

The Impact of Car Features on Price and Profitability

Project Description

This Project involves analysis of various car features of different brands to find their impact on price and profitability and give insights to manufacturer on how to optimize pricing, what decisions to make in product development to maximize profitability while meeting consumer demand.

The automotive industry has been rapidly evolving over the past few decades, with a growing focus on fuel efficiency, environmental sustainability, and technological innovation. With increasing competition among manufacturers and a changing consumer landscape, it has become more important than ever to understand the factors that drive consumer demand for cars.

In recent years, there has been a growing trend towards electric and hybrid vehicles and increased interest in alternative fuel sources such as hydrogen and natural gas. At the same time, traditional gasoline-powered cars remain dominant in the market, with varying fuel types and grades available to consumers.

As a data analyst, first I will understand the project requirements and purpose and then I will understand the data which is provided in the project attachment and will perform analysis to get the meaningful insights.

Analysis: involves understanding the relationships between different variables that describe car features, use of and statistical analysis, regression and segmented analysis to find the relation between car features, market categories and pricing which can help in decision making on pricing and product development.

For example, identifying which features and categories are most popular among consumers and most profitable for the manufacturer. This could help the manufacturer understand the customer preferences and willingness to pay for different types of cars, and adjust the pricing and product development strategies accordingly.

Approach

1. Data Cleaning: Checking for missing or inconsistent data. Handle outliers if necessary. This step ensures that the data is accurate and will not lead to misleading analysis results.

In the given Dataset below variables/columns as missing data (blank cells):

Column Name	Blank cell count
Engine Fuel Type	3
Engine HP	69
Engine Cylinders	30
Number of Doors	6
Market Category	3742

2. Handling Missing Data: Replace the missing data with the help appropriate descriptive statistics calculations. Below are the methods used to replace the missing data in the dataset of this project.

Column Name	Blank cell count	Blank cells replace method	Replaced value
Engine Fuel Type	3	most repeated value	premium unleaded (required)
Engine HP	69	most repeated value	200
Engine Cylinders	30	most repeated value	4
Number of Doors	6	most repeated value	4
Market Category	3742	most repeated value	Crossover

3. Descriptive Statistics: Generating basic statistics such as mean, min, max. This will give a general understanding of the dataset's distribution.
4. Correlation Analysis: Identifying relationships between different variables in the dataset. This can help in understanding how different attributes relate to the likelihood of default.
5. Visual Exploratory Data Analysis: Using plots and charts to visualize the data and identify relations. This could include:
 - o Stacked Column charts for categorical variables like Transmission types, Body style and Brand.
 - o Scatter and Bubble for numerical variables likes Price, Engine power and Fuel efficiency to understand their distribution and relationship.
 - o Line plots to understand the trends over time.

Throughout this process, my goal was to provide clear, accurate, and helpful insights and answers to questions.

Tech-Stack Used

The software and versions you used for the project:

Microsoft Excel
Pivot Tables to classify the data and to calculate Mean, etc.
Power Query to filter and sort the data
Data analysis tool pack for descriptive statistics
Tableau to create dashboards

Insights:

I will provide my insights in the form of answers for the questions posed by management team

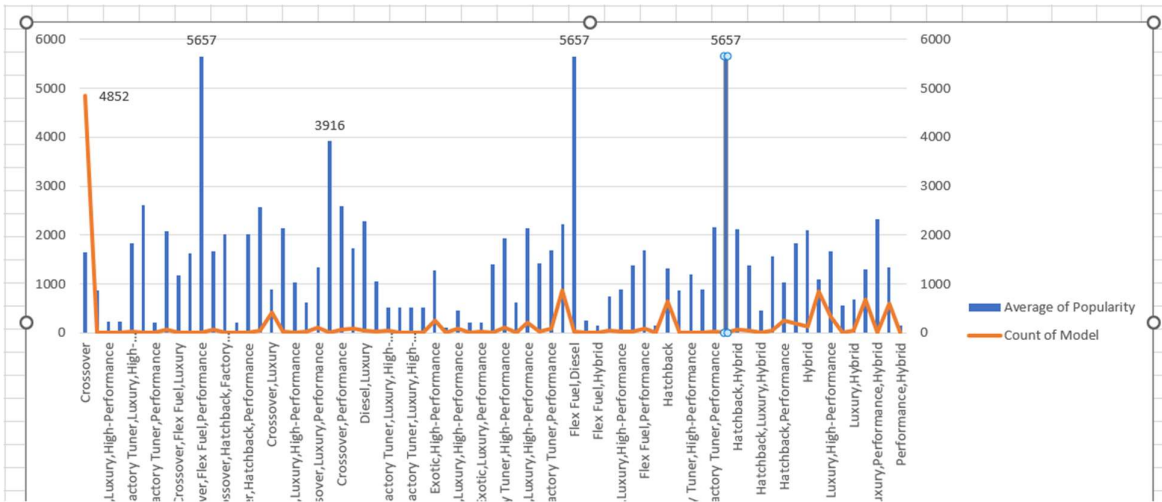
A. How does the popularity of a car model vary across different market categories?

