



Final Project-3

PROJECT 7

Analyzing the Impact of Car
Features on Price and Profitability

RAKSHA NAYAK

Project Description: The automotive industry has experienced rapid growth, with a focus on fuel efficiency, environmental sustainability, and technological innovation. As competition increases and consumer preferences shift, it is crucial to understand the factors driving car demand. Electric and hybrid vehicles are becoming popular as well as alternative fuel sources like hydrogen and natural gas are also becoming popular. Traditional gasoline-powered cars remain dominant, with varying fuel types and grades available. To optimize pricing and product development decisions, car manufacturers can analyze the relationship between features, market categories, and pricing, identifying popular features and profitable categories

This Project “**Analyzing the Impact of Car Features on Price and Profitability**” is designed to provide valuable insights to a car manufacturer and help them optimize their pricing and product development decisions to maximize profitability while meeting consumer demand.

By using data analysis techniques such as **Regression Analysis and Market Segmentation**, the manufacturer could develop a pricing strategy that balances consumer demand with profitability, and identify which product features to focus on in future product development efforts. This could help the manufacturer improve its competitiveness in the market and increase its profitability over time.

Approach: The dataset contains information on various car models and their specification. Below is the link of the dataset provided

<https://docs.google.com/spreadsheets/d/1iZzkW22BDBJLurKGBTizcl2vcjp4RqTs/edit?gid=1669281338#gid=1669281338>

The dataset after analysis with answers, insights and visualization is,

https://docs.google.com/spreadsheets/d/1Lg2cWRQWa1Nyxu2qE5jaEWMY_54HhSQ7/edit?usp=sharing&oid=108154584635151678812&rtpof=true&sd=true

The dataset contains,

- **Total data points/observations: 11,915**
- **Attributes/Columns: 16**
- **File type: CSV (Comma Separated Values)**

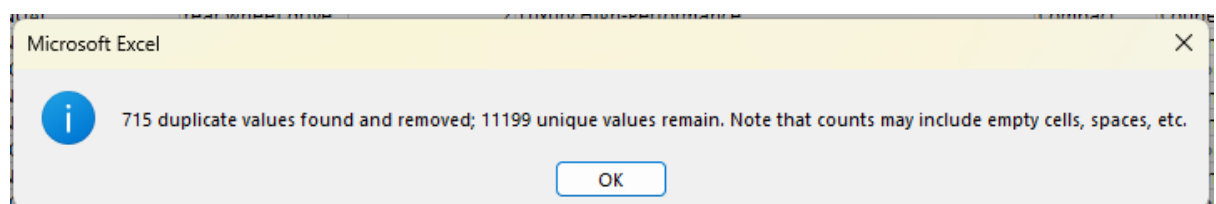
Column names:

1. **Make:** the make or brand of the car
2. **Model:** the specific model of the car
3. **Year:** the year the car was released
4. **Engine Fuel Type:** the type of fuel used by the car (gasoline, diesel, etc.)
5. **Engine HP:** the horsepower of the car's engine
6. **Engine Cylinders:** the number of cylinders in the car's engine
7. **Transmission Type:** the type of transmission (automatic or manual)
8. **Driven_Wheels:** the type of wheels driven by the car (front, rear, all)
9. **Number of Doors:** the number of doors the car has
10. **Market Category:** the market category the car belongs to (Luxury, Performance, etc.)
11. **Vehicle Size:** the size of the car
12. **Vehicle Style:** the style of the car (Sedan, Coupe, etc.)
13. **Highway MPG:** the estimated miles per gallon the car gets on the highway
14. **City MPG:** the estimated miles per gallon the car gets in the city
15. **Popularity:** a ranking of the popularity of the car (based on the number of times it has been viewed on Edmunds.com)
16. **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.

After downloading the data set, we need to pre-process and clean (Data Clean) it, this is one of the most important steps to perform before performing analysis.

1. **Omission/Dropping Columns/Removing unwanted columns:** All 16 columns are important so we are going to retain them.
2. **Handling Duplicate Values:** I have used '**DATA**' tabs, *Remove Duplicates* option on entire column range to look for a duplicate. 715 duplicates are found. After deleting them, total number of records/data points are 11200.

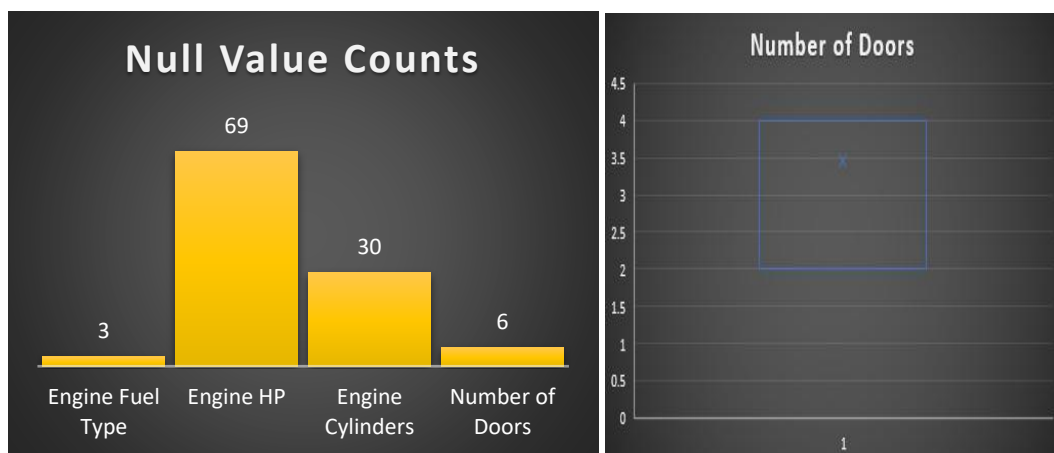


The records are **now reduced from 11915 to 11200**

1. Data Cleaning - Removing Duplicate Rows	
Initial number of Rows	11915
Duplicate Rows	715
After Removing Duplicate Rows	11200

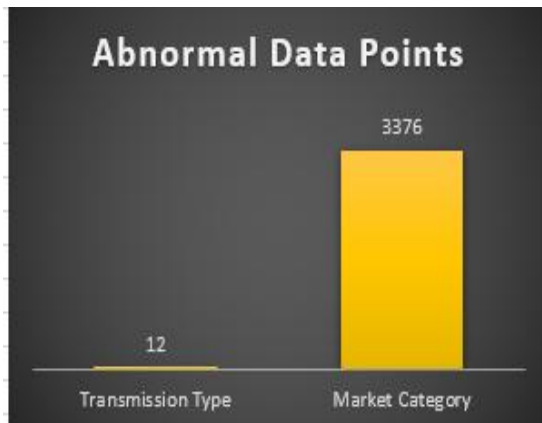
3. **Handling Missing Data:** We need to check if there are any missing values/blanks in the dataset.

Coulumns with Blanks	Number of Blanks	Imputation Method	Reasons	Value
Engine Fuel Type	3	Mode	Categorical Data	regular unleaded
Engine HP	69		0 As they are all Electric Cars	0
Engine Cylinders	30		0 As they are all Electric Cars	0
Number of Doors	6	Mean	They have no outliers so using Mean	4



- There are 3 null values in Engine Fuel Type column, which is a **categorical data** so imputing the blanks with the **MODE**.
- There are 69 null values in Engine HP column.
 - a) We will be filling the null values with **0** if they are electric cars.
 - b) We will use [AI](#) to search for a specific Make, Model, Year and Engine Fuel Type to fetch the Engine HP value and impute the null values with them.
- There are 30 null values in Engine Cylinders column.
 - a) We know that electric cars don't have any cylinders so the null values will be again filled with **0**.
 - b) For Mazda's(Petrol Engine) with Model RX-7 and RX-8 uses a rotary engine and thus do not have any cylinders. So, filling null values with 0.
- There are 6 null values in Number of Doors Column. The above box and whiskers plot shows that they have any outliers. Thus, the best measure of central tendency is **MEAN** in this scenario. Therefore, we shall impute the null values with the MEAN using excel formula, **=ROUNDUP(AVERAGE(I1:I11200),0)**

4. **Handling Abnormal Data Points:** Transmission Type and Market Category has abnormal values of Unknown and N/A.



