Impact of Gas Prices on Electric Vehicle Adoption

1.Introduction

The automotive industry is undergoing a transformative shift towards electric vehicles (EVs), driven by concerns about environmental sustainability and the rising costs of fossil fuels. This study examines the hypothesis that increasing gas prices encourage EV adoption in the United States due to the potential for cost savings. By employing a data pipeline approach, we analysed the relationship between historical gas prices and EV registrations from 2012 to 2021.

2. Used Data

2.1. Gas Data

- **Source:** U.S. Energy Information Administration (EIA) (link)
- **Format:** Weekly time-series data (XLS)
- **Period:** 1997 to present
- **Description:** Provides weekly gas prices across the United States, allowing for long-term trend analysis.
- **Challenges:**
    - Inconsistent weekly intervals
    - Multiple entries for some weeks

2.2. EV Registration Data

- **Source:** Atlas EV Hub (link)
- **Format:** Event-based data (CSV)
- **Period:** 2012 to 2021
- **Description:** Includes EV registration records, varying by state and time.
- **Challenges:**
    - Aggregation to monthly intervals
    - Missing data for specific states or months

2.3. Data License Compliance

Both datasets are publicly available and used under Open Database licenses

3. Analysis

3.1. Challenges with Data

The primary challenges encountered during the analysis included dealing with inconsistent weekly intervals in gas price data and multiple entries for the same time periods, which necessitated normalization and aggregation steps. The EV registration data presents issues with missing values for certain states and months, requiring careful handling to avoid skewing the analysis. Aligning the timeframes of both datasets to ensure compatibility also posed a technical hurdle.

3.2. Mitigating Challenges through Pipeline

The data pipeline integrates gas prices and EV registration data through the following stages:

3.2.1. Gas Price Transformation:

- To ensure consistency, the gas price data was chronologically sorted, allowing for the identification of deviations from the expected 7-day intervals. This step helped detect and log irregularities in the data.
- Weekly data with multiple entries for the same period was normalized by averaging the values. This process reduced redundancy and ensured a single consistent value for each week.
- Aggregation of weekly gas prices into monthly averages was conducted to align the dataset with the monthly granularity of EV registration data. This step provided a smoother trend for analysis and simplified correlation studies.

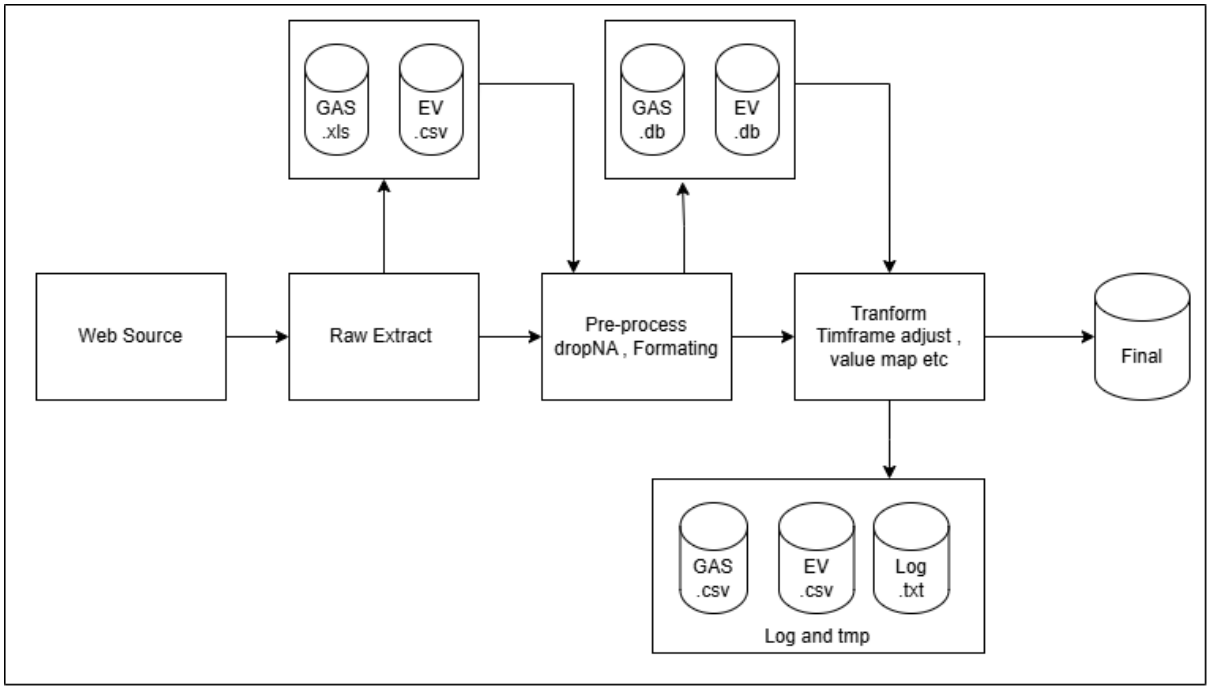
3.2.2. EV Data Transformation

- Relevant columns such as registration date and vehicle name were filtered to retain only necessary data fields, reducing unnecessary complexity and ensuring focus on key metrics.
- The filtered data was aggregated by month to calculate the total number of EV registrations (Aggregate irrespective of vehicle make), providing a clearer view of monthly trends in EV adoption across the analyzed timeframe.
- To facilitate a meaningful comparison with gas price data, the EV registration data was aligned to match the same 2012–2021 timeframe. This step ensured temporal consistency and accuracy in correlational analysis.

