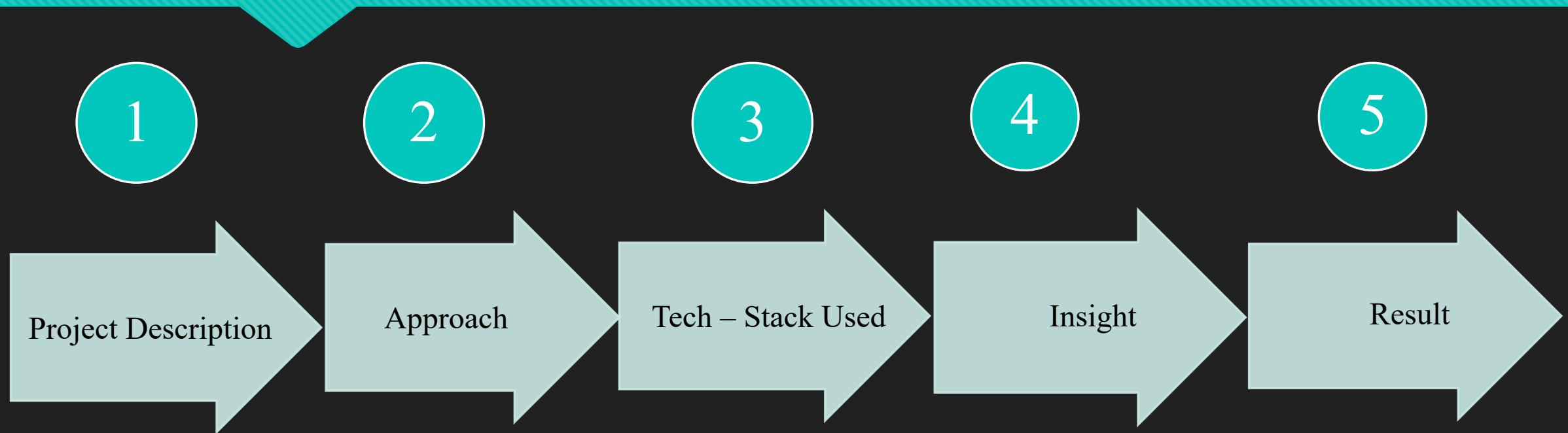


# Project: Analysing the Impact of Car Features on Price and Profitability



# Content



# Project Description

- The primary goal of this project is to assist a car manufacturer in optimizing pricing strategies and product development decisions to maximize profitability while meeting consumer demand.
- This involves analysing a comprehensive dataset encompassing various car attributes, market segments, and pricing information.

## About Dataset:

- The dataset "Car Features and MSRP" provides details about different car models, their specifications, and the Manufacturer's Suggested Retail Price (MSRP) collected by Cooper Union, a New York City-based private college.

The dataset offers information on:

- Observations: There are 11,159 entries (rows) in the dataset.
- Variables: It includes 16 different attributes (columns) describing various features of the cars, such as engine type, fuel efficiency, horsepower, transmission, fuel type, market category, and more.

# Approach

## 1. Data Preparation:

- Data Cleaning: Removal or imputation of null values, elimination of duplicates, handling outliers, and removal of irrelevant columns to ensure data quality.

## 2. Exploratory Data Analysis (EDA):

- Descriptive Statistics: Calculating summary statistics like mean, median, standard deviation, etc., to understand the central tendency and variability in the data.
- Pivot Tables and Functions: Using pivot tables and functions to aggregate and summarize data across different dimensions and variables.
- Graphs/Charts: Visual representation of data using graphs and charts to identify patterns, trends, and relationships within the dataset.

## 3. Statistical Analysis:

- Regression Analysis: Employing regression techniques to model relationships between variables, understanding how certain factors impact others, and predicting outcomes.
- Correlation Analysis: Determining the strength and direction of relationships between variables using correlation coefficients.

## 4. Dashboard Creation:

- Utilizing Tableau or similar tools to create interactive dashboards summarizing key findings and insights derived from the analysis.

## Tech- Stack Used

The analysis for this project has been performed using Microsoft Excel as it offers a user-friendly interface that allows users to work with data, perform calculations, and create visual representations without extensive programming knowledge.

# Descriptive Statistics

Categorical Attributes	
Column_name	Highest_Count
Make	Chevrolet
Make Model	Chevrolet Silverado 1500
Year	2014
Engine Fuel Type	regular unleaded
Transmission Type	AUTOMATIC
Driven_Wheels	front wheel drive
Market Category	N/A
Vehicle Size	Compact
Vehicle Style	Sedan

# Descriptive Statistics

	Engine HP	Engine Cylinders	Number of Doors	highway MPG	city mpg	Popularity	MSRP
mean	249.3860701	5.628828677	3.436093383	26.63748531	19.73325499	1554.911197	40594.73703
median	227	6	4	26	18	1385	29995
MODE	200	4	4	24	17	1385	2000
Standar Deviation	109.1918703	1.780559348	0.881315387	8.863000767	8.98779816	1441.855347	60109.1036
Variance	11922.86453	3.170391593	0.776716811	78.5527826	80.78051577	2078946.841	3613104336
Max	1001	16	4	354	137	5657	2065902
Min	55	0	2	12	7	2	2000

# Data Cleaning

- Data inspection
- Null values analysis & imputation
- Removal of duplicates
- Detection of outliers & suitable technique used
- Data engineering – columns transformed as per need (Market Category values separated into different columns. Also, N/A values replaced with highest count category i.e. Luxury.
- Data transformation – Standardization has been performed on numerical columns to transform data into similar scale.

## Null Value Analysis & Imputation

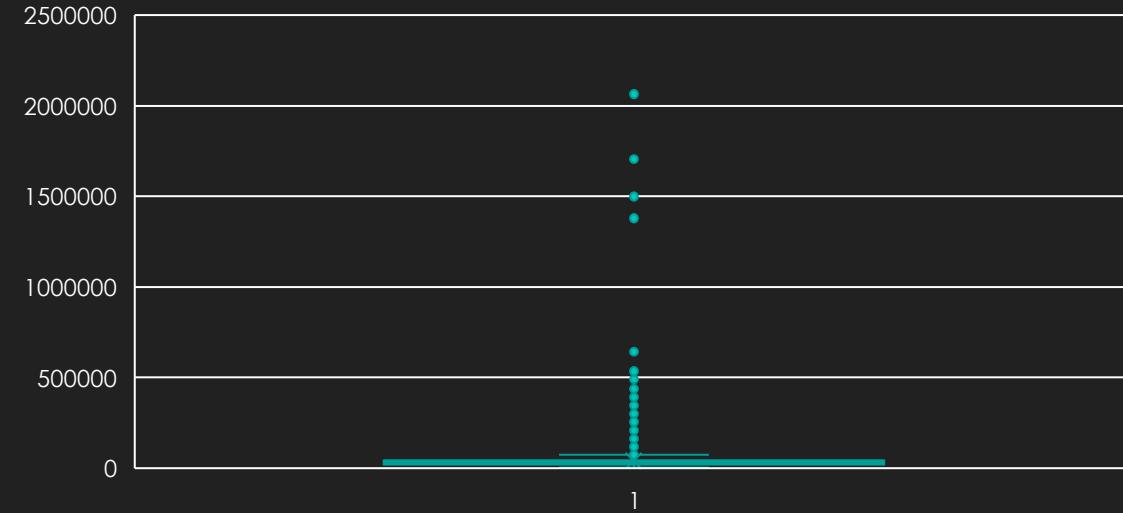
	Make	Model	Year	Engine Fuel Type	Engine HP	Engine Cylinders	Transmission Type	Driven_Wheels	Number of Doors	Market Category	Vehicle Size	Vehicle Style	Highway MPG	City MPG	Popularity	MSRP
Null Value (%)	0	0	0	0.025	0.579	0.251	0	0	0.050	0.008	0	0	0	0	0	0
Replaced with	-	-	-	Highest Count	Median	Median	-	-	Mean	Highest Count	-	-	-	-	-	-

# Outliers

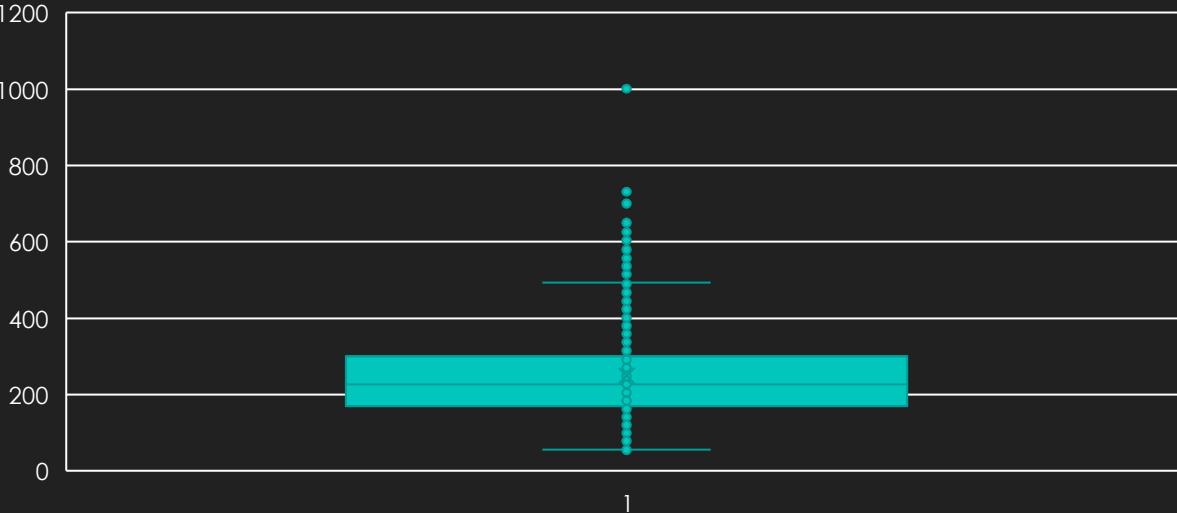
	Engine HP	Engine Cylinders	Number of Doors	Highway MPG	city MPG	Popularity	MSRP
Outliers (Y/N)	Y	Y	N	Y	Y	Y	Y
Replacement with	-	-	-	Median	-	-	-
				Not replaced as these outliers belong to electric cars which tend to higher MPG			

715 duplicates removed.