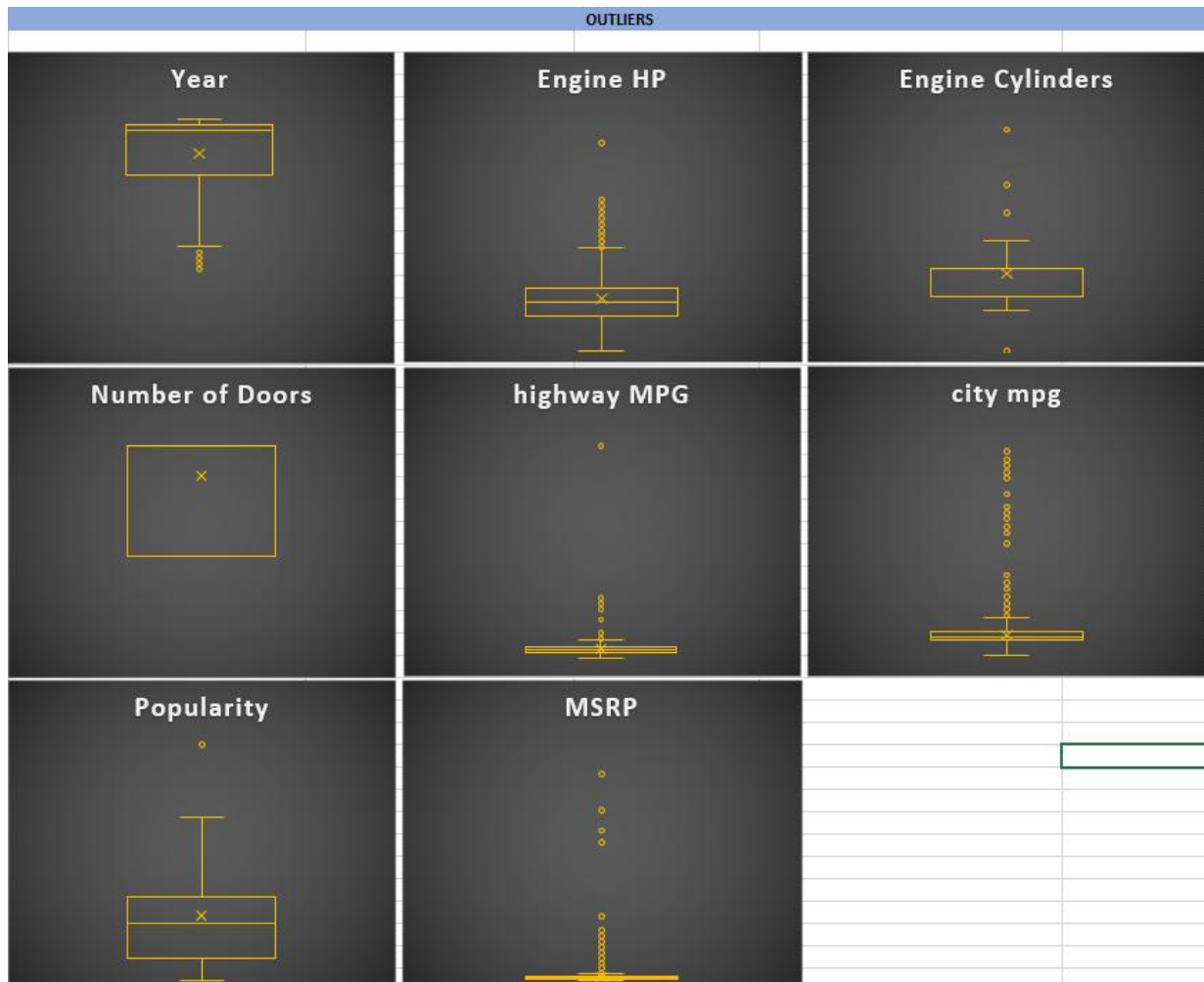
- We can use **MODE** to replace abnormal data value of UNKNOWN in column, Transmission Type as it is a Categorical Data.

2. Data Cleaning - Handling Abnormal Data Points			
Columns with Abnc Number of Abnormal Data Points	Abnormal Values	Imputation Method	Value
Transmission Type	12 Unknown	Mode as Categorical Data	Automatic
Market Category	3376 N/A	As no enough information	N/A
MODE COMPUTATION			
Transmission Type	Count		
MANUAL	2634		
AUTOMATIC	7944		
AUTOMATED_MANUAL	553		
DIRECT_DRIVE	68		
UNKNOWN	0		
	11199		

- N/A in Market Category are left as it is as enough information is not available to impute them.
5. **Outlier Detection:** Let's plot a box and whisker plot for each numerical variable to spot outliers.

Numerical Variable

- Year – Outlier Present
- Engine HP – Outlier Present
- Engine Cylinders – Outlier Present
- Number of Doors – No Outlier
- Highway MPG – Outlier Present
- City mpg – Outlier Present
- Popularity – Outlier Present
- MSRP – Outlier Present



6. Removing/Replacing Outliers:

- Engine Cylinders – Outlier Present – Outliers with 0 Cylinders are left as it is as they refer to Electric Cars with motors and no engines.
- Highway MPG – Outlier Present – Row 1085 with Audi A6 is removed as it is impossible to have 354 MPG.
- City mpg – Outlier Present – Left as it is as they are either electric or hybrid cars.
- MSRP – Outlier Present – Left as it is as they are luxury car brands

7. **Data Summary:** We have cleaned the data by managing missing values, errors and replacing the outliers with actual values. Summary of the cleaned data is as shown below,

DESCRIPTIVE STATISTICS					
Engine HP		Engine Cylinders		Number of Doors	
Mean	252.3550634	Mean	5.650919807	Mean	3.454366851
Standard Error	1.048761456	Standard Error	0.017183494	Standard Error	0.008248179
Median	237	Median	6	Median	4
Mode	200	Mode	4	Mode	4
Standard Deviation	110.9805695	Standard Deviation	1.818367717	Standard Deviation	0.872827248
Sample Variance	12316.68682	Sample Variance	3.306461153	Sample Variance	0.761827406
Kurtosis	2.146958778	Kurtosis	1.982211159	Kurtosis	-0.908002024
Skewness	1.192488217	Skewness	0.8458491	Skewness	-1.019594688
Range	1001	Range	16	Range	2
Minimum	0	Minimum	0	Minimum	2
Maximum	1001	Maximum	16	Maximum	4
Sum	2825872	Sum	63279	Sum	38682
Count	11198	Count	11198	Count	11198

highway MPG		city mpg		Popularity		MSRP	
Mean	26.58135381	Mean	19.7314699	Mean	1558.34524	Mean	41925.06323
Standard Error	0.079644587	Standard Error	0.08673052	Standard Error	13.6614185	Standard Error	581.5287948
Median	25	Median	18	Median	1385	Median	30675
Mode	24	Mode	17	Mode	1385	Mode	2000
Standard Deviation	8.428038251	Standard Deviation	9.1778763	Standard Deviation	1445.65954	Standard Deviation	61537.72767
Sample Variance	71.03182876	Sample Variance	84.2334134	Sample Variance	2089931.52	Sample Variance	3786891927
Kurtosis	33.8921754	Kurtosis	72.9304035	Kurtosis	2.20089362	Kurtosis	259.3544241
Skewness	4.154241858	Skewness	7.09366463	Skewness	1.62865906	Skewness	11.60927131
Range	99	Range	130	Range	5655	Range	2063902
Minimum	12	Minimum	7	Minimum	2	Minimum	2000
Maximum	111	Maximum	137	Maximum	5657	Maximum	2065902
Sum	297658	Sum	220953	Sum	17450350	Sum	469476858
Count	11198	Count	11198	Count	11198	Count	11198

Tech Stack Used:

- **Microsoft Excel Version 2407, 2019** – Excel is a spreadsheet editor developed by Microsoft. It features calculation or computation capabilities, graphing tools, pivot tables etc.
- **Tableau Public 2021.2** — A visualization tool to represent data in graphs and plots. Mainly used to create interactive Dashboards.

Approach:

Analytical Method: We are using different analytical methods like data visualization, descriptive statistics, regression analysis and correlation analysis. This analysis will help us to find the pattern, trend, relationships and insights.

Modelling Technique: We are using Regression Analysis to identify the key factors influencing car prices. This technique helps to quantify the relationship between independent variables (various car features) and the dependent variable (car price). By examining the coefficient values, we will be able to assess the relative importance of different features.

Data Analytics Tasks:

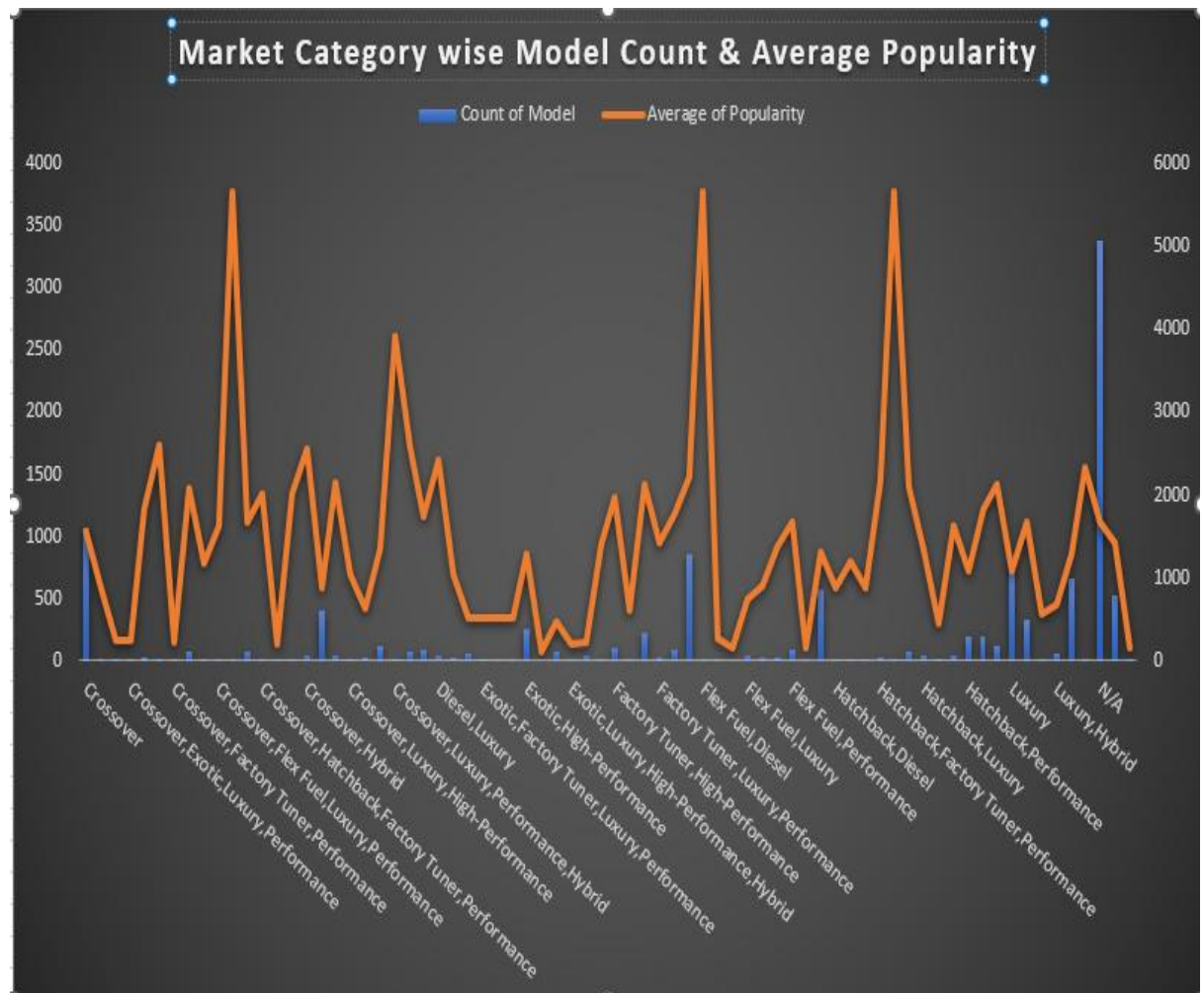
Insight Required: How does the popularity of a car model vary across different market categories?

