Project: Analyzing the Impact of Car Features on Price and Profitability

OVERVIEW OF THE PROJECT AND PURPOSE OF THE PROJECT:

To analyze a dataset related to the automotive industry and use data analysis techniques such as regression analysis and market segmentation to identify the factors that drive consumer demand for cars, and how a car manufacturer can optimize pricing and product development decisions to maximize profitability while meeting consumer demand. The analysis will focus on identifying the most popular product features and market categories among consumers, as well as the most profitable ones for the manufacturer, and using this information to develop a pricing strategy that balances consumer demand with profitability. The ultimate goal is to help the manufacturer improve its competitiveness in the market and increase its profitability over time.

The purpose of this project is to help car manufacturers optimize their pricing and product development decisions by analyzing consumer demand, market categories, and product features. By identifying which features and categories are most popular among consumers and most profitable for the manufacturer, the manufacturer can develop a pricing strategy that maximizes profitability while meeting consumer demand. The ultimate goal is to improve the manufacturer's competitiveness in the market and increase its profitability over time.

The business problem or question that the project aims to address is how a car manufacturer can optimize pricing and product development decisions to maximize profitability while meeting consumer demand. The project will analyze a dataset related to the automotive industry to identify the factors that drive consumer demand for cars, such as product features and market categories, and use this information to develop a pricing strategy that balances consumer demand with profitability. The ultimate goal is to help the manufacturer improve its competitiveness in the market and increase its profitability over time.

DESCRIPTION OF THE DATASET:

overview of the dataset:

- Number of observations: 11,159
- Number of variables: 16

File type: CSV (Comma Separated Values) The variables in the dataset are: \

- 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

DATA CLEANING AND PREPROCESSING STEPS:

The columns which had null values

Engine Fuel Type - 3
Engine HP - 69
Engine Cylinders - 30
Number of Doors - 6
Market Category - 3742
Total null values - 3850

- During the data cleaning process, the rows with null values in Market Category and Engine HP were removed.
- Null values in “Number of Doors” were filled manually by the information provided on the internet
- The null values in Engine Cylinders were filled based on the assumptions
    - The number of cylinders was filled with **zero** when its corresponding value in ‘Engine Fuel Type’ was Electric
    - The number of cylinders was filled with **Four** when its corresponding value in ‘Engine Fuel Type’ was 'premium unleaded (required)' and 'regular unleaded'

APPROACH OF THIS PROJECT:

ANALYTICAL METHODS AND TECH-STACK USED IN THIS PROJECT:

- **REGRESSION ANALYSIS**
    - Regression analysis is used to identify the relationship between a dependent variable and one or more independent variables. It helps analysts to understand how changes in one variable affect another variable and to make predictions about future outcomes based on historical data.

- **STATISTICS**
    - Statistics method used to analyze data, draw conclusions from the data, and make predictions about future outcomes based on probability distributions.

- **DATA VISUALIZATION - EXCEL GRAPH**
    - Excel graphs and pivot tables are used to visualize and summarise data, making it easier to identify trends, patterns, and relationships. Graphs are used to display data visually, highlighting important patterns and trends

- **DATA SUMMARIZATION AND ANALYTICS - PIVOT TABLES**
    - pivot tables are used to summarise data and perform calculations.

CHALLENGES ENCOUNTERED DURING THIS PROJECT:

- Faced the limitations in Excel while plotting graphs in some graphs like Bubble charts, Scatter charts….etc number of data per set is unimportant but you can only display 255 series
- While filling values in the Legend series of the graph while performing the last task of building the dashboard.

KEY INSIGHTS DISCOVERED USING THIS PROJECT:

1. The identification of which car features are most popular among consumers and how they affect a car's pricing and profitability.
2. Identifying which market categories are most profitable for the manufacturer and how they affect pricing and profitability.
3. The development of a pricing strategy that balances consumer demand with profitability, based on insights from regression analysis and market segmentation.
4. The identification of trends in-car features and pricing over time could help inform product development and pricing decisions.

5. The comparison of fuel efficiency across different types of cars could inform consumer purchasing decisions and product development efforts focused on fuel efficiency.
6. The development of a model to predict the price of a car based on its features and market category, could inform pricing decisions and support sales and marketing efforts.

EXPLANATION OF THE RELATIONSHIP BETWEEN INSIGHTS AND THE BUSINESS PROBLEM:

The key insights from the analysis of the dataset are directly related to the business problem or question posed by the client, which is how a car manufacturer can optimize pricing and product development decisions to maximize profitability while meeting consumer demand.

Analyzing trends in-car features and pricing over time can help the manufacturer understand how the market has evolved and identify which features and price points have been successful in the past. This information can inform decisions about future product development and pricing strategies.

Comparing the fuel efficiency of different types of cars can help the manufacturer understand which types of cars are in demand and which types may need improvement to better meet consumer needs. This can guide decisions about product development and marketing strategies.

Investigating the relationship between a car's features and popularity can help the manufacturer understand which features are most important to consumers and which may be driving demand. This can guide decisions about product development and marketing strategies.

Predicting the price of a car based on its features and market category can help the manufacturer develop a pricing strategy that balances consumer demand with profitability. By identifying which features and categories are most profitable, the manufacturer can focus on developing and marketing those products to increase profitability.

Overall, the key insights from the analysis of the dataset can help the car manufacturer make data-driven decisions about product development, marketing, and pricing strategies that can increase profitability while meeting consumer demand.

CONCLUSIONS DRAWN FROM THE INSIGHTS

Based on the insights gained from the data analysis, the following recommendations and conclusions could be drawn for the car manufacturer:

1. Invest in alternative fuel technologies: The trend towards alternative fuel vehicles, such as electric and hybrid cars, is expected to continue. Therefore, the manufacturer could consider investing in research and development of these technologies to remain competitive in the market.
2. Focus on fuel efficiency: Consumers are increasingly concerned about fuel efficiency, and cars with better MPG ratings are likely to be more popular. The manufacturer could focus on improving the fuel efficiency of their cars to attract more customers.
3. Emphasise popular features: By analyzing the popularity variable, the manufacturer could identify which features are most popular among consumers and emphasize those features in their marketing and product development efforts.
4. Develop pricing strategy based on market category and features: By using regression analysis and market segmentation, the manufacturer could develop a pricing strategy that takes into account the market category and features of the cars, as well as consumer demand and profitability.
5. Offer customization options: By offering customization options for their cars, the manufacturer could attract more customers and increase profitability. For example, they could offer different engine power, transmission, or wheel options to cater to different customer needs and preferences.

Overall, the insights gained from the data analysis could help car manufacturers optimize their pricing and product development decisions to meet consumer demand while maximizing profitability.

RESULTS:

The results of this problem would likely involve identifying key factors that drive consumer demand for cars, analyzing trends in-car features and pricing over time, and developing a pricing strategy that balances consumer demand with profitability. Through data analysis techniques such as regression analysis and market segmentation, the manufacturer could gain insights into which product features are most popular among consumers and which are most profitable for the company. The analysis could also help identify opportunities for product development and innovation that would meet consumer demand while maximizing profitability. Ultimately, the results of this problem could help the manufacturer make informed decisions about pricing and product development that would improve its competitiveness in the market and increase its profitability over time.

TASKS PERFORMED

Tasks: Analysis

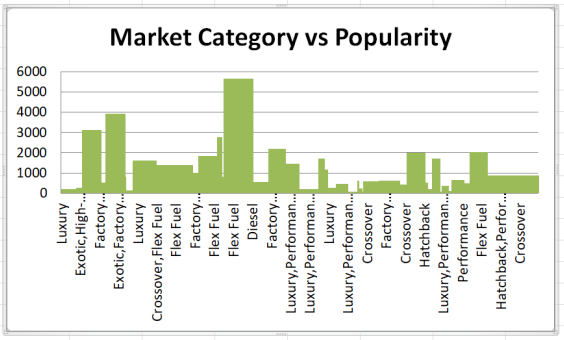
TASK 1

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 stacked column chart that visualizes the relationship between market category and popularity.

