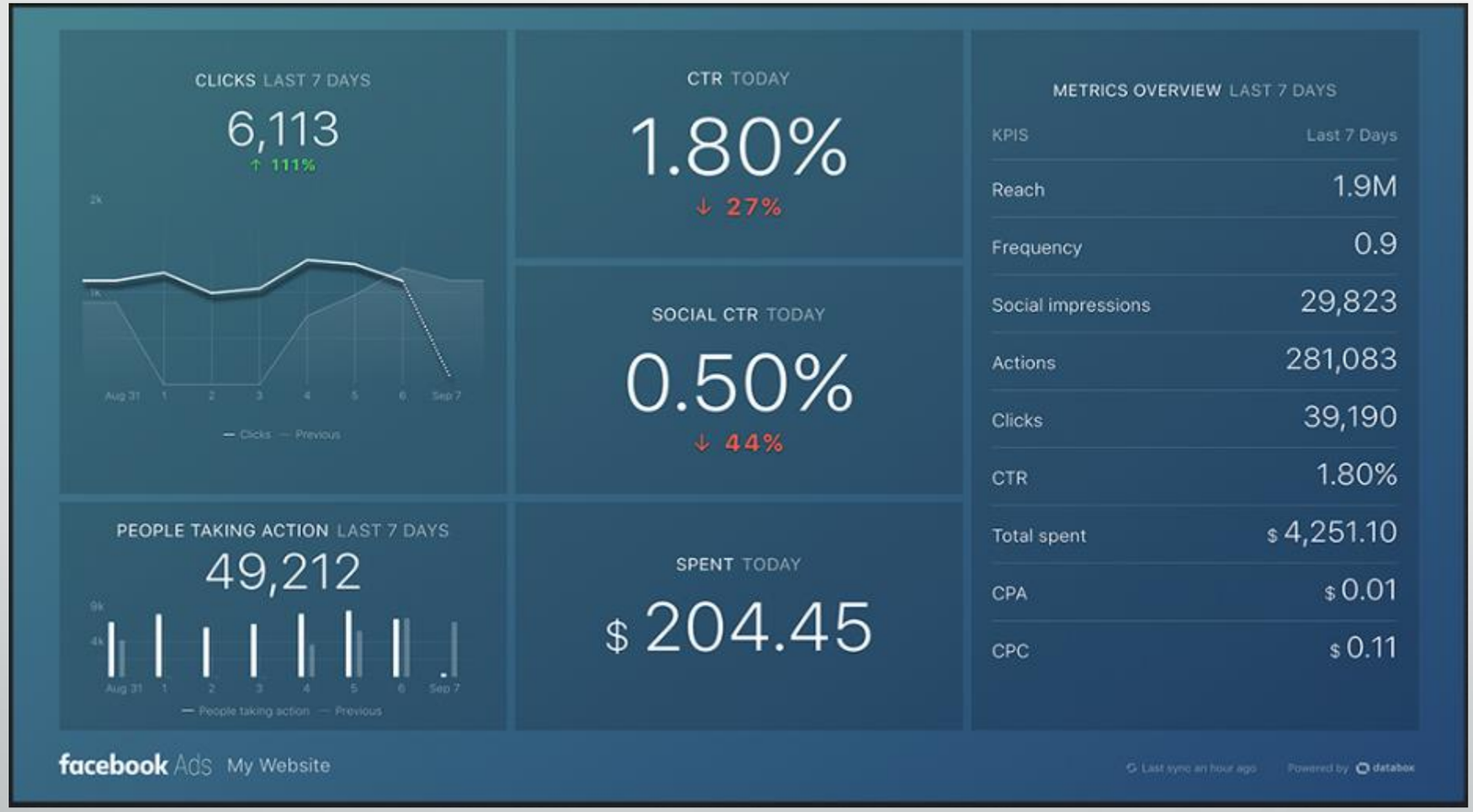
Impact of Car Features



By Hardi Palan

Project Description:

- 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.
- It is important to know impact of car features on price and profitability in the automotive industry.
- This dataset includes variables such as car's make, model, year, fuel type, engine power, transmission, wheels, number of doors, market category, size, style, estimated miles per gallon, popularity, and manufacturer's suggested retail price (MSRP).

Approach:

- First I have use =COUNTBLANK to count null values in each column. Then we will remove duplicate rows and use Median/Mode to remove null values.
- Regression analysis: Used to determine the relationship between variables and identify significant predictors.
- Various types of charts included scatter plots, line charts, column charts and bubble charts to visually represent the data and identify patterns and trends.
- Used pivot tables and formulas to filter the variables.

Tech-Stack Used:

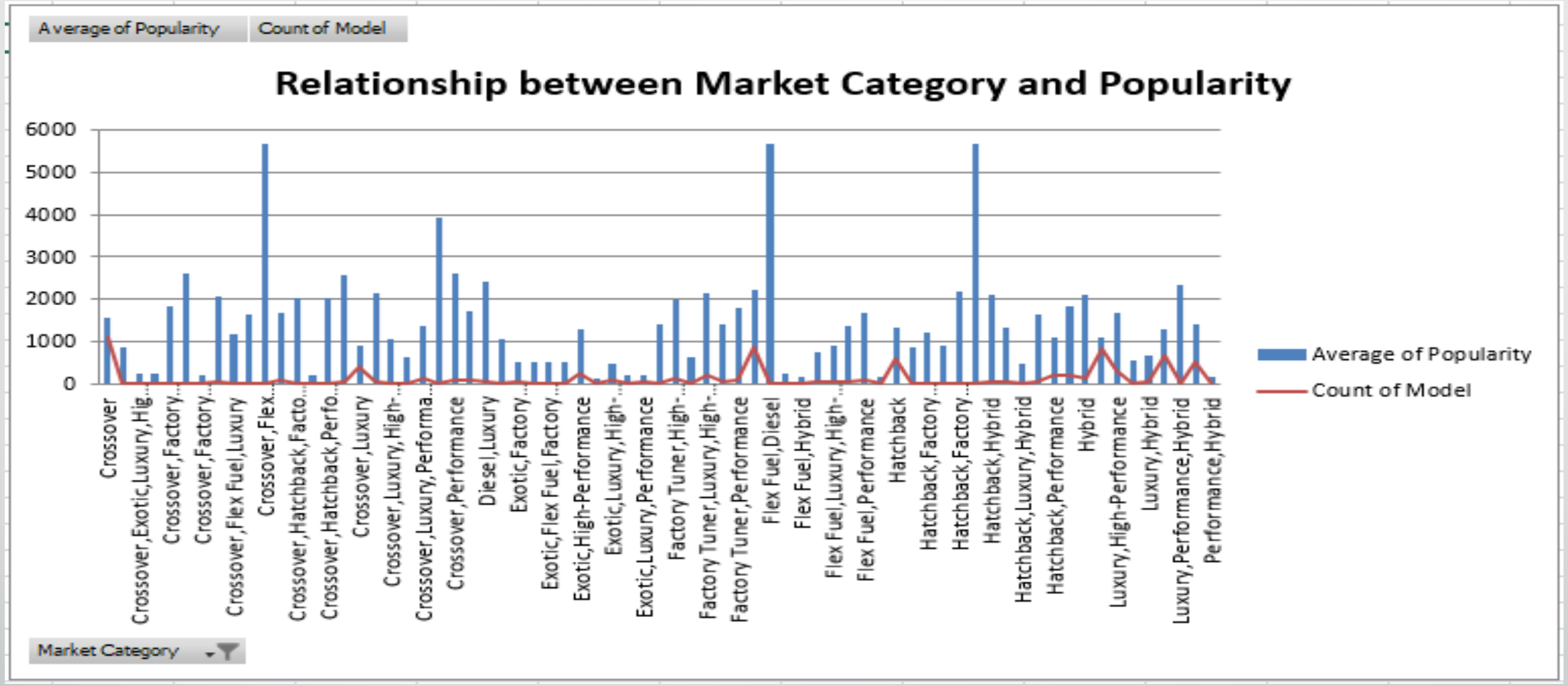
- Used Microsoft Excel for data cleaning, statistical analysis and visualization.