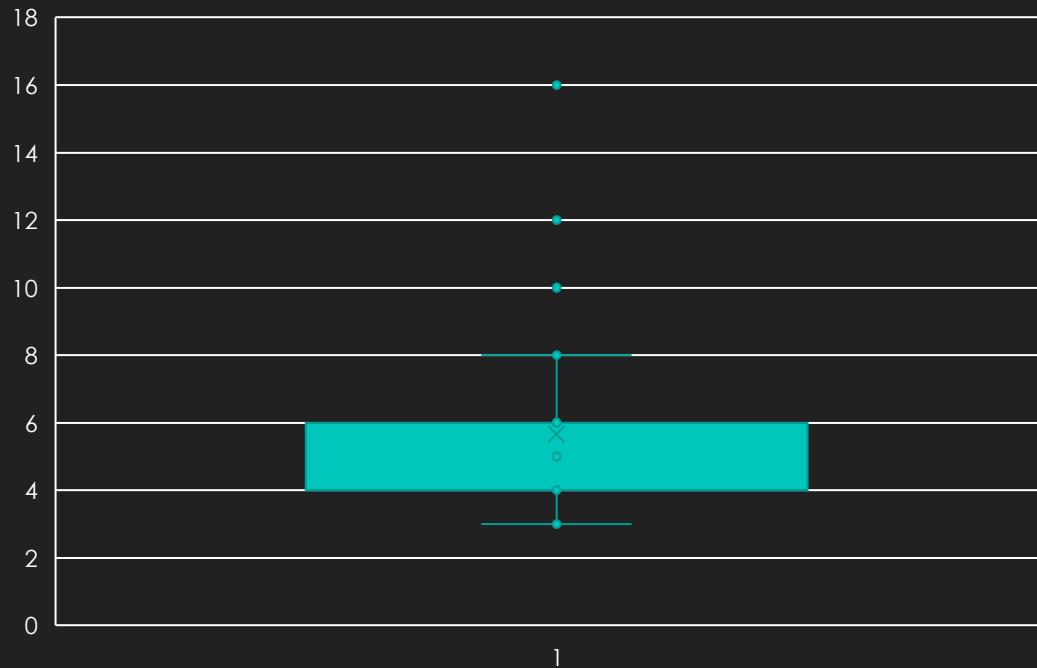
MSRP



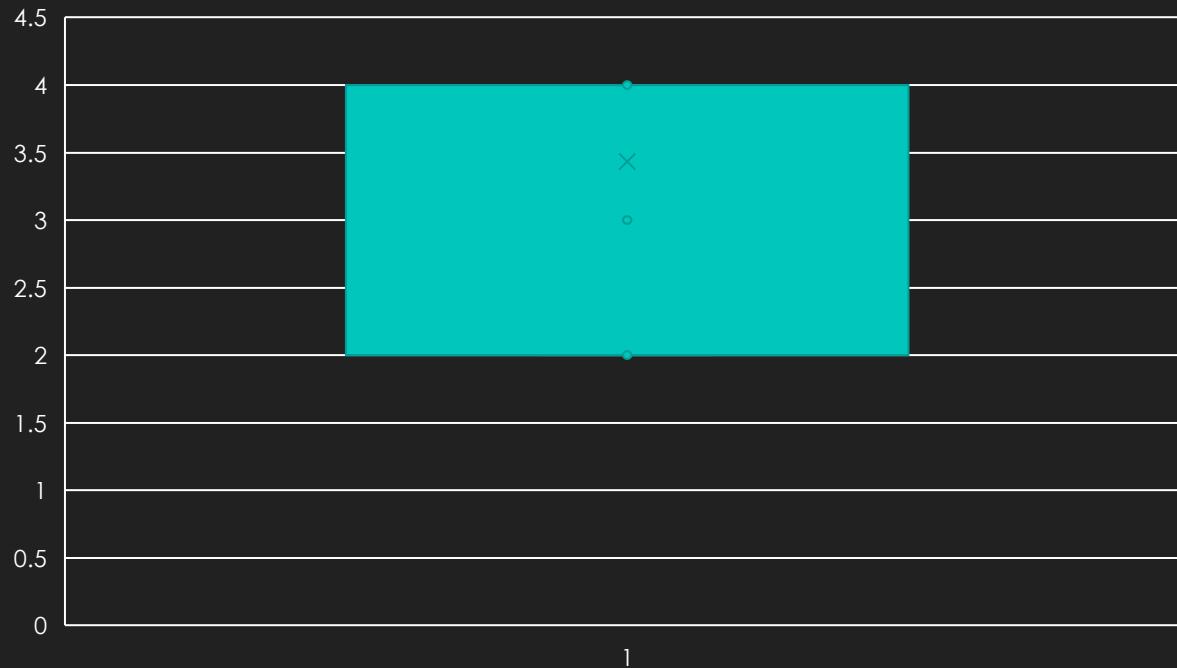
Engine HP



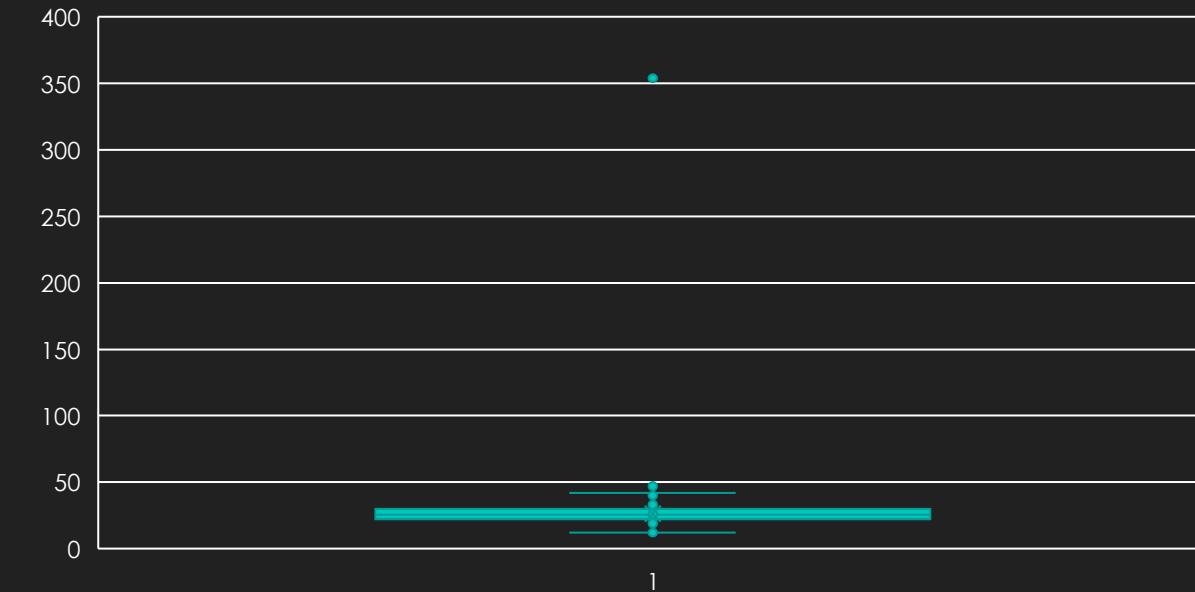
Engine Cylinders



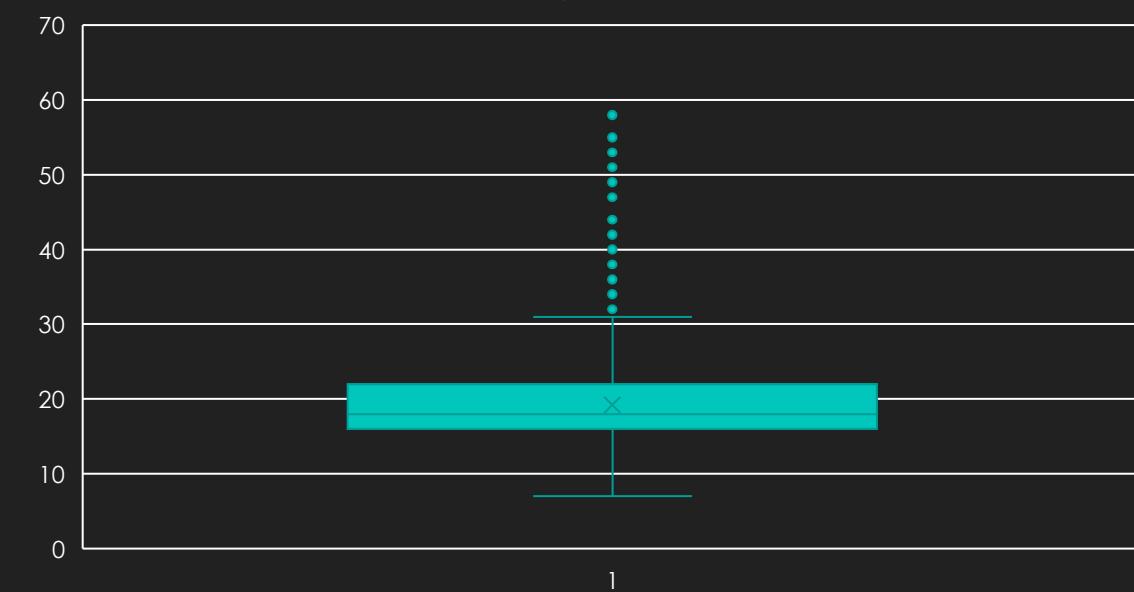
Number of Doors



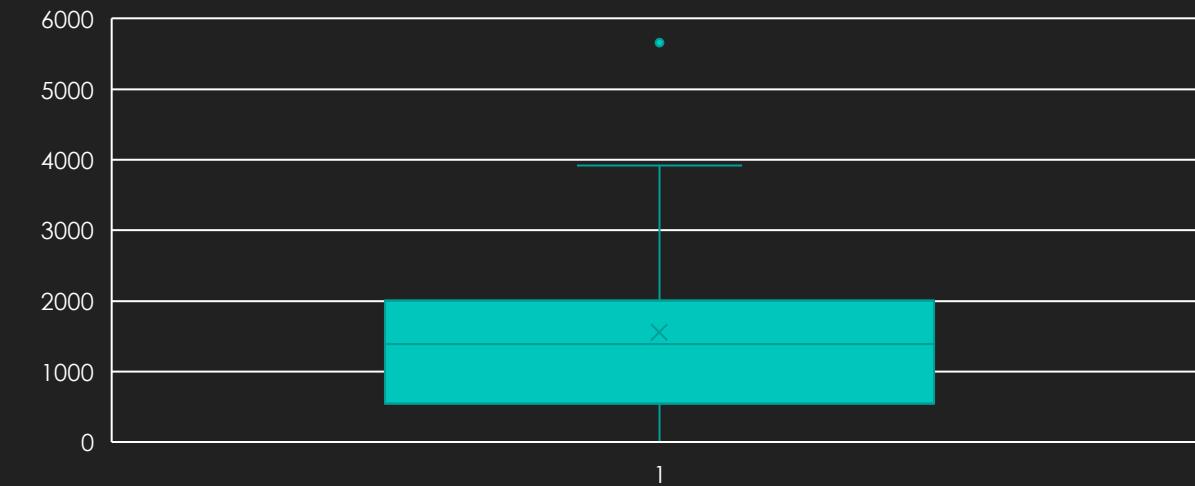
### Highway MPG



### City MPG



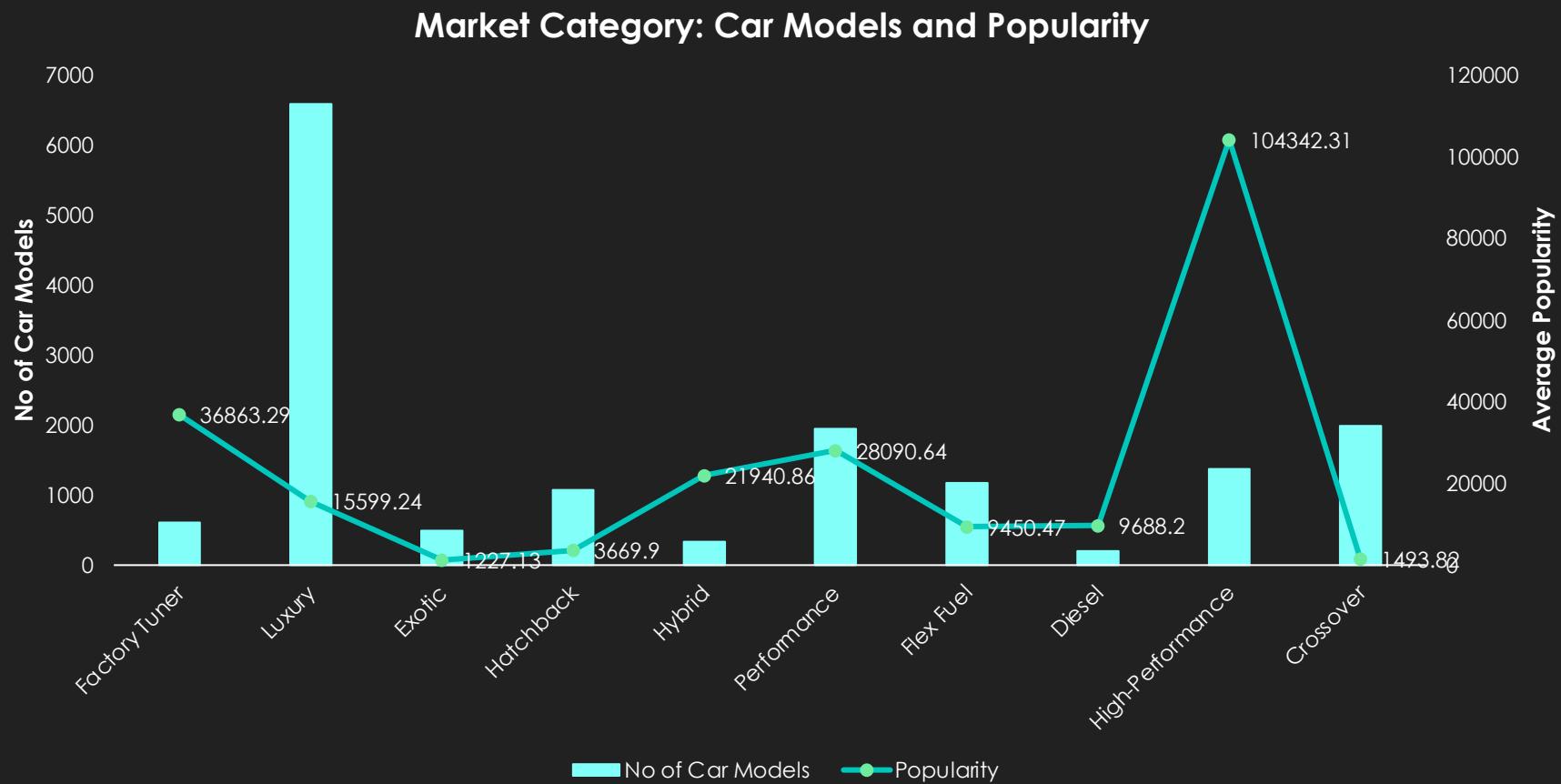
