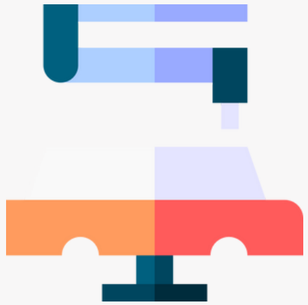
Fuel Economy Data Analysis

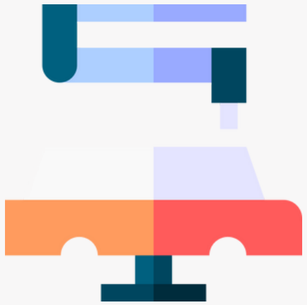
Yenying Chen

Report Overview



- Data Processing & Cleaning
- Key Findings & Insights
- Customer Segmentation Analysis
- Future Directions

Data Preprocessing & Cleaning



Dataset

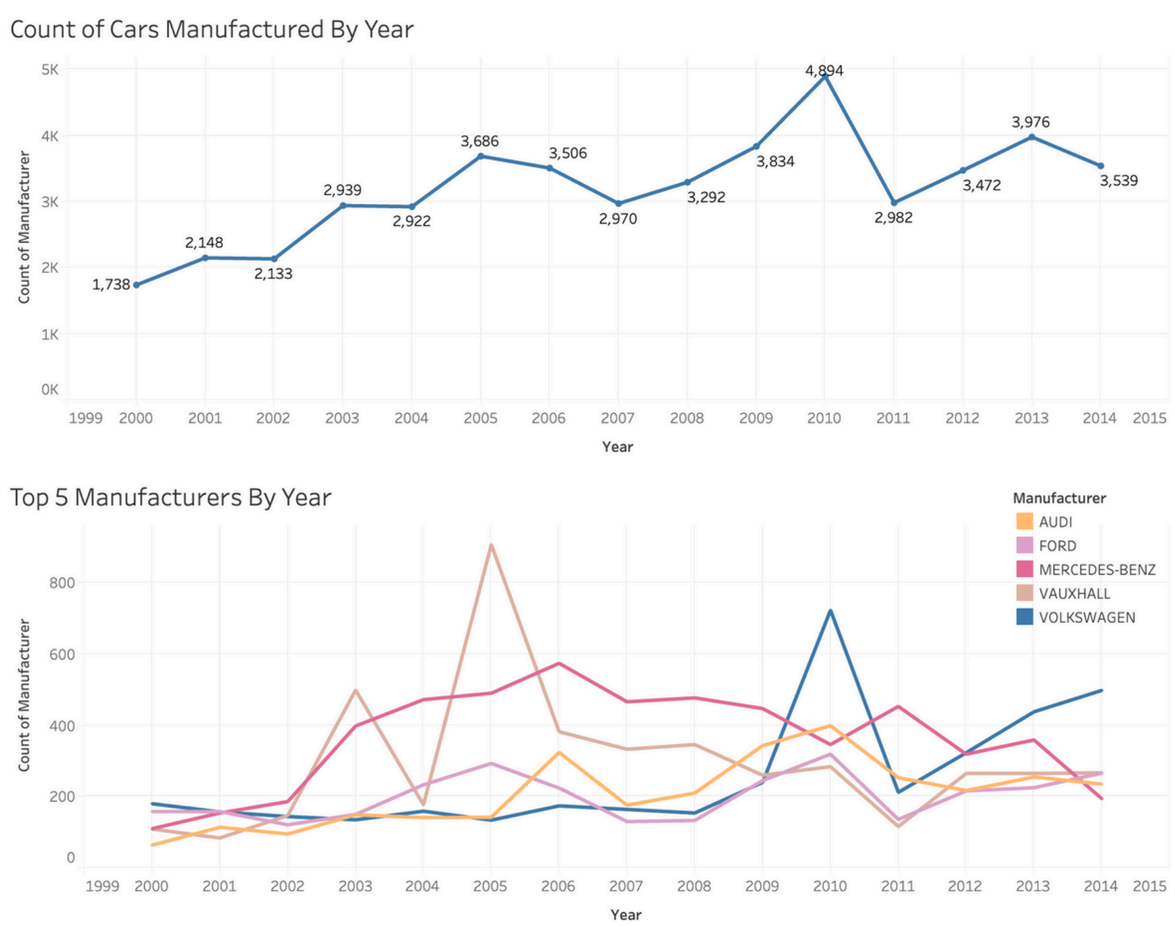
- Fuel economy data from the Vehicle Certification Agency (VCA) covering 2000-2014.
- shape: 23 columns, 48,071 rows

Preprocessing & Cleaning

- Renamed columns for Google BigQuery.
- Dropped unnecessary columns with high missing values.
- Removed duplicates (e.g., Euro Standard 1).
- Handled missing values.
- Converted data types (object → numeric).
- Adjusted Noise_Level_dBA.
- Focus on metric system (aligned with imperial).
- Created new column transmission_type based on transmission.
- Used linear regression to impute missing values for Fuel_Cost_6000_Miles using Metric_Combined.