Analysis

1) 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.
- **Result:**

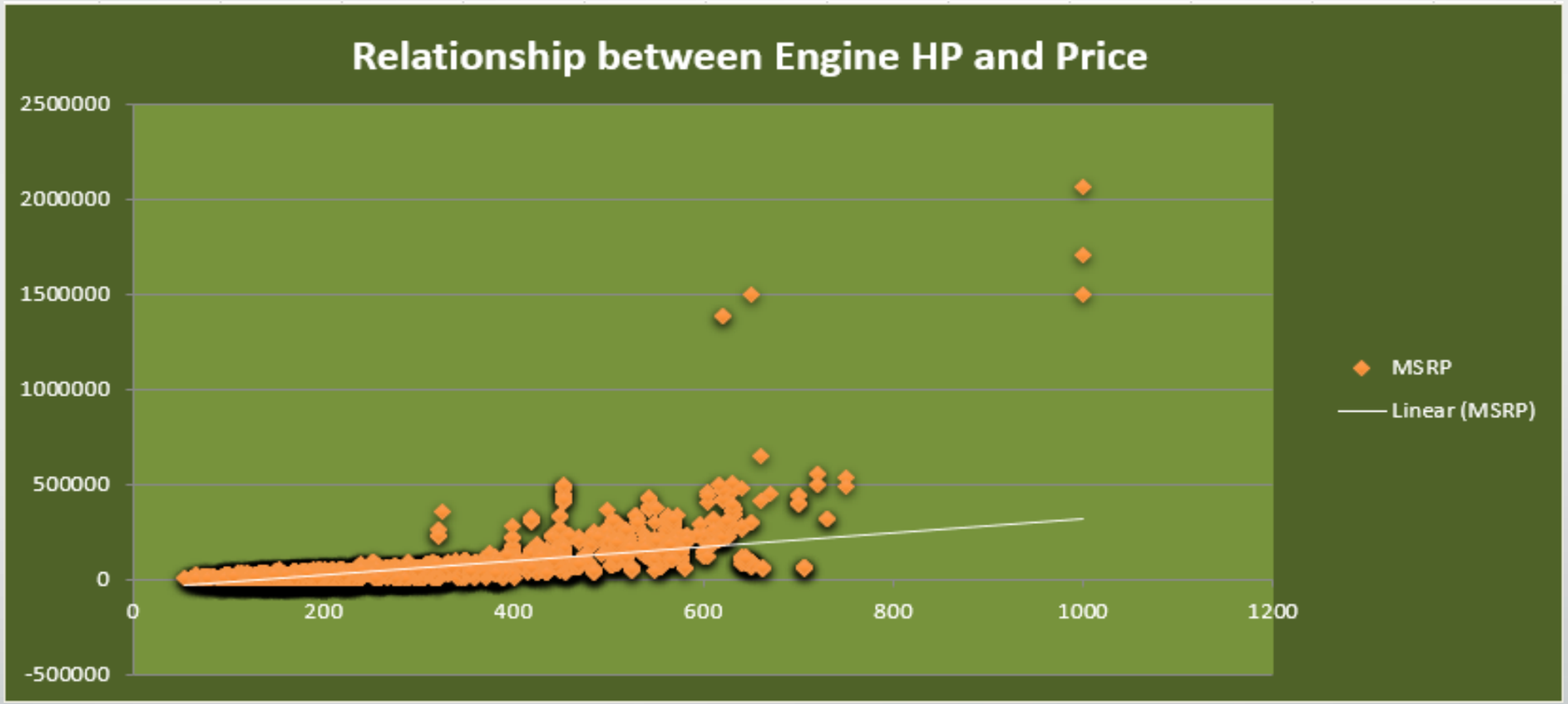
Row Labels	Average of Popularity	Count of Model
Crossover	1556.168372	1075
Crossover,Diesel	873	7
Crossover,Exotic,Luxury,High-Performance	238	1
Crossover,Exotic,Luxury,Performance	238	1
Crossover,Factory Tuner,Luxury,High-Performance	1823.461538	26
Crossover,Factory Tuner,Luxury,Performance	2607.4	5
Crossover,Factory Tuner,Performance	210	4
Crossover,Flex Fuel	2073.75	64
Crossover,Flex Fuel,Luxury	1173.2	10
Crossover,Flex Fuel,Luxury,Performance	1624	6
Crossover,Flex Fuel,Performance	5657	6
Crossover,Hatchback	1675.694444	72
Crossover,Hatchback,Factory Tuner,Performance	2009	6
Crossover,Hatchback,Luxury	204	7
Crossover,Hatchback,Performance	2009	6
Crossover,Hybrid	2563.380952	42
Crossover,Luxury	889.2142857	406
Crossover,Luxury,Diesel	2149.411765	34
Crossover,Luxury,High-Performance	1037.222222	9
Crossover,Luxury,Hybrid	630.9166667	24
Crossover,Luxury,Performance	1349.089286	112
Crossover,Luxury,Performance,Hybrid	3916	2
Crossover,Performance	2585.956522	69
Diesel	1730.904762	84
Diesel,Luxury	2416.106383	47
Exotic,Factory Tuner,High-Performance	1046.380952	21
Exotic,Factory Tuner,Luxury,High-Performance	523.0196078	51
Exotic,Factory Tuner,Luxury,Performance	520	3
Exotic,Flex Fuel,Factory Tuner,Luxury,High-Performance	520	13
Exotic,Flex Fuel,Luxury,High-Performance	520	11
Exotic,High-Performance	1280.047244	254
Exotic,Luxury	112.6666667	12
Exotic,Luxury,High-Performance	473.025974	77
Exotic,Luxury,High-Performance,Hybrid	204	1
Exotic,Luxury,Performance	217.0277778	36
Exotic,Performance	1391	10
Factory Tuner,High-Performance	1966.442308	104
Factory Tuner,Luxury	617	2
Factory Tuner,Luxury,High-Performance	2133.367442	215
Factory Tuner,Luxury,Performance	1413.419355	31
Factory Tuner,Performance	1774.047619	84
Flex Fuel	2225.71345	855
Flex Fuel,Diesel	5657	16
Flex Fuel,Factory Tuner,Luxury,High-Performance	258	1
Flex Fuel,Hybrid	155	2
Flex Fuel,Luxury	746.5384615	39
Flex Fuel,Luxury,High-Performance	898.3125	32
Flex Fuel,Luxury,Performance	1380.071429	28
Flex Fuel,Performance	1680.471264	87
Flex Fuel,Performance,Hybrid	155	2
Hatchback	1308.65331	574
Hatchback,Diesel	873	14
Hatchback,Factory Tuner,High-Performance	1205.153846	13
Hatchback,Factory Tuner,Luxury,Performance	886.8888889	9
Hatchback,Factory Tuner,Performance	2173.714286	21
Hatchback,Flex Fuel	5657	7
Hatchback,Hybrid	2111.15625	64
Hatchback,Luxury	1323.133333	45
Hatchback,Luxury,Hybrid	454	3
Hatchback,Luxury,Performance	1632.25	36
Hatchback,Performance	1073.661616	198
High-Performance	1823.378788	198
Hybrid	2116.586777	121
Luxury	1079.214896	819
Luxury,High-Performance	1668.017964	334
Luxury,High-Performance,Hybrid	568.8333333	12
Luxury,Hybrid	673.6346154	52
Luxury,Performance	1293.062215	659
Luxury,Performance,Hybrid	2333.181818	11
Performance	1415.209615	520
Performance,Hybrid	155	1
Grand Total	**1512.588393**	**7823**



