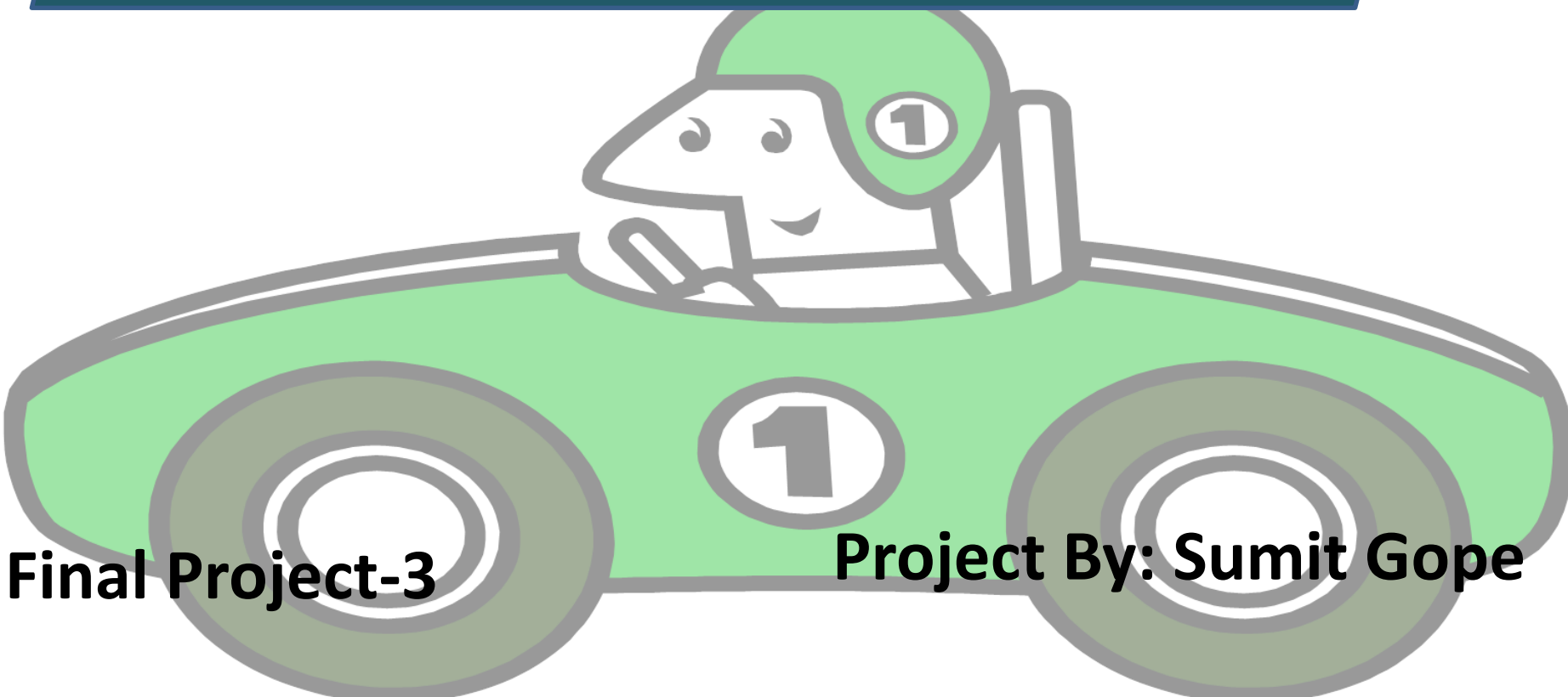
Impact Of Car Features



Final Project-3

Project By: Sumit Gope

Project Description

The purpose of this project is to help a car manufacturer optimize pricing and product development decisions to maximize profitability while meeting consumer demand. This will be done by analyzing a dataset containing information on various car models and their specifications.

The business problem- The project aims to address is how a car manufacturer can effectively balance consumer demand and profitability when making pricing and product development decisions.

The data used in this project is the "Car Features and MSRP" dataset available on Kaggle, which contains information on over 11,000 car models and their specifications.

Data preprocessing: Cleaning and handling missing data.

Approach

The analytical methods used in this project include visualization, and machine learning techniques such as regression analysis and market segmentation.

Regression models were developed to predict car prices based on their features and market category. Market segmentation was used to identify consumer groups with similar preferences and buying behaviors.

One limitation of this project is that the dataset was last updated in 2017, so it may not reflect current trends or prices in the automotive industry.

Tech-Stack Used

So I used Excel 2010 to gain insights from the dataset and for Data visualization. After that I used PowerPoint for presentation of the project.

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.

Solution- for this basically selected pivot table and then drag market category column into row label and drag model and popularity score column into values field. And then Calculated the count of model column and average of popularity column.

	Market Category	Count of Model	Popularity Score
2	Flex Fuel,Diesel	16	5657
3	Hatchback,Flex Fuel	7	5657
4	Crossover,Flex Fuel,Performance	6	5657
5	Crossover,Luxury,Performance,Hybrid	2	3916
6	Crossover,Factory Tuner,Luxury,Performance	5	2607
7	Crossover,Performance	69	2586
8	Crossover,Hybrid	42	2563
9	Diesel,Luxury	47	2416
10	Luxury,Performance,Hybrid	11	2333
11	Flex Fuel	855	2226
12	Hatchback,Factory Tuner,Performance	21	2174
13	Crossover,Luxury,Diesel	34	2149
14	Factory Tuner,Luxury,High-Performance	215	2133
15	Hybrid	121	2117
16	Hatchback,Hybrid	64	2111
17	Crossover,Flex Fuel	64	2074
18	Crossover,Hatchback,Factory Tuner,Performance	6	2009
19	Crossover,Hatchback,Performance	6	2009
20	Factory Tuner,High-Performance	104	1966
21	Crossover,Factory Tuner,Luxury,High-Performance	26	1823
22	High-Performance	198	1823
23	Factory Tuner,Performance	84	1774
24	Diesel	84	1731
25	Flex Fuel,Performance	87	1680

26	Crossover,Hatchback	72	1676
27	Luxury,High-Performance	334	1668
28	N/A	3376	1665
29	Hatchback,Luxury,Performance	36	1632
30	Crossover,Flex Fuel,Luxury,Performance	6	1624
31	Crossover	1075	1556
32	Performance	520	1415
33	Factory Tuner,Luxury,Performance	31	1413
34	Exotic,Performance	10	1391
35	Flex Fuel,Luxury,Performance	28	1380
36	Crossover,Luxury,Performance	112	1349
37	Hatchback,Luxury	45	1323
38	Hatchback	574	1309
39	Luxury,Performance	659	1293
40	Exotic,High-Performance	254	1280
41	Hatchback,Factory Tuner,High-Performance	13	1205
42	Crossover,Flex Fuel,Luxury	10	1173
43	Luxury	819	1079
44	Hatchback,Performance	198	1074
45	Exotic,Factory Tuner,High-Performance	21	1046
46	Crossover,Luxury,High-Performance	9	1037
47	Flex Fuel,Luxury,High-Performance	32	898
48	Crossover,Luxury	406	889
49	Hatchback,Factory Tuner,Luxury,Performance	9	887
50	Crossover,Diesel	7	873

51	Hatchback,Diesel	14	873
52	Flex Fuel,Luxury	39	747
53	Luxury,Hybrid	52	674
54	Crossover,Luxury,Hybrid	24	631
55	Factory Tuner,Luxury	2	617
56	Luxury,High-Performance,Hybrid	12	569
57	Exotic,Factory Tuner,Luxury,High-Performance	51	523
58	Exotic,Flex Fuel,Luxury,High-Performance	11	520
59	Exotic,Factory Tuner,Luxury,Performance	3	520
60	Exotic,Flex Fuel,Factory Tuner,Luxury,High-Performance	13	520
61	Exotic,Luxury,High-Performance	77	473
62	Hatchback,Luxury,Hybrid	3	454
63	Flex Fuel,Factory Tuner,Luxury,High-Performance	1	258
64	Crossover,Exotic,Luxury,High-Performance	1	238
65	Crossover,Exotic,Luxury,Performance	1	238
66	Exotic,Luxury,Performance	36	217
67	Crossover,Factory Tuner,Performance	4	210
68	Crossover,Hatchback,Luxury	7	204
69	Exotic,Luxury,High-Performance,Hybrid	1	204
70	Performance,Hybrid	1	155
71	Flex Fuel,Performance,Hybrid	2	155
72	Flex Fuel,Hybrid	2	155
73	Exotic,Luxury	12	113

From the above output we can see that market category “crossover” have highest number of car models and popularity score is also decent. It means in this category competition is high.
If we talk about highest Popularity score hatchback,flex fuel have that spot on the top also having same score in ‘flex fuel,diesel’ and ‘crossover,flex,fuel,performance’ .

Apart from this market category ‘crossover,luxury’ having good number of models but low popularity score. Needs some improvement in this category.

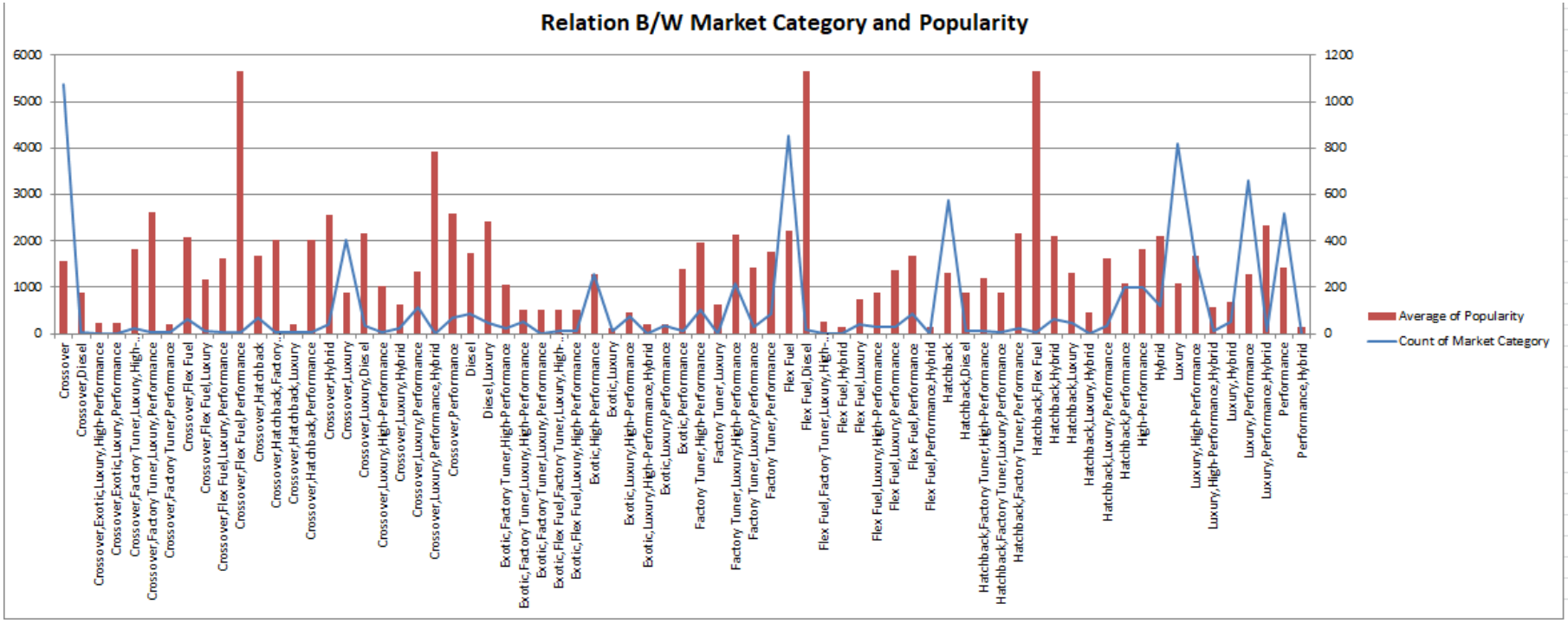
Market category ‘crossover,luxury,performance,hybrid’ is in top 4 in popularity score but have only 2 number of model. Which means companies should increase the number of models in this category.

