Creating a dimensional data model using MySQL Workbench

BY

HENRY OKAM

30 July 2024

1.0. Introduction

In the fast-paced automotive repair industry, data-driven insights are crucial for success. This report showcases the design and development of a dimensional data model that empowers car repair centers across western Canada to analyze sales performance and make informed decisions.

By examining a sample invoice, the analyst pinpointed key data points and identified essential analysis types, such as sales by customer, vehicle, service, parts, and location. A robust dimensional data model was then created, comprising a fact table for quantitative sales data and dimension tables for contextual information. The Entity-Relationship diagram, built in MySQL, seamlessly connects these tables, enabling detailed analysis and business insights.

This report offers a comprehensive guide to creating the dimensional model, from understanding business needs to designing and linking tables. The resulting model provides a powerful framework for sales data analysis, driving business growth and improvement. Stakeholders will find this report an invaluable resource, gaining a deeper understanding of the data model's design and functionality.

2.0. Business Requirements:

The business requires a comprehensive analysis of various aspects of its operations, including customer behavior, vehicle trends, job performance, parts usage, and overall financial health. By analyzing these factors, the business can optimize its operations, improve customer satisfaction, and increase profitability. The key information identified from the sample invoice will be used to create a dimensional data model, enabling detailed and flexible analysis to support these objectives.

3.0. Key Information Relevant for Sales Analysis

Based on the review of the sample invoice, the key pieces of information relevant for sales analysis will include:

1. Customer Information such as: *Customer Name, Customer Address, Customer Phone Number.*
2. Vehicle Information such as: *Make, Model, Year, Colour, Vehicle Identification Number (VIN), Registration Number, Mileage.*
3. Service Information such as: *Job Description, Hours Worked, Hourly Rate, Total Labor Cost.*
4. Parts Information such as: *Part Number, Part Name, Quantity, Unit Price, Total Parts Cost.*
5. Invoice Summary such as: *Invoice Number, Invoice Date, Total Service Charges, Total Parts Charges, Sales Tax, Total Amount Due.*
6. Location Information such as *Shop Location.*
7. Date Information such as: *Invoice Date*.

4.0. Types of Analysis the Business Will Need

Based on the sample invoice provided, the business will need the following types of analysis:

Customer Analysis

Top-Spending Customers: Identify the top 5 customers who have spent the most on vehicle repairs and parts.

Average Customer Spending: Determine the average spending of customers on repairs and parts.

Customer Visit Frequency: Analyse the frequency of customer visits and identify any patterns.

Vehicle Analysis

Average Mileage: Calculate the average mileage of vehicles serviced.

Common Vehicle Makes and Models: Identify the most common vehicle makes and models brought in for service.

Vehicle Age Trends: Analyse the distribution of vehicle ages and identify any trends in service requirements based on vehicle age.

Job Performance Analysis:

Common Job Types: Determine the most common types of jobs performed and their frequency.

Revenue by Job Type: Calculate the total revenue generated from each type of job.

Cost Analysis by Job Type: Identify the jobs with the highest and lowest average costs.

Parts Usage Analysis:

Frequently Used Parts: List the top 5 most frequently used parts and their total usage.

Average Part Cost: Calculate the average cost of parts used in repairs.

Revenue from Parts Sales: Determine the total revenue generated from parts sales.

Financial Analysis:

Monthly Revenue: Calculate the total revenue generated from labour and parts for each month.

Overall Profitability: Determine the overall profitability of the repair shop.