- **Insights:** Flex Fuel, Diesel, Hatchback, Crossover, Performance are the most popular market category for car models.

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

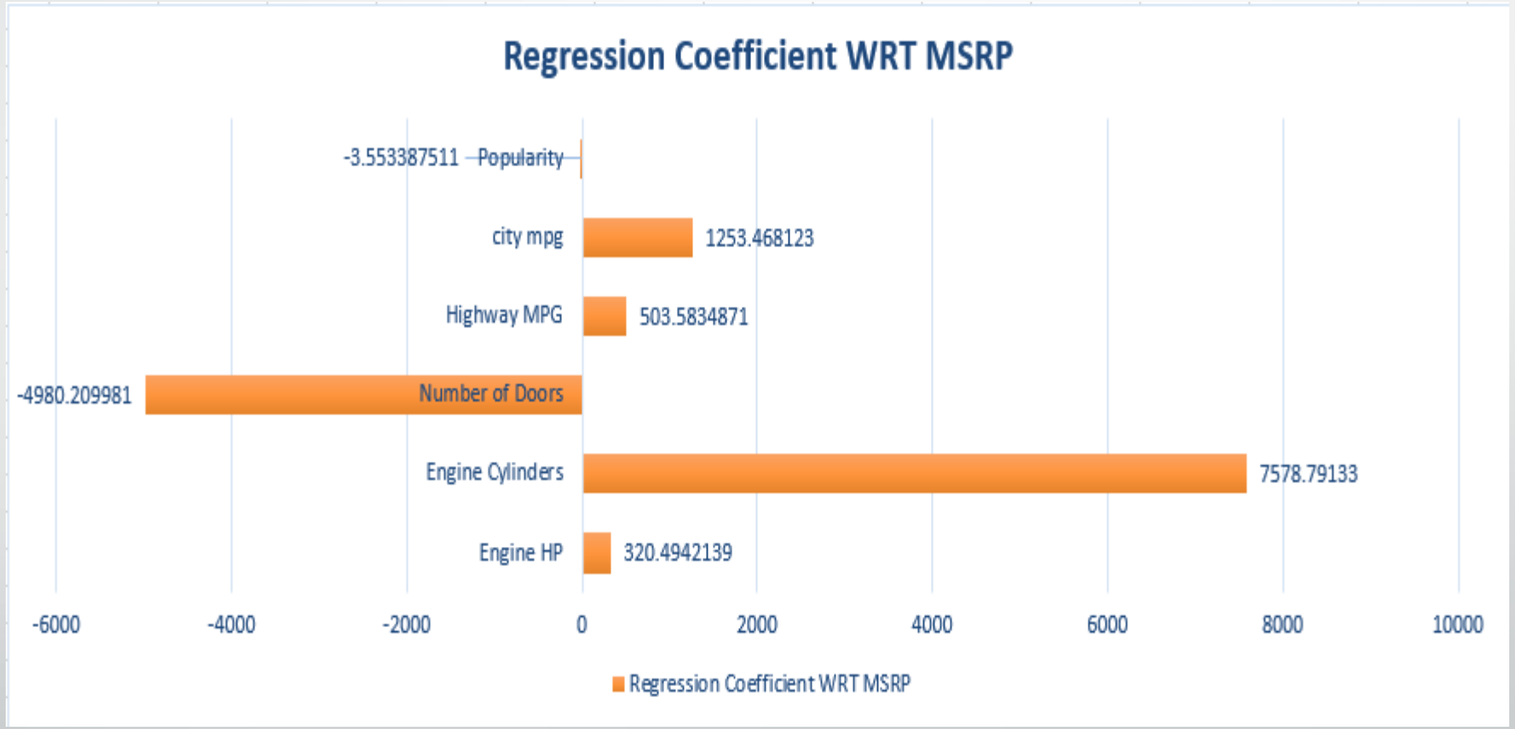
- **Task 2:** Create a scatter chart that plots engine power on the x-axis and price on the y-axis. Add a trendline to the chart to visualize the relationship between these variables.
- **Result:**



- Insights: If Engine power increases Price will also increase. So, it's positive relationship between both of them.

