# Colorado Motor Vehicle Sales Data Analysis Report

**Introduction**

This report presents an analysis of motor vehicle sales across various counties in Colorado, based on data segmented by year and quarter. The objective of this analysis is to identify key trends, highlight significant findings, and provide actionable insights to stakeholders. The dataset used in this project is named colorado_motor_vehicle_sales.csv and includes information on total sales, geographic distribution, and temporal patterns.

---

**Key Findings**

**Maximum and Minimum Sales**

- **Maximum Sales**: The highest motor vehicle sales event occurred in **Q3 of 2015** in **Arapahoe County**, with sales amounting **$916.91 million**.
- **Minimum Sales**: The lowest sales were recorded in **Q1 of 2009** in **Fremont County**, amounting to **$6.27 million**.

**Counties with Sales Performance**

**Counties with Average Sales Above the Overall Average:**

1. **Arapahoe**: $629.45 million
2. **El Paso**: $372.69 million
3. **Jefferson**: $283.08 million
4. **Adams**: $278.19 million
5. **Denver**: $211.36 million

These counties exhibit strong sales performance, likely attributed to larger populations or higher economic activity.

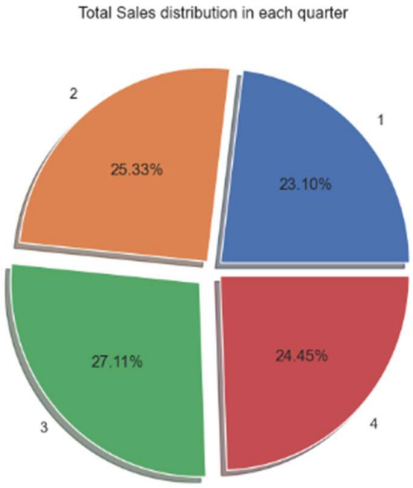
**Counties with Average Sales Below the Overall Average:**

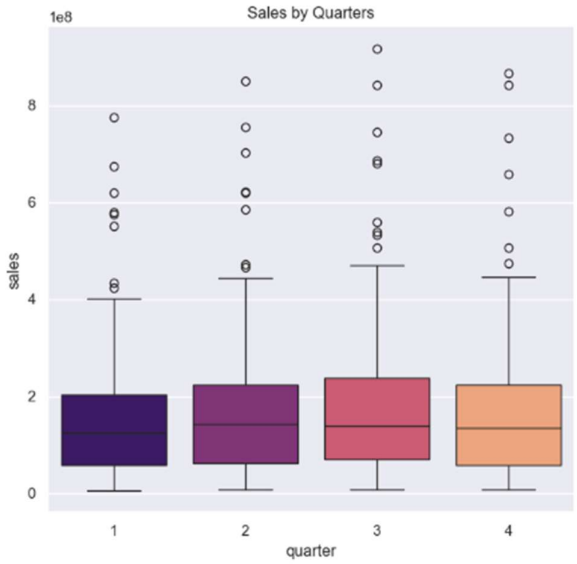
1. **Fremont**: $9.90 million
2. **La Plata**: $24.78 million
3. **Broomfield**: $36.69 million
4. **Garfield**: $50.02 million
5. **Pueblo**: $74.50 million

The lowest-performing counties tend to have smaller populations or lower economic activity, leading to reduced motor vehicle sales.

