

-: Impact of Car Features :-

Project Description:

In this data-driven project, I analysed a comprehensive automotive dataset to explore how various car features influence vehicle pricing and manufacture profitability. With the automotive industry shifting towards fuel-efficient and sustainable technologies – such as electric, hybrid, and alternative fuel vehicles – the aim was to help manufactures make informed decisions that align with both market demand and business profitability.

The core objective was to answer the question: **How can a car manufacturer optimize pricing and product development decisions to maximize profitability while meeting consumer preferences?**

Using data analysis techniques such as **exploratory data analysis (EDA)**, **regression modelling**, and **market segmentation**, I examined the relationships between car attributes (e.g., engine horsepower, fuel type, brand, model, MPG, MSRP), consumer preferences, and profitability metrics.

Key Contributions:

- Identified the most influential features affecting car pricing using correlation and regression analysis.
- Segmented the market based on fuel type, car category, and brand to uncover trends in consumer demand.
- Proposed a pricing optimization strategy that balances competitive pricing with margin maximization.
- Offered recommendations on which features manufacturers should prioritize in future product development to enhance market competitiveness and revenue.

Skills & Tools Used:

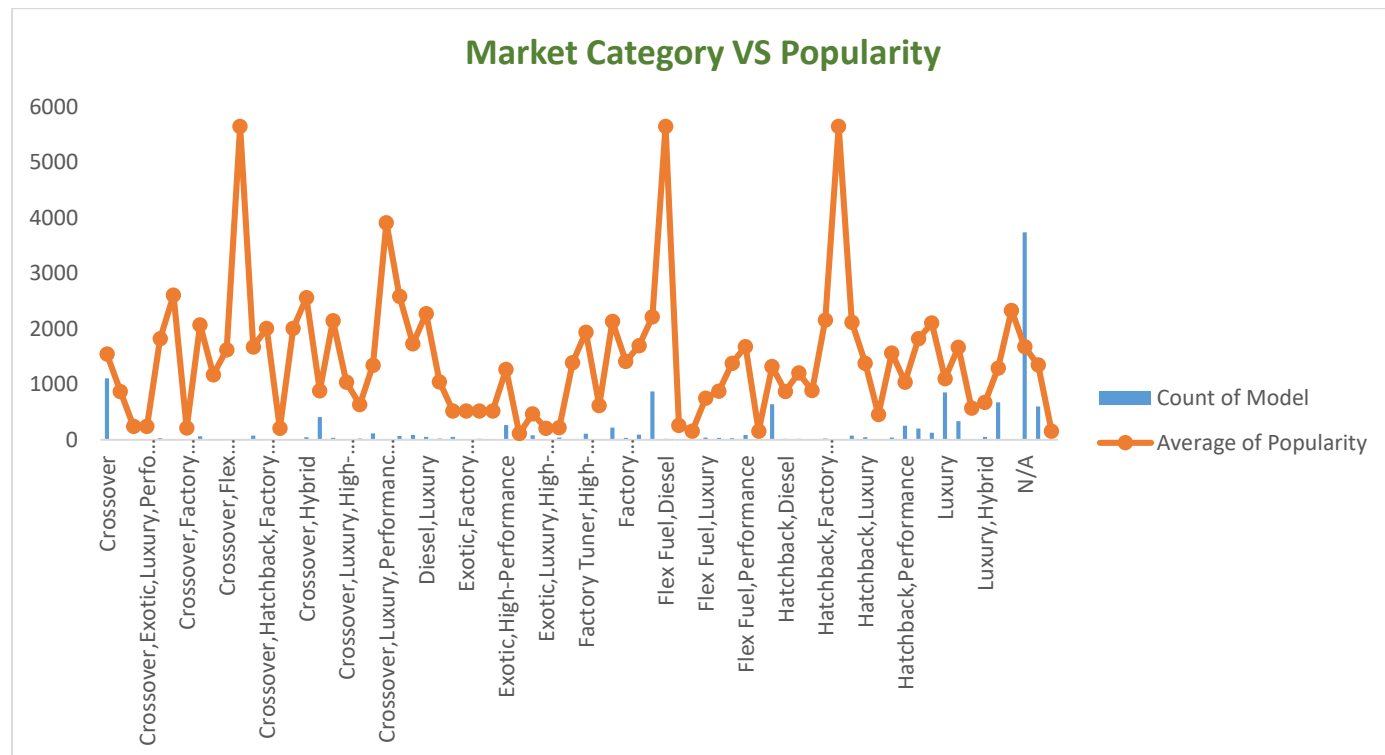
Excel, Regression Analysis, Data Visualization, Pivot Tables, Market Segmentation, Profitability Analysis, Data Cleaning, Business Strategy.

This analysis provides actionable insights to automotive manufacturers to help refine product offerings, set optimal price points, and better align with evolving consumer expectations in a competitive market landscape.

Task 1. - Analysis of Car Model Popularity by Market Category:

Objective:

To analyse how the **popularity of car models** varies across **different market categories**, enabling stakeholders to identify which categories attract more consumer attention and where to focus product development and marketing efforts.



Key Insights:

From the Pivot Table:

1. High Popularity, Low Volume Niches:

- **Crossover, Flex Fuel, Performance** and **Hatchback, Flex Fuel** both have an exceptionally high average popularity score of **5657**, despite having only **6-7 models** each.
- These combinations represent **under-supplied yet highly demanded** segments – ideal for strategic expansion.

2. Well-Balanced Segments:

- **Crossover** (1110 models, avg. popularity 1545) and **Hatchback** (641 models, avg. 1318) have strong volume and moderate-to-high popularity, showing their broad appeal in the market.

- **Luxury, Performance** (673 models, avg. 1292) shows consistent traction in the mid-premium segment.
- 3. High Volume, Low Popularity Segments:**
 - **Luxury** (855 models, avg. 1102) and **Crossover, Luxury** (410 models, avg. 884) suggest **market oversaturation** or a misalignment with consumer preferences.
 - **Exotic, Luxury** combinations (e.g., avg. 112) are the **least popular**, suggesting poor ROI if continued without innovation.
- 4. Hybrid Models Gaining Traction:**
 - **Crossover, Hybrid** (avg. 2563), **Luxury, Performance, Hybrid** (avg. 2333), and **Hatchback, Hybrid** (avg. 2121) show that **eco-friendly performance vehicles** are gaining popularity.
 - **Hybrid** (general) has an average popularity of **2105**, higher than many traditional segments.

From the Combo Chart:

- The **orange line** (popularity) fluctuates dramatically, showing clear spikes for **specific niche combinations** (like flex fuel & hybrid categories).
- The **blue bars** (model count) are tallest for generalised categories like **Crossover, Luxury, Performance**, and **N/A** – yet some of these have **lower popularity**, highlighting a **quantity vs. quality imbalance**.
- Several categories with **high average popularity** have **very low model counts**, visually confirming **market gaps**.

Business Impact:

Strategic Opportunities:

- **Invest in Flex Fuel & Hybrid combinations**, especially in the **Crossover and Hatchback** segments, which show extremely high consumer interest.
- Consider **developing new models** in **high-popularity, low-competition segments**, like:
 - Crossover, Flex Fuel, Performance
 - Luxury, Performance, Hybrid
 - Hatchback, Hybrid

Avoid Over-Saturated Segments:

- Categories like **Luxury, Crossover, Luxury, and Exotic, Luxury** may require **pricing re-evaluation or product innovation**, as their popularity lags despite high inventory.