3) 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.
- **Result:**

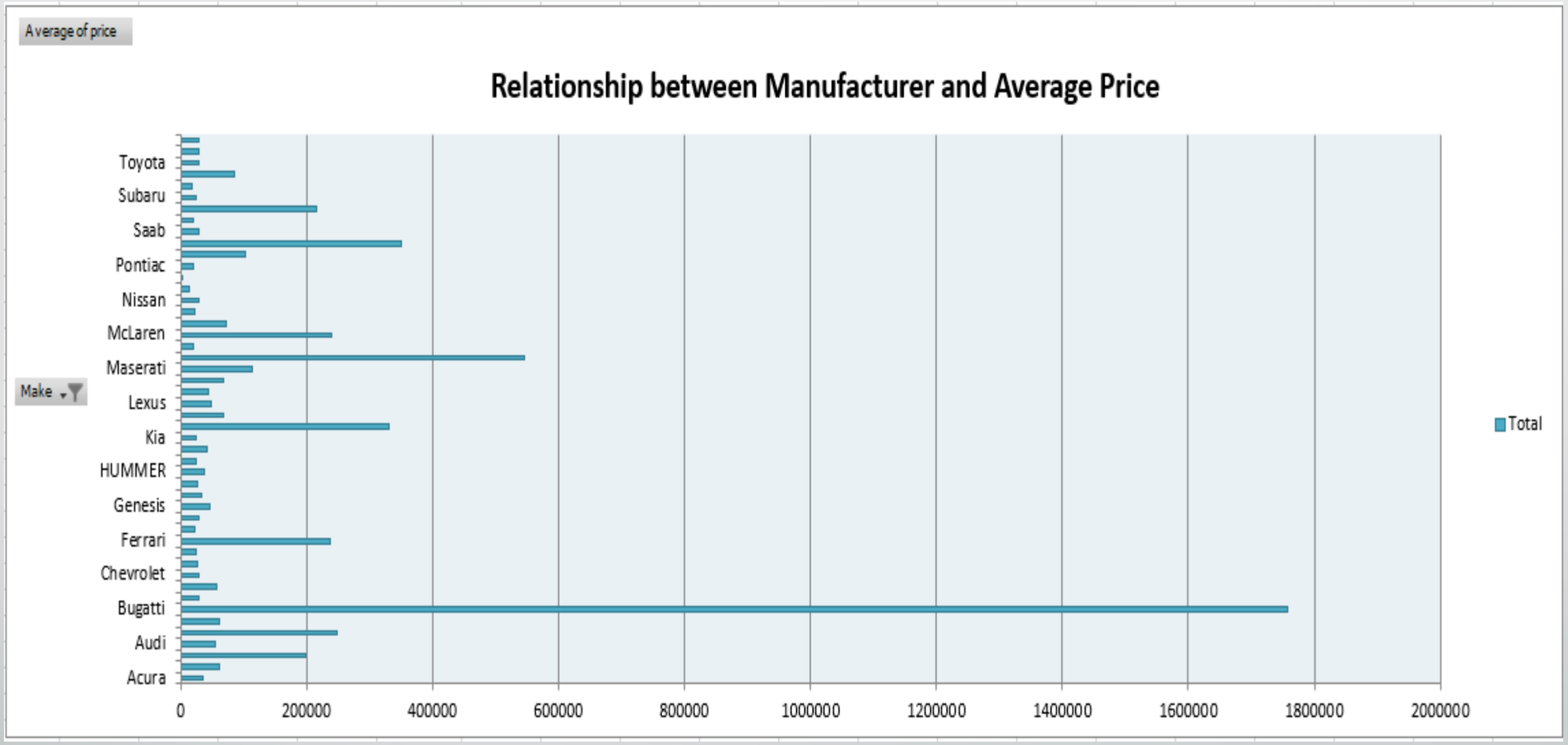


- **Insights:** Engine Cylinders are the most important features in determining a car's price.

4) 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.
- **Result:**

Row Labels	Average of price
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**

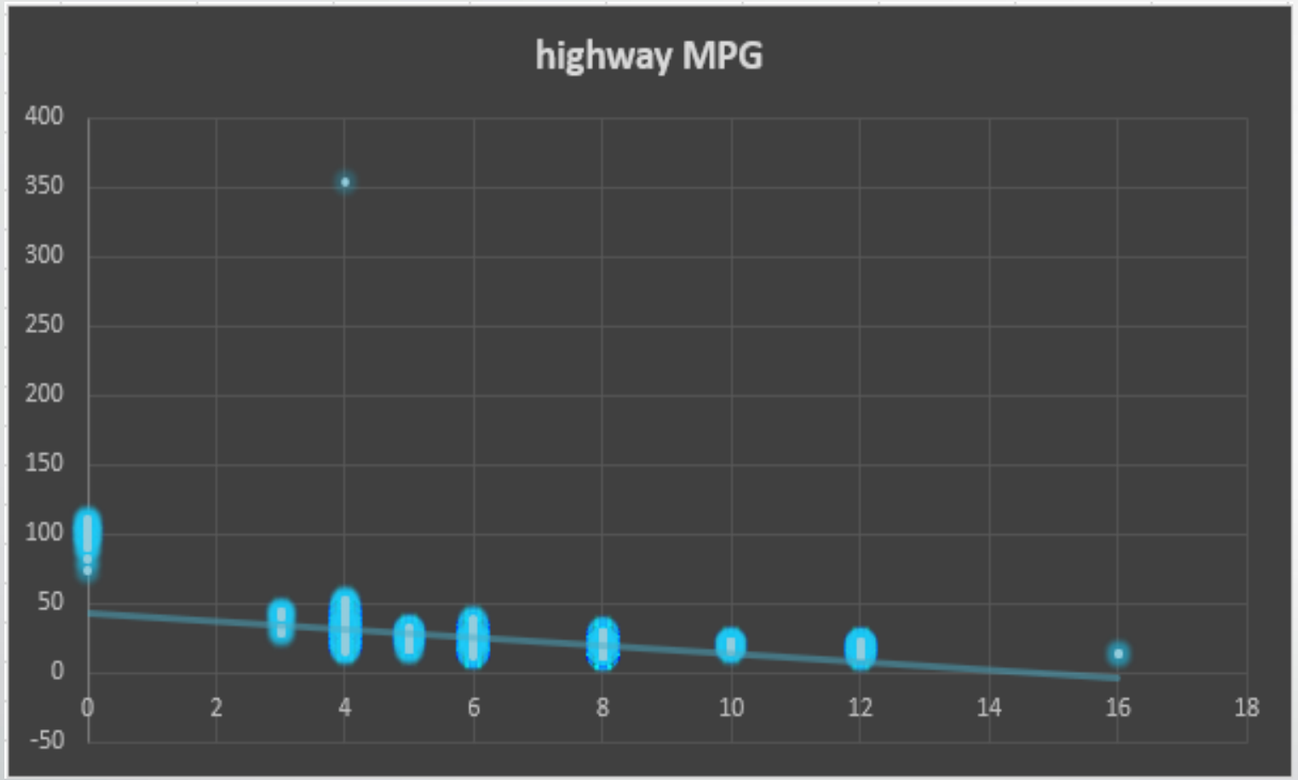


- **Insights:** Bugatti has the highest Average price and Plymouth has the lowest average price.