### Popularity



# Insight\_1

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

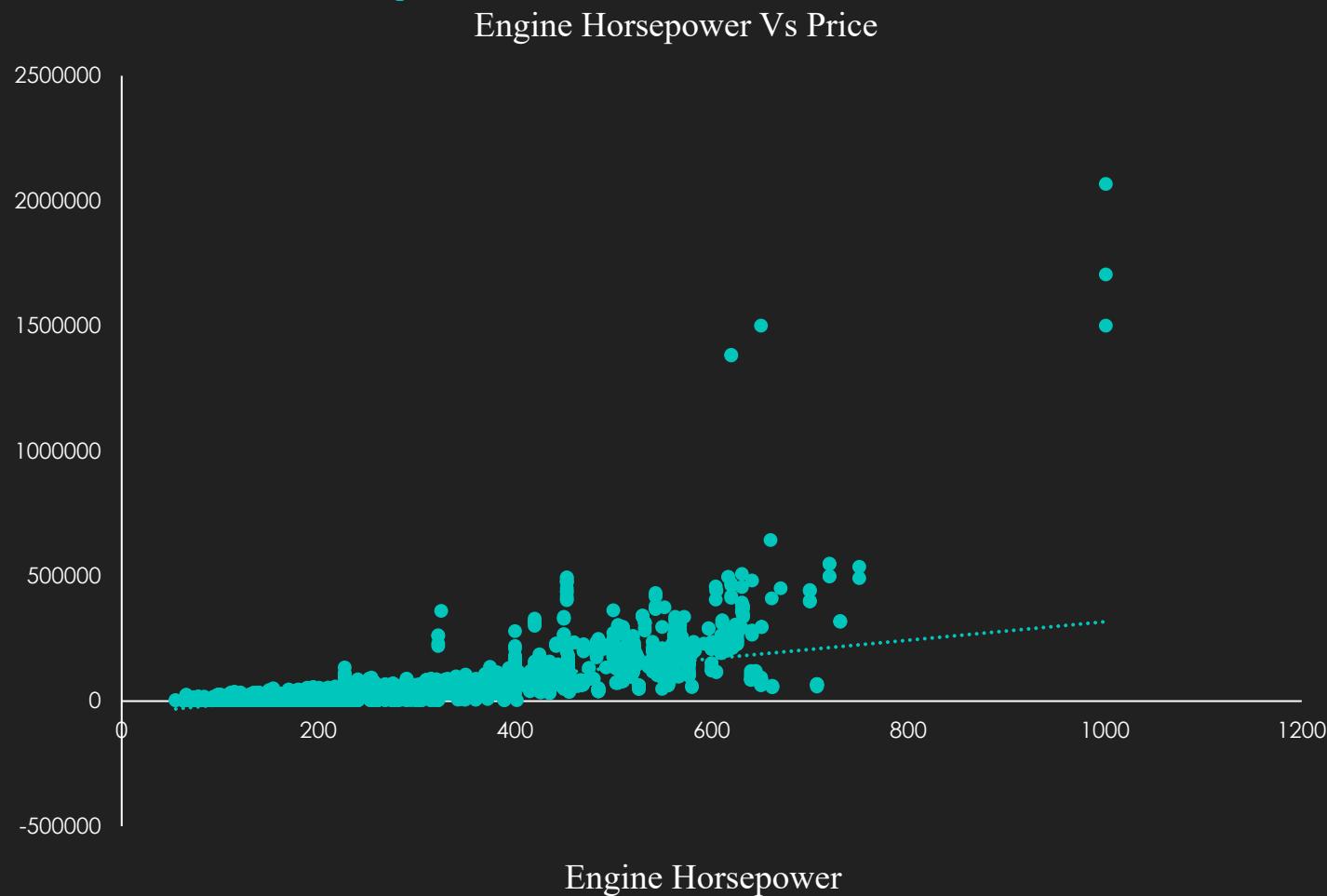
Market Category	No of Car Models	Popularity
Factory Tuner	609	36863.29
Luxury	6598	15599.24
Exotic	491	1227.13
Hatchback	1075	3669.9
Hybrid	337	21940.86
Performance	1953	28090.64
Flex Fuel	1179	9450.47
Diesel	202	9688.2
High-Performance	1373	104342.31
Crossover	1994	1493.82