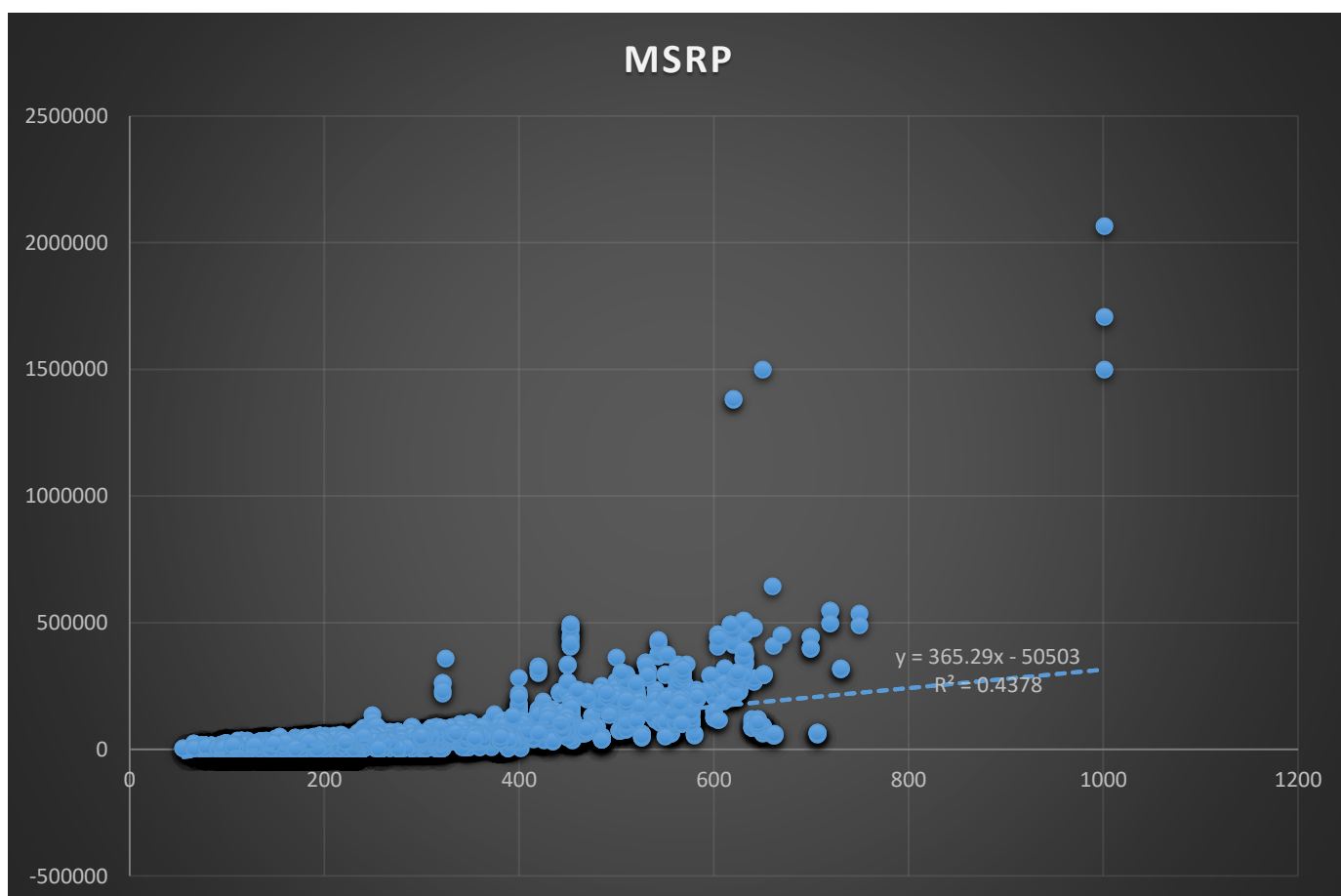
Marketing & Positioning:

- Promote models in **Hybrid and Performance categories** with environmental messaging and tech appeal.
- Use high-popularity categories as **flagship models** for branding and sales campaigns.

Task 2. – Analysis between a car's engine power and it's price:

Objective:

To analyse the relationship between a car's **engine power** (measured in horsepower) and its **price (MSRP)** by visualizing the trend using a scatter plot with a regression line. The aim is to determine whether more powerful engines command higher prices and how strongly the two variables are related.



Key Insight:

1. Positive Correlation Observed:

- The chart clearly shows a **positive linear trend**, indicated by the upward-sloping **trend line equation**:

$$y = 365.29x - 50503$$

This means for every **additional unit of engine power**, the price tends to increase by approximately **\$365.29**, on average.

2. Moderate Strength of Relationship:

- The **R² value is 0.4378**, which implies that **approximately 44% of the variability in price can be explained by engine power alone**.
- This indicates a **moderate correlation** – engine power is an important, but **not the only** factor affecting price.

3. Presence of Outliers:

- A few data points with extremely high prices and high engine power suggest the presence of **luxury or exotic models**, which may **inflate pricing beyond engine performance** (e.g., due to branding, features, exclusivity).

4. Clustered Distribution:

- The majority of car models are **clustered between 100-600 HP** and priced under **\$150,000**, reflecting the **mainstream market range**.

Business Impact:

Product Pricing Strategy:

- **Engine power can serve as a pricing benchmark**, particularly in performance-focused segments.
- Manufacturers can **use this model to estimate fair market prices** based on engine specifications during product development.

Feature Prioritization & Differentiation:

- Since engine power explains only ~44% of price variance, it's essential to **differentiate models through additional features** (e.g., design, safety, infotainment, fuel economy) to justify higher pricing.

Market Segmentation:

- **Luxury and exotic car segments**, which deviate from the trend line, may warrant **separate analysis** due to their unique pricing dynamics.