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

Row Lab	Count of Engine Cylinders
12	5
13	14
14	37
15	117
16	201
17	385
18	364
19	497
20	478
21	494
22	703
23	771
24	843
25	706
26	750
27	567
28	657
29	440
30	521
31	495
32	302
33	335
34	274
35	204
36	196
37	173
38	138
39	109
40	110
41	65
42	46
43	21
44	21
45	14
46	21
47	7
48	16
50	10
53	5
74	2
82	3
90	4
92	5
94	1
97	2
98	1
99	3
100	2
101	11
102	2
103	2
105	4
106	2
107	2
108	1
109	6
110	
111	3
354	1
Grand Tota	11169

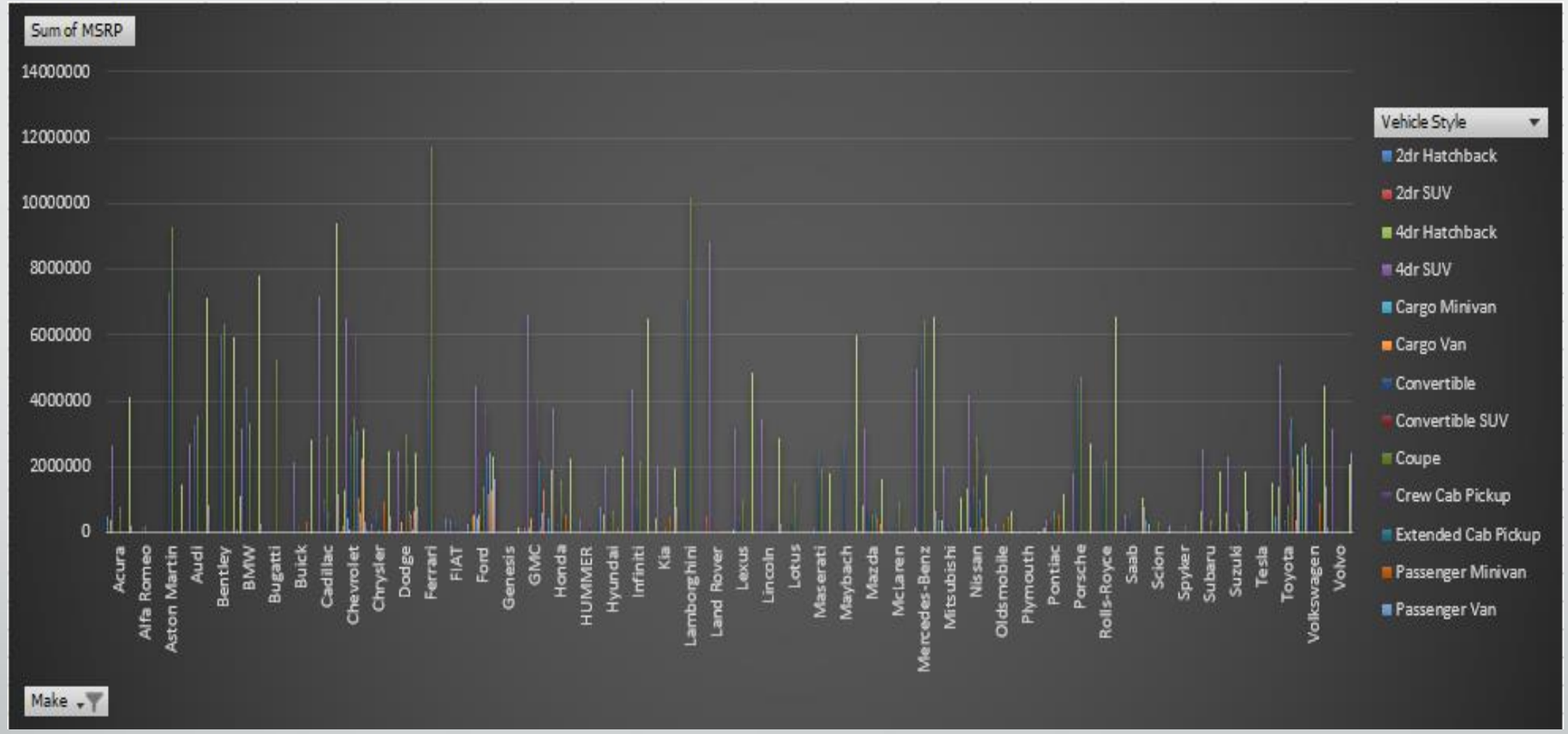


Correlation Coefficient	-0.617419591

- **Insights:** Number of Cylinders will increase then highway MPG will decrease. It's negative relationship between both of them.

Building the Dashboard

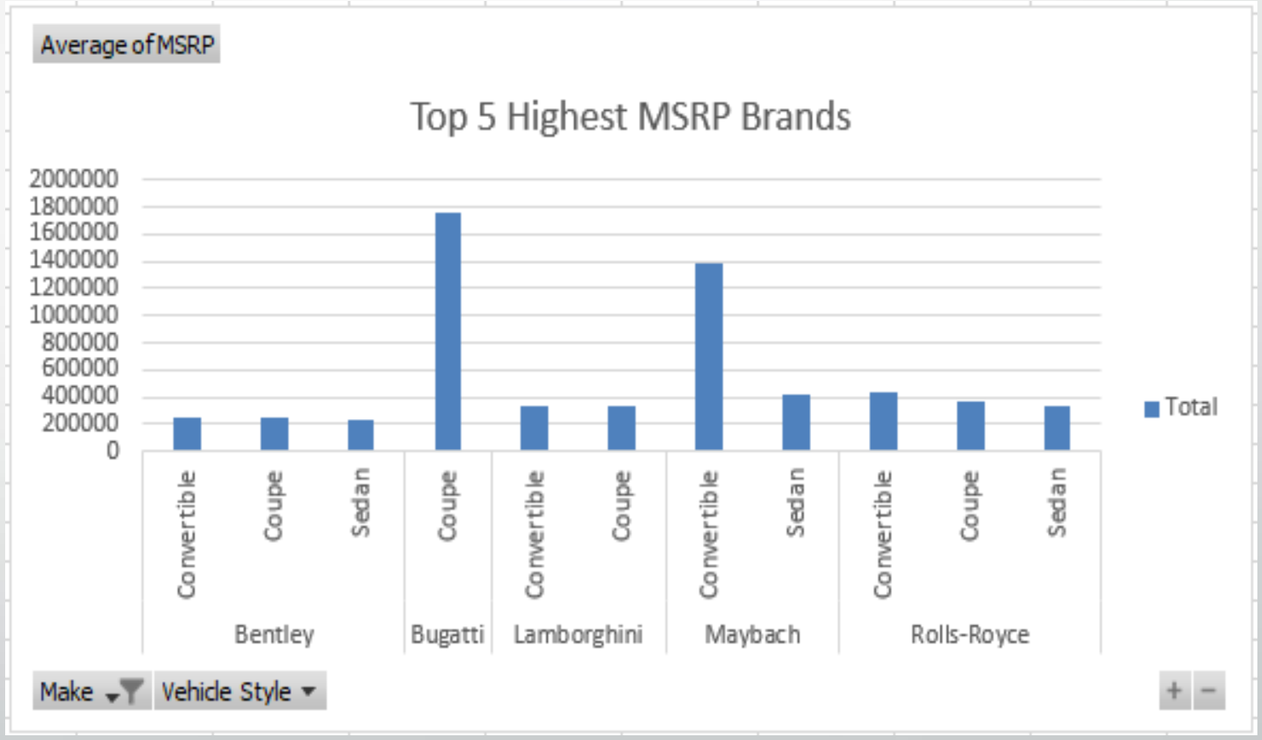
1) How does the distribution of car prices vary by brand and body style?



