Motorwatt Webscraping

Source: Motorwatt Electric Vehicle Marketplace
Toolchain: Google Colab, Selenium, Python, Pandas , PowerBi

Objective

The goal is to extract detailed electric vehicle data from the **Motorwatt marketplace website** and transform it into a clean, analysis-ready dataset. This enables comprehensive evaluation of electric car specifications, price-performance ratios, battery efficiency, and more for data-driven decision-making and dashboard visualization.

1. Web Scraping Phase

Setup and Configuration

- Used Google Colab as a cloud-based environment for coding and automation.
- Configured Selenium with headless ChromeDriver to handle JavaScript-heavy pages.
- Headless mode allowed the scraper to run efficiently without opening a browser window.

Scraping Strategy

1. Used **Selenium** to navigate through all pagination of the Motorwatt EV marketplace since it is java script heavy .

2. For each **car listing**, extracted the link to its detail page

3. Visited each **detail page** to collect the following fields:

    - **Car Name**
    - **Seller Location**
    - **Phone Number**
    - **E-car Producer**
    - **First Registration**
    - **Price ($)**
    - **Battery Capacity (kWh)**
    - **Engine Capacity (p.h.)**
    - **Charge Range**
    - **Current Mileage**
    - **Max Speed (km/h)**
    - **Car URL**

➢ Successfully scraped detailed data from **all available electric car listings**

➢ Exported raw data to electric_cars_all_pages.csv

2. Data Cleaning & Preprocessing

Cleaning Steps

- **Duplicate Removal**: Removed duplicate car entries
- **Column Renaming**: Renamed "Detailed Car Title" → "Car name"
- **Text Cleaning**:
    - Removed the word "Specification" from car names
    - Cleaned inconsistent phone number formats
- **Handling Missing Values**:
    - Filled missing **Seller Location** with "Unknown"

    - Replaced missing **E-car Producers** or "other" with inferred values from car names

- **Numerical Cleaning & Standardization**:

    - Converted prices, engine capacity, max speed, mileage, battery, and charge rates to numeric formats
    - Handled ranges like "100-120" by computing the average
    - Replaced invalid or missing entries with column-wise averages

Final Cleaning Output

- Exported cleaned data to cleaned_electric_cars_final.csv

3. Feature Engineering

- **Battery Efficiency** = Battery (kWh) / Engine Capacity (p.h.)
- **Price per kWh** = Price ($) / Battery (kWh)
- **Performance Index** = (Battery * Engine Capacity) / Price
- **Car Age** = 2025 - First Registration
- **Engine Category** = Binned into: Small, Medium, High, Extreme (based on engine capacity)

And performed some transformation needed for PowerBi;

- Transformed decimal values in key metrics (Battery Efficiency, Price per kWh, Performance Index) to European format (comma as decimal separator) so powerbi can work with the data .
- Created additional file electric_cars_dashboard_ready.csv with derived metrics like:
    - **MilesPerDollar** and **RangePerKWh**
- Saved as electric_cars_decimal_transformed.csv

3. Visualization

1. Key Metrics

- **Total Cars:** 278 electric vehicles are listed in the dataset.
- **Average Engine Capacity:** 343.65 p.h. indicates the average power capacity across the EVs in this marketplace .
- **Average Price:** $70,170 is the average selling price of the EVs.
- **Average Max Speed:** 186.60 km/h is the average top speed among the cars.
- **Total Sellers:** 44 different sellers are included in the data.
- **Average Current Mileage:** 5,750 km is the average mileage of the EVs currently available.

2. Filters and Slicers

- **First Registration Year:** Allows filtering cars by their year of first registration starting from 2016 to 2024 .
- **Engine Category:** Enables filtering by engine type or performance class.
- **Price Range:** Users can select a price range between $4,600 and $317,000 to filter the listings.
- **E-car Producer:** Dropdown to filter data by car manufacturer.

3. Battery Efficiency per Car

- This bar chart compares individual car models based on their **battery efficiency**.

- The chart helps identify which models are **more energy-efficient**.

---

4. Average Performance Index per Car

- Displays the average performance index for different car models.

- **Top performers** include: Yuanhang Y6, Li Xiang Auto, and NWONE, suggesting strong overall capability.

5. Seller Location Map

- Since the website is an international marketplace , A visual map was created to showi where sellers are located globally .
- Helps identify **geographical distribution** and possible regions with higher EV availability.

6. Best Value of Money Table

- Ranks **E-car producers** by their **average price per kWh**.

- Highlights cost-efficiency:

    - **Porsche** has the highest average cost per kWh.

    - **Total Motors** offers the lowest, indicating best value for budget-conscious buyers.

3. Conclusion

This project successfully collected and analyzed real-world electric vehicle data from the Motorwatt website. Using web scraping, data cleaning, and feature engineering, we built a clean and useful dataset.

The final Power BI dashboard helps users explore EV listings, compare prices, check performance, and find the best value. Overall, it shows how data can make car buying smarter and easier.