- It is evident from the table that luxury and crossover market categories have highest car models and diesel category has the least models.
- Also, performance and high- performance market categories are the most popular among customers and crossover and factory tuner are not so popular. Also, there is a large disparity in the popularity scale of most and least popular categories.

## Insight\_2

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

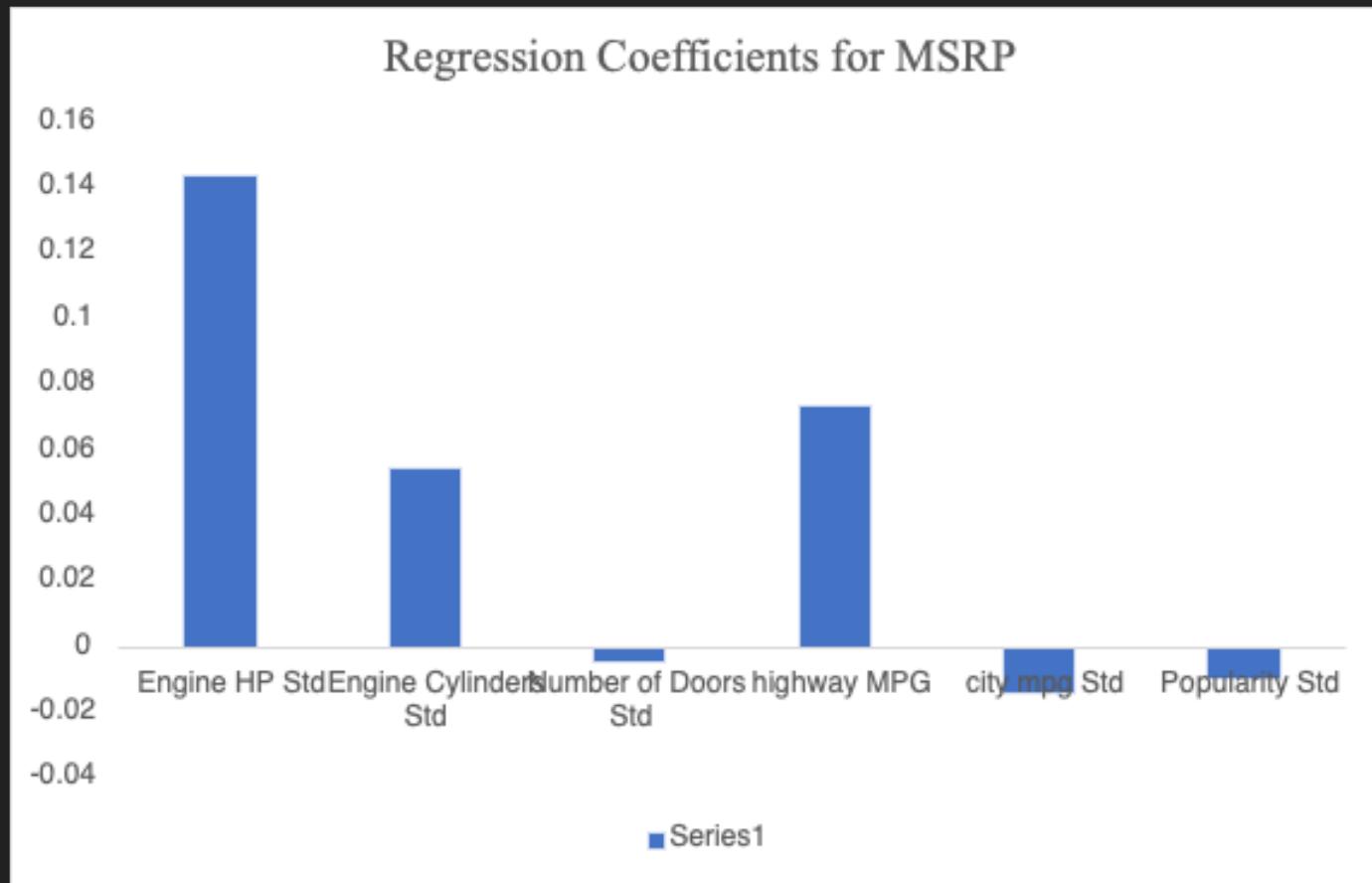


It is evident from the scatter plot that there is a positive linear relationship between the power of engine and the car's price. Also, the trend line indicates the same positive linear relation.

## Insight\_3

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

- It can be interpreted through the graph that there is a weak linear correlation between the horsepower of engine and the price of a car. Also, the cars with good highway mileage are offered at a higher price.
- As the number of cylinders in a car increases, the price also increases.
- Manufacturer's suggested retail price (MSRP) and variables such as number of doors, city mileage and popularity have a negative correlation which suggests that as the value of one of these variables increase, the value of other variable tends to decrease.



# Summary output

Regression Statistics	
Multiple R	0.67564978
R Square	0.456502625
Adjusted R Square	0.456211258
Standard Error	0.02208901
Observations	11199

	df	SS	MS	F	Significance F
Regression	6	4.586760361	0.76446006	1566.759296	0
Residual	11192	5.460849676	0.000487924		
Total	11198	10.04761004			

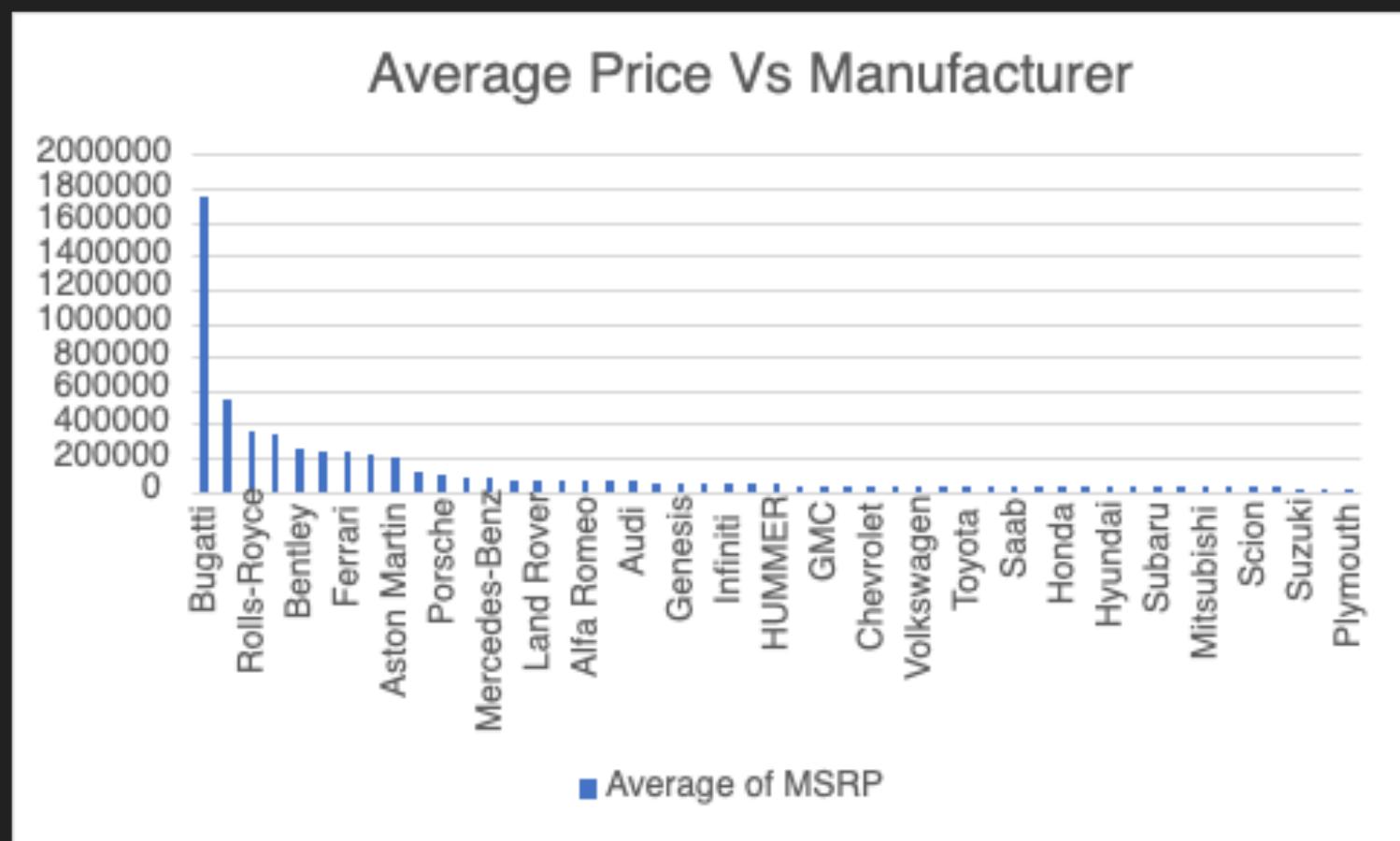
Column	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95%2	Upper 95%3
Intercept	-0.03388608	0.001420715	-23.85142363	1.09E-122	-0.036670931	-0.031101228	-0.036670931	-0.031101228
Engine HP Std	0.144236157	0.002917287	49.44187359	0	0.13851776	0.149954553	0.13851776	0.149954553
Engine Cylinders Std	0.05466031	0.003608305	15.1484724	2.50E-51	0.047587397	0.061733223	0.047587397	0.061733223
Number of Doors Std	-0.004507632	0.000484969	-9.294687637	1.75E-20	-0.005458256	-0.003557008	-0.005458256	-0.003557008
highway MPG	0.073769191	0.008077901	9.132222317	7.86E-20	0.057935084	0.089603299	0.057935084	0.089603299
city mpg Std	-0.013930196	0.009004464	-1.547032137	0.121883778	-0.03158053	0.003720138	-0.03158053	0.003720138
Popularity Std	-0.009463982	0.000819354	-11.55054714	1.09E-30	-0.011070059	-0.007857905	-0.011070059	-0.007857905