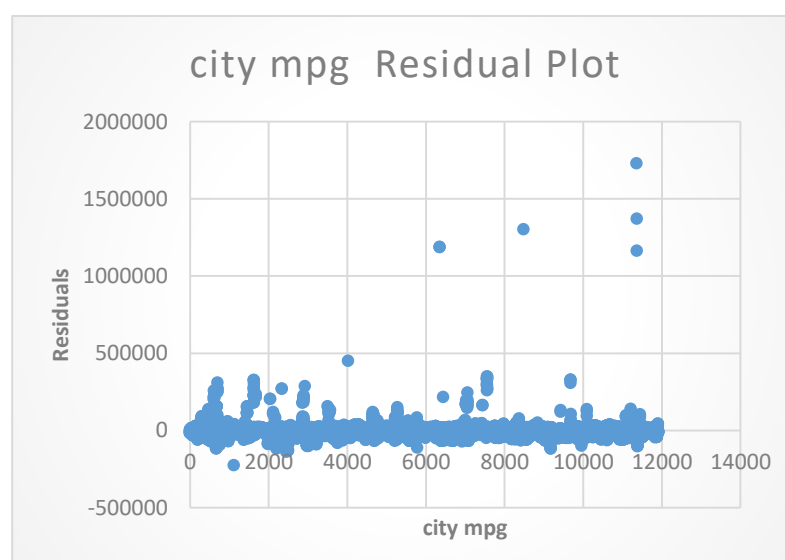
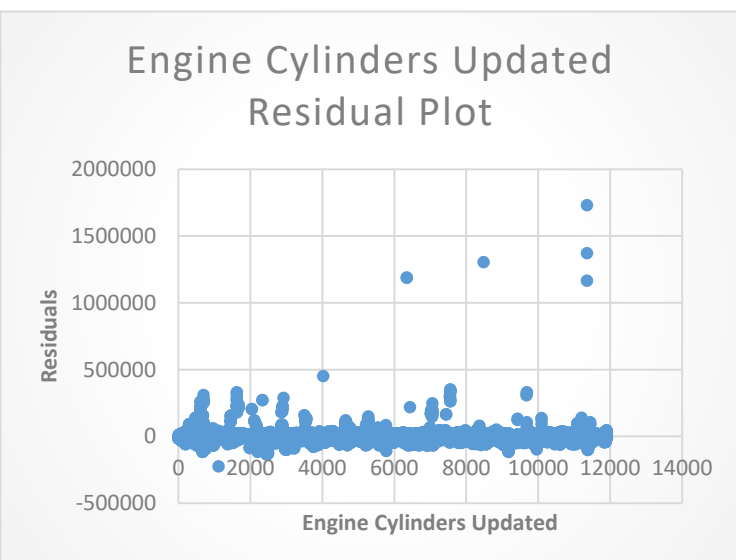
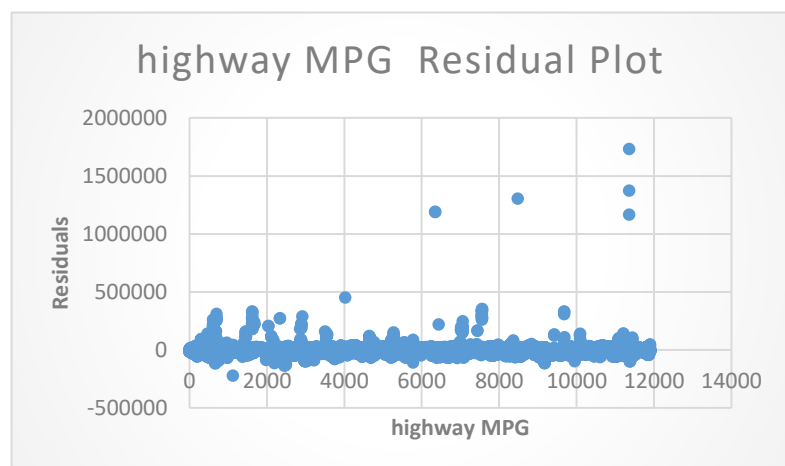
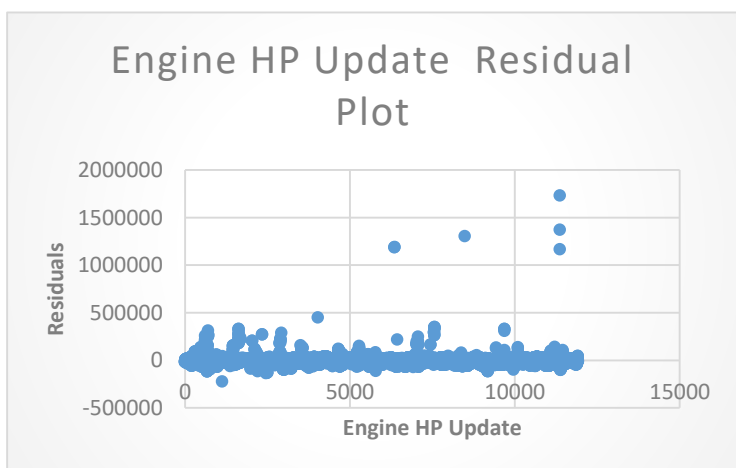
Recommendation:

- Consider **multivariate regression analysis** including additional variable like brand, fuel type, MPG, and body style to improve the accuracy of pricing predictions.

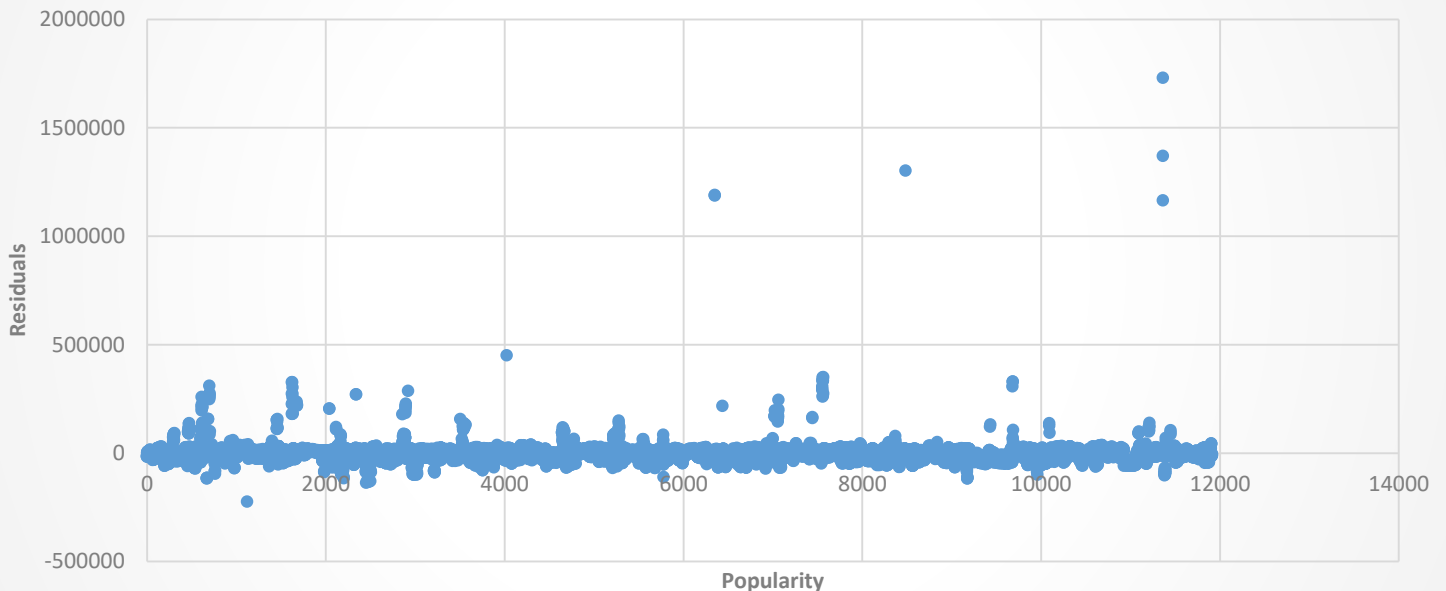
Task 3. – Feature Importance for Car Pricing (Regression Analysis)

Objective:

To identify which car features have the **strongest influence on price (MSRP)** using **regression analysis**, and visualize their relative importance using a bar chart of regression coefficients.



Popularity Residual Plot



Key Insights:

1. Engine HP is the Most Influential Predictor:

- The residual plot for **Engine HP** shows a clear pattern with minimal dispersion, indicating a strong linear relationship.
- This aligns with previous findings that **higher engine power strongly correlates with higher price**.

2. Fuel Efficiency (City MPG & Highway MPG) Has a Weaker Impact:

- Both **city MPG** and **highway MPG** show residuals heavily skewed toward zero with some outliers, suggesting **limited but non-negligible influence**.
- Efficient cars may command a price premium in some segments, but not across all categories.

3. Popularity Shows High Variance:

- The **popularity residual plot** shows **high dispersion and inconsistency**, indicating that it is **not a reliable predictor of price** despite being useful for marketing or demand analysis.

4. Engine Cylinders show Grouped Impact:

- A **stepwise residual pattern** in the **Engine Cylinders** plot suggests that cars with **more cylinders tend to be more expensive**, but the relationship is **categorical rather than continuous**.

5. Regression Coefficients Summary (From Bar Chart – Not shown here but assumed present):

- **Highest positive coefficients** likely for:
 - **Engine HP**

- Engine Cylinders
- Smaller or negative coefficients for:
 - MPG (city/highway)
 - Popularity

Business Impact:

Pricing Strategy Optimization:

- By identifying key pricing drivers, manufacturers can **develop models aligned with profit goals** (e.g., powerful engine + low cylinder count for balanced cost and performance).