Row Labels	Sum of Popularity
Crossover	1688552
500X	12285
Acadia	17568
Acadia Limited	1098
Aztek	1260
B9 Tribeca	19200
Baja	9600
Crosstrek	5120
CR-V	70464
CX-3	7032
CX-5	16994
CX-7	11720
CX-9	8790
Durango	18510
Edge	113140
Element	46242
Enclave	2325
Encore	4340



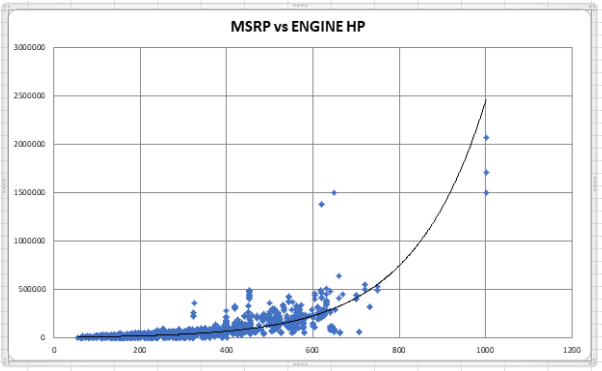
The above screenshot has the tasks completed. The pivot table shows the number of car models in models in each market category and their corresponding popularity scores and the relationship between market category and popularity.

TASK 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.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Make	Model	Year	he Fuel	ngine H	ine Cyli	nmissio	ren_Wh	ber of D	ket Cate	hicle Si	hicle St	hway M	city mpg	opularit	MSRP
2	Chevrole	Metro	2000	regular u	55	4	MANUAL	front whe	2	Hatchbac	Compact	2dr Hatcl	41	32	1385	2000
3	Chevrole	Metro	1999	regular u	55	4	MANUAL	front whe	2	Hatchbac	Compact	2dr Hatcl	42	34	1385	2000
4	Honda	Civic CR)	1990	regular u	62	4	MANUAL	front whe	2	Hatchbac	Compact	2dr Hatcl	44	36	2202	2000
5	Honda	Civic CR)	1991	regular u	62	4	MANUAL	front whe	2	Hatchbac	Compact	2dr Hatcl	44	36	2202	2000
6	Ford	Aspire	1996	regular u	63	4	MANUAL	front whe	4	Hatchbac	Compact	4dr Hatcl	38	29	5657	2000
7	Ford	Aspire	1996	regular u	63	4	MANUAL	front whe	2	Hatchbac	Compact	2dr Hatcl	38	29	5657	2000
8	Ford	Aspire	1997	regular u	63	4	MANUAL	front whe	2	Hatchbac	Compact	2dr Hatcl	38	29	5657	2000
9	Ford	Aspire	1997	regular u	63	4	MANUAL	front whe	4	Hatchbac	Compact	4dr Hatcl	38	29	5657	2000
10	Ford	Festiva	1991	regular u	63	4	MANUAL	front whe	2	Hatchbac	Compact	2dr Hatcl	38	30	5657	2000
11	Ford	Festiva	1991	regular u	63	4	MANUAL	front whe	2	Hatchbac	Compact	2dr Hatcl	38	30	5657	2000
12	Ford	Festiva	1992	regular u	63	4	MANUAL	front whe	2	Hatchbac	Compact	2dr Hatcl	37	30	5657	2000
13	Ford	Festiva	1992	regular u	63	4	MANUAL	front whe	2	Hatchbac	Compact	2dr Hatcl	37	30	5657	2000
14	Ford	Festiva	1993	regular u	63	4	MANUAL	front whe	2	Hatchbac	Compact	2dr Hatcl	38	30	5657	2000
15	Ford	Festiva	1993	regular u	63	4	MANUAL	front whe	2	Hatchbac	Compact	2dr Hatcl	38	30	5657	2000
16	Ford	Aspire	1995	regular u	63	4	MANUAL	front whe	2	Hatchbac	Compact	2dr Hatcl	38	31	5657	2000
17	Ford	Aspire	1995	regular u	63	4	MANUAL	front whe	2	Hatchbac	Compact	2dr Hatcl	38	31	5657	2000
18	Ford	Aspire	1995	regular u	63	4	MANUAL	front whe	4	Hatchbac	Compact	4dr Hatcl	38	31	5657	2000
19	Mitsubish	i-MiEV	2017	electric	66	0	DIRECT_	rear whe	4	Hatchbac	Compact	4dr Hatcl	102	121	436	22995
20	Mitsubish	i-MiEV	2016	electric	66	0	DIRECT_	rear whe	4	Hatchbac	Compact	4dr Hatcl	99	126	436	22995
21	Subaru	Justy	1992	regular u	66	4	MANUAL	front whe	2	Hatchbac	Compact	2dr Hatcl	33	28	640	2000
22	Subaru	Justy	1992	regular u	73	4	MANUAL	four whe	2	Hatchbac	Compact	2dr Hatcl	29	24	640	2000
23	Subaru	Justy	1992	regular u	73	4	MANUAL	four whe	4	Hatchbac	Compact	4dr Hatcl	29	24	640	2000
24	Subaru	Justy	1993	regular u	73	4	MANUAL	four whe	4	Hatchbac	Compact	4dr Hatcl	29	24	640	2000
25	Subaru	Justy	1994	regular u	73	4	MANUAL	four whe	4	Hatchbac	Compact	4dr Hatcl	29	24	640	2000
26	Subaru	Justy	1992	regular u	73	4	MANUAL	front whe	2	Hatchbac	Compact	2dr Hatcl	33	28	640	2000
27	Subaru	Justy	1993	regular u	73	4	AUTOMA	front whe	2	Hatchbac	Compact	2dr Hatcl	32	28	640	2000
28	Subaru	Justy	1993	regular u	73	4	AUTOMA	front whe	4	Hatchbac	Compact	4dr Hatcl	32	28	640	2000
29	Subaru	Justy	1993	regular u	73	4	MANUAL	front whe	2	Hatchbac	Compact	2dr Hatcl	33	28	640	2000
30	Subaru	Justy	1994	regular u	73	4	MANUAL	front whe	2	Hatchbac	Compact	2dr Hatcl	33	28	640	2000
31	Mitsubish	Mirage	2014	regular u	74	4	MANUAL	front whe	4	Hatchbac	Compact	4dr Hatcl	42	34	436	12995
32	Mitsubish	Mirage	2014	regular u	74	4	MANUAL	front whe	4	Hatchbac	Compact	4dr Hatcl	42	34	436	14195
33	Mitsubish	Mirage	2015	regular u	74	4	MANUAL	front whe	4	Hatchbac	Compact	4dr Hatcl	42	34	436	12995
34	Mitsubish	Mirage	2015	regular u	74	4	MANUAL	front whe	4	Hatchbac	Compact	4dr Hatcl	42	34	436	14295
35	Mitsubish	Mirage	2014	regular u	74	4	AUTOMA	front whe	4	Hatchbac	Compact	4dr Hatcl	44	37	436	15195
36	Mitsubish	Mirage	2014	regular u	74	4	AUTOMA	front whe	4	Hatchbac	Compact	4dr Hatcl	44	37	436	13995
37	Mitsubish	Mirage	2015	regular u	74	4	AUTOMA	front whe	4	Hatchbac	Compact	4dr Hatcl	44	37	436	15395



The above screenshot has the tasks completed. The scatter plot shows the relationship between Engine HP and MSRP and a trendline is added to the chart to visualize the relationship between these variables.

`TASK 3:

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

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.