Market Category	Count of Model	Popularity Score
Crossover	1075	1556
Flex Fuel	855	2226
Luxury	819	1079
Luxury,Performance	659	1293
Hatchback	574	1309
Performance	520	1415
Crossover,Luxury	406	889
Luxury,High-Performance	334	1668
Exotic,High-Performance	254	1280
Factory Tuner,Luxury,High-Performance	215	2133

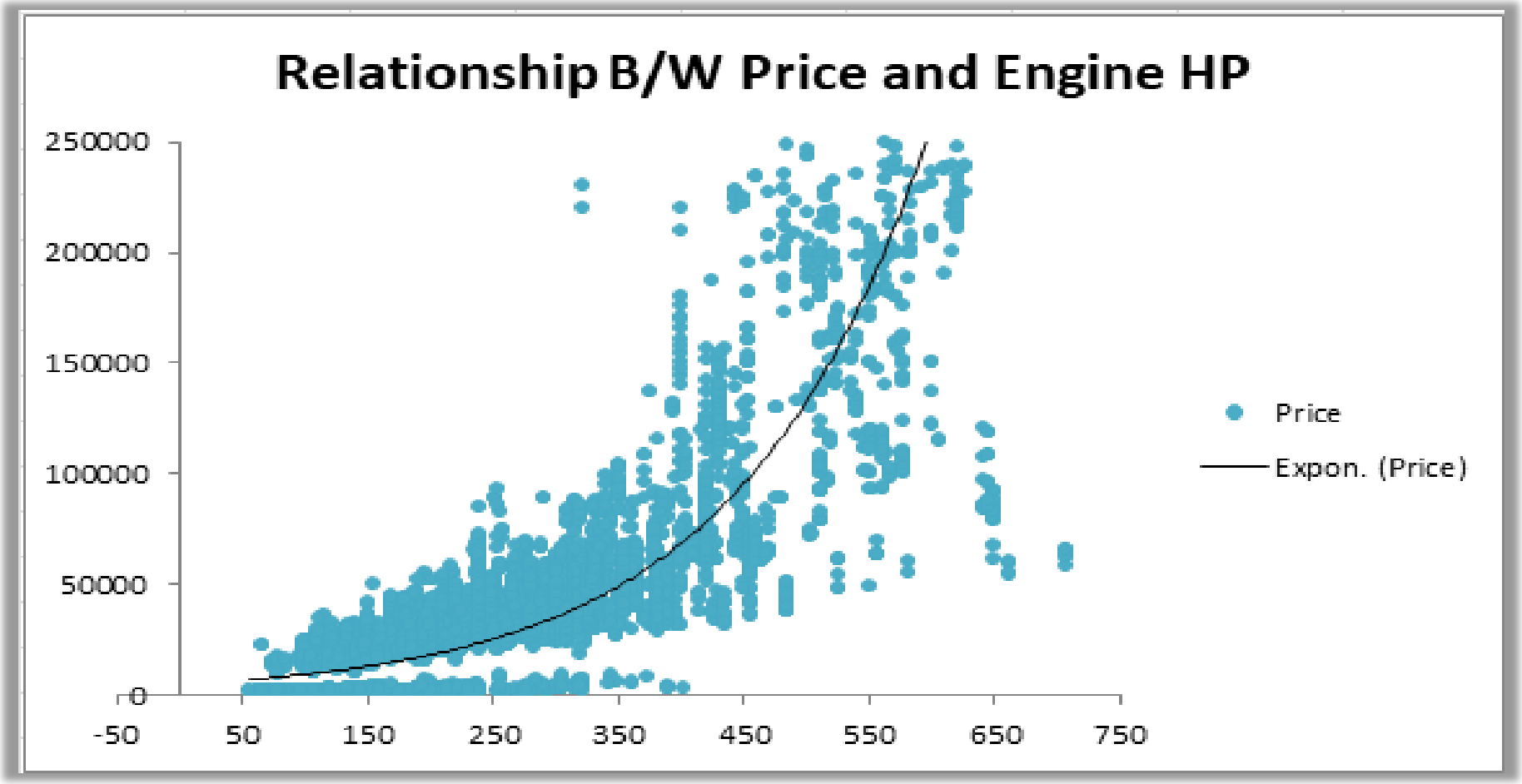
1	Market Category	Count of Model	Popularity Score
2	Hatchback,Flex Fuel	7	5657
3	Flex Fuel,Diesel	16	5657
4	Crossover,Flex Fuel,Performance	6	5657
5	Crossover,Luxury,Performance,Hybrid	2	3916
6	Crossover,Factory Tuner,Luxury,Performance	5	2607
7	Crossover,Performance	69	2586
8	Crossover,Hybrid	42	2563
9	Diesel,Luxury	47	2416
10	Luxury,Performance,Hybrid	11	2333
11	Flex Fuel	855	2226

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

Solution- for this we simply use combo chart form the chart option.



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.



As we can see this the positive correlation we get from the Car price and Engine HP. It means increases in Engine horse power increases Cars Price.

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.

Solution: for this I select the all numeric feature of cars and click on data analysis field from data section and then click on regression.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.678234739
R Square	0.460002361
Adjusted R Square	0.45971287
Standard Error	45230.86088
Observations	11199

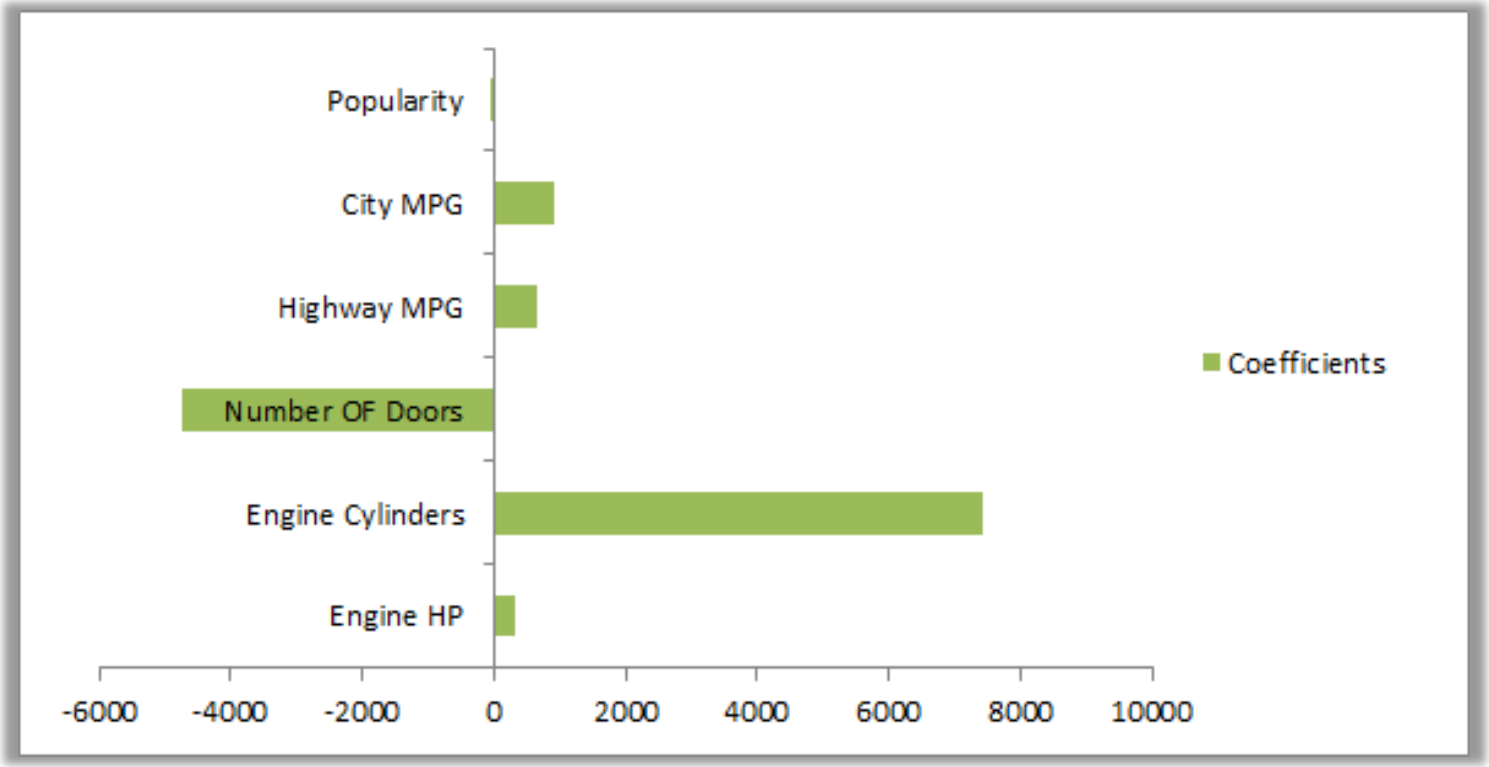
ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	6	1.9505E+13	3.25083E+12	1589.002757	0
Residual	11192	2.28969E+13	2045830776		
Total	11198	4.24019E+13			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-92250.68814	3550.862977	-25.97979385	1.532E-144	-99211.00442	-85290.37187	-99211.00442	-85290.37187
Engine HP	313.0702204	6.209763389	50.41580504	0	300.8979914	325.2424493	300.8979914	325.2424493
Engine Cylinders	7417.153571	435.3624475	17.03673253	2.82475E-64	6563.766563	8270.540578	6563.766563	8270.540578
Number OF Doors	-4741.393993	496.7198692	-9.545408362	1.63691E-21	-5715.052344	-3767.735643	-5715.052344	-3767.735643
Highway MPG	642.3041692	106.7539063	6.016680715	1.83599E-09	433.0477275	851.5606108	433.0477275	851.5606108
City MPG	914.327132	101.6463195	8.995181887	2.73768E-19	715.0824593	1113.571805	715.0824593	1113.571805
Popularity	-3.420848029	0.296575756	-11.53448304	1.31661E-30	-4.002188699	-2.839507359	-4.002188699	-2.839507359

The coefficient for number of cylinders is positive, indicating that there is a strong positive relationship between the number of cylinders in a car's engine and its price. This suggests that cars with more cylinders tend to be more expensive.

The coefficient for number of doors is negative, indicating that there is a strong negative relationship between the number of doors a car has and its price. This suggests that cars with fewer doors tend to be more expensive.



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.

From this table we can see the average highest price of manufacturer is 1.7M$ which is Bugatti and average lowest price of manufacturer is 3296$ which is plymouth.

