**Designing a Dimensional Model for Sales Analysis and SQL Queries**

**Part A - Model Creation**

#### **Table of Contents**

* Introduction
* Business Requirements
* Identification of Facts and Dimensions
* Dimensional Model Design
  + Fact Table
  + Dimension Tables
* Entity-Relationship (ER) Diagram
* Model Documentation
* Conclusion

**Introduction**

The creation of a dimensional model for evaluating a car repair shop's sales performance using sample invoices is shown in this analysis. The work entails building a solid and rational framework that enables the company to better understand its operations, streamline its workflow, and boost profitability. This dimensional model's main goal is to offer a thorough framework for examining several facets of sales, such as consumer purchasing, car maintenance trends, work performance, part usage, and financial data.

Sales analysis is crucial for the car repair shop as it enables the identification of top-spending customers, common vehicle issues, and the effectiveness of different services and parts. By leveraging this model, the shop can enhance customer satisfaction, streamline inventory management, and make data-driven decisions to boost overall business performance. This report details the creation of the dimensional model using MySQL Workbench, emphasising the logical business view of the data necessary for thorough sales analysis.

**Business Requirements**

The car repair shop operates across multiple locations in western Canada, catering to a diverse customer base with varying vehicle service needs. To optimise its operations and enhance its services, the shop requires a comprehensive analysis of its sales performance. This analysis will help the business understand customer behaviour, vehicle service trends, and the financial impact of its services and parts sales.

Key business questions that the dimensional model aims to answer include:

* Who are the top-spending customers and what are their spending patterns?
* Which vehicle brands, models, and years are most commonly serviced?
* What are the most frequent and profitable types of services and parts?
* How do sales vary across different shop locations?
* What is the overall financial performance of the shop, including monthly revenue and profitability?

The types of analysis needed to answer these questions involve:

Sales analysis by customer, including spending and visit frequency.

* Analysis of vehicle service patterns by brand, model, and year.
* Performance evaluation of various services and parts.
* Sales comparison across different shop locations.
* Financial analysis of revenue from labour and parts, and the impact of sales tax.

By addressing these requirements, the dimensional model will provide the car repair shop with actionable insights to improve its operational efficiency, enhance customer satisfaction, and increase profitability.

**Identification of Facts and Dimensions**

Based on the sample invoice and the task requirements, we designed a dimensional model for the car repair shop's sales analysis. We created the schema, differentiated between fact and dimension tables, created the ER diagram in MySQL Workbench.

**Fact vs. Dimension Tables:**

* Fact tables contain quantitative data (measurements or metrics) about business processes. In addition to numeric measures, a fact table always contains foreign keys for each of its associated dimensions, as well as optional degenerate dimension keys and date/time stamps. Fact tables are the primary target of computations and dynamic aggregations arising from queries.
* Dimension tables contain descriptive attributes that provide context for the facts.Every dimension table has a single primary key column. This primary key is embedded as a foreign key in any associated fact table where the dimension row’s descriptive context is exactly correct for that fact table row. Dimension tables are usually wide, ﬂat denormalized tables with many low-cardinality text attributes.

**Database Schema:**

1. Fact Table: Sales

* SalesID (Primary Key)
* CustomerID (Foreign Key)
* VehicleID (Foreign Key)
* LocationID (Foreign Key)
* DateID (Foreign Key)
* TotalLabor
* TotalParts
* SalesTax
* TotalAmount

2. Dimension Tables:

a. DimCustomer

* CustomerID (Primary Key)
* Name
* Address
* PhoneNumber

b. DimVehicle

* VehicleID (Primary Key)
* Make
* Model
* Year
* Colour
* VIN
* RegistrationNumber
* Mileage

c. DimLocation

* LocationID (Primary Key)
* LocationName
* Address

d. DimDate

* DateID (Primary Key)
* Date
* Day
* Month
* Year

e. DimService

* ServiceID (Primary Key)
* Description
* Rate

f. DimPart

* PartID (Primary Key)
* PartNumber
* PartName
* UnitPrice

g. FactServiceDetail

* ServiceDetailID (Primary Key)
* SalesID (Foreign Key)
* ServiceID (Foreign Key)
* Hours
* Amount

h. FactPartDetail

* PartDetailID (Primary Key)
* SalesID (Foreign Key)
* PartID (Foreign Key)
* Quantity
* Amount

**Explanation of the Schema:**

1. The Sales fact table captures the core transaction data, linking to various dimensions.

2. DimCustomer stores customer information.

3. DimVehicle contains details about the serviced vehicles.

4. DimLocation represents different shop locations.

5. DimDate allows for time-based analysis.

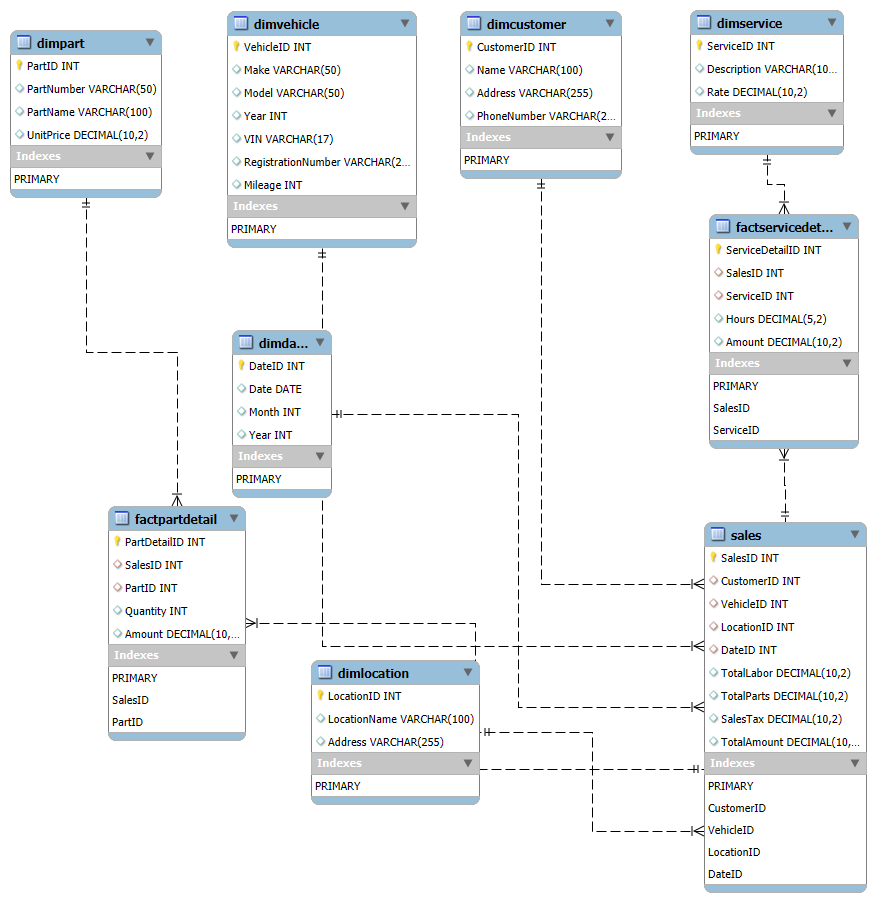
6. DimService and DimPart store information about services and parts respectively.