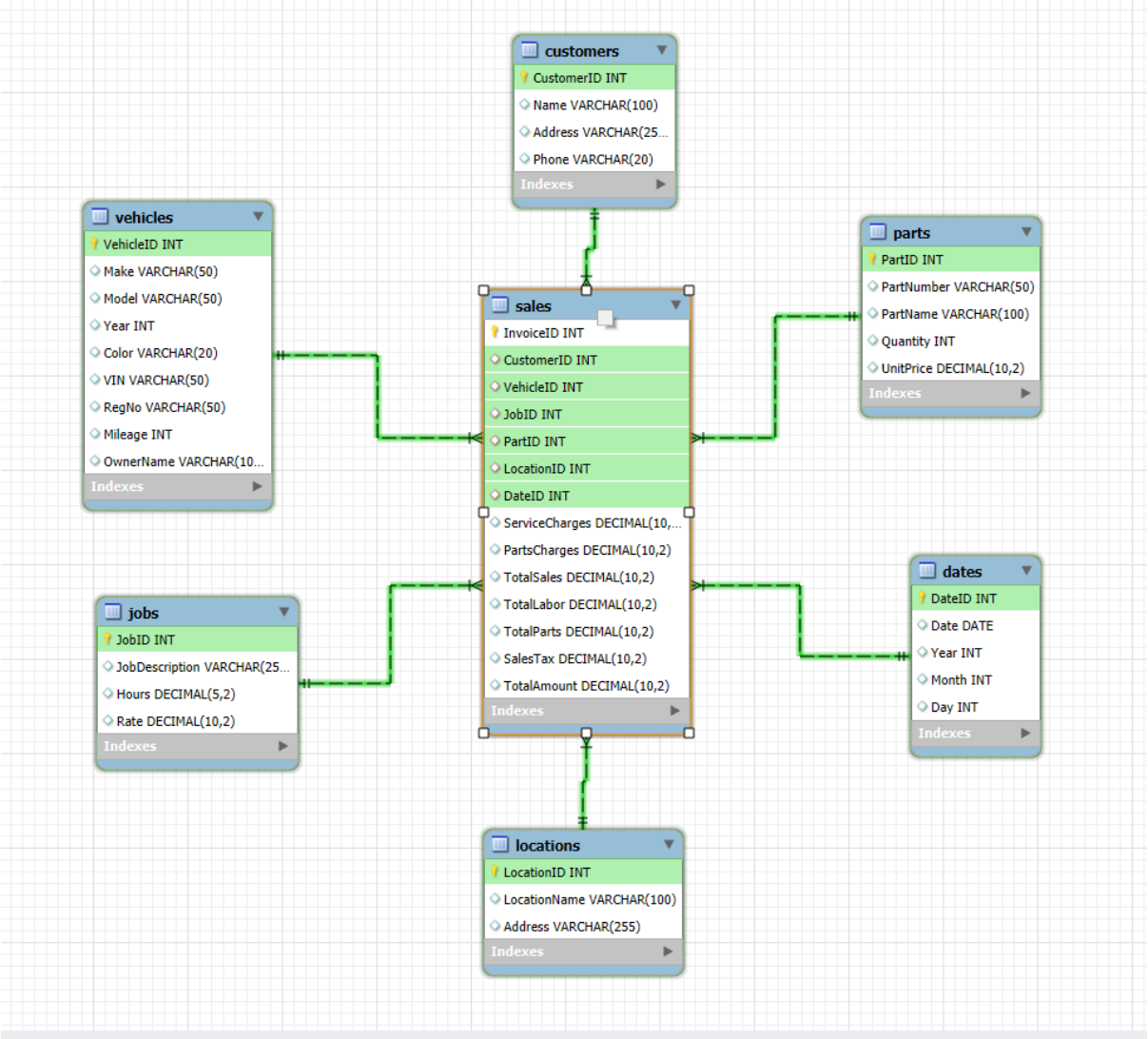
Sales Tax Impact: Analyse the impact of sales tax on the total revenue.

Location Analysis:

Performance by Location: Analyse the sales performance of different shop locations across western Canada.

5.0. The Entity Relationship Diagram (Star Schema)

A dimensional data model was constructed using MySQL and named 'DA_mchenryspagg' per the project's specifications. This data model adopts the star schema, consisting of a sales fact table and six additional dimension tables: 'jobs,' 'locations,' 'vehicles,' 'customers,' 'parts,' and 'dates.' The ER diagram screenshot below illustrates the relationships among these tables, connecting primary keys and foreign keys between the fact table and its associated dimensions.



6.0. Table Descriptions and Column Explanations

Based on the tables created on the database schema, the following descriptions and explanations will suffice for each table.

Dimension Tables

1. Customers Table

The purpose of this table is to store information about the customers.

Columns:

CustomerID (INT, Primary Key): A unique identifier for each customer.

Name (VARCHAR(100)): The customer's full name.

Address (VARCHAR(255)): The customer's address.

Phone (VARCHAR(20)): The customer's phone number.

Logical Explanation:

From the sales receipt, we have the customer’s name, address, and phone number. These attributes are essential for identifying and analysing customer behaviour and sales patterns.

2. Vehicles Table

The purpose of this table is to store information about the vehicles serviced.