The coefficient values indicate weak linear relationship between all the independent variables and MSRP. However, there is also weak negative correlation between Number of doors and City Mileage.

## Insight\_4

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

- The manufacturer Bugatti has the highest average price which is also quite high than the other high end manufacturers like Rolls-Royce, Bentley, Ferrari and others.
- Manufacturers like Scion, Suzuki and Plymouth have the lowest average price.

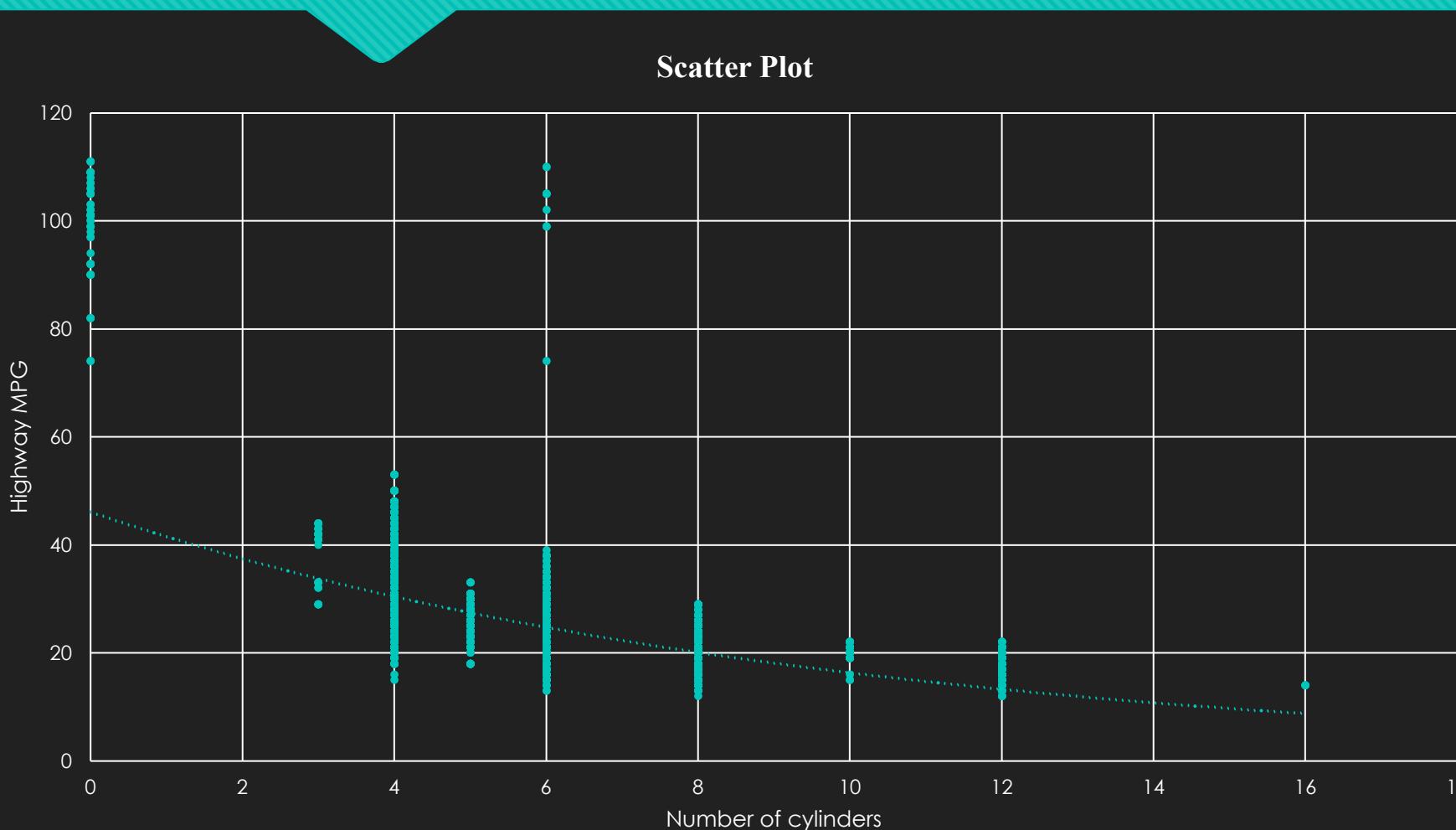


# Average MSRP per manufacturer

Make	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
Grand Total	41925.92714

# Insight\_5

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



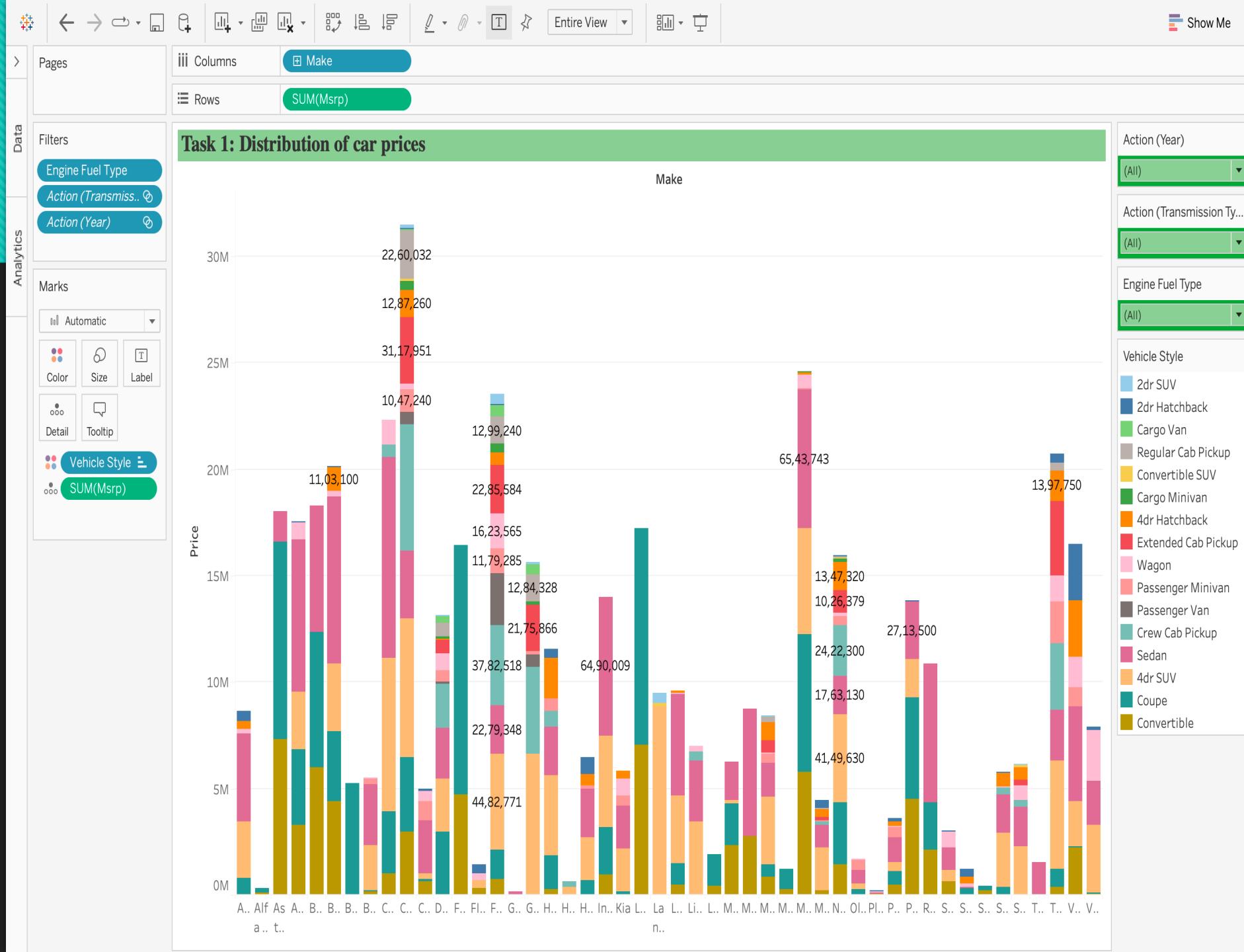
Correlation:  
-0.63193

There is a negative correlation between highway mileage and the number of cylinders, which is also corroborated by the scatter graph.

