Analysing the Impact of Car Features on Price and Profitability

By Atinderbir Pannu

DESCRIPTION

The project delves into analyzing the correlation between car features, consumer demand, and pricing in the automotive sector. Through regression analysis on a comprehensive dataset, it aims to pinpoint the most sought-after features influencing car prices. By considering shifting consumer preferences like the surge in electric and hybrid vehicles, the project seeks to advise manufacturers on strategic pricing and product development to bolster their market position and profitability effectively.

Data Description

- Make: the make or brand of the car
- Model: the specific model of the car
- Year: the year the car was released
- **Engine Fuel Type**: the type of fuel used by the car (gasoline, diesel, etc.)
- **Engine HP:** the horsepower of the car's engine
- **Engine Cylinders:** the number of cylinders in the car's engine
- **Transmission Type**: the type of transmission (automatic or manual)
- **Driven_Wheels:** the type of wheels driven by the car (front, rear, all)
- **Number of Doors:** the number of doors the car has
- Market Category: the market category the car belongs to (Luxury, Performance, etc.)
- **Vehicle Size:** the size of the car
- **Vehicle Style:** the style of the car (Sedan, Coupe, etc.)
- **Highway MPG:** the estimated miles per gallon the car gets on the highway
- **City MPG:** the estimated miles per gallon the car gets in the city
- **Popularity:** a ranking of the popularity of the car (based on the number of times it has been viewed on Edmunds.com)
- MSRP: the manufacturer's suggested retail price of the car

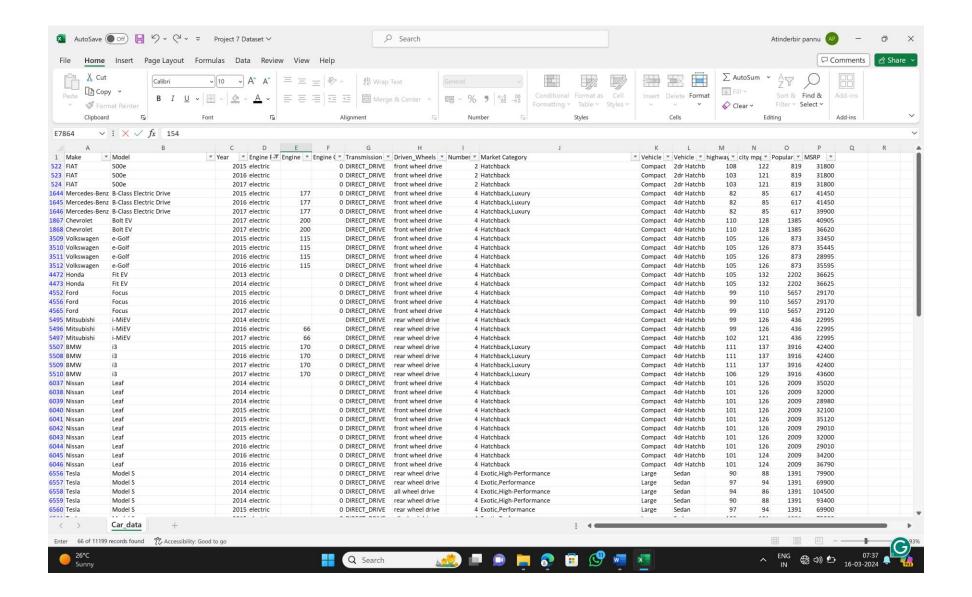
TECH-USED

- Ms-Excel 2021
- Ms-Powerpoint 2021

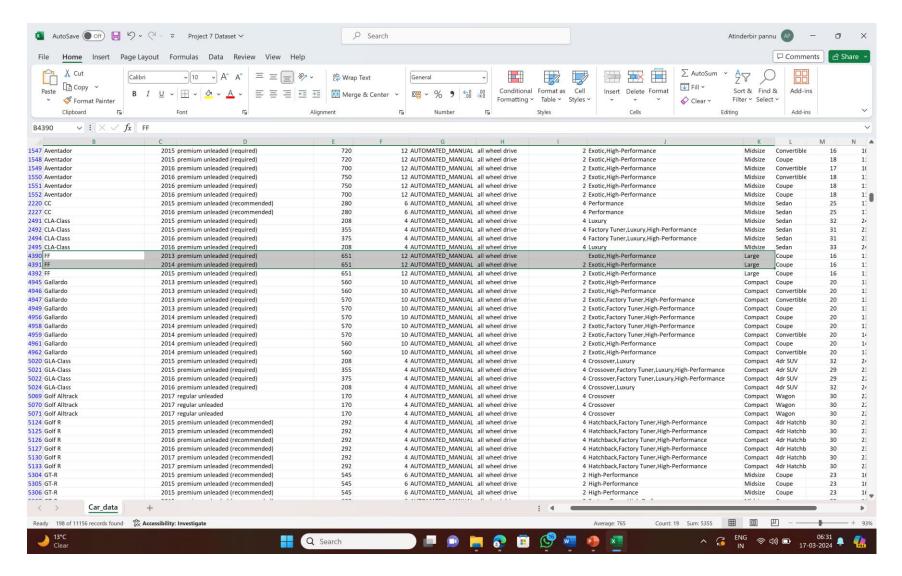
Approach

- The initial steps of the project involved data cleaning, including the removal duplicate entries.
- Missing values were then addressed through imputation, utilizing the median for numerical data (Engine HP) and categorical values (Number of Doors and Engine Cylinders) where appropriate.
- Furthermore, data cleaning was conducted, particularly for car model entries containing dates as values, which were deemed unsuitable for analysis and thus removed from the dataset.
- Pivot tables and Excel functions aided in extracting insights, while various charts visually represented the findings for enhanced understanding of the dataset.