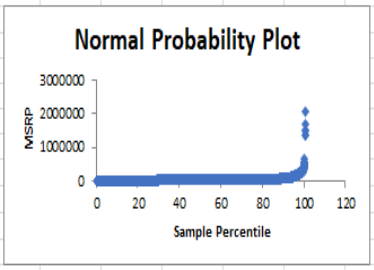
SUMMARY OUTPUT

Regression Statistics	
Multiple	0.683664
R Square	0.467397
Adjusted	0.467003
Standard	51271.95
Observati	8114

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regressic	6	1.87026E+13	3.1171E+12	1185.743733	0
Residual	8107	2.13118E+13	2628812553		
Total	8113	4.00144E+13			

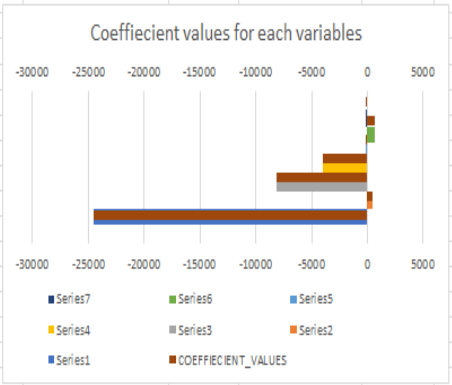
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-24528.5	4555.915596	-5.38388921	7.49339E-08	-33459.30864	-15597.78	-33459.30864	-15597.781
Engine HI	441.0854	5.846044641	75.4502336	0	429.6256859	452.54518	429.6256859	452.54518
Engine Cy	-8180.95	483.740684	-16.9118495	4.40324E-63	-9129.205514	-7232.694	-9129.205514	-7232.6937
Number c	-3947.74	651.7350228	-6.05727422	1.4462E-09	-5225.305662	-2670.17	-5225.305662	-2670.1698
highway I	-127.642	130.0279995	-0.98164933	0.326301928	-382.5301494	127.24635	-382.5301494	127.24635
city mpg	656.861	135.5353148	4.84641949	1.28044E-06	391.176989	922.54499	391.176989	922.54499
Popularit	-3.62336	0.40570433	-8.93104052	5.15346E-19	-4.418646423	-2.828077	-4.418646423	-2.8280772



ENGINE HP HAS HIGH t Stat value, it is most important in determining a car's price

PROBABILITY OUTPUT

Percentile	*MSRP*
0.006162	2000
0.018487	2000
0.030811	2000
0.043135	2000
0.05546	2000
0.067784	2000
0.080108	2000
0.092433	2000
0.104757	2000
0.117082	2000
0.129406	2000
0.14173	2000
0.154055	2000
0.166379	2000
0.178703	2000



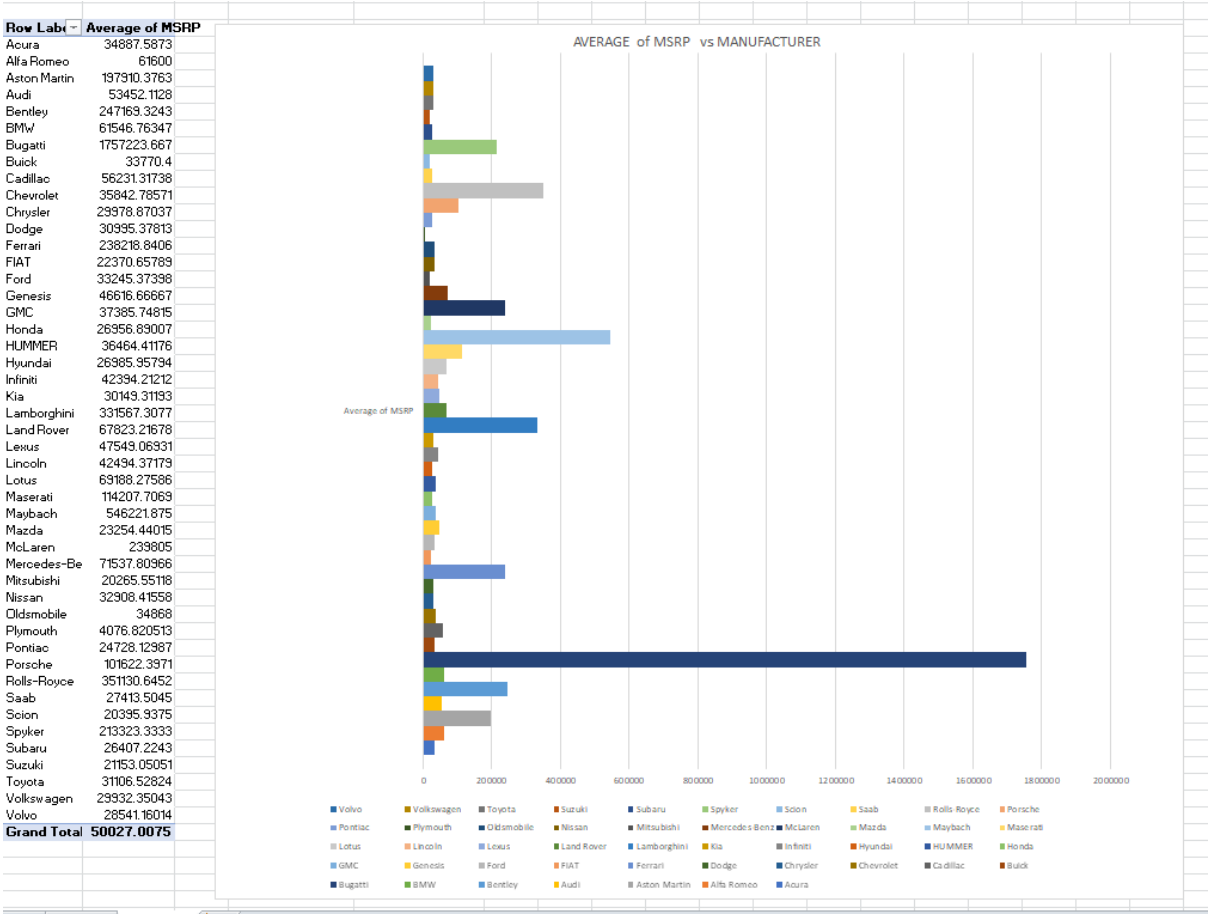
The above screenshot has the tasks completed. The strongest relationship between car price and other variables. Then the bar chart shows the coefficient values for each variable to visualize their relative importance

TASK 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 the manufacturer and average price.



Row Labels	Average of MSRP
Acura	34887.5873
Alfa Romeo	61600
Aston Martin	197910.3763
Audi	53452.1128
Bentley	247169.3243
BMW	61546.76347
Bugatti	1757223.667
Buick	33770.4
Cadillac	56231.31738
Chevrolet	35842.78571
Chrysler	29978.87037
Dodge	30995.37813
Ferrari	238218.8406
FIAT	22370.65789
Ford	33245.37398
Genesis	46616.66667
GMC	37385.74815
Honda	26956.89007
HUMMER	36464.41176
Hyundai	26985.95794
Infiniti	42394.21212
Kia	30149.31193
Lamborghini	331567.3077
Land Rover	67823.21678
Lexus	47549.06931
Lincoln	42494.37179
Lotus	69188.27586
Maserati	114207.7069
Maybach	546221.875
Mazda	23254.44015
McLaren	239805
Mercedes-Be	71537.80966
Mitsubishi	20265.55118
Nissan	32908.41558
Oldsmobile	34868
Plymouth	4076.820513
Pontiac	24728.12987
Porsche	101622.3971
Rolls-Royce	351130.6452
Saab	27413.5045
Scion	20395.9375
Spyker	213323.3333
Subaru	26407.2243
Suzuki	21153.05051
Toyota	31106.52824
Volkswagen	29932.35043
Volvo	28541.16014
Grand Total	50027.0075

