Impact of Car Features on Price and Profitability

Link for Excel sheet:

https://docs.google.com/spreadsheets/d/10QceKTy49wcBmaG8W8PhtQBoad8L1MI_/edit?usp=sharing&ouid=107365393175079460343&rtpof=true&sd=true

Project Description:

This project aims to analyze a dataset containing information about various Car Brands, Car models they make and their respective car features along with their prices. The goal is to gain insights about impact of car features on price and profitability, performing various analysis tasks and also build a dashboard to better visualize the insights. The data provided has various missing or null Data, our task is to handle those missing values appropriately, by either deleting or imputing these data. There are various outliers in data, we have to find these outliers. We utilize various excel features such as pivot tables and charts to better represent data. We find trends in car features and their popularities by implementing various methodologies and data analysis techniques such as regression. Thus, by employing statistics and Excel formulas, we will extract meaningful conclusions to help understand the factors that contribute to popularity and profitability of particular cars.

Approach:

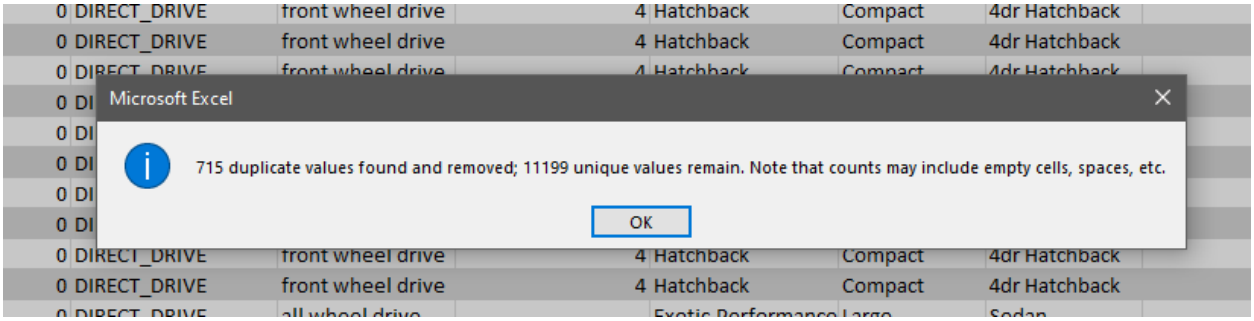
As an individual working on this project, I followed a structured approach to analyze data about Car Brands, models and features. I began by carefully examining the provided database and familiarizing myself with its structure and columns. I tried to find columns which had the most significance in the dataset. I handled missing values by eliminating columns which had most empty cells, and were not significant. And imputed data into cells that were necessary for analysis. Then, I utilized Excel fundamentals to retrieve the necessary information for each task, employing appropriate functions and statistical methods. I focused on data accuracy and quality throughout the project, ensuring reliable results. By leveraging my Excel skills and maintaining a systematic workflow, I successfully executed the project and created a comprehensive report that fulfilled the objectives of providing marketing insights and investor metrics.

Tech-Stack Used:

For this project, I utilized Microsoft Excel as the primary software tool.

Data Cleaning:

Given Data had various missing and duplicate values. For accurate analysis we need to handle this missing data, and eliminate the duplicate data as it is redundant and might skew the results. For Removing the duplicate data we used excel's Remove Duplicates feature in the Data Tools Tab. We had 715 duplicate rows, which were removed completely.

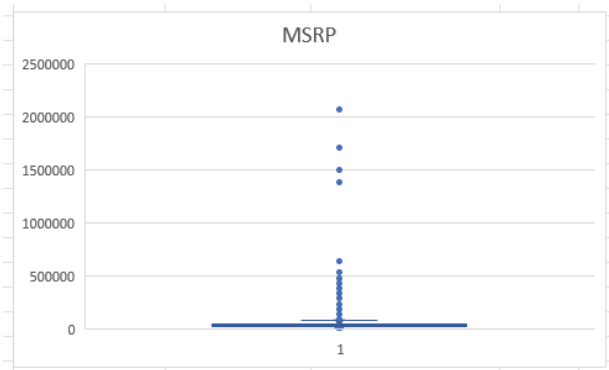


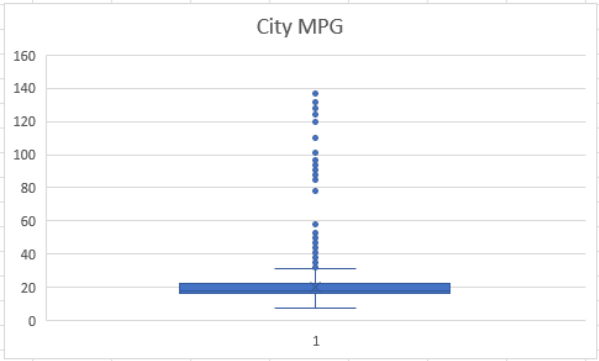
To find missing values we used the COUNTBLANK formula in excel.
=COUNTBLANK(A$2:A$11160)

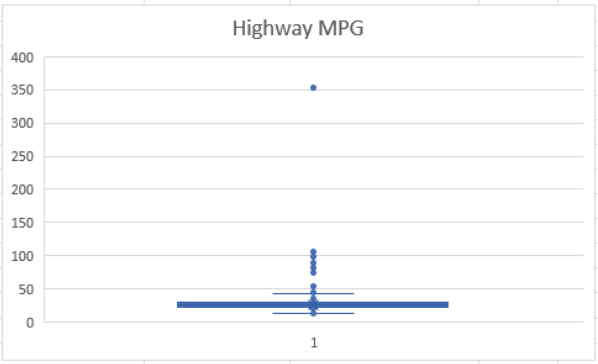
Columns	No. of Null values	Count N/A or Unknown
Make	0	0
Model	0	0
Year	0	0
Engine Fuel Type	3	0
Engine HP	69	0
Engine Cylinders	30	0
Transmission Type	0	0
Driven_Wheels	0	0
Number of Doors	6	0
Market Category	0	3376
Vehicle Size	0	0
Vehicle Style	0	0
highway MPG	0	0
city mpg	0	0
Popularity	0	0
MSRP	0	0

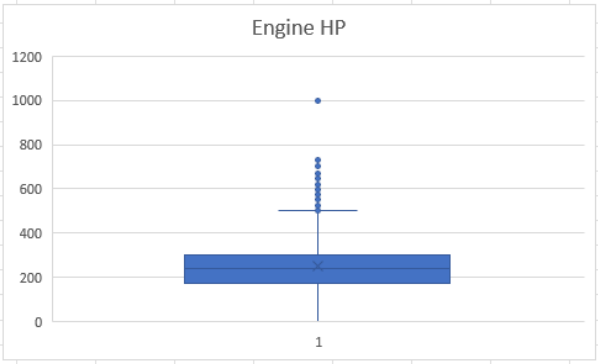
We removed rows which had less no. of nulls and imputed values in columns such as Engine HP and Engine Cylinders according to the given data.

Data also had some outliers or false values, which needed to be handled. We plotted these outliers using BOX and Whisker chart type.
As seen in the The chart below the features have outliers, some of which are justified but, feature Highway MPG has value which is a bit out of range. So we check with the data of similar Cars and adjust it accordingly.


MSRP
2500000
2000000
1500000
1000000
500000
0
1


City MPG
160
140
120
100
80
60
40
20
0
1


Highway MPG
400
350
300
250
200
150
100
50
0
1


Engine HP
1200
1000
800
600
400
200
0
1

Insights:

Analysis:

Task 1:

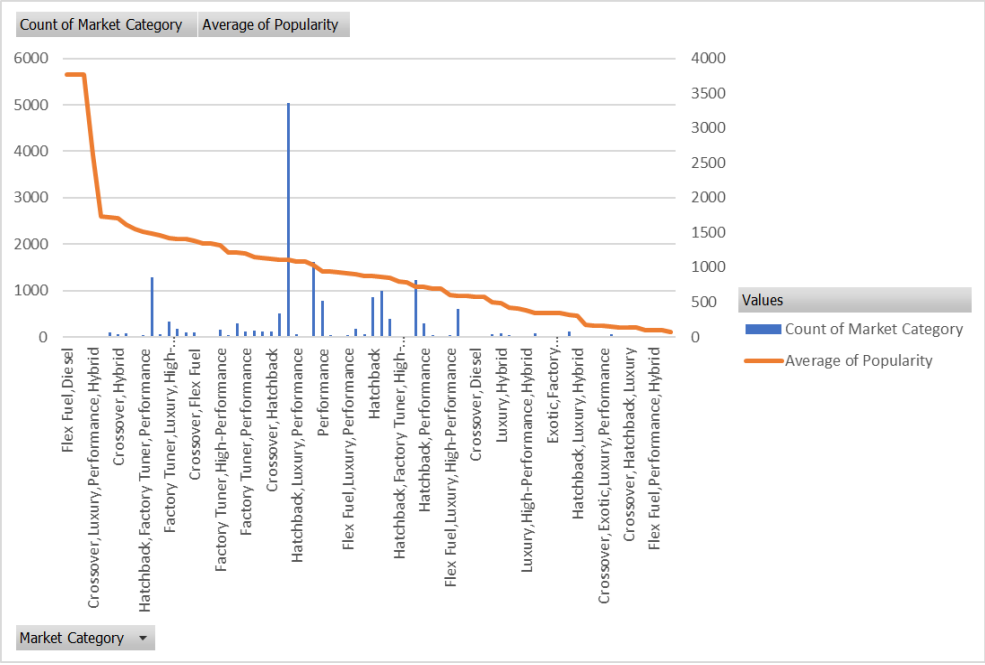
Insight Required: How does the popularity of a car model vary across different market categories?