Missing values in the data imputed by using median



Missing data that is imputed using the dataset (categorically)



INSIGHTS

These are the insights that have been obtained from the dataset:

- There exists a direct correlation between Engine HP (horsepower) and MSRP (manufacturer's suggested retail price), implying that vehicles with higher horsepower generally command higher prices.
- Conversely, as the number of cylinders increases, Highway MPG (miles per gallon) tends to decrease, showcasing an inverse relationship between these variables.
- Higher the number of cylinders typically corresponds to a higher manufacturer's suggested retail price (MSRP), indicating a direct correlation between the two factors in the automotive industry.
- Results for the obtained insights are given below:

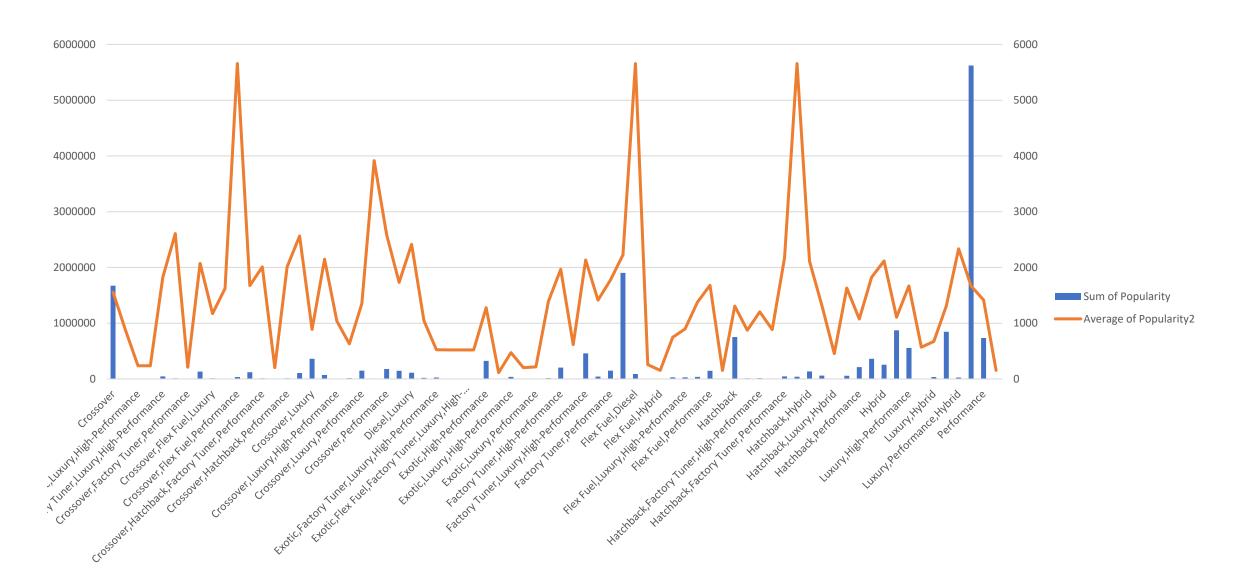
Task 1.A: Create a pivot table that shows the number of car models in each market category and their corresponding popularity scores.

Row Labels	Count of Model	Sum of Popularity
Crossover	1075	
Crossover, Diesel	7	6111
Crossover,Exotic,Luxury,High-Performance	1	
Crossover, Exotic, Luxury, Performance	_1	
Crossover, Factory Tuner, Luxury, High-Performance	26	
Crossover, Factory Tuner, Luxury, Performance	5	
Crossover, Factory Tuner, Performance	4 64	
Crossover,Flex Fuel Crossover,Flex Fuel,Luxury	10	132720 11732
Crossover, Flex Fuel, Luxury, Performance	6	
Crossover, Flex Fuel, Performance	6	
Crossover.Hatchback	72	
Crossover, Hatchback, Factory Tuner, Performance	6	12054
Crossover, Hatchback, Luxury	7	1428
Crossover, Hatchback, Performance	6	12054
Crossover, Hybrid	42	
Crossover,Luxury	406	
Crossover,Luxury,Diesel	34	73080
Crossover, Luxury, High-Performance	.9	
Crossover, Luxury, Hybrid	24	
Crossover, Luxury, Performance	112	
Crossover, Luxury, Performance, Hybrid	2	
Crossover,Performance	69	178431 145396
Diesel Diesel,Luxury	84 47	113557
Exotic,Factory Tuner,High-Performance	21	
Exotic,Factory Tuner,Luxury,High-Performance	51	
Exotic, Factory Tuner, Luxury, Performance	3	
Exotic, Flex Fuel, Factory Tuner, Luxury, High-Performanc		
Exotic,Flex Fuel,Luxury,High-Performance	11	
Exotic, High-Performance	254	
Exotic, Luxury	12	1352
Exotic,Luxury,High-Performance	77	36423
Exotic,Luxury,High-Performance,Hybrid	1	
Exotic,Luxury,Performance	36	
Exotic,Performance	10	
Factory Tuner, High-Performance	104	204510
Factory Tuner, Luxury	2	
Factory Tuner, Luxury, High-Performance	215	
Factory Tuner,Luxury,Performance Factory Tuner,Performance	31 84	43816 149020
Flex Fuel	855	
Flex Fuel Diesel	16	
Flex Fuel, Factory Tuner, Luxury, High-Performance	1	
Flex Fuel Hybrid	2	
Flex Fuel,Luxury	39	
Flex Fuel Luxury High-Performance	32	28746
Flex Fuel, Luxury, Performance	28	38642
Flex Fuel, Performance	87	146201
Flex Fuel,Performance,Hybrid	2	
Hatchback	574	
Hatchback,Diesel	14	
Hatchback, Factory Tuner, High-Performance	13	
Hatchback, Factory Tuner, Luxury, Performance	9	
Hatchback,Factory Tuner,Performance Hatchback,Flex Fuel	21 7	45648 39599
Hatchback,Hybrid	64	
Hatchback, Luxurv	45	
Hatchback,Luxury,Hybrid	3	
Hatchback, Luxury, Performance	36	
Hatchback, Performance	198	
High-Performance	198	
Hybrid	121	
Luxury	788	
Luxury,High-Performance	334	557118
Luxury,High-Performance,Hybrid	12	
Luxury,Hybrid	52	
Luxury,Performance	647	847616
Luxury,Performance,Hybrid	11	
NA Performance	3376 520	5620476 735909
Performance	520	/30309