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

Output of Task 1.A:

Row Labels	Count of Model	Average of Popularity
Crossover	1075	1556
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
Crossover,Factory Tuner,Luxury,Performance	5	2607
Crossover,Factory Tuner,Performance	4	210
Crossover,Flex Fuel	64	2074
Crossover,Flex Fuel,Luxury	10	1173
Crossover,Flex Fuel,Luxury,Performance	6	1624
Crossover,Flex Fuel,Performance	6	5657
Crossover,Hatchback	72	1676
Crossover,Hatchback,Factory Tuner,Performance	6	2009
Crossover,Hatchback,Luxury	7	204
Crossover,Hatchback,Performance	6	2009
Crossover,Hybrid	42	2563
Crossover,Luxury	406	889
Crossover,Luxury,Diesel	34	2149
Crossover,Luxury,High-Performance	9	1037
Crossover,Luxury,Hybrid	24	631
Crossover,Luxury,Performance	112	1349
Crossover,Luxury,Performance,Hybrid	2	3916
Crossover,Performance	69	2586
Diesel	84	1731
Diesel,Luxury	47	2416
Exotic,Factory Tuner,High-Performance	21	1046
Exotic,Factory Tuner,Luxury,High-Performance	51	523
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	254	1280
Exotic,Luxury	12	113
Exotic,Luxury,High-Performance	77	473
Exotic,Luxury,High-Performance,Hybrid	1	204
Exotic,Luxury,Performance	36	217
Exotic,Performance	10	1391
Factory Tuner,High-Performance	104	1966

Visual Representation: Output of Task 1.B



Insight: From above combo chart, we can identify the most popular and least popular market category along with most selling model/car as below.

MOST POPULAR MARKET CATEGORY	LEAST POPULAR MARKET CATEGORY	HIGHEST SELLING MODEL/CAR
Crossover,Flex Fuel,Performance	Exotic,Luxury	Crossover
Flex Fuel,Diesel		
Hatchback,Flex Fuel		

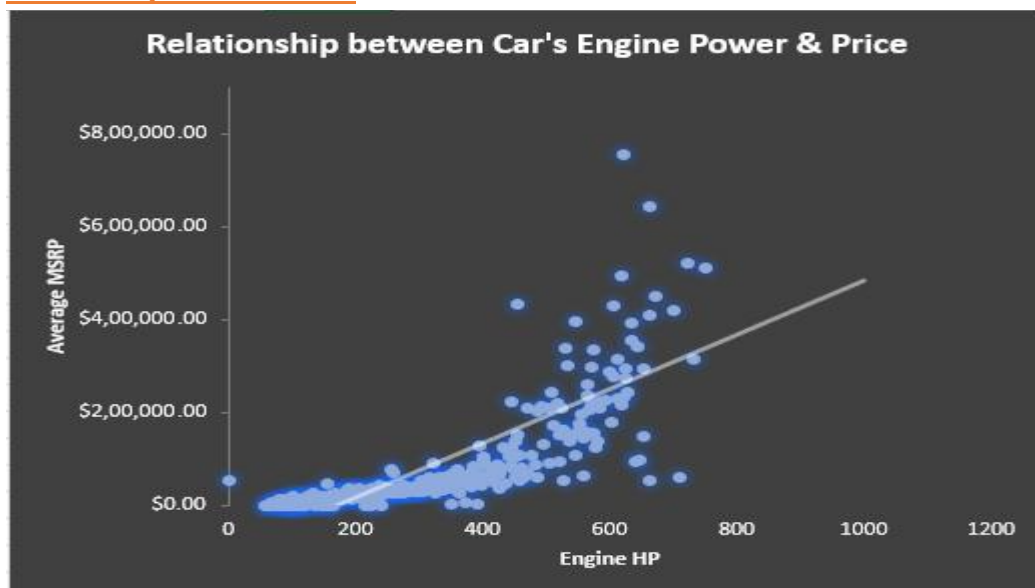
Insight Required: What is the relationship between a car's engine power and its price?

- **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.**

Output:

Engine HP	Average of MSRP
0	54745.11364
55	2000
62	2000
63	2000
66	7998.571429
73	2000
74	8116.944444
78	15082.5
79	5910.090909
81	2000
82	2000
84	14493.33333
88	2000
90	2000
92	2000
93	2015.791667
94	10096.42857
95	2006.3
96	2000
97	2000
98	18189.58333
99	21377.69231
100	6146.5
101	18061.36364
102	2000
103	13152.5

Visual Representation:



Insight: The above scatter plot shows a positive slope trendline. Thus, displaying a **positive linear relationship** between Car's Engine Power (Engine HP) and Price.

Higher the Engine power, Higher the price due to more complex level of design, engineering and expensive subparts to give a better performance.

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.**

Regression Analysis: Regression analysis is a set of statistical methods used for the estimation of relationships between a dependent variable and one or more independent variables. It can be utilized to assess the strength of the relationship between variables and for modelling the future relationship between them.

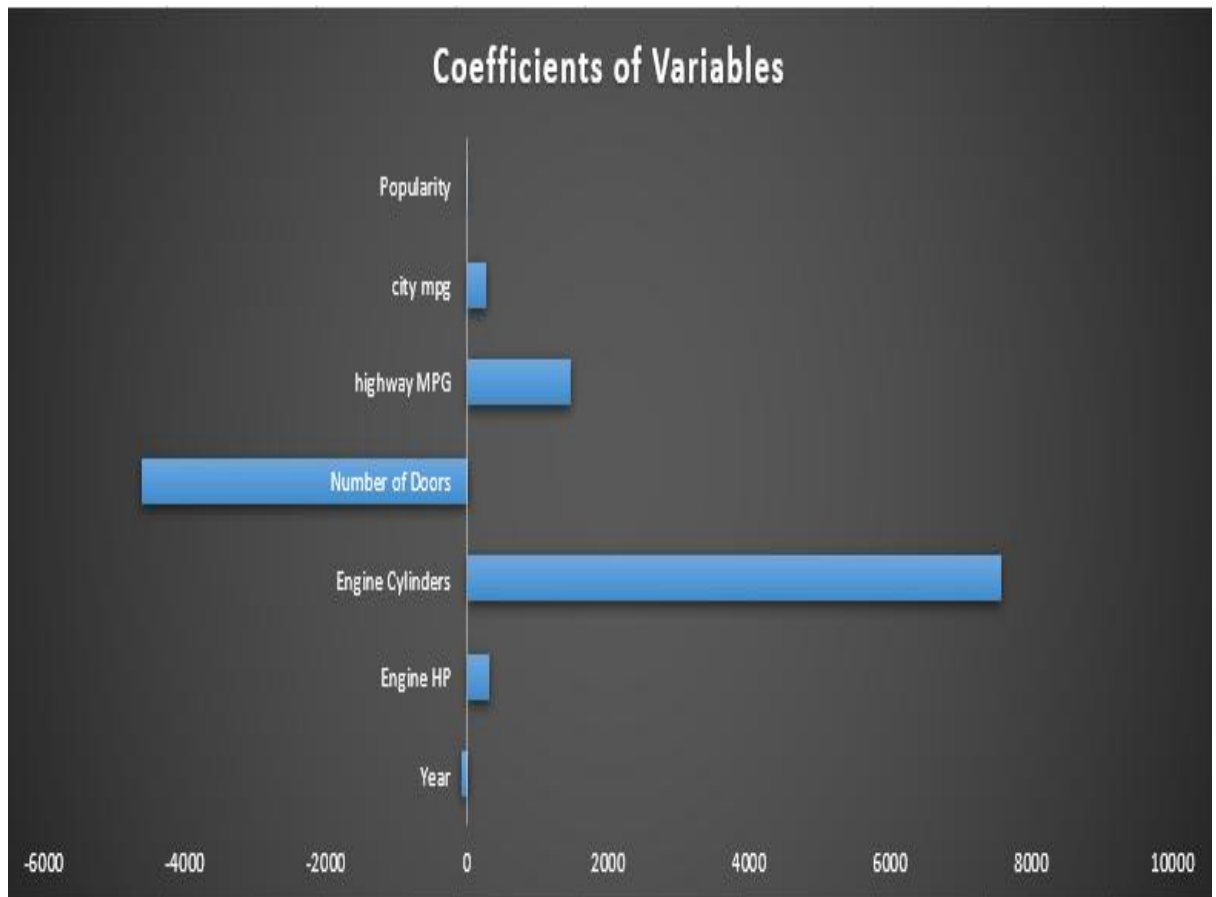
In regression analysis, a coefficient is a numerical value that represents the relationship between an independent variable and a dependent variable.

A larger regression coefficient indicates that the variable has a larger impact on the model output.

Output:

SUMMARY OUTPUT								
<i>Regression Statistics</i>								
Multiple R	0.682493293							
R Square	0.465797095							
Adjusted R Square	0.46546292							
Standard Error	44991.48979							
Observations	11198							
<i>ANOVA</i>								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	7	1.97506E+13	2.82152E+12	1393.870982	0			
Residual	11190	2.26512E+13	2024234153					
Total	11197	4.24018E+13						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	50775.97868	161532.2048	0.314339662	0.753268938	-265855.5733	367407.5307	-265855.5733	367407.5307
Year	-77.60045589	81.19720486	-0.955703537	0.339242681	-236.7612686	81.56035685	-236.7612686	81.56035685
Engine HP	321.7728466	7.771144247	41.40610911	0	306.5400361	337.0056571	306.5400361	337.0056571
Engine Cylinders	7578.741706	474.8725152	15.95952906	1.03124E-56	6647.907995	8509.575416	6647.907995	8509.575416
Number of Doors	-4604.306842	524.5605435	-8.777455527	1.91463E-18	-5632.537834	-3576.075851	-5632.537834	-3576.075851
highway MPG	1461.089722	170.6349986	8.562661438	1.24705E-17	1126.615092	1795.564352	1126.615092	1795.564352
city mpg	276.5423035	142.9131013	1.935038152	0.053010782	-3.592528694	556.6771357	-3.592528694	556.6771357
Popularity	-3.359863807	0.298220636	-11.26636928	2.76478E-29	-3.944428741	-2.775298872	-3.944428741	-2.775298872

Visual Representation:



Insight: From the above bar chart, it is evident that city mpg, highway MPG, Engine Cylinders and Engine HP have positive relationship in determining car's price.