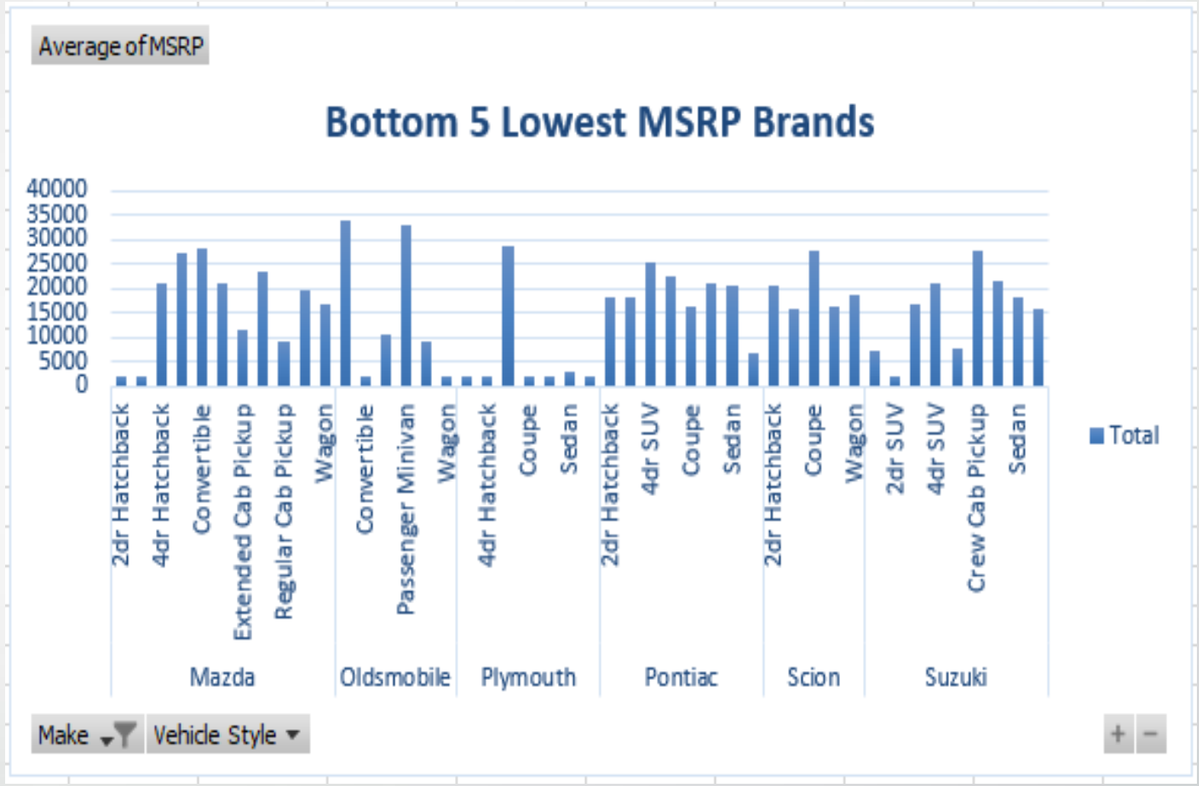
- **Insights:** Chevrolet has the highest price distribution by body style.

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

Row Labels	Average of MSRP
Bentley	**247169.3243**
Convertible	250536.25
Coupe	254270.4
Sedan	236836
Bugatti	**1757223.667**
Coupe	1757223.667
Lamborghini	**331567.3077**
Convertible	336402.381
Coupe	328291.9355
Maybach	**546221.875**
Convertible	1381375
Sedan	426914.2857
Rolls-Royce	**351130.6452**
Convertible	428273
Coupe	367445.8333
Sedan	326950.5
Grand Total	**343342.6193**