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

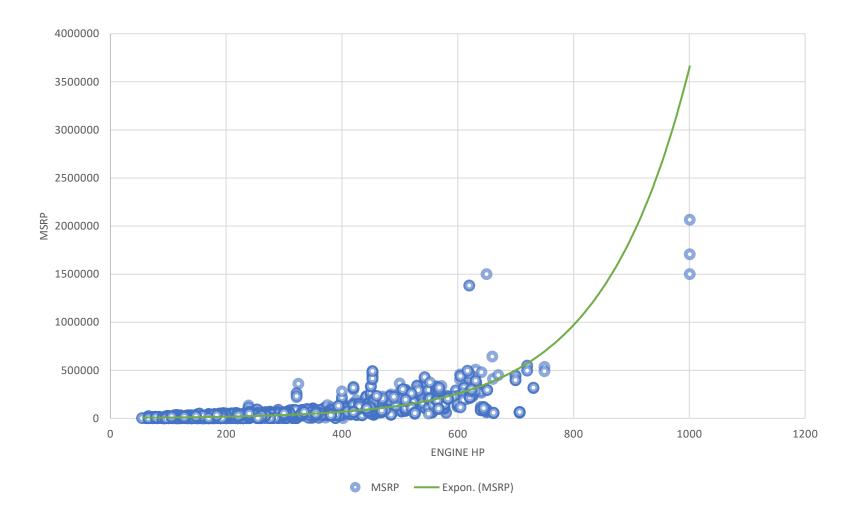
Row Labels	▼ Sum of Popularity	Average of Popularity2
Crossover	1672881	1556.168372
Crossover, Diesel	6111	873
Crossover, Exotic, Luxury, High-Performance	238	238
Crossover, Exotic, Luxury, Performance	238	238
Crossover, Factory Tuner, Luxury, High-Performance	47410	1823.461538
Crossover, Factory Tuner, Luxury, Performance	13037	2607.4
Crossover, Factory Tuner, Performance	840	210
Crossover,Flex Fuel	132720	2073.75
Crossover,Flex Fuel,Luxury	11732	1173.2
Crossover, Flex Fuel, Luxury, Performance	9744	1624
Crossover, Flex Fuel, Performance	33942	5657
Crossover, Hatchback	120650	1675.694444
Crossover, Hatchback, Factory Tuner, Performance	12054	2009
Crossover, Hatchback, Luxury	1428	204
Crossover, Hatchback, Performance	12054	2009
Crossover,Hybrid	107662	2563.380952
Crossover,Luxury	361021	889.2142857
Crossover,Luxury,Diesel	73080	2149.411765
Crossover, Luxury, High-Performance	9335	1037.222222
Crossover, Luxury, Hybrid	15142	630.9166667
Crossover, Luxury, Performance	151098	1349.089286
Crossover, Luxury, Performance, Hybrid	7832	3916
Crossover, Performance	178431	2585.956522
Diesel	145396	1730.904762
Diesel,Luxury	113557	2416.106383
Exotic, Factory Tuner, High-Performance	21974	1046.380952
Exotic, Factory Tuner, Luxury, High-Performance	26674	523.0196078
Exotic, Factory Tuner, Luxury, Performance	1560	520
Exotic, Flex Fuel, Factory Tuner, Luxury, High-Performan	ce 6760	520
Exotic,Flex Fuel,Luxury,High-Performance	5720	520
Exotic, High-Performance	325132	1280.047244
Exotic,Luxury	1352	112.6666667
Exotic, Luxury, High-Performance	36423	473.025974
Exotic, Luxury, High-Performance, Hybrid	204	204
Exotic, Luxury, Performance	7813	217.0277778
Exotic, Performance	13910	1391
Factory Tuner.High-Performance	204510	1966.442308

Chart for market category and popularity

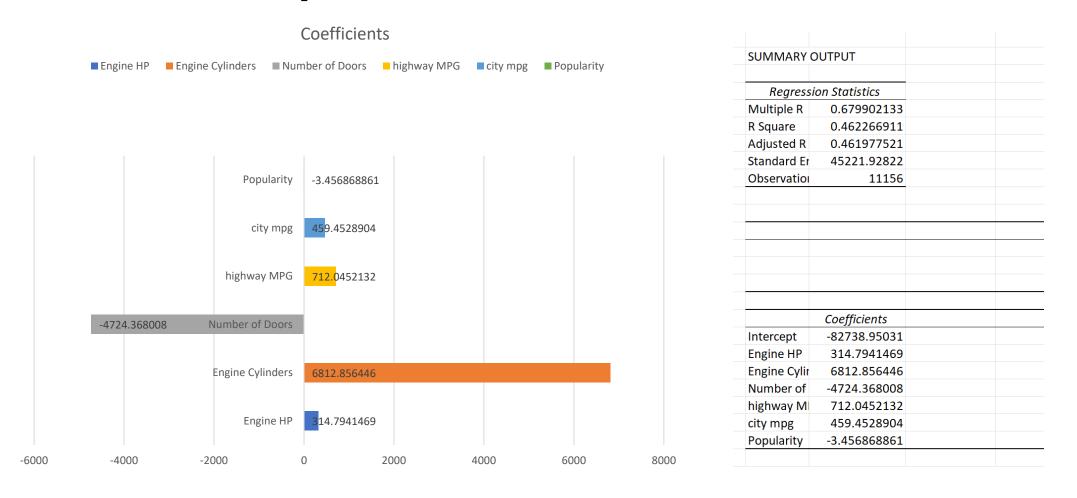


Task 2: Create a scatter chart that plots engine power on the x-axis and price on the y-axis. Add a trendline to the chart to visualize the relationship between these variables.