To perform this task we utilized a pivot table in excel that shows the number of car models in each market category and their corresponding popularity scores.

	Market Category	Count of Market Category	Average of Popularity
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	2607.4
7	Crossover,Performance	69	2585.956522
8	Crossover,Hybrid	42	2563.380952
9	Diesel,Luxury	47	2416.106383
10	Luxury,Performance,Hybrid	11	2333.181818
11	Hatchback,Factory Tuner,Performance	20	2271.9
12	Flex Fuel	855	2225.71345
13	Crossover,Luxury,Diesel	33	2195.848485
14	Factory Tuner,Luxury,High-Performance	215	2133.367442
15	Hybrid	121	2116.586777
16	Hatchback,Hybrid	64	2111.15625
17	Crossover,Flex Fuel	64	2073.75
18	Crossover,Hatchback,Factory Tuner,Performance	6	2009

This pivot table shows Market Category with its count and average popularity for each.

From the above pivot table we plot a combo chart of column-line charts. We select a secondary axis for count to better visualize the chart.



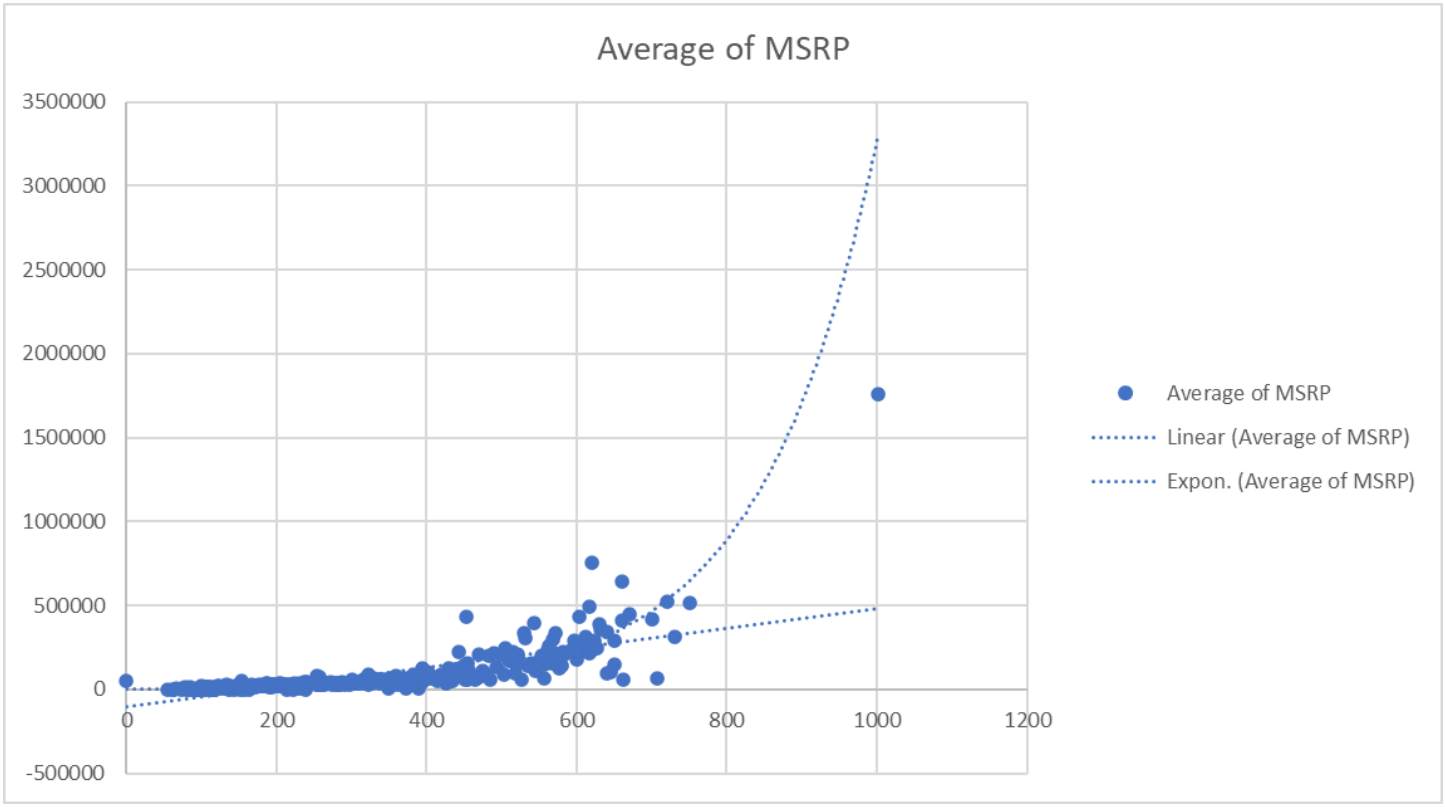
Task 2:

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

To find the relationship between a car's engine power that is Engine HP and its MSRP, we utilize power pivot to find average MSRP for each Engine HP. We then copy this data into a new table and then create a scatter plot of Engine HP vs average MSRP.

	A	B	C	D	E	F
1	Engine HP	Average of MSRP		Engine HP	Average of MSRP	
2	163	2000		163	2000	
3	114	2000		114	2000	
4	102	2000		102	2000	
5	105	2000		105	2000	
6	63	2000		63	2000	
7	113	2000		113	2000	
8	73	2000		73	2000	
9	62	2000		62	2000	
10	96	2000		96	2000	
11	97	2000		97	2000	
12	82	2000		82	2000	
13	81	2000		81	2000	
14	90	2000		90	2000	
15	118	2000		118	2000	
16	92	2000		92	2000	
17	55	2000		55	2000	
18	214	2000		214	2000	

We now Create a scatter plot for the above table.



We have also added trendlines to understand how MSRP is changing according to the change inEngine HP. Trend seems to increase exponentially rather than linearly, but to predict more accurately we need to have more data available.

Task 3:

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

To perform this analysis, we need to consider every feature which is correlated with the price of a car. For this we need to perform regression analysis and then plot coefficients of each feature to check which have most impact on MSRP. But for regression analysis we need to have numerical data, so we first convert the data into numerical data by converting categorical data into encoded data.

Vehicle Size	Encoding
Compact	1
Large	3
Midsize	2

Vehicle Style	Encoding
Coupe	1
Sedan	2
Convertible	3
4dr SUV	4
Wagon	5
Crew Cab Pickup	6
Extended Cab Pickup	7
4dr Hatchback	8
Regular Cab Pickup	9

We use this type of conversion to encode data into numerical values.

Make (Encoded)	Year	Engine Fuel Type (Encoded)	Engine HP	Engine Cylinders	Transmission Type (Encoded)	Driven Wheels (Encoded	Number of Doors	Vehicl
1	2008	1	1001	16	3	3	2	
1	2009	1	1001	16	3	3	2	
1	2008	1	1001	16	3	3	2	
2	2016	2	1000	0	1	3	4	
2	2016	2	1000	0	1	3	4	
2	2015	2	1000	0	1	3	4	
2	2014	2	1000	0	1	3	4	
2	2014	2	1000	0	1	1	4	
2	2016	2	1000	0	1	3	4	
2	2015	2	1000	0	1	3	4	
2	2015	2	1000	0	1	1	4	

We get this type of data. But we have to normalize it first. As the parameters have very large differences in their ranges.

To normalize we find maximum and minimum values in each column and then, normalize them using the following formula.

=(Analysis_Task3!$F2-Analysis_Task3!F$11199)/Analysis_Task3!F$11201

Here we subtract minimum values from each value and then divide with the difference between maximum and minimum values, to get normalized values between 0 and 1.

Thus all of the values will get converted into range from 0 to 1.

