

# Impact of Car Features on Price and Profitability

## Project Description:

This project aims to analyse the relationship between a car's features, market category, and pricing, and identifying which features and categories are most popular among consumers and most profitable for the manufacturer.

This could help the manufacturer improve its competitiveness in the market and increase its profitability over time.

## Problem Statement:

The automotive industry has been rapidly evolving over the past few decades.

With increasing competition among manufacturers and a changing consumer landscape, it's important to understand the factors that drive consumer demand for cars.

For the given dataset, as a Data Analyst, the client has asked How can a car manufacturer optimize pricing and product development decisions to maximize profitability while meeting consumer demand?

## Dataset Description:

The dataset was collected and made available on Kaggle by Cooper Union, a private college located in New York City.

Brief overview of the dataset:

- **Number of observations:** 11,914
- **Number of variables:** 16

The variables in the dataset are:

- **Make:** the make or brand of the car
- **Model:** the specific model of the car
- **Year:** the year the car was released
- **Engine Fuel Type:** the type of fuel used by the car (gasoline, diesel, etc.)
- **Engine HP:** the horsepower of the car's engine
- **Engine Cylinders:** the number of cylinders in the car's engine
- **Transmission Type:** the type of transmission (automatic or manual)
- **Driven\_Wheels:** the type of wheels driven by the car (front, rear, all)
- **Number of Doors:** the number of doors the car has
- **Market Category:** the market category the car belongs to (Luxury, Performance, etc.)
- **Vehicle Size:** the size of the car
- **Vehicle Style:** the style of the car (Sedan, Coupe, etc.)
- **Highway MPG:** the estimated miles per gallon the car gets on the highway
- **City MPG:** the estimated miles per gallon the car gets in the city
- **Popularity:** a ranking of the popularity of the car (based on the number of times it has been viewed on Edmunds.com)
- **MSRP:** the manufacturer's suggested retail price of the car

Note: the dataset was last updated in 2017, so it may not reflect current trends or prices in the automotive industry.

## Dataset Cleaning Steps performed:

- 1) Remove duplicates

## 2) Check null % in columns

Columns	Engine Fuel Type	Engine HP	Engine Cylinders	Number of Doors
original data count	11911	11845	11884	11908
after removing duplicates	11196	11130	11169	11193
null count	3	69	30	6
null%	0.026788	0.616126	0.267881	0.053576

## 3) Check count of null rows: 102

## 4) Identify and impute missing values:

- Engine Fuel Type: categorical column : impute with mode
- Engine HP: Numerical column : check outliers
- Engine Cylinders : Numerical column : check outliers
- Number of Doors: Numerical column: only 6 data points are missing, replace with common values based on other columns

Cleaned dataset description:

- Number of observations: 11,199
- Number of variables: 16

## Approach:

After cleaning the dataset, by using data analysis techniques such as regression analysis and market segmentation, analysing the car features on price & profitability.

## Tech-Stack Used:

- MS EXCEL 2019 for analysis and creating interactive dashboard

## Insights:

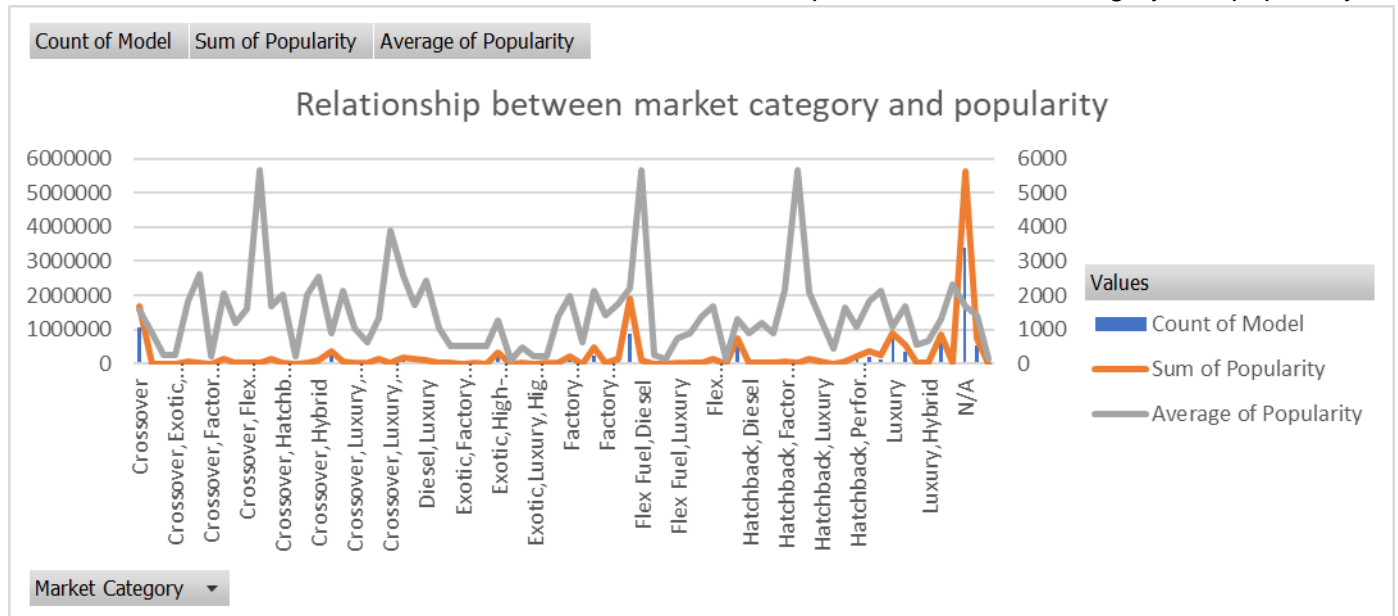
**Insight Required:** How does the popularity of a car model vary across different market categories?

