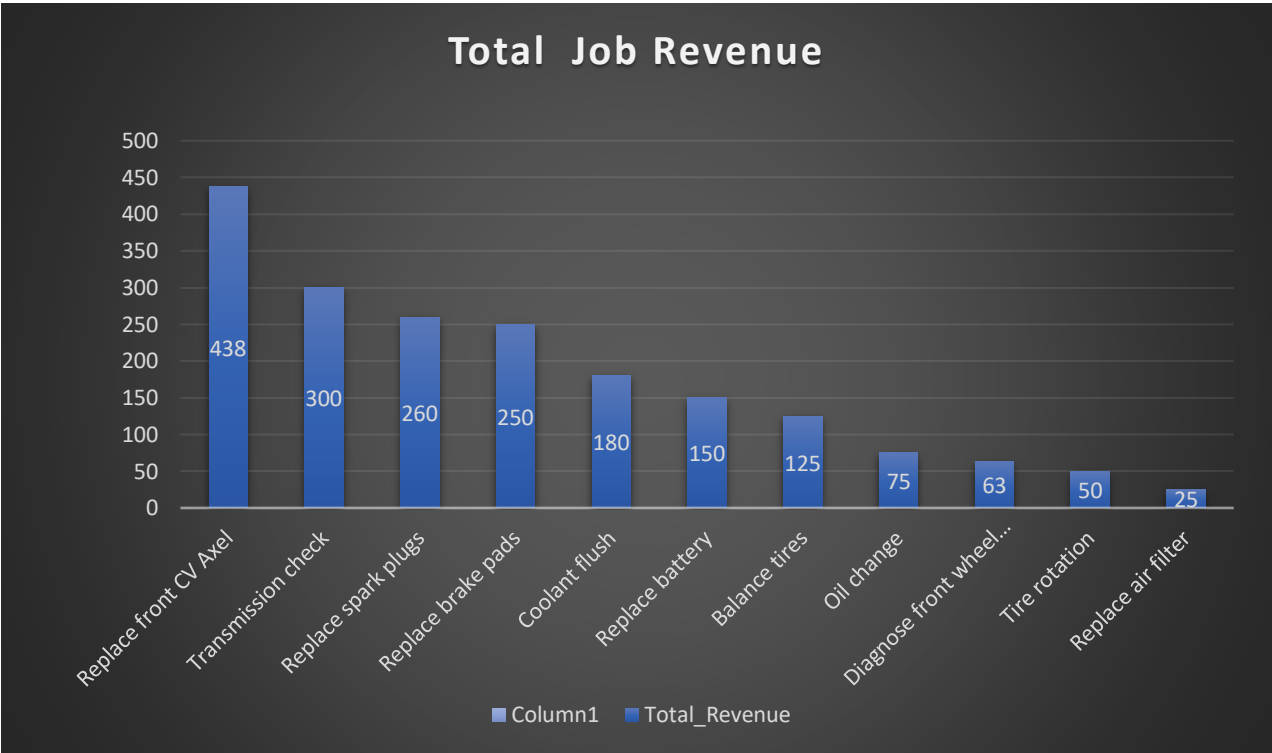
Result Grid | ▦ | ↻ Filter Rows: [   ] | Export: ▦ | Wrap Cell Content: ↕A

Description	JobIDs	TotalRevenue
Replace front CV Axel	2	438
Transmission check	8	300
Replace spark plugs	11	260
Replace brake pads	5	250	260
Coolant flush	10	180

esult 12 ⌄

Total Job Revenue

Job With Highest and Lowest Average Costs

The jobs with the highest and lowest average costs were determined by calculating the average amount for each job description. This analysis provides insights into the cost structure of different services and helps in pricing and cost management.

```
101
102     #the jobs with the highest average costs
103  •  SELECT j.Description, AVG(j.Amount) AS Highest_AverageCost
104     FROM job j
105     GROUP BY j.Description
106     ORDER BY Highest_AverageCost DESC
107     LIMIT 1;
108
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Description	Highest_AverageCost
Replace front CV Axel	438.0000

```
L09     #the jobs with the lowest average costs
L10 •   SELECT j.Description, AVG(j.Amount) AS lowest_AverageCost
L11     FROM job j
L12     GROUP BY j.Description
L13     ORDER BY lowest_AverageCost ASC
L14     LIMIT 1;
L15
L16     #the top 5 most frequently used parts and their total usage
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: A