3.2.3. Merging Datasets

- The gas price and EV registration datasets were combined by matching their time periods, ensuring that monthly data from both sources aligned. This allowed for a direct comparison between the two variables over the same timeframe.
- Any gaps or missing entries were carefully logged and flagged for future reference. This step ensured transparency in the analysis and provided a clear record of data inconsistencies.
- Rows that lacked data for both gas prices and EV registrations were removed, simplifying the dataset and focusing the analysis on periods with meaningful information. This cleaning step helped ensure the results reflected reliable and complete data.

3.2.4. Overall arch of Pipeline

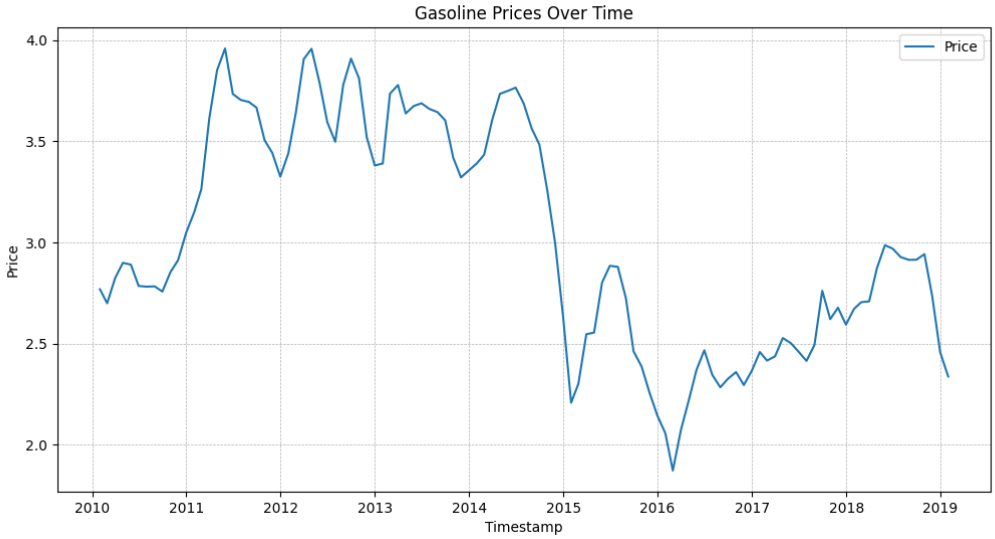


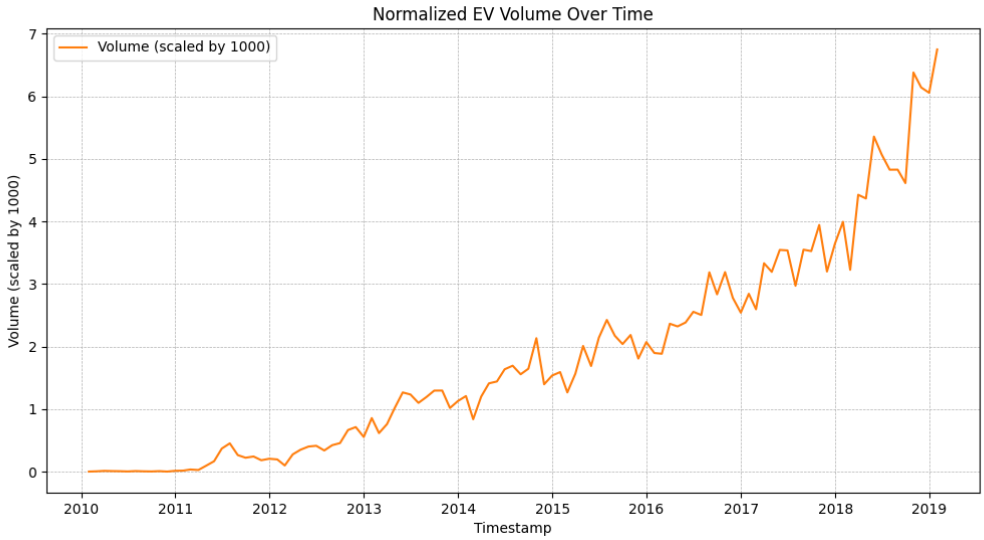
3.2.5. Error Logging and Data Quality Issues

To ensure transparency and reliability, an error logging system was implemented during the data pipeline process. This system flagged missing or inconsistent data points and recorded these issues for future review. Common data quality problems included gaps in weekly gas prices, which were either interpolated or excluded based on the severity of the inconsistency. Similarly, missing entries in EV registration data for certain states were documented to avoid skewing the analysis. The logged errors provided insights into the limitations of the datasets and helped in refining the cleaning and transformation processes.

3.2.6. Analysis Outcomes

The merged dataset was primarily visualized to observe broad trends and relationships between gas prices and EV registrations over time. No extensive analytical techniques were employed; instead, emphasis was placed on normalizing the data and ensuring its integrity for basic trend evaluation. This minimalistic approach allowed for a foundational understanding of the dataset while leaving room for future, more complex analyses.





4. Conclusion

This project primarily focused on processing data to create a reliable and structured foundation for future analysis, rather than performing extensive statistical evaluations. While minimal analysis, such as plotting data trends and visual comparisons, was conducted, the emphasis remained on preparing the data pipeline and ensuring data quality.

The findings indicate that EV sales have a limited impact on gas prices, failing to support the initial hypothesis. However, this conclusion is tentative, as more advanced analysis, incorporating additional factors and methods like ANOVA or nonlinear models, is necessary to provide a comprehensive understanding of the relationship between gas prices and EV adoption.