The graph displays a positive slope on the trendline, indicating a direct relationship between a car's engine power and its price. In simpler terms, cars with more powerful engines tend to have higher prices.



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.

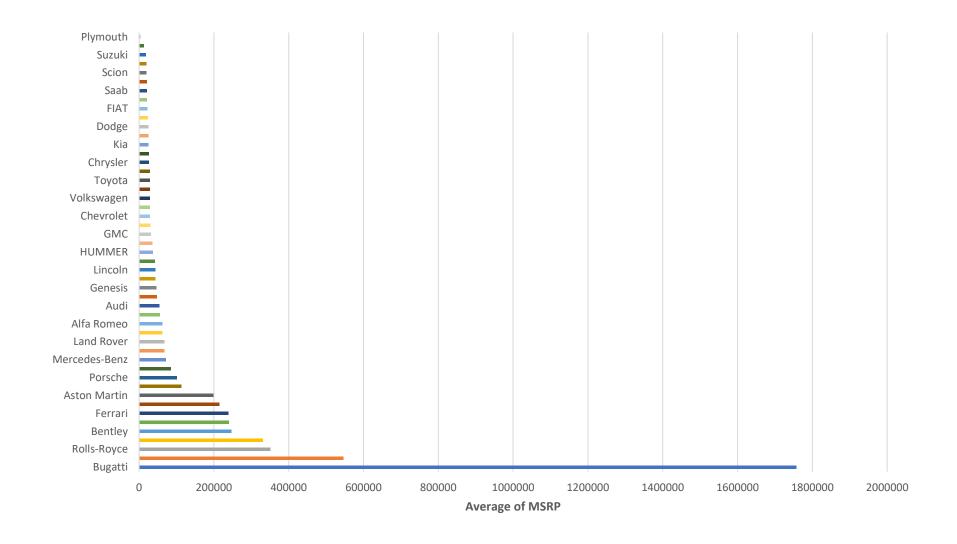


• More the number of cylinders typically corresponds to a higher manufacturer's suggested retail price (MSRP), indicating a direct correlation between the two factors in the automotive industry.

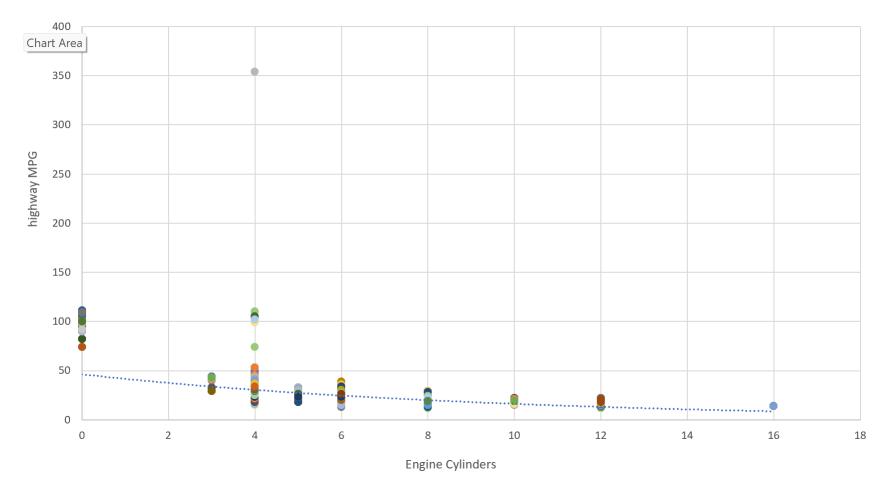
Task 4.A: Create a pivot table that shows the average price of cars for each manufacturer.

Row Labels	Average of MSRP
Acura	35087.4878
Alfa Romeo	61600
Aston Martin	198123.4615
Audi	54574.1215
Bentley	247169.3243
BMW	62218.47988
Bugatti	1757223.667
Buick	29034.18947
Cadillac	56368.26515
Chevrolet	29074.72576
Chrysler	26722.96257
d	44100
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-Ber	
Mitsubishi	21340.5625
Nissan	28921.15245
Oldsmobile	12843.79545
Plymouth	3296.873239
Pontiac	19800.0442
Porsche	101622.3971
	351130.6452
Rolls-Royce	
Saab	20657.71212
Scion	19932.5
Spyker	214990
Subaru	24240.67364
Suzuki	18026.4152
Tesla	85255.55556
Toyota	28846.5605
Volkswagen	28978.52289
Volvo	29724.68421

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



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.



• More cylinders in the engine lead to lower highway MPG, as shown by a negative correlation in the analysis. The trendline slope on the graph confirms this relationship.

