Electric Vehicle Population Dashboard: Process Book

Overview and Motivation

Project Overview

Electric vehicles (EVs) represent the forefront of sustainable transportation, providing a cleaner alternative to traditional gasoline-powered vehicles. As global interest in renewable energy and eco-friendly technology grows, understanding EV adoption trends is crucial for policymakers, manufacturers, and consumers alike.

This project aims to create a comprehensive dashboard visualizing the distribution, trends, and performance of electric vehicles using static D3.js visualizations. By leveraging a publicly available dataset, the project explores key questions about EV growth, manufacturer dominance, and technological advancements in electric range.

Motivation

The motivation for this project is rooted in addressing knowledge gaps in EV adoption:

- **For Policymakers**: Identifying regions that need infrastructure like charging stations or additional incentives.
- **For Manufacturers**: Understanding market trends and opportunities for innovation.
- **For Consumers**: Providing insights into the most popular or high-performing EV models.

The focus on detailed, aggregated, and interactive-ready visualizations ensures the project is accessible to a broad audience while maintaining analytical depth.

Questions

Initial Questions

1. **How have electric vehicles' ranges evolved over time?**
2. **Which manufacturers dominate the EV market?**
3. **What trends exist in EV adoption across different regions and years?**

Evolved Questions

During the analysis, the following questions emerged:

1. **Are BEVs or PHEVs more prevalent in certain years?**
2. **Which manufacturers consistently lead in electric range performance?**
3. **What patterns can be identified in the growth of EV adoption by year?**

The project evolved to explore both granular details and broader aggregated trends, enabling a multi-level understanding of EV adoption.

Data

Source

The dataset, **Electric Vehicle Population Data**, provides comprehensive details on EV registrations, including attributes such as make, model, type, year, electric range, and geographical location.

Preprocessing

To prepare the dataset for visualization:

1. **Cleaning Missing Values**: Rows with missing `Electric Range` or `Model Year` data were handled by removing or imputing based on averages.
2. **Aggregating Data**: Data was aggregated by key dimensions such as `Model Year`, `Make`, and `Electric Vehicle Type` for high-level visualizations.
3. **Geospatial Formatting**: The `Vehicle Location` column was formatted for use in mapping tools.

Challenges: Missing Data in Electric Range

- **Electric Range Variations**: The **Electric Vehicle Population Data** dataset provided a wealth of information, but it also posed certain challenges, particularly with missing values in the `Electric Range` column. Addressing these missing values was critical for ensuring the reliability and accuracy of the visualizations. Below is a detailed explanation of how this challenge was managed.\

Approaches to Address Missing Data

To handle the missing data, the following strategies were considered and implemented:

A. Removal of Records (For Non-Critical Cases)

- **When Used**: Records were removed only if the missing data was minimal and other attributes (like `Make` or `Model Year`) provided little value for analysis.
- **Reason**: This ensured that incomplete or non-representative data did not distort visualizations.
- **Drawback**: Potential loss of some data points, especially for older or uncommon vehicles.

B. Imputation with Averages (For Heatmap and Scatter Plot)

- Approach
  - Missing `Electric Range`

    values were imputed using the average range of similar vehicles. For example:
    - If a `Make` and `Model` had missing range data, the average range for that `Make` was used.
    - If no `Make`-specific average was available, a global average for the same `Electric Vehicle Type` (BEV or PHEV) was used.

- **Reason**: Imputation preserved the dataset's completeness, ensuring visualizations retained their interpretability.
- **Drawback**: Imputed values might oversimplify actual variations in range performance.

Steps Taken to Handle Missing Data

Data Cleaning Steps

1. **Identify Missing Values**:
    - Rows with `NaN` or blank values in the `Electric Range` column were identified.
    - Summary statistics helped determine the proportion of missing data.
2. **Impute or Remove**:
    - Records missing multiple critical attributes (e.g., `Make` and `Electric Range`) were removed.
    - For remaining records, missing `Electric Range` values were imputed using calculated averages.
3. **Validation**:
    - Post-imputation, scatter plots and heatmaps were re-evaluated to ensure that the imputed data did not introduce bias.