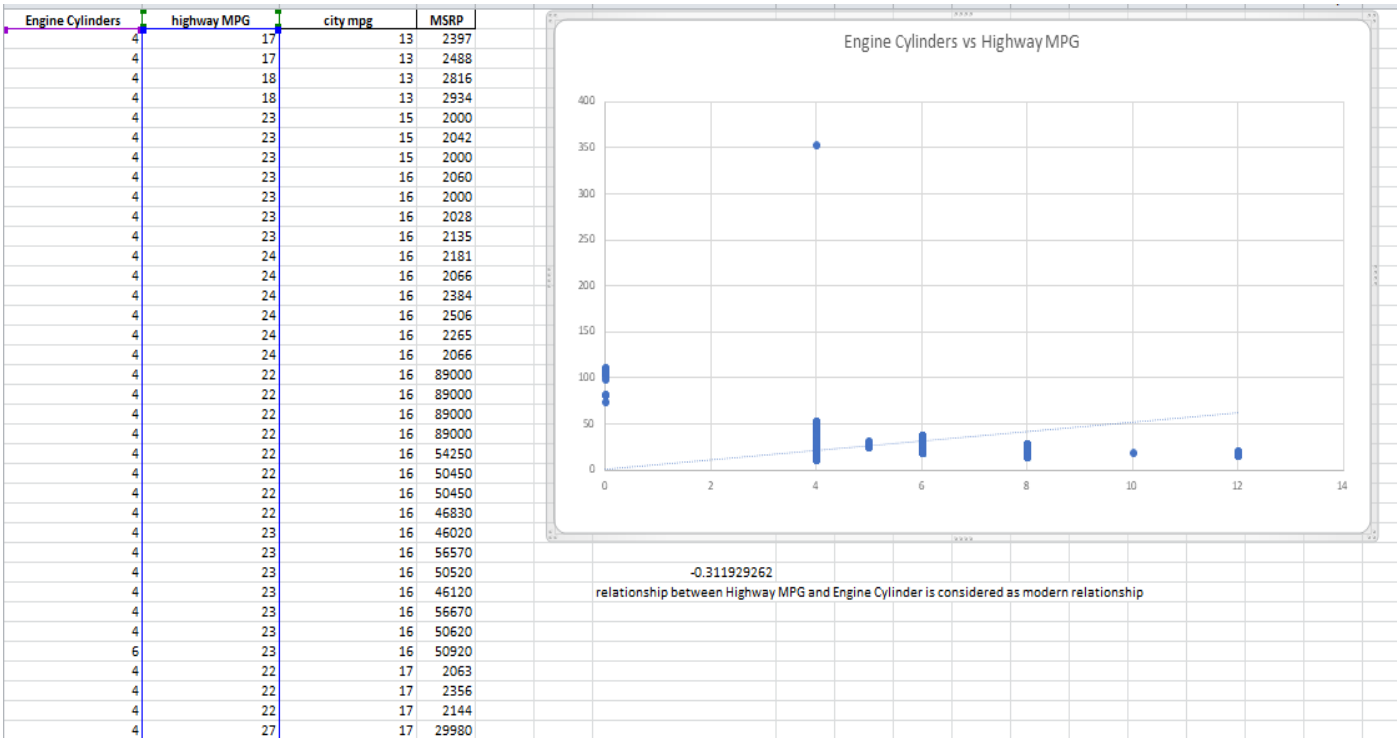
The above screenshot has the tasks completed. The pivot table shows the average price of cars for each manufacturer and the bar chart visualizes the relationship between the manufacturer and the average price.

TASK 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.



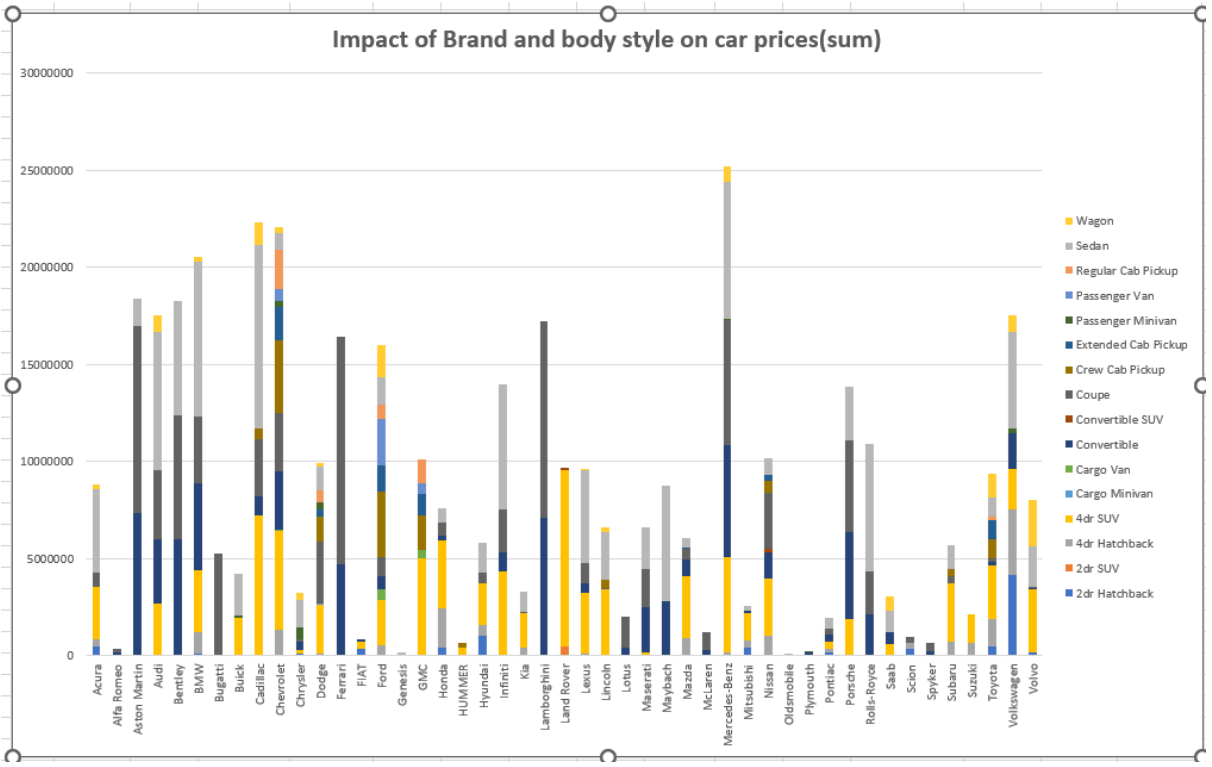
Engine Cylinders	highway MPG	city mpg	MSRP
4	17	13	2397
4	17	13	2488
4	18	13	2816
4	18	13	2934
4	23	15	2000
4	23	15	2042
4	23	15	2000
4	23	16	2060
4	23	16	2000
4	23	16	2028
4	23	16	2135
4	24	16	2181
4	24	16	2066
4	24	16	2384
4	24	16	2506
4	24	16	2265
4	24	16	2066
4	22	16	89000
4	22	16	89000
4	22	16	89000
4	22	16	89000
4	22	16	54250
4	22	16	50450
4	22	16	50450
4	22	16	46830
4	23	16	46020
4	23	16	56570
4	23	16	50520
4	23	16	46120
4	23	16	56670
4	23	16	50620
6	23	16	50920
4	22	17	2063
4	22	17	2356
4	22	17	2144
4	27	17	29980

The above screenshot has the tasks completed. the scatter plot is the relationship between the number of cylinders on the x-axis and highway MPG on the y-axis and the trendline on the scatter plot to visually estimate the slope of the relationship and assess its significance. The correlation has been the number of cylinders and highway MPG is **-0.311929262** which has been mentioned in the above screenshot.

The next portion of the project i.e Building the dashboard:

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

● Hints: Stacked column chart to show the distribution of car prices by brand and body style. Use filters and slicers to make the chart interactive. Calculate the total MSRP for each brand and body style using Pivot Tables.