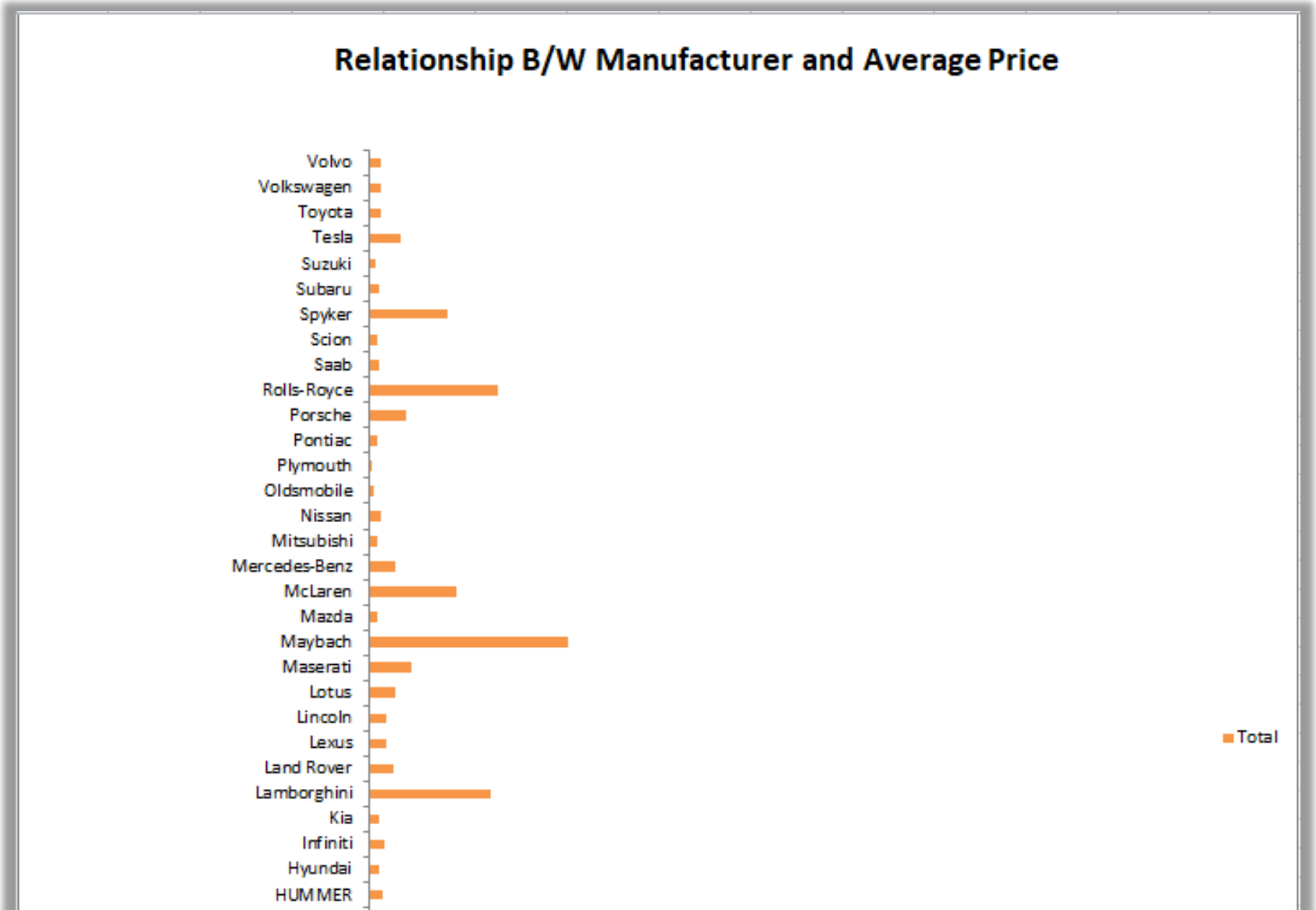
Highest

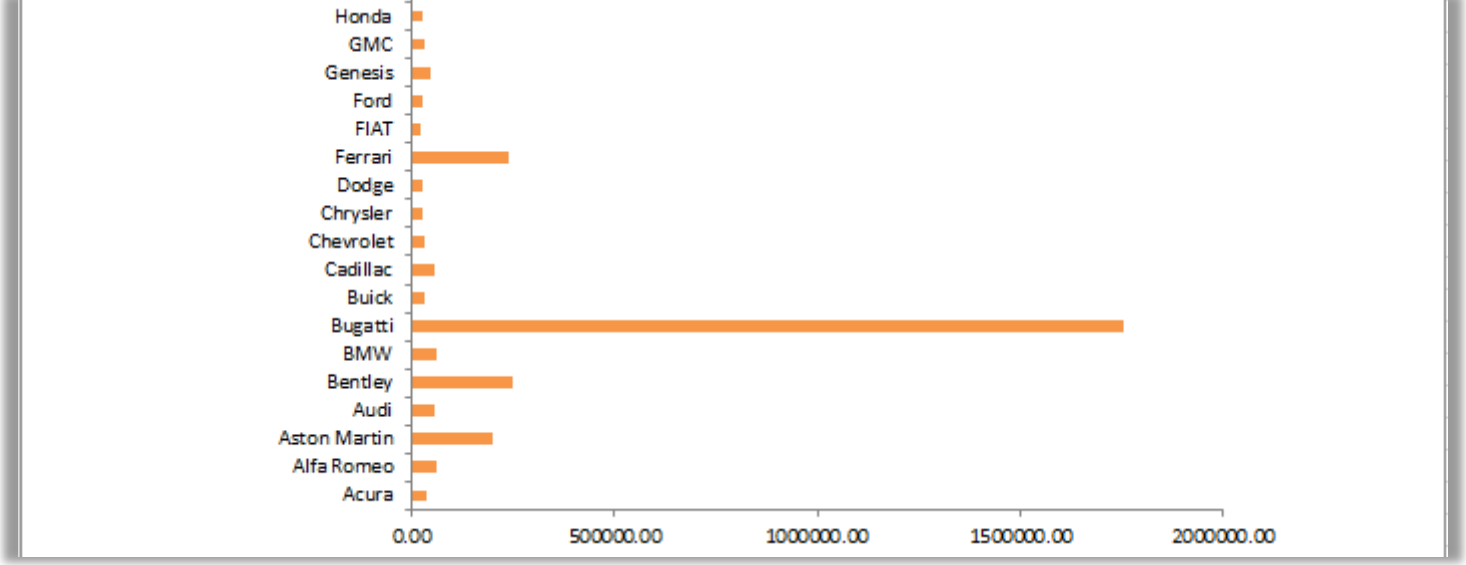
Manufacturer	Average Price
Acura	35087.49
Alfa Romeo	61600.00
Aston Martin	198123.46
Audi	54574.12
Bentley	247169.32
BMW	62162.56
Bugatti	1757223.67
Buick	29034.19
Cadillac	56368.27
Chevrolet	29074.73
Chrysler	26722.96
Dodge	24857.05
Ferrari	238218.84
FIAT	22670.24
Ford	28511.31
Genesis	46616.67
GMC	32444.09
Honda	26655.15
HUMMER	36464.41
Hyundai	24926.26
Infiniti	42640.27
Kia	25513.76
Lamborghini	331567.31
Land Rover	68067.09
Lexus	47549.07
Lincoln	43860.83
Lotus	68377.14
Maserati	113684.49
Maybach	546221.88
Mazda	20416.62

Lowest

32	McLaren	239805.00
33	Mercedes-Benz	72069.53
34	Mitsubishi	21340.56
35	Nissan	28921.15
36	Oldsmobile	12843.80
37	Plymouth	3296.87
38	Pontiac	19800.04
39	Porsche	101622.40
40	Rolls-Royce	351130.65
41	Saab	27879.81
42	Scion	19932.50
43	Spyker	214990.00
44	Subaru	24240.67
45	Suzuki	18026.42
46	Tesla	85255.56
47	Toyota	28846.56
48	Volkswagen	28978.52
49	Volvo	29724.68

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



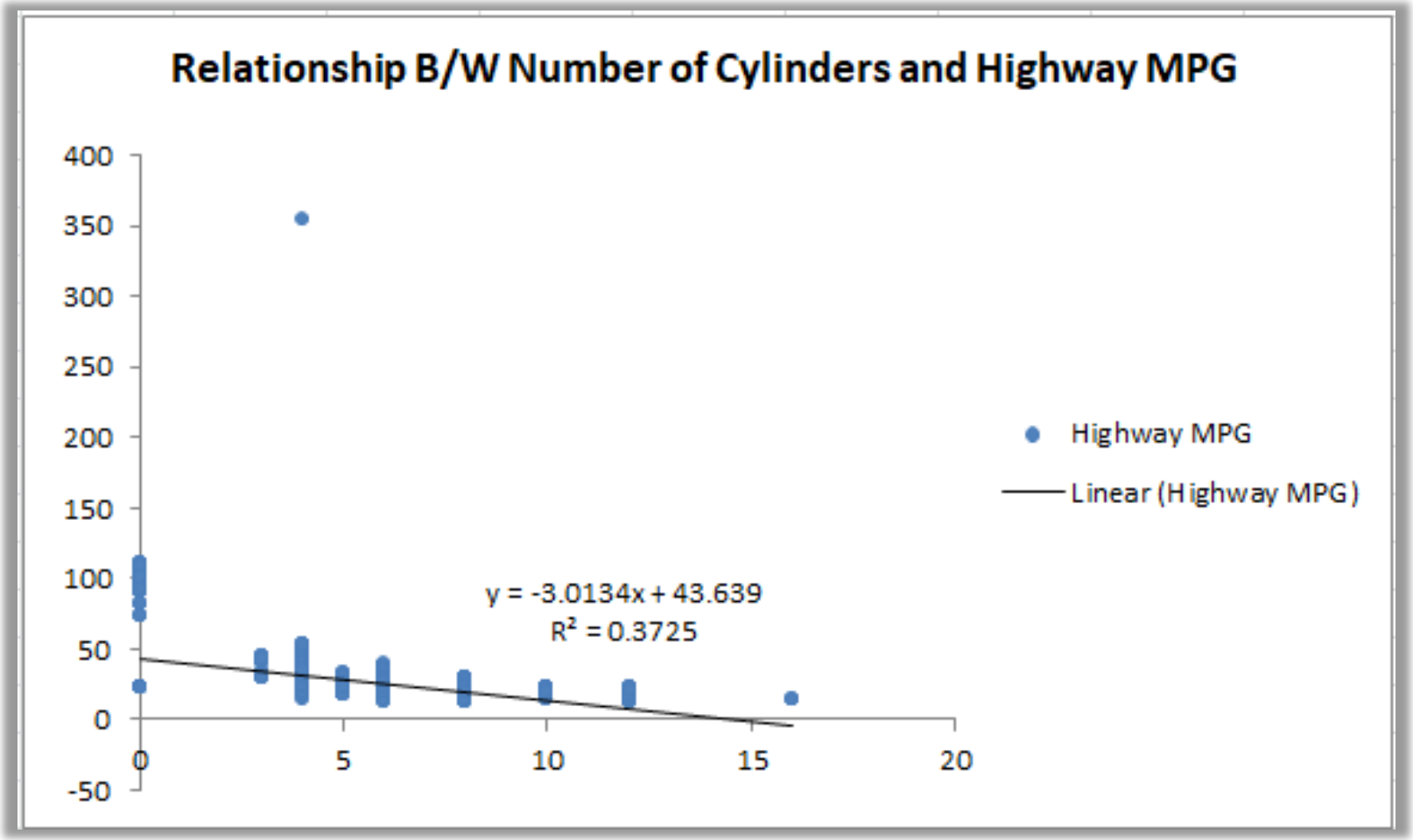


Plymouth is a brand that is associated with more affordable or budget-friendly cars, while Bugatti is a luxury brand that produces high-end, expensive vehicles. This suggests that car prices can vary widely across different manufacturers, depending on the brand's positioning in the market.

The range between the lowest and highest average prices is quite large, indicating that car prices can vary significantly even within the same market segment. This can be due to a variety of factors such as differences in manufacturing costs, brand reputation, features, and technology.

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 slope of the linear equation is negative, which indicates an inverse relationship between the x and y variables. In other words, as the value of x increases, the value of y decreases, and vice versa.

The magnitude of the slope (-3.0134) indicates the strength of the relationship between the two variables. A larger magnitude suggests a stronger relationship between the variables.

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

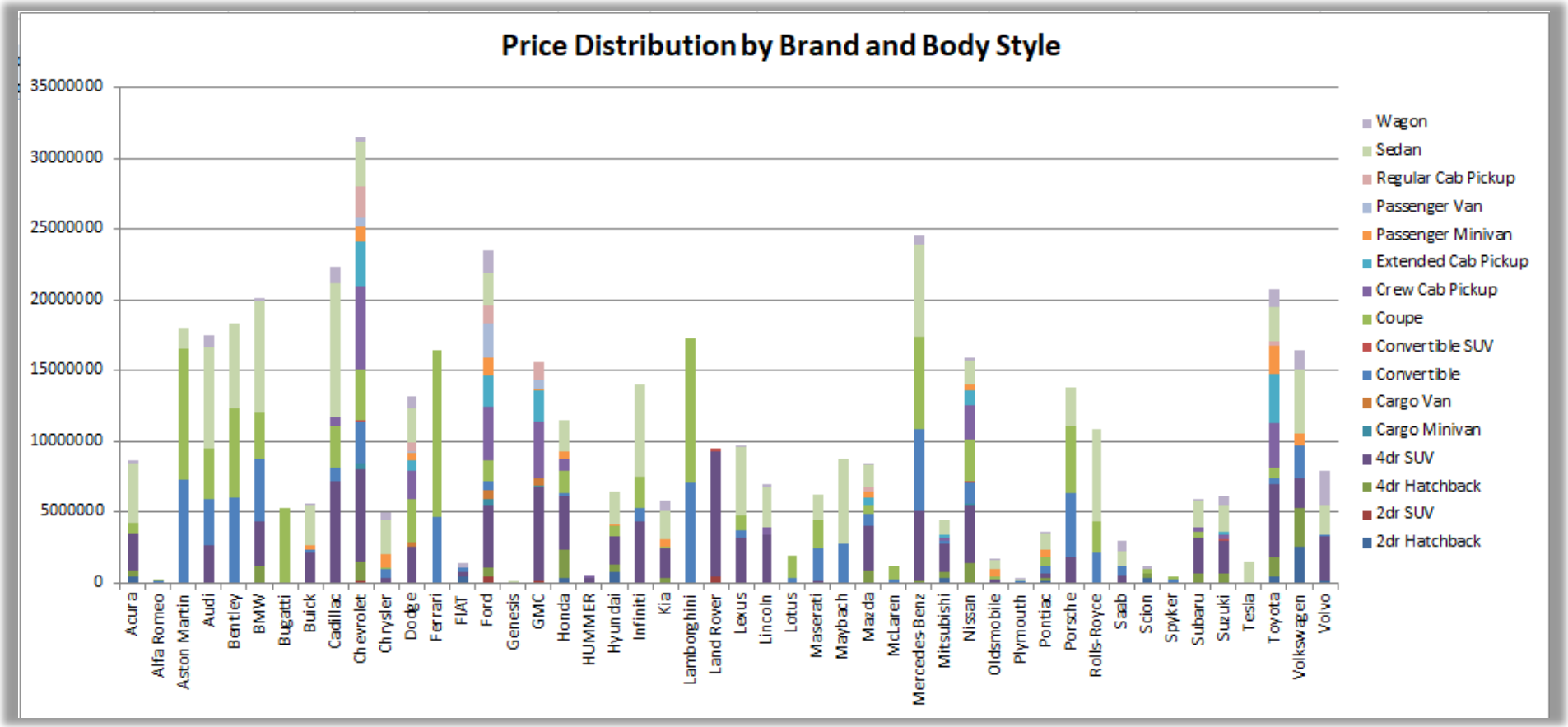
	Number of Cylinders	*Highway MPG*
Number of Cylinders	1	
Highway MPG	-0.610337793	1

A correlation coefficient value of -0.61033 indicates a moderate negative correlation between the number of cylinders and highway MPG. This means that as the number of cylinders in a car's engine increases, its highway MPG generally decreases, and vice versa.

