# **Data Engineering Lead at Craigslist: Used Vehicles Data Strategy**

By: Patriciah Nici Mboya

Date: 18/02/2024

## **Time-Series Model**

### **Data Enrichment Recommendations**

To drive more valuable analyses, there are several areas where additional data can be beneficial. Here are a few recommendations:

1. **Demographic Data**: Including demographic information such as age, gender, income, and education level of vehicle buyers can provide insights into the preferences and purchasing power of different population segments. This data can enable analyses of the correlation between demographics and vehicle preferences, helping to identify target markets and tailor marketing strategies accordingly.

2. **Economic Indicators**: Including economic indicators such as GDP, unemployment rates, and consumer sentiment can provide insights into the overall economic climate and its influence on vehicle sales. This data would enable us to identify correlations between economic factors and vehicle demand, helping us make informed decisions about inventory management and pricing strategies.

3. **Weather Data**: By integrating weather data into the analysis, we can explore the impact of weather conditions on vehicle listings. This data would enable us to identify patterns such as increased demand for certain vehicle types during specific weather conditions (e.g., SUVs during winter) and adjust pricing or marketing strategies accordingly.

4. **Competitor Data:** Gathering data on competitor listings, pricing, and customer reviews can provide valuable insights into market dynamics and customer preferences. By analyzing this data, we can identify opportunities to differentiate our offerings, optimize pricing, and improve customer satisfaction.

5. **Vehicle Features**: Incorporating detailed information about vehicle features, such as engine type, fuel efficiency, safety ratings, and technology options, can facilitate analyses of the popularity of specific features and their impact on vehicle demand. This data can help identify trends in consumer preferences and guide product development decisions.

By incorporating these additional data sources into the analysis, Craigslist can better understand the used vehicle market, identify emerging trends, and make data-driven decisions to optimize its platform and services.

### **Data Warehouse Structure**

This proposed data warehouse structure integrates the existing dataset with the additional data sources mentioned earlier. It consists of a fact table, **vehicle\_listings**, which contains the primary data related to vehicle listings. The dimension tables provide additional context and allow for efficient querying and analysis.

The **regions table** stores information about different regions, enabling segmentation and analysis based on geographical factors. The vehicle\_types table categorizes vehicles, facilitating analysis by vehicle type.

The **weather\_data** table contains weather-related information, allowing us to analyze weather patterns and their impact on vehicle listings. The economic\_indicators table stores economic data, enabling analysis of the relationship between economic factors and vehicle demand.

The **economic indicator** table captures economic data such as GDP, unemployment rates, and consumer sentiment. This dimension contributes valuable insights into market dynamics and customer preferences helping to identify competitive advantages and areas for improvement.

Finally, the **competitor\_data** table includes data on competitor listings, pricing, and customer reviews, providing insights into market dynamics and customer preferences.

Appendix