Output:

Pivot Table:

Market Category	Average of Popularity	Count of Model
Flex Fuel,Diesel	5657	16
Hatchback,Flex Fuel	5657	7
Crossover,Flex Fuel,Performance	5657	6
Crossover,Luxury,Performance,Hybrid	3916	2
Crossover,Factory Tuner,Luxury,Performance	2607.4	5
Crossover,Performance	2585.956522	69
Crossover,Hybrid	2563.380952	42
Luxury,Performance,Hybrid	2333.181818	11
Diesel,Luxury	2275	51
Flex Fuel	2217.302752	872
Hatchback,Factory Tuner,Performance	2159.045455	22
Crossover,Luxury,Diesel	2149.411765	34
Factory Tuner,Luxury,High-Performance	2133.367442	215
Hatchback,Hybrid	2121.25	72
Hybrid	2105.569106	123
Crossover,Flex Fuel	2073.75	64

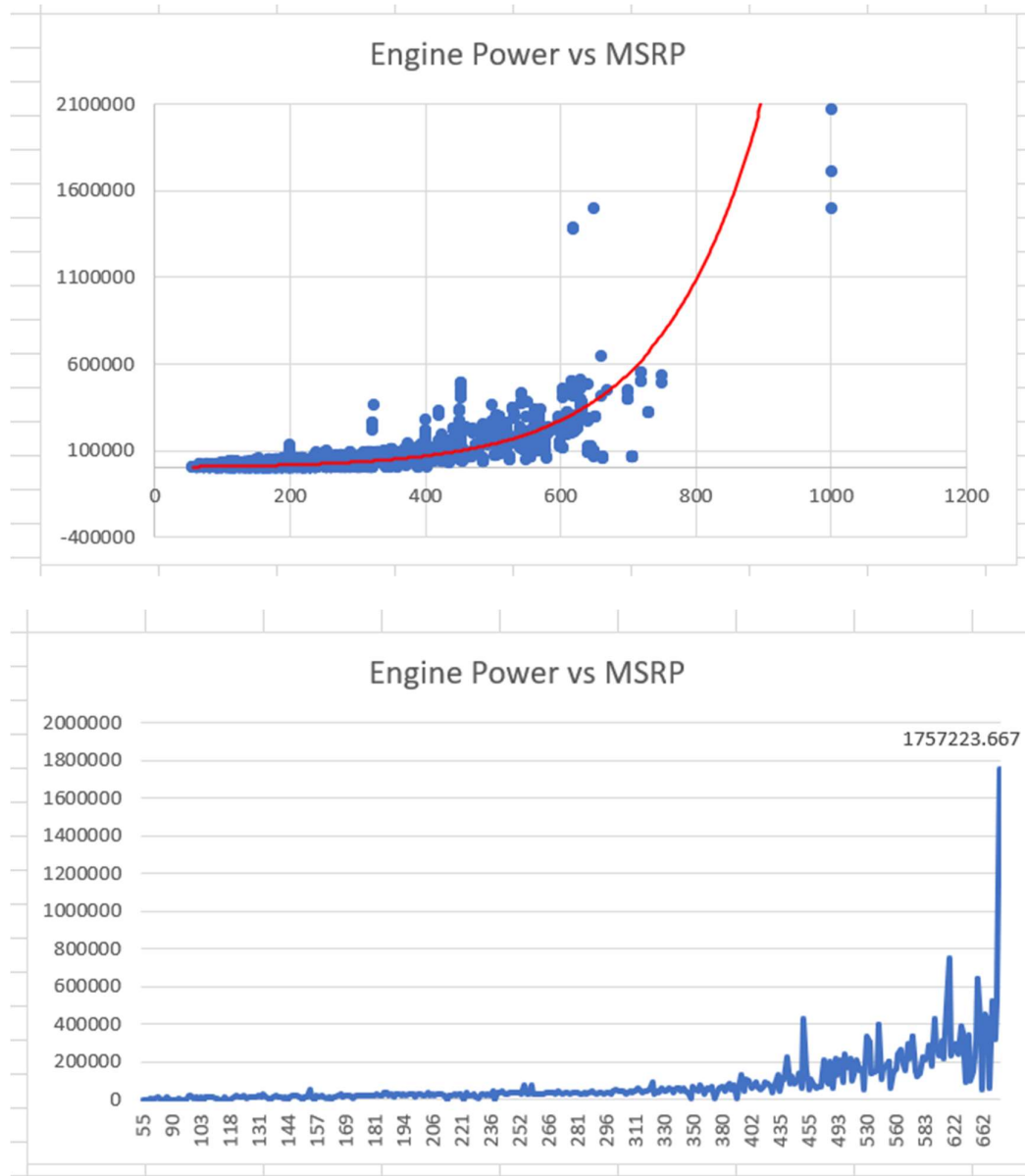
Graph:



Above Pivot table and Graph shows that “Flex Fuel” is the most popular market category while the “Crossover” market segment has the more number of models available. Please see “Question 1” sheet in Project_7 dataset excel for complete graph and details