Make (Encoded)	Engine Fuel Type (Encoded)	Engine HP	Engine Cylinders	Transmission Type (Encoded)	Driven Wheels (Encoded)	Number of Doors	Vehicle Size (Encoded)	Vehicle Style (Encoded)	highway MPG	city mpg	Popularity	MSRP
0	0	1	1	0.666666667	0.666666667	0	0	0	0.005847953	0.0076923	0.144650752	1
0	0	1	1	0.666666667	0.666666667	0	0	0	0.005847953	0.0076923	0.144650752	0.825509
0	0	1	1	0.666666667	0.666666667	0	0	0	0.005847953	0.0076923	0.144650752	0.72581
0.021276596	0.2	0.998942918	0	0	0.666666667	1	1	0.066666667	0.271929825	0.6538462	0.245623342	0.064199
0.021276596	0.2	0.998942918	0	0	0.666666667	1	1	0.066666667	0.257309942	0.6461538	0.245623342	0.053297
0.021276596	0.2	0.998942918	0	0	0.666666667	1	1	0.066666667	0.251461988	0.6307692	0.245623342	0.049905
0.021276596	0.2	0.998942918	0	0	0.666666667	1	1	0.066666667	0.239766082	0.6076923	0.245623342	0.049663
0.021276596	0.2	0.998942918	0	0	0	1	1	0.066666667	0.228070175	0.6230769	0.245623342	0.044285
0.021276596	0.2	0.998942918	0	0	0.666666667	1	1	0.066666667	0.277777778	0.7230769	0.245623342	0.042395
0.021276596	0.2	0.998942918	0	0	0.666666667	1	1	0.066666667	0.274853801	0.6769231	0.245623342	0.040215
0.021276596	0.2	0.998942918	0	0	0	1	1	0.066666667	0.228070175	0.6230769	0.245623342	0.037792
0.021276596	0.2	0.998942918	0	0	0	1	1	0.066666667	0.228070175	0.6230769	0.245623342	0.037744
0.021276596	0.2	0.998942918	0	0	0.666666667	1	1	0.066666667	0.271929825	0.7307692	0.245623342	0.03755
0.021276596	0.2	0.998942918	0	0	0.666666667	1	1	0.066666667	0.263157895	0.7230769	0.245623342	0.03537
0.021276596	0.2	0.998942918	0	0	0.666666667	1	1	0.066666667	0.263157895	0.7230769	0.245623342	0.03537
0.021276596	0.2	0.998942918	0	0	0	1	1	0.066666667	0.257309942	0.6923077	0.245623342	0.035128
0.021276596	0.2	0.998942918	0	0	0.666666667	1	1	0.066666667	0.277777778	0.7230769	0.245623342	0.033432
0.021276596	0.2	0.998942918	0	0	0	1	1	0.066666667	0.228070175	0.6230769	0.245623342	0.032947
0.021276596	0.2	0.998942918	0	0	0	1	1	0.066666667	0.248538012	0.6692308	0.245623342	0.032899
0.021276596	0.2	0.998942918	0	0	0	1	1	0.066666667	0.248538012	0.6692308	0.245623342	0.032899
0.021276596	0.2	0.998942918	0	0	0.666666667	1	1	0.066666667	0.260233918	0.7	0.245623342	0.031009
0.042553191	0	0.734672304	0.75	0.666666667	0.666666667	0	0.5	0.133333333	0.01754386	0.0307692	0.204420866	0.258491
0.042553191	0	0.734672304	0.75	0.666666667	0.666666667	0	0.5	0	0.01754386	0.0307692	0.204420866	0.236784
0.063829787	0	0.714587738	0.75	0.666666667	0	0	0.5	0	0.011695906	0.0307692	0.490185676	0.154075
0.063829787	0	0.714587738	0.75	0.666666667	0	0	0.5	0	0.011695906	0.0307692	0.490185676	0.152085
0.063829787	0	0.714587738	0.75	0.666666667	0	0	0.5	0	0.011695906	0.0307692	0.490185676	0.152085
0.042553191	0	0.702959831	0.75	0.666666667	0.666666667	0	0.5	0.133333333	0.011695906	0.0230769	0.204420866	0.264935
0.042553191	0	0.702959831	0.75	0.666666667	0.666666667	0	0.5	0.133333333	0.011695906	0.0230769	0.204420866	0.264935
0.042553191	0	0.702959831	0.75	0.666666667	0.666666667	0	0.5	0	0.01754386	0.0307692	0.204420866	0.240152
0.042553191	0	0.702959831	0.75	0.666666667	0.666666667	0	0.5	0	0.01754386	0.0307692	0.204420866	0.240152
0.085106383	0	0.689217759	0.5	0.333333333	0	0	1	0	0.026315789	0.0461538	0.326967286	0.030983
0.085106383	0	0.689217759	0.5	0.333333333	0	1	1	0.066666667	0.029239766	0.0461538	0.326967286	0.030983
0.085106383	0	0.689217759	0.5	0.333333333	0	1	1	0.066666667	0.029239766	0.0461538	0.326967286	0.030983
0.085106383	0	0.689217759	0.5	1	0	0	1	0	0.026315789	0.0461538	0.326967286	0.029311

We now use this data to perform regression analysis by using the Data Analysis feature in the Data menu. We get following output:

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.693768659
R Square	0.481314952
Adjusted R Square	0.480758373
Standard Error	0.021486617
Observations	11196

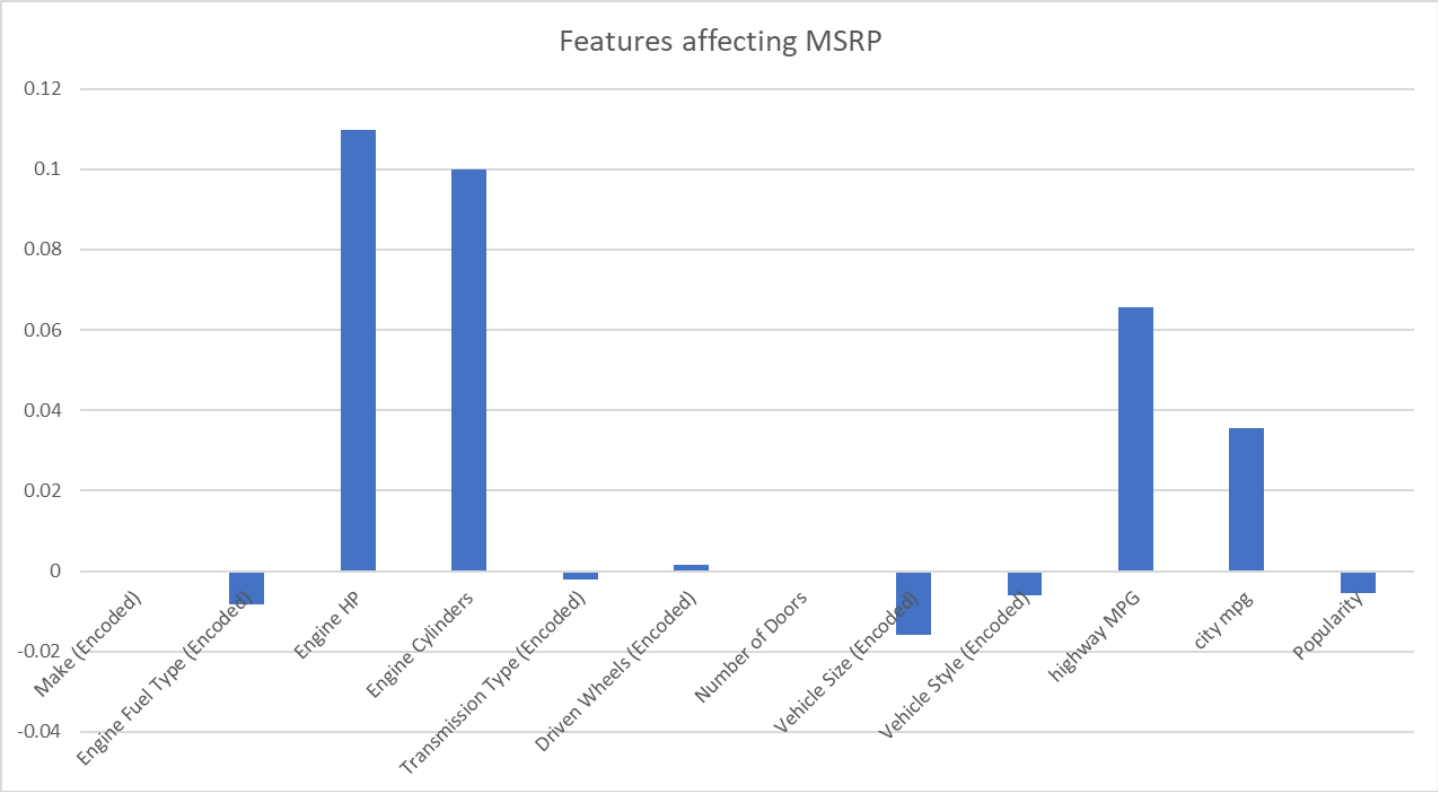
ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	12	4.790932299	0.399244358	864.7741572	0
Residual	11183	5.162908281	0.000461675		
Total	11195	9.95384058			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-0.031589947	0.001778814	-17.75899174	1.32073E-69	-0.035076737	-0.028103158	-0.035076737	-0.028103158
Make (Encoded)	-0.000126829	0.000898814	-0.141107322	0.887787701	-0.001888663	0.001635004	-0.001888663	0.001635004
Engine Fuel Type (Encoded)	-0.008170822	0.000930401	-8.782046676	1.83872E-18	-0.009994572	-0.006347073	-0.009994572	-0.006347073
Engine HP	0.109738052	0.003028023	36.24082452	5.1054E-272	0.103802594	0.115673511	0.103802594	0.115673511
Engine Cylinders	0.099772838	0.003461532	28.82331736	2.7635E-176	0.092987626	0.106558051	0.092987626	0.106558051
Transmission Type (Encoded)	-0.002025972	0.00084897	-2.386386485	0.017031389	-0.003690103	-0.00036184	-0.003690103	-0.00036184
Driven Wheels (Encoded)	0.001653045	0.000658457	2.510481034	0.012070678	0.000362352	0.002943737	0.000362352	0.002943737
Number of Doors	-0.000224093	0.000559842	-0.400279344	0.688958426	-0.001321482	0.000873295	-0.001321482	0.000873295
Vehicle Size (Encoded)	-0.016006733	0.000670551	-23.87101698	7.0368E-123	-0.017321131	-0.014692335	-0.017321131	-0.014692335
Vehicle Style (Encoded)	-0.006067337	0.001019628	-5.950541158	2.75268E-09	-0.008065987	-0.004068687	-0.008065987	-0.004068687
highway MPG	0.065650438	0.017952962	3.65680265	0.000256553	0.030459471	0.100841405	0.030459471	0.100841405
city mpg	0.035422796	0.006433244	5.506210878	3.74738E-08	0.022812505	0.048033086	0.022812505	0.048033086
Popularity	-0.005513007	0.000923333	-5.97076887	2.43311E-09	-0.007322902	-0.003703112	-0.007322902	-0.003703112

