Trainity

Analyzing the Impact of Car Features on Price and Profitability

Final Project – 3

Done by

Gaddam Lakshmi Deepak

Data Analytics Trainee

Description

This project is used to analyzing the dataset of “Car_data.csv” to draw meaningful insights from it. This analysis helps the client to get how can a car manufacturer optimize pricing and product development decisions to maximize profitability while meeting consumer demand. By doing analysis of cars data the relationship between a car's features, market category, and pricing, and identifying which features and categories are most popular among consumers and most profitable for the manufacturer.

Dataset used: Car_data.csv

Insight Required:

A. How does the popularity of a car model vary across different market categories?
B. What is the relationship between a car's engine power and its price?
C. Which car features are most important in determining a car's price?
D. How does the average price of a car vary across different manufacturers?
E. What is the relationship between fuel efficiency and the number of cylinders in a car's engine?

Dashboard Tasks:

1. How does the distribution of car prices vary by brand and body style?
2. Which car brands have the highest and lowest average MSRPs, and how does this vary by body style?
3. How do the different feature such as transmission type affect the MSRP, and how does this vary by body style?
4. How does the fuel efficiency of cars vary across different body styles and model years?
5. How does the car's horsepower, MPG, and price vary across different Brands?

My Approach

Firstly, I have taken the dataset and studied the whole dataset to gets some knowledge and get some idea to make decisions. After that, cleaned the dataset removed unnecessary things.

After analysis, I have started reading tasks understand the statement and identified required columns to solve the tasks. Started creating tables, pivot tables based on task and create graphs or charts to represent visualization of data. This is my approach to solve the given tasks.

Software used to do this project

Microsoft Office Excel

Microsoft Word Document

Data Cleaning

Find the count of blanks for each column using **COUNTBLANK()** function.

Engine Fuel Type having 3 blanks filled with natural gas.

Number of doors having 6 blanks filled with searched from internet.

Engine Cylinder having 30 blanks filled with 0 as some are electric and some has no cylinders.

Engine HP having 69 blanks filled with N/A.