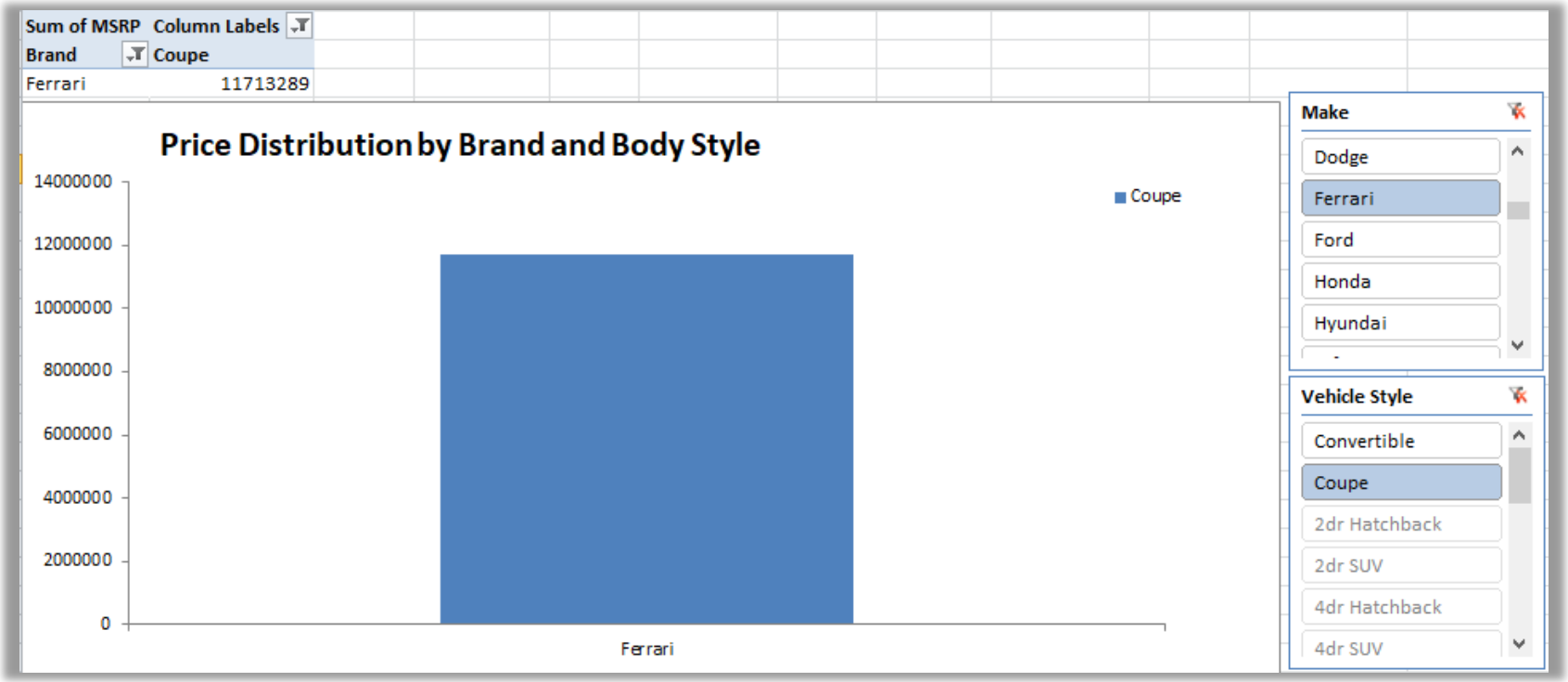
The absolute value of -0.61033 is closer to 1 than to 0, indicating a moderate strength of correlation between the two variables. Therefore, we can conclude that there is a significant relationship between the number of cylinders and highway MPG.

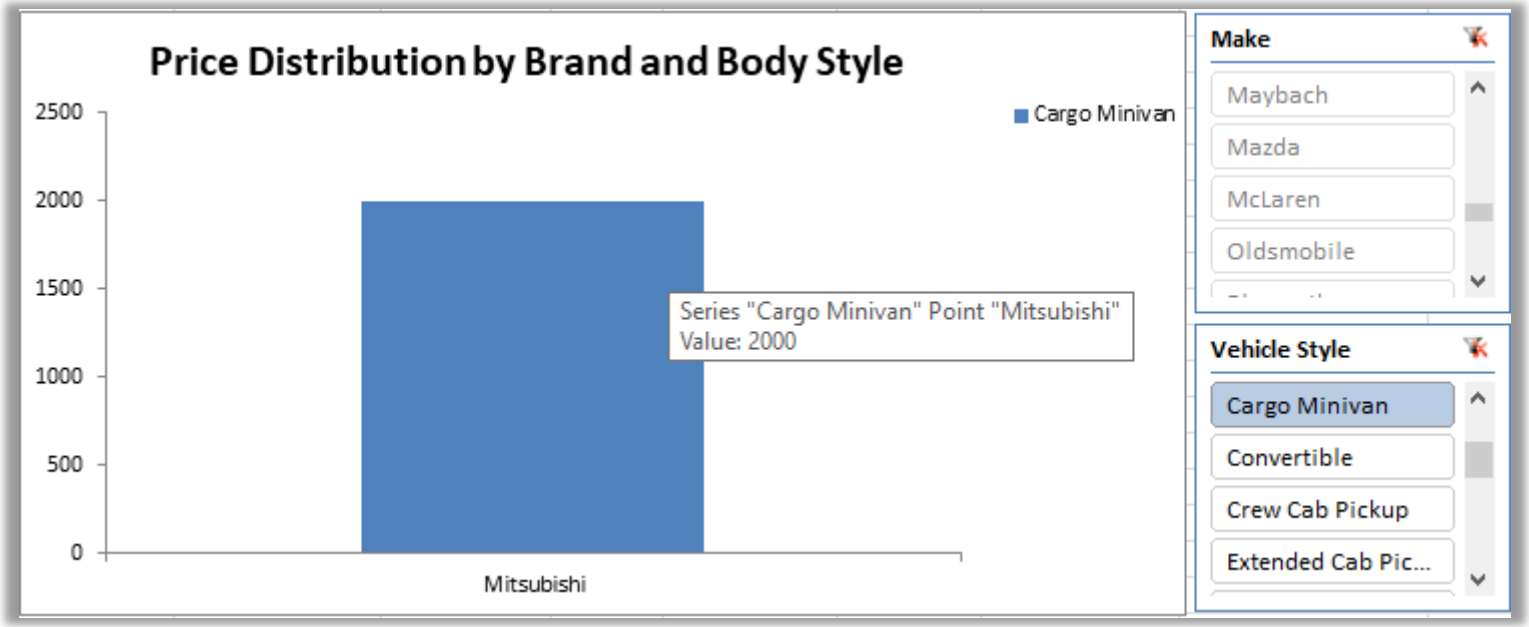
Dashboard Tasks

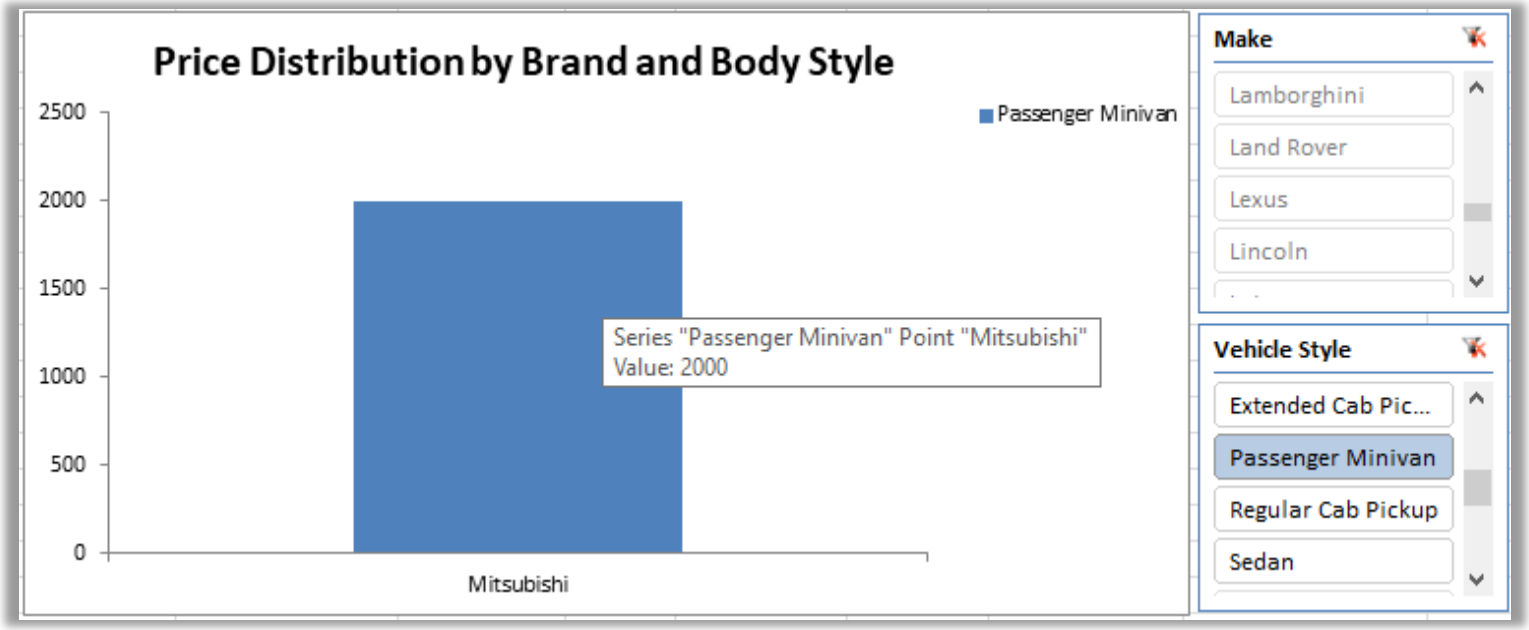
Task 1: How does the distribution of car prices vary by brand and body style?



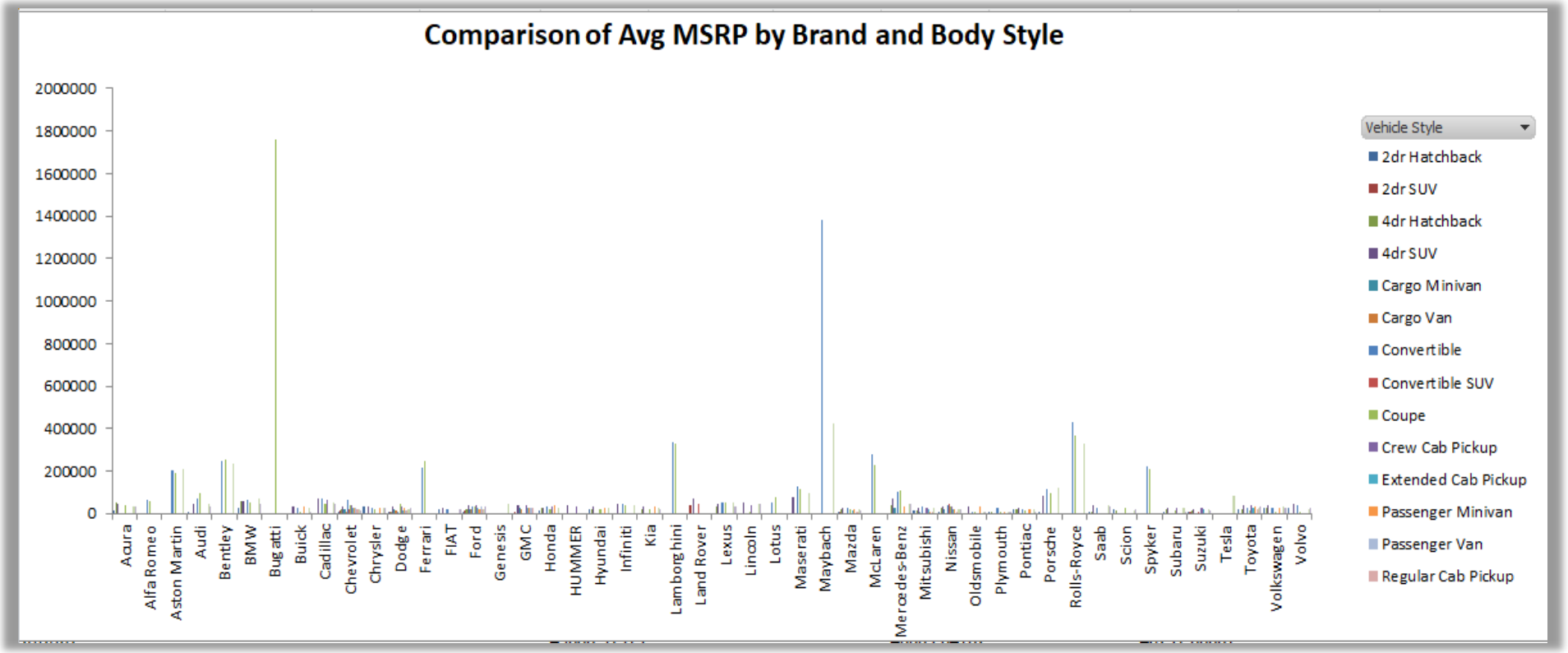
The fact that the brand Ferrari with the body style coupe has the highest price and the brand Mitsubishi with the body styles cargo minivan and passenger minivan has the lowest price indicates that there is a wide range of prices across different brands and body styles.