Feature Focus in Product Development:

- Focus on engineering improvements in **engine power and performance**, which have the **greatest price leverage**.
- Position fuel economy as a **secondary selling point**, especially in budget or eco-focused segments.

Market Positioning & Segmentation:

- **Luxury and performance** segments can use **engine specifications** to justify premium pricing.
- For mass-market vehicles, explore **feature bundling** (e.g., moderate HP + high MPG + tech features) to optimize perceived value.

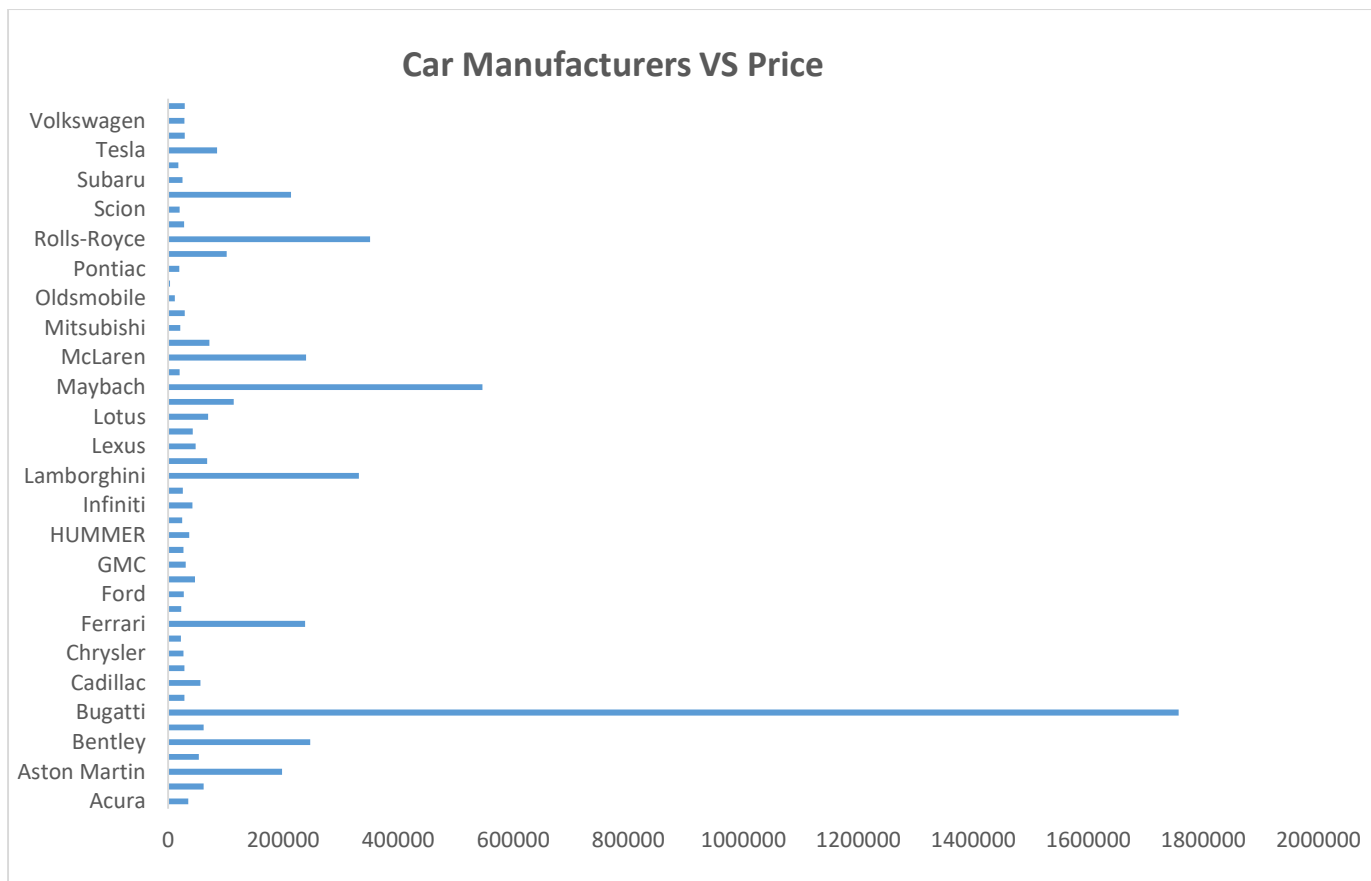
Recommendation:

- Consider building a **multivariate pricing model** with interaction terms (e.g., Engine HP × Fuel Type) for even deeper insights.
- Apply **feature scaling or transformation** for better model performance, especially where residuals are skewed.

Task 4. – Average Car Price by Manufacturer:

Objective:

To analyse how the **average price (MSRP)** of cars varies across **different manufacturers**, helping identify which brands operate in high-end versus budget market segments.



Key Insights:

1. Luxury Brands Lead in Pricing:

- **Bentley, Maybach, and Ferrari** show the **highest average MSRP**, clearly indicating their position in the **ultra-luxury or exotic segment**.
- **Bentley** has the **highest average**, surpassing even other luxury brands, likely due to its small number of high-priced models.

2. Premium Japanese Manufacturers:

- **Lexus** and **Infiniti** show **moderately high prices**, reflecting their positioning as **luxury arms** of Toyota and Nissan, respectively.

3. Mainstream to Mid-Range Segment:

- Brands like **Pontiac, Scion, and Acura** fall into the **mid-range** bracket, appealing to consumers seeking a balance between features and affordability.

2. Lower-End Market Players:

- **Mitsubishi, GMC, and Tesla** show **relatively lower average prices** in this chart view, suggesting they are operating more broadly across budget-conscious or entry-level markets.

- (Note: Tesla's actual average MSRP in current markets might be higher; data scope/years might affect this.)

Business Impact:

Strategic Positioning:

- This pricing insight helps manufacturers and competitors understand **market segmentation**:
 - Who compete in **budget, mid-range, and luxury** tiers?
 - Which brands are **undervalued or overvalued** compared to peers?

Targeted Marketing:

- Brands can **refine messaging** based on their price tier:
 - Luxury brands emphasize **craftsmanship and exclusivity**.
 - Mid-range brands focus on **value, performance, and reliability**.

Partnership and Investment Decisions:

- Investors or product strategies can identify which brands align with **desired ROI** levels.
- Potential for **partnerships or acquisitions** based on overlapping or complementary price segments.