By using this we can plot a bar graph to see which features are affecting the MSRP most.

We select coefficients of each features and create a bar graph as below:



As we can see from above graph that Engine HP, Cylinders and MPG are some of the deciding factors for the MSRP of a car.

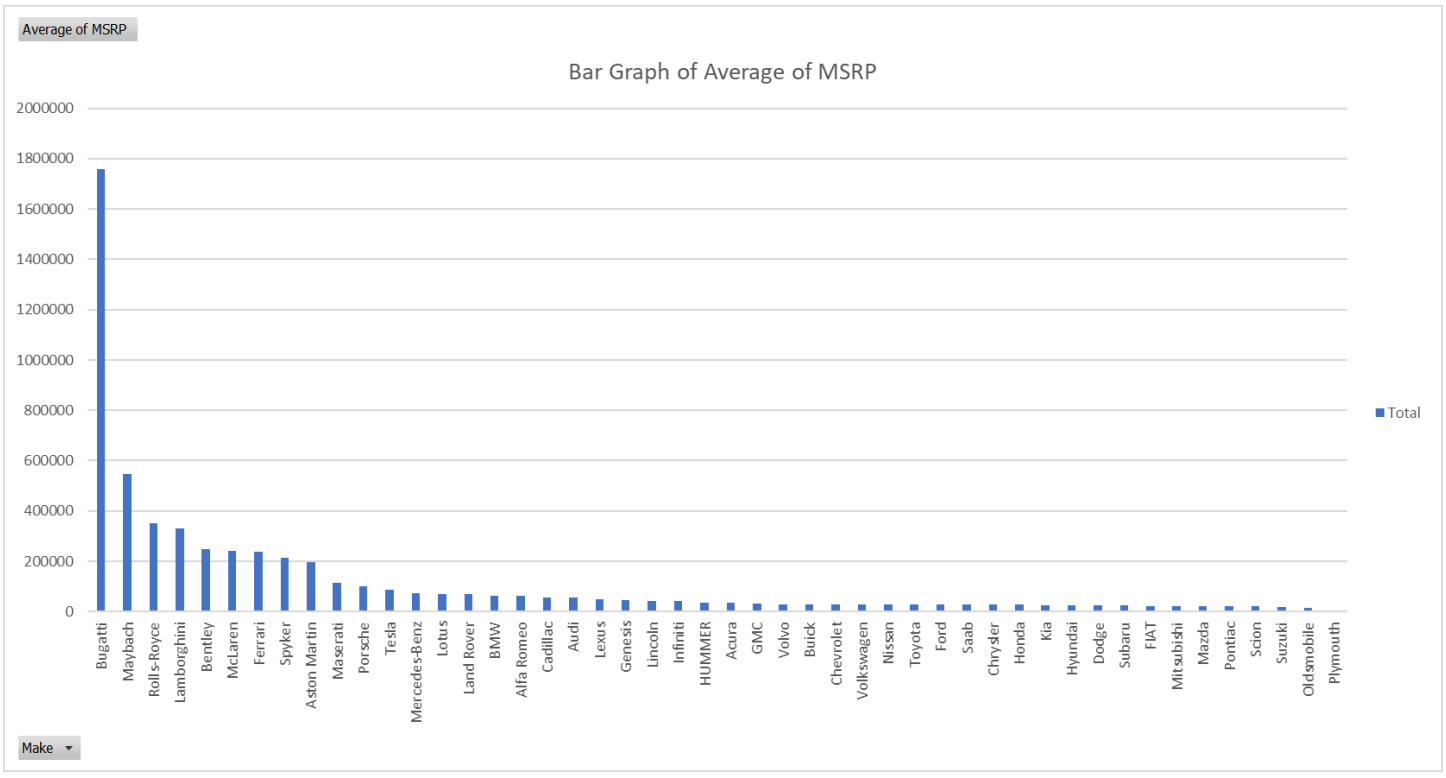
Task 4:

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

For doing this task we utilized pivot tables and found the average price of a car for each car manufacturer.

	Car Manufactures	Average of MSRP
2	Bugatti	1757223.667
3	Maybach	546221.875
4	Rolls-Royce	351130.6452
5	Lamborghini	331567.3077
6	Bentley	247169.3243
7	McLaren	239805
8	Ferrari	238218.8406
9	Spyker	214990
10	Aston Martin	198123.4615
11	Maserati	113684.4909
12	Porsche	101622.3971
13	Tesla	85255.55556
14	Mercedes-Benz	72135.02647
15	Lotus	68377.14286
16	Land Rover	68067.08633

By using this we can plot a column chart to visualize this data.



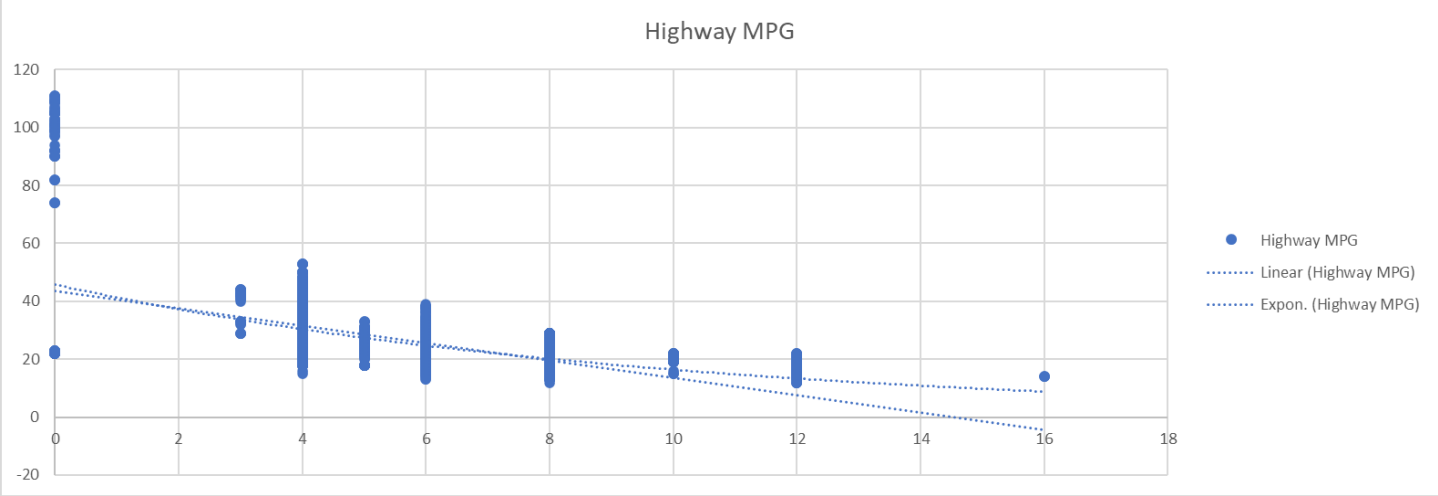
We can see here that brands such as Bugatti have very high average MSRP as they are into high end cars and don't have any cars in lower price segments.