Cleaned Columns

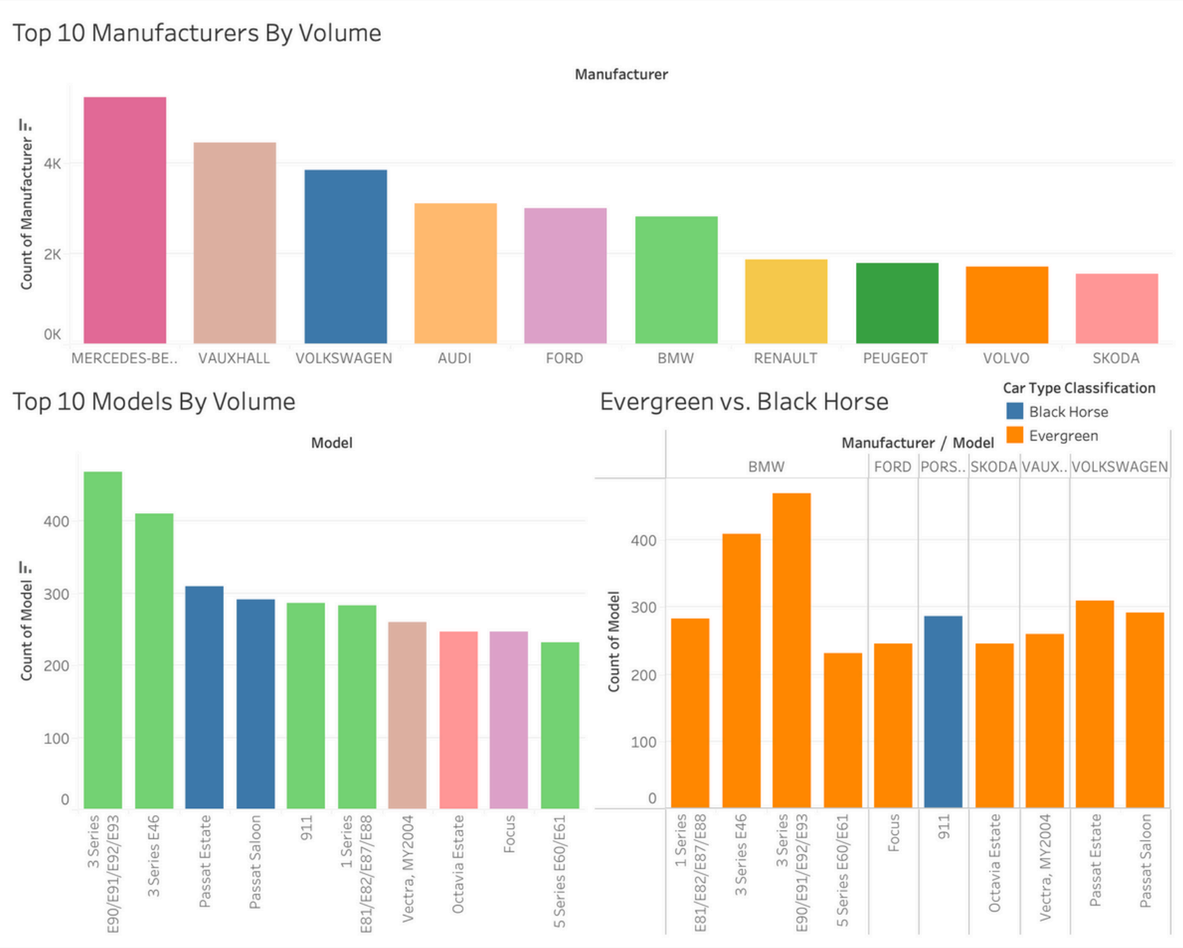
'year', 'Euro_Standard', 'Manufacturer', 'Model', 'Description',
'Transmission', 'Engine_Capacity', 'Fuel_Type',
'Metric_Urban_cold', 'Metric_Extra-urban',
'Metric_Combined',
'CO2_g/km','Fuel_Cost_6000_Miles', 'Noise_Level_dBA'