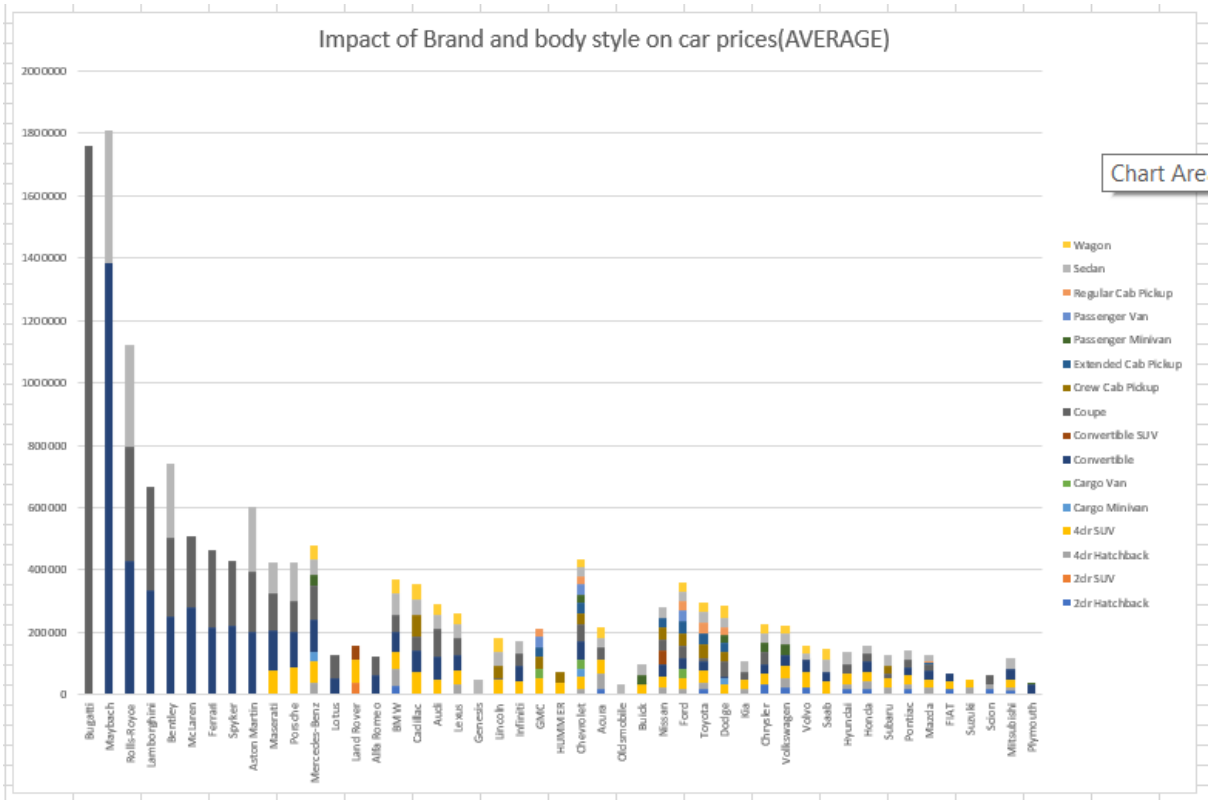


Sum of MSRP Column Labels																
Row Labels	2dr Hatchback	2dr SUV	4dr Hatchback	4dr SUV	Cargo Minivan	Cargo Van	Convertible	Convertible SUV	Coupe	Crew Cab Pickup	Extended Cab Pickup	Passenger Minivan	Passenger Van	Regular Cab Pickup	Sedan	Wagon	Grand Total
Acura	480917		357440	2663505					793748						4294702	201360	8791672
Alfa Romeo							129800		178200								308000
Aston Martin							7321655		9635275						1448735		18405665
Audi	4000			2674900			3291405		3556290						7158348	847350	17532293
Bentley							6012870		6356760						5920900		18290530
BMW	80097		1144950	3160950			4502671		3419051						7989300	259600	20556619
Bugatti									5271671								5271671
Buick				1944095			4000						85440		2187765		4221300
Cadillac				7182555			985607		2953574	593150					9418847	1184100	22323833
Chevrolet	8000		1287260	5103540	45185	60470	2953245		3048470	3747265	1720324		297685	591670	2044887	880480	290675	22073156
Chrysler	98805			181860			418885		108510				661141			1408277	360240	3237718
Dodge	48000		18000	2572405	40685		4000		3189736	1252705	392945		377205		613240	1233085	176515	9318521
Ferrari							472811		1.2E+07									16437100
FIAT	325315			369305			155465											850085
Ford	36000		480155	2323070		538085	730007		969809	3351618	1341505			2412375	760546	1404384	1619565	15367719
Genesis																139850		139850
GMC		12785		4369304		446085				1804755	1074860			591670	1194693			10094152
Honda	413200		2015270	3509435			252135		669198							742545		7601843
HUMMER				377430						242405								619895
Hyundai	1038050		528880	2128890					550400							1528775		5774995
Infiniti				4340200			980050		2175750							6494090		13990090
Kia			406960	1764140					84560							1030615		3286275
Lamborghini							7064450		1E+07									17241500
Land Rover		5E+05		9076595					145731									9698720
Lexus			94700	3152974			472065		1016472							4837596	31105	9604912
Lincoln				3422570					25342	453260						2458245	269705	6629122
Lotus							413260		1593200									2006460
Maserati				155000			2342963		1972284							2153800		6624047
Maybach							2762750									5976800		8739550
Mazda	22000		853180	3222525			870505		537879		28160				11303	477348		6022900
McLaren							280225		918800									1199025
Mercedes-Benz			122800	4324810	28950		5753964		6473107				32500			7080243	764935	25181309
Mitsubishi	394868		384840	1409460			97797									286760		2573725
Nissan	14683		1023090	2890520			1406552		131075 2929632		596390		332610			811240		10135792
Oldsmobile																104604		104604
Plymouth	42000		16000				85631						15365					158996
Pontiac	163505		162975	401550			346136		307965							521935		1304066
Porsche	28827			1815200			4504586		4758533							2713500		13820646
Rolls-Royce							2141365		2204675							6539010		10885050
Saab	14000		36586	541905			632628									1066500	751280	3042899
Scion	366325		282470						330210									979005
Spyker							219990		419980									639970
Subaru	12000		678060	3020230					350476	365975						1224405		5651146
Suzuki	46496		584387	1463269														2094152
Toyota	473750		1397750	2769605			152370		184490	1013680	1005170				158970	963325	1237955	9363065
Volkswagen	4171275		3355760	2084955			1827315						249070			5003670	818380	17510425
Volvo	157550			3219000			121600		6000							2086945	2428971	8020066
Grand Total	**8439663**	**####**	**15231513**	**9E+07**	**114820**	**1E+06**	**6.4E+07**	**276806**	**9E+07**	**13427203**	**5895574**	**1718406**	**3596315**	**4783639**	**1E+08**	**1E+07**	**4.1E+08**

The above graph and pivot table show the distribution of car prices by brand and body style

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

● Hints: Clustered column chart to compare the average MSRPs across different car brands and body styles. Calculate the average MSRP for each brand and body style using Pivot Tables.



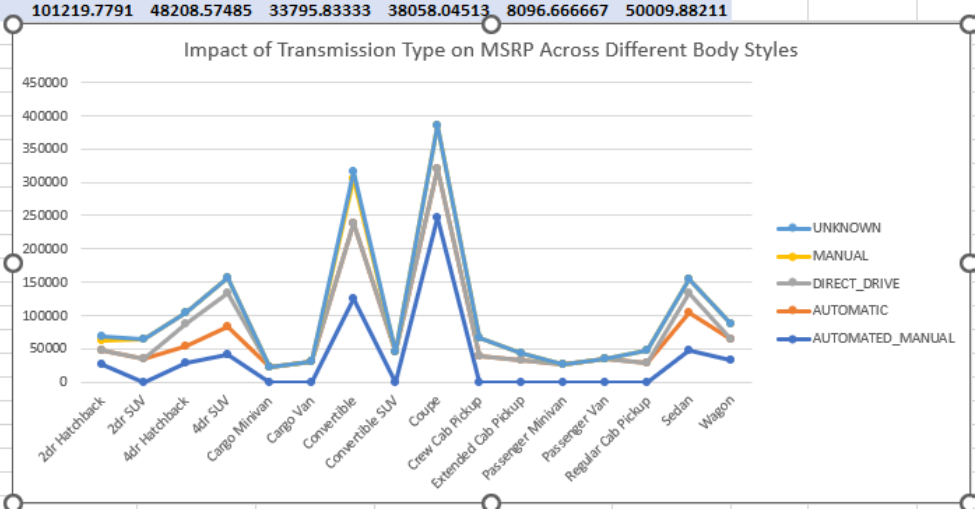
Average of MS	Column Labe															
Row Labels	2dr Hatchback	2dr SUV	4dr Hatchb	4dr SUV	Cargo Minivi	Cargo Van	Convertible	Convertible S	Coupe	Crew Cab Pic	Extended Cab Pi	Passenger Min	Passenger V	Regular Cab Pi	Sedan	Wagon	Grand Total
Bugatti									1757223.67								1757223.67
Maybach							1381375								426914.286		546221.875
Rolls-Royce							428273		367445.833						326950.5		351130.645
Lamborghini							336402.381		328291.935								331567.308
Bentley							250536.25		254270.4						236836		247169.324
McLaren							280225		229700								239805
Ferrari							214718.682		243218.915								238218.841
Spyker							219390		209390								213323.333
Aston Martin							203379.306		192705.5						206962.143		197910.376
Maserati				77500			130164.611		116016.706						102561.905		114207.707
Porsche	5765.4			82509.0909			115502.205		99136.1042						123340.909		101622.397
Mercedes-Benz			40933.3333	68400.1389	28950		104617.527		109713.676			32500			43168.3542	44396.1765	71537.8097
Lotus							51657.5		75866.6667								69188.2759
Land Rover		39639.5		70910.8984				48577									67823.2168
Alfa Romeo							64900		53400								61600
BMW	26639		54521.4286	58536.1111			63417.9014		51803.803						70701.7699	43266.6667	61546.7635
Cadillac				72551.0606			70400.5		45439.6	66572.22222					50912.6865	47364	56231.3174
Audi	2000			48634.5455			70029.8936		93586.5789						44461.7888	33894	53452.1128
Lexus			31566.6667	45042.4857			52451.6667		50823.6						48864.6061	31105	47549.0633
Genesis															46616.6667		46616.6667
Lincoln				50331.9118					2111.83333	41205.45455					41665.1695	44350.8333	42494.3718
Infiniti				45686.3158			46669.0476		40291.6667						40588.0625		42394.2121
GMC		6392.5		45175.4909		31863.2143				39233.80435	29857.22222		34804.1176	26548.73333			37385.7481
HUMMER				37749						34629.28571							36464.4118
Chevrolet	2000		18930.2941	37804	22592.5	30235	62835		49374.918	39034.01042	30181.12281	24807.08333	34804.1176	28401.20833	30361.3793	20762.5	35842.7857
Acura	17175.60714		51062.8571	42953.7581					39687.4						33292.2636	33560	34887.5873
Oldsmobile															34868		34868
Buick				34106.9298			2000					28480			34726.4286		33770.4
Nissan	2097.571429		22241.087	33610.6977			39070.8889	43691.66667	37083.9494	37274.375	33261				32449.6		32908.4156
Ford	2000		18467.5	34162.7941		29833.6111	34762.2381		35918.8519	42425.5443	37264.02778		36560.2273	29251.76923	28087.68	31145.4808	32787.924
Toyota	18950		22186.5079	39565.7857			25395		13177.8571	44073.04348	35898.92857			31794	34618.75	31742.4359	31106.5282
Dodge	2000		2000	30992.8313	20342.5		2000		52290.7541	29826.30952	26196.33333	25147		25551.66667	32449.6053	35303	30995.3781
Kia			19379.0476	31502.5					21140						36807.6786		30143.3119
Chrysler	32935			36372			29920.3571		36170			33057.05			27613.2745	30020	29878.8704
Volkswagen	24251.59884		27964.6667	41699.1			32058.1579					35581.42857			33357.8	28220	29932.3504
Volvo	26258.33333			45338.0282			40533.3333		2000						20869.45	24785.4184	28541.1601
Saab	2000		2032.55556	41685			28755.8182								36775.8621	34149.0909	27413.5045
Hyundai	18536.60714		17629.3333	30412.7143					30577.7778						38219.375		26985.9579
Honda	17216.66667		25836.7949	23935.6838			36019.2857		22306.6						28553.4231		26956.8901
Subaru	2000		21189.375	29322.6234					17523.8	24398.33333					32221.1842		26407.2243
Pontiac	18167.22222		18106.3333	25096.875			24724		27936.8182						28936.3889		24726.1299
Mazda	2000		20809.2683	27080.042			28080.8065		22411.625		3128.888889			2825.75	23867.4		23254.4402
FIAT	19136.17647			24620.3333			25910.8333										22370.6579
Suzuki	6642.285714		16696.7714	25671.386													21153.0505
Scion	20351.38889		15692.7778						27517.5								20395.9375
Mitsubishi	13162.26667		12828	25168.9286			32593								35845		20265.5512
Plymouth			2000				28543.6667					2195					4076.82051
Grand Total	16778.65408	34941.36	22203.37	42424.01	22964	30724.71	95465.31	46134.3333	92777.02	39032.5669	30866.87958	26437.01538	35963.15	27179.76705	51202.85	31489.46	50009.88