Price Distribution by Brand and Body Style
2500
2000
1500
1000
500
0
Cargo Minivan
Series "Cargo Minivan" Point "Mitsubishi"
Value: 2000
Mitsubishi
Make
Maybach
Mazda
McLaren
Oldsmobile
Vehicle Style
Cargo Minivan
Convertible
Crew Cab Pickup
Extended Cab Pic...


Price Distribution by Brand and Body Style
2500
2000
1500
1000
500
0
Passenger Minivan
Series "Passenger Minivan" Point "Mitsubishi"
Value: 2000
Mitsubishi
Make
Lamborghini
Land Rover
Lexus
Lincoln
Vehicle Style
Extended Cab Pic...
Passenger Minivan
Regular Cab Pickup
Sedan

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



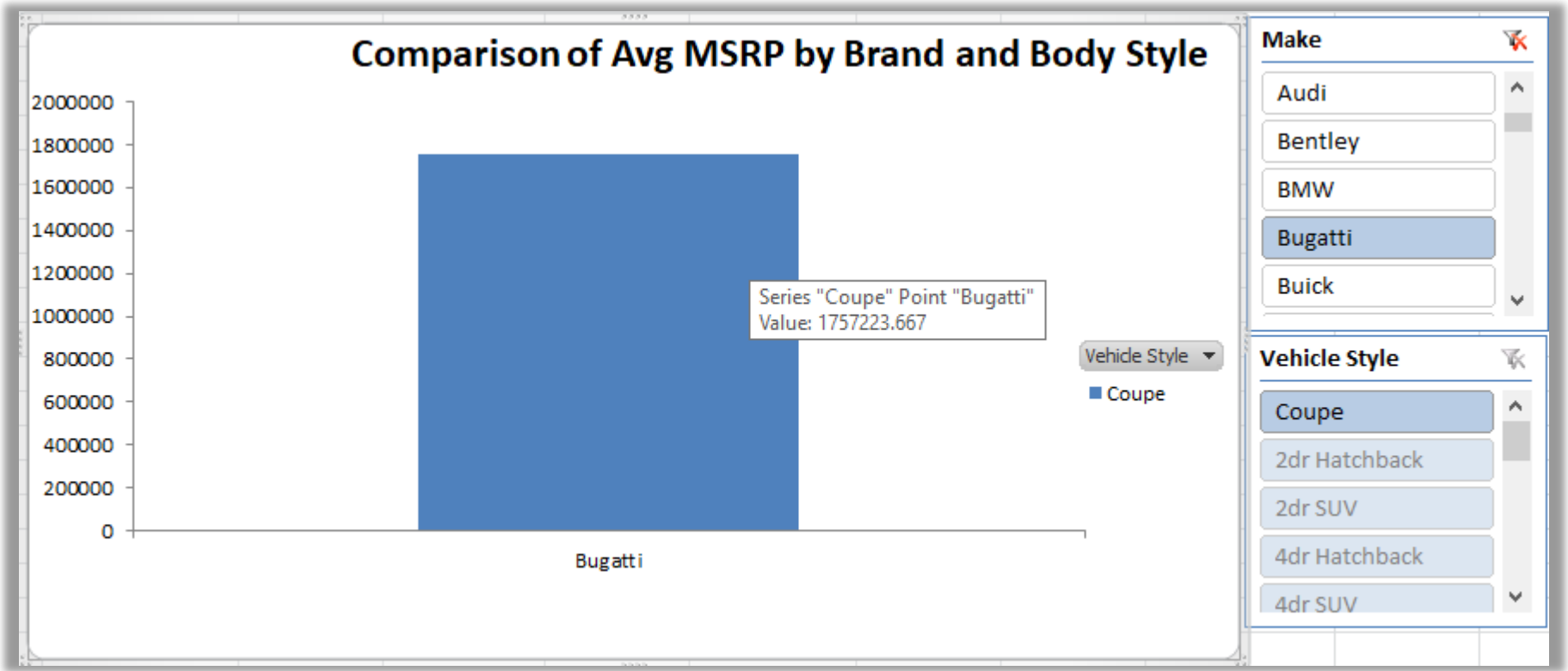
So the Bugatti have highest average MSRP which has only one body style which is Coupe.

Highest MSRP

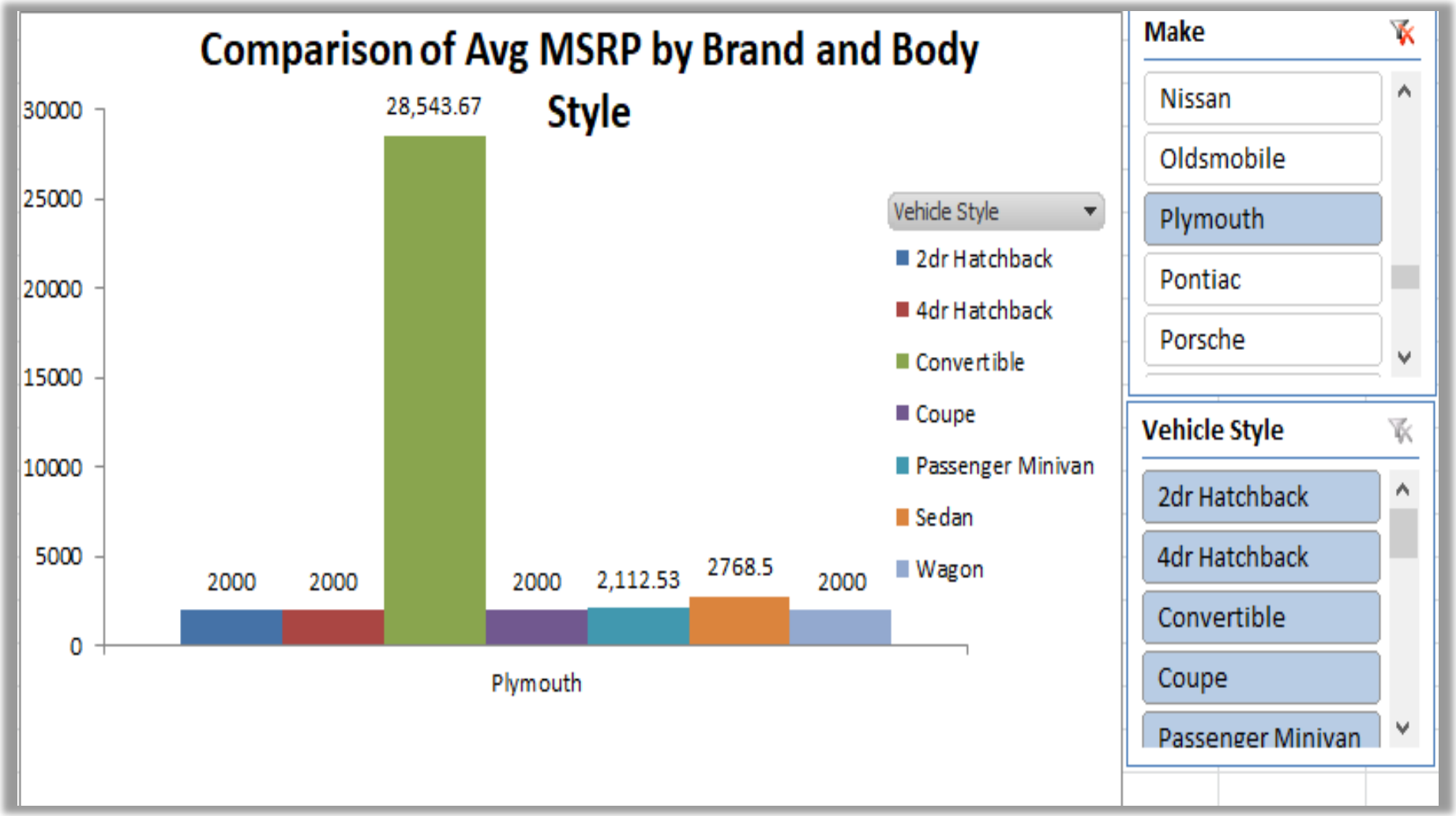
Brand	Average of MSRP
Bugatti	1757223.667
Maybach	546221.875
Rolls-Royce	351130.6452
Lamborghini	331567.3077
Bentley	247169.3243
McLaren	239805
Ferrari	238218.8406
Spyker	214990
Aston Martin	198123.4615
Maserati	113684.4909
Porsche	101622.3971

Lowest MSRP

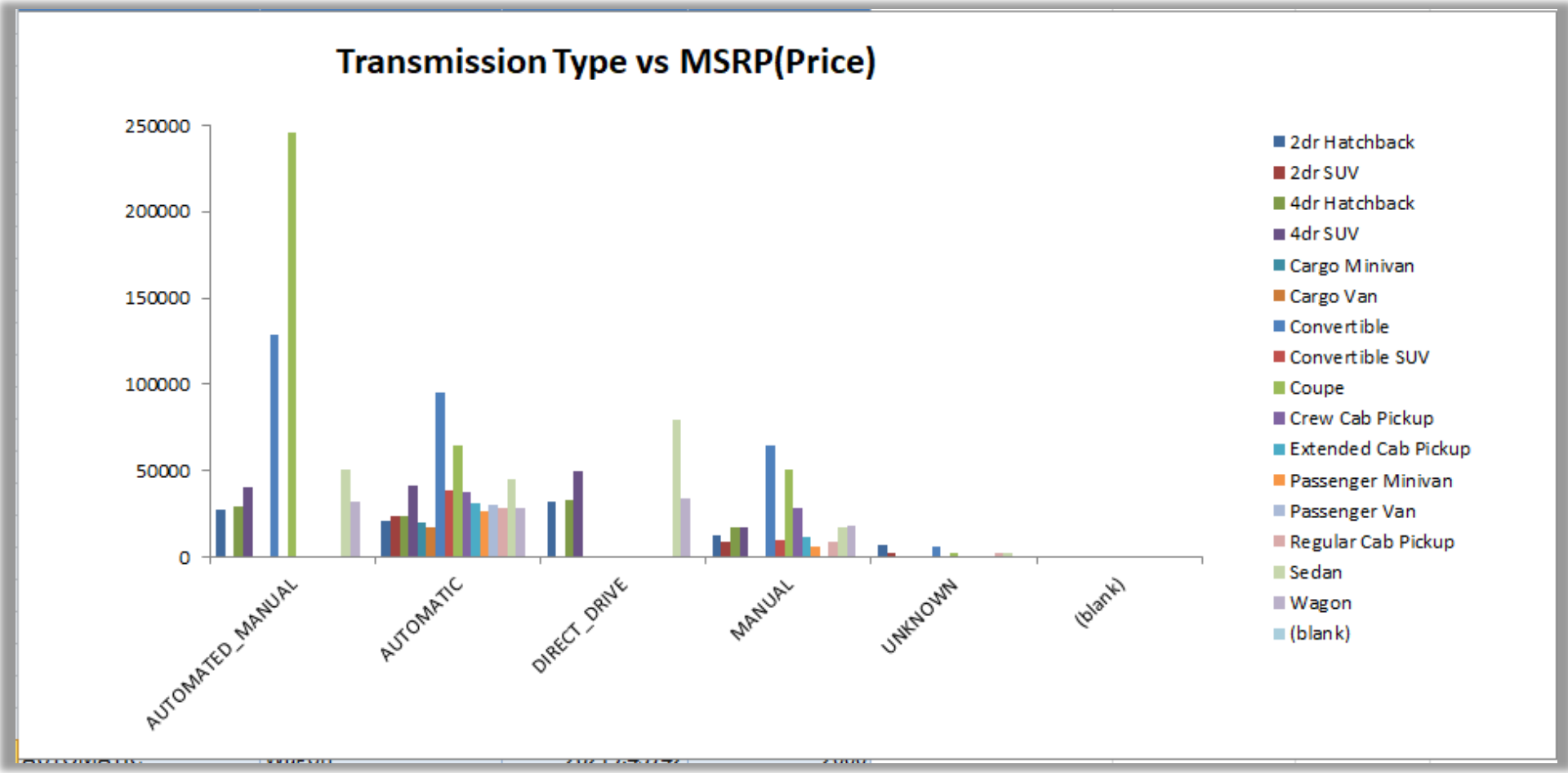
Brand	Average of MSRP
Plymouth	3296.873239
Oldsmobile	12843.79545
Suzuki	18026.4152
Pontiac	19800.0442
Scion	19932.5
Mazda	20416.62379
Mitsubishi	21340.5625
FIAT	22670.24194
Subaru	24240.67364
Dodge	24857.04537
Hyundai	24926.26255