B. What is the relationship between a car's engine power and its price?

OUTPUT:

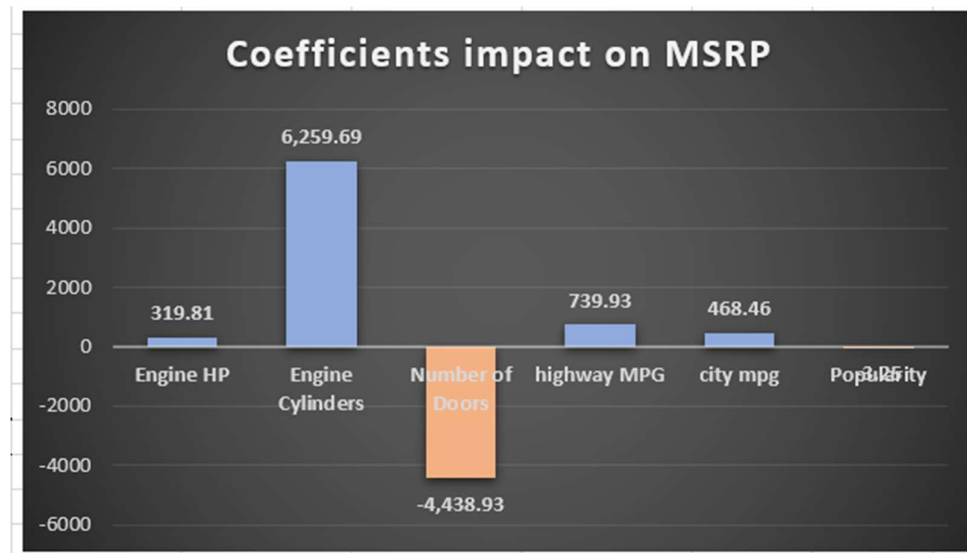


Both Scatter and Line trend graph shows strong positive relationship between Engine HP and Car Price. Higher the engine HP, higher is the price of car becomes. For more details, please see "Question 2" sheet in Project_7 dataset excel.

C. Which car features are most important in determining a car's price?

OUTPUT:

Regression Analysis:



Correlation Analysis:

	MSRP	Engine HP	Engine Cylinders	Number of Doors	highway MPG	city mpg	Popularity
MSRP	1						
Engine HP	0.660896077	1					
Engine Cylinder	0.531170755	0.770187557	1				
Number of Doo	-0.126081408	-0.103662312	-0.141053659	1			
highway MPG	-0.160042679	-0.366710381	-0.6074225	0.118924554	1		
city mpg	-0.157675722	-0.362922348	-0.574381831	0.121206426	0.886829496	1	
Popularity	-0.048476232	0.036217414	0.042430719	-0.048250854	-0.020991024	-0.00322	1

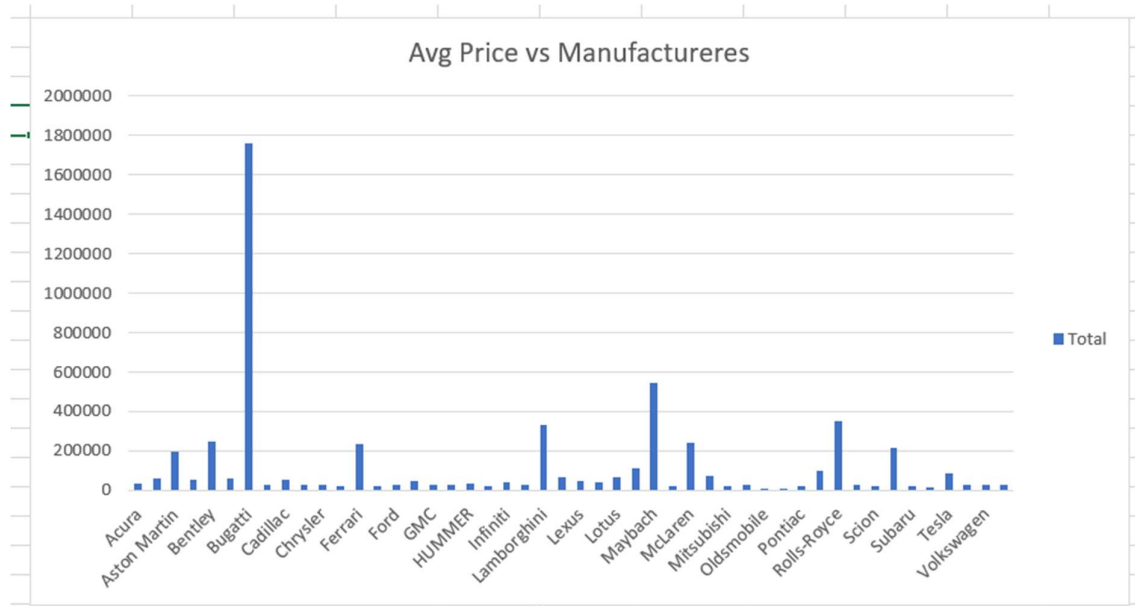
Above first graph is result of Regression analysis and it shows that Engine Cylinders feature has the highest importance in price determination.