The **strongest impact on car's price is dependent on Engine Cylinders** followed by highway MPG.

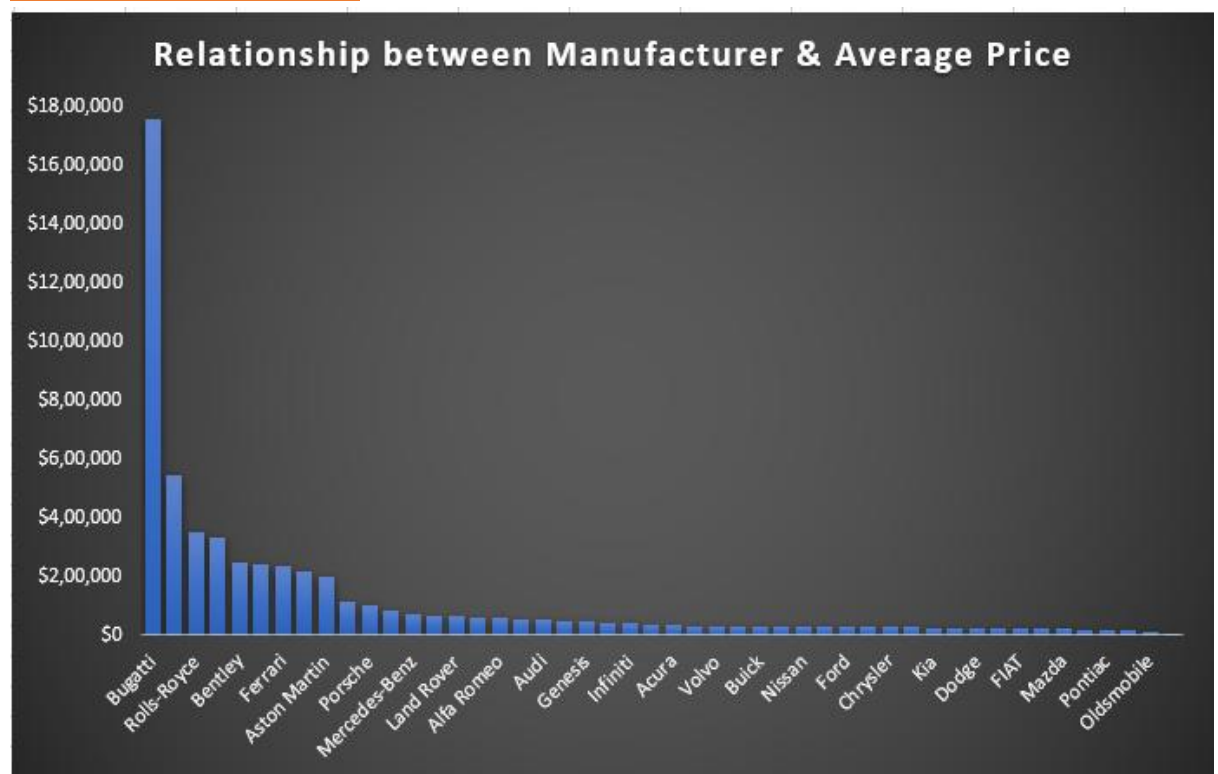
Insight Required: How does the average price of a car vary across different manufacturers?

- **Task 4.A:** Create a pivot table that shows the average price of cars for each manufacturer.
- **Task 4.B:** Create a bar chart or a horizontal stacked bar chart that visualizes the relationship between manufacturer and average price.

Output:

Row Labels	Average of MSRP
Bugatti	\$17,57,224
Maybach	\$5,46,222
Rolls-Royce	\$3,51,131
Lamborghini	\$3,31,567
Bentley	\$2,47,169
McLaren	\$2,39,805
Ferrari	\$2,38,219
Spyker	\$2,14,990
Aston Martin	\$1,98,123
Maserati	\$1,13,684
Porsche	\$1,01,622
Tesla	\$85,256
Mercedes-Benz	\$72,070
Lotus	\$68,377
Land Rover	\$68,067
BMW	\$62,163
Alfa Romeo	\$61,600
Cadillac	\$56,368
Audi	\$54,583
Lexus	\$47,549
Genesis	\$46,617
Lincoln	\$43,861
Infiniti	\$42,640
HUMMER	\$36,464
Acura	\$35,087
GMC	\$32,444
Volvo	\$29,725
Chevrolet	\$29,075
Buick	\$29,034
Volkswagen	\$28,979
Nissan	\$28,921
Toyota	\$28,847

Visual Representation:



Insight: From the above bar chart, we understand that Bugatti is the most expensive car followed by Rolls-Royce, Bentley, Ferrari. All these cars are luxury brands with high performance.

Plymouth is the cheapest car out of all.

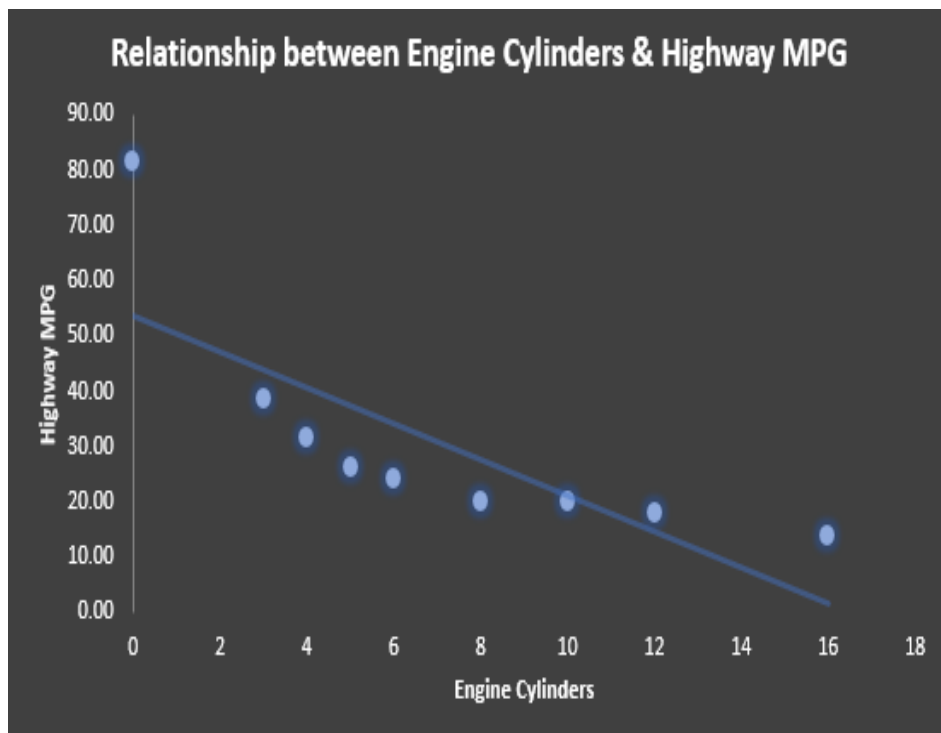
Insight Required: What is the relationship between fuel efficiency and the number of cylinders in a car's engine?

- **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.**

Output:

Engine Cylinders	Average of highway MPG
0	81.66
3	38.67
4	31.43
5	26.07
6	24.00
8	20.18
10	20.00
12	17.74
16	14.00

Visual Representation:

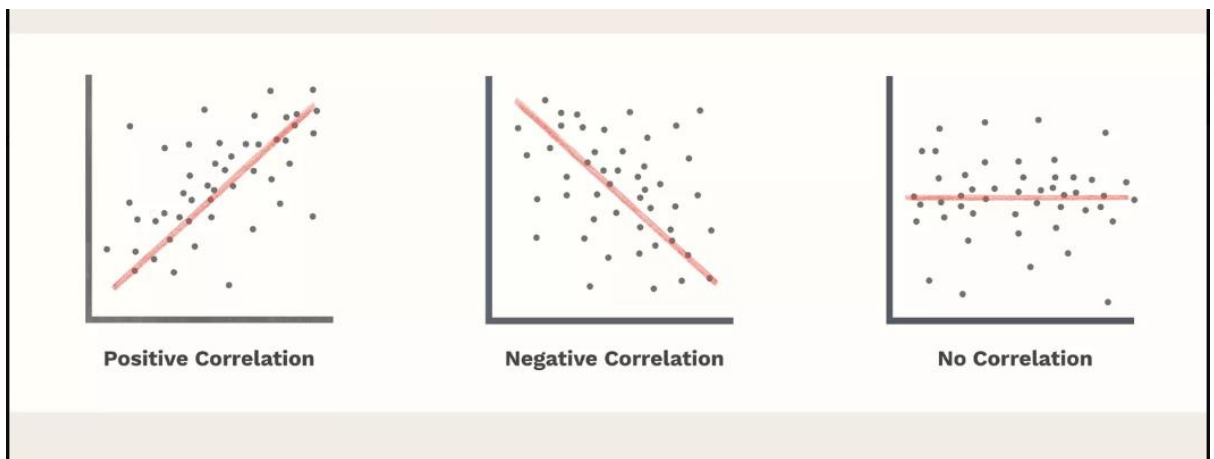


Insight: From the scatter plot, it is evident that as the number of cylinders increases, the fuel efficiency decreases. The trendline indicates downward/negative trend indicating the same.

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

Note: and A correlation coefficient is a number between -1 and 1 that tells us the strength and direction of a relationship between variables, tells us how similar the measurements of two or more variables are across a dataset.