Columns

VehicleID (INT, Primary Key): A unique identifier for each vehicle.

Make (VARCHAR(50)): The make of the vehicle (e.g., BMW).

Model (VARCHAR(50)): The model of the vehicle (e.g., X5).

Year (INT): The year of manufacture.

Color (VARCHAR(20)): The color of the vehicle.

VIN (VARCHAR(50)): The Vehicle Identification Number.

RegNo (VARCHAR(50)): The vehicle's registration number.

Mileage (INT): The mileage of the vehicle at the time of service.

Logical Explanation

The sales receipt includes details about the vehicle, which are crucial for analysing service patterns, identifying common issues with specific makes/models, and tracking vehicle service history.

3. Jobs Table

The purpose of this table is to store information about the jobs performed during service.

Columns:

JobID (INT, Primary Key): A unique identifier for each job.

JobDescription (VARCHAR(255)): A description of the job performed (e.g., "Diagnose front wheel vibration").

Hours (DECIMAL(5, 2)): The number of hours spent on the job.

Rate (DECIMAL(10, 2)): The hourly rate for the job.

Logical Explanation:

Each job performed during a service is detailed on the sales receipt. This table helps in analysing job performance, frequency of specific jobs, and revenue generated from different types of jobs.

4. Parts Table

The purpose of this table is to store information about the parts used in repairs.

Columns:

PartID (INT, Primary Key): A unique identifier for each part.

PartNumber (VARCHAR(50)): The part number.

PartName (VARCHAR(100)): The name of the part.

Quantity (INT): The quantity of the part used.

UnitPrice (DECIMAL(10, 2)): The price per unit of the part.

Logical Explanation:

The sales receipt lists parts used in repairs. This table allows tracking parts usage, inventory management, and analysing the cost and revenue associated with parts.

5. Locations Table

The purpose of this table is to store information about the shop locations.

Columns:

LocationID (INT, Primary Key): A unique identifier for each location.

LocationName (VARCHAR(100)): The name of the location.

Address (VARCHAR(255)): The address of the location.

Logical Explanation:

Having location information is essential for analysing sales and service performance across different shop locations, which can help in operational optimization and strategic planning.

6. Dates Table

The purpose of this table is to store information about the date of transactions.

Columns:

DateID (INT, Primary Key): A unique identifier for each date.

Date (DATE): The actual date of the transaction.

Year (INT): The year of the transaction.

Month (INT): The month of the transaction.

Day (INT): The day of the transaction.

Logical Explanation:

Breaking down the invoice date into year, month, and day allows for detailed time-based analysis of sales, helping to identify trends and seasonal patterns.

Fact Table

7. Sales Table

The purpose of this table is to store quantitative data related to sales, linking to dimension tables for detailed analysis.

Columns:

InvoiceID (INT, Primary Key): A unique identifier for each invoice.

CustomerID (INT, Foreign Key): Links to the Customers table.

VehicleID (INT, Foreign Key): Links to the Vehicles table.

JobID (INT, Foreign Key): Links to the Jobs table.

PartID (INT, Foreign Key): Links to the Parts table.

LocationID (INT, Foreign Key): Links to the Locations table.

DateID (INT, Foreign Key): Links to the Dates table.

ServiceCharges (DECIMAL(10, 2)): The total service charges.

PartsCharges (DECIMAL(10, 2)): The total charges for parts.

TotalSales (DECIMAL(10, 2)): The total sales amount.

TotalLabor (DECIMAL(10, 2)): The total labor charges.

TotalParts (DECIMAL(10, 2)): The total parts charges.

SalesTax (DECIMAL(10, 2)): The sales tax amount.

TotalAmount (DECIMAL(10, 2)): The total invoice amount.

Logical Explanation:

The sales receipt is the primary source of data for this table. It aggregates financial data for comprehensive sales analysis, linking to all dimension tables to provide a detailed and flexible analysis framework.

7.0. Summary

The dimensional model consists of clearly defined fact and dimension tables, each serving a specific purpose in sales analysis. The tables and columns were designed based on the data available in the sales receipt, ensuring all necessary details for thorough analysis are captured. This structure supports a wide range of analyses, from customer behaviour and vehicle trends to job performance and parts usage, enabling the business to make informed decisions and optimize operations.