7. FactServiceDetail and FactPartDetail are additional fact tables that provide detailed breakdowns of services and parts for each sale.

This schema allows for flexible analysis of sales by customer, vehicle, location, time, services, and parts, meeting the requirements specified in the task.

**Entity-Relationship (ER) Diagram**

ER Diagram illustrating relationships between fact and dimension tables:

****

The diagram visually represents the structure of the database, highlighting how different entities (tables) are connected and interact with each other.

Key elements in the ER diagram include the fact table, which captures quantitative sales data such as the Total Labour, Sales Tax etc, and multiple dimension tables, which store qualitative attributes for customers, vehicles, services, parts, locations, and dates. Each table is defined with primary keys to uniquely identify records within the table.

Primary keys are essential for ensuring the uniqueness of each record in a table. In the fact table, foreign keys link to the primary keys in the dimension tables, establishing relationships between the tables. These foreign key relationships ensure that each fact (such as a sales transaction) can be associated with the relevant dimensions (such as customer details, vehicle information, and service type).

Referential integrity between the tables is maintained by these primary and foreign key relationships. This integrity ensures that the data remains consistent and accurate, preventing orphan records and ensuring that each reference in the fact table corresponds to an existing record in the related dimension tables.

The ER diagram is a critical component of the dimensional model, providing a clear and organised view of the data relationships. It serves as a blueprint for the database structure, facilitating efficient data retrieval and analysis for the car repair shop’s sales performance.

**Model Documentation**

Below are the descriptions for each table and column, along with explanations for the decisions made based on the sales receipt.

1. DimCustomer

* CustomerID: Unique identifier for each customer
* Name: Customer's full name
* Address: Customer's mailing address
* PhoneNumber: Customer's contact number

Explanation: This table captures essential customer information from the invoice. It allows for customer-specific analysis and segmentation.

2. DimVehicle

* VehicleID: Unique identifier for each vehicle
* Make: Brand of the vehicle (e.g., BMW)
* Model: Specific model of the vehicle (e.g., X5)
* Year: Manufacturing year of the vehicle
* Colour: Colour of the vehicle
* VIN: Vehicle Identification Number
* RegistrationNumber: Vehicle's registration number
* Mileage: Current mileage of the vehicle

Explanation: This table stores detailed vehicle information, enabling analysis based on vehicle characteristics. The VIN and registration number allow for unique vehicle identification.

3. DimLocation

* LocationID: Unique identifier for each repair shop location
* LocationName: Name of the repair shop (e.g., Latino Garage Winnipeg North)
* Address: Full address of the repair shop

Explanation: Although only one location is shown on the invoice, this dimension allows for future expansion to multiple locations across western Canada.

4. DimDate

* DateID: Unique identifier for each date
* Date: Actual date of the transaction
* Month: Month of the transaction
* Year: Year of the transaction

Explanation: This dimension enables time-based analysis, allowing for tracking trends over different time periods.

5. DimService

* ServiceID: Unique identifier for each service type
* Description: Description of the service (e.g., "Replace front CV Axel")
* Rate: Standard rate charged for the service

Explanation: This table captures the various services offered, allowing for analysis of service popularity and profitability.

6. DimPart

* PartID: Unique identifier for each part
* PartNumber: Manufacturer's part number
* PartName: Description of the part
* UnitPrice: Standard price per unit of the part

Explanation: This dimension stores information about parts used in repairs, enabling inventory analysis and parts sales tracking.

7. Sales (Fact Table)

* SalesID: Unique identifier for each sales transaction
* CustomerID, VehicleID, LocationID, DateID: Foreign keys to respective dimensions
* TotalLabor: Total amount charged for labour
* TotalParts: Total amount charged for parts
* SalesTax: Sales tax amount
* TotalAmount: Total invoice amount

Explanation: This fact table represents the core transaction, linking all relevant dimensions and storing key financial metrics.

8. FactServiceDetail

* ServiceDetailID: Unique identifier for each service line item
* SalesID: Foreign key to the Sales fact table
* ServiceID: Foreign key to the DimService table
* Hours: Number of hours billed for the service
* Amount: Total amount charged for this service

Explanation: This table provides a detailed breakdown of services for each sale, allowing for granular analysis of service performance.

9. FactPartDetail

* PartDetailID: Unique identifier for each part line item
* SalesID: Foreign key to the Sales fact table
* PartID: Foreign key to the DimPart table
* Quantity: Number of units of the part used
* Amount: Total amount charged for this part

Explanation: Similar to FactServiceDetail, this table offers a detailed view of parts used in each transaction.

**Logical Explanations for Table Creation Decisions:**

1. Separation of Dimensions: By creating separate dimension tables for customers, vehicles, locations, dates, services, and parts, we enable flexible and efficient analysis across these different aspects of the business.

2. Granular Fact Tables: The main Sales fact table captures the overall transaction, while FactServiceDetail and FactPartDetail provide more granular information. This structure allows for both high-level and detailed analysis without overly complicating the main fact table.

