Analyzing the Impact of Car Features on Price and Profitability

Project Description:

In a competitive automotive market where consumer preferences are constantly evolving, understanding the relationship between car features and their impact on pricing and profitability is essential for manufacturers. This project aims to analyze a range of car features—such as horsepower, fuel efficiency, body style, and market category—and determine how these features affect car prices and consumer demand. The goal is to equip car manufacturers with actionable insights that will guide pricing strategies, product development, and market positioning.

Problem Statement

With increasing competition and the rise of new technologies (such as electric and hybrid vehicles), car manufacturers must understand which features drive consumer demand and how these features correlate with car prices. By examining factors like horsepower, fuel type, MPG, and brand, we aim to determine which attributes contribute most to the pricing and popularity of vehicles across different market segments. These insights will help manufacturers align their product offerings with consumer preferences, optimize pricing strategies, and enhance profitability.

Approach:

This project is designed to guide car manufacturers in making data-driven decisions that enhance profitability and strengthen their market presence in the evolving automotive industry.

To achieve the objective of this project, the approach will involve structured steps focusing on data exploration, analysis, and insights extraction using Exploratory Data Analysis (EDA).

1. **Data Understanding and Cleaning**
    - Load and review the dataset for completeness and accuracy.
    - Check for missing, incorrect, or inconsistent values, especially in crucial columns like MSRP, Engine HP, MPG, and popularity.

- Clean the data by filling or removing missing values, standardizing formats, and removing duplicates.

2. **Exploratory Data Analysis (EDA)**
   - **Descriptive Statistics**: Summarize key statistics (mean, median, max, min) for numerical variables such as MSRP, horsepower, and MPG to understand the data's general distribution.
   - **Feature Distribution Analysis**: Plot distributions for MSRP, engine HP, and popularity to see common ranges, variations, and potential outliers.
   - **Correlations**: Use correlation matrices to identify relationships between variables (e.g., horsepower and MSRP), guiding feature selection for deeper analysis.
3. **Feature Engineering**
   - **Create MSRP Ranges**: Group MSRP values into ranges (e.g., <$20,000, $20,000-$40,000) to facilitate category-based analyses.
   - **Year Binning**: For trend analysis, group car models by release year to observe changes in pricing and features over time.
   - **Interaction Terms**: Create combinations of variables (e.g., "Body Style x Transmission Type") for exploring specific segment trends.
4. **Analysis and Visualization**
   - **Trend Analysis**:
     - Pivot table and line chart to show changes in average MSRP, MPG, and horsepower over time.
     - Explore trends in fuel efficiency by body style and model year to understand the evolution of consumer preferences for fuel economy.
   - **Feature-Price Analysis**:
     - Perform a **regression analysis** with MSRP as the dependent variable to determine which features (e.g., horsepower, transmission type, engine cylinders) impact price the most.
     - Visualize regression coefficients in a bar chart to highlight features with the highest price impact.

- **Brand and Body Style Analysis**:
  - Calculate average MSRP for each brand and body style and visualize this in a clustered column chart to compare pricing across brands and styles.
  - Create a **scatter plot** to examine how features like engine HP, MPG, and popularity correlate with MSRP, differentiating by body style and brand.
- **Fuel Efficiency Analysis**:
  - Use a line chart to show how fuel efficiency (MPG) varies across different body styles over time.
  - Explore fuel type distribution by body style to understand consumer preferences and how they may impact future pricing strategies.

5. **Dashboard Development**
   - **Interactive Filters**: Implement slicers for brand, body style, model year, and other key features.
   - **Chart Integration**: Combine scatter plots, line charts, and pivot tables for interactive exploration of the relationships between car features and MSRP.
   - **Dynamic Summary Tables**: Create summary tables to show averages for each segment (brand, body style) based on selected filters.
6. **Insight Extraction**
   - Summarize key insights regarding high-impact features on pricing, most popular body styles and fuel types, and brand performance in the luxury and economy segments.

This structured approach allows for a comprehensive analysis of car features and their impact on MSRP, equipping car manufacturers with actionable insights to guide future product development and pricing strategies.

Tech-Stack Used:

1. Microsoft excels

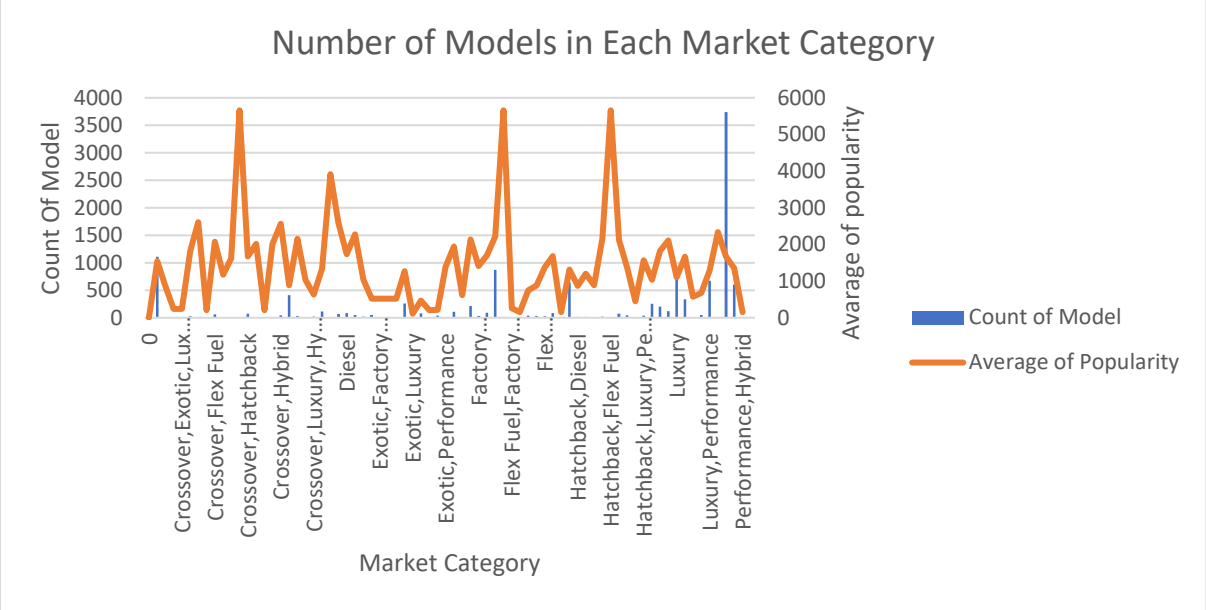
Insights:

Tasks: Analysis

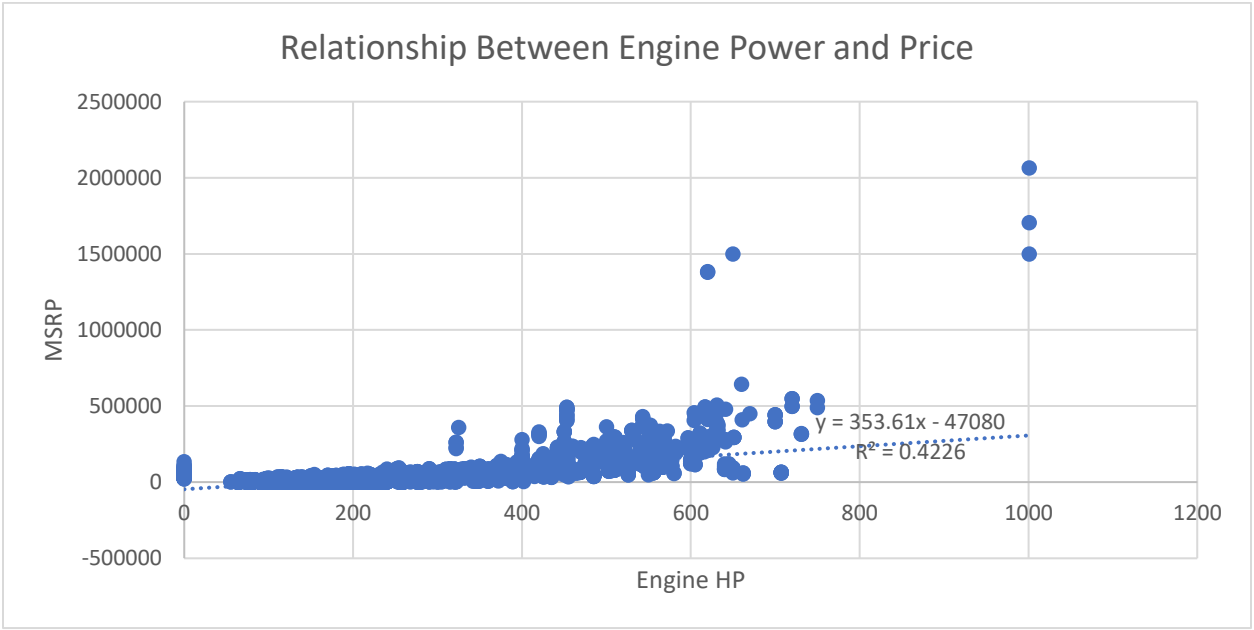
- **Task 1.A:** Create a pivot table that shows the number of car models in each market category and their corresponding popularity scores.