**Quarterly Sales Trends**

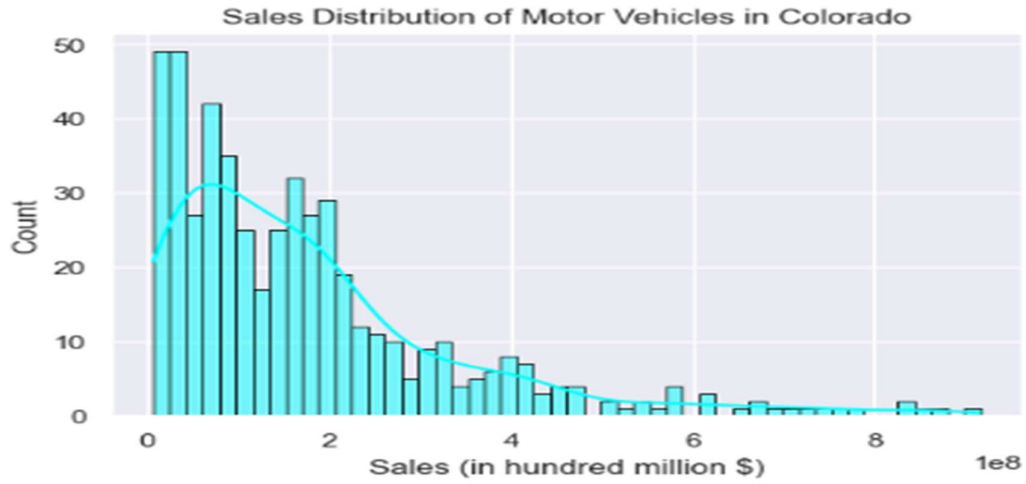
- **Peak Quarter**: Sales consistently peaked in **Q3** across the dataset, potentially driven by seasonal demand or economic conditions.
- **Low Quarter**: **Q1** exhibited the lowest sales, possibly due to post-holiday slowdowns or adverse weather conditions affecting vehicle purchases.




**Distribution Insights**

The sales data shows a **right-skewed distribution**, with most counties having relatively low sales, while a few counties, such as Arapahoe and El Paso, contribute disproportionately high sales. This indicates significant economic or population disparities among counties.



