Automobile Sales and Economic Trends Using Data Science To

GitHub Repository:

https://github.com/amxxxmelhem/coursera-projects

1. Executive Summary

This project explores automobile sales data alongside economic indicators such as GDP and unemployment rates. The objective is to uncover sales trends, understand impacts of economic cycles, and build predictive models for sales classification.

2. Introduction

The automotive industry is a key economic sector influenced by multiple factors. This analysis aims to identify sales trends by vehicle type and region, especially during recession periods, and to develop a data-driven dashboard.

3. Data Collection & Wrangling Methodology

- Data from multiple CSV sources containing sales, GDP, unemployment, and advertising expenditure.
- Data cleaning involved handling missing values, merging datasets, and feature engineering.
- Tools used include pandas and SQL queries.

4. EDA & Interactive Visual Analytics Methodology

- Exploratory Data Analysis used visualization libraries including Matplotlib, Seaborn, Plotly.

Capstone Project Presentation

- Interactive plots allowed dynamic exploration of trends by vehicle types and regions.

- SQL was used for data aggregation and complex filtering.

5. EDA Visualization Results

- Line plots showed year-over-year fluctuations in sales.

- Bar charts highlighted differences by vehicle categories.

- Scatter plots revealed relationships between advertising spend and sales.

6. EDA with SQL Results

- SQL queries aggregated sales data by region and year.

- Complex joins combined economic indicators with sales.

- Results highlighted regions most affected by economic recessions.

7. Interactive Map with Folium

- Folium map visualized sales density by geographic location.

- Color-coded markers indicated sales volumes and trends.

- Interactive popups provided additional sales details per location.

8. Plotly Dash Dashboard Results

- Dashboard included filters for vehicle types, years, and regions.

- Interactive graphs updated dynamically based on user selection.

- Provided a comprehensive tool for stakeholders to explore sales data.

9. Predictive Analysis Methodology

- Classification models were developed to predict sales category based on features.

- Data preprocessing included normalization and encoding.

- Models used included Logistic Regression and Random Forest.

Capstone Project Presentation

10. Predictive Analysis Results

- Model accuracy reached approximately 85%.

- Confusion matrix indicated balanced precision and recall.

- Feature importance analysis showed advertising spend and GDP as key predictors.

11. Conclusion

- The analysis provided actionable insights on automobile sales trends.

- Economic indicators strongly correlate with sales performance.

- Future work includes improving model accuracy and expanding datasets.

12. Creativity & Innovation

- Combined multiple data sources and tools (SQL, Folium, Dash) effectively.

- Developed an interactive dashboard enhancing data exploration.

- Provided novel visual insights on the impact of economic cycles.