Row Labels	Count of Model	Average of Popularity
0	1	0
Crossover	1110	1545.263063
Crossover,Diesel	7	873
Crossover,Exotic,Luxury,High-Performance	1	238
Crossover,Exotic,Luxury,Performance	1	238
Crossover,Factory Tuner,Luxury,High-Performance	26	1823.461538
Crossover,Factory Tuner,Luxury,Performance	5	2607.4
Crossover,Factory Tuner,Performance	4	210
Crossover,Flex Fuel	64	2073.75
Crossover,Flex Fuel,Luxury	10	1173.2
Crossover,Flex Fuel,Luxury,Performance	6	1624
Crossover,Flex Fuel,Performance	6	5657
Crossover,Hatchback	72	1675.694444
Crossover,Hatchback,Factory Tuner,Performance	6	2009
Crossover,Hatchback,Luxury	7	204
Crossover,Hatchback,Performance	6	2009
Crossover,Hybrid	42	2563.380952
Crossover,Luxury	410	884.5487805
Crossover,Luxury,Diesel	34	2149.411765
Crossover,Luxury,High-Performance	9	1037.222222
Crossover,Luxury,Hybrid	24	630.9166667
Crossover,Luxury,Performance	113	1344.849558
Crossover,Luxury,Performance,Hybrid	2	3916
Crossover,Performance	69	2585.956522
Diesel	84	1730.904762
Diesel,Luxury	51	2275
Exotic,Factory Tuner,High-Performance	21	1046.380952
Exotic,Factory Tuner,Luxury,High-Performance	52	517.5384615
Exotic,Factory Tuner,Luxury,Performance	3	520
Exotic,Flex Fuel,Factory Tuner,Luxury,High-Performance	13	520
Exotic,Flex Fuel,Luxury,High-Performance	11	520
Exotic,High-Performance	261	1271.333333
Exotic,Luxury	12	112.6666667
Exotic,Luxury,High-Performance	79	467.0759494
Exotic,Luxury,High-Performance,Hybrid	1	204
Exotic,Luxury,Performance	36	217.0277778
Exotic,Performance	10	1391
Factory Tuner,High-Performance	106	1941.415094
Factory Tuner,Luxury	2	617
Factory Tuner,Luxury,High-Performance	215	2133.367442
Factory Tuner,Luxury,Performance	31	1413.419355
Factory Tuner,Performance	92	1695.695652
Flex Fuel	872	2217.302752
Flex Fuel,Diesel	16	5657
Flex Fuel,Factory Tuner,Luxury,High-Performance	1	258
Flex Fuel,Hybrid	2	155
Flex Fuel,Luxury	39	746.5384615
Flex Fuel,Luxury,High-Performance	33	878.9090909
Flex Fuel,Luxury,Performance	28	1380.071429
Flex Fuel,Performance	87	1680.471264
Flex Fuel,Performance,Hybrid	2	155
Hatchback	641	1318.865835
Hatchback,Diesel	14	873
Hatchback,Factory Tuner,High-Performance	13	1205.153846
Hatchback,Factory Tuner,Luxury,Performance	9	886.8888889
Hatchback,Factory Tuner,Performance	22	2159.045455
Hatchback,Flex Fuel	7	5657
Hatchback,Hybrid	72	2121.25
Hatchback,Luxury	46	1379.5
Hatchback,Luxury,Hybrid	3	454
Hatchback,Luxury,Performance	38	1566.131579
Hatchback,Performance	252	1039.646825
High-Performance	199	1821.447236
Hybrid	123	2105.569106
Luxury	855	1102.65731
Luxury,High-Performance	334	1668.017964
Luxury,High-Performance,Hybrid	12	568.8333333
Luxury,Hybrid	52	673.6346154
Luxury,Performance	673	1292.615156
Luxury,Performance,Hybrid	11	2333.181818
N/A	3742	1676.889364
Performance	601	1348.873544
Performance,Hybrid	1	155

- **Task 1.B**: Create a combo chart that visualizes the relationship between market category and popularity.



- **Task 2**: Create a scatter chart that plots engine power on the x-axis and price on the y-axis. Add a trendline to the chart to visualize the relationship between these variables.



- **Task 3**: Use regression analysis to identify the variables that have the strongest relationship with a car's price. Then create a bar chart that shows the coefficient values for each variable to visualize their relative importance.

SUMMARY OUTPUT

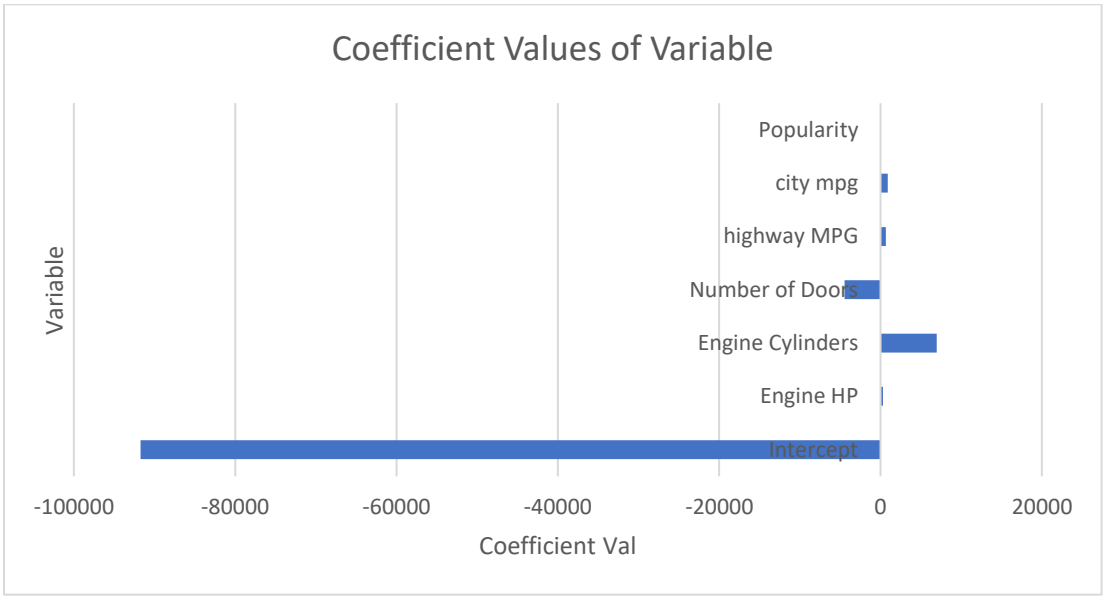
Regression Statistics	
Multiple R	0.680299618
R Square	0.46280757
Adjusted R Square	0.462536876
Standard Error	44067.11182
Observations	11914

