Colorado Motor Vehicle Sales SQL Data Analysis

Author: Pranjal Waghmare
Date: November 11, 2025
Database Platform: MySQL Workbench

---

1. Project Overview

This project focuses on analyzing motor vehicle sales data across various counties in Colorado using **MySQL Workbench**.
The goal is to **clean, transform, and explore** the dataset through SQL queries — identifying key sales trends, top-performing counties, and generating insights useful for **business strategy and policy-making**.

---

2. Dataset Description

The dataset `colorado_motor_vehicle_sales_cleaned.csv` contains **yearly and quarterly sales data** of motor vehicles for multiple Colorado counties.
It helps analyze automotive market trends, performance variations, and regional differences.

Columns:

- **Year:** Year of recorded sales.
- **Quarter:** Quarter of the year (Q1–Q4).

- **County:** Name of the Colorado county.
- **Sales:** Total vehicle sales amount.

---

3. Database Setup Instructions

1. Open **MySQL Workbench**.
2. Create a new database using the provided SQL script: `colorado_motor_vehicle_sales_analysis.sql`.
3. Use **Table Data Import Wizard** to import the CSV file into the table `raw_motor_sales`.
4. Execute all SQL sections in order to perform **data cleaning**, **EDA**, and **insight generation**.

---

4. SQL Data Cleaning & Transformation

Data cleaning ensures that all fields are consistent, duplicates are removed, and county names are standardized.

Key Steps:

- Create a raw data table for CSV import.
- Clean data using string manipulation (`TRIM`, `REPLACE`, `CONCAT`).
- Convert data types to ensure numeric consistency.

- Remove duplicates using `ROW_NUMBER()` window functions.
- Create indexes for efficient query performance.

---

5. Exploratory Data Analysis (EDA)

EDA is performed with SQL queries to reveal major trends and distributions.

Analyses include:

- Yearly and quarterly sales totals
- Year-over-year (YoY) growth analysis
- Top and bottom performing counties
- County-wise contribution to state sales
- Quarterly comparison trends

📸 *Insert Query Output Screenshot: Yearly Sales Trend Here*
📸 *Insert Query Output Screenshot: Top Counties by Sales Here*

---

6. Views and Stored Procedures

To simplify analysis and dashboard creation, **views** and a **stored procedure** are created.

Views:

- `v_yearly_sales` — Summarized yearly sales totals
- `v_quarterly_sales` — Aggregated quarterly sales totals
- `v_county_sales` — County-wise total sales

Stored Procedure Example:

```
CALL get_top_counties(10);
```

Retrieves the **top 10 counties** by total vehicle sales.

📸 *Insert Screenshot: Stored Procedure Output Here*

---

7. Key Insights Summary

- Vehicle sales show **steady year-over-year growth** in major counties.
- **Quarter 2 and Quarter 4** record the highest activity.
- Counties such as **Denver** and **El Paso** dominate overall sales.
- Smaller rural counties contribute **less than 2%** of total vehicle sales.
- The overall trend aligns with **economic growth cycles** in Colorado.

---

8. Execution Instructions for MySQL Workbench

1. Open MySQL Workbench.

2. Navigate to: **File → Open SQL Script → colorado_motor_vehicle_sales_analysis.sql**.

3. Run database creation and cleaning commands.

4. Execute EDA queries sequentially and capture screenshots for visualization.

5. Use the provided views and stored procedure for quick insights.

---

9. Author

Prepared by: Pranjal Waghmare
Project: Colorado Motor Vehicle Sales SQL Data Analysis
Tool Used: MySQL Workbench