3. Date Dimension: Although the invoice only shows a single date, creating a separate date dimension allows for easy time-based analysis and reporting.

4. Location Dimension: Despite having only one location on the invoice, this dimension was included to support the requirement of analysing performance across multiple locations in western Canada.

5. Exclusion of Certain Details: Some details from the invoice, such as the "COLOUR" of the vehicle, were omitted as they were deemed less relevant for sales analysis. However, these could be added if found to be important for specific analyses.

6. Use of Surrogate Keys: Each table uses an auto-incrementing ID as its primary key, which simplifies relationships and improves query performance compared to using natural keys.

This structure provides a balance between capturing the necessary detail from the sales receipt and creating a flexible, performance-oriented model for sales analysis across various dimensions of the business.

**Conclusion**

In summary, the design of the dimensional model for the car repair shop’s sales analysis involved a thorough examination of the sample invoice data to identify relevant facts and dimensions. The process included creating a fact table to capture key sales metrics and several dimension tables to allow for detailed analysis of customer information, vehicle details, service types, parts, locations, and transaction dates. The resulting model provides a logical and structured view of the business data, tailored to address the specific analytical needs of the shop.

This dimensional model effectively meets the business requirements by enabling the shop to answer critical questions about customer spending patterns, vehicle service trends, service and parts performance, and financial metrics across different locations. The model supports flexible and detailed analysis, helping the shop to gain deeper insights into its operations.

The potential benefits for the car repair shop’s sales analysis are significant. With this model, the shop can identify top-spending customers, understand common vehicle issues, and evaluate the profitability of various services and parts. This information will help the shop optimise its inventory management, improve customer satisfaction through targeted marketing and service offerings, and make data-driven decisions to enhance overall business performance.

**Appendices**

Sample invoice: [salesreceipt.png](https://drive.google.com/file/d/1enihc1uFFW9dHPYFiBC27wpk92QodBC-/view)

Query for Model Creation: [Model Creation](https://drive.google.com/file/d/1rdgG4vND7EHPxF0oFjxrOWOeAbvikhfV/view?usp=sharing)

**References**

[Dimensional Modelling Techniques](https://www.kimballgroup.com/data-warehouse-business-intelligence-resources/kimball-techniques/dimensional-modeling-techniques/)

**Designing a Dimensional Model for Sales Analysis and SQL Queries**

**Data Analysis of Car Repair Shop’s Operations**

**Part B - SQL Queries and Insights**

#### **Table of Contents**

* Introduction
* Data Ingestion and Preparation
* Customer Analysis
* Vehicle Analysis
* Job Performance Analysis
* Parts Usage Analysis
* Financial Analysis
* Optimization Recommendations
* Conclusion

**Introduction**

In this report, we present the data analysis of the car repair shop’s operations, focusing on insights derived from an extended set of sales receipts. The task involves importing CSV files into a relational database, cleaning and preparing the data, and performing detailed SQL queries to analyse various aspects of the shop's performance.

The purpose of this data analysis is to extract meaningful insights that can help the car repair shop optimise its operations, improve customer satisfaction, and increase profitability. By examining customer spending patterns, vehicle service trends, job performance, parts usage, and financial metrics, we aim to provide actionable recommendations for the business.

Deriving insights from the car repair shop’s data is crucial for making informed decisions. Understanding customer behaviour, identifying common vehicle issues, and evaluating the effectiveness of services and parts can lead to better inventory management, targeted marketing efforts, and enhanced operational efficiency. This analysis will enable the shop to make data-driven decisions that drive growth and profitability.

**Data Ingestion and Preparation**

In this section, we outline the process of cleaning, properly formatting, and indexing the provided CSV files for efficient querying, As Well as the importation/imputation of the cleaned files into our relational database.

The process began with cleaning and formatting the provided files. This included adding more tables, standardising data formats (e.g., Number formats and text casing), and handling missing data, also making sure foreign keys in the fact tables were properly linked to the dimension tables.

Nulls must be avoided in the fact table’s foreign keys because these nulls would automatically cause a referential integrity violation. Rather than a null foreign key, the associated dimension table must have a default row (and surrogate key) representing the unknown or not applicable condition. Our dimensional location table only has one entry and it is the location of the car repair shop provided from the sample invoice, as no location dimension was given the original CSV files and in order to maintain referential integrity, all sales entries were assumed to have happened at the same location. These steps ensured that the data was accurate and ready for importation/ imputation.

Once the data was cleaned we attempted to import our data into our database but ran into `--secure-file-priv` issues and after much trouble shooting none of the solutions were feasible. Hence we populated our relational database with data from the CSV files using simple insert queries as our data was not large. This was achieved using SQL scripts to create the necessary tables and populate them with data from the CSV files.

To optimise the database for efficient querying, indexing strategies were implemented. Indexes were auto incremented on key columns such as CustomerID, Vehicle ID, ServiceID, ParTID, DateID etc.. These indexes improved the speed of data retrieval, allowing for faster and more efficient execution of SQL queries during the analysis phase.

**Customer Analysis**

**SQL QUERY:**

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

SELECT c.Name, SUM(s.TotalAmount) as TotalSpent

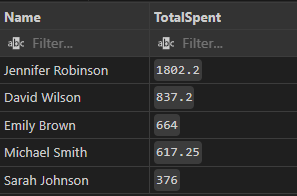
FROM Sales s

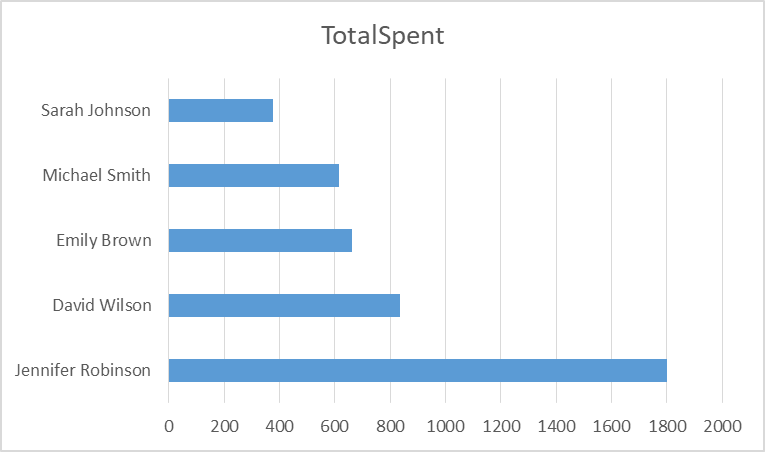
JOIN DimCustomer c ON s.CustomerID = c.CustomerID

GROUP BY c.CustomerID, c.Name

ORDER BY TotalSpent DESC

LIMIT 5;





This bar chart shows the total amount spent by each customer on vehicle repairs and parts. Jennifer Robinson stands out as the highest spender, with a total expenditure significantly higher than the other customers, followed by David Wilson. Sarah Johnson, Emily Brown, and Michael Smith have comparatively lower total spending.