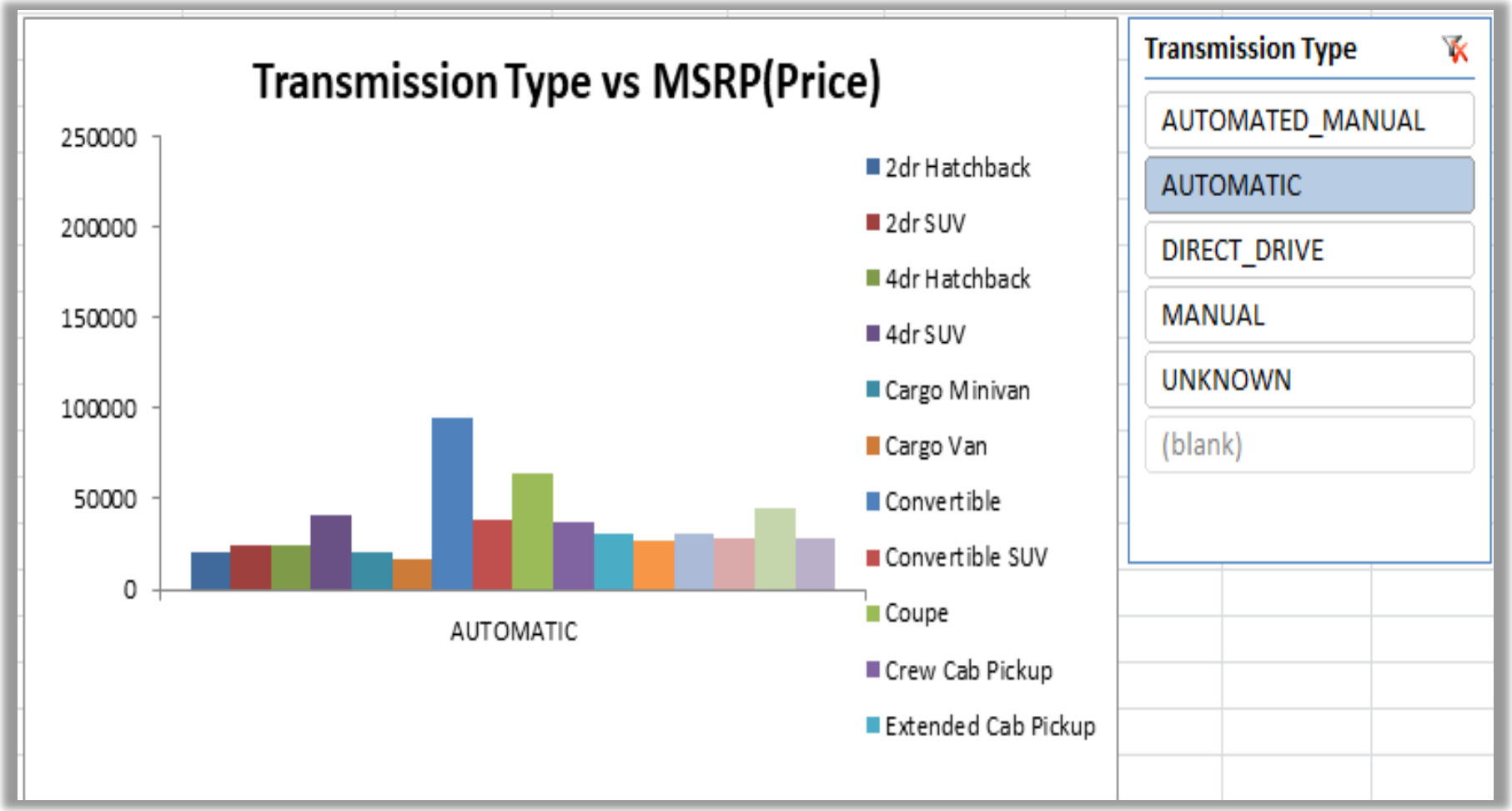
So the Plymouth have Lowest average MSRP which has 7 body style which are 2dr hatchback, 4dr hatchback, convertible, coupe, passenger minivan, sedan and wagon.



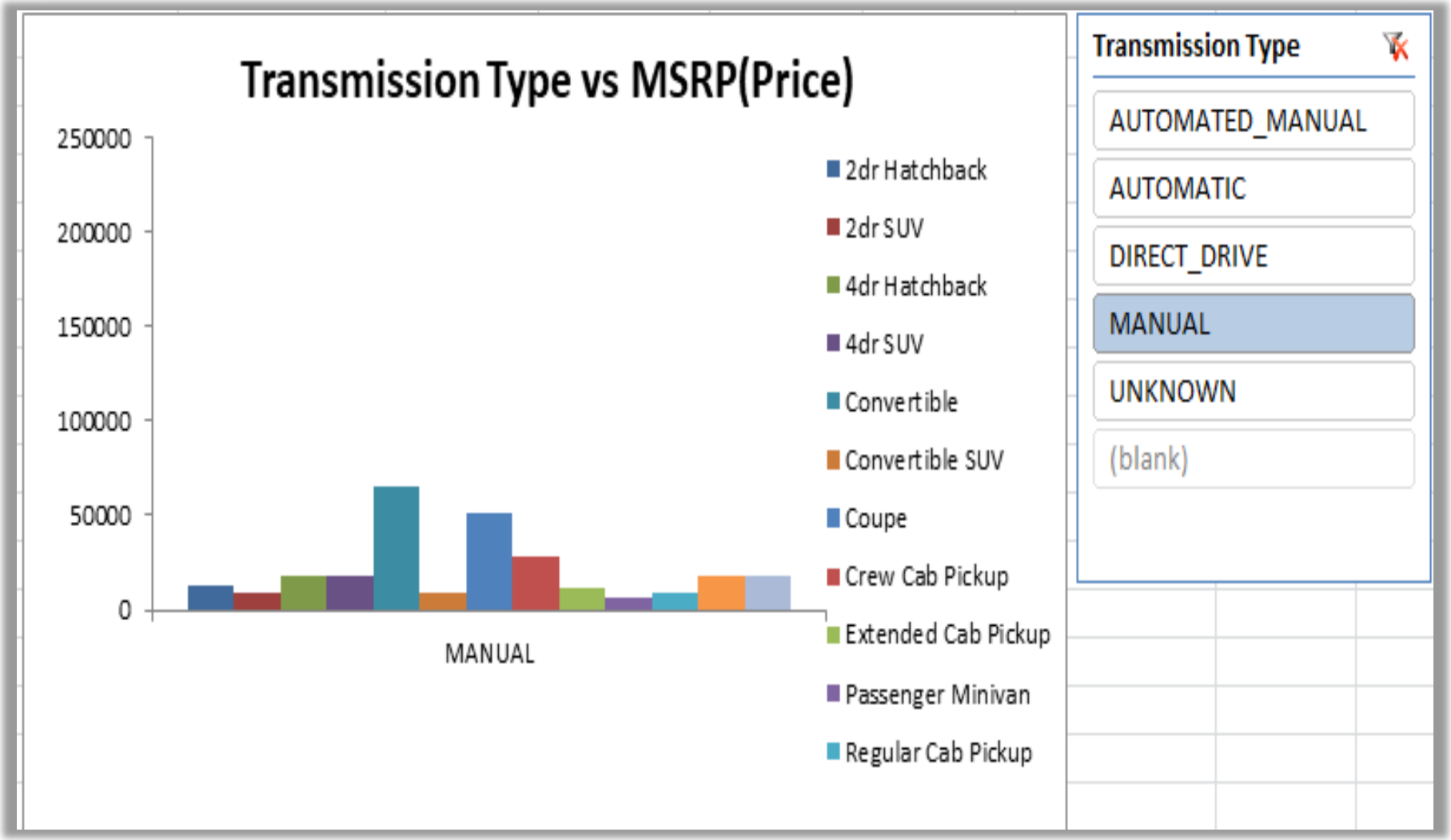
Task 3: How do the different feature such as transmission type affect the MSRP, and how does this vary by body style?



The data shows that on average, cars with automatic transmission have a higher MSRP compared to cars with manual transmission.



The price difference between automatic and manual transmission cars varies across body styles. For example, the price difference is larger for cargo van and minivan compared to Convertible and coupes.