- 1 value indicates a **perfect positive correlation** – all data points align in a straight line
- -1 value indicates a **perfect negative/inverse correlation** - all data points align in a straight line
- 0 value indicates **no linear relationship** or a weak correlation.
- Closer to 0 – weaker correlation
- Closer to 1 or -1 – stronger correlation



Output:

	Engine Cylinders	Average of highway MPG
Engine Cylinders	1	
Average of highway MPG	-0.776843095	1

Insight: The Correlation between Engine Cylinders and Highway MPG is -0.77, which is a negative correlation. This holds good as number of cylinder increases, the amount of fuel needed increases, thus decreasing the highway MPG/Mileage.

Building the Dashboard:

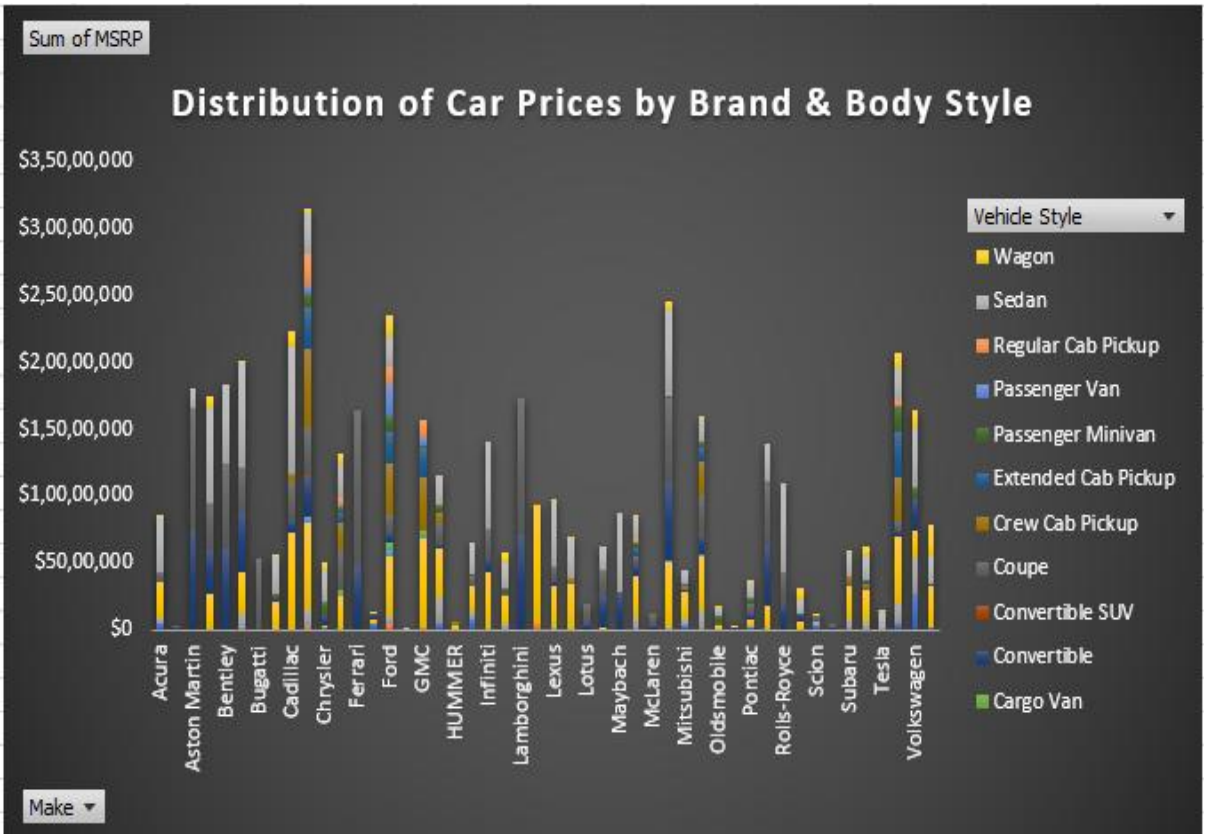
Task 1: How does the distribution of car prices vary by brand and body style?

- 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.

Output:

Sum of MSRP	Column Labels																	Grand Total
Row Labels	2dr SUV	4dr Hatchback	4dr SUV	Cargo Minivan	Cargo Van	Convertible	Convertible SUV	Coupe	Crew Cab Pickup	Extended Cab Pickup	Passenger Minivan	Passenger Van	Regular Cab Pickup	Sedan	Wagon			
Acura	\$4,80,917		\$5,57,440	\$26,63,505			\$1,29,800	\$7,93,748						\$41,34,552	\$2,01,360	\$86,31,522		
Alfa Romeo							\$1,78,200	\$1,78,200								\$3,08,000		
Aston Martin							\$73,21,655	\$92,58,845								\$1,80,29,235		
Audi	\$4,000			\$26,74,900			\$32,91,405	\$35,56,290						\$70,92,748	\$8,47,350	\$1,74,66,693		
Bentley							\$60,12,870	\$63,56,760						\$59,20,900		\$1,82,90,530		
BMW	\$80,097		\$11,03,100	\$31,60,950			\$44,03,171	\$33,04,051						\$78,29,700	\$2,59,600	\$2,01,40,669		
Bugatti							\$52,71,671	\$52,71,671								\$52,71,671		
Buick				\$21,41,770			\$1,79,325	\$18,534			\$3,30,065			\$28,38,590	\$8,212	\$55,16,496		
Cadillac				\$71,82,555			\$9,85,607	\$29,53,574		\$5,99,150				\$94,16,847	\$11,84,100	\$2,23,21,833		
Chevrolet	\$8,000	\$1,93,310	\$12,87,260	\$65,09,468	\$4,20,150	\$74,688	\$29,53,245	\$1,06,300	\$35,04,525	\$59,27,617	\$31,17,951	\$10,47,240	\$5,99,670	\$22,60,032	\$31,77,797	\$3,00,675	\$3,14,87,928	
Chrysler	\$98,805			\$2,50,545			\$6,30,105	\$1,14,510						\$9,22,295	\$24,79,859	\$5,01,075	\$49,97,194	
Dodge	\$38,000	\$12,000	\$16,000	\$24,62,875	\$60,520	\$3,38,497	\$6,000	\$29,73,842	\$20,72,780		\$6,84,682		\$5,57,425	\$70,708	\$6,53,408	\$24,09,585	\$7,93,055	\$1,31,49,377
Ferrari							\$47,23,811	\$1,17,13,289									\$1,64,37,100	
FIAT	\$4,20,715			\$3,69,305			\$5,27,965								\$2,87,570		\$14,05,555	
Ford	\$24,000	\$4,67,873	\$5,67,615	\$44,82,771	\$4,15,630	\$5,56,351	\$7,30,007	\$13,98,144	\$37,82,518	\$22,85,584	\$11,79,285	\$24,29,898		\$12,99,240	\$22,79,348	\$16,23,565	\$2,35,21,829	
Genesis															\$1,39,850		\$1,39,850	
GMC		\$1,28,319		\$66,33,919	\$1,42,750	\$4,60,085			\$40,62,482	\$21,75,866		\$1,50,630	\$5,99,670	\$12,84,328			\$1,56,38,049	
Honda	\$4,13,200		\$19,19,260	\$38,00,589			\$2,52,135		\$15,88,705	\$7,50,215		\$5,53,185		\$22,64,390			\$1,15,41,679	
HUMMER				\$5,77,420					\$2,42,405								\$6,19,895	
Hyundai	\$7,89,650			\$19,94,390					\$6,85,920					\$1,33,075	\$23,23,987		\$64,55,902	
Infiniti				\$43,40,200			\$9,80,050		\$21,75,750					\$64,90,009			\$1,39,86,009	
Kia			\$4,06,960	\$20,49,645			\$1,42,630		\$1,01,77,050					\$4,94,650		\$7,72,405	\$1,72,41,500	
Lamborghini							\$70,64,450										\$94,61,325	
Land Rover	\$4,76,394			\$88,39,200				\$1,45,751	\$10,16,472						\$48,37,596	\$31,105	\$96,04,912	
Lexus			\$94,700	\$31,52,974			\$4,72,065								\$28,54,855	\$2,69,705	\$70,17,732	
Lincoln				\$34,22,570					\$17,342	\$4,53,260							\$19,14,560	
Lotus				\$1,55,000			\$4,13,260		\$15,01,300						\$17,82,400		\$62,52,847	
Maserati							\$27,62,750								\$59,76,800		\$87,39,550	
Maybach							\$28,42,983		\$19,72,284								\$11,99,025	
Mazda	\$18,000	\$12,000	\$8,53,180	\$31,75,515			\$8,70,505		\$5,41,879		\$5,80,033	\$4,43,130		\$2,65,486	\$16,18,571	\$33,350	\$84,11,649	
McLaren							\$2,80,225		\$9,18,800								\$2,45,75,709	
Mercedes-Benz			\$1,22,800	\$49,74,610	\$28,950		\$57,53,964	\$64,73,107							\$65,43,743	\$6,46,035	\$44,38,837	
Mitsubishi	\$3,70,169		\$4,03,835	\$20,09,807	\$2,000		\$2,09,899		\$2,40,210	\$1,34,360		\$2,000		\$8,000	\$10,58,563	\$8,000	\$1,59,35,555	
Nissan	\$14,683		\$13,47,320	\$41,49,630	\$1,28,620		\$14,06,552	\$1,31,075	\$29,37,632	\$24,22,300	\$10,26,379		\$4,13,320	\$19,914	\$17,63,130	\$1,75,000	\$1,59,35,555	
Oldsmobile				\$2,38,150			\$2,000		\$2,76,015					\$4,92,055	\$6,67,161	\$20,000	\$16,95,381	
Plymouth	\$40,000		\$14,000				\$85,631		\$8,000					\$38,759	\$16,000		\$2,34,078	
Pontiac	\$1,63,505		\$1,62,975	\$4,01,550			\$4,73,481		\$6,63,715					\$11,56,535	\$27,13,500	\$20,855	\$35,83,808	
Porsche	\$28,827			\$18,15,200			\$45,04,586		\$47,58,533					\$5,41,192			\$1,38,20,646	
Rolls-Royce							\$21,41,365		\$22,04,675						\$65,39,010		\$1,08,85,050	