Insights and Results

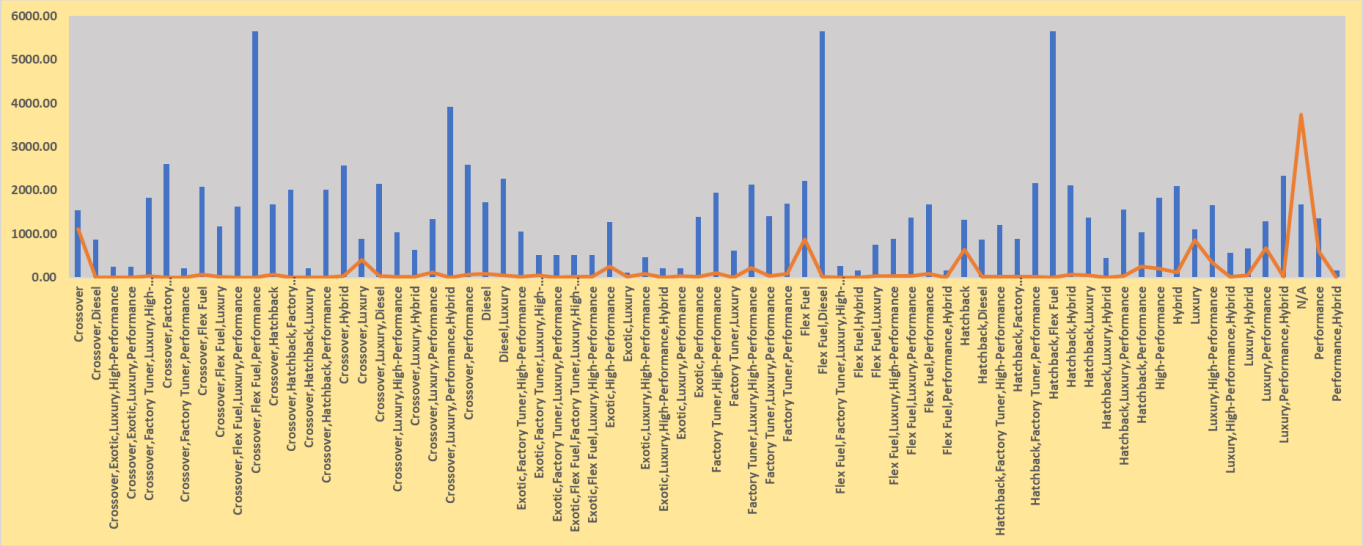
Required Insights

A. How does the popularity of a car model vary across different market categories?

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

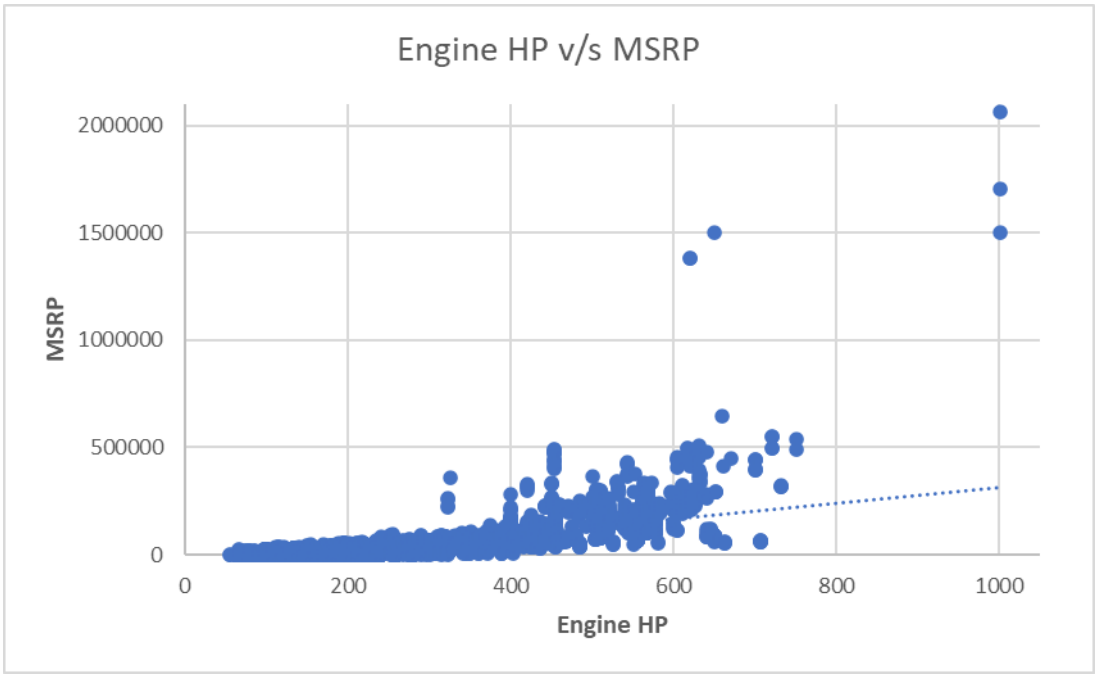
Market Category	Average of Popularity	Count of Model
Crossover	1545.26	1110
Crossover,Diesel	873.00	7
Crossover,Exotic,Luxury,High-Performance	238.00	1
Crossover,Exotic,Luxury,Performance	238.00	1
Crossover,Factory Tuner,Luxury,High-Performance	1823.46	26
Crossover,Factory Tuner,Luxury,Performance	2607.40	5
Crossover,Factory Tuner,Performance	210.00	4
Crossover,Flex Fuel	2073.75	64
Crossover,Flex Fuel,Luxury	1173.20	10
Crossover,Flex Fuel,Luxury,Performance	1624.00	6
Crossover,Flex Fuel,Performance	5657.00	6
Crossover,Hatchback	1675.69	72
Crossover,Hatchback,Factory Tuner,Performance	2009.00	6
Crossover,Hatchback,Luxury	204.00	7
Crossover,Hatchback,Performance	2009.00	6
Crossover,Hybrid	2563.38	42
Crossover,Luxury	884.55	410
Crossover,Luxury,Diesel	2149.41	34
Crossover,Luxury,High-Performance	1037.22	9
Crossover,Luxury,Hybrid	630.92	24
Crossover,Luxury,Performance	1344.85	113
Crossover,Luxury,Performance,Hybrid	3916.00	2
Crossover,Performance	2585.96	69
Diesel	1730.90	84
Diesel,Luxury	2275.00	51
Exotic,Factory Tuner,High-Performance	1046.38	21
Exotic,Factory Tuner,Luxury,High-Performance	517.54	52
Exotic,Factory Tuner,Luxury,Performance	520.00	3
Exotic,Flex Fuel,Factory Tuner,Luxury,High-Performance	520.00	13
Exotic,Flex Fuel,Luxury,High-Performance	520.00	11
Exotic,High-Performance	1271.33	261
Exotic,Luxury	112.67	12
Exotic,Luxury,High-Performance	467.08	79
Exotic,Luxury,High-Performance,Hybrid	204.00	1
Exotic,Luxury,Performance	217.03	36
Exotic,Performance	1391.00	10
Factory Tuner,High-Performance	1941.42	106
Factory Tuner,Luxury	617.00	2
Factory Tuner,Luxury,High-Performance	2133.37	215
Factory Tuner,Luxury,Performance	1413.42	31
Factory Tuner,Performance	1695.70	92
Flex Fuel	2217.30	872
Flex Fuel,Diesel	5657.00	16
Flex Fuel,Factory Tuner,Luxury,High-Performance	258.00	1
Flex Fuel,Hybrid	155.00	2
Flex Fuel,Luxury	746.54	39
Flex Fuel,Luxury,High-Performance	878.91	33
Flex Fuel,Luxury,Performance	1380.07	28
Flex Fuel,Performance	1680.47	87
Flex Fuel,Performance,Hybrid	155.00	2
Hatchback	1318.87	641
Hatchback,Diesel	873.00	14
Hatchback,Factory Tuner,High-Performance	1205.15	13
Hatchback,Factory Tuner,Luxury,Performance	886.89	9
Hatchback,Factory Tuner,Performance	2159.05	22
Hatchback,Flex Fuel	5657.00	7
Hatchback,Hybrid	2121.25	72
Hatchback,Luxury	1379.50	46
Hatchback,Luxury,Hybrid	454.00	3
Hatchback,Luxury,Performance	1566.13	38
Hatchback,Performance	1039.65	252
High-Performance	1821.45	199
Hybrid	2105.57	123
Luxury	1102.66	855
Luxury,High-Performance	1668.02	334
Luxury,High-Performance,Hybrid	568.83	12
Luxury,Hybrid	673.63	52
Luxury,Performance	1292.62	673
Luxury,Performance,Hybrid	2333.18	11
N/A	1676.89	3742
Performance	1348.87	601
Performance,Hybrid	155.00	1

Task A.2: Create a combo chart that visualizes the relationship between market category and popularity.