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

	Engine Cylinders ighway MPG
Engine Cylinders	
highway MPG	=CORREL(A:A,B:B)
	CORREL(array1, array2)

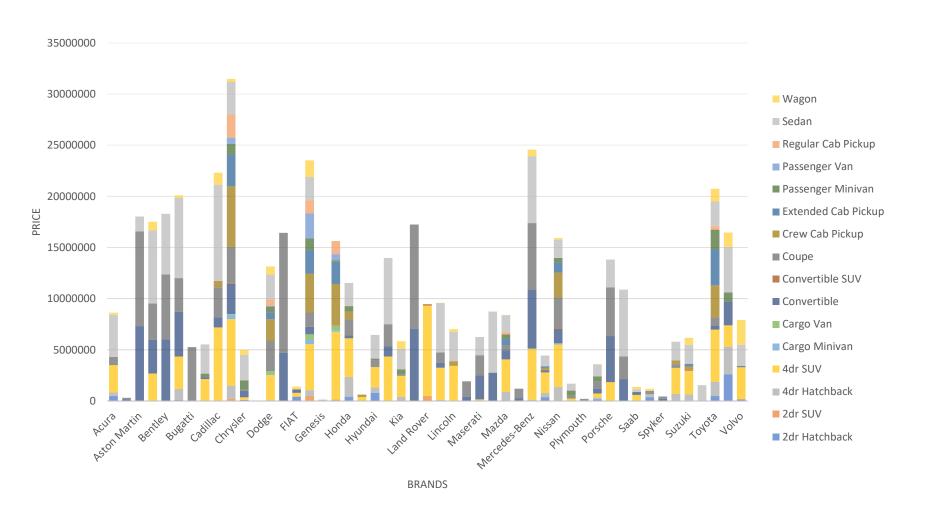
	Engine Cylinders	highway MPG
Engine Cylinders	1	
highway MPG	-0.604333097	1

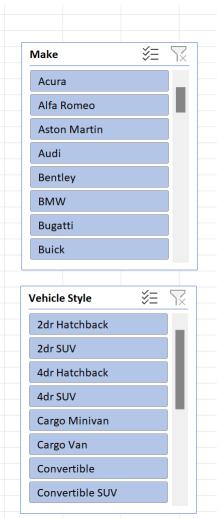
Dashboards

Part-2

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

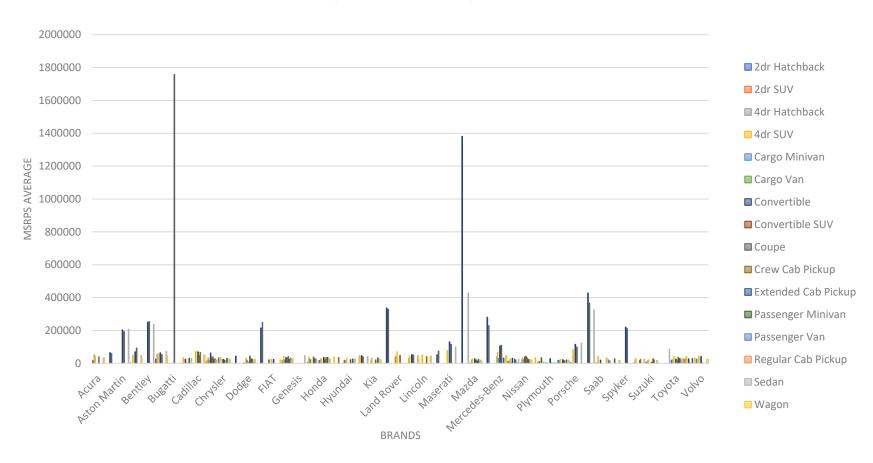
		zar suv 4			cargo ivilnivan C	argo van (Lonvertible	Convertible SUV	•	стем сар искир	Extended Cab Pickup	Passenger Minivar	rassenger van Re	винаг саю Ріскир		Wagor
cura	480917		357440	2663505			400000		793748						4134552	2013
lfa Romeo							129800		178200						1440705	
ston Martin	4000			2674000			7321655		9258845						1448735	
udi	4000			2674900			3291405		3556290						7144348	
entley	7,000		4400400	24 50050			6012870		6356760						5920900	
MW	80097		1103100	3100950			4359071		3304051						7829700	259
ugatti							470005		5271671			22225				
uick				2141770			179325		18534	500450		330065)		2838590	
adillac				7182555			985607		2953574	599150					9416847	
nevrolet		193310	1287260		420150	74688	2953245	106300		5927617	3117951			2260032	3177797	
nrysler	98805			250545			630105		114510			922295			2479859	501
							44100									
odge	38000	12000	16000	2462875	60520	338497	6000		2973842	2072780	684682	557425	70708	653408	2409585	793
errari							4723811		11713289							
AT	420715			369305			327965									287
ord	24000	467873	567615	4482771	415630	556351	730007		1398144	3782518	2285584	1179285	2429898	1299240	2279348	
enesis															139850)
MC		128319		6633919	142750	460085				4062482	2175866	150630	599670	1284328		
onda	413200		1919260	3800589			252135		1588705	750215		553185	5		2264390)
UMMER				377490						242405						
yundai	789650		528880	1994390					685920			133075	5		2323987	•
finiti				4340200			980050		2175750						6490009	
ia			406960	2049645					142630			494650)		1976360	772
amborghini							7064450		10177050							
and Rover		476394		8839200				145731								
exus			94700	3152974			472065		1016472						4837596	31
incoln				3422570					17342	453260					2854855	269
otus							413260		1501300							
Maserati				155000			2342963		1972284						1782400)
1aybach							2762750								5976800	
1azda	18000	12000	853180	3175515			870505		541879		580033	443130)	265486	1618571	
1cLaren							280225		918800							
lercedes-Benz			122800	4974610	28950		5753964		6473107			32500)		6543743	646
litsubishi	370169			2009807	2000		209893		0170207	240210	134360			8000		
issan	14683			4149630	128620		1406552	131075	2937632	2422300	1026379				1763130	
ldsmobile	14000		1047020	238150	120020		2000	151075	276015	2422300	102007	492055		15514	667161	
ymouth	40000		14000	250150			85631		8000			31688			38759	
ontiac	163505		162975	401550			473481		663715			541192			1156535	
orsche	28827		1029/3	1815200			4504586		4758533			541154	<u>-</u>		2713500	
olls-Royce	2002/			1013200			2141365		2204675						6539010	
	12000		34586	541905			277228		22040/3						266850	
aab				341303			2//228		220210							
cion	366325		282470				210000		330210						32500	184
pyker	10000		670650	2520000			219990		209990	205075					1000110	
ubaru	12000	40000		2539900				400:5:	354476	365975	2				1833110	
ızuki	44496	12000	584387	2303493				120194		304131	259659	,			1852967	
esla															1534600	
oyota	473750		1397750				386668		811995	3131895	3491424			369446	2380826	
olkswagen	2606540		2699540				2296916		6000			906430)		4434595	
olvo	157550			3131700			121600		6000						2072945	241