But the Correlation Analysis results shows that Engine HP has the strongest positive relationship with the Price (MSRP), which is matching with the results of Task 2.

Please see "Question 3" sheet in Project_7 dataset excel file.

D. How does the average price of a car vary across different manufacturers?

OUTPUT:

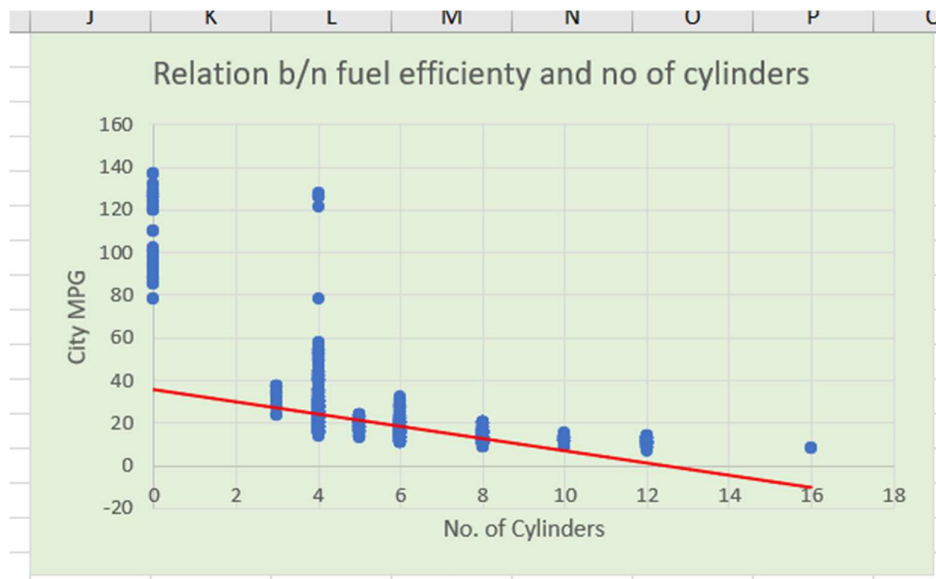
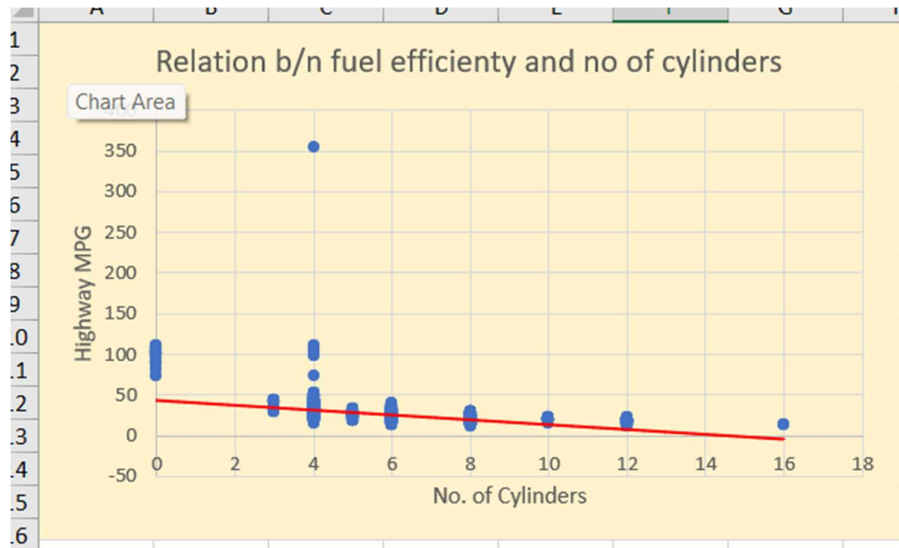


"Bugatti", "Maybach", "Rolls Royce," and "Lamborghini" are the makers with the highest average automobile pricing.

Among the makers, "Bugatti" has the greatest average price range and "Plymouth" the lowest average price range.

E. What is the relationship between fuel efficiency and the number of cylinders in a car's engine?

OUTPUT:



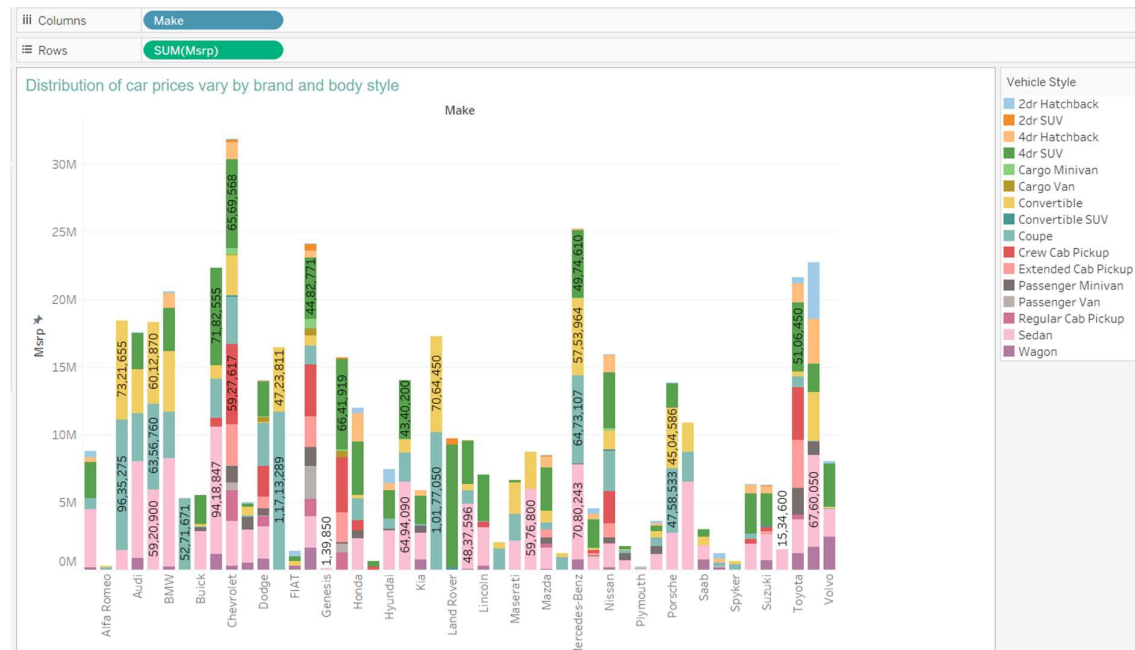
Both graphs of Highway fuel efficiency and City fuel efficiency against no of cylinders show negative relationship between fuel efficiency and no. of cylinders.

Vehicles with fewer engine cylinders often have higher average Highway and City MPG, whereas vehicles with more engine cylinders typically have lower average highway and City MPG.

DASHBOARD TASKS

A. How does the distribution of car prices vary by brand and body style?

OUTPUT:

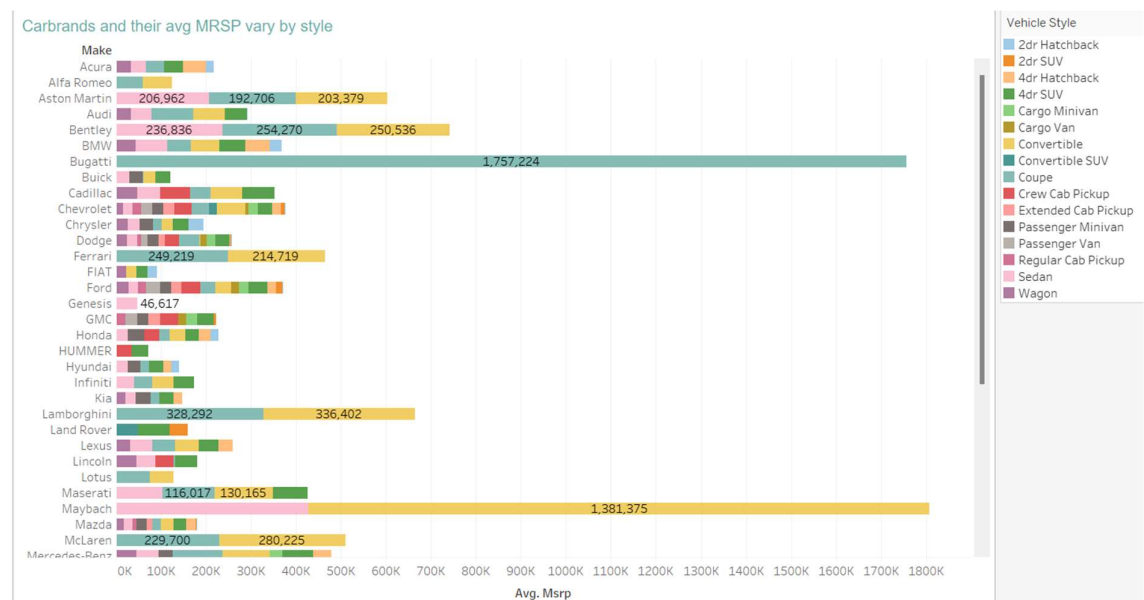


Make	
Chevrolet	3,18,37,483
Mercedes-Benz	2,52,31,109
Ford	2,41,38,754
Volkswagen	2,27,34,826
Cadillac	2,23,23,833
Toyota	2,16,56,392
BMW	2,05,56,619
Aston Martin	1,84,05,665
Bentley	1,82,90,530
Audi	1,75,32,293
Lamborghini	1,72,41,500
Ferrari	1,64,37,100
Nissan	1,59,49,555
GMC	1,57,04,049
Dodge	1,40,16,177
Infiniti	1,39,90,090
Porsche	1,38,20,646
Honda	1,19,76,779
Rolls-Royce	1,08,85,050
Land Rover	96,98,720
Lexus	96,04,912

Vehicle Style	
Sedan	11,96,97,062
4dr SUV	10,05,69,617
Coupe	9,23,36,718
Convertible	6,67,89,858
Crew Cab Pickup	2,53,47,138
4dr Hatchback	1,57,39,448
Wagon	1,51,30,300
Extended Cab Pickup	1,40,10,508
Passenger Minivan	1,06,83,978
2dr Hatchback	85,35,063
Regular Cab Pickup	62,53,854
Passenger Van	37,13,946
Cargo Minivan	14,85,390
Cargo Van	14,51,621
2dr SUV	13,95,896
Convertible SUV	5,05,300

- For instance, "Chevrolet" brand highest total MSRP as it provides more variety of body types.
- . and Sedan body style has highest total MSRP.