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

Market Category	Count of Model	Sum of Popularity	Average of Popularity
Crossover	1075	1672881	1556.17
Crossover,Diesel	7	6111	873.00
Crossover,Exotic,Luxury,High-Performance	1	238	238.00
Crossover,Exotic,Luxury,Performance	1	238	238.00
Crossover,Factory Tuner,Luxury,High-Performance	26	47410	1823.46
Crossover,Factory Tuner,Luxury,Performance	5	13037	2607.40
Crossover,Factory Tuner,Performance	4	840	210.00
Crossover,Flex Fuel	64	132720	2073.75
Crossover,Flex Fuel,Luxury	10	11732	1173.20
Crossover,Flex Fuel,Luxury,Performance	6	9744	1624.00
Crossover,Flex Fuel,Performance	6	33942	5657.00
Crossover,Hatchback	72	120650	1675.69
Crossover,Hatchback,Factory Tuner,Performance	6	12054	2009.00
Crossover,Hatchback,Luxury	7	1428	204.00
Crossover,Hatchback,Performance	6	12054	2009.00
Crossover,Hybrid	42	107662	2563.38
Crossover,Luxury	406	361021	889.21
Crossover,Luxury,Diesel	34	73080	2149.41
Crossover,Luxury,High-Performance	9	9335	1037.22
Crossover,Luxury,Hybrid	24	15142	630.92
Crossover,Luxury,Performance	112	151098	1349.09
Crossover,Luxury,Performance,Hybrid	2	7832	3916.00
Crossover,Performance	69	178431	2585.96
Diesel	84	145396	1730.90
Diesel,Luxury	47	113557	2416.11
Exotic,Factory Tuner,High-Performance	21	21974	1046.38
Exotic,Factory Tuner,Luxury,High-Performance	51	26674	523.02
Exotic,Factory Tuner,Luxury,Performance	3	1560	520.00
Exotic,Flex Fuel,Factory Tuner,Luxury,High-Performance	13	6760	520.00
Exotic,Flex Fuel,Luxury,High-Performance	11	5720	520.00
Exotic,High-Performance	254	325132	1280.05
Exotic,Luxury	12	1352	112.67
Exotic,Luxury,High-Performance	77	36423	473.03

Exotic,Luxury,High-Performance,Hybrid	1	204	204.00
Exotic,Luxury,Performance	36	7813	217.03
Exotic,Performance	10	13910	1391.00
Factory Tuner,High-Performance	104	204510	1966.44
Factory Tuner,Luxury	2	1234	617.00
Factory Tuner,Luxury,High-Performance	215	458674	2133.37
Factory Tuner,Luxury,Performance	31	43816	1413.42
Factory Tuner,Performance	84	149020	1774.05
Flex Fuel	855	1902985	2225.71
Flex Fuel,Diesel	16	90512	5657.00
Flex Fuel,Factory Tuner,Luxury,High-Performance	1	258	258.00
Flex Fuel,Hybrid	2	310	155.00
Flex Fuel,Luxury	39	29115	746.54
Flex Fuel,Luxury,High-Performance	32	28746	898.31
Flex Fuel,Luxury,Performance	28	38642	1380.07
Flex Fuel,Performance	87	146201	1680.47
Flex Fuel,Performance,Hybrid	2	310	155.00
Hatchback	574	751167	1308.65
Hatchback,Diesel	14	12222	873.00
Hatchback,Factory Tuner,High-Performance	13	15667	1205.15
Hatchback,Factory Tuner,Luxury,Performance	9	7982	886.89
Hatchback,Factory Tuner,Performance	21	45648	2173.71
Hatchback,Flex Fuel	7	39599	5657.00
Hatchback,Hybrid	64	135114	2111.16
Hatchback,Luxury	45	59541	1323.13
Hatchback,Luxury,Hybrid	3	1362	454.00
Hatchback,Luxury,Performance	36	58761	1632.25
Hatchback,Performance	198	212585	1073.66
High-Performance	198	361029	1823.38
Hybrid	121	256107	2116.59
Luxury	819	883877	1079.21
Luxury,High-Performance	334	557118	1668.02
Luxury,High-Performance,Hybrid	12	6826	568.83
Luxury,Hybrid	52	35029	673.63
Luxury,Performance	659	852128	1293.06
Luxury,Performance,Hybrid	11	25665	2333.18
N/A	3376	5620476	1664.83
Performance	520	735909	1415.21
Performance,Hybrid	1	155	155.00
<b>Grand Total</b>	<b>11199</b>	<b>17453455</b>	<b>1558.48</b>