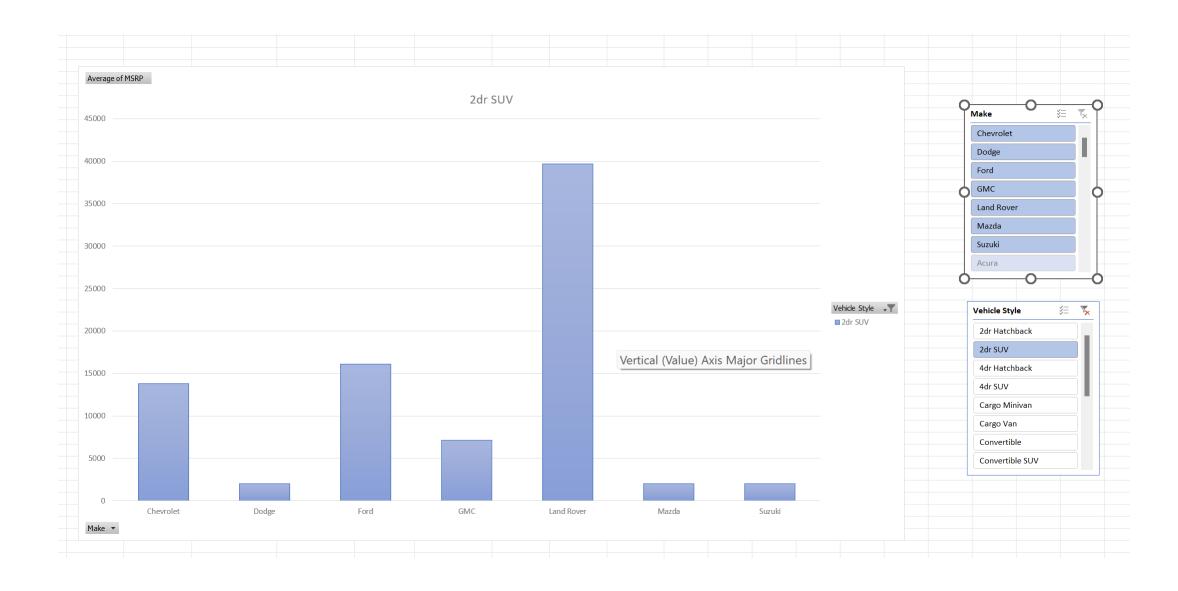




Task 2: Which car brands have the highest and lowest average MSRPs, and how does this vary by body style?

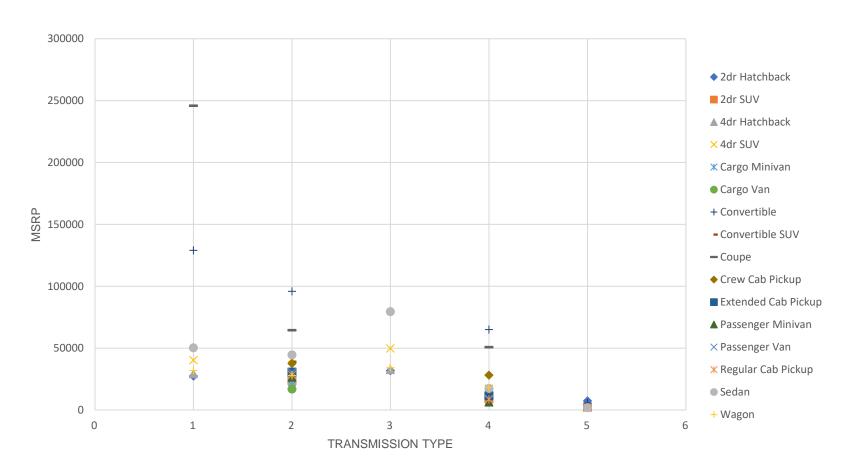
	Brand name	Average MSRP
Maximum	Bugatti	1757223.667
Minimum	Plymouth	3296.873239