Task 5:

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

To find relationships between fuel efficiency and number of cylinders in a car's engine we have to create a scatter plot of number of cylinders vs its MPG and see if there exists any trend by plotting a trendline.

We select two columns Engine Cylinders and Highway MPG and create a scatter plot.



We also create a correlation matrix to check if there exist any correlation between them.

We create a pivot table of no. of Cylinders and Highway MPG and City MPG. We create a correlation matrix by using the CORREL function in excel and conditional formatting.

No. of Cylinders	Average of highway MPG	Average of city mpg
0	81.6627907	90.1744186
3	38.66666667	32.03333333
4	31.50057484	23.9029662
5	26.06508876	18.77514793
6	24.00679634	17.13452074
8	20.17278287	14.18399592
10	20	12.56923077
12	17.73684211	11.25
16	14	8
Grand Total	26.61403352	19.73214446

No. of Cylinders	Average of highway MPG	Average of city mpg
0	81.6627907	90.1744186
3	38.66666667	32.03333333
4	31.50057484	23.9029662
5	26.06508876	18.77514793
6	24.00679634	17.13452074
8	20.17278287	14.18399592
10	20	12.56923077
12	17.73684211	11.25
16	14	8

Correlation Between Cylinders and Highway MPG

	No. of Cylinders	Average of highway MPG	Average of city mpg
No. of Cylinders	1		
Average of highway MPG	-0.777122379	1	
Average of city mpg	-0.729775621	0.996412646	1

As we can see there is less correlation between MPG and no. cylinders. But there is high correlation between highway MPG and city MPG.

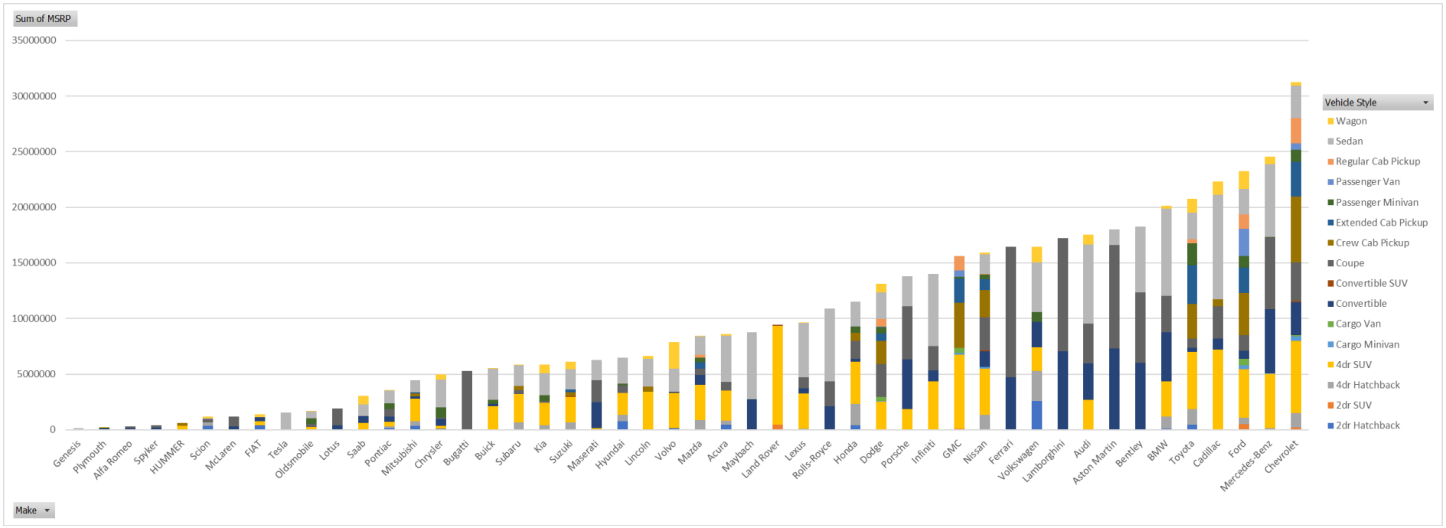
Dashboard:

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

We created a stacked column chart of car price for each brand and each body style in that column. To create this chart, we first need to create a pivot table consisting of Sum of MSRP for each category. These categories being, brand of car in row and body style in columns and we get an interactive table consisting of total MSRP for each body style that each car brand makes. Table looks like below:

Sum of MSRP	Body Style																
Car Brand	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
Genesis															139850		139850
Plymouth	40000		14000				85631		8000			31688			38759	16000	234078
Alfa Romeo							129800		178200								308000
Spyker							219990		209990								429980
HUMMER				377490						242405							619895
Scion	366325		282470						330210						32500	184445	1195950
McLaren							280225		918800								1199025
FIAT	420715			369305			327965									287570	1405555
Tesla															1534600		1534600
Oldsmobile				238150			2000		274015			492055			665161	20000	1691381
Lotus							413260		1501300								1914560
Saab	12000		34586	541905			632628								1066500	751280	3038899
Pontiac	148782		162975	401550			463914		663715			541192			1156535	20855	3559518
Mitsubishi	370169		403835	2009807	2000		209893			240210	134360	2000		8000	1058563		4438837
Chrysler	98805			250545			628105		112510			922295			2479859	501075	4993194
Bugatti									5271671								5271671
Buick				2141770			179325		18534			330065			2838590	8212	5516496
Subaru	12000		678060	2539900					354476	365975					1833110	10000	5793521
Kia			406960	2049645					142630			494650			1976360	772405	5842650
Suzuki	44496	12000	584387	2303493				120194		304131	259659				1797070	683707	6109137
Maserati				155000			2342963		1972284						1782400		6252647
Hyundai	789650		528880	1994390					685920			133075			2323987		6455902
Lincoln				3422570					17342	453260					2458245	269705	6621122
Volvo	157550			3131700			121600		6000						2072945	2416971	7906766
Mazda	18000	12000	853180	3175515			870505		541879		580033	443130		265486	1618571	33350	8411649
Acura	480917		357440	2663505					793748						4134552	201360	8631522
Maybach							2762750								5976800		8739550
Land Rover		476394		8839200				145731									9461325
Lexus			94700	3152974			472065		1016472						4837596	31105	9604912
Rolls-Royce							2141365		2204675						6539010		10885050
Honda	413200		1919260	3800589			252135		1588705	750215		553185			2264390		11541679
Dodge	38000	12000	16000	2462875	60520	338497	6000		2973842	2072780	684682	557425	70708	651408	2409585	793055	13147377
Porsche	28827			1815200			4504586		4758533						2713500		13820646
Infiniti				4340200			980050		2175750						6490009		13986009
GMC		118835		6633919	142750	460085				4062482	2175866	150630	599670	1284328			15628565
Nissan	14683		1347320	4149630	128620		1406552	131075	2937632	2422300	1026379	413320		19914	1763130	175000	15935555
Ferrari							4723811		11713289								16437100

We can plot a stacked column chart from this table where each column in the chart for a particular brand would have the total sum of MSRP of all body styles and sections would have different colors to identify each body style. Chart looks like below:



We can also change chart features by changing the filter in the pivot table.