Finding1 – Market Trends

Insight

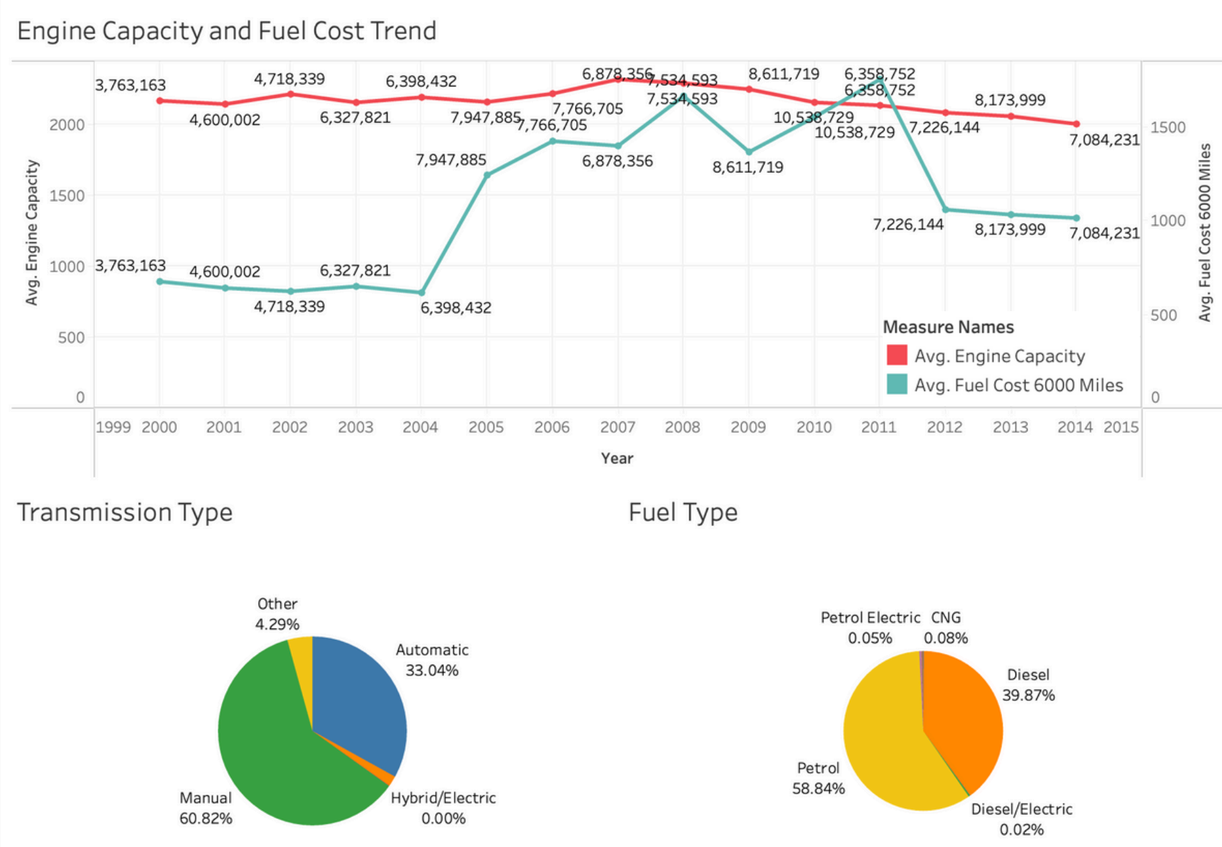
- Vehicle production peaked in **2010**, followed by a fluctuating decline, possibly due to economic downturns, policy changes, or shifts in consumer demand.
- Individual manufacturers showed **varying trends**, indicating differences in market strategy, model popularity, and global expansion efforts.