OUTPUT:



Make	MSRP
Bugatti	1,757,224
Maybach	546,222
Rolls-Royce	351,131
Lamborghini	331,567
Bentley	247,169
McLaren	239,805
Ferrari	238,219
Spyker	213,323
Aston Martin	197,910
Maserati	114,208
Porsche	101,622
Tesla	85,256
Mercedes-Benz	71,476
Lotus	69,188
Land Rover	67,823
Alfa Romeo	61,600
BMW	61,547
Cadillac	56,221

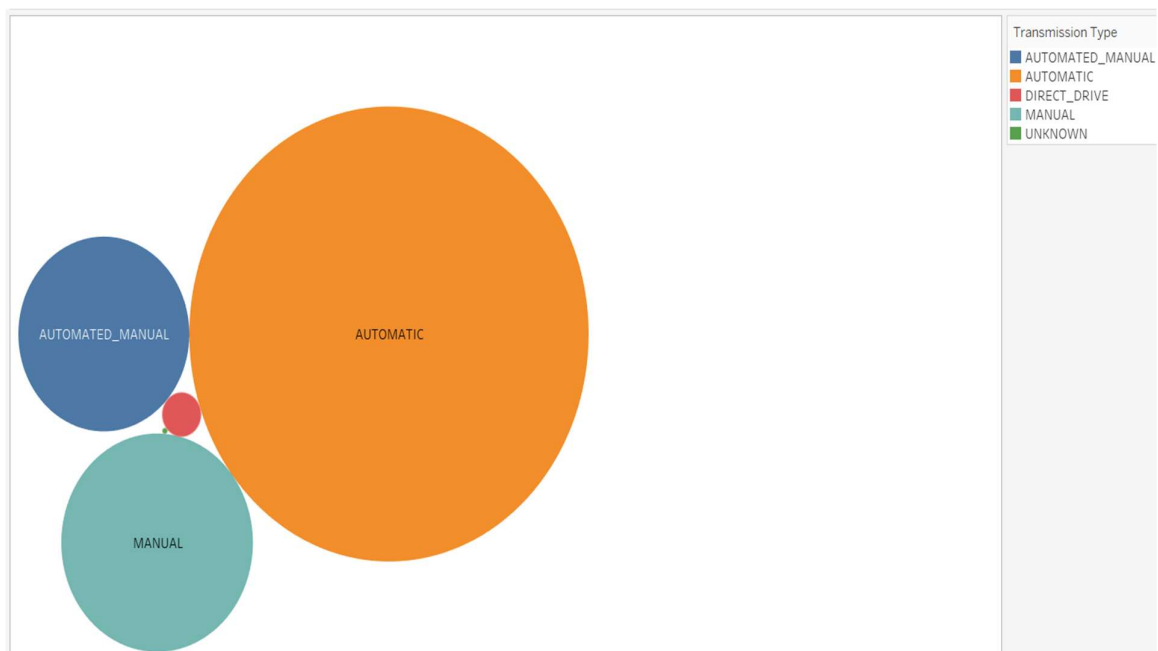
Make	MSRP
Bugatti	1,757,224
Plymouth	3,123

As per above graph and pivot table results Bugatti has the highest MSRP average and Plymouth with lowest MSRP average but with limited styles.

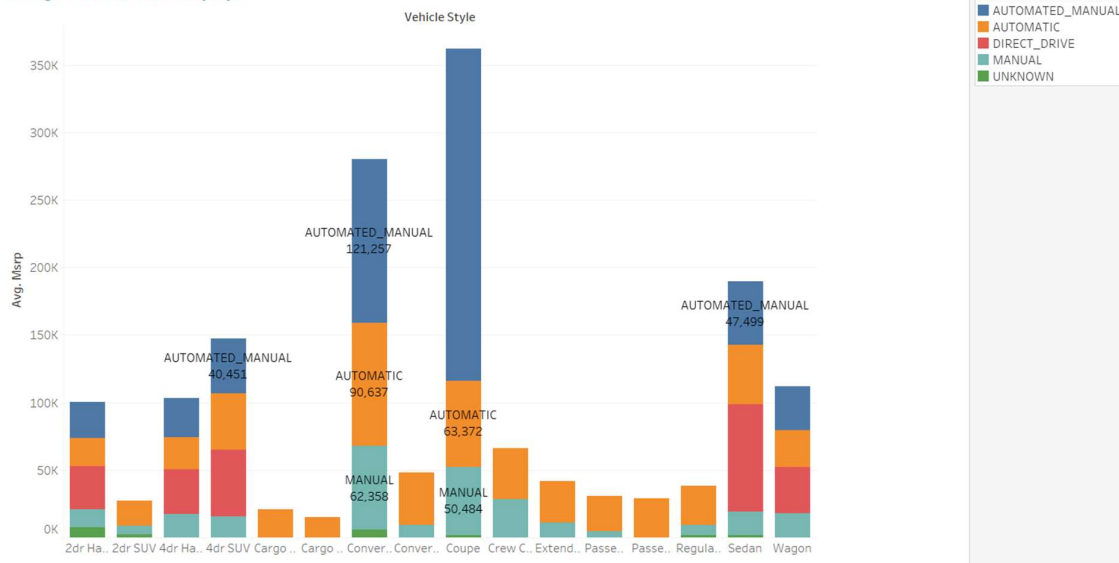
In the brands with more styles "Chevrolet" has good MSRP average.