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

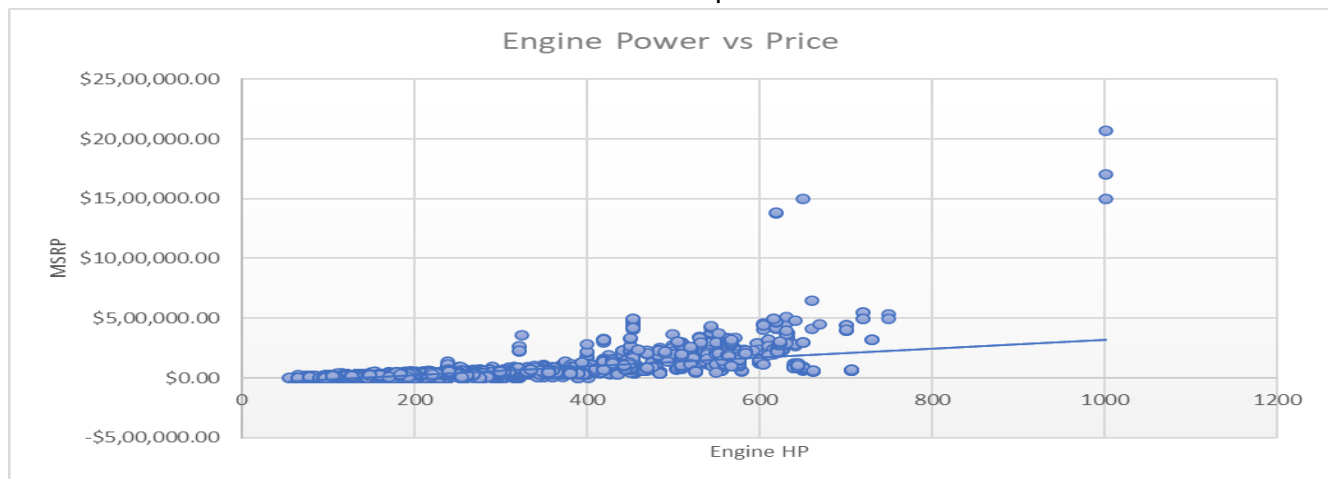


From the above pivot table and combo chart, it is seen that:

- “Flex Fuel,Diesel” Category is most popular with average popularity score of 5657
- “Exotic,Luxury” Category is least popular with average popularity score of 112.67

**Insight Required:** What is the relationship between a car's engine power and its price?

- **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.



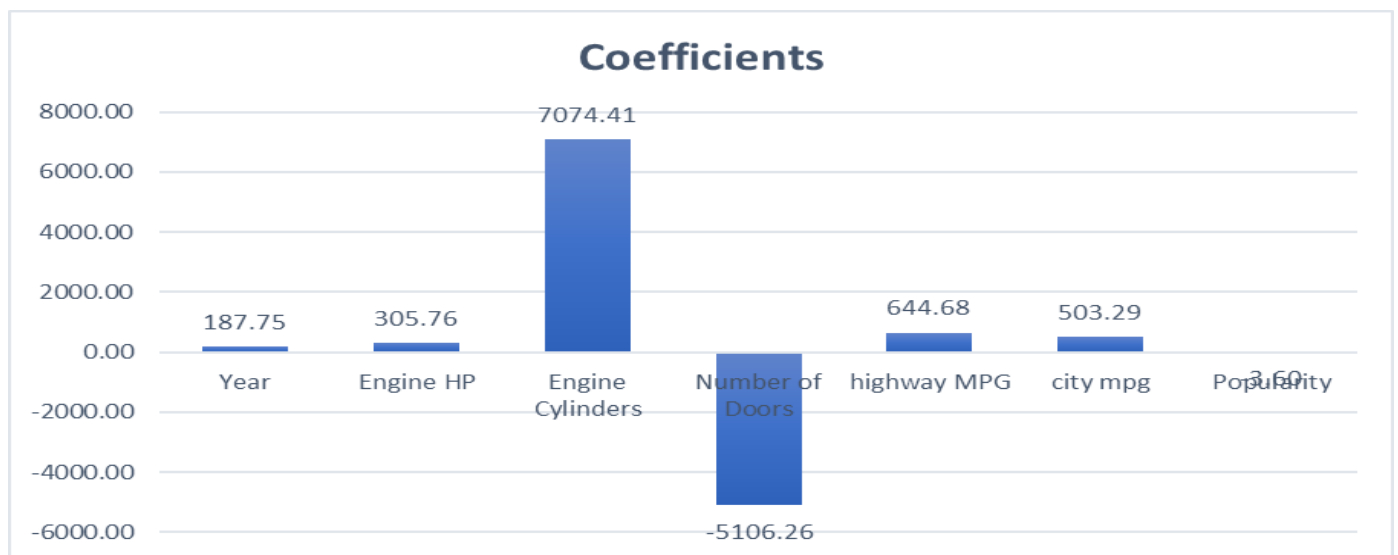
From above scatter plot it is seen that:

- The trendline slopes upward from left to right, it shows a positive correlation between Engine Power and its price.
- This states that cars with more powerful engines tend to be more expensive.
- This insight is valuable for market segmentation and target specific customers who want cars within specific engine power and price range combination.

**Insight Required:** Which car features are most important in determining a car's price?

- **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.