8.0. Appendix

8.1. SQL Script

```
-- Step 1: Create Database
CREATE DATABASE DA_mchenryspagg;
USE DA_mchenryspagg;

-- Step 2 : CREATE DIMENSION TABLES
-- (a)  Customers Table:
CREATE TABLE Customers (
    CustomerID INT AUTO_INCREMENT PRIMARY KEY,
    Name VARCHAR(100),
    Address VARCHAR(255),
    Phone VARCHAR(20)
);

-- Create the Vehicles Table:
CREATE TABLE Vehicles (
    VehicleID INT AUTO_INCREMENT PRIMARY KEY,
    Make VARCHAR(50),
    Model VARCHAR(50),
```

```
    Year INT,
    Color VARCHAR(20),
    VIN VARCHAR(50),
    RegNo VARCHAR(50),
    Mileage INT,
    OwnerName VARCHAR(100)
);

-- Create the Jobs Table
CREATE TABLE Jobs (
    JobID INT AUTO_INCREMENT PRIMARY KEY,
    JobDescription VARCHAR(255),
    Hours DECIMAL(5, 2),
    Rate DECIMAL(10, 2)
);

-- Create the Parts Table:
CREATE TABLE Parts (
    PartID INT AUTO_INCREMENT PRIMARY KEY,
    PartNumber VARCHAR(50),
    PartName VARCHAR(100),
    Quantity INT,
    UnitPrice DECIMAL(10, 2)
);

-- Create the Locations Table
CREATE TABLE Locations (
    LocationID INT AUTO_INCREMENT PRIMARY KEY,
    LocationName VARCHAR(100),
    Address VARCHAR(255)
```

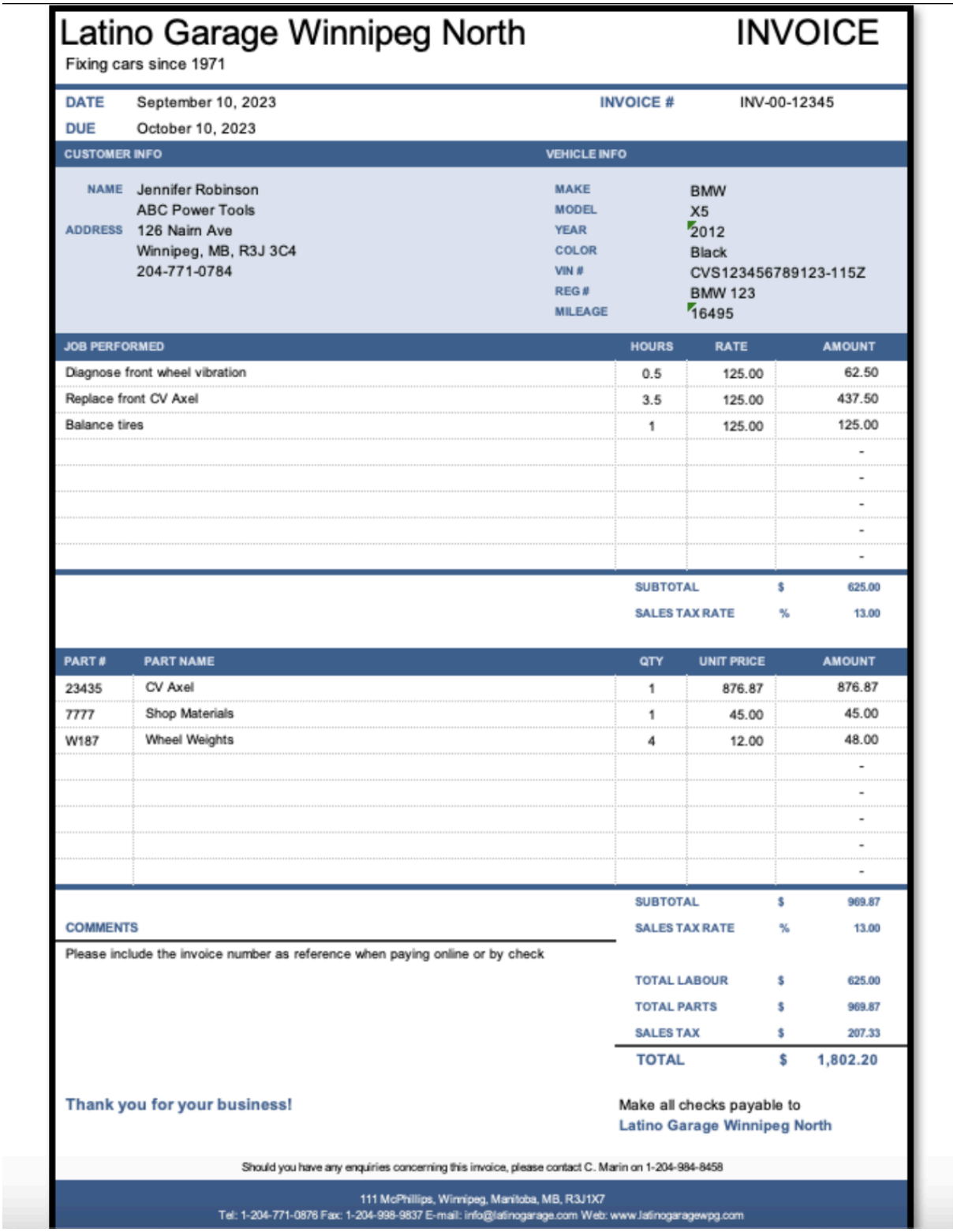
```
);

-- Create the Dates Table
CREATE TABLE Dates (
    DateID INT AUTO_INCREMENT PRIMARY KEY,
    Date DATE,
    Year INT,
    Month INT,
    Day INT
);

-- Step 3: Create the Fact Table
CREATE TABLE Sales (
    InvoiceID INT AUTO_INCREMENT PRIMARY KEY,
    CustomerID INT,
    VehicleID INT,
    JobID INT,
    PartID INT,
    LocationID INT,
    DateID INT,
    ServiceCharges DECIMAL(10, 2),
    PartsCharges DECIMAL(10, 2),
    TotalSales DECIMAL(10, 2),
    TotalLabor DECIMAL(10, 2),
    TotalParts DECIMAL(10, 2),
    SalesTax DECIMAL(10, 2),
    TotalAmount DECIMAL(10, 2),
    FOREIGN KEY (CustomerID) REFERENCES Customers(CustomerID),
    FOREIGN KEY (VehicleID) REFERENCES Vehicles(VehicleID),
    FOREIGN KEY (JobID) REFERENCES Jobs(JobID),
```

```
    FOREIGN KEY (PartID) REFERENCES Parts(PartID),
    FOREIGN KEY (LocationID) REFERENCES Locations(LocationID),
    FOREIGN KEY (DateID) REFERENCES Dates(DateID)
);
```