B. What is the relationship between a car's engine power and its price?

Task: 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.



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

Task: 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.

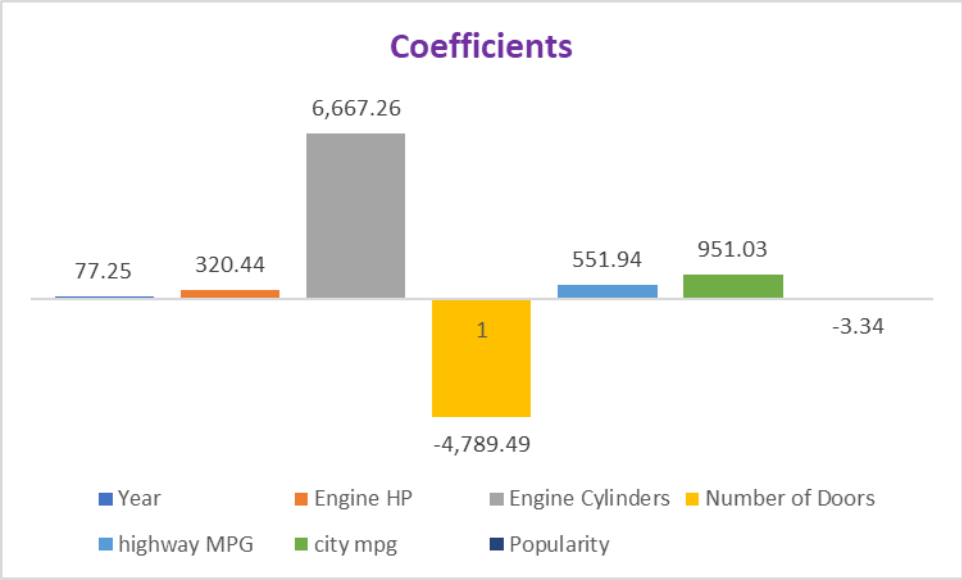
Columns: Year, Engine HP, Engine Cylinders, Number of Doors, highway MPG, city mpg, Popularity and MSRP

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.684243771
R Square	0.468189539
Adjusted R Square	0.467875044
Standard Error	43951.11856
Observations	11845

ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	7	2.01301E+13	2.87573E+12	1488.704279	0
Residual	11837	2.28655E+13	1931700823		
Total	11844	4.29957E+13			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-243615.2999	147711.1249	-1.649268463	0.099119183	-533153.3907	45922.79093	-533153.3907	45922.79093
Year	77.24920862	74.16223475	1.041624607	0.297607029	-68.12096497	222.6193822	-68.12096497	222.6193822
Engine HP	320.4390728	7.476153103	42.86149151	0	305.7845835	335.0935621	305.7845835	335.0935621
Engine Cylinders	6667.261006	450.7705081	14.79081015	4.59962E-49	5783.676697	7550.845316	5783.676697	7550.845316
Number of Doors	-4789.493815	490.285796	-9.768779464	1.86349E-22	-5750.534586	-3828.453044	-5750.534586	-3828.453044
highway MPG	551.9380247	106.0631618	5.203861694	1.98455E-07	344.0367891	759.8392603	344.0367891	759.8392603
city mpg	951.0287144	112.9543644	8.4195836	4.21201E-17	729.6195886	1172.43784	729.6195886	1172.43784
Popularity	-3.341987382	0.283470311	-11.78954991	6.66308E-32	-3.8976358	-2.786338965	-3.8976358	-2.786338965