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-456858.41	155407.7154	-2.939741	0.00329158	-761484.8821	-152232	-761485	-152232
Year	187.75	77.84762973	2.4117339	0.01589286	35.15271151	340.3428	35.15271	340.3428
Engine HP	305.76	7.567317878	40.405299	0	290.9264696	320.593	290.9265	320.593
Engine Cylinder	7074.41	475.0066455	14.893277	1.089E-49	6143.309065	8005.502	6143.309	8005.502
Number of Doors	-5106.26	523.4883008	-9.75429	2.1713E-22	-6132.386041	-4080.13	-6132.39	-4080.13
highway MPG	644.68	108.3586982	5.9494989	2.7701E-09	432.2778391	857.0821	432.2778	857.0821
city mpg	503.29	101.6322301	4.9520219	7.4518E-07	304.0679705	702.5021	304.068	702.5021
Popularity	-3.60	0.29822306	-12.06366	2.6437E-33	-4.182230506	-3.01309	-4.18223	-3.01309



The above chart shows that “Engine Cylinder” has highest coefficient value of 7074.41 making it the most important car feature in determining a car's price, followed by highway MPG, city mpg and Engine HP

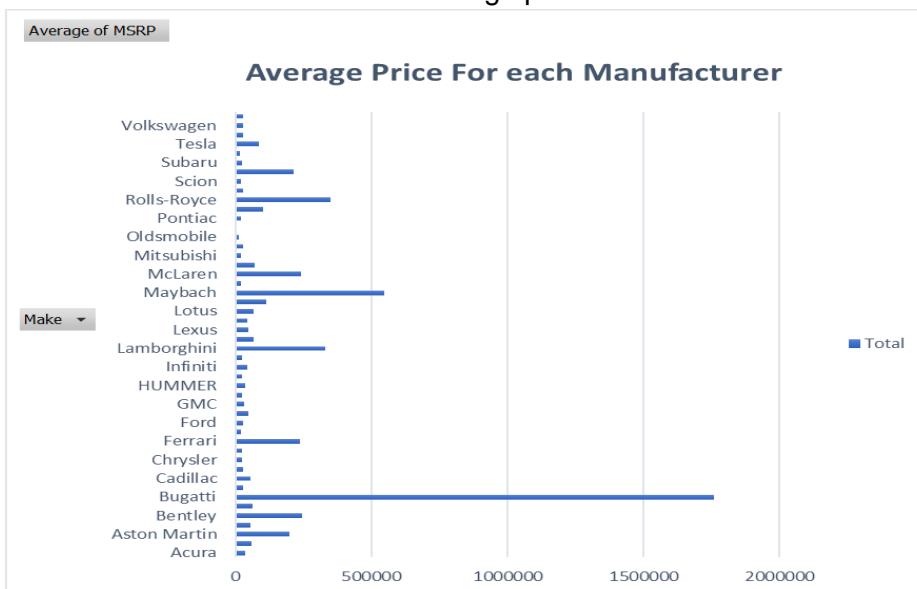
**Insight Required:** How does the average price of a car vary across different manufacturers?

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

Manufacture	Average of MSRP
Acura	35087.4878
Alfa Romeo	61600
Aston Martin	198123.4615
Audi	54574.1215
Bentley	247169.3243
BMW	62162.55864
Bugatti	1757223.667
Buick	29034.18947
Cadillac	56368.26515
Chevrolet	29074.72576
Chrysler	26722.96257
Dodge	24857.04537
Ferrari	238218.8406
FIAT	22670.24194
Ford	28511.30788
Genesis	46616.66667
GMC	32444.08506
Honda	26655.14781

HUMMER	36464.41176
Hyundai	24926.26255
Infiniti	42640.27134
Kia	25513.75546
Lamborghini	331567.3077
Land Rover	68067.08633
Lexus	47549.06931
Lincoln	43860.825
Lotus	68377.14286
Maserati	113684.4909
Maybach	546221.875
Mazda	20416.62379
McLaren	239805
Mercedes-Benz	72069.52786
Mitsubishi	21340.5625
Nissan	28921.15245
Oldsmobile	12843.79545
Plymouth	3296.873239
Pontiac	19800.0442
Porsche	101622.3971
Rolls-Royce	351130.6452
Saab	27879.80734
Scion	19932.5
Spyker	214990
Subaru	24240.67364
Suzuki	18026.4152
Tesla	85255.55556
Toyota	28846.5605
Volkswagen	28978.52289
Volvo	29724.68421
<b>Grand Total</b>	<b>41925.92714</b>

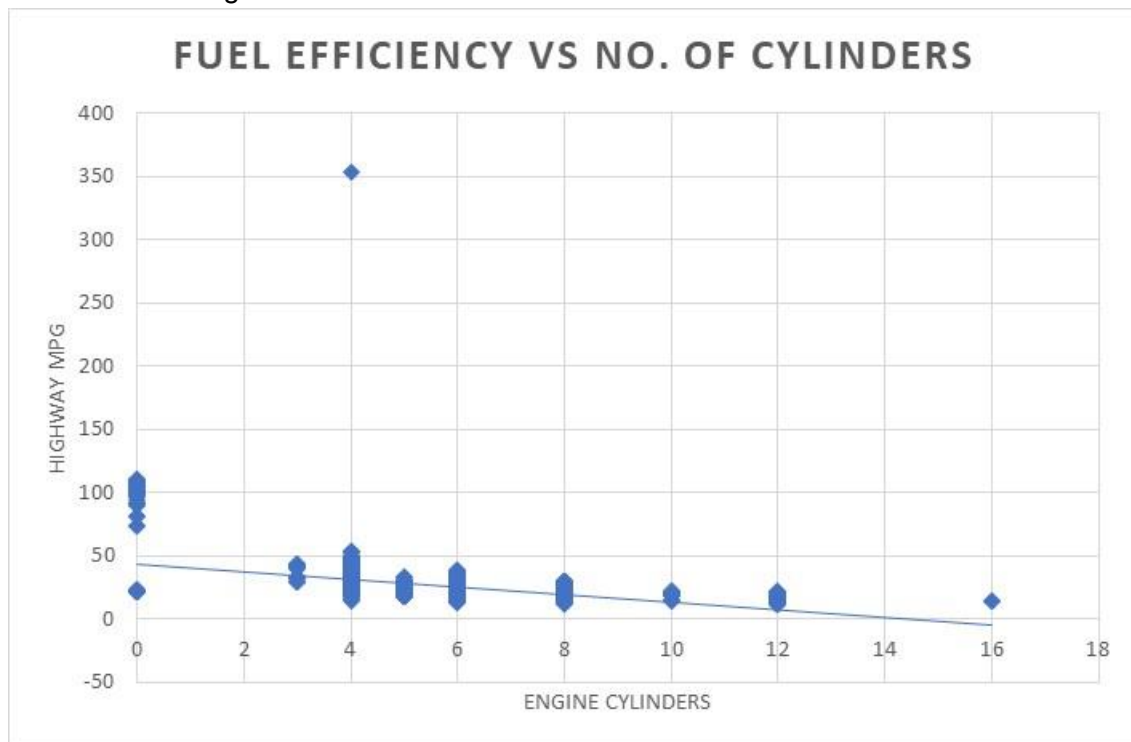
- **Task 4.B:** Create a bar chart or a horizontal stacked bar chart that visualizes the relationship between manufacturer and average price.