Visual Representation:



Insight: Above is the stacked column chart that shows distribution of total car prices (MSRP) by brand(make) and body style (Vehicle Style)

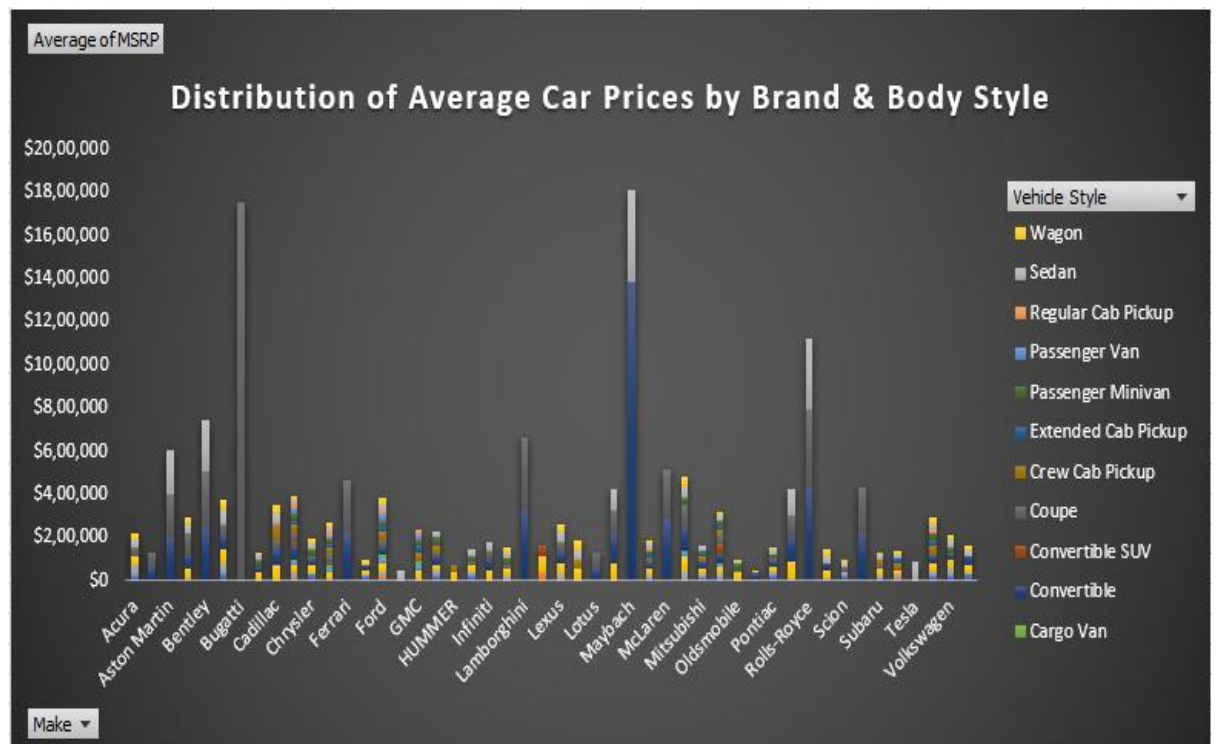
Task 2: Which car brands have the highest and lowest average MSRPs, and how does this vary by body style?

- **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.

Output:

Row Labels	2dr Hatchback	2dr SUV	4dr Hatchback	4dr SUV	Cargo Minivan	Cargo Van	Convertible	Convertible SUV	Coupe	Crew Cab Pickup	Extended Cab Pickup	Passenger Minivan	Passenger Van	Regular Cab Pickup	Sedan	Wagon	Grand Total
Acura	\$17,176		\$51,063	\$42,960					\$39,687						\$33,614	\$33,560	\$35,087
Alfa Romeo							\$64,900		\$59,400								\$61,600
Aston Martin							\$2,03,379		\$1,92,893						\$2,06,962		\$1,98,123
Audi	\$2,000			\$48,635			\$70,030		\$93,587						\$46,358	\$33,894	\$54,583
Bentley							\$2,50,536		\$2,54,270						\$2,36,836		\$2,47,169
BMW	\$26,699		\$55,155	\$58,536			\$63,814		\$52,445						\$71,832	\$43,267	\$62,163
Bugatti									\$17,57,224								\$17,57,224
Buick				\$33,996			\$25,618		\$2,059				\$30,006		\$29,569	\$2,053	\$29,034
Cadillac				\$72,551			\$70,401		\$45,440		\$66,572				\$51,179	\$47,364	\$56,368
Chevrolet	\$2,000	\$13,808	\$18,930	\$35,554	\$20,007	\$8,299	\$62,835	\$17,717	\$38,939	\$39,256		\$24,170	\$24,994	\$28,556	\$19,825	\$20,635	\$15,825
Chrysler	\$32,935			\$35,792			\$24,235		\$19,085				\$29,751		\$26,104	\$26,372	\$26,723
Dodge	\$2,000	\$2,000	\$2,000	\$31,176	\$20,173	\$12,537	\$2,000		\$45,058	\$31,406	\$16,302		\$25,338	\$14,142	\$14,850	\$22,519	\$24,783
Ferrari							\$2,14,719		\$2,49,219								\$2,38,219
FIAT	\$21,036			\$24,620			\$23,426								\$22,121		\$22,670
Ford	\$2,000	\$16,134	\$19,573	\$41,507	\$19,792	\$20,606	\$94,762		\$34,101	\$41,566	\$23,808	\$23,123	\$32,836	\$17,798	\$23,259	\$30,066	\$38,511
Genesis															\$46,617		\$46,617
GMC		\$7,129		\$37,480	\$23,792	\$21,909									\$25,183		\$32,444
Honda	\$17,217		\$26,656	\$28,576			\$36,019		\$21,763		\$39,062	\$27,896	\$25,105	\$28,556	\$25,183	\$26,027	\$36,655
HUMMER				\$37,749							\$34,629		\$36,879				\$36,464
Hyundai	\$18,364		\$17,629	\$30,218					\$22,126				\$26,615				\$24,926
Infiniti				\$45,686			\$46,669		\$40,292						\$41,076		\$42,640
Kia			\$19,379	\$31,533					\$20,376						\$23,812	\$20,326	\$25,514
Lamborghini							\$3,36,402		\$3,28,292								\$3,31,567
Land Rover	\$39,700			\$71,284				\$48,577									\$68,067
Lexus			\$31,567	\$45,042			\$52,452		\$50,824						\$48,865	\$31,105	\$47,549
Lincoln				\$50,332					\$2,168		\$41,205				\$42,610	\$44,951	\$43,861
Lotus							\$51,658		\$75,065								\$68,377
Maserati				\$77,500			\$1,30,165		\$1,16,017						\$99,022		\$1,13,684
Maybach							\$13,81,375								\$4,26,914		\$5,46,222
Mazda	\$2,000	\$2,000	\$20,809	\$27,141			\$28,081		\$20,842					\$9,155	\$19,739	\$16,675	\$20,417
McLaren							\$2,80,225		\$2,29,700								\$2,39,805
Mercedes-Benz			\$40,933	\$68,145	\$28,950		\$1,04,618		\$1,09,714				\$32,500		\$48,834	\$43,069	\$72,070
Mitsubishi	\$12,764		\$13,925	\$26,101	\$2,000		\$29,985			\$26,690	\$19,194		\$2,000		\$2,000		\$21,341
Nissan	\$2,098		\$24,059	\$34,294	\$21,437		\$39,071	\$43,692	\$35,393	\$32,734	\$20,528		\$22,962		\$2,213	\$21,604	\$17,500
Oldsmobile			\$34,021				\$2,000		\$10,616				\$32,804		\$9,139	\$2,000	\$12,844
Plymouth	\$2,000		\$2,000				\$28,544		\$2,000				\$2,113		\$2,769	\$2,000	\$3,297
Pontiac	\$18,167		\$18,108	\$25,097			\$22,547		\$16,188				\$20,815		\$20,652	\$6,952	\$19,800
Porsche	\$5,765		\$82,509				\$1,15,502		\$99,136						\$1,23,341		\$1,01,622

Visual Representation:



Insight: Above is the stacked column chart that shows distribution of average car prices (MSRP) by brand(make) and body style (Vehicle Style)

Highest Average MSRP	\$17,57,224	Bugatti
Lowest Average MSRP	\$3,297	Plymouth

The average MSRP varies based on different body styles for all brands.

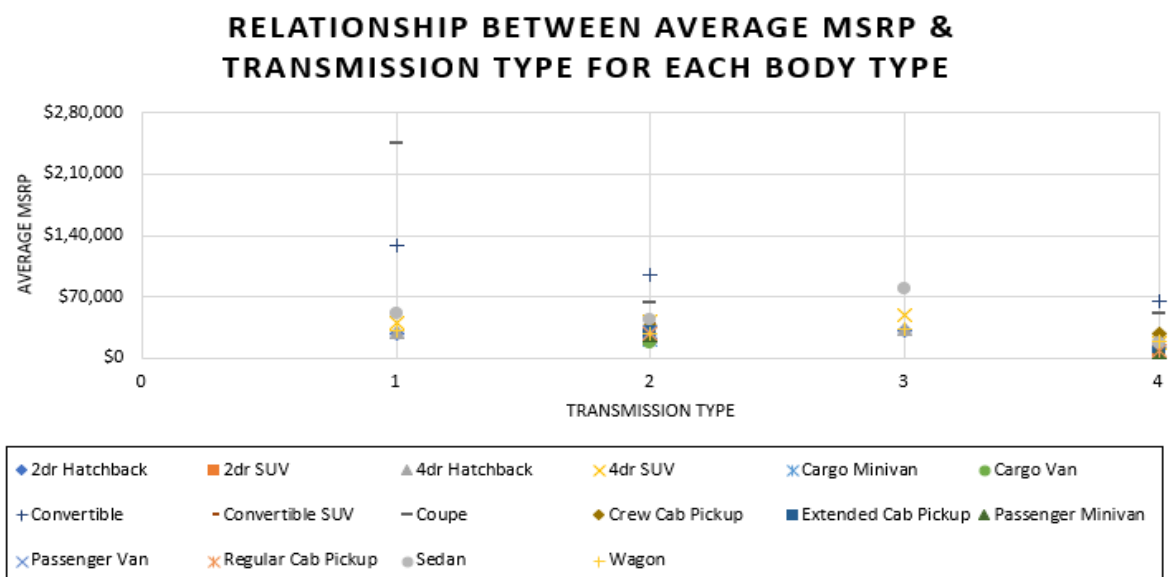
Task 3: How do the different feature such as transmission type affect the MSRP, and how does this vary by body style?