# Task 1

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

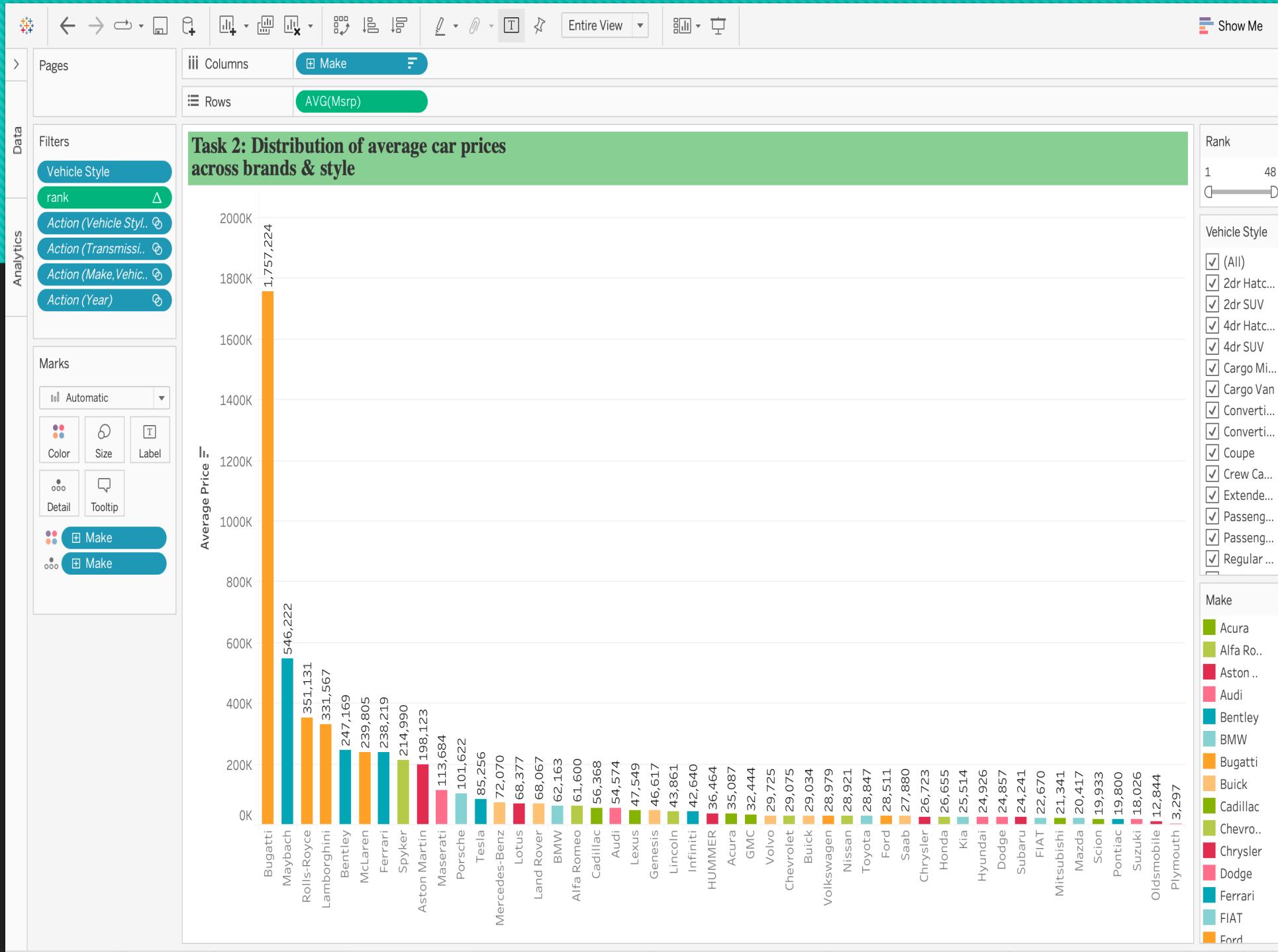
- The car brand Chevrolet holds the highest sum of MSRP, which is followed by Mercedes-Benz.
- Among the vehicle style category, Coupe holds the highest sum of MSRPs.



# Task 2

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

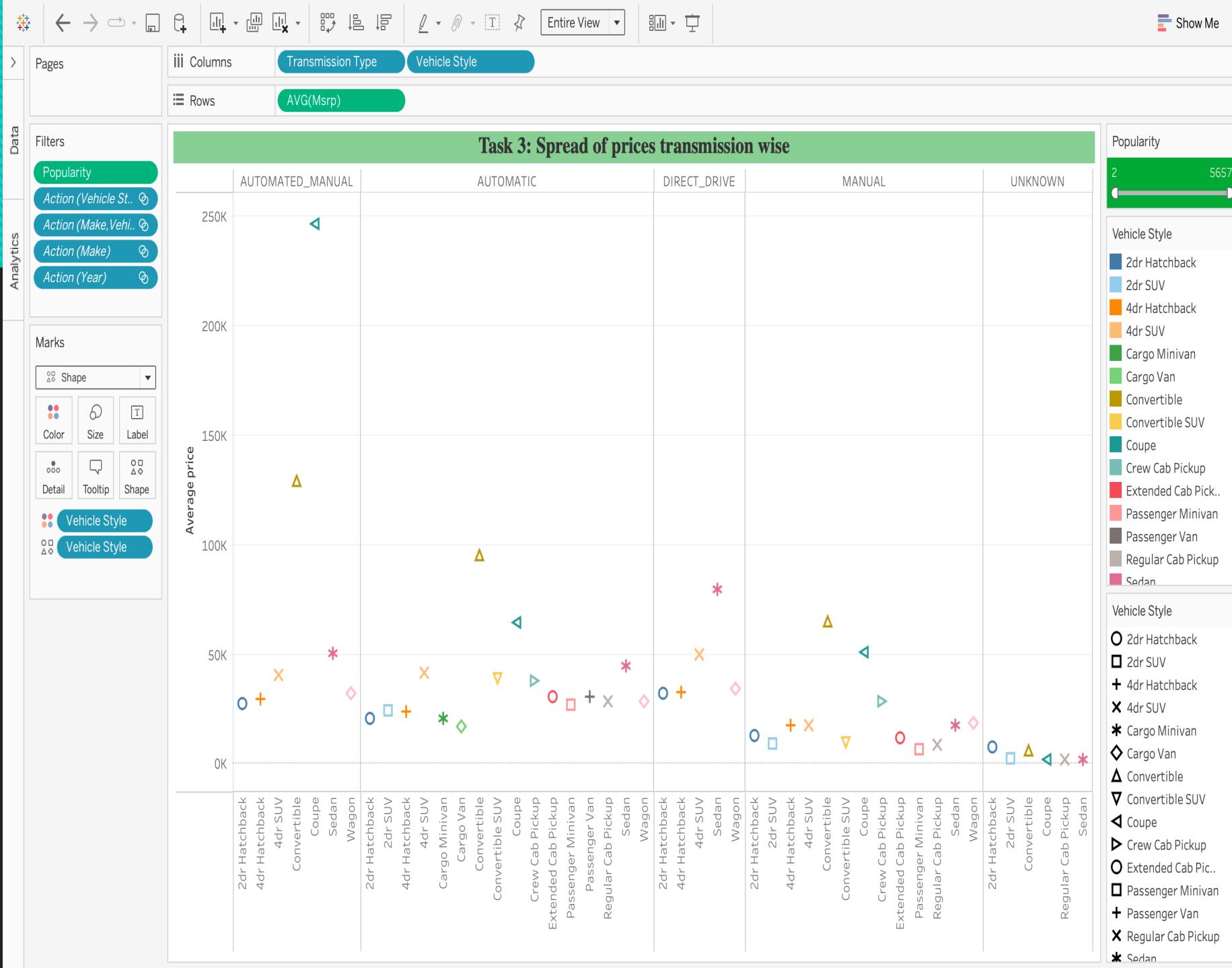
- Among all the manufacturers, Bugatti has the highest average MSRP, followed by Maybach, Roll Royce, while Plymouth has the lowest average price for its 2dr Hatchback.
- Coupe style of Bugatti has the highest average MSRP, followed by Maybach's convertible.