The above pivot table and chart shows that Bugatti manufacturer has the highest average price of 1757223.67 and Plymouth has least average price of 3296.87

**Insight Required:** What is the relationship between fuel efficiency and the number of cylinders in a car's engine?

- **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.



The above chart shows a negative trendline which means that as no. of cylinder increases the highway MPG decreases.

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

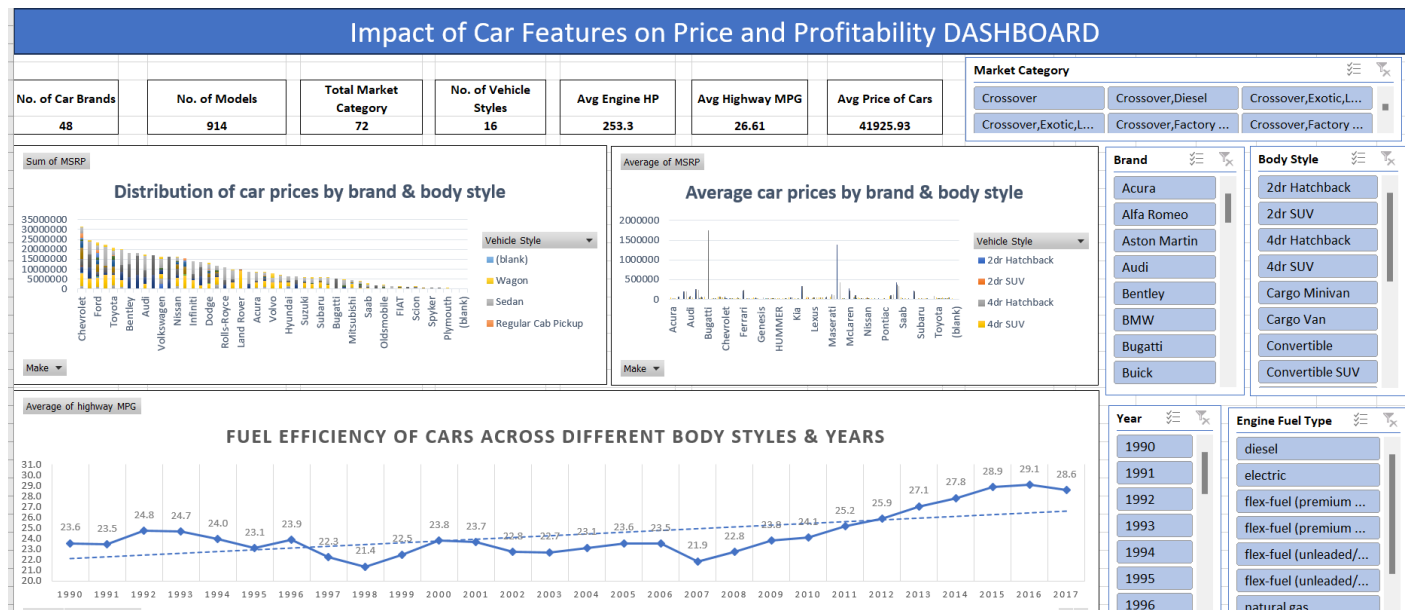
Correlation co-efficient