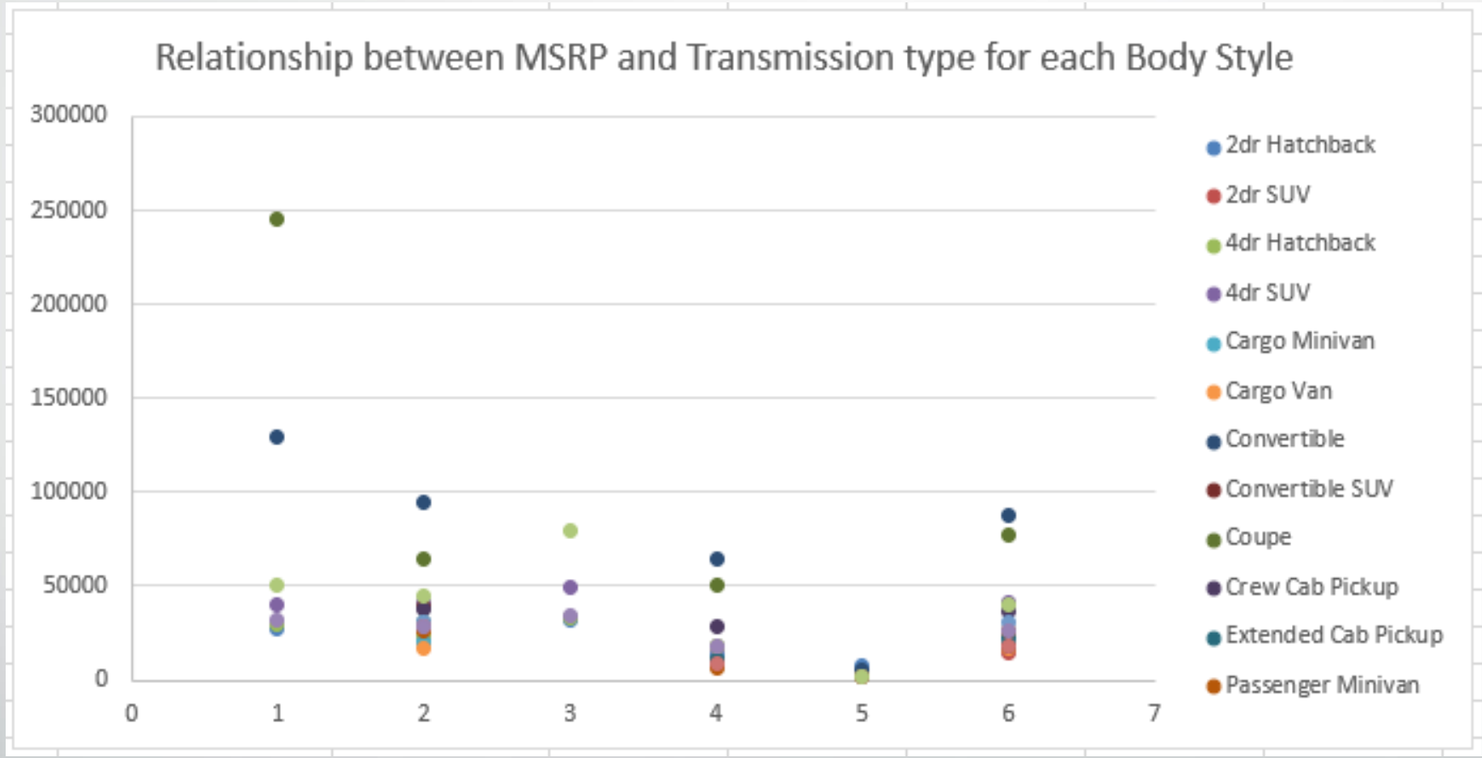


Row Labels	Average of MSRP
Mazda	**20416.62379**
2dr Hatchb	2000
2dr SUV	2000
4dr Hatchb	20809.26829
4dr SUV	27141.15385
Convertible	28080.80645
Coupe	20841.5
Extended C	11600.66
Passenger I	23322.63158
Regular Cat	9154.689655
Sedan	19738.67073
Wagon	16675
Oldsmobile	**12843.79545**
4dr SUV	34021.42857
Convertible	2000
Coupe	10615.96154
Passenger I	32803.66667
Sedan	9139.191781
Wagon	2000
Plymouth	**3296.873239**
2dr Hatchb	2000
4dr Hatchb	2000
Convertible	28543.66667
Coupe	2000
Passenger I	2112.533333
Sedan	2768.5
Wagon	2000
Pontiac	**19800.0442**
2dr Hatchb	18167.22222
4dr Hatchb	18108.33333
4dr SUV	25096.875
Convertible	22546.71429
Coupe	16188.17073
Passenger I	20815.07692
Sedan	20652.41071
Wagon	6951.666667
Scion	**19932.5**
2dr Hatchb	20351.38889
4dr Hatchb	15692.77778
Coupe	27517.5
Sedan	16250
Wagon	18444.5
Suzuki	**18026.4152**
2dr Hatchb	7416
2dr SUV	2000
4dr Hatchb	16696.77143
4dr SUV	21132.9633
Convertible	7512.125
Crew Cab P	27648.27273
Extended C	21638.25
Sedan	17989.97087
Wagon	15538.79545
Grand Total	**17767.86311**



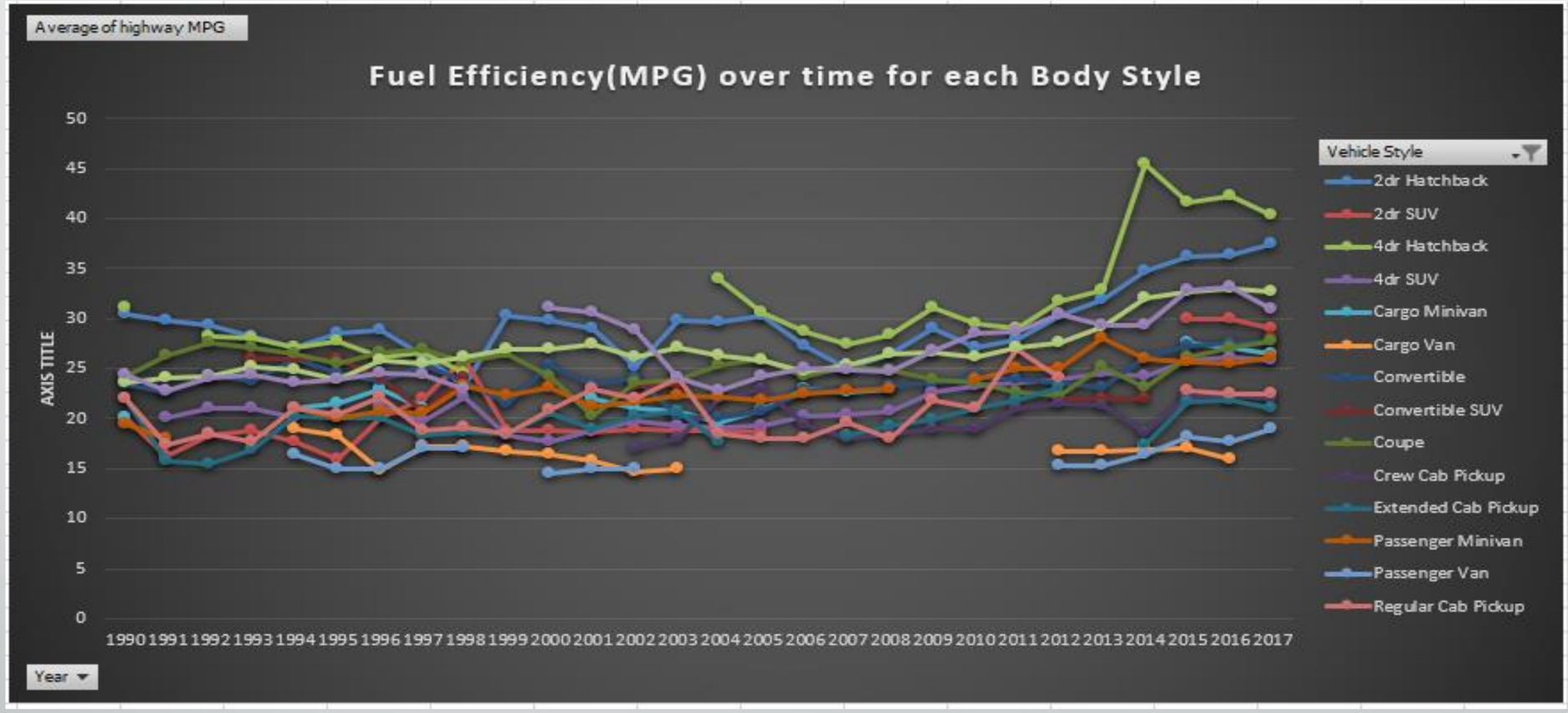
- **Insights:** Bugatti has the highest average MSRPs and Plymouth has the lowest average MSRPs by body style.

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



- **Insights:** AUTOMATED_MANUAL with Coupe body style is the most expensive transmission.