Finding2 – Manufacturers & Models

Insight

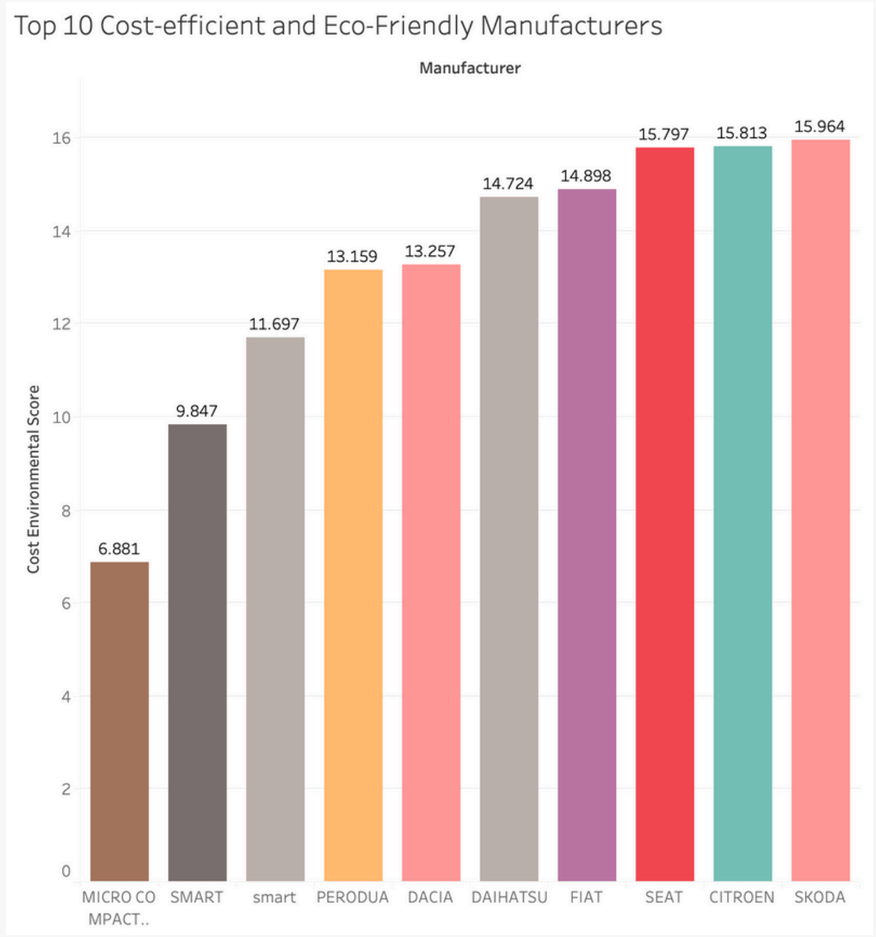
- **Mercedes-Benz, Vauxhall, and Volkswagen** dominate the market in volume, showing strong brand loyalty and widespread adoption.
- **BMW's 3 Series and Volkswagen**'s Passat are "**Evergreen**" models, maintaining steady popularity, while the **Porsche 911** stands out as a "**Black Horse**", achieving exceptional success despite lower overall manufacturer volume.



Finding3 – Performance

Insight

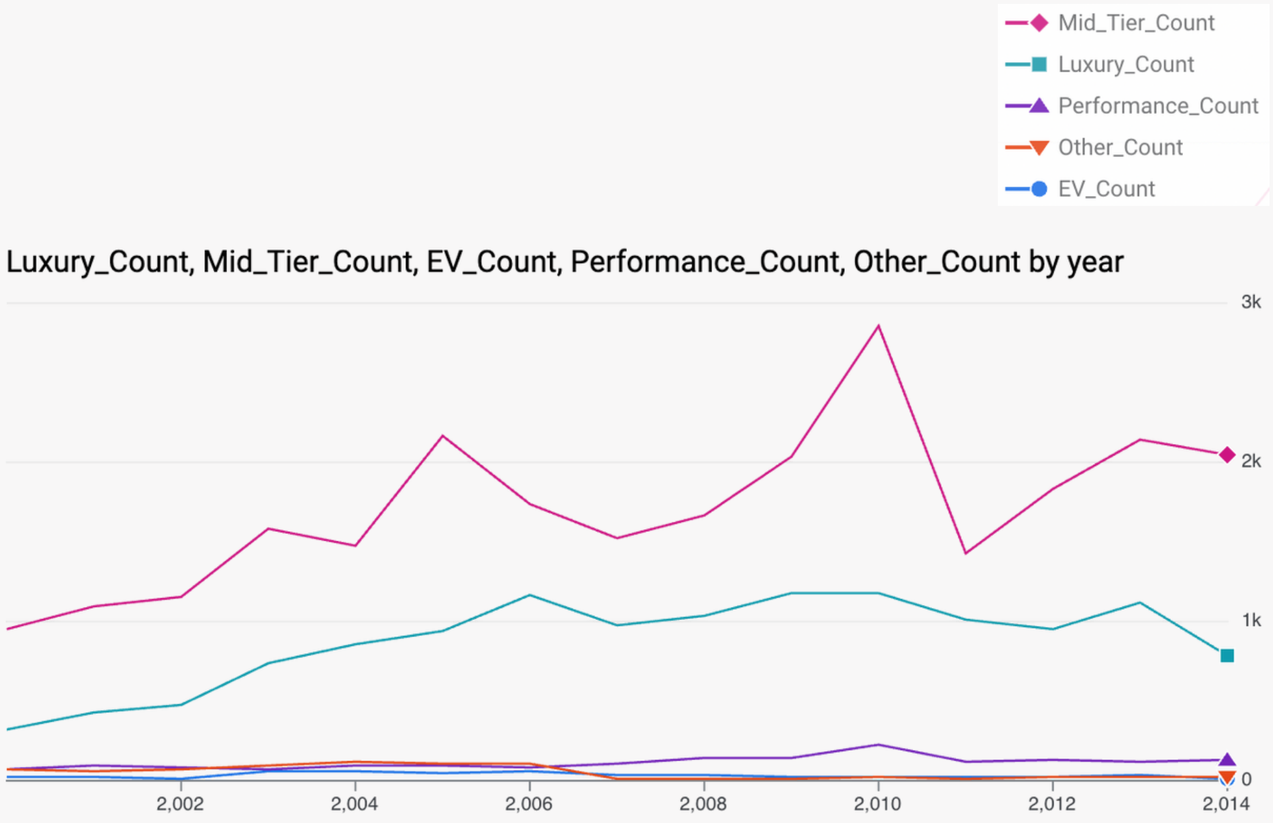
- Engine capacity remains stable, and fuel costs show no clear decline, indicating limited cost reductions.
- Manual (60.82%) and automatic (33.04%) transmissions both maintain strong presence.
- Petrol (58.84%) and diesel (39.87%) dominate, with minimal hybrid/electric adoption.