Description	lowest_AverageCost
Replace air filter	25.0000

Part Usage Analysis

Top 5 most frequently used parts and their total usage

The most frequently used parts were identified by counting the occurrences of each part name. This analysis helps in inventory management, ensuring that the most commonly used parts are adequately stocked. From the table shown, the total usage of all part were used once.

```
L15
L16     #the top 5 most frequently used parts and their total usage
L17 •   SELECT p.PartName, COUNT(*) AS UsageCount
L18     FROM part p
L19     GROUP BY p.PartName
L20     ORDER BY UsageCount DESC
L21     LIMIT 5;
```
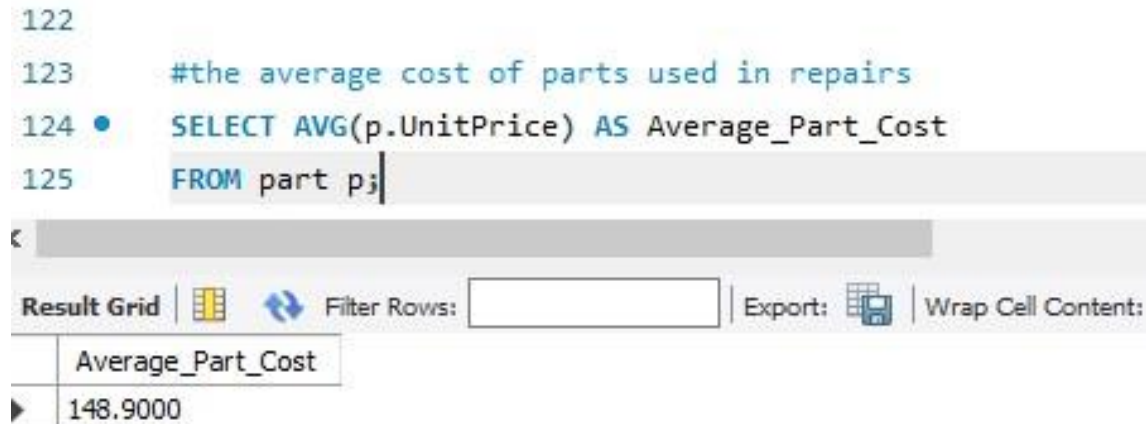
Result Grid | Filter Rows: | Export: | Wrap Cell Content: A

PartName	UsageCount
CV Axel	1
Shop Materials	1
Wheel Weights	1
Brake Pads	1
Battery	1

Average Cost of Parts

The average cost of parts used in repairs was calculated to understand the typical expenditure on parts. This metric helps in cost management and pricing strategies for parts.

```
122
123       #the average cost of parts used in repairs
124 •     SELECT AVG(p.UnitPrice) AS Average_Part_Cost
125       FROM part p;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Average_Part_Cost
148.9000

Total Revenue from Parts Sales

The total revenue generated from parts sales was calculated by summing the amounts associated with each part. This analysis provides insights into the contribution of parts sales to the overall revenue and helps in strategic planning for parts inventory and sales.

```
:7      #the total revenue generated from parts sales
:8 ●    SELECT SUM(p.Amount) AS Total_Part_Revenue
:9      FROM part p;
:0
:1      #the total revenue generated from labor and parts for each m
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: IA