# Building the Dashboard:

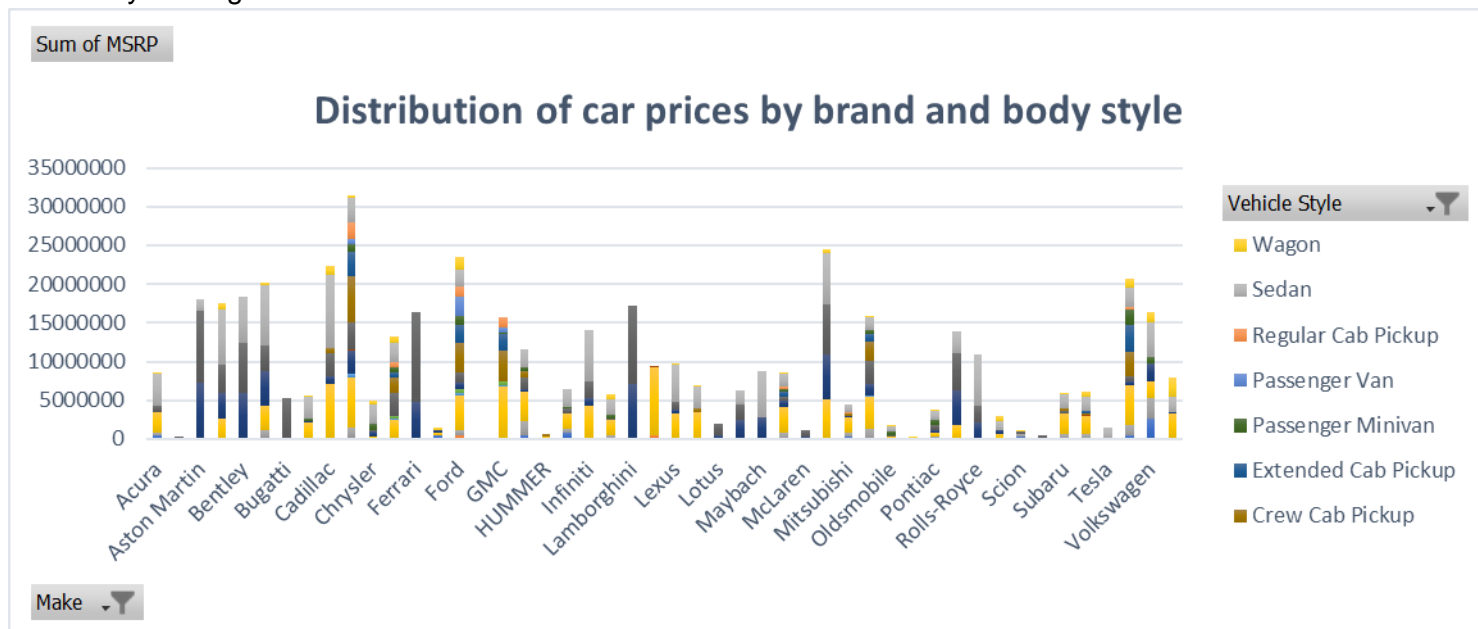
The Interactive Dashboard created using MS EXCEL:

<https://drive.google.com/uc?export=download&id=1dwX6k24mLE-S9bMnnLhUG7jYimOj4BC>



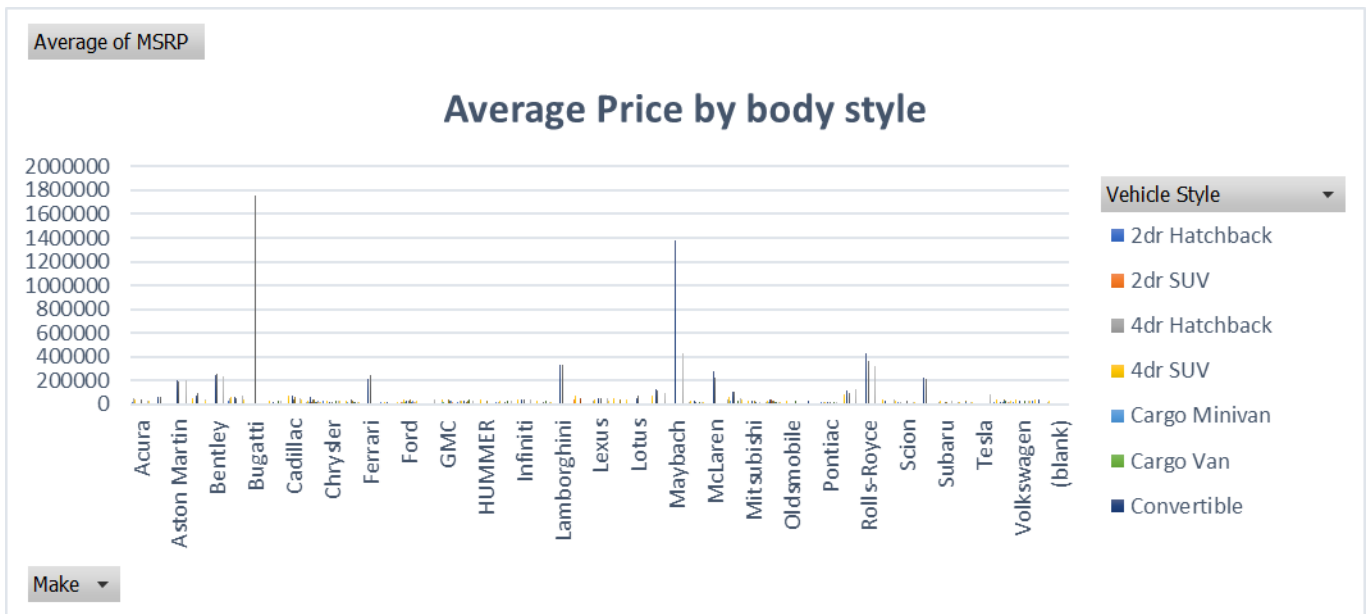
**Task 1:** How does the distribution of car prices vary by brand and body style?

- Hints:** Stacked column chart to show the distribution of car prices by brand and body style. Use filters and slicers to make the chart interactive. Calculate the total MSRP for each brand and body style using SUMIF or Pivot Tables.