**SQL QUERY:**

-- Average spending of customers on repairs and parts

SELECT AVG(TotalSpent) as AverageSpending

FROM (

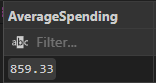
SELECT c.CustomerID, SUM(s.TotalAmount) as TotalSpent

FROM Sales s

JOIN DimCustomer c ON s.CustomerID = c.CustomerID

GROUP BY c.CustomerID

) as CustomerSpending;



The average spending of customers on repairs and parts is calculated to be $859.33. This metric helps in understanding the general spending behavior of the shop's customers. It provides a benchmark to identify high and low spenders and evaluate overall customer spending trends.

**SQL QUERY:**

-- Frequency of customer visists

SELECT c.Name, COUNT(\*) as VisitCount

FROM Sales s

JOIN DimCustomer c ON s.CustomerID = c.CustomerID

GROUP BY c.CustomerID, c.Name

ORDER BY VisitCount DESC;



**SQL QUERY:**

--enhanced query using date table for context

SELECT c.Name, COUNT(\*) as VisitCount, MIN(d.Date) as FirstVisit, MAX(d.Date) as LastVisit

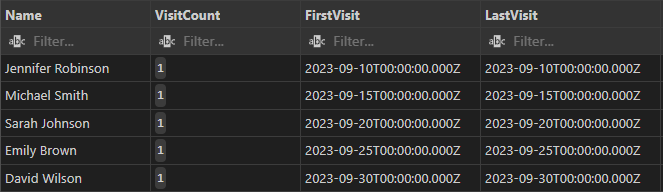
FROM Sales s

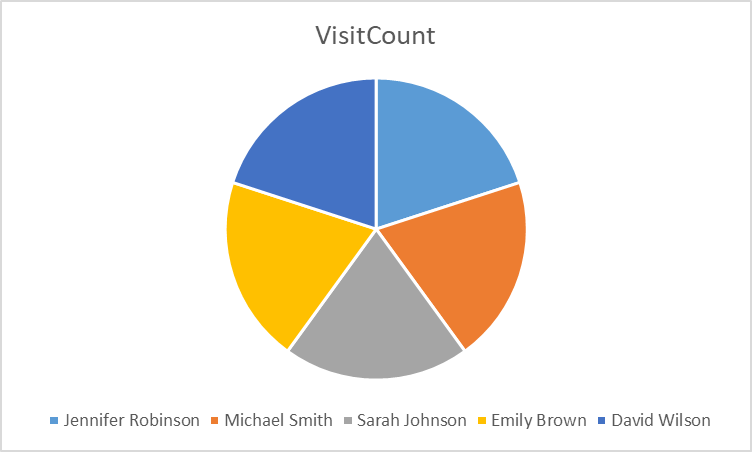
JOIN DimCustomer c ON s.CustomerID = c.CustomerID

JOIN DimDate d ON s.DateID = d.DateID

GROUP BY c.CustomerID, c.Name

ORDER BY VisitCount DESC;





The pie chart illustrates the visit frequency of each customer. The distribution shows that all customers, including Jennifer Robinson, Michael Smith, Sarah Johnson, Emily Brown, and David Wilson, have an even distribution of visit counts. From the query results that includes the dates they came, we see they each came once to the repair shop.

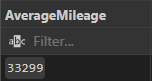
**Vehicle Analysis:**

**SQL QUERY:**

-- Average mileage of vehicles serviced

SELECT AVG(Mileage) as AverageMileage

FROM DimVehicle;



The average mileage of vehicles serviced at the car repair shop is calculated to be 33,299 miles. This metric helps in understanding the typical usage of vehicles that come in for repairs and can be useful in predicting the types of services required based on mileage.

**SQL QUERY:**

-- Most common vehicle makes and models brought in for service

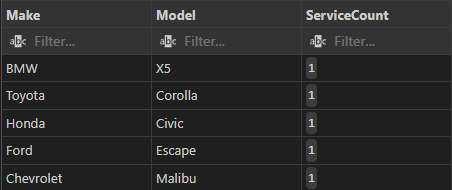
SELECT Make, Model, COUNT(\*) as ServiceCount

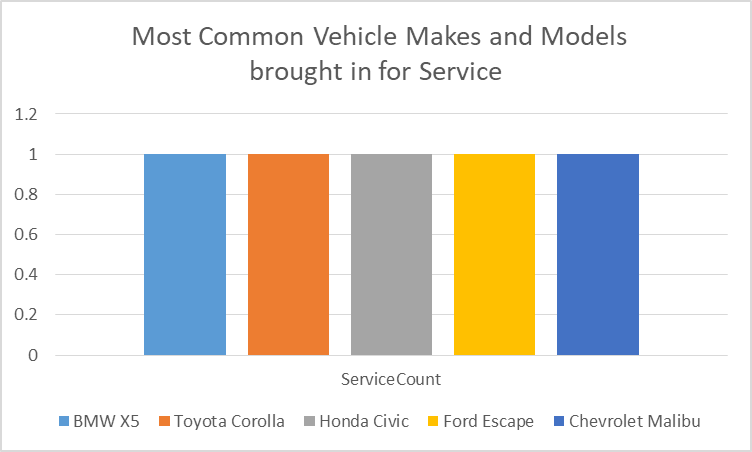
FROM DimVehicle

GROUP BY Make, Model

ORDER BY ServiceCount DESC

LIMIT 5;





The service count for all vehicles is 1, indicating that each vehicle make and model in the dataset has been serviced once. This result suggests a diverse range of vehicle makes and models being brought in for service, without any particular make or model standing out as the most common.