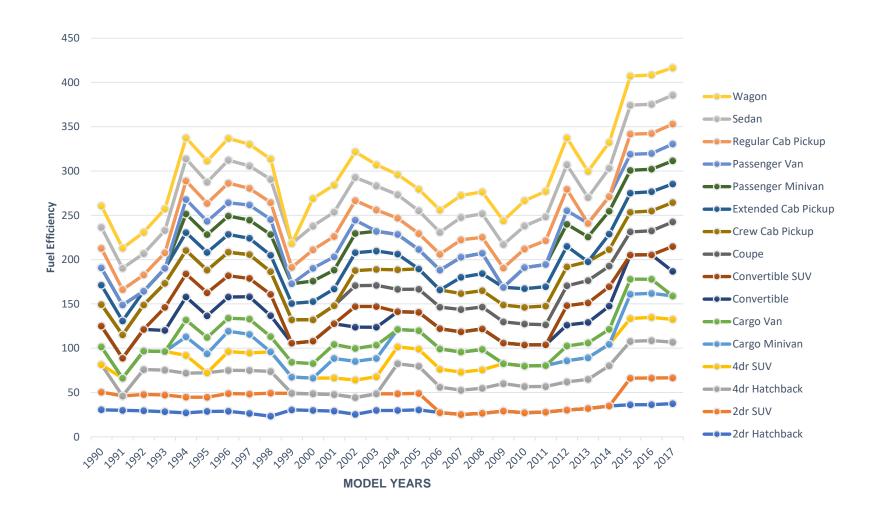


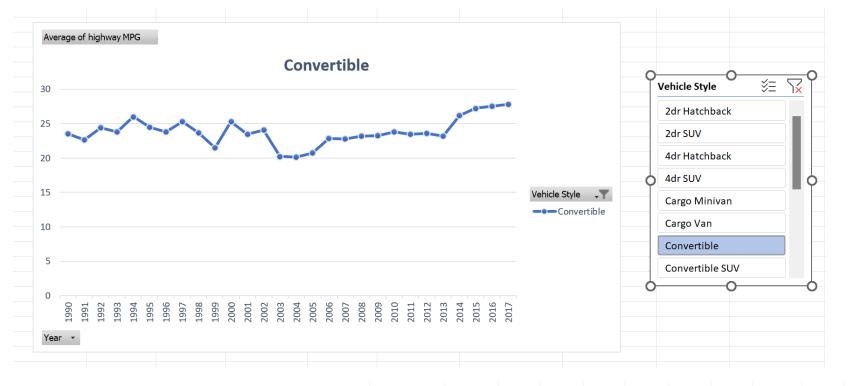
Task 3: How do the different feature such as transmission type affect the MSRP, and how does this vary by body style?

Average of MSRP	Column Labels 💌															
Row Labels	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
AUTOMATED_MANUAL	27470.41667		29347.04545	40451.15385			129082.2339		245977.4252						50385.39326	31985.2777
AUTOMATIC	20784.09901	24153.60606	23888.73529	41638.26534	20315.59322	17019.29762	95940.57468	38925.5	64523.41955	37718.95307	30711.45251	26589.50919	30578.06612	28536.8239	44712.05697	27925.51
DIRECT_DRIVE	31800		32799.72973	49800											79512.25	342
MANUAL	12840.65556	9173.018519	17500.36364	17422.08791			65036.0602	9594.8	50901.4973	28233.10811	11553.29707	6510)	8759.454054	17444.12625	18100.5
UNKNOWN	7361.5	2371					5783.5		2000					2000	2000	



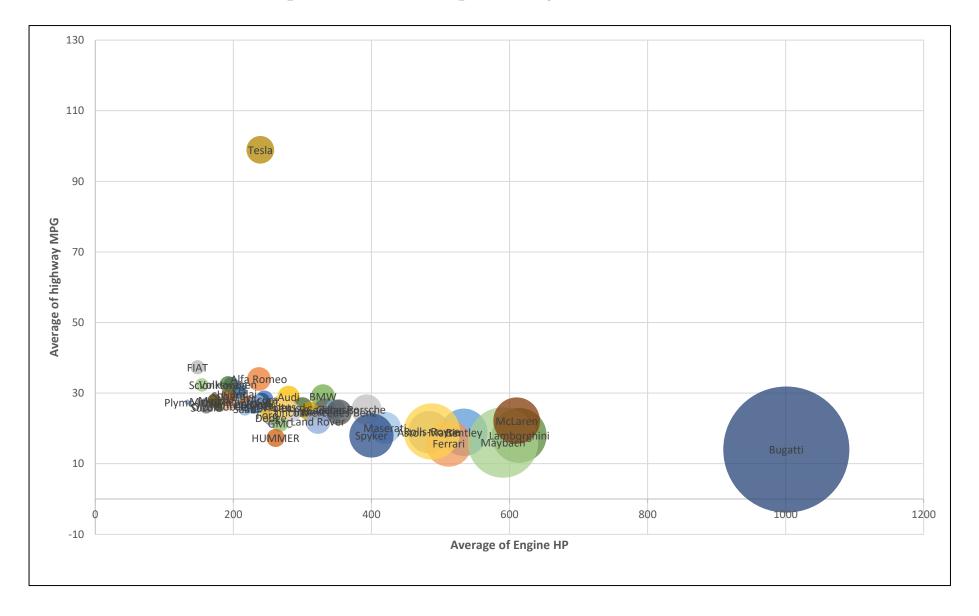
Task 4: How does the fuel efficiency of cars vary across different body styles and model years?