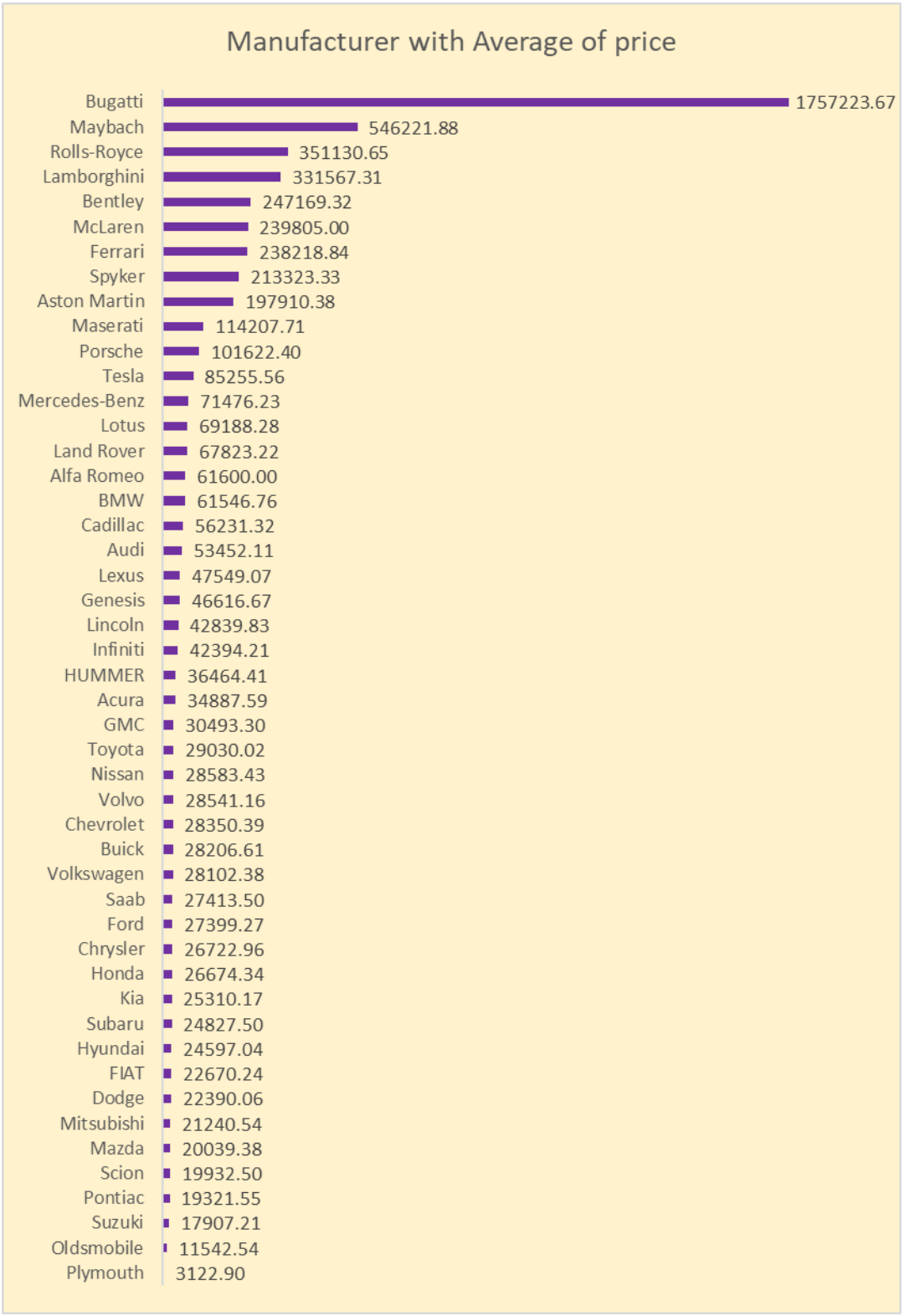


D. How does the average price of a car vary across different manufacturers?

Task D.1: Create a pivot table that shows the average price of cars for each manufacturer.

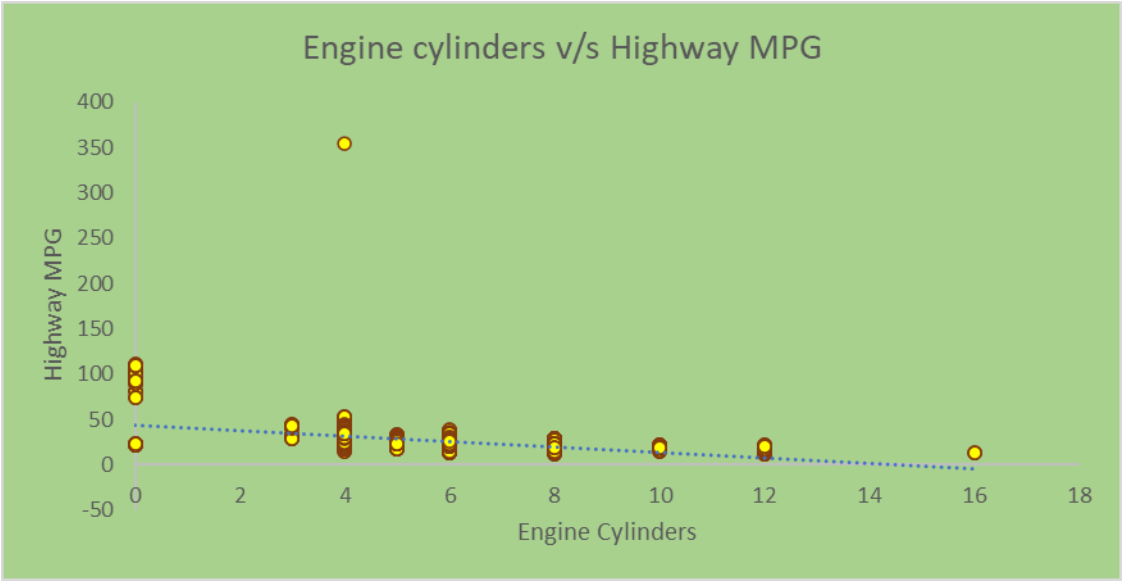
Manufacturer	Average of MSRP
Plymouth	3122.90
Oldsmobile	11542.54
Suzuki	17907.21
Pontiac	19321.55
Scion	19932.50
Mazda	20039.38
Mitsubishi	21240.54
Dodge	22390.06
FIAT	22670.24
Hyundai	24597.04
Subaru	24827.50
Kia	25310.17
Honda	26674.34
Chrysler	26722.96
Ford	27399.27
Saab	27413.50
Volkswagen	28102.38
Buick	28206.61
Chevrolet	28350.39
Volvo	28541.16
Nissan	28583.43
Toyota	29030.02
GMC	30493.30
Acura	34887.59
HUMMER	36464.41
Infiniti	42394.21
Lincoln	42839.83
Genesis	46616.67
Lexus	47549.07
Audi	53452.11
Cadillac	56231.32
BMW	61546.76
Alfa Romeo	61600.00
Land Rover	67823.22
Lotus	69188.28
Mercedes-Benz	71476.23
Tesla	85255.56
Porsche	101622.40
Maserati	114207.71
Aston Martin	197910.38
Spyker	213323.33
Ferrari	238218.84
McLaren	239805.00
Bentley	247169.32
Lamborghini	331567.31
Rolls-Royce	351130.65
Maybach	546221.88
Bugatti	1757223.67