Finding4 – Most Cost-Efficient & Eco-Friendly Brands

Insight

- Consider CO2 g/km and Fuel Cost 6000 Miles to calculate Cost Environmental Score.
- **Micro Compact, Smart, and Perodua** lead in cost-efficiency and eco-friendliness with low fuel costs and CO_2 emissions.
- European brands like Citroën, Fiat, and Skoda rank high, reflecting **regional** efficiency trends.



Finding5 – Vehicle Segments

Insight

- Reclassify manufacturers.
- **Mid-tier** vehicles remain the most produced segment, showing stable demand.
- EV production remains low during the period.

Customer Segmentation Analysis

	Characteristics	Strategies
Segment 1	• **Mass-market cars**, likely compact or mid-size vehicles from popular brands.	• Highlight cost-effectiveness, reliability, and maintenance benefits.
Segment 2	• **Luxury** sedans & SUVs, including premium hybrid & **diesel** vehicles.	• Promote comfort, advanced technology.
Segment 3	• **Eco-friendly** budget cars with small engines, high fuel efficiency, and low CO2 emissions.	• Best fuel economy, low running costs, and government incentives for fuel-efficient cars.
Segment 4	• High-end sports cars & **performance** vehicles, **large engine sizes, high emissions.Interested in EV.**	• **Price is not the main concern**—focus on exclusivity & brand image.

Future Directions

Improving Market Trend Analysis

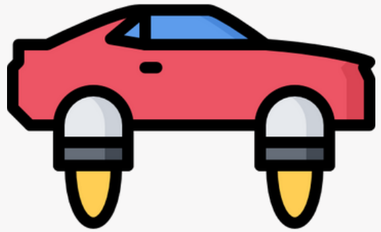
- Incorporate **economic indicators** (e.g., fuel prices, policy shifts) to explain production fluctuations**.**
- Analyze **manufacturer strategies** to understand shifts in market dominance.

Refining Performance & Efficiency Insights

- Explore how engine downsizing and hybrid adoption impact fuel costs.
- Investigate whether emission **regulations** correlate with declining performance car production.

Enhancing Clustering Accuracy

- Improve **feature selection** with fuel efficiency ratios.
- Validate clusters with silhouette scores & PCA visualizations.



Exploring Alternative Clustering Techniques

- Compare K-Means vs. Hierarchical vs. DBSCAN to find better segmentation.
- Detect **outliers** or unique sub-groups in fuel economy.

Predicting Future Market Trends

- Apply **time-series forecasting** on fuel costs to predict future demand.
- Identify which car segments (EVs, Diesel, Budget Cars) will dominate.

Thank you!

Appendix

SQL Query

Python Code

Tableau Dashboard