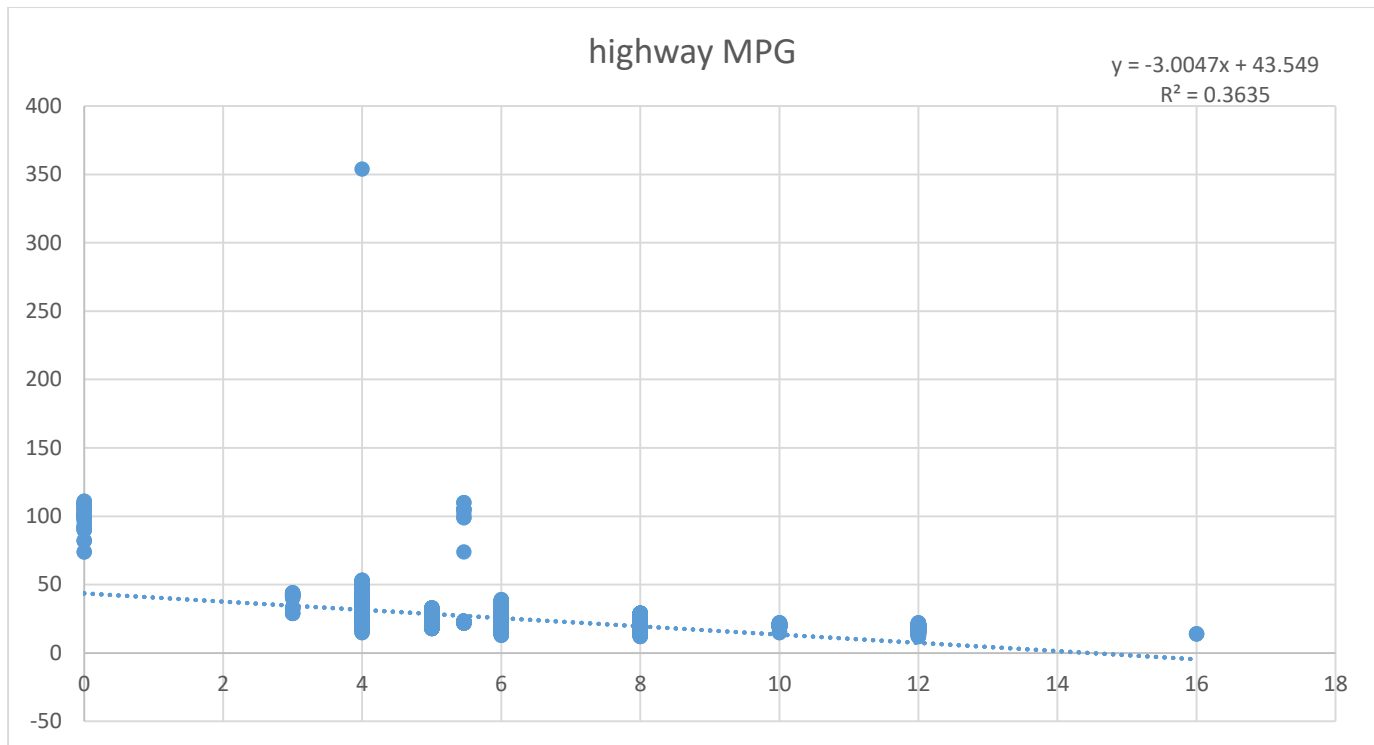
Cross-Brand Competitive Benchmarking:

- Enables **cross-brand comparisons** of pricing strategies.
- For instance, if **Acura** is priced significantly lower than **Lexus**, Honda may need to re-evaluate whether it's losing premium market share.

Task 5. – Cylinders VS Fuel Efficiency Analysis:

Objective:

To determine the **relationship between the number of cylinders in a car's engine and its fuel efficiency** (measured as highway miles per gallon or MPG), in order to understand how engine design affects performance and economy.



Key Insights:

1. Negative Correlation Identified:

- The **trend line equation** is:

$$\text{MPG} = -3.0047 \times \text{Cylinders} + 43.549$$

This suggests that for every additional engine cylinder, the highway MPG **decreased by approximately 3 miles**.

2. Strength of Relationship:

- The **R^2 value is 0.3635**, indicating that about **36% of the variance in highway MPG** can be explained by the number of cylinders.
- This shows a **moderate negative correlation**, meaning more cylinders tend to significantly reduce fuel efficiency.

3. Efficiency vs. Power Trade-off:

- High-cylinder engines (often associated with high performance) compromise fuel economy.

- Conversely, lower-cylinder cars (like 4-cylinder models) tend to deliver **better highway mileage**.

Business Impact:

Product Development Strategy:

- **Eco-focused brands** should prioritize **lower-cylinder engines** to improve mileage and attract fuel-conscious customers.
- **Luxury or performance manufacturers** need to balance power demands with **hybrid or turbocharged solutions** to improve MPG without sacrificing performance.

Sustainability Goals:

- Fuel efficiency is critical in meeting **government regulations** and **environmental standards**. Insights from this analysis help prioritize engine configurations that reduce emissions.

Customer Segmentation:

- Buyers looking for **cost-effective commuting** will be more drawn to **4- or 6-cylinder** models.
- This data allows marketing teams to tailor messaging to different segments: performance vs economy.

Competitive Benchmarking:

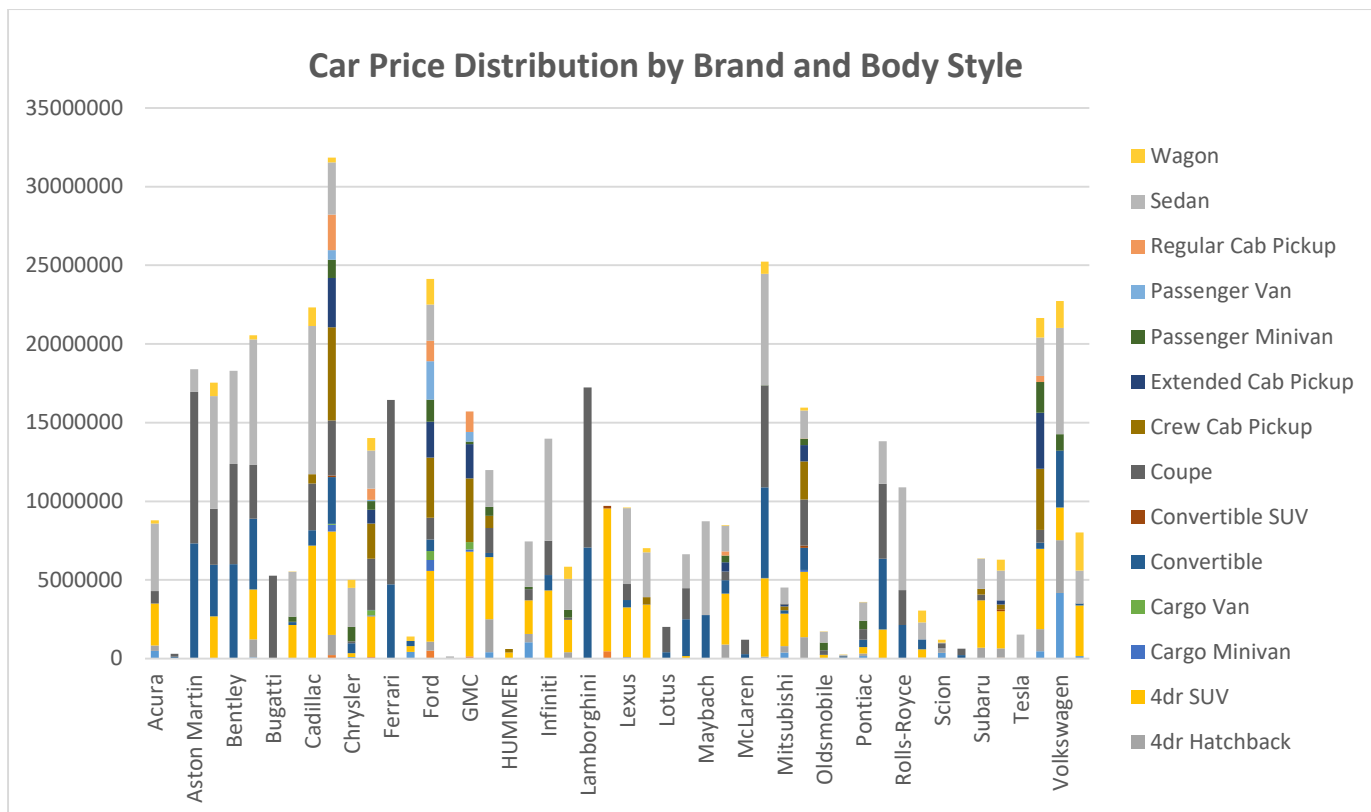
- Automakers can use this insight to compare against rivals and **optimize engine designs** that strike a better balance between performance and fuel efficiency.