Task D.2: Create a bar chart or a horizontal stacked bar chart that visualizes the relationship between manufacturer and average price.



E. What is the relationship between fuel efficiency and the number of cylinders in a car's engine?

Task E.1: 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.



Significance is -3.025

Task E.2: Calculate the correlation coefficient between the number of cylinders and highway MPG to quantify the strength and direction of the relationship.

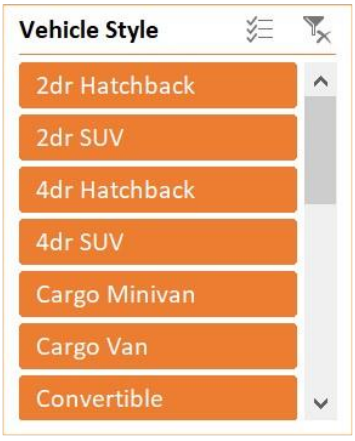
Correlation coefficient	-0.61454

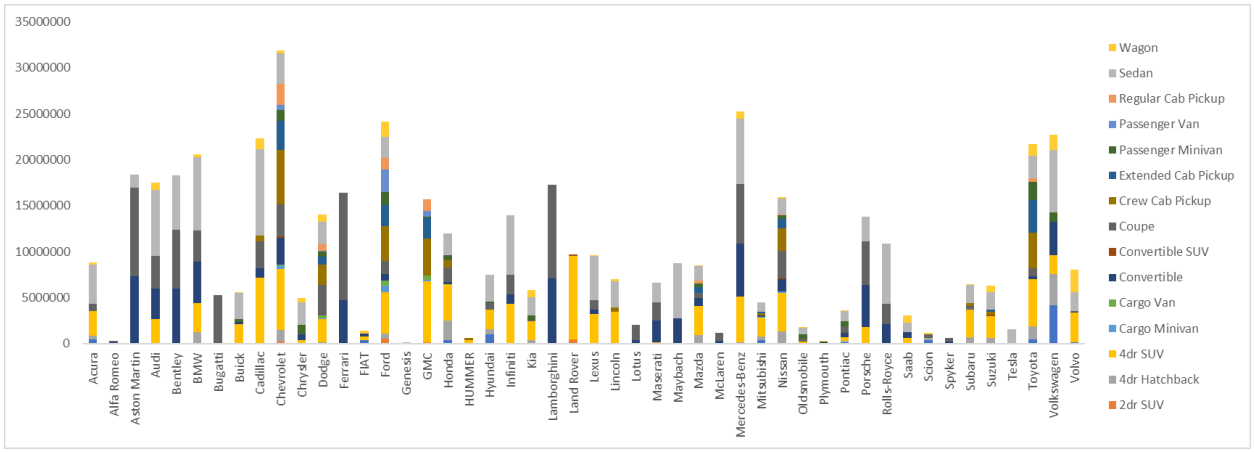
Direction of relationship is negative.

It indicates a strong negative linear relationship (as the number of cylinders increases, highway MPG tends to decrease)