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	6	1.99206E+13	3.3201E+12	1709.706934	0
Residual	11907	2.31223E+13	1941910344		
Total	11913	4.30429E+13			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-91719.25488	3358.756319	-27.30750497	2.6294E-159	-98302.96554	-85135.54422	-98302.96554	-85135.54422
Engine HP	316.0413332	5.838045105	54.13478784	0	304.5978118	327.4848546	304.5978118	327.4848546
Engine Cylinders	6953.672063	413.0781906	16.83379133	7.34049E-63	6143.97138	7763.372747	6143.97138	7763.372747
Number of Doors	-4475.59031	463.8183583	-9.649446232	5.94632E-22	-5384.750005	-3566.430615	-5384.750005	-3566.430615
highway MPG	675.1195202	102.7132839	6.572854985	5.14195E-11	473.7847171	876.4543234	473.7847171	876.4543234
city mpg	876.0024934	98.23660262	8.917271873	5.47255E-19	683.4427164	1068.562271	683.4427164	1068.562271
Popularity	-3.202189173	0.280791469	-11.40415407	5.70219E-30	-3.752586287	-2.651792058	-3.752586287	-2.651792058

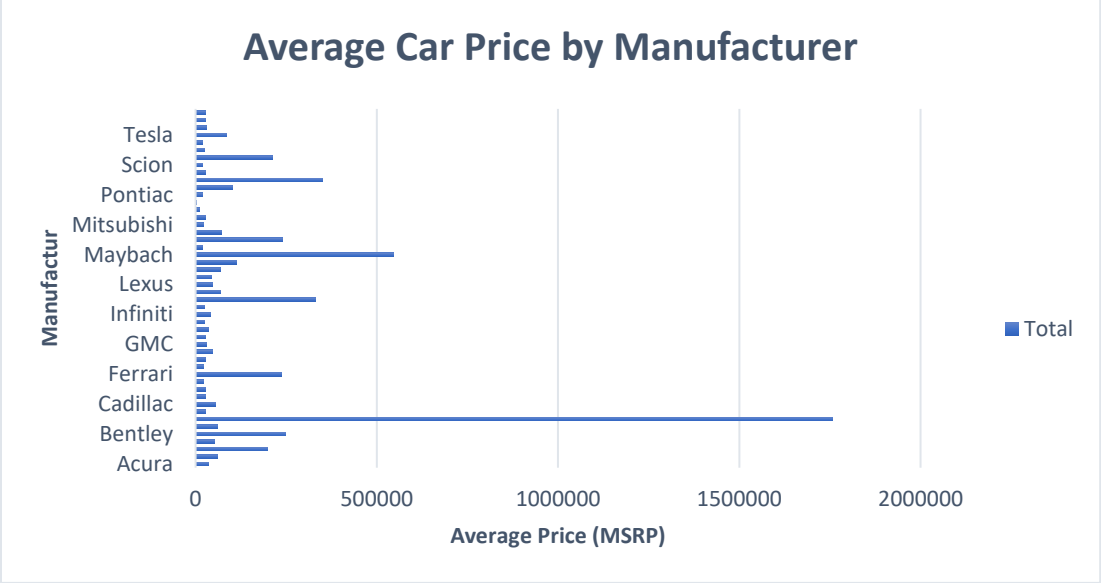
Variable	*Coefficients*
Intercept	-91719.25488
Engine HP	316.0413332
Engine Cylinders	6953.672063
Number of Doors	-4475.59031
highway MPG	675.1195202
city mpg	876.0024934
Popularity	-3.202189173



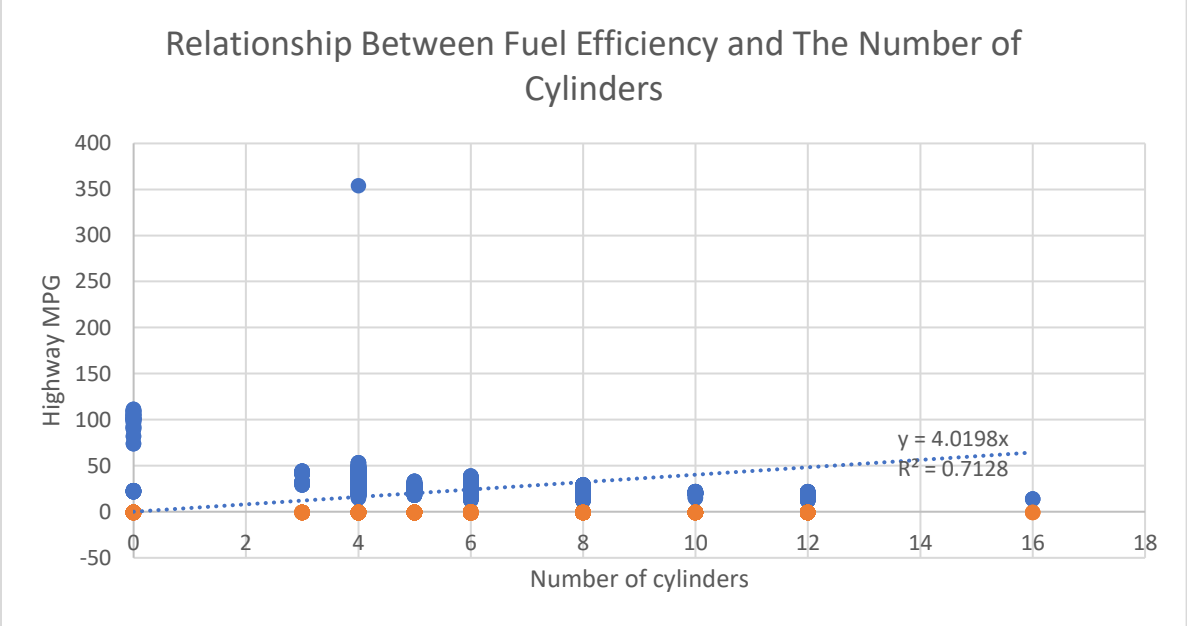
- **Task 4.A**: Create a pivot table that shows the average price of cars for each manufacturer.

