

Data Cleaning Report: Electric Vehicle Population Dataset

1. Column Name Cleaning

Renamed columns by stripping spaces, converting to lowercase with underscores, and removing special characters.

2. Duplicate Removal

Removed duplicate entries based on the 'VIN (1-10)' column to ensure each vehicle is uniquely represented.

3. Handling Missing Values

- 'Postal_Code': Replaced missing values with 'Unknown' and converted to text format.
- 'City' and 'County': Filled missing values with 'Unknown'.

4. Data Type Conversion

Converted the following columns to integers:

- 'Model_Year'
- 'Electric_Range'
- 'Base_MSRP'
- 'Legislative_District'

5. Text Standardization

Standardized text columns to uppercase and stripped whitespaces for:

- 'Make', 'Model', 'Electric_Vehicle_Type', 'Electric_Utility'

6. Splitting 'Vehicle_Location'

Extracted latitude and longitude from the 'Vehicle_Location' column into two new columns: 'Latitude' and 'Longitude'.

7. Columns Dropped

Dropped 'Vehicle_Location' after extracting coordinates to reduce redundancy.