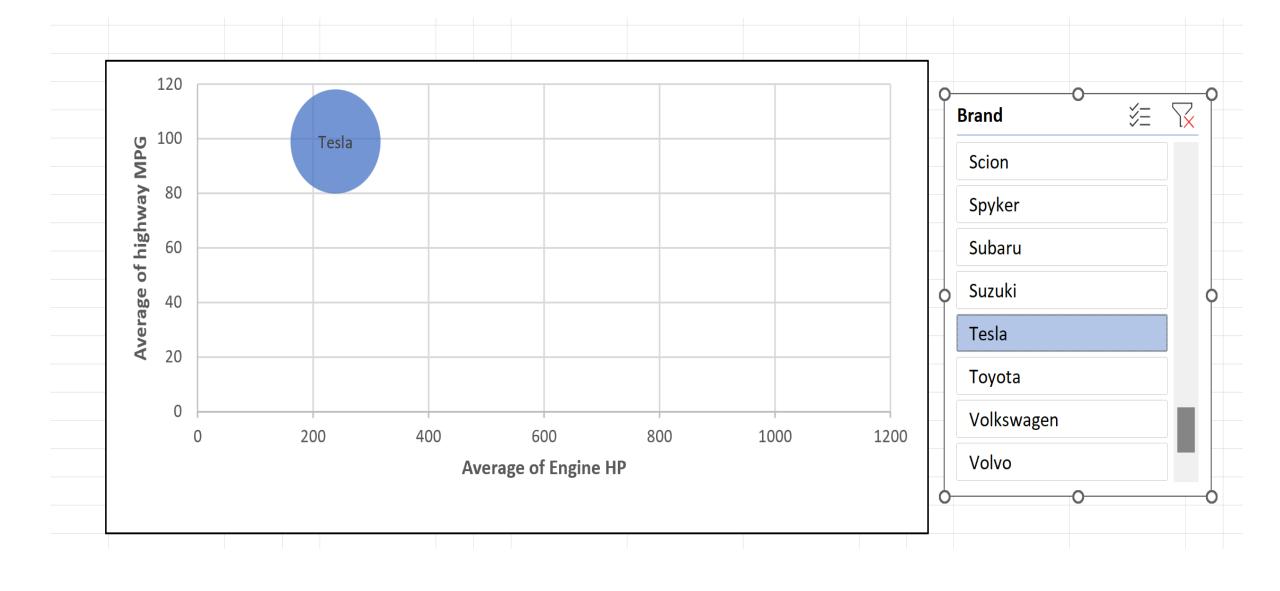




Row Labels 1990 1991 1992 1993 1994 1995 1996 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011		30.4 29.83333333 29.39285714	20 16.25			Cargo Minivan		Convertible	Convertible SUV	Counc	Comment Malana	Establish Nation		Decrease Ven	Dogular Cab Bishup	Sodon	
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011	▼ ;	30.4 29.83333333 29.39285714	20 16.25	31		•		Convertible	Convertible SUV	Counc	Control of the production	Englanded Calcottation	B	December Man	Deguler Cele Dieleum	Codon	****
1991 1992 1993 1994 1995 1996 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2010		29.83333333 29.39285714	16.25			20				coupe	Crew Cab Pickup	Extended Cab Pickup	Passenger Minivan	Passenger van	Regular Cab Pickup	Seuari	Wagon
1992 1993 1994 1995 1996 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011		29.39285714				20		23.5		24.27272727		22	19.5		22	23.65384615	24.3
1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011			40 20574420		20			22.625		26.25		15.83333333	18		17.28571429	24.11111111	22.72727273
1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011		28.25925926	18.285/1429	28.16666667	21			24.375		27.63636364		15.4			18.42857143	24.175	24
1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011		LOILOSLOSLO	18.85714286	28.125	21			23.81818182	26	27.13636364		16.90909091			17.625	25.2	24.41666667
1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011		27.05263158	17.625	27.14285714	20	21	19	26	26	26.42857143		20.28571429	21	16.33333333	21	24.90625	23.63636364
1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011		28.6	16	27.66666667		21.5	18.33333333	24.5	26	25.51724138		20	20.1	15	20.375	23.88461538	23.88888889
1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011		28.8	20		21.25	23	14.8	23.8	24	26.6		20	20.7777778	15	22.2	25.83333333	24.57142857
1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011		26.25	22	26.66666667	19.7	21	17.2	25.28571429	20.66666667	26.92857143		18.35714286	20.5555556	17	18.78571429	25.42105263	24.4
2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011		23.2	26	24.5	22.11111111		17.2	23.66666667	24	25.64285714		18.625	23.33333333	17	19.15151515	26.1	23
2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011		30.33333333	18.75		18.3		16.66666667	21.5		26.5		18.42307692	22.33333333		18.42857143	26.875	
2002 2003 2004 2005 2006 2007 2008 2009 2010 2011		29.72727273	18.75		17.73333333		16.4	25.28571429		24.16666667		20.5	23.16666667	14.5	20.83333333	26.86363636	31
2003 2004 2005 2006 2007 2008 2009 2010 2011		29	18.66666667		18.72727273	22	15.8	23.4375		20.29411765		19	21.2	15	23	27.37735849	30.625
2004 2005 2006 2007 2008 2009 2010 2011		25.25			19.79411765	21		24.07142857	23.28571429			20.2222222	21.6875		22.06666667		28.88888889
2005 2006 2007 2008 2009 2010 2011		29.75				20.66666667	15	20.23076923	23.4	23.87878788			22.2972973			27.05769231	24
2006 2007 2008 2009 2010 2011		29.71428571	18.75		19.04081633	19.6		20.1		25.26666667						26.36231884	22.8
2007 2008 2009 2010 2011		30.33333333			19.33333333	20.85714286		20.72727273		26			21.88888889		18	25.75409836	
2008 2009 2010 2011		27.25			20.19444444	23		22.85714286		24.25925926			22.45		18	24.75	
2009 2010 2011		25.09090909			20.46296296	22.66666667		22.76		25.2		18.38983051	22.75			25.30769231	24.8
2010 2011		26.42857143		28.33333333		23		23.19047619		24.78947368	18.43181818	19.2222222	23			26.52173913	
2011		29			22.59139785			23.25		23.89473684	19.02941176					26.47826087	26.84
		27.125			23.25454545			23.82352941		23.52173913	18.94594595	21	23.85714286			25.96666667	28.5
		27.83333333			23.58333333			23.4375		22.67857143						26.84057971	
2012		30.21428571			23.8444444			23.57692308		22.37142857	21.43333333	23.0625		15.33333333	24.125	27.60493827	
2013		31.90909091			24.47368421			23.18181818		25.18604651	21.31818182		28				29.39473684
2014		34.75		45.46808511				26.15789474	22	23.16883117	18.71428571	17.4				31.99065421	29.25
2015		36.10294118	30	12107 000000		27.5		27.22377622		26.16931217	22.12121212	21.65934066	25.65384615			32.64007092	
2016		36.26530612 37.4375	30			27.11111111	16	27.51666667		27.10989011	22.37593985	21.78409091		17.71428571	22.52941176		33.08333333
2017			20	40.29411765	25.73739496	26.5		27.80263158	28	27.71724138	21.96153846	20.98684211	26.05555556	19	22.52941176	32.61589404	30.86486486

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





THANK YOU

Excelsheet <u>link</u>