**SQL QUERY:**

-- Distribution of vehicle ages

SELECT

v.Make,

YEAR(CURDATE()) - v.Year as VehicleAge,

COUNT(\*) as VehicleCount,

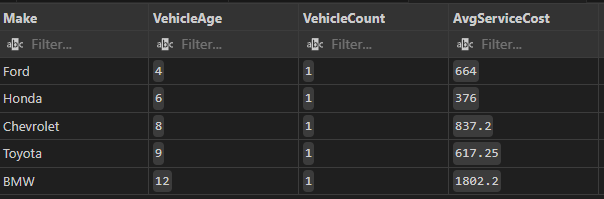
AVG(s.TotalAmount) as AvgServiceCost

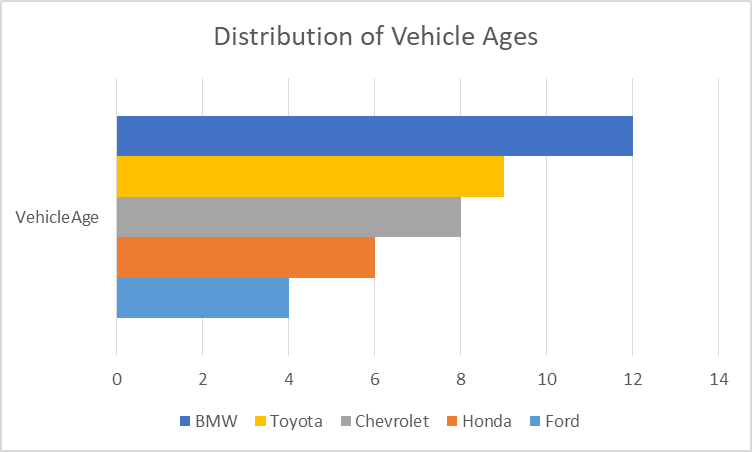
FROM DimVehicle v

JOIN Sales s ON v.VehicleID = s.VehicleID

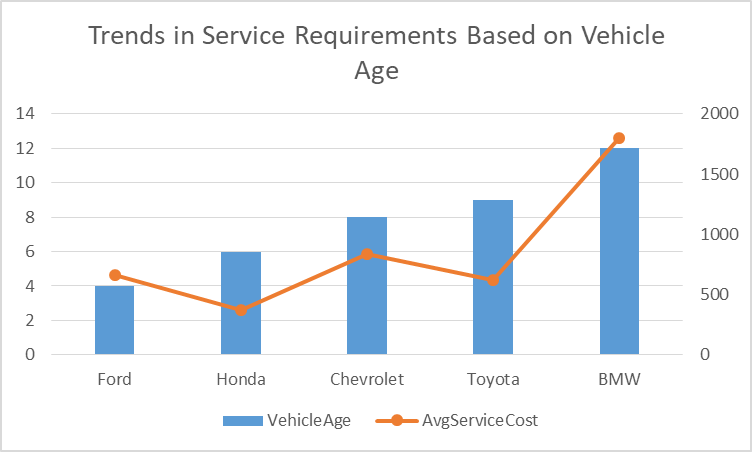
GROUP BY v.Make, VehicleAge

ORDER BY VehicleAge;





This bar chart shows the distribution of vehicle ages for different makes, including BMW, Toyota, Chevrolet, Honda, and Ford. BMW and Toyota vehicles have the highest representation in the older age brackets, indicating they are more likely to be serviced at the shop. Chevrolet and Honda have a moderate distribution across various ages, while Ford has a lower representation, particularly in older vehicle ages.



The combined line and bar chart illustrates the relationship between vehicle age and average service cost. The bar chart shows the ages for different vehicle makes, while the line chart represents the average service cost associated with these ages.

The analysis reveals that older BMW and Toyota vehicles tend to have higher service requirements and costs, indicating that these older models are likely to need more extensive repairs. Honda and Chevrolet show a steady service requirement trend, but their average service costs do not fluctuate as significantly as BMW and Toyota. Ford has fewer service requirements, which may suggest either lower maintenance needs or fewer older models being serviced.

**Job Performance Analysis:**

**SQL QUERY:**

-- Most common types of jobs and their frequency

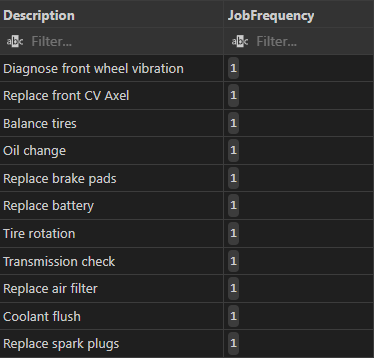
SELECT Description, COUNT(\*) as JobFrequency

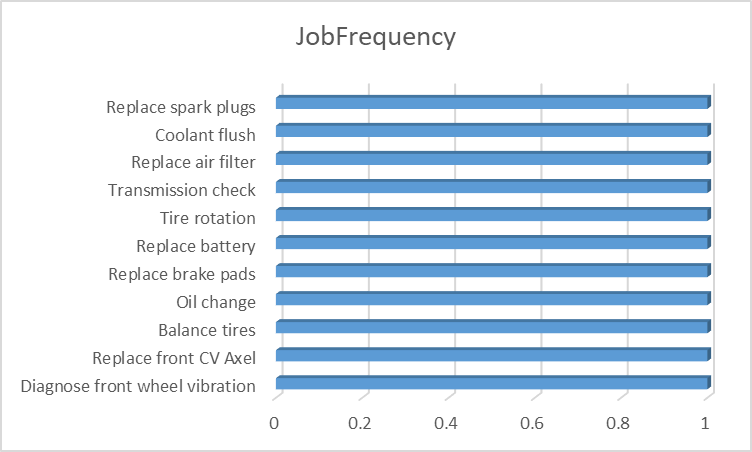
FROM FactServiceDetail fsd

JOIN DimService ds ON fsd.ServiceID = ds.ServiceID

GROUP BY Description

ORDER BY JobFrequency DESC;





The job frequency chart shows that all listed jobs have the same frequency of 1. This suggests that in the given dataset:

* Each job type has been performed exactly once.
* There is no clear "most common" job type, as all are equally frequent.
* This could indicate a limited dataset, perhaps covering a short time period or a sample where each job happened to occur once.