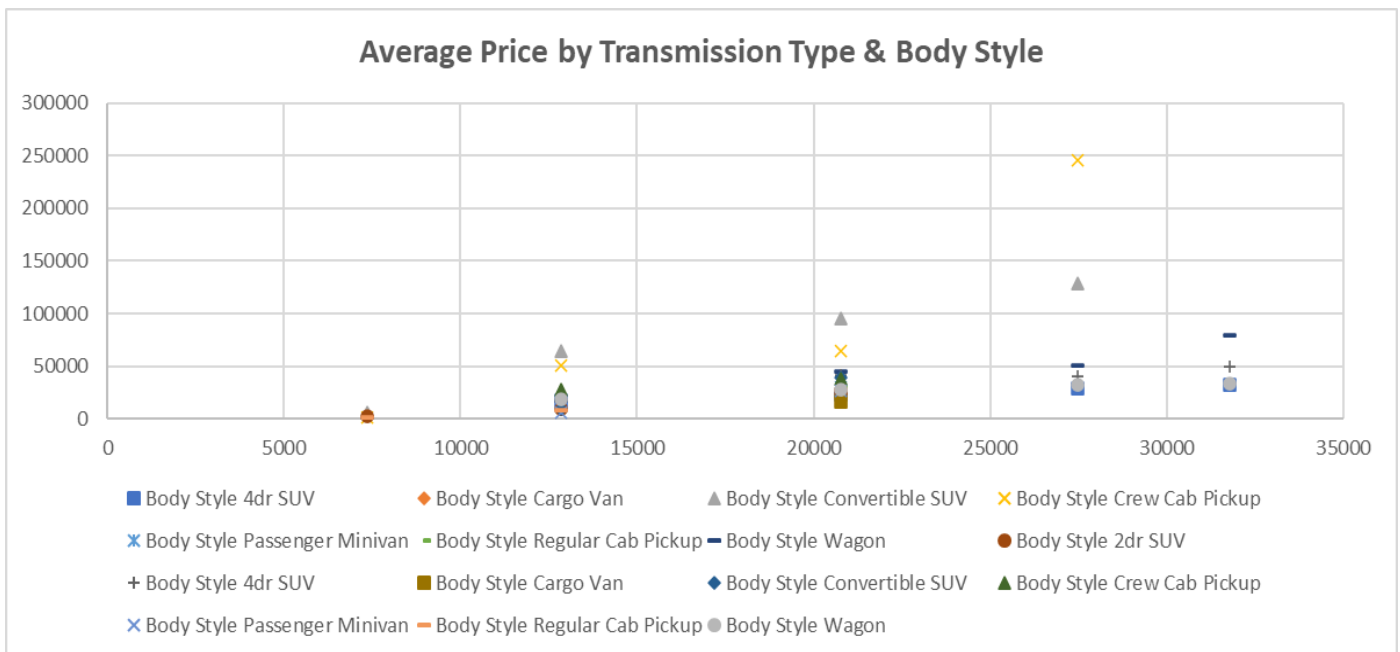
**Task 2:** Which car brands have the highest and lowest average MSRPs, and how does this vary by body style?

- Hints:** Clustered column chart to compare the average MSRPs across different car brands and body styles. Calculate the average MSRP for each brand and body style using AVERAGEIF or Pivot Tables.



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

- Hints:** Scatter plot chart to visualize the relationship between MSRP and transmission type, with different symbols for each body style. Calculate the average MSRP for each combination of transmission type and body style using AVERAGEIFS or Pivot Tables.

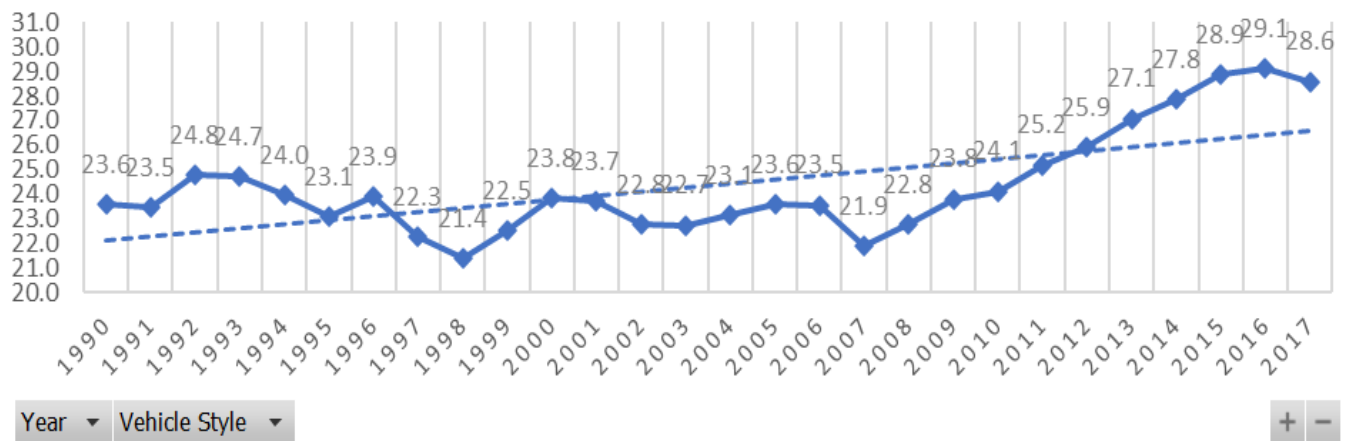


**Task 4:** How does the fuel efficiency of cars vary across different body styles and model years?

- Hints:** Line chart to show the trend of fuel efficiency (MPG) over time for each body style. Calculate the average MPG for each combination of body style and model year using AVERAGEIFS or Pivot Tables.