Make	Average of MSRP
Acura	34888
Alfa Romeo	61600
Aston Martin	197910
Audi	53452
Bentley	247169
BMW	61547
Bugatti	1757224
Buick	28207
Cadillac	56231
Chevrolet	28350
Chrysler	26723
Dodge	22390
Ferrari	238219
FIAT	22670
Ford	27399
Genesis	46617
GMC	30493
Honda	26674
HUMMER	36464
Hyundai	24597
Infiniti	42394
Kia	25310
Lamborghini	331567
Land Rover	67823
Lexus	47549
Lincoln	42840
Lotus	69188
Maserati	114208
Maybach	546222
Mazda	20039
McLaren	239805
Mercedes-Benz	71476
Mitsubishi	21241
Nissan	28583
Oldsmobile	11543
Plymouth	3123
Pontiac	19322
Porsche	101622
Rolls-Royce	351131
Saab	27414
Scion	19933
Spyker	213323
Subaru	24828
Suzuki	17907
Tesla	85256
Toyota	29030
Volkswagen	28102
Volvo	28541

- **Task 4.B:** Create a bar chart or a horizontal stacked bar chart that visualizes the relationship between manufacturer and average price.



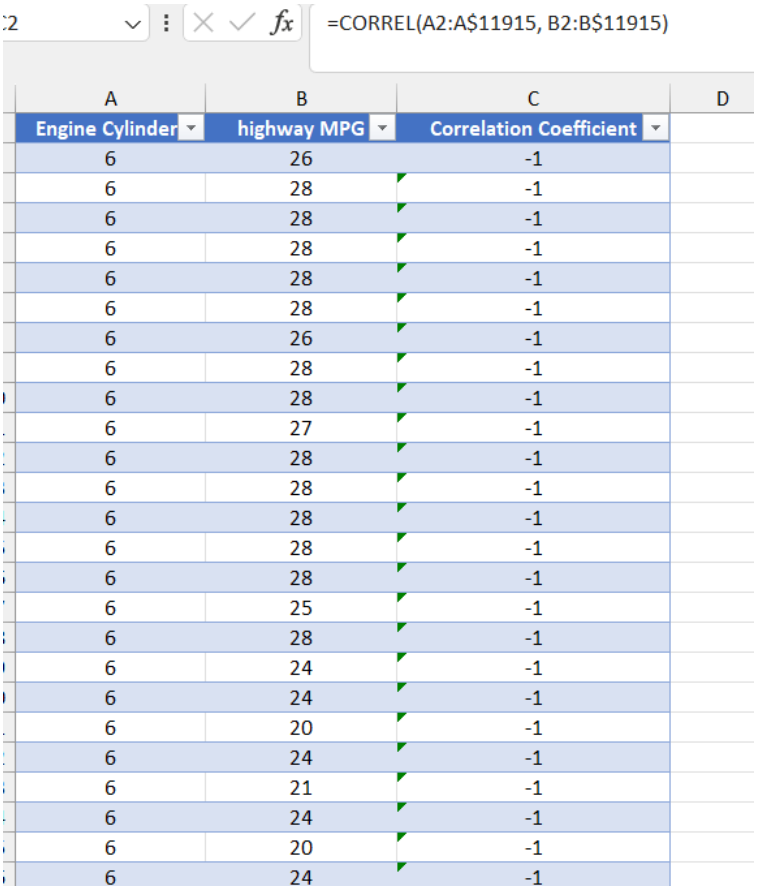
- **Task 5.A:** Create a scatter plot with the number of cylinders on the x-axis and highway MPG on the y-axis. Then create a trendline on the scatter plot to visually estimate the slope of the relationship and assess its significance.



- **Task 5.B:** Calculate the correlation coefficient between the number of cylinders and highway MPG to quantify the strength and direction of the relationship.

  The correlation coefficient (r) will range from -1 to +1.