- 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.

Output:

Symbols for Transmission Type	2dr Hatch	2dr SUV	4dr Hatch	4dr SUV	Cargo Mi	Cargo Va	Convertible	Convertibl	Coupe	Crew Cab	Extended	Passenge	Passenge	Regular C	Sedan	Wagon
1 AUTOMATED_MAN	\$27,470		\$29,347	\$40,451			\$1,29,082		\$2,45,977						\$50,379	\$31,985
2 AUTOMATIC	\$20,523	\$21,799	\$23,889	\$41,638	\$20,316	\$17,019	\$94,586	\$38,926	\$64,270	\$37,719	\$30,711	\$26,590	\$30,578	\$28,371	\$44,651	\$28,219
3 DIRECT_DRIVE	\$31,800		\$32,800	\$49,800											\$79,512	\$34,250
4 MANUAL	\$12,841	\$9,173	\$17,500	\$17,422			\$64,794	\$9,595	\$50,901	\$28,233	\$11,553	\$6,510		\$8,759	\$17,557	\$18,399

Visual Representation:



Insight: Above is the scatter plot that shows distribution of average car prices (MSRP) by Transmission Type with different symbols for each body style.

The average MSRP varies depending on different body styles and Transmission Types.

	Value	Transmission type	Body Style
Highest Average MSRP	\$2,45,977	AUTOMATED_MANUAL	Coupe
Lowest Average MSRP	\$6,510	MANUAL	Passenger Minivan

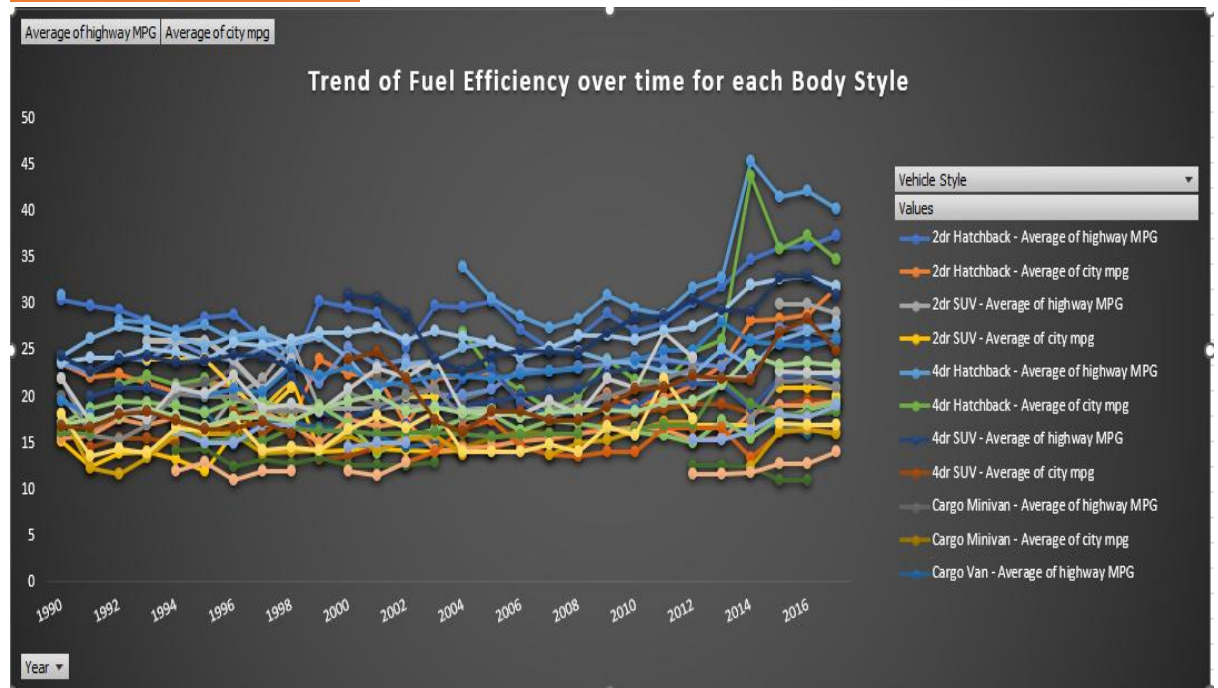
Task 4: How does the fuel efficiency of cars vary across different body styles and model years?

- **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.

Output:

	Col ▾		2dr Hatchback		2dr SUV		4dr Hatchback		4dr SUV		Cargo Minivan		Cargo Van		Convertible		Convertible St Coupe		Crew Cab Pick		Extended Cab		Passenger Mir		Passenger Van		Regular Cab		Pi Sedan		Wagon		Total A Total A		
Row ▾	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average	Average of city mpg	
1990	30	24	20	15	31	22					20	18			24	16			24	17			22	18	20	16		22	18	24	17	24	17	24	17
1991	30	22	16	13			20	15							23	16			26	19			16	12	18	16		17	14	24	18	23	17	23	17
1992	29	22	18	14	28	21	21	16							24	18			28	19			15	12				18	14	24	18	24	18	25	18
1993	28	21	19	14	28	22	21	16							24	17	26	24	27	19			17	13				18	14	25	19	24	18	25	18
1994	27	20	18	13	27	21	20	15	21	17	19	14	26	18	26	24	26	19			20	16	21	17	16	12	21	17	25	18	24	17	24	18	
1995	29	22	16	12	28	22			22	17	18	14	25	17	26	24	26	18			20	16	20	15	15	13	20	16	24	17	24	16	23	17	
1996	29	21	20	16	26	19	21	19	23	17	15	12	24	17	24	21	27	20			20	16	21	15	15	11	22	18	26	18	25	17	24	18	
1997	26	20	22	19	27	19	20	16	21	15	17	13	25	18	21	18	27	19			18	14	21	15	17	12	19	14	25	18	24	17	22	16	
1998	23	17	26	22	25	18	22	18			17	13	24	17	24	21	26	19			19	14	23	17	17	12	19	15	26	17	23	16	21	16	
1999	30	24	19	14			18	13			17	13	22	15			27	19			18	14	22	16		18	14	27	19			23	17		
2000	30	22	19	14			18	14			16	13	25	18			24	17			21	16	23	16	15	12	21	17	27	20	31	24	24	18	
2001	29	22	19	14			19	14	22	16	16	13	23	17			20	14			19	14	21	16	15	12	23	18	27	20	31	25	24	18	
2002	25	17	19	14			20	16	21	16	15	13	24	17	23	20	24	16	17	13	20	15	22	16	15	13	22	17	26	18	29	22	23	17	
2003	30	22	19	14			19	15	21	15	15	13	20	14	23	20	24	16	18	14	21	16	22	16		24	18	27	19	24	18	23	17		
2004	30	22	19	14	34	27	19	15	20	15			20	14			25	18	22	17	18	14	22	16		18	14	26	18	23	16	23	17		
2005	30	23	19	14	31	23	19	14	21	15			21	15			26	18	23	17			22	16		18	14	26	19	24	18	24	17		
2006	27	20			29	21	20	16	23	16			23	15			24	16	19	15			22	16		18	14	25	18	25	18	24	17		
2007	25	18			27	19	20	15	23	16			23	16			25	18	18	14	18	14	23	16		20	15	25	18	25	18	22	16		
2008	26	19			28	20	21	16	23	16			23	16			25	17	18	14	19	15	23	16		18	14	27	19	25	17	23	16		
2009	29	20			31	24	23	17					24	16			24	17	19	14	20	15				22	17	27	19	27	19	24	17		
2010	27	19			30	22	23	18					24	17			24	16	19	14	21	16	24	16		21	16	26	18	28	21	24	18		
2011	28	20			29	21	24	19					24	17			23	16	21	17	22	17	25	17		27	22	27	19	29	21	25	19		
2012	30	21			32	25	24	19			17	13	24	15	22	17	22	15	21	17	23	17	25	17	15	12	24	18	28	20	30	22	26	19	
2013	32	23			33	26	24	19			17	13	23	16	22	17	25	17	21	17			28	22	15	12		28	21	29	22	27	20		
2014	35	28			45	44	24	18			17	12	26	18	22	17	23	15	19	13	17	12	26	19	16	12		32	25	29	22	28	21		
2015	36	28	30	21	42	36	26	19	28	23	17	11	27	19			26	18	22	16	22	16	26	18	18	13	23	17	33	23	33	27	29	21	
2016	36	29	30	21	42	37	26	20	27	22	16	11	28	19			27	19	22	16	22	16	26	18	18	13	23	17	33	24	33	28	29	22	
2017	37	32	29	21	40	35	26	19	27	22			28	20	28	20	28	19	22	16	21	16	26	18	19	14	23	17	32	23	31	25	28	21	
Grand	31	24	20	15	38	32	25	18	24	18	17	13	25	18	24	20	26	18	21	16	20	15	24	17	17	13	21	16	30	22	28	22	27	20	

Visual Representation:



Insight: Above Pivot table calculates the average MPG(Highway and City) for each combination of body style and model year.

The above line chart shows the trend of fuel efficiency (MPG) over time for each body style.