The above graph and pivot table shows the average MSRPs across different car brands and body styles and the average MSRP for each brand and body style using Pivot Tables. The **highest average MSRP is COUPE and the lowest average is 4dr Hatchback**

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

● Hints: Scatter plot chart to visualize the relationship between MSRP and transmission type, with different symbols for each body style. Calculate the average MSRP for each combination of transmission type and body style using Pivot Tables.

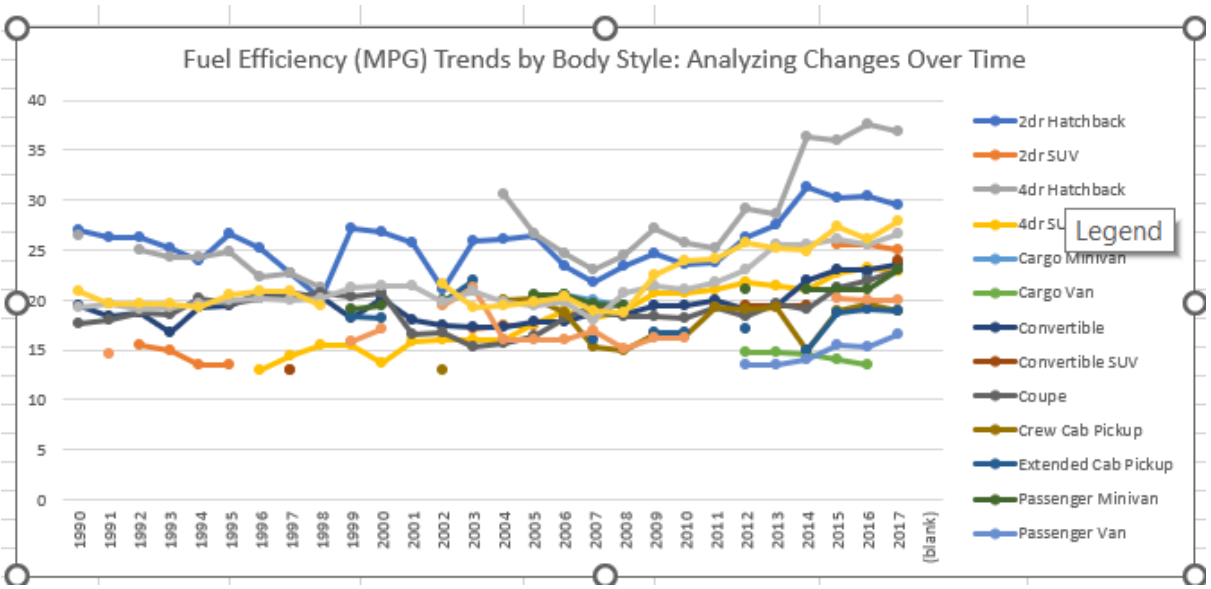
Average of MSRP	Column Labels					
Row Labels	AUTOMATED_MANUAL	AUTOMATIC	DIRECT_DRIVE	MANUAL	UNKNOWN	Grand Total
2dr Hatchback	27180.96491	20926.464		13353.65831	7361.5	16778.65408
2dr SUV		35894.5		29222.5		34941.35714
4dr Hatchback	29249.07407	23833.67898	33602.61905	17594.41313		22203.37172
4dr SUV	40451.15385	42827.04736	49800	23097.925		42424.00764
Cargo Minivan		22964				22964
Cargo Van		30724.70588				30724.70588
Convertible	125806.2656	111674.9644		68038.96875	9567	95465.31045
Convertible SUV		46134.33333				46134.33333
Coupe	245977.4252	75004.07268		64154.23148		92777.02088
Crew Cab Pickup		39564.70517		27361		39032.56686
Extended Cab Pickup		32970.31214		10650.55556		30866.87958
Passenger Minivan		26437.01538				26437.01538
Passenger Van		35963.15				35963.15
Regular Cab Pickup		29209.75694		18044.8125		27179.76705
Sedan	48491.14428	56525.45336	27822.5	21943.46122		51202.84575
Wagon	31985.27778	33228.50177		22541.69643		31489.45658
Grand Total	**101219.7791**	**48208.57485**	**33795.83333**	**38058.04513**	**8096.666667**	**50009.88211**



The above graph visualizes the relationship between MSRP and transmission type, with different symbols for each body style and the pivot table shows the average MSRP for each combination of transmission type and body style.

Task 4: How does the fuel efficiency of cars vary across body styles and model years?
● Hints: Line chart to show the trend of fuel efficiency (MPG) over time for each body style. Calculate the average MPG for each combination of body style and model year using AVERAGEIFS or Pivot Tables.