Exploratory Data Analysis

Visualizations

1. **Scatter Plot**: Initial scatter plots revealed clusters of vehicles with similar electric ranges and Model Years.
2. **Bar Chart**: A bar chart depicting EV counts by year highlighted Tesla's dominance in recent years.
3. **Heatmap**: A heatmap provided insights into average electric range performance across manufacturers.

Design Evolution

Initial Sketches

1. Scatter Plot
    - Simple scatter plot for Model Year vs. Electric Range.
    - Tooltips proposed for detailed exploration.
2. Bar Chart
    - Stacked bar chart comparing BEV and PHEV adoption by year.
    - Considered an area chart but opted for clarity in year-over-year trends.
3. Heatmap
    - Heatmap showing average electric range by Make and Model Year.
    - Early designs included a bubble chart but shifted to a heatmap for readability.

Feedback and Iteration

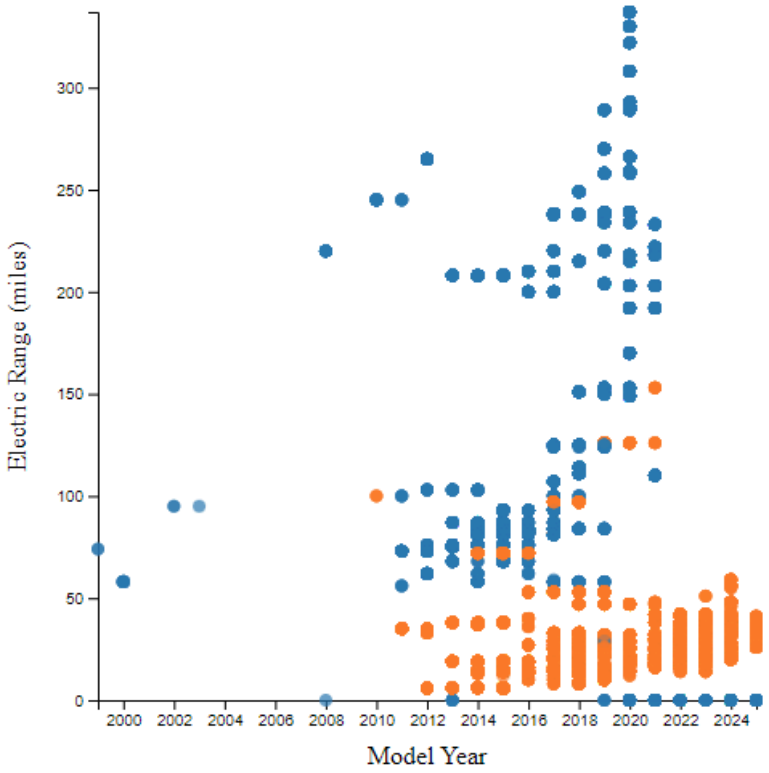
1. **Color Adjustments**: Early versions used excessive colors, leading to confusion. A more subdued color palette improved readability.
2. **Tooltip Enhancement**: Detailed tooltips were added to provide contextual information on hover.
3. **Axes and Labels**: Clear labels and legends were incorporated to ensure self-explanatory visuals.

Final Designs

1. A scatter plot with detailed tooltips for individual vehicles.
2. A bar chart showing yearly adoption trends segmented by vehicle type.
3. A heatmap highlighting manufacturer performance in electric range.

