**Situation:**

Our team, **Team 404. Gas Not Found**, participated in a hackathon hosted by Rice University’s School of Engineering and Computing. The challenge, presented by Chevron, required us to predict the on-road vehicle population for 2024 and analyze future fuel and energy demand. This was a critical task as it would help Chevron understand the evolving trends in fuel consumption and electric vehicle (EV) adoption, which are shaping the future of the energy industry.

**Task:**

The primary task was to develop a predictive model that could accurately forecast the vehicle population for 2024 and beyond. This involved analyzing historical vehicle data, understanding the relationships between various features (e.g., fuel type, vehicle category, and model year), and predicting future trends in fuel and energy demand. Additionally, we needed to ensure that our insights aligned with Chevron’s business goals and provided actionable value.

**Action:**

1. **Data Exploration and Cleaning:**
   * We began by exploring the dataset, which contained information about vehicles registered between 2019 and 2024, including details like model year, vehicle type, fuel type, and the number of vehicles registered at the same address.
   * We cleaned the data by handling missing values, inconsistencies, and performing exploratory data analysis (EDA) using Power BI to uncover key insights and patterns.
2. **Feature Engineering:**
   * We created new features from existing ones to improve model performance. For example, we used one-hot encoding to convert categorical variables like fuel type and fuel technology into numerical values.
   * We also engineered features such as vehicle weight categories (light, medium, heavy) based on the GVWR (Gross Vehicle Weight Rating) class.
3. **Model Selection and Training:**
   * We experimented with several machine learning models, including CatBoost, Random Forest, and XGBoost, to predict the vehicle population.
   * We trained these models using the prepared data and optimized them by scaling and feature selection.
4. **Model Evaluation and Prediction:**
   * We evaluated the models using metrics like RMSE (Root Mean Squared Error) to ensure accuracy.
   * The best-performing models were used to predict the vehicle population for 2024 and future years.
5. **Visualization and Reporting:**
   * We visualized the results using charts and graphs to make the insights more accessible.
   * We prepared a comprehensive report summarizing our findings, including key insights such as:
     + Gas-powered cars still make up 90% of the market, but their share is declining.
     + Economic downturns influence car-buying behavior and fuel technology adoption.
     + Electric vehicles are on the rise, which will impact future fuel demand.
6. **Collaboration and Alignment with Business Goals:**
   * We worked closely with our stakeholder/mentor, Timothy Angeles, to ensure our insights aligned with Chevron’s business objectives.
   * We explored how various features drive fuel and energy demand, particularly in the context of EV adoption and changing fuel consumption trends.

**Result:**

* We successfully built a predictive model that provided valuable insights into future vehicle populations and fuel demand, helping Chevron prepare for industry shifts.

Future Goals to drive more business impact

* Improve RMSE score by using techniques like model stacking, adding features like exponential decay feature, and implementing polynomial feature engineering.