# Task\_3

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

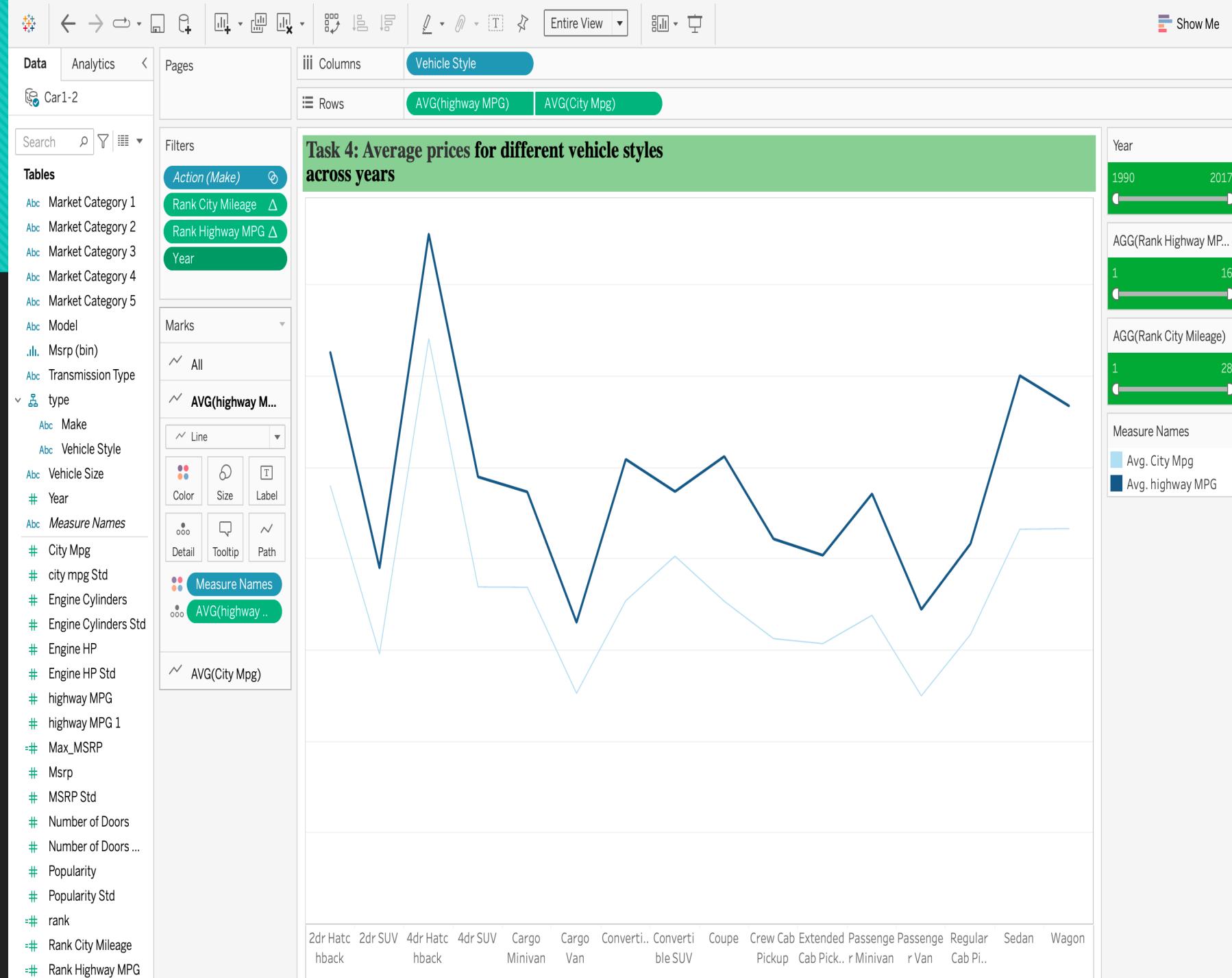
- Among all the transmission types, automatic manual Coupe had the highest average MSRP, followed by automatic manual convertible.
- Manual transmission type cars have the lowest average MSRPs.
- Also, Sedans have the lowest MSRPs.



# Task 4

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

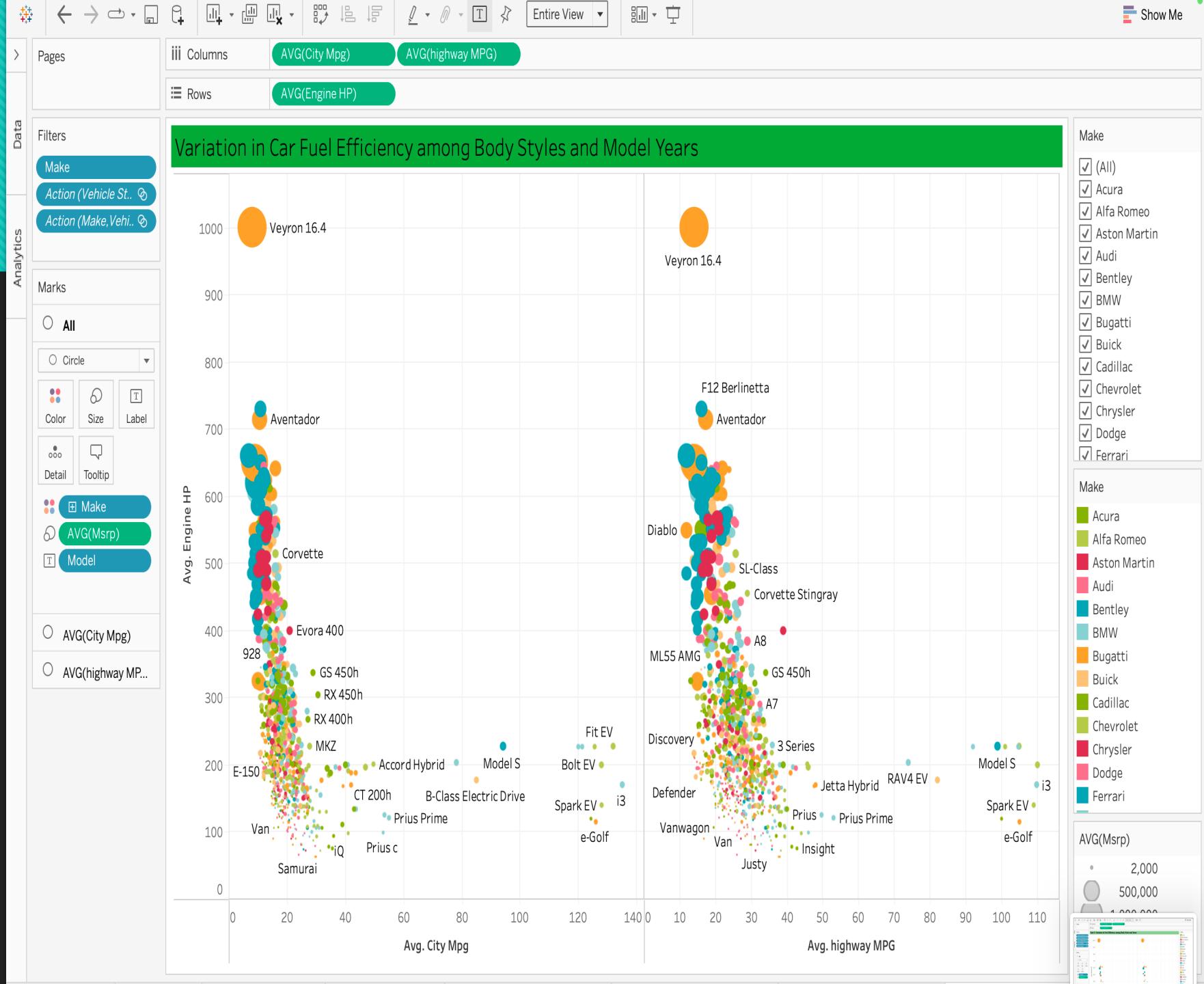
- The average mileage for city as well as highways have been increasing over the years, there have been some phases of dip in mileage but the overall trend is increasing.
- Also, 4dr Hatchback vehicle style has the highest mileage i.e. most fuel efficient.
- Whereas, Cargo Van has the lowest mileage.



# Task\_5

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

- Car model Veyron 16.4 of Bugatti has exhibited the highest horsepower as well as the highest average MSRPs but showcased the lowest mileage for city and highways, followed by Ferrari's F12 Berlinetta and Lamborghini's Aventador.
- i3 BMW is the most fuel efficient car model in the city mileage segment, while Chevrolet's Bolt EV had the best highway mileage.



# Task\_5

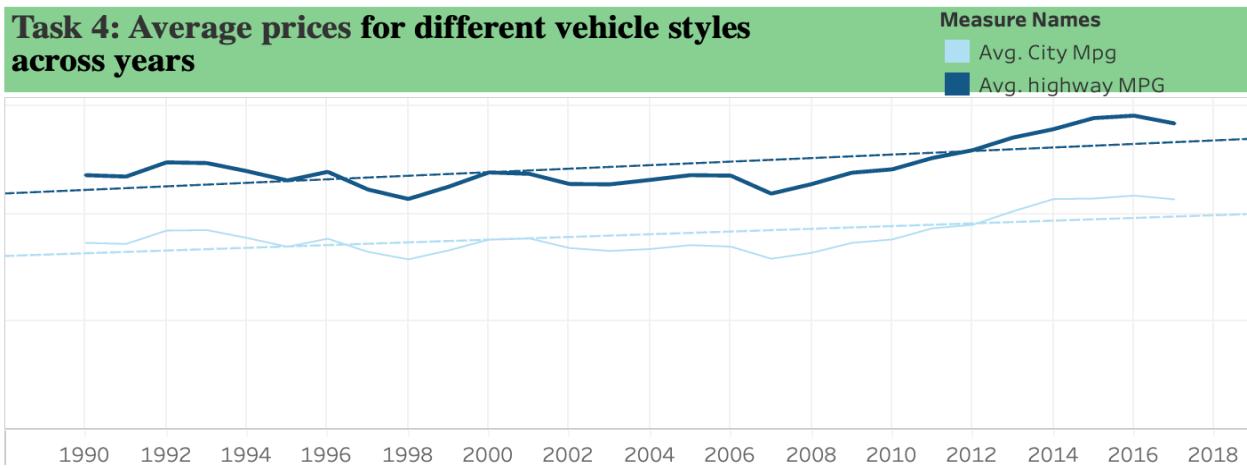
How does the car's horsepower, MPG, and price vary across different Brands?