  - Positive Value (0 to +1): Indicates a positive relationship (as one variable increases, the other tends to increase).
  - Negative Value (0 to -1): Indicates a negative relationship (as one variable increases, the other tends to decrease).
  - Close to 0: Indicates little to no relationship between the variables

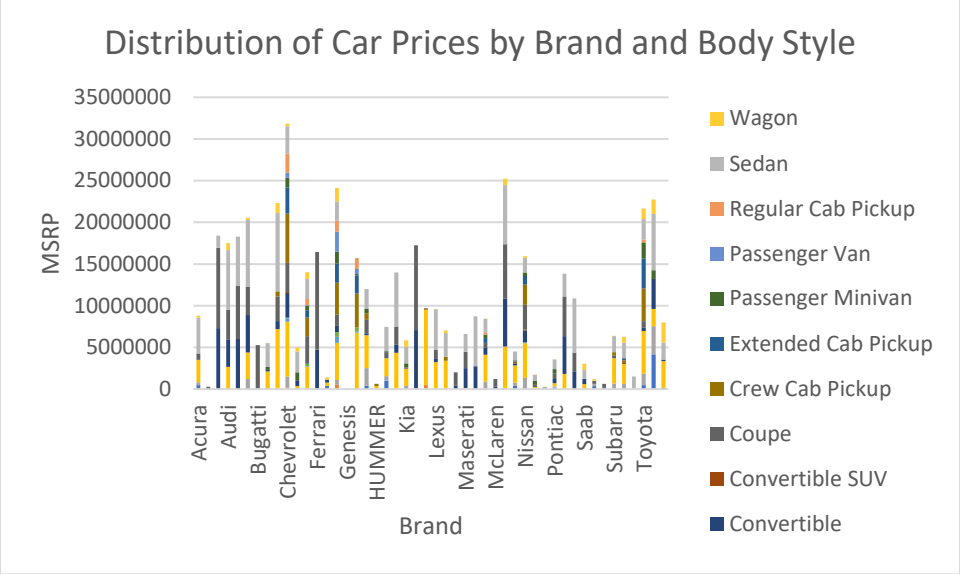


	A	B	C	D
	Engine Cylinder	highway MPG	Correlation Coefficient	
	6	26	-1	
	6	28	-1	
	6	28	-1	
	6	28	-1	
	6	28	-1	
	6	28	-1	
	6	26	-1	
	6	28	-1	
	6	28	-1	
	6	27	-1	
	6	28	-1	
	6	28	-1	
	6	28	-1	
	6	28	-1	
	6	28	-1	
	6	28	-1	
	6	25	-1	
	6	28	-1	
	6	24	-1	
	6	24	-1	
	6	20	-1	
	6	24	-1	
	6	21	-1	
	6	24	-1	
	6	20	-1	
	6	24	-1	

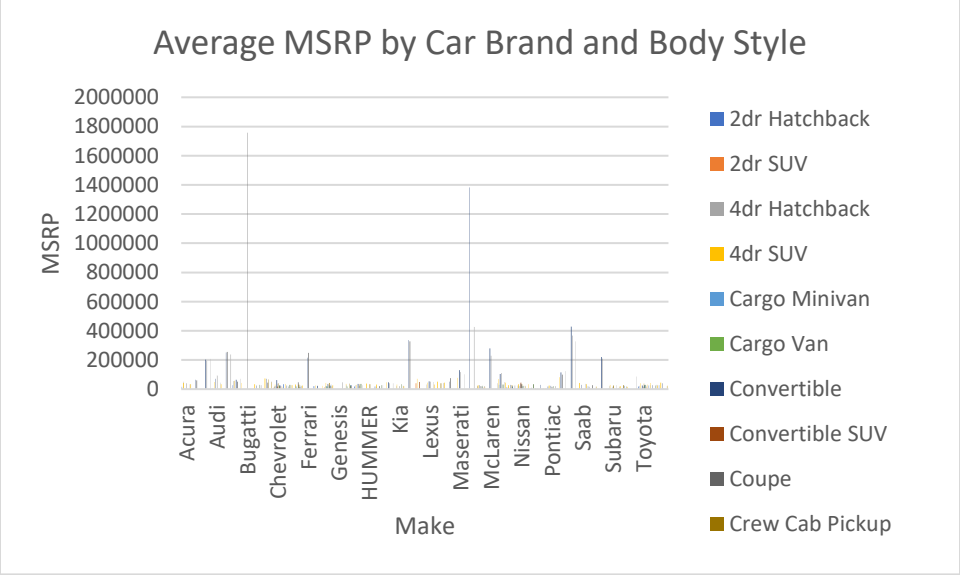
Task:Building the Dashboard

Created the Interactive Dashboard. Used filters and slicers to make the chart interactive.