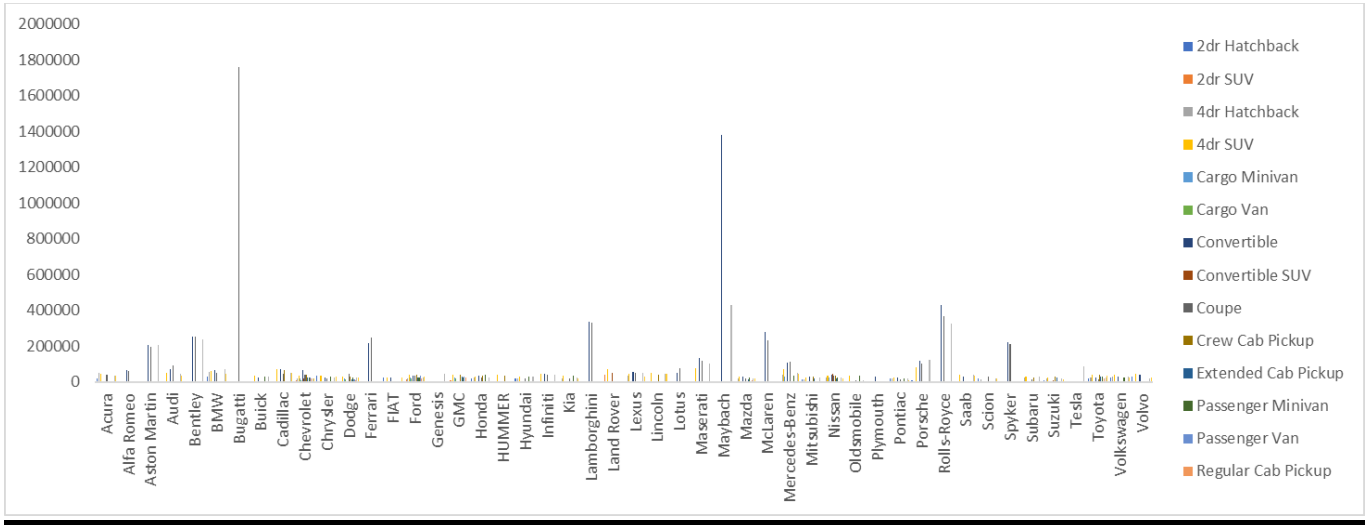
Dashboard Tasks

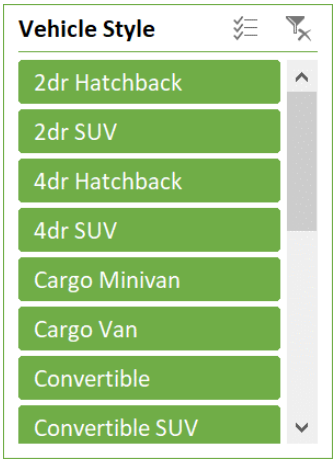
1. How does the distribution of car prices vary by brand and body style?





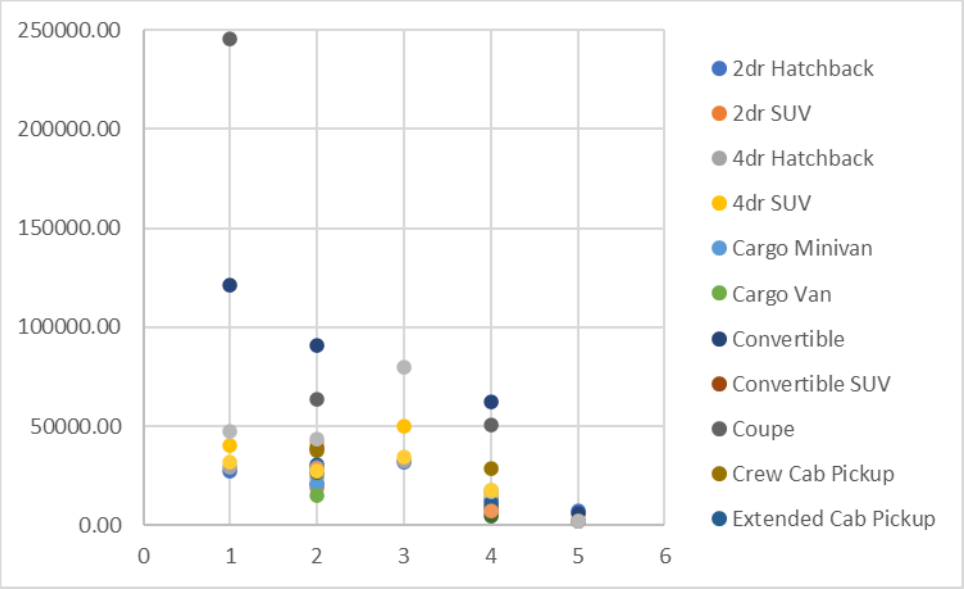
2. Which car brands have the highest and lowest average MSRPs, and how does this vary by body style?