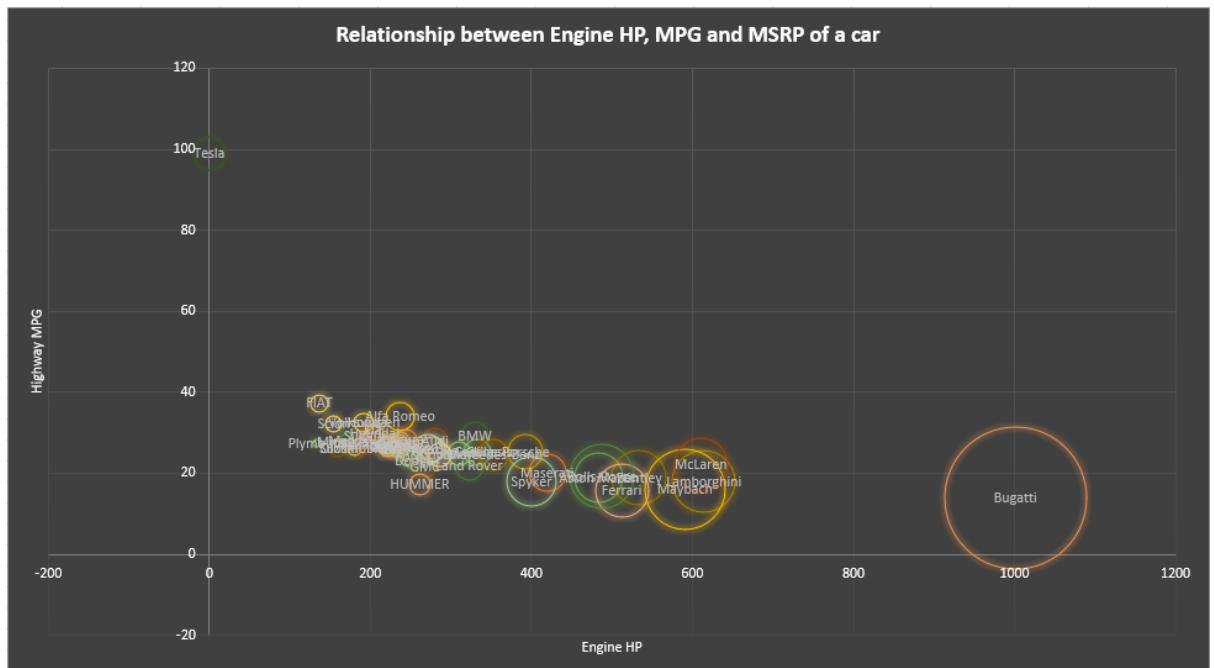
Task 5: How does the car's horsepower, MPG, and price vary across different Brands?

- **Hints:** Bubble chart to visualize the relationship between horsepower, MPG, and price across different car brands. Assign different colours 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.

Output:

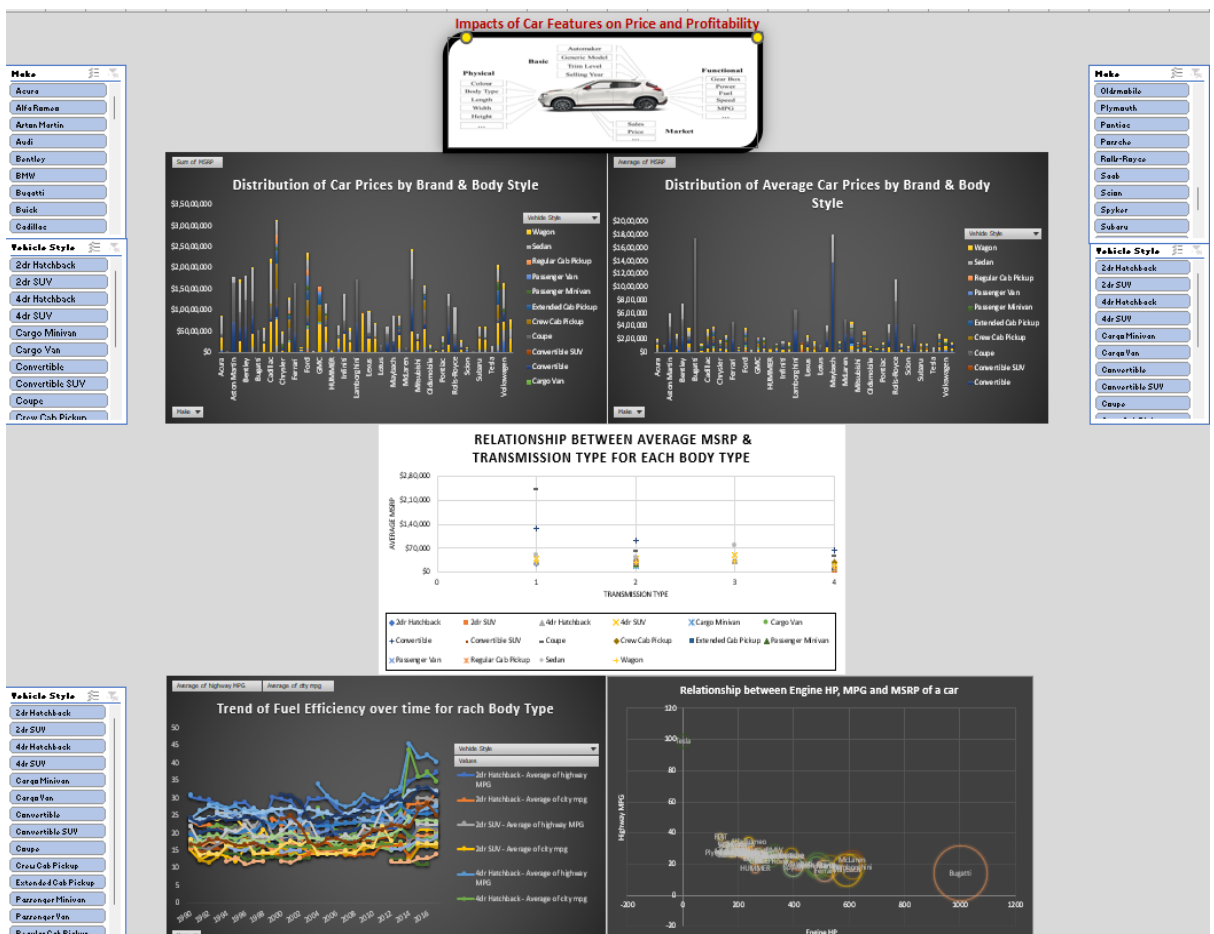
Row Labels	Average of Engine HP	Average of highway MPG	Average of MSRP
Acura	245	28	\$35,087
Alfa Romeo	237	34	\$61,600
Aston Martin	484	19	\$1,98,123
Audi	280	28	\$54,583
Bentley	534	19	\$2,47,169
BMW	330	29	\$62,163
Bugatti	1001	14	\$17,57,224
Buick	220	27	\$29,034
Cadillac	333	25	\$56,368
Chevrolet	249	26	\$29,075
Chrysler	229	26	\$26,723
Dodge	254	23	\$24,857
Ferrari	512	16	\$2,38,219
FIAT	137	37	\$22,670
Ford	248	24	\$28,511
Genesis	347	25	\$46,617
GMC	268	21	\$32,444
Honda	196	32	\$26,655
HUMMER	261	17	\$36,464
Hyundai	205	30	\$24,926
Infiniti	311	25	\$42,640
Kia	203	31	\$25,514
Lamborghini	614	18	\$3,31,567
Land Rover	323	22	\$68,067
Lexus	277	26	\$47,549
Lincoln	286	25	\$43,861
Lotus	272	26	\$68,377
Maserati	420	20	\$1,13,684
Maybach	591	16	\$5,46,222
Mazda	173	28	\$20,417

Visual Representation:



Insight: The above Bubble chart visualizes the relationship between horsepower (Engine HP), MPG, and price across different car brands.

Final Dashboard:



Insights:

a) Key Insights:

- Higher the Engine power, Higher the price.
- The **strongest impact on car's price is dependent on Engine Cylinders** followed by highway MPG.
- The majority of the car buyers prefer more budget-friendly options. Luxury and exotic cars being least popular as they are most expensive.
- With increase in Engine Cylinders, fuel efficiency decreases.

b) Relevance to Business Problem:

These insights directly address the business problem by providing actionable information for car manufacturers to optimize pricing strategies, identify profitable market segments, and prioritizes product development efforts.

c) Recommendations:

Based on the insights gained, we recommend that car manufacturers can focus on developing fuel-efficient models, strategically price cars based on feature importance and create marketing strategies to target popular market categories.

Result:

- a) Visualization: We used visualizations such as pivot tables, scatter plots, and bar charts to display the results obtained from the analysis. These visualizations will help stakeholders to understand the findings more effectively and facilitate decision making.
- b) Discussion: The insights obtained from the analysis have significant implications for car manufacturers, providing valuable insights into consumer preferences, pricing dynamics, and market trends and patterns. These findings can be used to create strategic decisions aimed at enhancing profitability and competitiveness.
- c) Limitations or uncertainties regarding the results: The insights obtained is based on the dataset which was last updated in 2017, so it may not reflect current trends or prices in the automotive industry.
- d) Future Directions: Additional analysis could explore dynamic pricing strategies, incorporate real-time market data, and leverage **advanced machine learning techniques for predictive modeling**. Furthermore, ongoing monitoring of market trends and consumer preferences are essential for staying competitive in the automotive industry.

Overall, the insights gained from this analysis provide valuable guidance for car manufacturers seeking to optimize pricing and product development strategies in a rapidly changing market landscape.

Links: My excel worksheet link with different sheets for each task,

https://docs.google.com/spreadsheets/d/1Lg2cWRQWa1Nyxu2qE5jaEWMY_54HhSQ7/edit?usp=sharing&oid=108154584635151678812&rtpof=true&sd=true

You can connect with me on LinkedIn account,

<https://www.linkedin.com/in/raksha-nayak-41578738/>

Loom Video Link,

<https://go.screenpal.com/watch/cZX3iLnVz5D>

Tableau Dashboard Link and Screenshot,

https://public.tableau.com/views/Project7-ImpactofCarFeaturesonPriceProfitability/Dashboard1?:language=en-US&publish=yes&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link