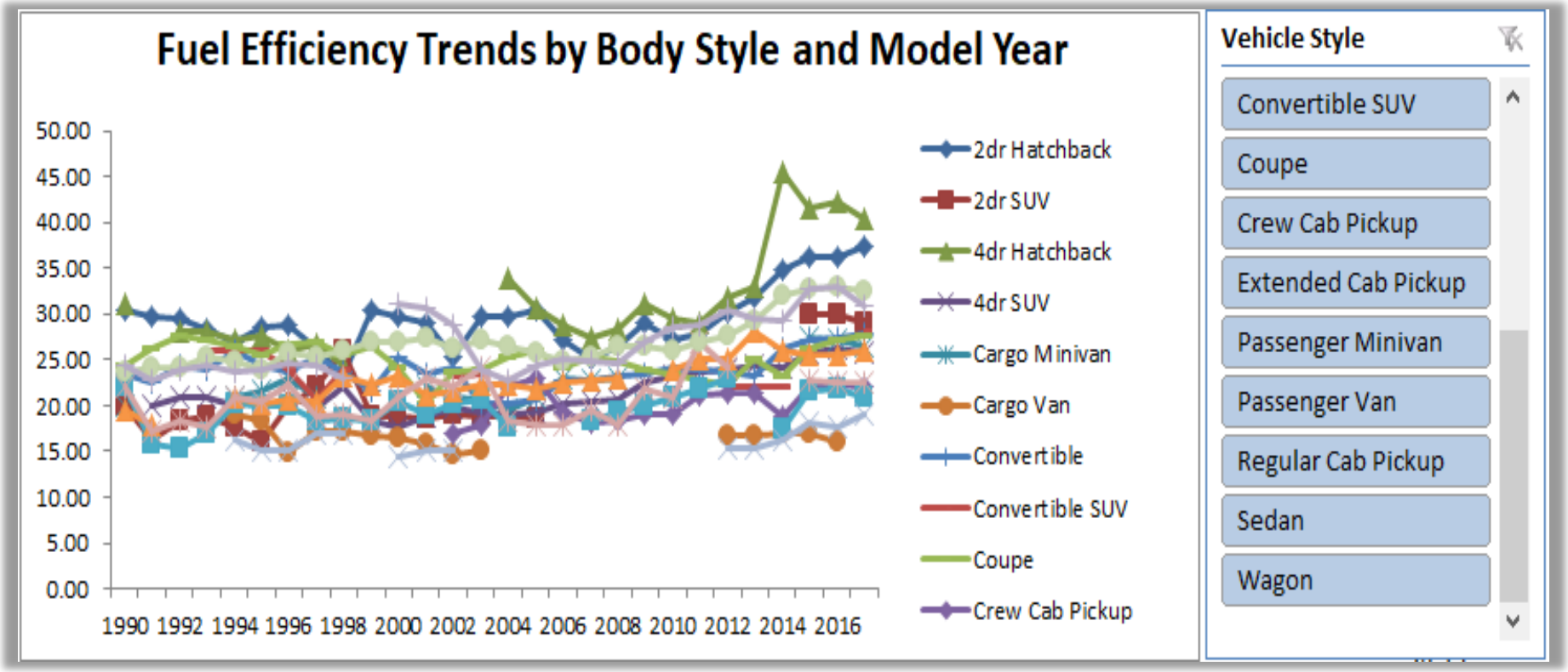


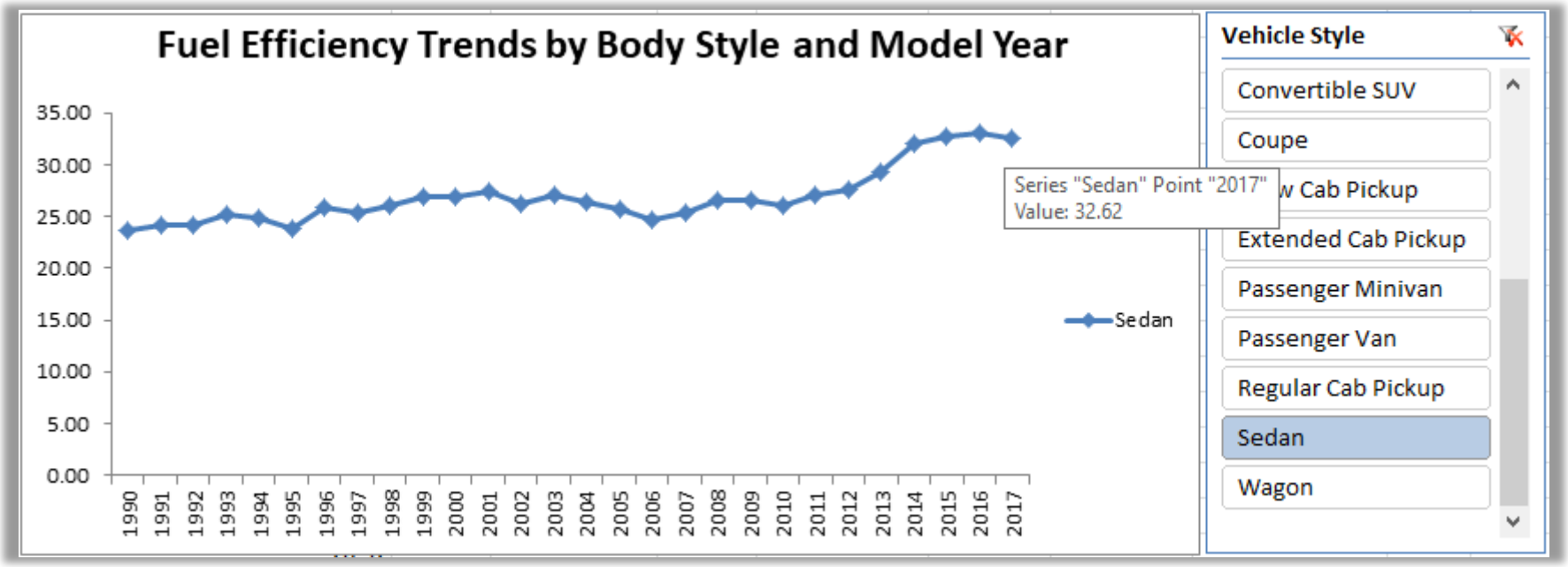
The combination of body style and transmission type has a notable impact on the MSRP. Body styles associated with higher price points, such as convertibles and coupes, tend to command a premium when equipped with automated manual transmissions.

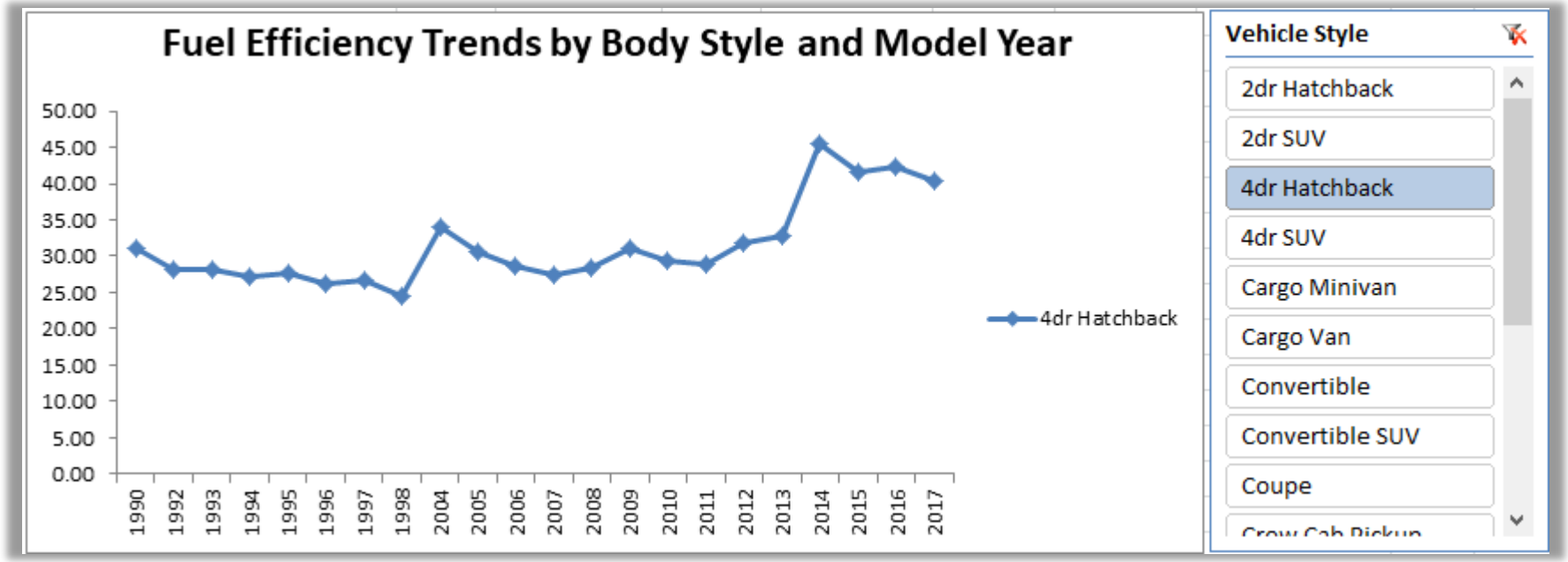
On the other hand, body styles like 2-door hatchbacks and SUVs generally have lower average prices, appealing to cost-conscious consumers.

The 4-door hatchback, sedan, and wagon fall in the mid-range, offering a balance between price, functionality, and performance.

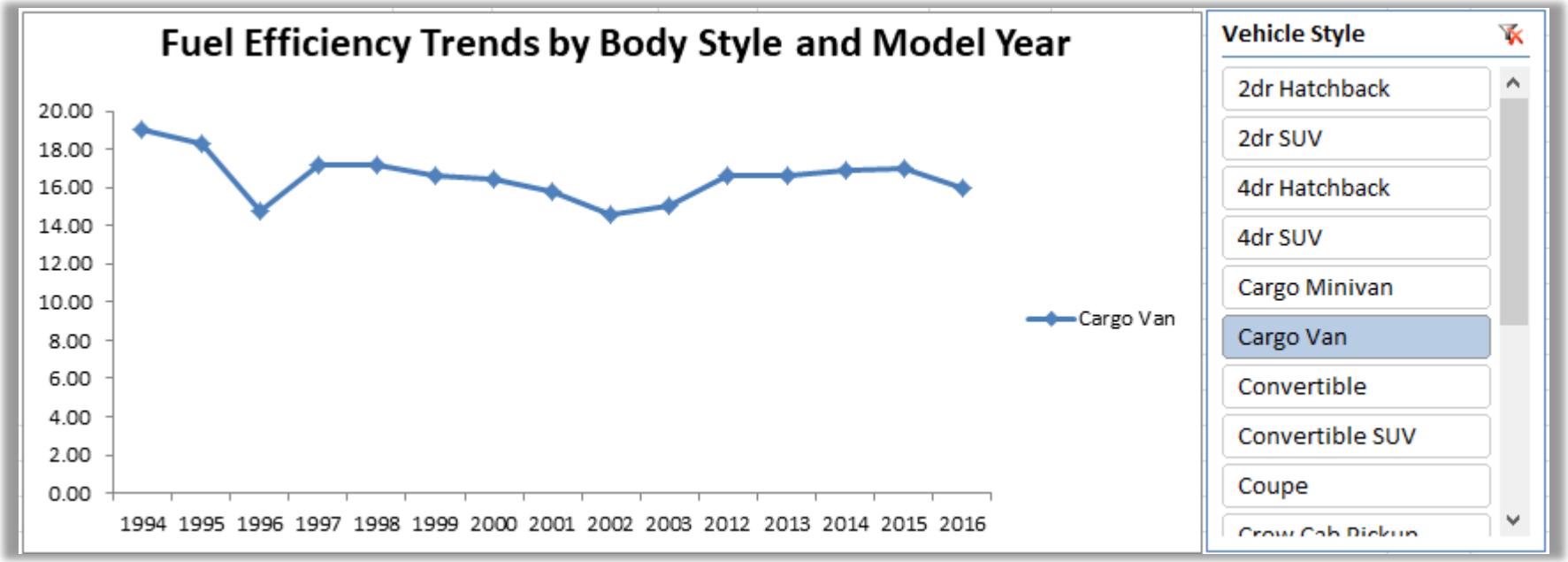
Task 4: How does the fuel efficiency of cars vary across different body styles and model years?




Fuel Efficiency Trends by Body Style and Model Year
35.00
30.00
25.00
20.00
15.00
10.00
5.00
0.00
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017
Series "Sedan" Point "2017"
Value: 32.62
Sedan
Vehicle Style
Convertible SUV
Coupe
w Cab Pickup
Extended Cab Pickup
Passenger Minivan
Passenger Van
Regular Cab Pickup
Sedan
Wagon

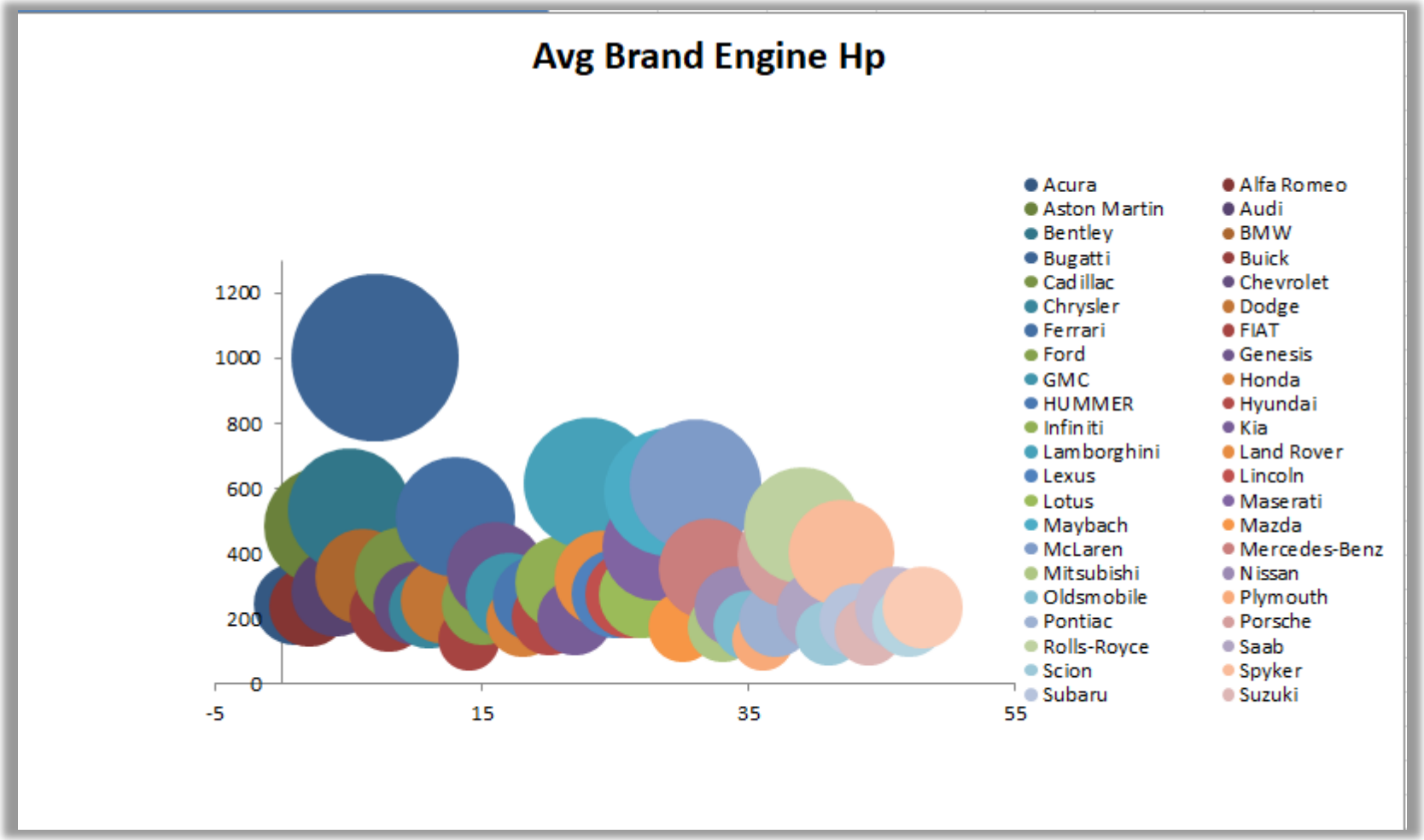

Fuel Efficiency Trends by Body Style and Model Year
50.00
45.00
40.00
35.00
30.00
25.00
20.00
15.00
10.00
5.00
0.00
1990 1992 1993 1994 1995 1996 1997 1998 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017
4dr Hatchback
Vehicle Style
2dr Hatchback
2dr SUV
4dr Hatchback
4dr SUV
Cargo Minivan
Cargo Van
Convertible
Convertible SUV
Coupe