D. How does the Transmission Type affect the MSRP and how does this vary by Body Style:

OUTPUT:



Average MSRP for each body style

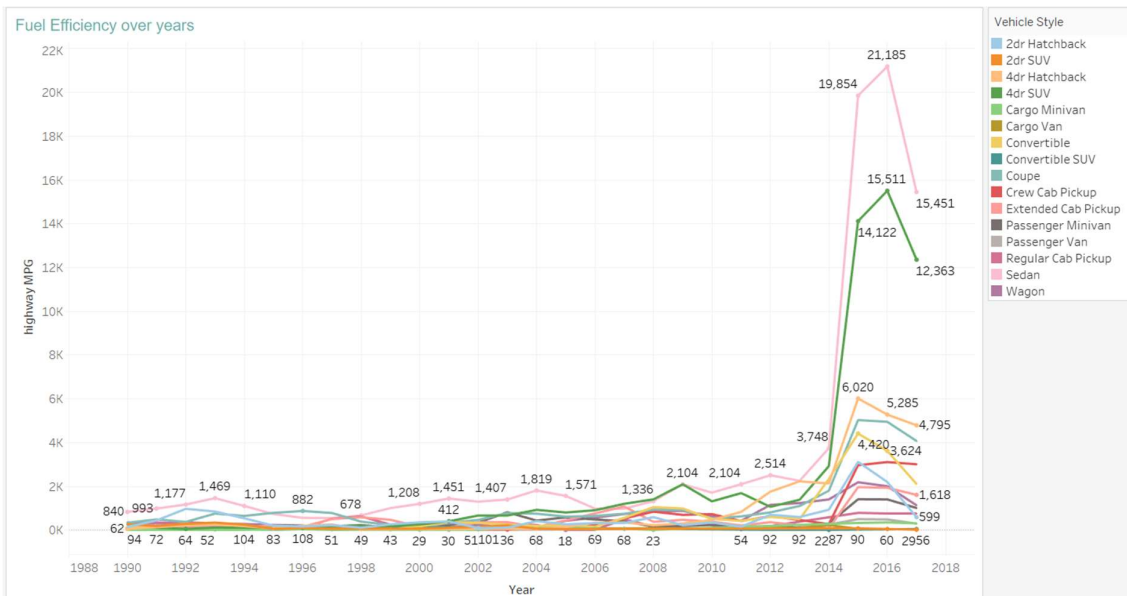


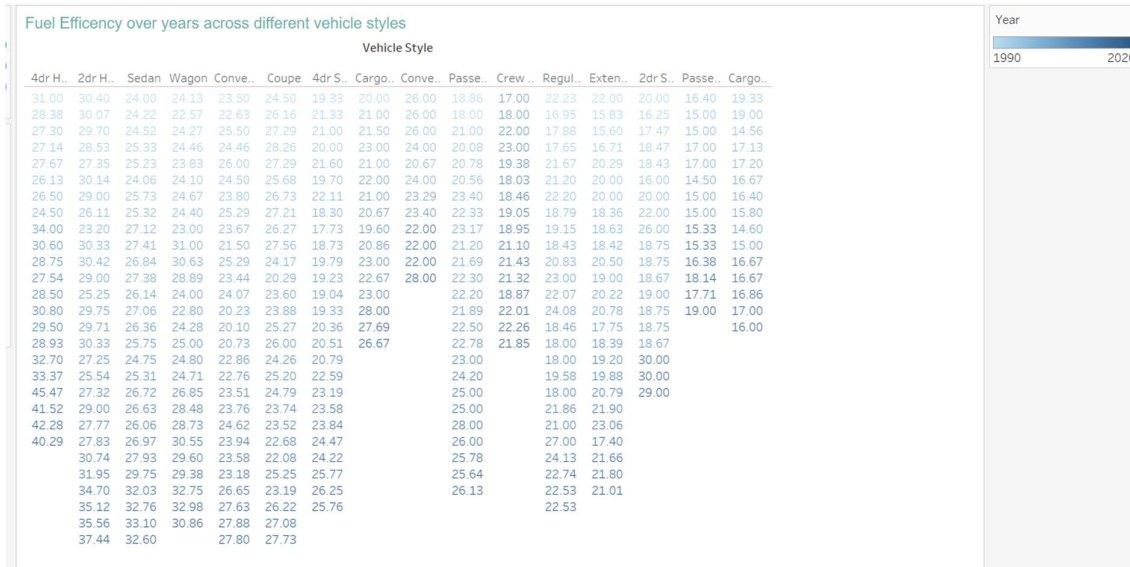
Above graphs results shows that “Automatic” transmission type has highest MSRP average.