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

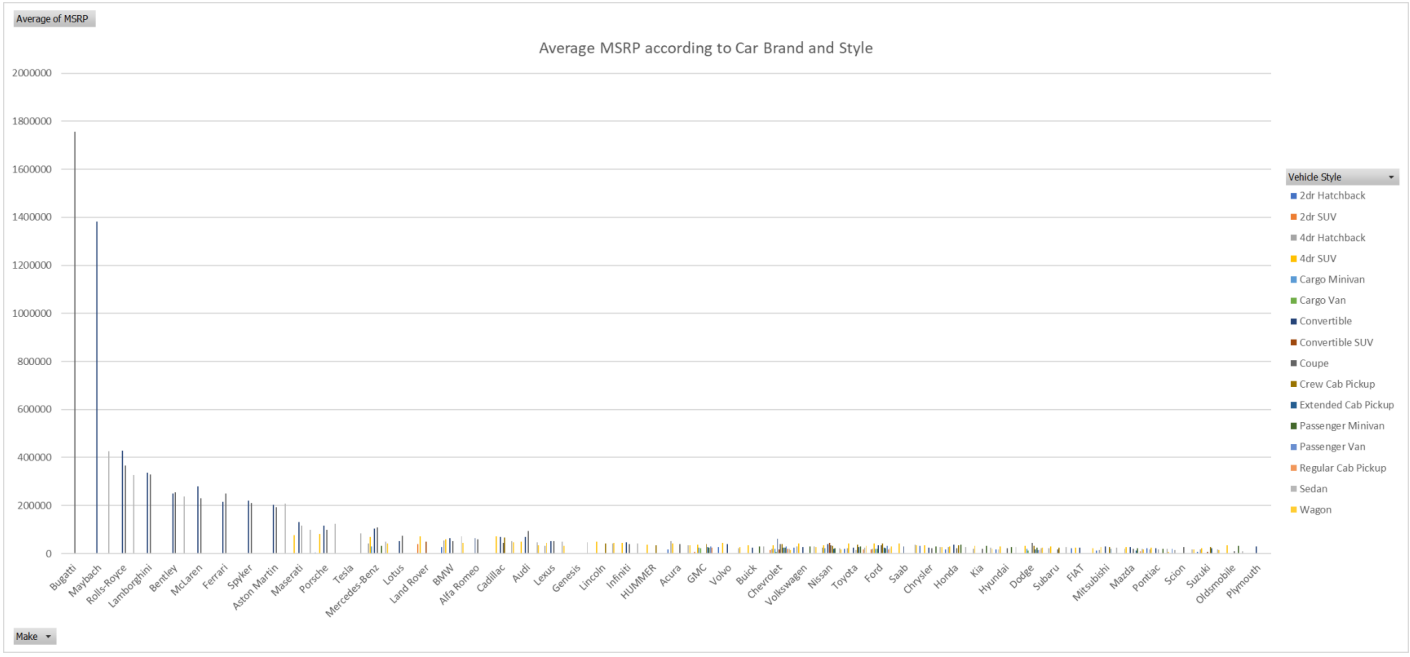
We created a clustered column chart of average car price for each brand and each body style in that column. To create this chart, we first need to create a pivot table consisting of the average of MSRP for each category. These categories being, brand of car in row and body style in columns and we get an interactive table consisting of average MSRP for each body style that each car brand makes.

Pivot table looks like below:

Average of MSRP	Body Style															
Car Brands	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
Bugatti									1757223.667								1757223.667
Maybach							1381375								426914.2857		546221.875
Rolls-Royce							428273		367445.8333						326950.5		351130.6452
Lamborghini							336402.381		328291.9355								331567.3077
Bentley							250536.25		254270.4						236836		247169.3243
McLaren							280225		229700								239805
Ferrari							214718.6818		249218.9149								238218.8406
Spyker							219990		209990								214990
Aston Martin							203379.3056		192892.6042						206962.1429		198123.4615
Maserati				77500			130164.6111		116016.7059						99022.22222		113684.4909
Porsche	5765.4			82509.09091			115502.2051		99136.10417						123340.9091		101622.3971
Tesla															85255.55556		85255.55556
Mercedes-Benz			40933.33333	68400.13889	28950		104617.5273		109713.678			32500			48833.90299	43069	72135.02647
Lotus							51657.5		75065								68377.14286
Land Rover		39699.5		71283.87097				48577									68067.08633
BMW	26699		55155	58536.11111			63814.07246		52445.25397						71832.11009	43266.66667	62162.55864
Alfa Romeo							64900		59400								61600
Cadillac				72551.06061			70400.5		45439.6	66572.22222					51178.5163	47364	56368.26515
Audi	2000			48634.54545			70029.89362		93586.57895						46391.87013	33894	54574.1215
Lexus			31566.66667	45042.48571			52451.66667		50823.6						48864.60606	31105	47549.06931
Genesis															46616.66667		46616.66667
Lincoln				50331.91176					2167.75	41205.45455					41665.16949	44950.83333	43560.01316
Infiniti				45686.31579			46669.04762		40291.66667						41076.00633		42640.27134
HUMMER				37749						34629.28571							36464.41176
Acura	17175.60714		51062.85714	42959.75806					39687.4						33614.2439	33560	35087.4878
GMC		8488.214286		37479.76836	23791.66667	21908.80952				39062.32692	27895.71795	25105	28555.71429	25182.90196			32695.74268
Volvo	26258.33333			45386.95652			40533.33333		2000						22289.73118	26271.42391	29724.68421
Buick				33996.34921			25617.85714		2059.333333			30005.90909			29568.64583	2053	29034.18947
Chevrolet	2000	13807.85714	18930.29412	33553.95876	20007.14286	8298.666667	62835	17716.66667	38939.16667	39255.74172	24170.16279	24934.28571	28555.71429	19824.84211	19882.64865	15825	29018.35005
Volkswagen	24134.62963		28416.21053	41699.1			27673.68675		2000			29239.67742			30795.79861	26385.64815	28978.52289
Nissan	2097.571429		24059.28571	34294.46281	21436.66667		39070.88889	43691.66667	35393.15663	32733.78378	20527.58	22962.22222		2212.666667	22604.23077	17500	28921.15245
Toyota	18950		22186.50794	40851.6			25777.86667		15615.28846	36845.82353	26251.30827	30038.73846		17592.66667	24800.27083	31742.4359	28846.5605
Ford	2000	16133.55172	19572.93103	42027.60577	19700	20605.59259	34762.2381		34101.07317	41566.13187	23808.16667	22587.17391	32836.45946	17797.80822	23258.65306	30066.01852	28525.18282
Saab	2000		2034.470588	41685			28755.81818								36775.86207	34149.09091	27879.80734
Chrysler	32935			35792.14286			25124.2		22502			29751.45161			26103.77895	26372.36842	26990.23784
Honda	17216.66667		26656.38889	28575.85714			36019.28571		21763.08219	34100.68182		36879			26027.47126		26655.14781
Kia			19379.04762	31533					20375.71429			32976.66667			23811.56627	20326.44737	25513.75546

Analysis_Task2 | Analysis_Task3 | Normalization of data in Task3 | Analysis_Task4 | Analysis_Task5 | Dashboard_Task1 | Dashboard_Task2

We can plot a clustered column chart from this table where each cluster in the chart for a particular brand would have the column of average MSRP of all body styles and sections would have different colors to identify each body style. Chart looks like below:

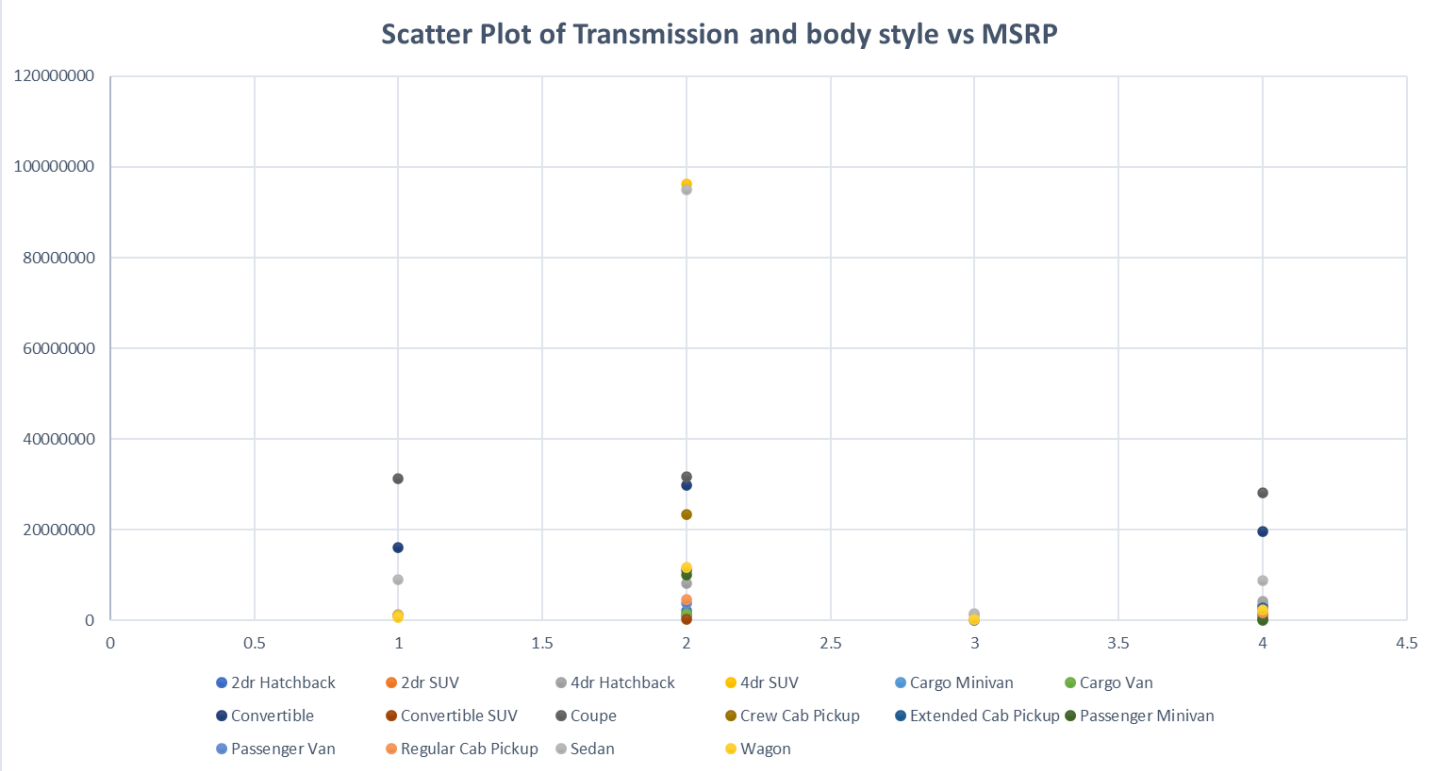


Task 3: How do the different features such as transmission type affect the MSRP, and how does this vary by body style?

To find the effect of transmission type on MSRP we have to create a pivot table and add body style as column and transmission type as row. We consider the average of MSRP to better visualize the data. We get the following pivot table.

Average of MSRP	Body Style							
Transmission Type	2dr Hatchback	2dr SUV	4dr Hatchback	4dr SUV	Cargo Minivan	Cargo Van	Convertible	Conve
AUTOMATED_MANUAL	27470.41667		29347.04545	40451.15385			129082.2339	
AUTOMATIC	20784.09901	24153.60606	23888.73529	41658.40017	20292.93103	17019.29762	95153.3131	
DIRECT_DRIVE	31800		32799.72973	49800				
MANUAL	12840.65556	9173.018519	17500.36364	17422.08791			64794.34437	
Grand Total	**16220.74634**	**14855.31034**	**22416.46757**	**40747.54467**	**20292.93103**	**17019.29762**	**88439.88633**	

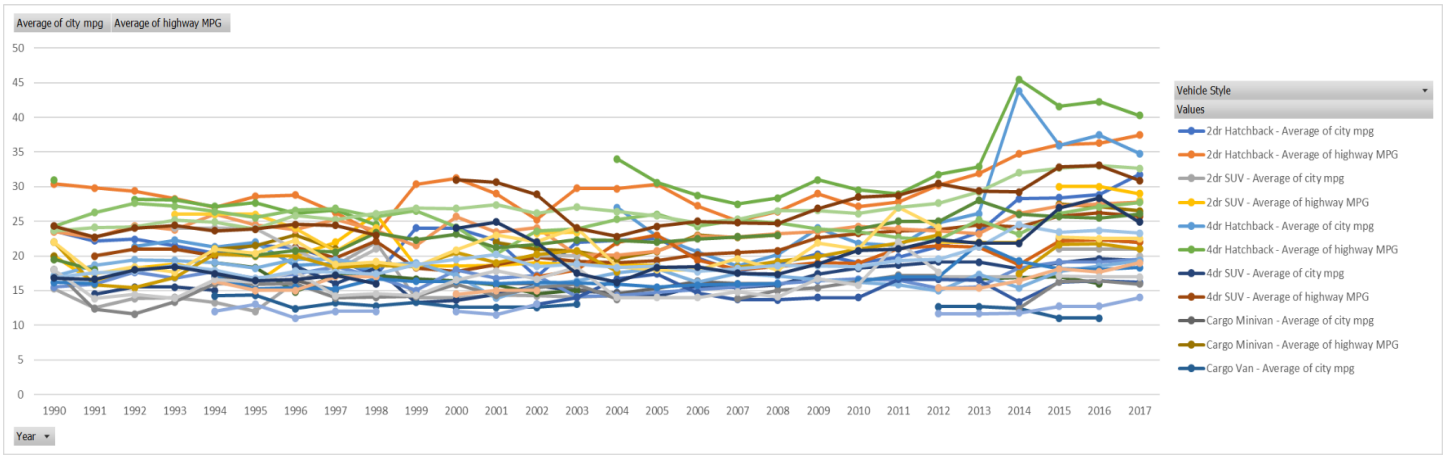
We copy down the contents of the pivot table to create a scatter plot. Scatter plot look like below:



Task 4: How does the fuel efficiency of cars vary across different body styles and model years? To find how fuel efficiency of a car varies across different body styles across different years we create a pivot table consisting of average highway MPG and City MPG across body styles as columns and years as rows. We get following table:

Body Style	2dr Hatchback		2dr SUV		4dr Hatchback		4dr SUV		Cargo Minivan	
Year	Average of city mpg	Average of highway MPG	Average of city mpg	Average of highway MPG	Average of city mpg	Average of highway MPG	Average of city mpg	Average of highway MPG	Average of city mp	
1990	23.6	30.4	15.25	20	22	31				
1991	22.16666667	29.83333333	12.5	16.25			14.5	20		
1992	22.39285714	29.39285714	13.85714286	18.28571429	21.33333333	28.16666667	15.5	21		
1993	21.48148148	28.25925926	14	18.85714286	22.25	28.125	15.5	21		
1994	20.42105263	27.05263158	13.25	17.625	21.28571429	27.14285714	15	20		
1995	21.6	28.6	12	16	22	27.66666667				16
1996	21.2	28.8	16.2	20	18.625	26.125	18.5	21.25		16
1997	19.5	26.25	18.66666667	22	18.88888889	26.66666667	16	19.7		
1998	17.2	23.2	22	26	18	24.5	18.22222222	22.11111111		
1999	24	30.33333333	14	18.5			13.3	18.3		
2000	24	31.22222222	14	18.5			13.6	17.73333333		
2001	22.28571429	29	14.33333333	18.66666667			14.45454545	18.72727273		
2002	17	25.25	14.25	19			15.73529412	19.79411765		
2003	22	29.75	14.08333333	18.75			14.97142857	19.22857143	15.166666	
2004	22.28571429	29.71428571	14.25	18.75	27	34	14.65306122	19.04081633	14	
2005	22.55555556	30.33333333	14.33333333	18.66666667	22.8	30.6	14.19047619	19.33333333	15.333333	
2006	19.66666667	27.25			20.58333333	28.75	15.58333333	20.19444444	16.333333	
2007	17.72727273	25.09090909			18.54545455	27.45454545	15.38888889	20.46296296		
2008	18.85714286	26.42857143			20.16666667	28.33333333	15.78125	20.765625		
2009	20.25	29			24	31	17.39784946	22.59139785		
2010	19	27.125			21.8125	29.5	18.21818182	23.25454545		
2011	19.83333333	27.83333333			21.44827586	28.93103448	18.68055556	23.58333333		
2012	21.35714286	30.21428571			24.78571429	31.76190476	19.15555556	23.84444444		
2013	23.45454545	31.90909091			26.11764706	32.8627451	19.12280702	24.47368421		
2014	28.25	34.75			43.82978723	45.46808511	18.15702479	24.2231405		
2015	28.41176471	36.10294118	21	30	35.95138889	41.57638889	19.04283054	25.76350093	22	
2016	28.85714286	36.26530612	21	30	37.456	42.28	19.61025641	26.1965812	22.333333	
2017	31.75	37.4375	21	29	34.75630252	40.29411765	19.36016949	25.70974576		
Grand Total	24.0804878	31.37804878	14.85057471	19.55172414	32.08898944	37.81146305	18.48456155	24.508028	18.517241	