Dashboard Analysis:

Dashboard Task 1. – Car Price Distribution by Brand and Body Style

Objective:

To analyse how **car prices (MSRP)** are distributed across **different manufacturers (brands)** and **vehicle body styles**, helping stakeholders understand which body styles and brands dominate the market in terms of total sales value.



Key Insights:

1. Brand Dominance in MSRP Contribution:

- **Cadillac, Ferrari, and Ford** show the **highest total MSRP**, indicating a broad and/or premium line-up.
- **Luxury brands** like **Bentley, Bugatti, Maybach, and McLaren** contribute large MSRPs even with fewer models, reflecting **high-value cars**.

2. Body Style Trends:

- **Coupes** (dark grey bars) are the **dominant body style** across luxury brands (e.g., Ferrari, Lamborghini, McLaren).
- **Sedans and SUVs** (e.g., **Extended Cab Pickup**) are prominent in mainstream brands like **Ford, GM, Chrysler, and Volkswagen**, suggesting consumer preference for practicality.

3. Niche Segment Presence:

- Brands like **Subaru, Scion, and Pontiac** have relatively **lower total MSRP** contributions, indicating either fewer models or targeting budget-conscious consumers.

4. Brand Specialization by Body Style:

- Some brands show strong alignment with specific body styles:
 - **Ferrari and Lamborghini:** Mostly coupes.

- **Ford and GMC:** More pickups and SUVs.
- **Maybach, Rolls-Royce:** Heavily sedan-based, targeting luxury transportation.

Business Impact:

Strategic Product Planning:

- Identifies which body styles are most **profitable** for specific brands.
- Helps **automakers decide future investments** in sedan, coupe, SUV, or pickup segments.

Marketing & Positioning:

- Tailor **targeted campaigns:** luxury brands can focus on high-end coupe buyers; mainstream brands can target family sedan and pickup users.
- Supports **brand identity positioning** (e.g., performance vs utility).

Revenue Optimization:

- Highlights opportunities to **rebalance product portfolios:**
 - For example, brands with low coupe representation but high demand might **expand offerings** in that segment.

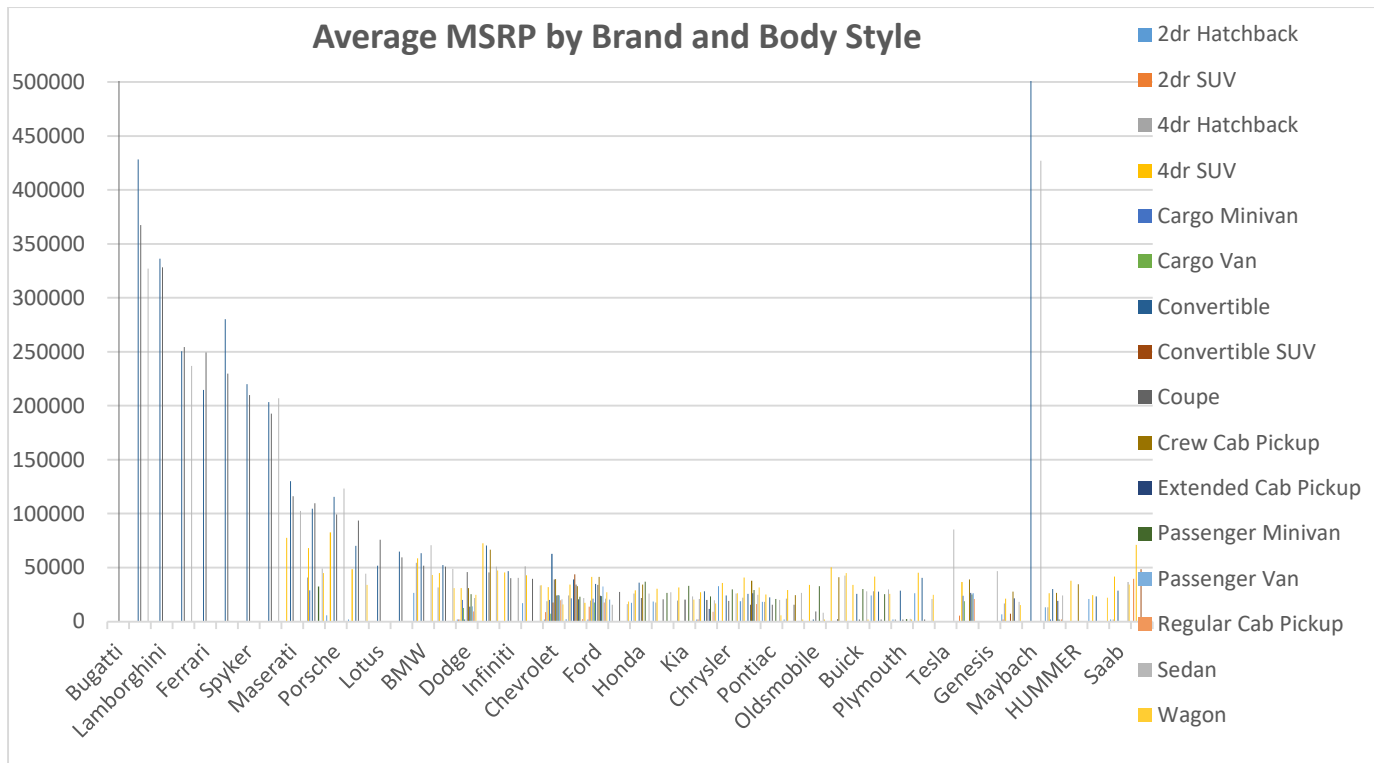
Competitive Intelligence:

- Understand how brands **stack up against competitors** in total MSRP and product diversity.
- Helps dealerships and partners with **inventory management** and **pricing strategies**.

Dashboard Task 2. – Brand & Body Style Pricing Analysis

Objective:

To analyse **which car brands have the highest and lowest average prices (MSRPs)** and assess how **vehicle body style** contributes to those price differences. This helps identify premium vs. budget brands and style-based pricing trends.



Key Insights:

1. Top Premium Brands:

- Bugatti, Lamborghini, Ferrari, and Spyker lead with **average MSRPs well above \$200,000**, dominated by body styles like:
 - 2dr Hatchback
 - Convertible
- These brands focus heavily on **luxury sports and performance vehicles**, explaining their elevated prices.

2. Mid-Tier Brands:

- Brands like **BMW, Porsche, Lotus, and Maserati** fall in the **\$75,000-\$150,000 range**, offering a mix of **premium sedans, coupes, and SUVs**.

3. Mainstream and Budget Brands:

- Brands like **Chevrolet, Ford, Honda, Kia, and Chrysler** show much **lower average MSRPs, generally below \$40,000**.
- These brands primarily offer **4-door SUVs, hatchbacks, and minivans** catering to the **mass market**.

4. Body Style Trends Across Brands:

- **2dr Hatchback and Convertible** dominate in high-priced luxury brands.

- **4dr SUVs and Hatchbacks** are common in mid-tier and budget segments.
- Some brands (e.g., **Maybach, Genesis**) show **price spikes in specific body styles**, indicating niche luxury offerings.

Business Impact:

Strategic Positioning & Market Segmentation:

- Helps automotive companies **benchmark their brand's price tier** relative to competitors.
- Informs decisions on **whether to enter or expand** in luxury, mid-range, or economy segments.

Product Line Optimization:

- Reveals which **body styles drive premium pricing**.
- Assists manufacturers in **fine-tuning model mixes** (e.g., increasing focus on convertibles for high-margin sales).

Sales Strategy & Marketing:

- Enables creation of **targeted marketing campaigns**:
 - Luxury brands can emphasize performance and exclusivity.
 - Mainstream brands can highlight affordability and practicality.