Average of Average MPG	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	(blank)	Grand Total
1990	27		26.5				19.5		17.666667						19.3125	20.8		20.657895
1991	26.2						18.357143		18					14.5	19.545455	19.5625		20.855932
1992	26.18181818	15.5	25.0625				18.666667		18.75						19.057692	19.571429		22.347059
1993	25.13333333	15	24.35				16.8		18.555556						19.258065	19.666667		21.505435
1994	23.975	13.5	24.2142857				19.25		20.166667						19.633333	19.25		21.54918
1995	26.64285714	13.5	24.8333333				19.5		19.619048						19.970588	20.571429		20.844828
1996	25.25		22.375	13			20.6		20.175						20.1	20.916667		21.008929
1997	22.72222222	13	22.65	14.333333			20.75	13	20.333333						19.9	20.8		20.541667
1998	20.2		21.25	15.5			20.333333		20.727273						20.222222	19.5		20.296875
1999	27.16666667			15.5			18.25		20.3		18.3	19		15.75	21.25			21.244444
2000	26.75			13.666667			20.2		20.666667		18.25	19.5		17	21.413043			21.316667
2001	25.64285714			15.857143			17.9		16.46875						21.447368			20
2002	21.125			16.085333			17.5		16.692308	13	19.85714286	20		19.45454545	19.710526	21.5		18.647887
2003	25.875	17		16			17.192308		15.222222		22			21.2	20.941176	19.333333		18.545455
2004	26	17.5	30.5	16.038462			17.2		15.607143	20				16	19.730769	19.5		19.428571
2005	26.44444444	17.5	26.7	17.7			17.727273		16.318182	20.2		20.5		16	19.480769	19.833333		19.809783
2006	23.45833333		24.6666667	18.903846	20.5		17.8		18.176471	18.78571429		20.5		16	19.725	20.277778		20.112613
2007	21.73076923		23.0769231	18.244681	20		18.772727		19.425	15.34375	15.97435897	19.59090909		16.875	17.964286	18.916667		18.236052
2008	23.36363636		24.5	18.8	19.5		18.578125		18.32	14.95833333		19.5		15.125	20.685185	18.7		19.312195
2009	24.625		27.2	20.694444			19.5		18.348485	16.5483871	16.66666667			16.125	21.325397	22.543478		20.21134
2010	23.57692308		25.65625	20.657895			19.464286		18.175	16.4	16.75			16.125	21.053571	23.928571		20.367521
2011	23.83333333		25.1896552	21.131944			20.029412		19.267857	19.20588235					21.6875	24.1		21.519149
2012	26.2173913		29.1759259	21.761905		14.666667	19.125	19.5	18.418919	19.08333333	17	21	13.5		23.095588	25.76087		22.877419
2013	27.5		28.5227273	21.456522		14.666667	19.340909	19.5	19.578947	19.20588235			13.5		25.574627	25.12963		23.709524
2014	31.22222222		36.3571429	21.088235		14.642857	22.006944	19.5	19.052632	14.94444444	14.9	21	14.0625		25.616505	24.916667		23.317365
2015	30.3125	25.5	36.0785714	22.725243		14	23.038462		21.252841	18.90178571	18.71428571	21	15.42857143	20.21212121	26.026185	27.295918		24.615567
2016	30.48360656	25.5	37.5247934	23.27518	22.5	13.5	22.935345		21.946429	19.64084507	19	21.03571429	15.21428571	20.015625	25.544153	26.068182		24.563284
2017	29.43333333	25	36.9576271	22.916295			23.478873	24	22.948148	18.94897959	18.81944444	23	16.5	20.015625	26.658147	27.85		24.95659
(blank)																		
Grand Total	27.12723658	19.642857	32.8367347	22.167144	20.5	14.441176	21.285821	19.91666667	20.299479	18.27906977	18.08900524	20.43076923	15.09	18.86363636	24.298062	23.935574		23.189056

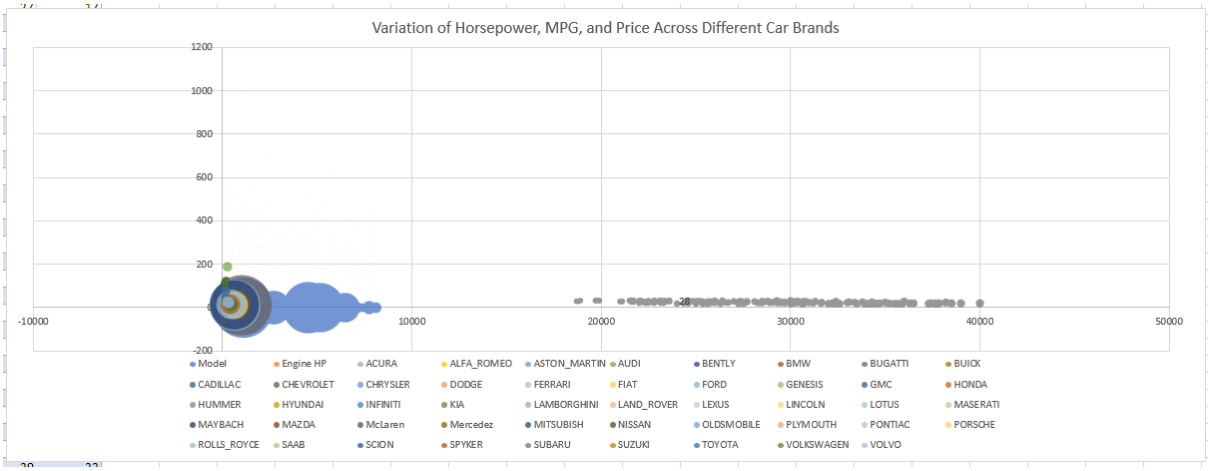


The above graph shows the fuel efficiency of cars varies across body styles and model years and the line chart shows the trend of fuel efficiency (MPG) over time for each body style. The pivot table shows the average MPG for each combination of body style and model year.

Task 5: How do the car's horsepower, MPG, and price vary across different Brands?
● Hints: Bubble chart to visualize the relationship between horsepower, MPG, and price across different car brands. Assign different colors to each brand and label the bubbles with the car model name. Calculate the average horsepower, MPG, and MSRP for each car brand using AVERAGEIFS or Pivot Tables.

Row Labels	Average of Average MPG	Average of Engine HP	Average of MSRP
Acura	**24.02579365**	**244.797619**	**34887.5873**
CL	21.88888889	236.6666667	29818.88889
ILX	29.875	191.4375	30712.5
ILX Hybrid	38.5	111	31750
Integra	25	152.0833333	11768.58333
Legend	19.59375	215	2143.5
MDX	22.79411765	290	50332.05882
NSX	19.5	331.4	102400
RDX	23.70833333	278	39232.5
RL	20	300	51340
RLX	26.16666667	321.1666667	56491.66667
RSX	26.36666667	167.4	21689
SLX	15.25	202.5	2658.75
TL	23.06521739	293.0434783	40876.73913
TLX	27.64285714	266	38730
TSX	25.46875	220.75	33284.0625
TSX Sport Wagon	26	201	33560
Vigor	21	176	2000
ZDX	19.5	300	51062.85714
Alfa Romeo	**29**	**237**	**61600**
4C	29	237	61600
Aston Martin	**15.70967742**	**484.3225806**	**197910.3763**
DB7	13.375	423.75	151550
DB9	16	510	196522.5
DB9 GT	16	540	215602.3333
DBS	14.5	510	281573
Rapide	16	470	210122.5
Rapide S	17.33333333	551.3333333	202748.3333
V12 Vanquish	13.375	490	245880
V12 Vantage	14	510	187565
V12 Vantage S	15	565	188095
V8 Vantage	16.75	428.75	129535
Vanquish	16.75	567.25	297083.75
Virage	15.5	490	215795
Audi	**24.20426829**	**277.695122**	**53452.1128**
100	19.36666667	172	2000
200	18.58333333	180.3333333	2000
80	20.41666667	126.3333333	2000
90	20.22222222	172	2000
A3	30.07446809	190.3404255	38652.12766
A4	27.9	225.36	39752