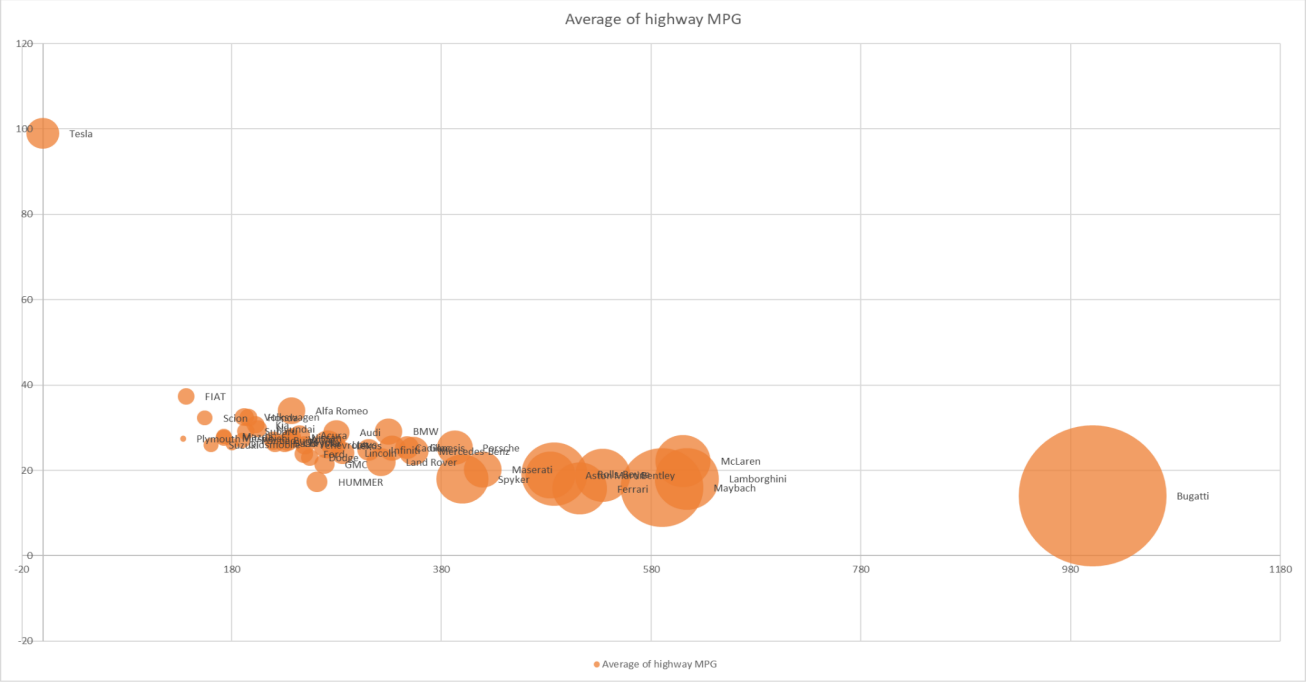
We use this table to create a line plot with markers to visualize the data as a timeline and across different body types.



Task 5: How does the car's horsepower, MPG, and price vary across different Brands?
To find relationships between a car's horsepower and MPG and price across different brands we can create a bubble plot to better visualize, for this we find average MPG , price and horsepower across each brand and create a pivot table. This pivot table looks like below:

Row Labels	Average of Engine HP	Average of city mpg	Average of highway MPG	Average of MSRP
Acura	244.9634146	20.00406504	28.2195122	35087.4878
Alfa Romeo	237	24	34	61600
Aston Martin	483.7582418	12.56043956	18.93406593	198123.4615
Audi	280	19.63551402	28.92834891	54574.1215
Bentley	533.8513514	11.55405405	18.90540541	247169.3243
BMW	329.6203704	20.70061728	29.12654321	62162.55864
Bugatti	1001	8	14	1757223.667
Buick	220.0105263	18.78421053	27.01052632	29034.18947
Cadillac	332.7954545	17.36111111	25.24494949	56368.26515
Chevrolet	249.4837512	19.12070566	25.93221913	29018.35005
Chrysler	230.5351351	17.74054054	26.38378378	26990.23784
Dodge	254.5984848	16.45643939	22.99810606	24900.33523
Ferrari	511.9565217	10.56521739	15.72463768	238218.8406
FIAT	136.6129032	30.64516129	37.33870968	22670.24194
Ford	248.7730061	17.89815951	23.87730061	28525.18282
Genesis	347.3333333	16.33333333	25.33333333	46616.66667
GMC	268.2949791	15.79916318	21.47698745	32695.74268
Honda	195.8637413	25.2147806	32.39953811	26655.14781
HUMMER	261.2352941	13.52941176	17.29411765	36464.41176

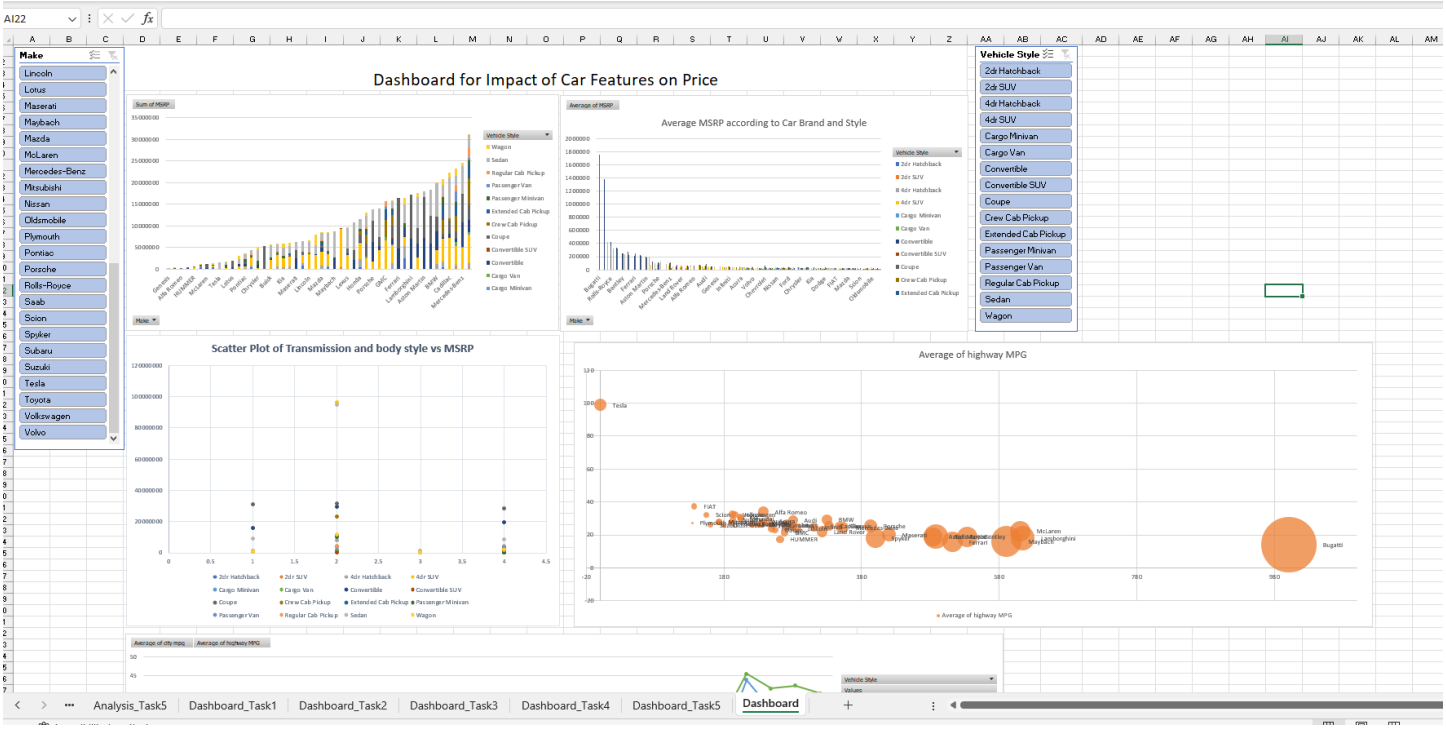
We copy this contents to create a table to create a bubble plot.
On the X axis there would be average horsepower and on the Y axis average MPG. Each bubble would represent each car brand and would be labeled to better identify the brands



Making Dashboard:

We have created each chart to visualize different parameters and relationships and trends, we can now create a dashboard by combining all these charts into one single worksheet to have a better understanding of the data. We copy all these charts into one worksheet and add slicers to change parameters which are shown the data point for. We add two slicers in the worksheet, make, and vehicle style, which represent car brand and body style. We then make connections with these slicers with all the charts to make the dashboard functional. Now we can easily find different trends and relationships between price and parameters.

Dashboard looks like below:



Results:

While working on this project, I have gained a better understanding of Impact of Car Features on Price and Profitability as well as popularity of the Car. I have improved my understanding of Advanced Excel methodologies. By analyzing Car features Data, I was able to provide insights on various aspects such as Features most affecting MSRP, Outliers in the Data, relation between Engine HP and MSRP, Regression Analysis, average MSRP across different brands and relation between no. of cylinders and fuel efficiency. I was also able to create different visualizations to improve data understanding and create a dashboard for ease of understanding between various parameters in Car Features.

This project has helped me enhance my Excel skills, particularly in data visualization and creating pivot tables and charts to derive meaningful insights. It has also improved my ability to interpret data and provide actionable recommendations based on the analysis.