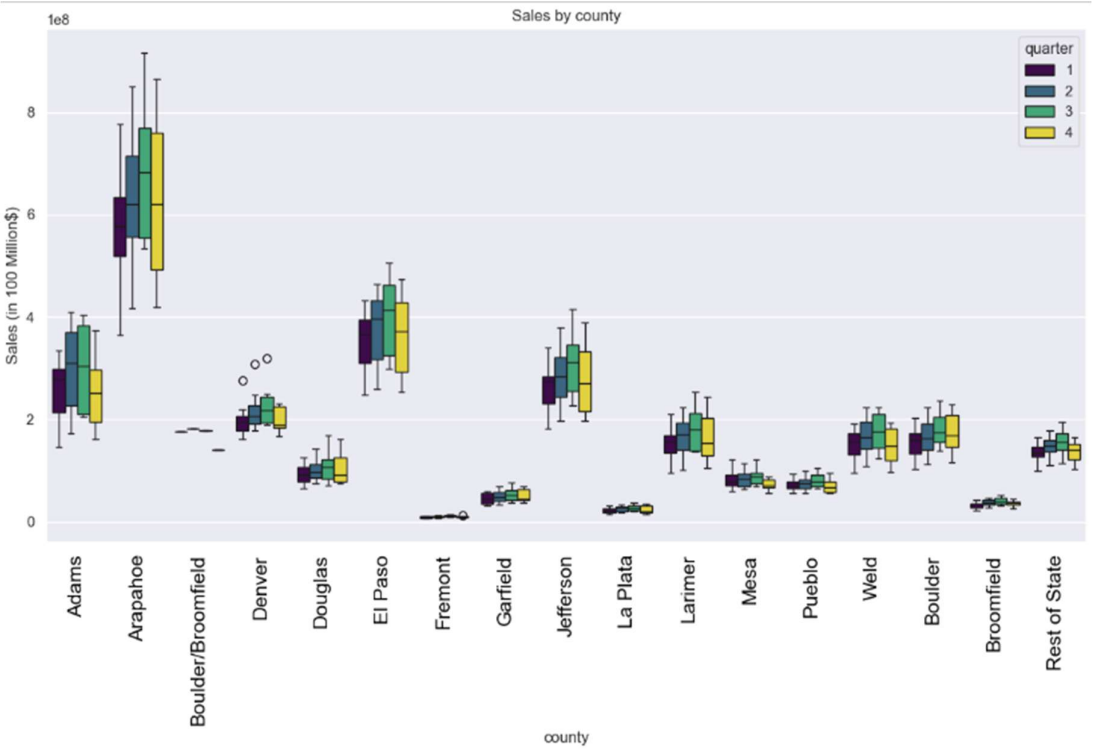

---

**Statistical and Predictive Analysis**

**Sales Summary by County**

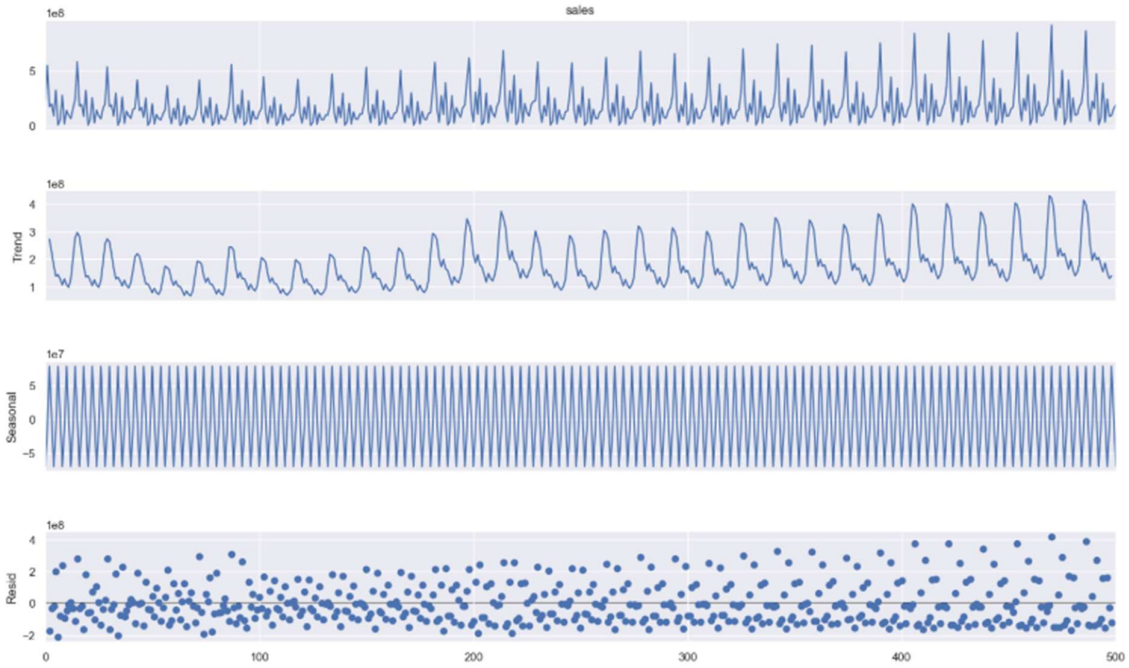
| County | Total Sales (Sum) | Average Sales (Mean) | Maximum Sales | Minimum Sales |
|---|---|---|---|---|
| Arapahoe | $20.14 billion | $629.45 million | $916.91M | $365.06M |
| Fremont | $316.94 million | $9.90 million | $13.91M | $6.27M |
| El Paso | $11.93 billion | $372.69 million | $506.45M | $247.02M |
| Rest of State | $3.58 billion | $143.29 million | $195.50M | $99.93M |

- **Arapahoe County** leads in all major metrics, highlighting its economic prominence.
- **Fremont County** shows the lowest figures, indicating limited vehicle transactions.



**Time Series Analysis**

- The sales data was tested for stationarity using statistical methods.
    - **Test Statistic**: -0.76
    - **p-value**: 0.83
    - **Conclusion**: The data is non-stationary, meaning trends and patterns change over time without a constant mean or variance.




- Predictive modeling with ARIMA or SARIMA is recommended to account for non-stationarity and forecast future sales effectively.