- The manufacturer Bugatti has the highest horsepower as well as the highest average MSRP but has a low city and highway mileage.
- On the other hand, the manufacturer Tesla has low horsepower cars but these cars have high highway as well as city mileage.

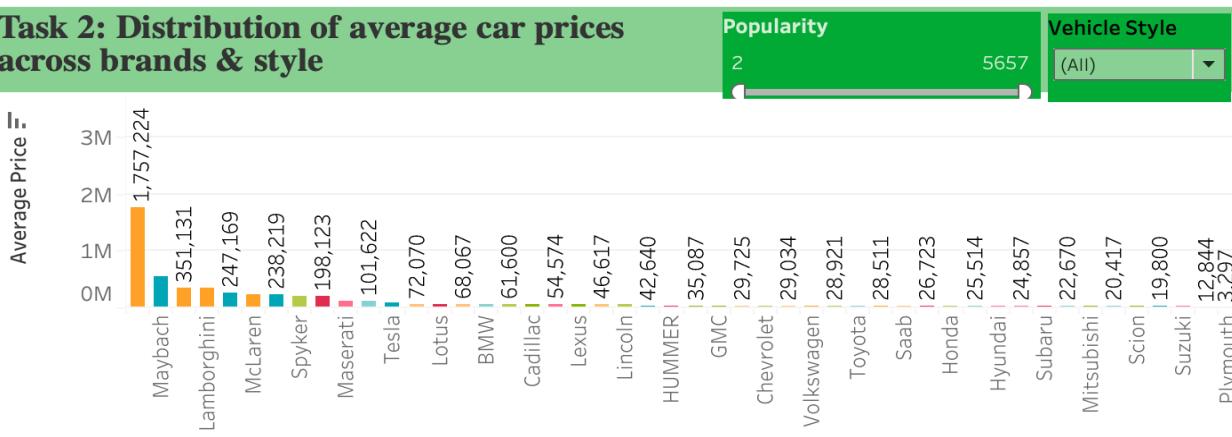


# Impact of Car Features on Price and Profitability

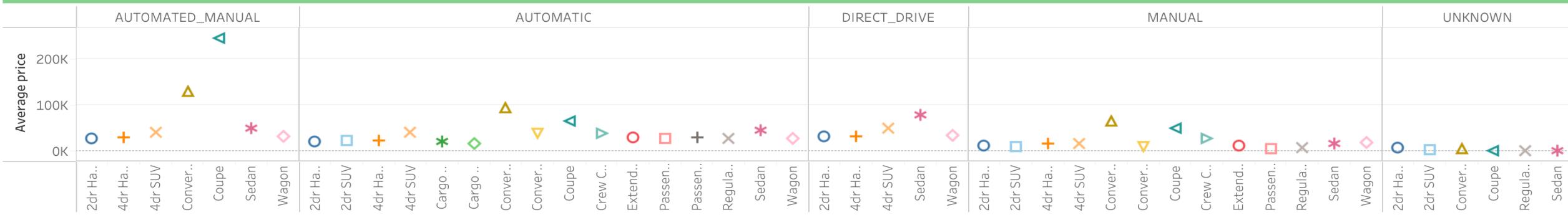
## Task 4: Average prices for different vehicle styles across years



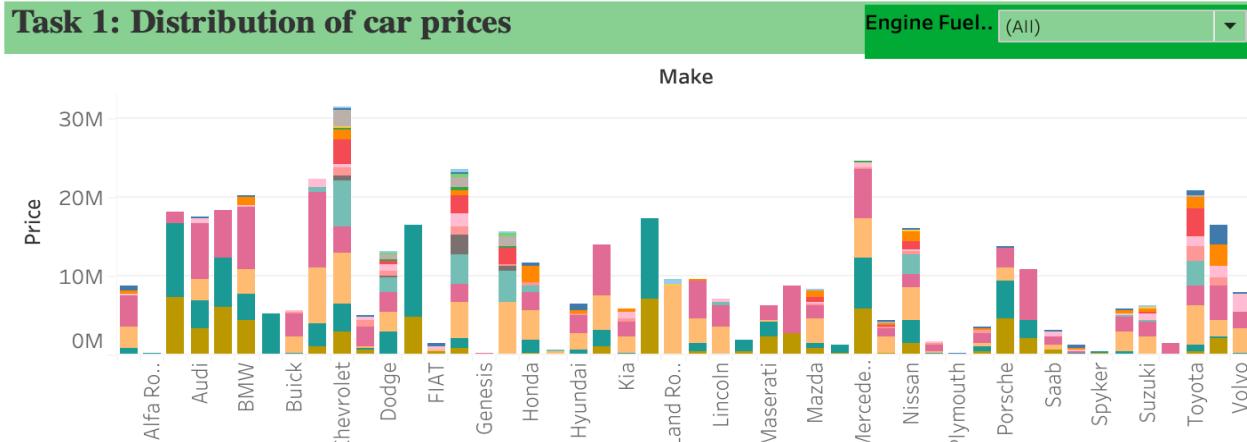
## Task 2: Distribution of average car prices across brands & style



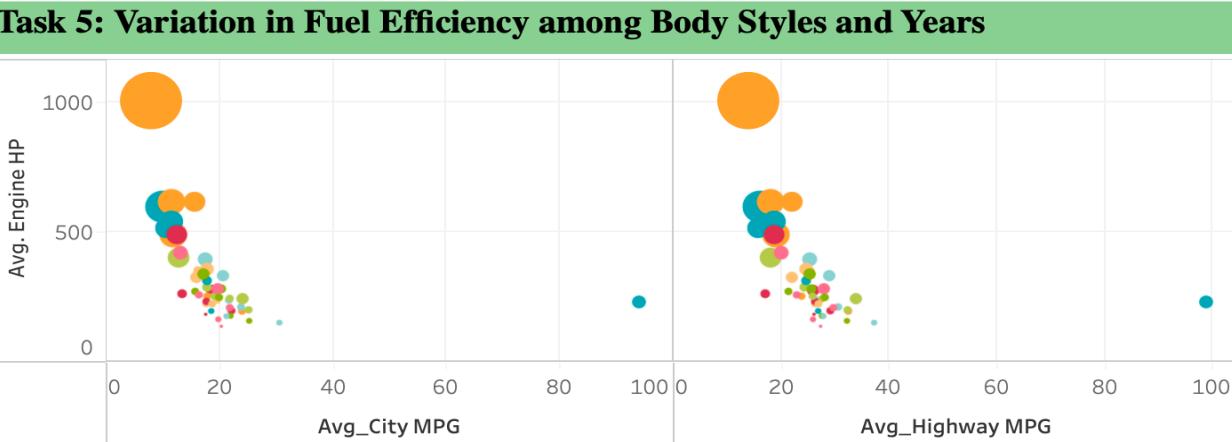
## Task 3: Spread of prices transmission wise



## Task 1: Distribution of car prices



## Task 5: Variation in Fuel Efficiency among Body Styles and Years



# Dashboard Insights

- Chevrolet secures the leading position in total MSRP (Manufacturer's Suggested Retail Price) sum, signifying a substantial overall value across its vehicles. Conversely, Bugatti boasts the highest average MSRP, implying premium pricing per vehicle. This discrepancy suggests Chevrolet's wider market presence, as the brand's collective MSRP outpaces Bugatti's despite the latter's higher average pricing.
- The vehicle styles of coupes and convertibles exhibit the highest average MSRP (Manufacturer's Suggested Retail Price) compared to other categories.
- 4dr Hatchback vehicle style is the most fuel efficient, whereas, Cargo Van has the lowest efficiency.
- Bugatti's Veyron 16.4 model stands out with the highest horsepower and average MSRP among cars. However, it records the lowest city and highway mileage. Following this trend are Ferrari's F12 Berlinetta and Lamborghini's Aventador.
- BMW i3 emerges as the most fuel-efficient car model for city mileage, while Chevrolet's Bolt EV leads in highway mileage efficiency.
- Tesla, as a manufacturer, produces cars with lower horsepower. However, these vehicles showcase commendable efficiency with high mileage ratings both on highways and in city driving conditions.

## Excel File Link

[https://docs.google.com/spreadsheets/d/17Z07ES1pqNTrXFTCPHuFP3jV7A4LCBO3/edit?usp=share\\_link&oid=101949921485202693908&rtpof=true&sd=true](https://docs.google.com/spreadsheets/d/17Z07ES1pqNTrXFTCPHuFP3jV7A4LCBO3/edit?usp=share_link&oid=101949921485202693908&rtpof=true&sd=true)

## Loom Video Links

- <https://www.loom.com/share/f4fabec8255f4c16848aa92c008c1392?sid=9a38d9ed-9518-4fdf-a172-72118c414959>
- <https://www.loom.com/share/c021f78c1e6d410888722dc319a8e0cc?sid=f3f2d334-a652-40a8-879a-261ae146d9ff>
- <https://www.loom.com/share/0d4fa05b16f647c4a23b2faeb4480071?sid=440a485e-d2d5-4102-a42b-acf1d7173dbc>
- <https://www.loom.com/share/6b57a57c3c6b4861ab44f060c04ce1fb?sid=c2ef3ea8-f982-4b85-90b9-7c0b21164ec6>
- <https://www.loom.com/share/711e59436e5843b68a6a85b8491f3254?sid=755674d1-1e67-41ec-91a0-e8c91157534a>