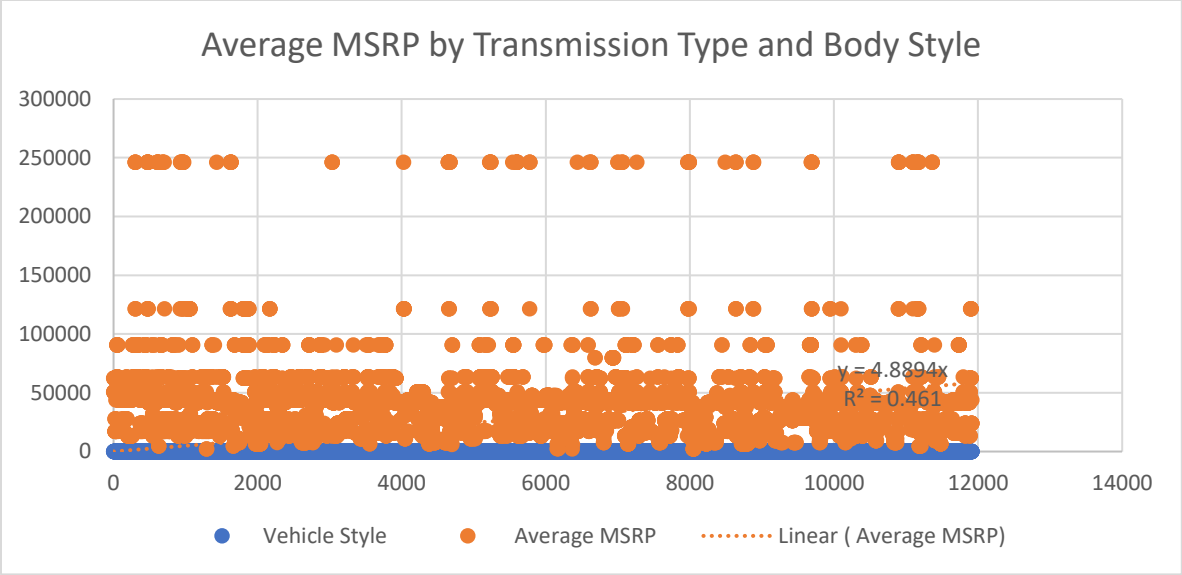
Task 1: How does the distribution of car prices vary by brand and body style?



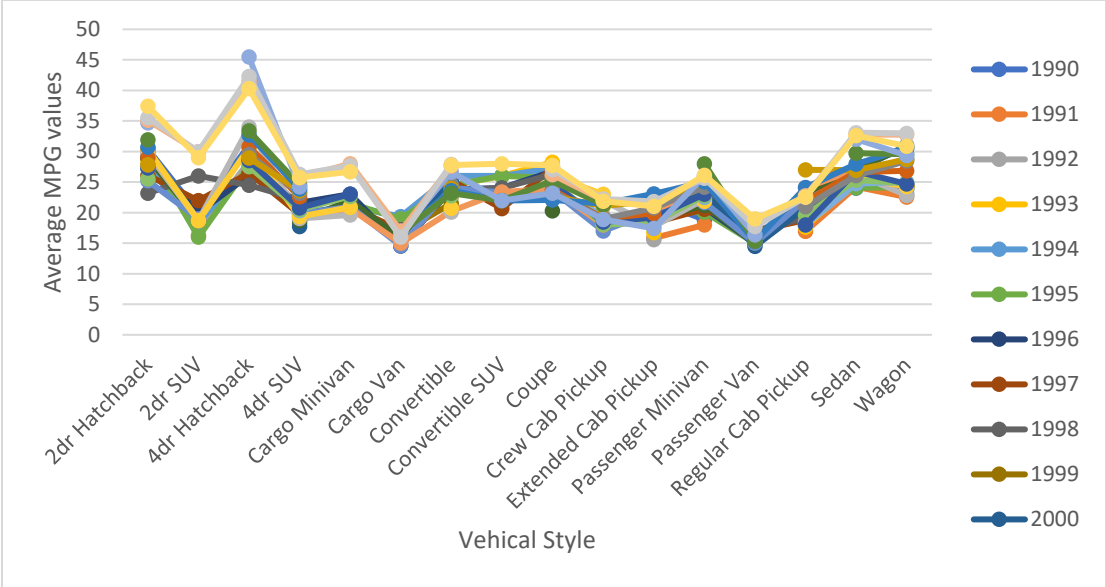
Task 2: Which car brands have the highest and lowest average MSRPs, and how does this vary by body style?



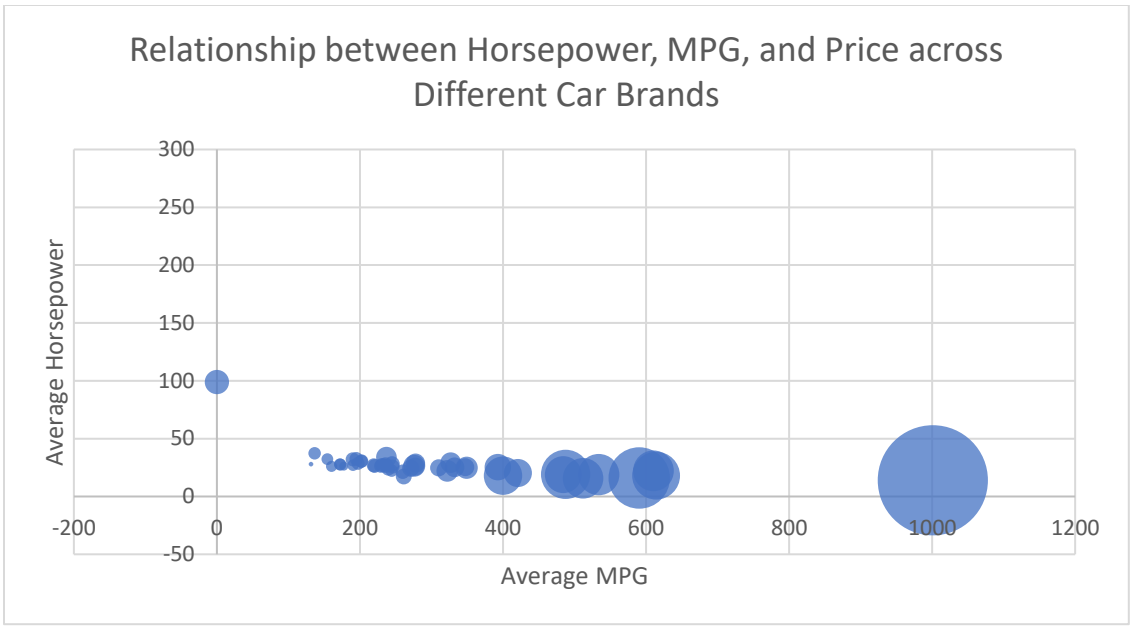
Task 3: How do the different feature such as transmission type affect the MSRP, and how does this vary by body style?