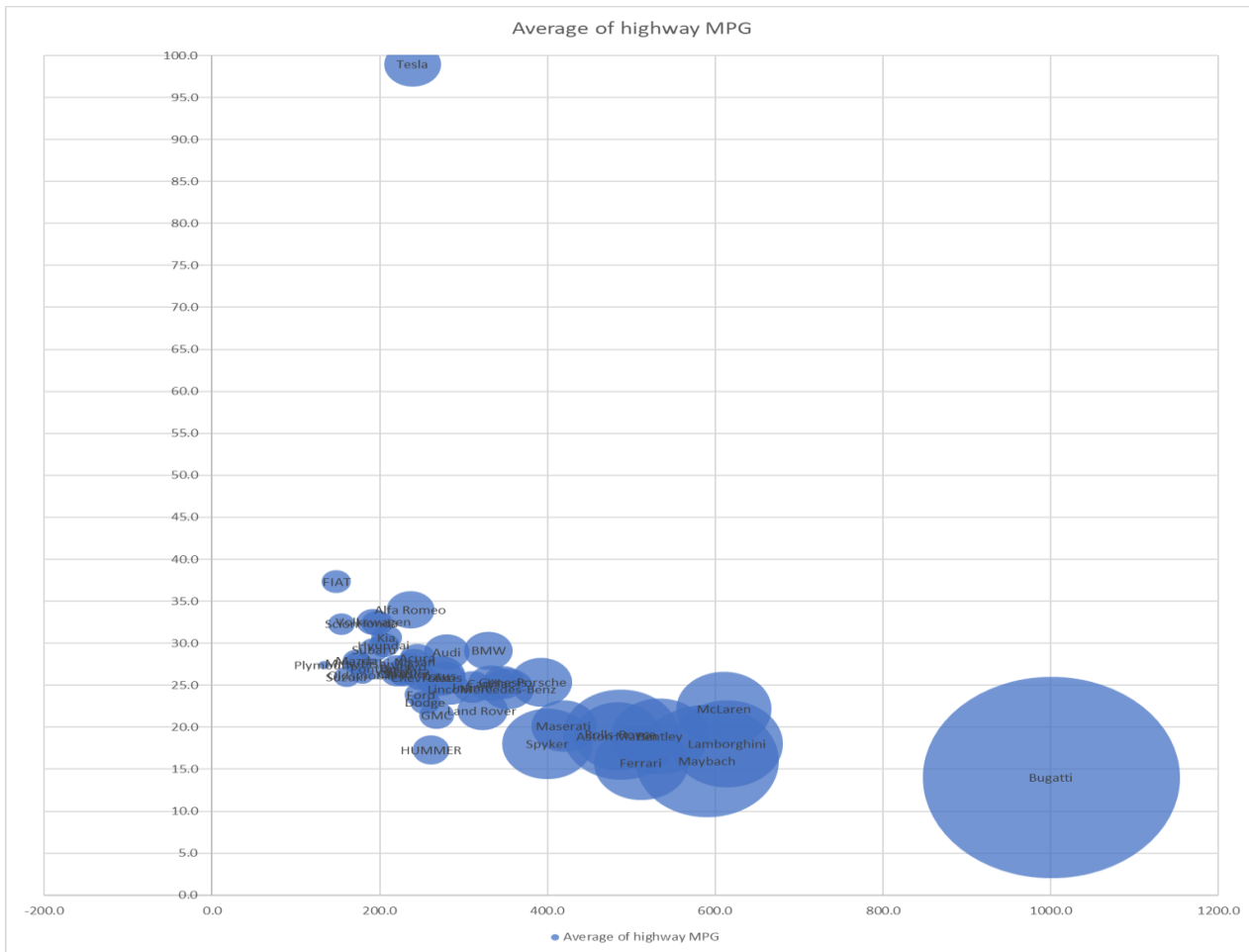
Average of highway MPG

## FUEL EFFICIENCY OF CARS ACROSS DIFFERENT BODY STYLES & YEARS



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

- **Hints:** Bubble chart to visualize the relationship between horsepower, MPG, and price across different car brands. Assign different colors to each brand and label the bubbles with the car model name. Calculate the average horsepower, MPG, and MSRP for each car brand using AVERAGEIFS or Pivot Tables.



### Result:

From above charts shown in the dashboard following insights are drawn:

- Chevrolet brand has highest car price having highest share of 4dr SUV vehicle style
- Bugatti has highest average price followed by Maybach
- Fuel efficiency has an upward trend over the years
- Tesla has highest highway MPG, but Bugatti has highest engine power and price

### Links:

#### Raw Dataset Link:

[https://drive.google.com/file/d/1HgHNOQx\\_wKnvApgBQ5DuKS1KlXmkDWPC/view?usp=share\\_link](https://drive.google.com/file/d/1HgHNOQx_wKnvApgBQ5DuKS1KlXmkDWPC/view?usp=share_link)

#### Dashboard Link:

<https://drive.google.com/uc?export=download&id=1dwX6k24mLE-S9bMnnLIhUG7jYjmOj4BC>

#### Analysis Folder having workfile link:

[https://drive.google.com/drive/folders/1TBIHX8eHyRIKPxeYXBKMg4wdOIDjldnp?usp=drive\\_link](https://drive.google.com/drive/folders/1TBIHX8eHyRIKPxeYXBKMg4wdOIDjldnp?usp=drive_link)