While this doesn't provide insight into which jobs are typically more common, it does show the range of services offered by the shop.

**SQL QUERY:**

-- Total revenue from each type of job

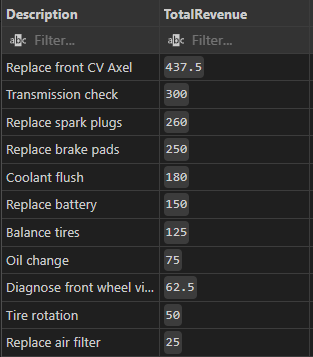
SELECT Description, SUM(Amount) as TotalRevenue

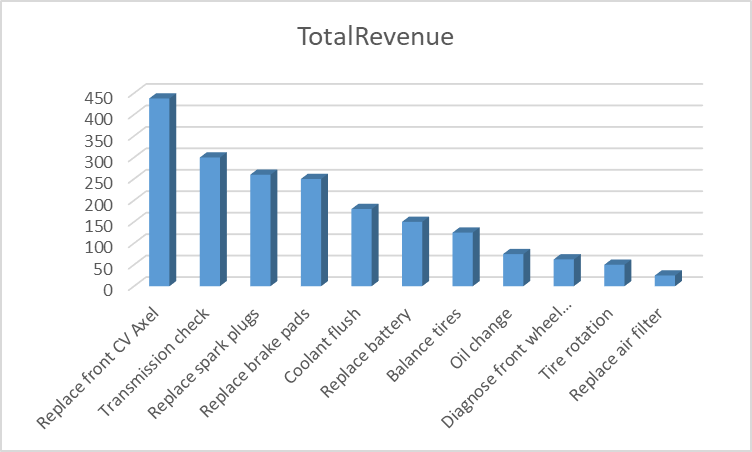
FROM FactServiceDetail fsd

JOIN DimService ds ON fsd.ServiceID = ds.ServiceID

GROUP BY Description

ORDER BY TotalRevenue DESC;





The total revenue chart and the corresponding table from the query results clearly show the revenue for each job type:

* Replace front CV Axel is the highest revenue generator at $437.5
* Followed by Transmission check at $300
* Replace spark plugs at $268
* Replace brake pads at $258
* The lowest revenue generators are Tire rotation at $50 and Replace air filter at $25

This data allows us to rank jobs by their contribution to overall revenue. It's clear that complex jobs like CV Axel replacement and transmission checks are significant revenue drivers for the shop.

**SQL QUERY:**

-- Jobs with highest and lowest average costs

-- Highest Average Cost

(SELECT Description, AVG(Amount) as AvgCost

FROM FactServiceDetail fsd

JOIN DimService ds ON fsd.ServiceID = ds.ServiceID

GROUP BY Description

ORDER BY AvgCost DESC

LIMIT 1)

UNION ALL

-- Lowest Average Cost

(SELECT Description, AVG(Amount) as AvgCost

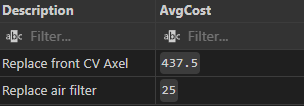
FROM FactServiceDetail fsd

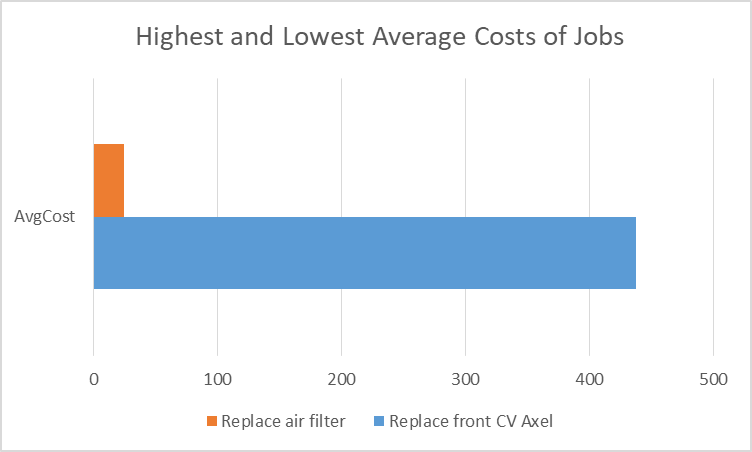
JOIN DimService ds ON fsd.ServiceID = ds.ServiceID

GROUP BY Description

ORDER BY AvgCost ASC

LIMIT 1);





The "Highest and Lowest Average Costs of Jobs" chart (bottom of Image 1) and the AvgCost table (bottom right of Image 2) provide this information:

* Highest average cost: Replace front CV Axel at $437.5
* Lowest average cost: Replace air filter at $35

It's worth noting that in this dataset, the average cost seems to directly correlate with the total revenue for each job. This suggests that either: a) Each job was only performed once, so the average cost equals the total revenue, or b) The pricing is very consistent for each job type.

Overall insights:

* There's a wide range in job costs and revenues, from $35 to $437.5.
* More complex jobs (like CV Axel replacement) generate significantly more revenue than simpler tasks (like air filter replacement).
* The consistency in job frequency limits our ability to identify truly common jobs, but provides a clear picture of the revenue and cost structure for various services.
* The shop might want to focus on promoting higher-revenue services or find ways to increase the frequency of these more profitable jobs.
* For lower-revenue jobs, the shop could consider bundling them with other services to increase overall transaction value.