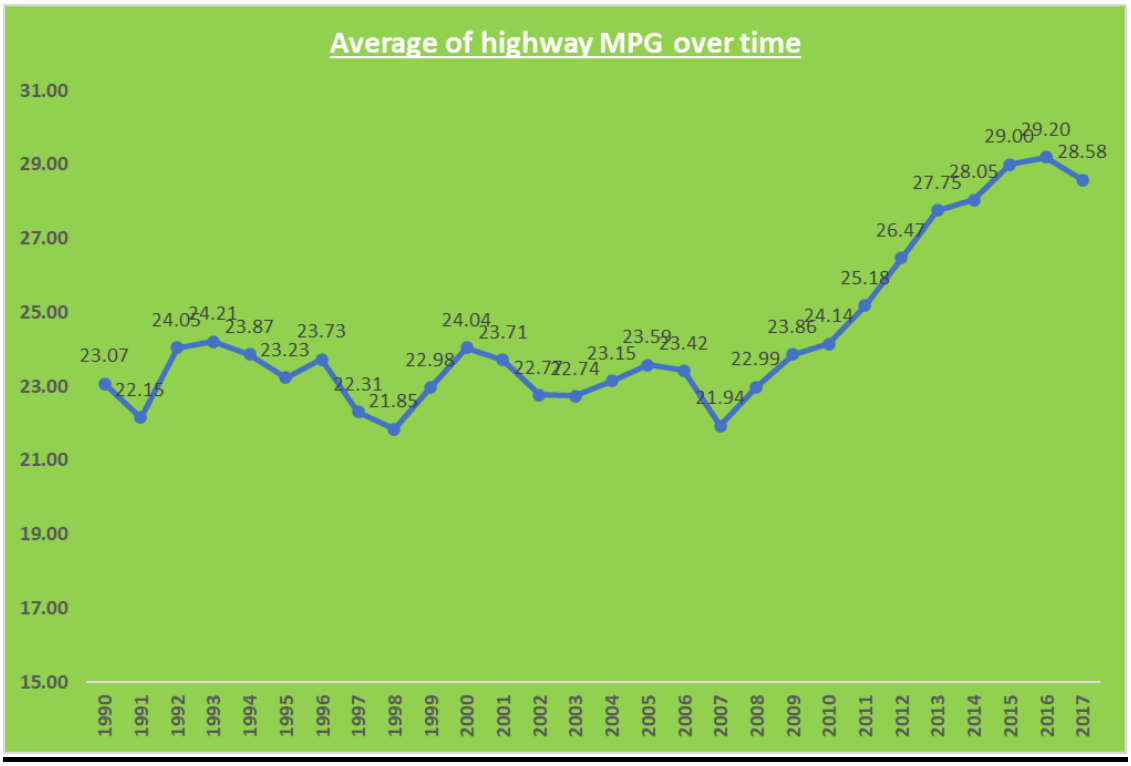
Highest- Bugati coupe, Lowest- Audi, Dodge, Ford, etc.

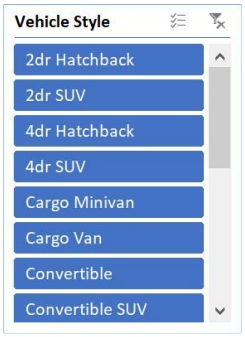
3. How do the different feature such as transmission type affect the MSRP, and how does this vary by body style?



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

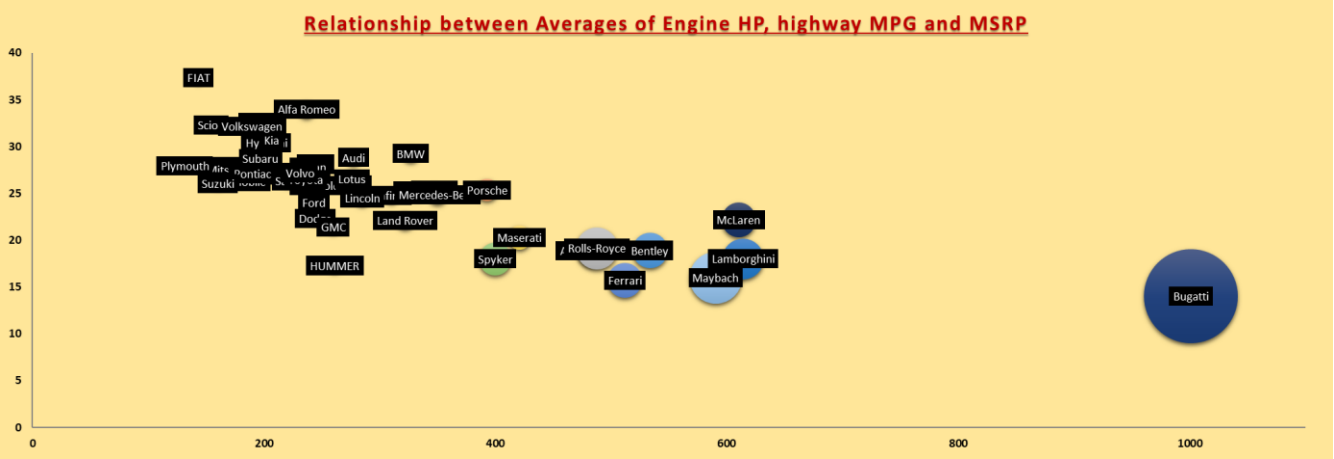
Vehicle Style	Average of highway MPG
1990	23.07
1991	22.15
1992	24.05
1993	24.21
1994	23.87
1995	23.23
1996	23.73
1997	22.31
1998	21.85
1999	22.98
2000	24.04
2001	23.71
2002	22.77
2003	22.74
2004	23.15
2005	23.59
2006	23.42
2007	21.94
2008	22.99
2009	23.86
2010	24.14
2011	25.18
2012	26.47
2013	27.75
2014	28.05
2015	29.00
2016	29.20
2017	28.58


Average of highway MPG over time
31.00
29.00
27.00
25.00
23.00
21.00
19.00
17.00
15.00
23.07
22.15
24.05
24.21
23.87
23.23
23.73
22.31
21.85
22.98
24.04
23.71
22.77
22.74
23.15
23.59
23.42
21.94
22.99
23.86
24.14
25.18
26.47
27.75
28.05
29.00
29.20
28.58
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017