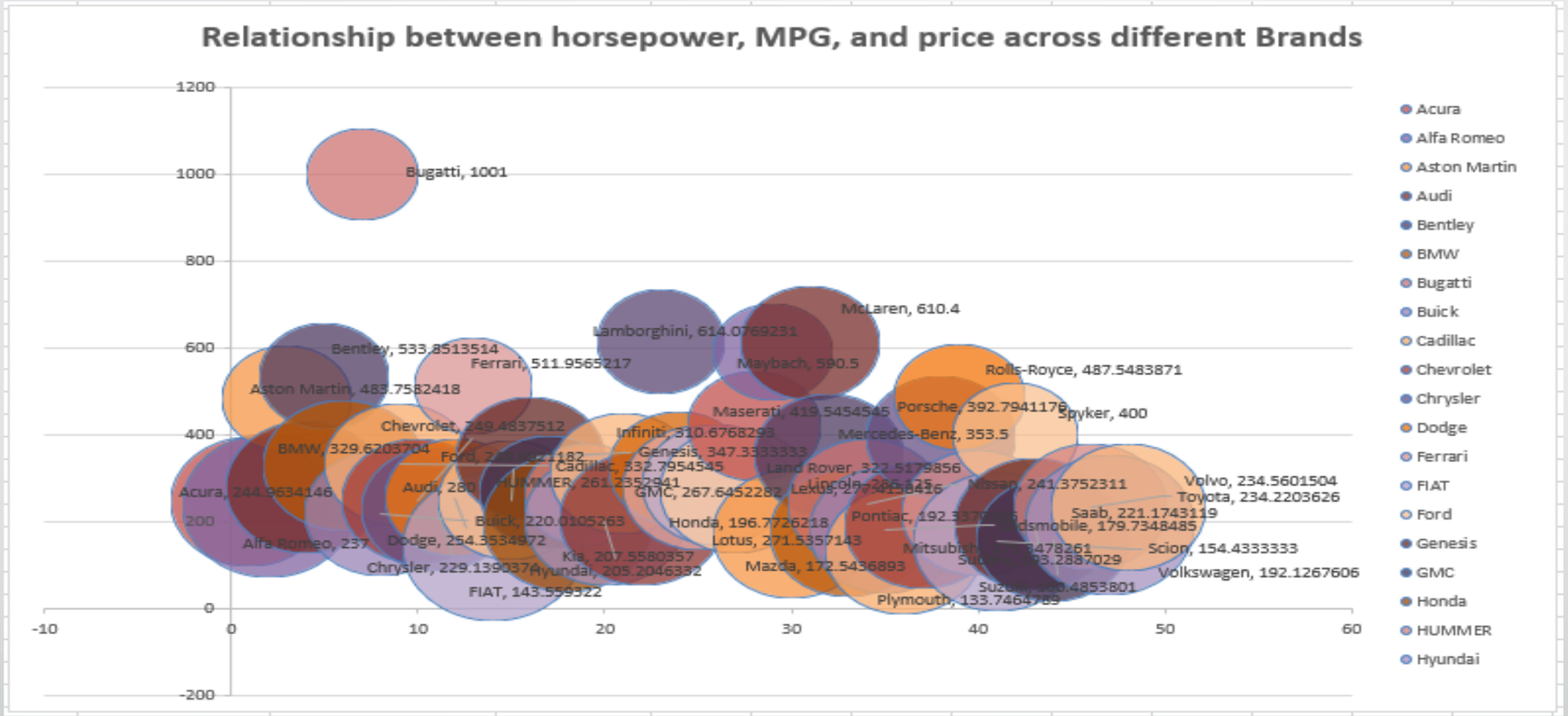
4) How does the fuel efficiency of cars vary across different body styles and model years?



- **Insights**: Fuel efficiency of cars increased across different body styles and model years. Wagon body style has the highest fuel efficiency in 2017.

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

Row Labels	Average of Engine HP	Average of highway MPG	Average of MSRP
Acura	244.9634146	28.2195122	35087.4878
Alfa Romeo	237	34	61600
Aston Martin	483.7582418	18.93406593	198123.4615
Audi	280	28.92834891	54574.1215
Bentley	533.8513514	18.90540541	247169.3243
BMW	329.6203704	29.12654321	62162.55864
Bugatti	1001	14	1757223.667
Buick	220.0105263	27.01052632	29034.18947
Cadillac	332.7954545	25.24494949	56368.26515
Chevrolet	249.4837512	25.92705448	29074.72576
Chrysler	229.1390374	26.36898396	26722.96257
Dodge	254.3534972	22.98676749	24857.04537
Ferrari	511.9565217	15.72463768	238218.8406
FIAT	143.559322	37.33870968	22670.24194
Ford	249.6921182	23.88606061	28511.30788
Genesis	347.3333333	25.33333333	46616.66667
GMC	267.6452282	21.45643154	32444.08506
Honda	196.7726218	32.39953811	26655.14781
HUMMER	261.2352941	17.29411765	36464.41176
Hyundai	205.2046332	29.77220077	24926.26255
Infiniti	310.6768293	24.79573171	42640.27134
Kia	207.5580357	30.68558952	25513.75546
Lamborghini	614.0769231	18.01923077	331567.3077
Land Rover	322.5179856	21.97841727	68067.08633
Lexus	277.4158416	25.87623762	47549.06931
Lincoln	286.125	24.5375	43860.825
Lotus	271.5357143	26.10714286	68377.14286
Maserati	419.5454545	20.16363636	113684.4909
Maybach	590.5	16	546221.875
Mazda	172.5436893	27.93932039	20416.62379
McLaren	610.4	22.2	239805
Mercedes-Benz	353.5	24.56891496	72069.52786
Mitsubishi	173.3478261	27.64423077	21340.5625
Nissan	241.3752311	27.77495463	28921.15245
Oldsmobile	179.7348485	26.18939394	12843.79545
Plymouth	133.7464789	27.4084507	3296.873239
Pontiac	192.3370166	26.96132597	19800.0442
Porsche	392.7941176	25.36764706	101622.3971
Rolls-Royce	487.5483871	19.12903226	351130.6452
Saab	221.1743119	26.37614679	27879.80734
Scion	154.4333333	32.3	19932.5
Spyker	400	18	214990
Subaru	193.2887029	29.20502092	24240.67364
Suzuki	160.4853801	26.00292398	18026.4152
Tesla		98.94444444	85255.55556
Toyota	234.2203626	26.62447844	28846.5605
Volkswagen	192.1267606	32.52992958	28978.52289
Volvo	234.5601504	27.26315789	29724.68421
Grand Total	**253.3888589**	**26.61059023**	**41925.92714**



- **Insights:** If Engine HP goes up, Highway MPG goes down but price increases.

- **Conclusion:** This project helps in handling the large datasets. How exploratory data analysis can be applied to large datasets. When dealing with the large datasets it is also important to select only those columns which are extremely useful to our analysis. Finding correlations columns can become very convenient while dealing with large datasets as it saves time selecting which columns should be considered for analysis.

- **Link for Resultant Dataset:** https://docs.google.com/spreadsheets/d/1TCv-qZbSWC4fzEUgQR2MgK2txNgZ3hNV/edit?usp=sharing&ouid=101431809048624548912&rtpof=true&sd=true

Thank You