But average MSRP of transmission types varies under different styles. Based on above graph “Coupe” style with “Automated_Manual” transmission type are the most sold vehicles with highest MSRP average.

F. Fuel Efficiency Across Body Styles and Model Years:

OUTPUT:



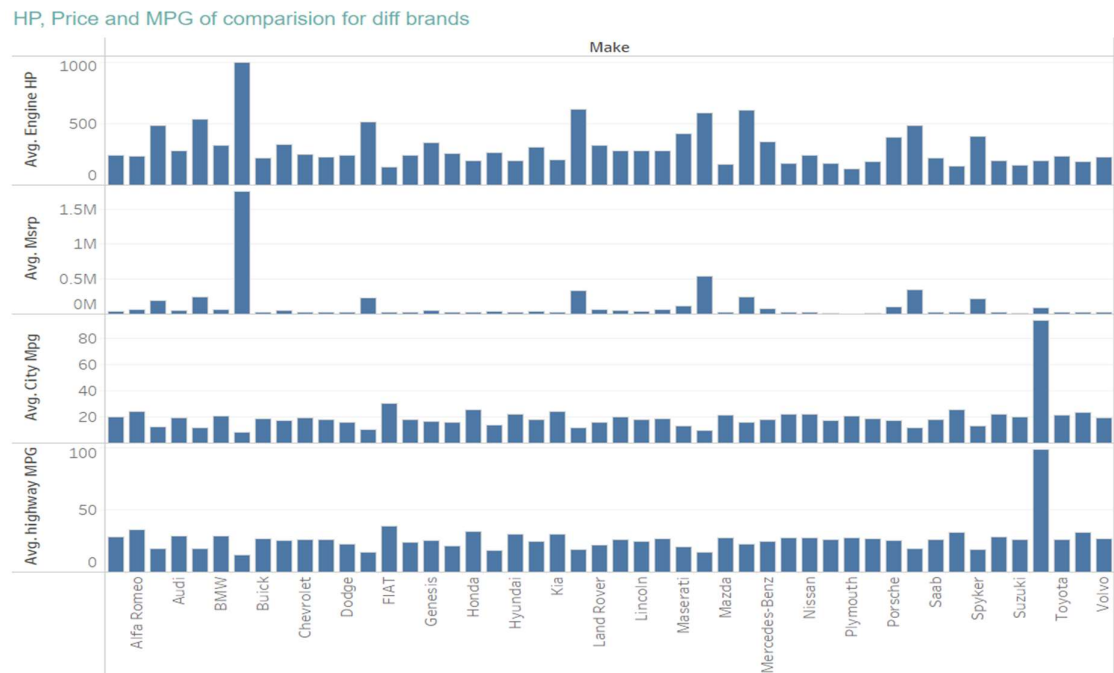


The city and highway MPG, which are used to calculate fuel efficiency, varied for various model years and body types. Transmission with a somewhat smaller the mean.

There are fluctuations in fuel efficiency over the years, but overall, there is a trend of improvement in city and highway MPG.

G. Car's horsepower, MPG, and price vary across different Brand

OUTPUT:



As per comparison of car features in above graph, brands that manufacturing vehicles with high engine horse power has high average MSRP compare to the brands that manufacture vehicles with high fuel efficiency.

Result:

Numerous significant insights on car pricing, market segments, automotive characteristics, and fuel efficiency were found by analyzing the car dataset. The findings allow for the important conclusions to be reached as follows:

1. Brands and body styles have a big impact on car costs. The typical MSRP of some brands, such as Rolls Royce, Maybach, and Bugatti, is greater than that of other brands.
2. Engine power, expressed in horsepower (HP), is a key factor in deciding automobile costs.
3. Automated and Automated_Manual transmission type vehicles also have high MSRP average.
4. The choice of body style also affects the price range of cars, with certain body types fetching higher prices. Higher horsepower vehicles often have higher average prices. This suggests that consumers are prepared to spend more for more powerful engines.
5. Features like no. of doors, vehicle size are least interested features among consumers.

Drive Link:

For Excel files:

https://docs.google.com/spreadsheets/d/1-RQvi_mVs1DtXE0jzJviCkZrWSiFYC5/edit?usp=sharing&ouid=106678072727235537296&rtpof=true&sd=true

For Project Dashboard:

<https://drive.google.com/file/d/1m8MMbrztpjimp7ECoGrffCtCTW3rLNbt/view?usp=sharing>

For Tableau Public dashboard:

https://public.tableau.com/views/Project_7_17032494664010/Dashboard1?:language=en-US&publish=yes&:display_count=n&:origin=viz_share_link