**Parts Usage Analysis:**

**SQL QUERY:**

-- Top 5 most frequently used parts

SELECT PartName, SUM(Quantity) as TotalUsage

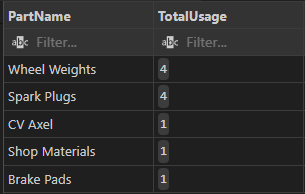
FROM FactPartDetail fpd

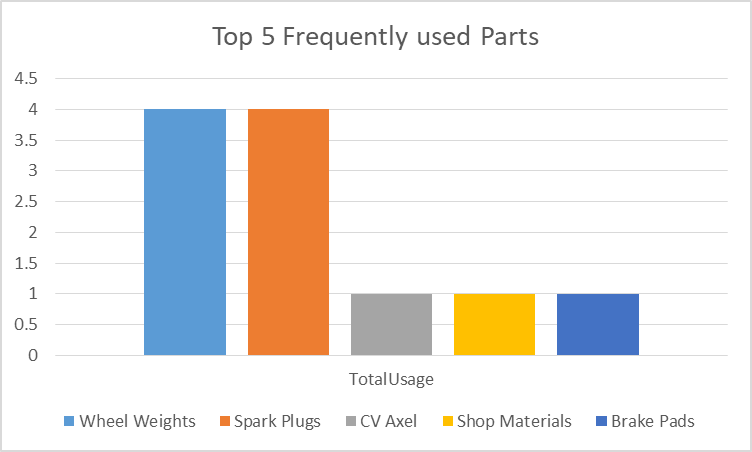
JOIN DimPart dp ON fpd.PartID = dp.PartID

GROUP BY PartName

ORDER BY TotalUsage DESC

LIMIT 5;





Top 5 most frequently used parts:

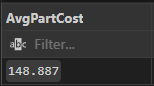
* Wheel Weights and Spark Plugs are tied for most frequent use (4 each)
* CV Axel, Shop Materials, and Brake Pads are next (1 each)

**SQL QUERY:**

-- Average cost of parts used in repairs

SELECT AVG(UnitPrice) as AvgPartCost

FROM DimPart;



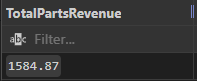
The average part cost is $148.887

**SQL QUERY:**

-- Total revenue from parts sales

SELECT SUM(Amount) as TotalPartsRevenue

FROM FactPartDetail;



Total parts revenue is $1584.87

This analysis shows a mix of frequently used low-cost items (like wheel weights) and less frequent but likely more expensive parts (like CV Axel). The parts sales contribute significantly to the overall revenue.

**Financial Analysis:**

**SQL QUERY:**

-- Monthly revenue from labor and parts

SELECT

CONCAT(dd.Year, '-', dd.Month) as YearMonth,

SUM(s.TotalLabor) as LaborRevenue,

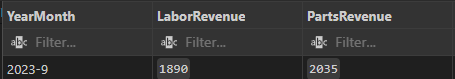
SUM(s.TotalParts) as PartsRevenue

FROM Sales s

JOIN DimDate dd ON s.DateID = dd.DateID

GROUP BY dd.Year, dd.Month

ORDER BY dd.Year, dd.Month;



For September 2023: Labor Revenue is $1598, Parts Revenue is $2035

**SQL QUERY:**

-- Overall profitability (Assuming some costs (e.g., 20% of labor costs and 30% of parts costs))

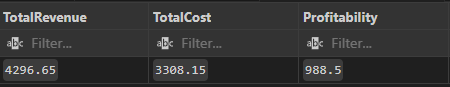
SELECT

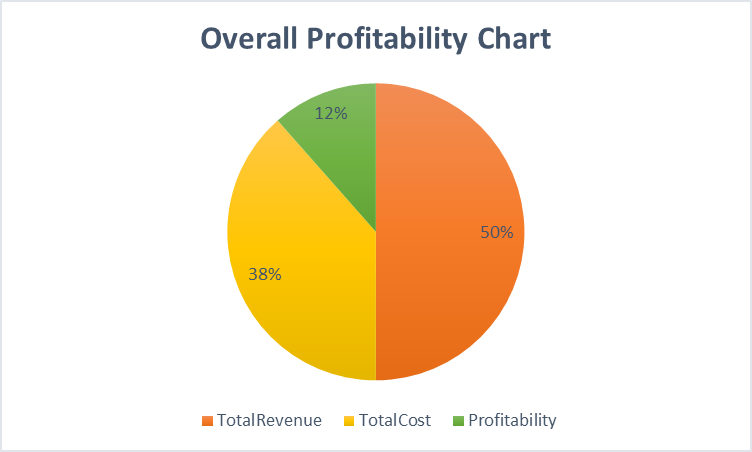
SUM(s.TotalAmount) as TotalRevenue,

SUM(s.TotalAmount- (s.TotalLabor \* 0.2) - (s.TotalParts \* 0.3)) as TotalCost,

SUM(s.TotalAmount) - SUM(s.TotalAmount - (s.TotalLabor \* 0.2) - (s.TotalParts \* 0.3)) as Profitability

FROM Sales s;





Overall profitability:

* Total Revenue: 50% of the pie chart
* Total Cost: 38% of the pie chart
* Profitability: 12% of the pie chart This indicates a healthy profit margin of 12%

**SQL QUERY:**

-- Impact of sales tax on total revenue

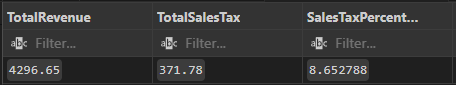
SELECT

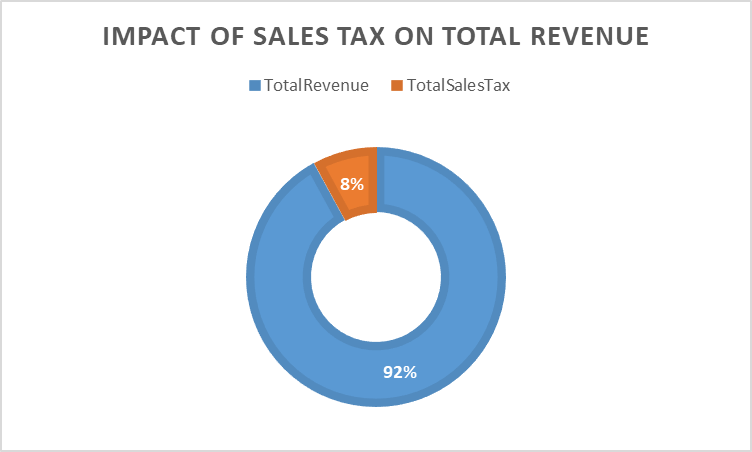
SUM(TotalAmount) as TotalRevenue,

SUM(SalesTax) as TotalSalesTax,

(SUM(SalesTax) / SUM(TotalAmount)) \* 100 as SalesTaxPercentage

FROM Sales;





Impact of sales tax:

* + Total Revenue: 92% of the circle
  + Total Sales Tax: 8% of the circle Sales tax accounts for 8% of the total revenue, which is a significant portion that needs to be accounted for in pricing and financial planning.