Total_Part_Revenue
1585

Financial Analysis

Monthly Revenue from Labor and Parts

The total revenue generated from labor and parts for each month was calculated by aggregating the amounts from invoices, jobs, and parts. This analysis helps in understanding the shop's monthly performance and identifying seasonal trends in revenue.

```
31      #the total revenue generated from labor and parts for each month
32 ●    SELECT DATE_FORMAT(i.InvoiceDate, '%Y-%m') AS Month,
33          SUM(j.Amount) AS Total_Labour_Revenue,
34          SUM(p.Amount) AS Total_Part_Revenue
35      FROM invoices i
36      JOIN job j ON i.InvoiceID = j.InvoiceID
37      JOIN part p ON j.JobID = p.JobID
38      GROUP BY Month
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: IA

Month	Total_Labour_Revenue	Total_Part_Revenue
2023-09	2216	1585

Visualization

Overall Profitability

The overall profitability of the repair shop is to be calculated by the difference between total revenue and total costs. However, from the invoice table, total cost was not provided, hence total profitability can also be equivalent to summation of total labor and total parts.

Impact of Sales Tax on Revenue

The impact of sales tax on the total revenue was analyzed by calculating the total sales tax collected and the revenue excluding sales tax. This analysis helps in understanding the contribution of sales tax to the total revenue and its impact on profitability which can be significant to the company income.
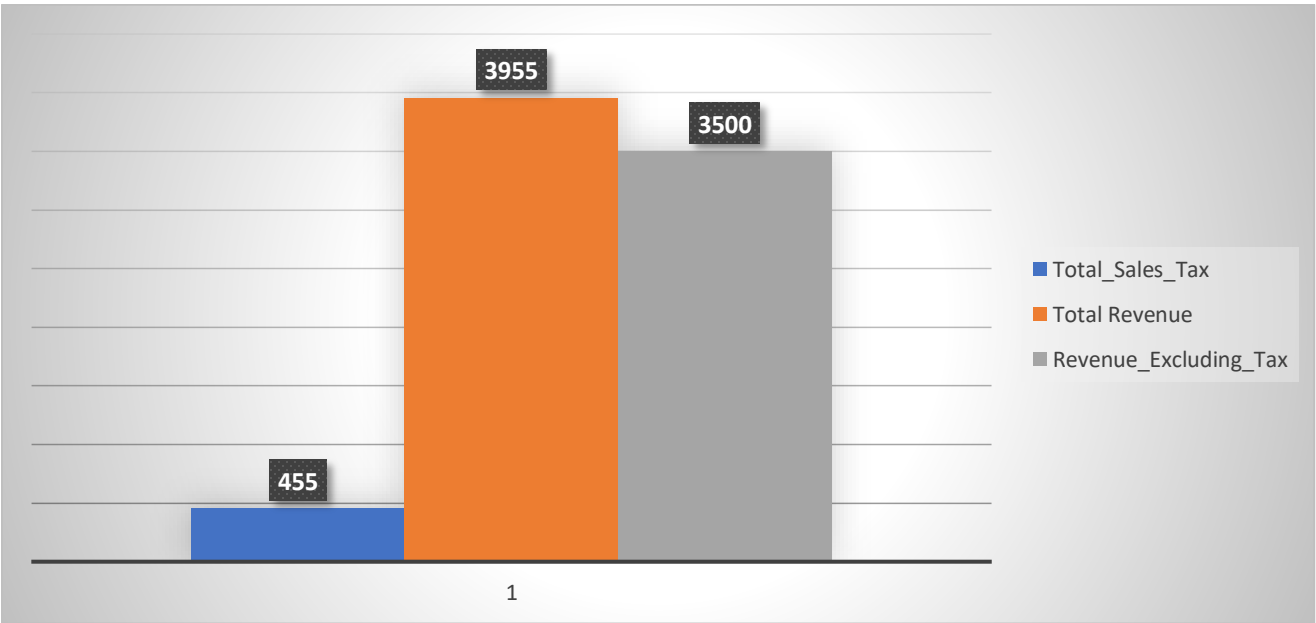
```
144
145     #the impact of sales tax on the total revenue
146 •   SELECT SUM(i.SalesTax) AS Total_Sales_Tax, SUM(i.Total) AS Total_Revenue,
147           (SUM(i.Total) - SUM(i.SalesTax)) AS Revenue_Excluding_Tax
148     FROM invoices i;
```

Total_Sales_Tax	Total_Revenue	Revenue_Excluding_Tax
455	3955	3500

Visualization

Optimization Recommendations

Identifying Underperforming Services

Based on the analysis of the provided job data table, it is clear that certain services are underperforming. To enhance operational efficiency and profitability, specific recommendations have been identified. The analysis highlights that the service of replacing the air filter is the least performing service, generating a total revenue of only $25. This indicates a need for a strategic reassessment to optimize this service.

Recommendations

1. **Evaluate Operations and Pricing Strategy**: The current demand and pricing strategy for the air filter replacement service should be thoroughly evaluated. This involves understanding customer needs and preferences, as well as benchmarking against competitors. Adjusting the price to better reflect the value and cost of the service may help increase its appeal and profitability.

2. **Offer as Part of a Package Deal**: To enhance the attractiveness of the air filter replacement service, it should be considered for inclusion in a comprehensive maintenance package. Bundling this service with other routine car maintenance tasks, such as oil changes or tire rotations, can provide added value to customers. This approach can drive higher customer engagement and increase overall service uptake.

Effective Stock Management: Effective stock management is crucial for ensuring that the most frequently used parts are readily available, minimizing downtime and improving service efficiency.

Recommendations

Based on the frequency of use, the following parts should be prioritized for higher stock levels:

1. **Spark Plugs:** Due to their high quantity usage, spark plugs should be kept in higher stock to ensure availability for frequent replacements.

2. **Wheel Weights:** Similarly, wheel weights have a high total quantity used and should be stocked in larger quantities to meet demand.

Proposing Customer Loyalty Programs for Top Spending Customer: Implementing customer loyalty programs can significantly enhance customer retention and increase repeat business.

Recommendations

1. **Exclusive Discounts and Offers**: Provide top-spending customers with exclusive discounts on services and parts. This can include percentage discounts, fixed-amount discounts, or buy-one-get-one-free offers.

2. **Personalized Communication**: Use the provided customer information to send personalized communication, such as birthday greetings, service reminders, and special offers tailored to their preferences and history.

3. **Referral Bonuses**: Encourage top-spending customers to refer friends and family by offering referral bonuses, such as discounts on future services for both the referrer and the referred customer.

Deliverables

Relational Database containing imported data from Csv files

https://drive.google.com/drive/folders/1qPH-AFxQwdH9g4aaeBnkLwAOjNl1Gfd2?usp=sharing

SQL Script used to create table and import files

https://drive.google.com/file/d/15rqg7FV_qgVSV9KIICFb35MWotRRlyZW/view?usp=sharing