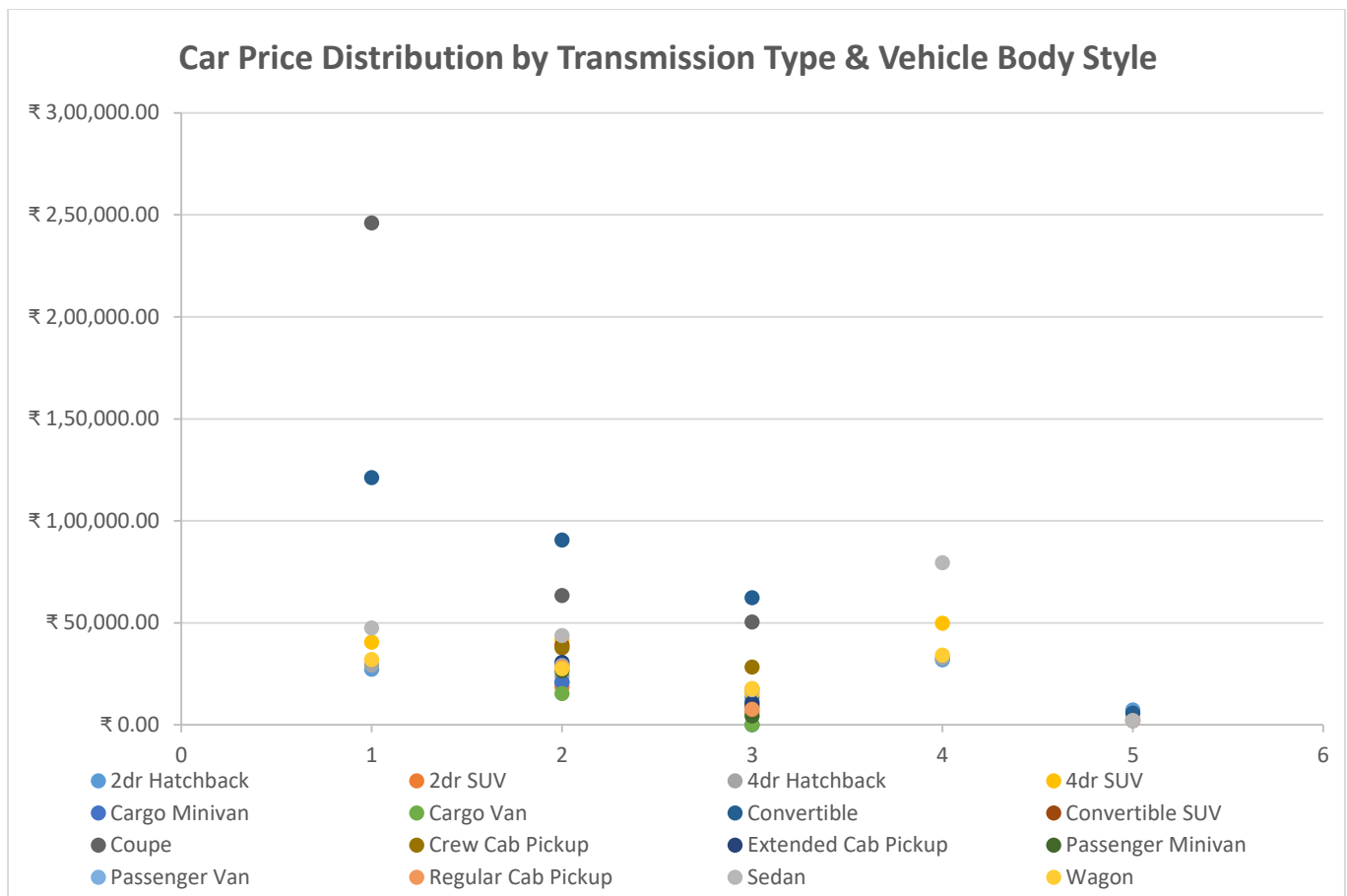
Inventory and Dealership Planning:

- Dealerships can use this to **adjust stock levels** based on local demand and pricing trends by brand and style.

Dashboard Task 3. – Car Price Distribution by Transmission Type & Body Style

Objective:

To analyse how **transmission type** influences car prices (MSRP) across various **vehicle body styles**, and to visualize the interaction between these two features in affecting pricing.



Key Insights:

1. Transmission Type Matters:

- Certain transmission types are consistently associated with **higher MSRP values**, especially for performance-oriented or luxury body styles like **Convertibles** and **Coupes**.
- Manual or less common transmission types are often linked with either **niche performance vehicles** or **budget models**, depending on the body style.

2. Body Style Dependency:

- High-priced outliers (e.g., **Coupes** and **Convertibles**) tend to cluster under specific transmission types.
- Mainstream body styles such as **Sedans**, **SUVs**, and **Minivans** exhibit more balanced pricing across all transmission types.

3. Wide Price Range:

- There is a **significant variance in price** within the same transmission type, indicating that body style heavily modulated the impact of transmission on pricing.

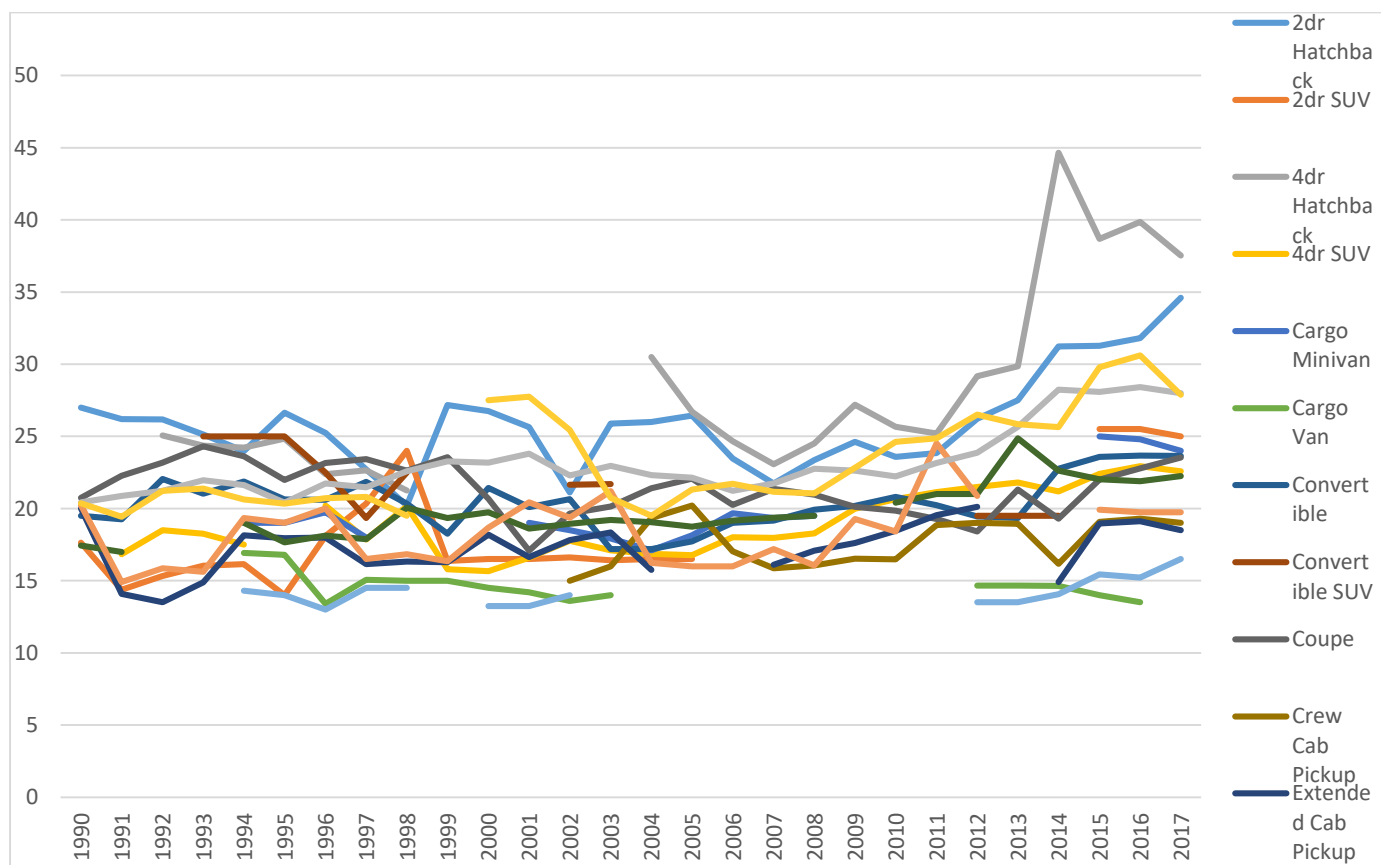
Business Impact:

- **Strategic Product Positioning:** Insights from this analysis help car manufacturers determine optimal **feature combinations** (e.g., transmission + body style) to target specific **price segments**.
- **Customer Segmentation:** Automotive marketers can **bundle and promote** specific configurations (e.g., automatic convertibles) to match pricing expectations of **premium vs budget-conscious** buyers.
- **Inventory & Pricing Optimization:** Helps dealerships adjust pricing based on **market trends** and **consumer preference patterns** by body style and transmission pairings.

Dashboard Task 4. – Fuel Efficiency Trend by Body Style and Year

Objective:

To analyse how **fuel efficiency** (measured as combined MPG) has changed over time across different **vehicle body styles**, and to identify patterns, improvements, or inconsistencies from **1990 to 2017**.



Key Insights:

1. Upward Trend in Fuel Efficiency:

- Several body styles such as **Cargo Minivans, 2dr Hatchbacks, and 4dr Hatchbacks** show a consistent **increase in average MPG**, especially post-2010.

2. Diverse Performance Across Styles:

- **Pickup trucks, Crew Cab, and SUVs** tend to show **lower fuel efficiency**, often hovering around 15-20 MPG throughout the years.
- In contrast, hatchbacks and minivans reach **over 30 MPG** in recent years, indicating improvements in design and technology.

3. Significant Gains Post-2010:

- Marked improvements in MPG post-2010 may reflect the impact of **fuel economy regulations, technological advances** like hybrid engines, and **market pressure** for efficiency.

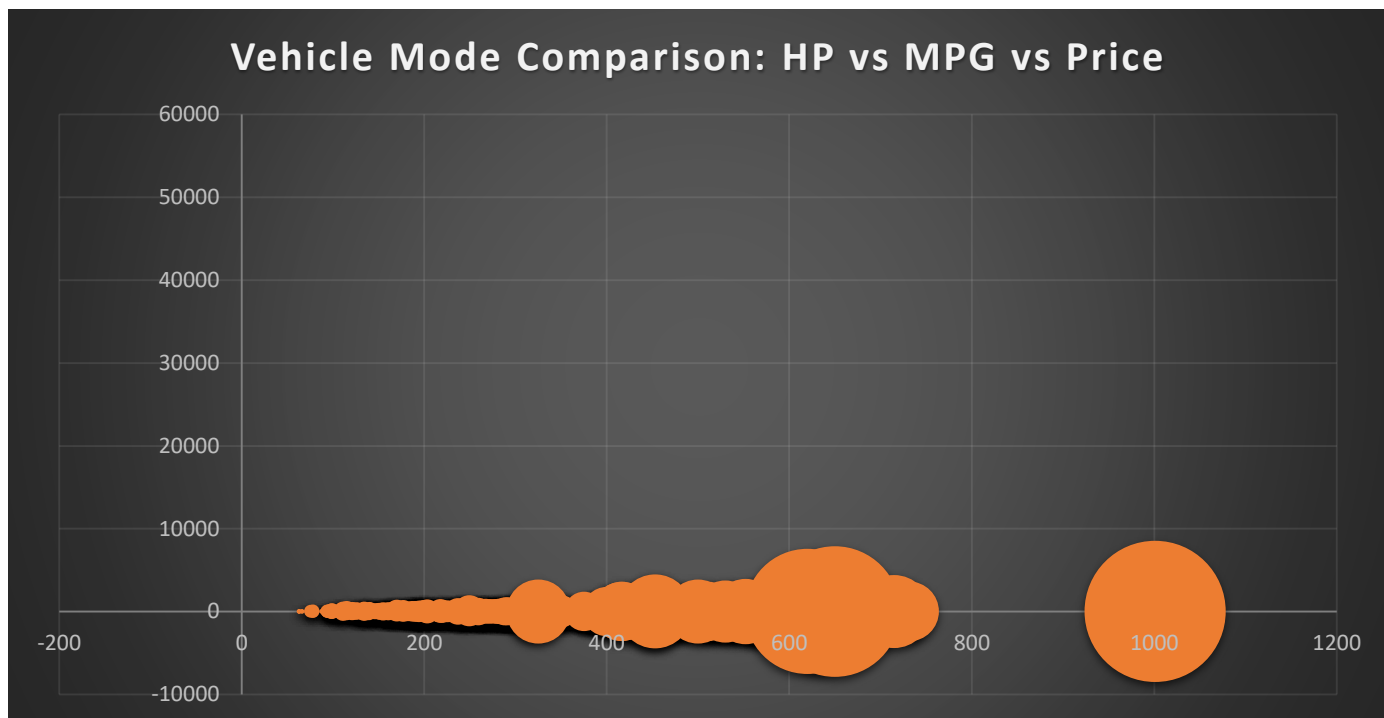
Business Impact:

- **Product Development:** Helps **automakers optimize design choices** to meet efficiency standards, especially in low-performing body styles (like pickups).
- **Marketing Strategy:** Enables **eco-conscious branding** for vehicle segments that show improved efficiency trends (e.g., hatchbacks and minivans).
- **Regulatory Compliance:** Assists in tracking the impact of **fuel economy legislation** over time and informs future development strategy

Dashboard Task 5. – HP vs MPG vs Price Bubble Analysis

Objective:

To examine the **relationship among horsepower (HP), fuel efficiency (MPG), and price (MSRP)** across different car models/brands using a bubble chart. The **x-axis** represents **horsepower**, the **y-axis** represents **MPG**, and the **bubble size** represents **price**.



Key Insights:

1. Inverse Relationship Between HP and MPG:

- As horsepower increases, **fuel efficiency tends to decrease**.
- Vehicles with high HP (e.g., >500 HP) mostly cluster around **lower MPG values** (under 20 MPG), consistent with performance or luxury vehicles.

2. Luxury or High-Performance Vehicles Are Expensive:

- The **largest bubbles** (highest prices) appear at the **extreme high horsepower** and **low MPG** ends, indicating that **supercars and luxury models** are typically high in HP but low in fuel efficiency.

3. Efficient, Budget-Friendly Cars Cluster Together:

- Vehicles with **moderate HP and higher MPG** (e.g., 100-200 HP, 30-40 MPG) tend to have **smaller bubbles**, showing affordability and efficiency (compact or hybrid cars).

Business Impact:

- **Product Positioning:** Helps identify **market segments** – performance-focused, fuel-efficient, or affordable – so manufacturers can tailor design and pricing strategies accordingly.
- **Pricing Strategy:** Guides pricing decisions by showing how **horsepower and efficiency affect value perception**.

- **Marketing Insight:** Reveals which models deliver the best **value mix of performance and efficiency**, allowing targeted promotion to specific buyer personas.