**Predictive Modeling**

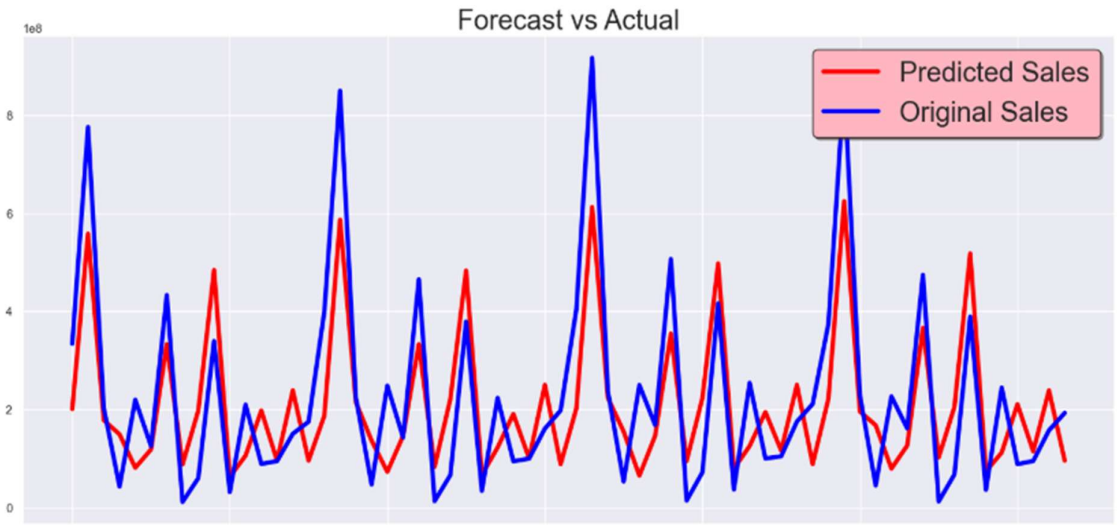
**Approach and Techniques**

1. **Model Selection**: Based on the non-stationary nature of the data, ARIMA and SARIMA models were selected for forecasting future sales. These models are well-suited for handling time series data with trends and seasonal patterns.
2. **Data Preparation**:
    - Differencing was applied to the data to make it stationary.
    - Seasonal patterns were identified and incorporated into the SARIMA model.
3. **Model Training**:
    - ARIMA and SARIMA models were trained using historical sales data.
    - Model parameters were optimized using grid search to identify the best configuration (e.g., p, d, q values).
4. **Evaluation**:

- Models were evaluated using metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE).
- SARIMA outperformed ARIMA in capturing seasonal trends, with lower error metrics.

**Forecast Results**

- The SARIMA model provided accurate forecasts for future quarters, highlighting expected peaks in Q3 and lower sales in Q1.
- Forecast visualizations showed a strong alignment with historical patterns, validating the model's reliability.




---

**Insights and Recommendations**

**Key Insights:**

1. **High-Sales Counties**: Arapahoe, El Paso, Jefferson, Adams, and Denver consistently outperform other counties in vehicle sales.
2. **Seasonal Trends**: Peak sales in Q3 suggest a need for increased inventory and marketing efforts during this period.
3. **Economic Disparities**: Significant differences in sales figures highlight the need for targeted economic development in underperforming counties.

**Recommendations:**

- **Business Strategy**: Automotive businesses should focus on high-sales counties for expansion and allocate more resources during Q3.
- **Policy Making**: State officials could investigate barriers to vehicle sales in low-performing counties, such as Fremont and La Plata, to foster economic growth.
- **Forecasting Models**: Continue using SARIMA models for planning inventory and marketing strategies, as they accurately capture seasonal and trend components.

---

**Conclusion**

This analysis highlights the diverse economic landscape of Colorado's counties in terms of motor vehicle sales. Understanding these trends provides valuable insights for businesses, policymakers, and other stakeholders. Future analyses could incorporate additional factors, such as population growth or economic indicators, to enhance the understanding of motor vehicle sales dynamics in the state.