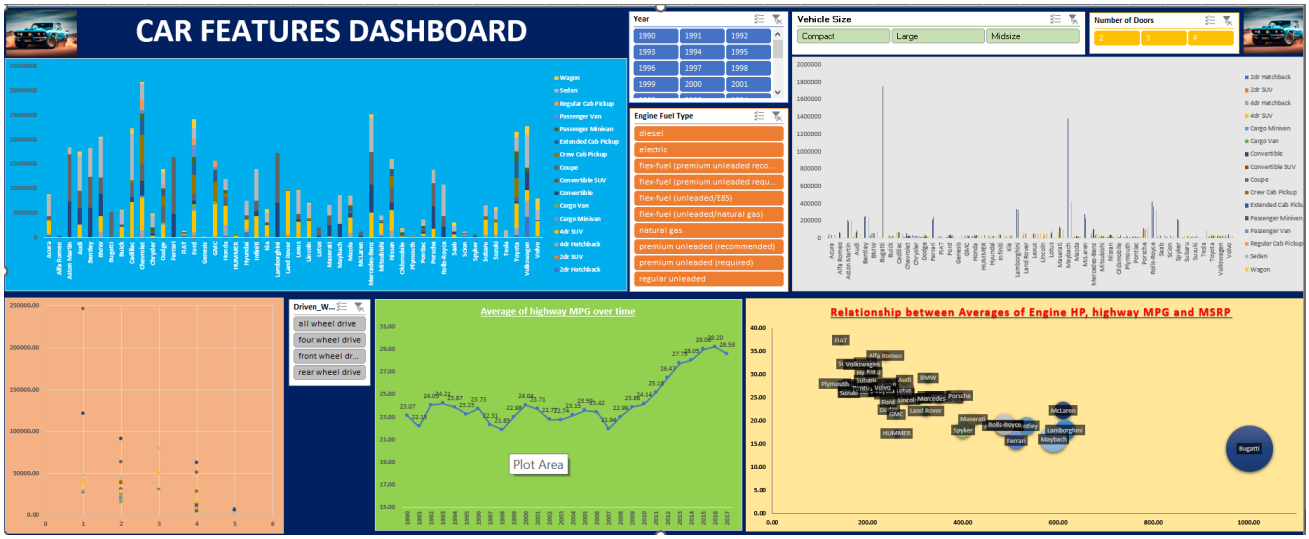

Vehicle Style
2dr Hatchback
2dr SUV
4dr Hatchback
4dr SUV
Cargo Minivan
Cargo Van
Convertible
Convertible SUV

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

Manufacture	Average of Engine HP	Average of highway MPG	Average of MSRP
Acura	244.80	28.11	34887.59
Alfa Romeo	237.00	34.00	61600.00
Aston Martin	484.32	18.89	197910.38
Audi	277.70	28.82	53452.11
Bentley	533.85	18.91	247169.32
BMW	326.91	29.25	61546.76
Bugatti	1001.00	14.00	1757223.67
Buick	219.24	26.95	28206.61
Cadillac	332.31	25.24	56231.32
Chevrolet	246.97	25.82	28350.39
Chrysler	229.14	26.37	26722.96
Dodge	244.42	22.35	22390.06
Ferrari	511.96	15.72	238218.84
FIAT	143.56	37.34	22670.24
Ford	243.10	24.01	27399.27
Genesis	347.33	25.33	46616.67
GMC	259.84	21.40	30493.30
Honda	195.75	32.57	26674.34
HUMMER	261.24	17.29	36464.41
Hyundai	201.92	30.39	24597.04
Infiniti	310.07	24.78	42394.21
Kia	206.83	30.65	25310.17
Lamborghini	614.08	18.02	331567.31
Land Rover	322.10	22.13	67823.22
Lexus	277.42	25.88	47549.07
Lincoln	284.91	24.49	42839.83
Lotus	275.97	26.55	69188.28
Maserati	420.79	20.29	114207.71
Maybach	590.50	16.00	546221.88
Mazda	171.99	27.85	20039.38
McLaren	610.40	22.20	239805.00
Mercedes-Benz	350.18	24.83	71476.23
Mitsubishi	173.43	27.54	21240.54
Nissan	239.92	27.80	28583.43
Oldsmobile	177.47	26.23	11542.54
Plymouth	131.56	27.96	3122.90
Pontiac	190.30	27.07	19321.55
Porsche	392.79	25.37	101622.40
Rolls-Royce	487.55	19.13	351130.65
Saab	220.52	26.35	27413.50
Scion	154.43	32.30	19932.50
Spyker	400.00	18.00	213323.33
Subaru	197.31	28.68	24827.50
Suzuki	160.29	26.03	17907.21
Toyota	236.15	26.45	29030.02
Volkswagen	189.76	32.13	28102.38
Volvo	230.97	27.20	28541.16



DASHBOARD



Dataset Drive Link

Analysis File - car_data analysis file click here

Video Link - Video link click here

Conclusion

Here are the solutions given for the given tasks need to do as the data analyst.

In this task, used the concepts of statistics and advanced excel with basic and advanced topics to create tables, charts and dashboard. These tables, charts and dashboard were implemented using the Microsoft Office Excel.

THANK YOU.