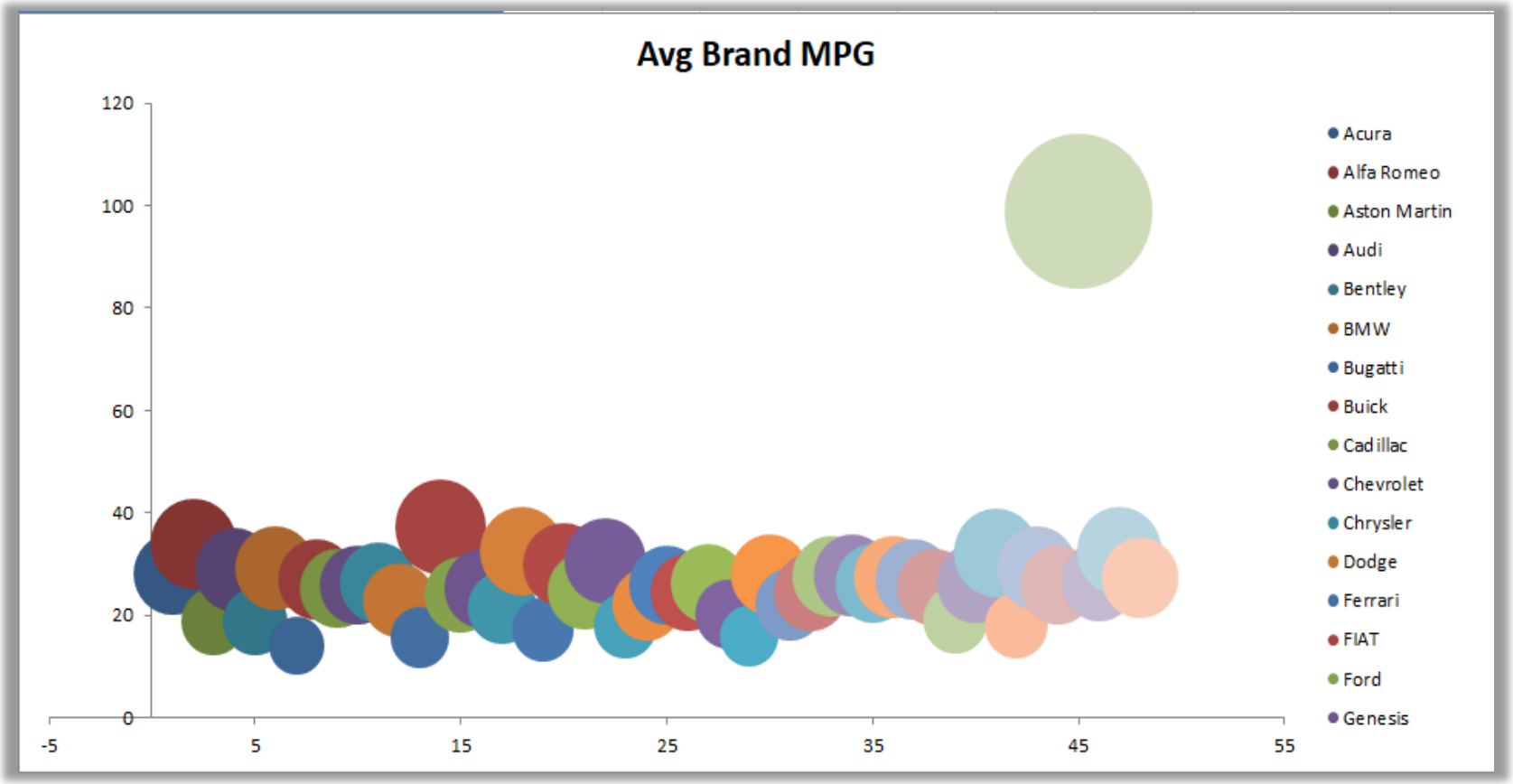
One key insight is that Sedan and 4dr Hatchback body styles have significantly improved their fuel efficiency over time. The line chart shows a clear upward trend in fuel efficiency for both body styles across different model years.

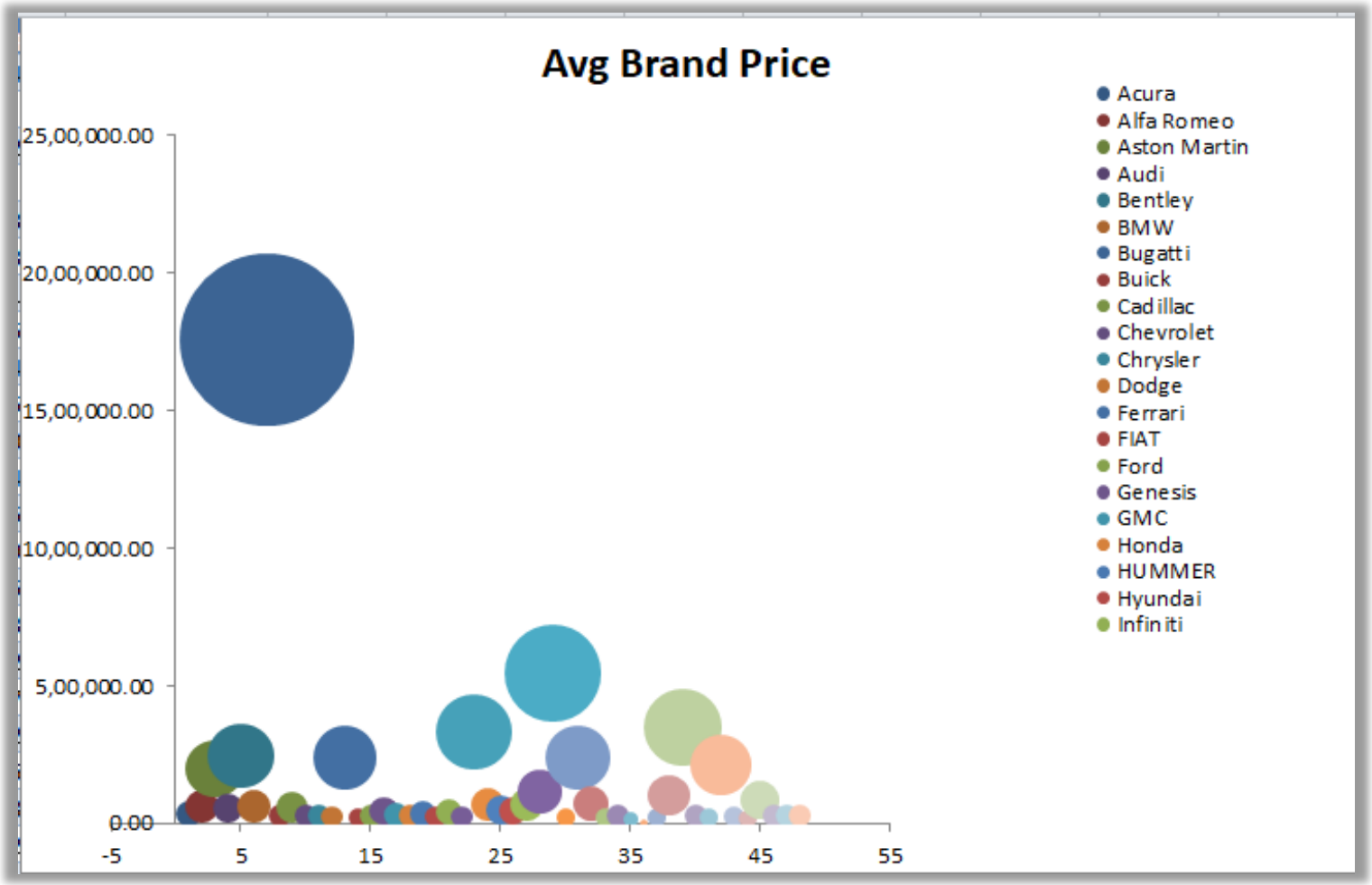


The fuel efficiency of Cargo van has actually decreased slightly over time. The line chart shows a slight downward trend in fuel efficiency for cargo vans across different model years, with the average fuel efficiency dropping from 19 mpg to 16 mpg. This trend may be concerning for consumers who prioritize fuel efficiency.

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




Avg Brand MPG
120
100
80
60
40
20
0
-5
5
15
25
35
45
55
Acura
Alfa Romeo
Aston Martin
Audi
Bentley
BMW
Bugatti
Buick
Cadillac
Chevrolet
Chrysler
Dodge
Ferrari
FIAT
Ford
Genesis


Avg Brand Price
25,00,000.00
20,00,000.00
15,00,000.00
10,00,000.00
5,00,000.00
0.00
-5
5
15
25
35
45
55
Acura
Alfa Romeo
Aston Martin
Audi
Bentley
BMW
Bugatti
Buick
Cadillac
Chevrolet
Chrysler
Dodge
Ferrari
FIAT
Ford
Genesis
GMC
Honda
HUMMER
Hyundai
Infiniti

Result

Based on the data analysis, the following recommendations can be made to enhance the business:

Market Category Enhancement:

Companies should focus on the "crossover" category, which has the highest number of car models but also high competition. They should strive to differentiate their offerings and provide unique features or value propositions to stand out in this crowded market.

Improving Crossover Luxury Category

The market category "crossover, luxury" has a good number of models but a low popularity score. To improve their business in this category, companies should invest in enhancing the luxury features and overall appeal of their crossover models. This could involve incorporating premium materials, advanced technology, and superior performance to attract luxury car buyers.

Expanding Models in Performance Hybrid Category

Although the market category "crossover, luxury, performance, hybrid" ranks high in terms of popularity score, it has a limited number of models. Companies should consider expanding their offerings in this category to capitalize on the growing demand for high-performance hybrid vehicles. By providing more choices, they can attract customers seeking both luxury and environmentally friendly options.

Leveraging Engine Horsepower for Pricing

The positive correlation between car price and engine horsepower suggests that companies can position their higher-priced models as performance-oriented vehicles. By investing in engine development and showcasing the power and performance capabilities of their cars, companies can justify premium pricing and target customers who prioritize performance.

Strategic Brand Positioning

Analyzing the variation in average prices across different manufacturers and body styles, it is essential for companies to strategically position their brands. For affordable or budget-friendly brands like Plymouth, emphasizing value for money and practicality can attract price-conscious consumers. Luxury brands like Bugatti should focus on exclusivity, craftsmanship, and cutting-edge technology to justify their premium prices.

Optimizing Transmission Options

Companies should consider the impact of transmission types on prices. Automatic transmissions generally command higher prices, so it is important to offer a range of transmission options that align with customer preferences and budget constraints. Additionally, for body styles where the price difference between transmission types is significant (such as cargo vans and minivans), providing both options can cater to a broader customer base.

Continuous Improvement in Fuel Efficiency:

While Sedan and 4-door Hatchback body styles have shown improved fuel efficiency over time, companies should prioritize innovation in fuel-efficient technologies across all body styles. This can include developing hybrid or electric variants, optimizing engine performance, and reducing overall vehicle weight to appeal to environmentally conscious consumers.

Project Workbook Link:

https://docs.google.com/spreadsheets/d/1po0oY6qg_Bp0J8cbpeGpB0Rk8QciryNr/edit?usp=share_link&ouid=112692218028277537705&rtpof=true&sd=true


THE END