8.1. Sample Invoice Receipt



Latino Garage Winnipeg North

Fixing cars since 1971

INVOICE

DATE September 10, 2023
DUE October 10, 2023
INVOICE # INV-00-12345

CUSTOMER INFO

NAME Jennifer Robinson
ABC Power Tools
ADDRESS 126 Nairn Ave
Winnipeg, MB, R3J 3C4
204-771-0784

VEHICLE INFO

MAKE BMW
MODEL X5
YEAR 2012
COLOR Black
VIN # CVS123456789123-115Z
REG # BMW 123
MILEAGE 16495

JOB PERFORMED	HOURS	RATE	AMOUNT
Diagnose front wheel vibration	0.5	125.00	62.50
Replace front CV Axel	3.5	125.00	437.50
Balance tires	1	125.00	125.00
			-
			-
			-
			-
			-

SUBTOTAL	$	625.00
SALES TAX RATE	%	13.00

PART #	PART NAME	QTY	UNIT PRICE	AMOUNT
23435	CV Axel	1	876.87	876.87
7777	Shop Materials	1	45.00	45.00
W187	Wheel Weights	4	12.00	48.00
				-
				-
				-
				-
				-

SUBTOTAL	$	969.87
SALES TAX RATE	%	13.00
TOTAL LABOUR	$	625.00
TOTAL PARTS	$	969.87
SALES TAX	$	207.33
TOTAL	$	1,802.20

COMMENTS

Please include the invoice number as reference when paying online or by check

Thank you for your business!

Make all checks payable to
Latino Garage Winnipeg North

Should you have any enquiries concerning this invoice, please contact C. Marin on 1-204-984-8458

111 McPhillips, Winnipeg, Manitoba, MB, R3J1X7
Tel: 1-204-771-0876 Fax: 1-204-998-9837 E-mail: info@latinogarage.com Web: www.latinogaragewpg.com