These analyses provide insights into inventory management, pricing strategies, and overall financial health of the repair shop.

**Optimization Recommendations**

Based on the analysis of the car repair shop’s data, here are comprehensive recommendations to optimize operations, enhance overall business performance, and improve profitability:

1. **Identifying Underperforming Services:**

* **Focus on Improving or Marketing Lower-Revenue Services**: Services like tire rotation and air filter replacement are identified as underperforming. Enhance their profitability by bundling them with higher-value services. Investigate the reasons behind their low frequency or revenue, which could be due to pricing, demand, or efficiency issues.
* **Marketing Efforts**: Increase marketing efforts for these services through promotions or package deals, emphasising their importance in vehicle maintenance.

2. **Parts Inventory Management:**

* **Higher Stock of Frequently Used Parts**: Ensure a higher inventory of frequently used parts like wheel weights and spark plugs. This strategy will reduce downtime and improve service efficiency.
* **Balanced Inventory of Expensive Part**s: Maintain a balanced stock of more expensive, less frequently used parts like CV axles to ensure availability without overinvesting.
* **Just-in-Time Inventory System**: Implement a just-in-time inventory system for less frequently used parts to reduce carrying costs and improve inventory turnover.

3. **Customer Loyalty Program:**

* **Tiered Loyalty Program**: Implement a tiered loyalty program based on total spending or visit frequency. Offer discounts on high-margin services or parts, priority scheduling, and extended warranties for top-tier customers to encourage repeat business.

4. **Scheduling Adjustments:**

* **Prioritise High-Revenue Services**: While the current data shows equal frequency for all jobs, prioritise scheduling for high-revenue services like CV axle replacement and transmission checks. Allocate more time slots for these complex jobs.
* **Efficient Workflow for Quick Jobs**: Develop efficient workflow processes for quicker jobs like tire rotations and air filter replacements to fit them between longer services, maximising the use of available time slots.

5. **Profitability Improvements:**

* Increase Operational Efficiency: With a current profit margin of 12%, look for ways to increase efficiency in operations to boost this margin. Analyse the cost structure (38% of revenue) to identify areas for potential cost reduction without compromising quality.

6. **Pricing Strategy:**

* **Adjust Prices for Sales Tax Impact**: Consider adjusting prices to account for the 8% sales tax impact, ensuring it doesn't reduce profits. Implement dynamic pricing for high-demand services or during peak times to maximise revenue.

7. **Service Mix Optimization:**

* **Promote High-Revenue Services**: Actively promote high-revenue services like CV axle replacement and transmission checks through targeted marketing campaigns. Train staff to upsell these services during routine maintenance visits.

8. **Customer Education:**

* **Develop an Education Program**: Educate customers on the importance of regular maintenance, focusing on high-revenue services. This can drive more frequent visits and increase the uptake of preventive services, contributing to overall revenue growth.

9. **Technician Training:**

* **Invest in Training Programs**: Invest in training programs to improve technician efficiency, particularly in performing high-value services. This can potentially increase the number of these jobs completed daily, enhancing overall service capacity.

10. **Data Collection and Analysis:**

* **Implement Robust Data Collection Systems**: Develop systems to collect detailed data on customer behaviour, job frequencies, and parts usage. This will enable more nuanced analysis and data-driven decision-making in the future.

By implementing these recommendations, the car repair shop can optimise its operations, improve inventory management, enhance customer loyalty, and ensure better resource allocation. These changes, driven by data analysis, will contribute to increased profitability and a more efficient service delivery model.

### **Conclusion**

### In conclusion, the analysis of the car repair shop's data has yielded valuable insights that can significantly benefit the business. Key findings include:

* **Customer Analysis**: Identifying top-spending customers and average spending patterns can help tailor marketing strategies and loyalty programs to enhance customer retention.
* **Vehicle Analysis**: Understanding the distribution of vehicle ages and average mileage serviced allows for better inventory and service planning, ensuring the shop can meet customer needs efficiently.
* **Job Performance and Parts Usage Analysis**: Highlighting the most common jobs performed and frequently used parts helps optimise inventory management and prioritise high-revenue services.
* **Financial Analysis**: Evaluating total revenue and profitability metrics provides a clear picture of the shop's financial health, guiding strategic decisions to improve margins and operational efficiency.

These insights can help the car repair shop optimise its operations, improve customer satisfaction, and increase profitability. Implementing the recommendations, such as focusing on underperforming services, enhancing parts inventory management, and developing customer loyalty programs, will drive business growth and efficiency.

### **Future Steps**

To further leverage the data analysis:

* **Continuous Data Collection and Analysis**: Implement robust systems for ongoing data collection to enable continuous improvement and more nuanced future analysis.
* **Refinement of Inventory Management**: Develop and refine inventory management strategies based on usage patterns and demand forecasts.
* **Advanced Customer Analytics**: Utilise advanced analytics to gain deeper insights into customer behaviour and preferences, enabling more targeted marketing efforts.
* **Operational Efficiency Improvements**: Invest in training and process optimization to improve the efficiency and quality of service delivery.

By taking these future steps, the car repair shop can maintain a competitive edge and continue to enhance its service offerings, ultimately leading to sustained business success.

**Data**

Cleaned CSV files**:** [DB CSV Files](https://drive.google.com/drive/folders/1wPwQ6Wy5CZfarcHxnNPkCXTFhHI0OeB-?usp=sharing)

SQL Queries:  [Dimensional Model for Sales Analysis and SQL Queries](https://drive.google.com/file/d/1hbWdRWSyCu8KWNtAxv8X_1GQ8cRN93yx/view?usp=sharing)