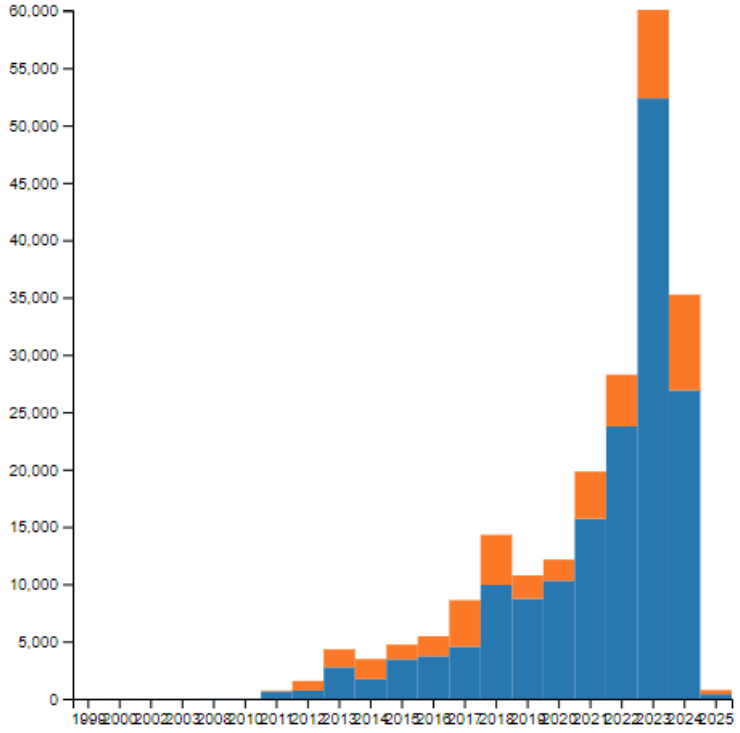
Implementation

Visualization 1: Scatter Plot



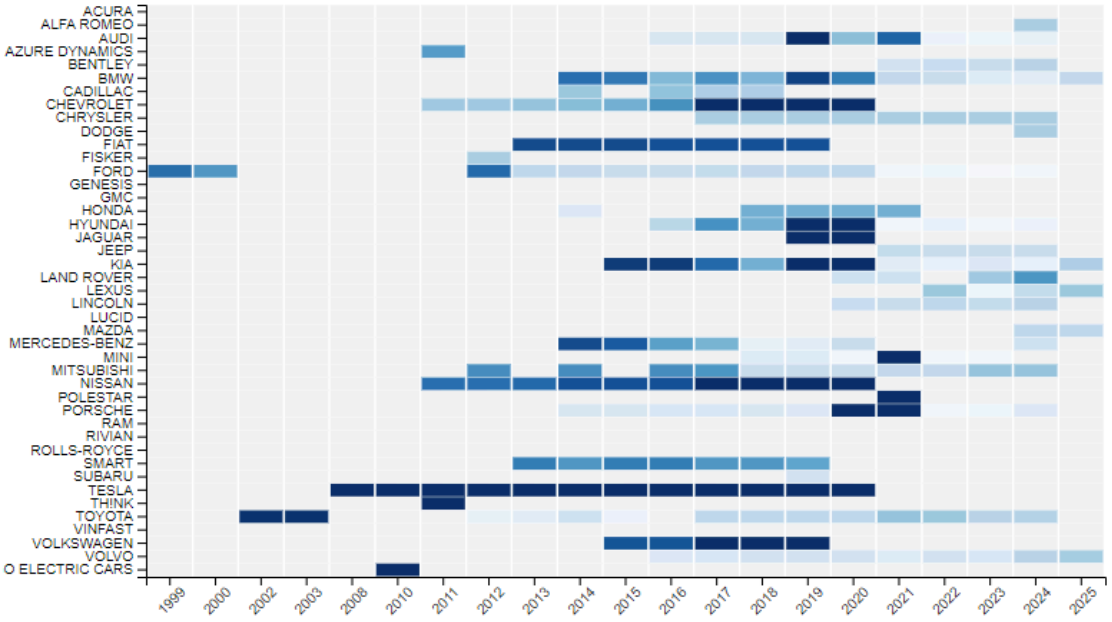
- Attributes
  - X-axis: Model Year
  - Y-axis: Electric Range
  - Color: Electric Vehicle Type
  - Tooltip: Displays Make, Model, Year, Type, and Range
- Expected Insights
  - Technological improvements over time.
  - Distribution of ranges by EV type and year.

Visualization 2: Bar Chart



- Attributes
  - X-axis: Model Year
  - Y-axis: Count of vehicles
  - Color: Vehicle type (BEV vs. PHEV)
- Expected Insights
  - Trends in EV adoption.
  - Growth in BEV dominance over time.

Visualization 3: Heatmap



- Attributes
  - X-axis: Model Year
  - Y-axis: Make
  - Color Intensity: Average Electric Range
- Expected Insights
  - Manufacturers leading in electric range technology.
  - Temporal patterns in range performance.

Improvements

1. **Add Interactivity**: Future versions could allow filtering by region or manufacturer.
2. **Incorporate Additional Data**: Charging station availability and cost data could add context to adoption trends.

Future Work

- Enhance interactivity by allowing filters for years, regions, and manufacturers.