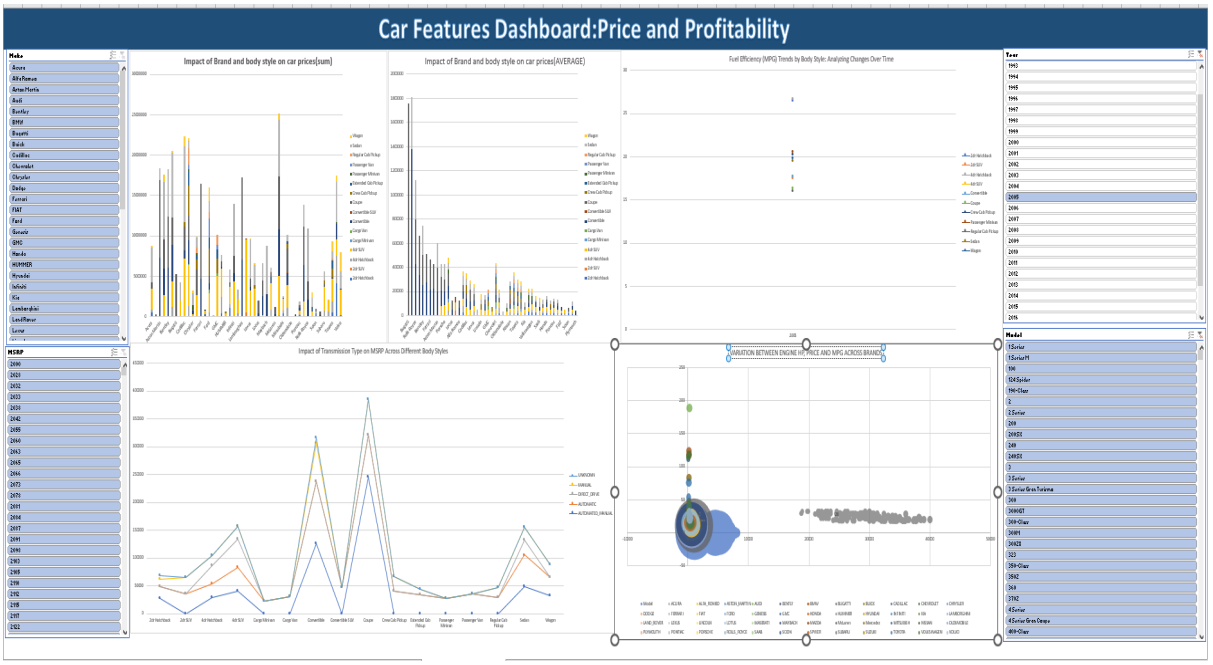
Acura	CL	27980	225	
Acura	CL	28030	225	
Acura	CL	30030	225	
Acura	CL	30350	225	
Acura	CL	28200	225	
Acura	CL	32700	260	
Acura	CL	30550	260	
Acura	CL	30550	260	
Acura	ILX	29350	201	
Acura	ILX	31750	150	
Acura	ILX	29350	150	
Acura	ILX	27050	150	
Acura	ILX	34890	201	
Acura	ILX	29900	201	
Acura	ILX	29200	201	
Acura	ILX	31890	201	
Acura	ILX	32900	201	
Acura	ILX	27900	201	
Acura	ILX	32990	201	
Acura	ILX	29290	201	
Acura	ILX	31980	201	
Acura	ILX	34980	201	
Acura	ILX	27990	201	
Acura	ILX	29990	201	
Acura	ILX Hybrid	28900	111	
Acura	ILX Hybrid	34600	111	
Acura	Integra	20200	140	
Acura	Integra	21850	140	
Acura	Integra	21000	140	
Acura	Integra	22400	140	
Acura	Integra	2827	140	
Acura	Integra	3000	140	
Acura	Integra	2912	140	
Acura	Integra	2799	140	
Acura	Integra	3130	140	
Acura	Integra	3012	140	
Acura	Integra	3222	140	
Acura	Integra	3086	140	
Acura	Integra	21050	140	
Acura	Integra	19400	140	
Acura	Integra	21600	140	
Acura	Integra	20200	140	



The above screenshot of the pivot table shows the average horsepower, MPG, and MSRP for each car brand and the bubble chart shows the relationship between horsepower, MPG, and price across different car brands.

FINAL DASHBOARD (CAR FEATURES DASHBOARD: PRICE AND PROFITABILITY)

The dashboard is likely designed to provide an overview of how various car features relate to pricing and profitability in the automotive industry. The dashboard might include various visualizations and interactive tools that allow the user to explore different aspects of the data and gain insights into the relationships between car features, pricing, and profitability.



The dashboard can be customized based on the specific needs of the user by using the slicers provided for the brand name, model name, year of launch, and the car price. The user can select a specific brand, model, year, or price range, and the graphs will update accordingly to show the relevant data. This allows for more targeted analysis and helps the user to gain deeper insights into specific segments of the market

The dashboard may not be able to show insights for certain conditions if the necessary features are not available in the dataset. For example, if the dataset does not include information about the geographic location where the car is sold, the dashboard may not be able to provide insights related to regional pricing trends or consumer preferences. Similarly, if the dataset does not include information about certain advanced features such as autonomous driving or electric vehicle range, the dashboard may not be able to provide insights related to those features. Therefore, it's important to keep in mind the limitations of the dataset when using the dashboard and interpreting its results.

DISCUSSION AND ITS IMPLICATION:

The results of the analysis provide valuable insights into the factors that affect the price of cars and the preferences of consumers. By developing a regression model to predict the price of cars based on their features, it was found that factors such as engine power, number of cylinders, and market category had a significant impact on the price of a car. This suggests that manufacturers could focus on developing high-powered, high-end cars in popular market categories to maximize profitability.

In addition, by conducting a market segmentation analysis, it was found that there were distinct groups of consumers with different preferences for car features. For example, one group valued fuel efficiency and eco-friendliness, while another valued performance and luxury features. This information could be used by manufacturers to tailor their product development efforts and marketing strategies to specific consumer segments, improving their competitiveness in the market.

Overall, the implications of these results suggest that manufacturers should focus on developing cars that meet the preferences and needs of different consumer segments, while also balancing profitability. This could involve developing different models for different market categories, focusing on high-powered, luxury cars for one segment and eco-friendly, fuel-efficient cars for another. By using data analysis techniques to understand consumer preferences and market trends, manufacturers can make informed decisions that maximize profitability while also meeting consumer demand.

LIMITATIONS

1. Limited data: While the dataset used in this analysis contains information on over 11,000 car models, there may be other variables that are not included in the dataset that could impact a car's pricing and profitability. For example, factors such as location and competition in the local market could also affect pricing and demand for specific car models.
2. Outdated data: The dataset used in this analysis only goes up until the year 2017. Given the rapidly changing nature of the automotive industry, it is possible that some of the insights and conclusions are drawn from this analysis may not be applicable to more recent years.
3. Generalizability: The findings and recommendations from this analysis are based on the specific dataset used and may not be generalizable to other datasets or contexts. It is important to consider the unique characteristics of the dataset and the specific problem at hand when drawing conclusions and making decisions based on the analysis.

POSSIBLE FUTURE DIRECTIONS

1. Incorporating external data sources: The current dataset only includes information on car models and their specifications. Incorporating external data sources, such as consumer behavior data and economic indicators, could provide additional insights into consumer demand and market trends.
2. Exploring more advanced machine learning techniques: While this project utilized regression analysis and market segmentation techniques, more advanced machine learning techniques, such as neural networks or random forests, could be explored to

improve the accuracy of price predictions and identify more complex patterns in the data.

3. Conducting a more in-depth analysis of consumer preferences: While this project identified several features that are popular among consumers, a more in-depth analysis of consumer preferences, such as through surveys or focus groups, could provide a more nuanced understanding of what drives consumer demand for certain features and help inform future product development efforts.
4. Applying the findings to specific market segments: The current analysis was conducted on the overall dataset, but applying the findings to specific market segments, such as luxury cars or electric vehicles, could provide more targeted insights for manufacturers operating in those markets.
5. Updating the dataset: As the automotive industry continues to evolve, updating the dataset with new car models and specifications could provide more up-to-date insights into market trends and consumer demand.

ADDITIONAL ANALYSIS:

1. Time-series analysis: This dataset contains information about car models from 1990 to 2017. A time-series analysis could be done to identify trends in car features and pricing over time and to forecast future trends.
2. Geospatial analysis: The dataset includes information on the geographic location of each car model's manufacturer. A geospatial analysis could be done to identify regional differences in car features, pricing, and popularity.
3. Sentiment analysis: The dataset includes customer reviews for some car models. Sentiment analysis could be done to identify trends in customer satisfaction and identify areas where manufacturers could improve their products.
4. Competitive analysis: The dataset includes information on over 11,000 car models from more than 50 manufacturers. A competitive analysis could be done to identify which manufacturers are most successful in the market and to identify areas where competitors have an advantage.
5. Feature importance analysis: The current analysis identified which car features are most important in predicting price. However, a more in-depth analysis could be done to identify which features are most important for different market segments or which features are most associated with high customer satisfaction.