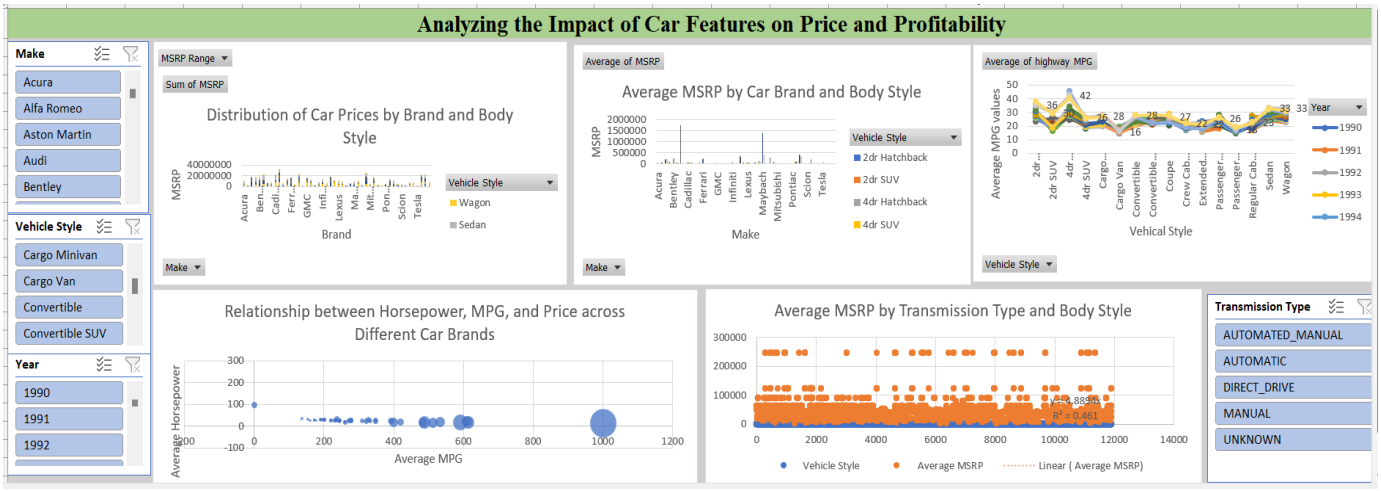
Task 4: How does the fuel efficiency of cars vary across different body styles and model years?



Task 5: How does the car's horsepower, MPG, and price vary across different Brands?



Dashboard:



Result:

This analysis offers valuable insights into the complex relationships between car features, brand positioning, and pricing in the automotive industry. By examining data on horsepower, fuel

efficiency, transmission type, body style, and MSRP across brands, several key trends have emerged:

1. **Premium Pricing Justifications**: Features such as high horsepower, luxury body styles, and brand prestige significantly impact MSRP, justifying premium pricing. Brands can leverage these features to strengthen their positioning in high-end segments, especially in luxury sedans and SUVs.
2. **Fuel Efficiency Demand**: Fuel-efficient models—particularly hybrids and compact vehicles—are increasingly popular among consumers. Investments in fuel-efficient technology, especially for compact and SUV body styles, align with evolving consumer preferences and sustainability trends.
3. **Electric and Hybrid Growth Potential**: Although traditional gasoline vehicles dominate, there is a notable upward trend in hybrid and electric vehicle popularity, especially for brands focusing on innovation. Expanding these offerings can enhance brand perception and cater to eco-conscious consumers.
4. **Market-Specific Strategies**:
    - **Luxury Brands**: Should continue focusing on performance features such as high horsepower and advanced transmission options, as these are key determinants of higher MSRP and consumer demand in the luxury market.
    - **Mainstream Brands**: Benefit from targeting affordability and fuel efficiency in compact and crossover SUVs to capture a broader consumer base, especially in mid-range and family-oriented segments.
5. **Interactive Dashboard Utility**: The developed dashboard allows stakeholders to dynamically analyze relationships among car features, brands, and pricing across different segments. By using filters for features like body style and transmission, manufacturers gain a versatile tool to adapt their pricing strategies and monitor changing consumer demands.