

# Analyzing the Impact of Car Features on Price and Profitability



By Aayushi Singh

# PROJECT DESCRIPTION

The automotive industry has been rapidly evolving over the past few decades, with a growing focus on fuel efficiency, environmental sustainability, and technological innovation. With increasing competition among manufacturers and a changing consumer landscape, it has become more important than ever to understand the factors that drive consumer demand for cars.

In recent years, there has been a growing trend towards electric and hybrid vehicles and increased interest in alternative fuel sources such as hydrogen and natural gas. At the same time, traditional gasoline-powered cars remain dominant in the market, with varying fuel types and grades available to consumers.

The major question is how can a car manufacturer optimize pricing and product development decisions to maximize profitability while meeting consumer demand?

This problem could be approached by analyzing 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. By using data analysis techniques such as regression analysis and market segmentation, the manufacturer could develop a pricing strategy that balances consumer demand with profitability, and identify which product features to focus on in future product development efforts. This could help the manufacturer improve its competitiveness in the market and increase its profitability over time.

The dataset contains information on various car models and their specifications, and is titled "Car Features and MSRP". It was collected and made available on Kaggle by Cooper Union, a private college located in New York City.

Dataset Link: [Car Features and MSRP](#)

brief 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

The process of Data Cleaning is a vital process for analysis to get accurate analysis, including removing the blank cells, removing the outliers from the data, making sure the datatype of column is according to data stored in the respective cell and deleting the duplicate values from the table. Assuming that dataset is untouched and contains errors.

## APPROACH:

To Analyze this dataset by using data analysis techniques such as regression analysis comparing different segments with others to understand how they factor each other and understanding to correlations between the variables in the dataset this lead up to analyse various questions such as Exploring trends in car features and pricing over time ,Comparing the fuel efficiency of different types of cars investigating the relationship between a car's features and its popularity predicting the price of a car based on its features and market category. This helps in analyzing different components of variables of the dataset helps us to analyse which factors will affect the sales of a product in the market and what steps can be taken in that variable to further increase the sales in the future .

## TECH STACK USED

Software used for the project is Excel , MS Word, MS Sheets

To conclude insights on the problem statement on dataset we require data analysis techniques such as regression analysis, pivot tables, sensitivity analysis, optimization, and time series analysis. Excel also provides an interactive dashboard, which through the help of filter and slicer can help in extensively analyzing charts or graphs

## INSIGHTS

The Analysis : [Car Dataset](#)

Analyzing trends in car features and pricing over time: By examining the variables in the dataset, a data analyst could identify how car features and prices have changed over time, which could help manufacturers make informed decisions about product development and pricing.

Comparing the fuel efficiency of different types of cars: By looking at the MPG variables in the dataset, a data analyst could compare the fuel efficiency of different types of cars and identify which types are the most efficient. This could help consumers make informed decisions about which car to purchase.

Investigating the relationship between a car's features and its popularity: By examining the popularity variable in the dataset, a data analyst could identify which features are most popular among consumers and how they affect a car's popularity. This could help manufacturers make informed decisions about product development and marketing.

Predicting the price of a car based on its features and market category: By using the various features and market category variables in the dataset, a data analyst could develop a model to predict the price of a car. This could help manufacturers and consumers understand how different features affect the price of a car and make informed decisions about pricing and purchasing.

Overall, this dataset could be a valuable resource for data analysts interested in exploring various aspects of the automotive industry and could provide insights that could inform decisions related to product development, marketing, and pricing.

By analysing the popularity of a car model vary across different market categories, through which can infer that if by ignoring the Unknown category the most purchase in terms of bought models are crossover and luxury and market category in terms of popularity are fuel Flex and crossover

By analysing the Car's Engine Power and its price we can observe that the engine power in the range of 225-304 are the high on the price range with it trendline in terms of moving average with its peak at 225-304 and then decline in price

By observing the Regression Table we observe that it has the lowest correlation with popularity and highest with city MPG if not considering engine cylinders then with highway MPG

through the bar chart that by the average MSRP and Manufacturers , the companies with highest average MSRP are Maybach,Rolls-Royce, Lamborghini, Spyker

As by observing the scatter plot between highway mpg and number of cylinder we observe that its graph peaks when number of cylinder are 4 and after that its a steady decline in highway mpg as the number of cylinders increases . the correlation coefficient between number of cylinder and highway mpg is -0.123884693 since its value is close to 0 we can conclude that it has a very weak correlation

# RESULTS

## 1. Cleaning the data

A	B	C	D	E	F	G	H	I	J	K	L	M	N	
Make	Model	Year	Engine Fuel Type	Engine HP	Engine Cylinders	Transmission Type	Driven_Wheels	Number of Doors	Market Category	Vehicle Size	Vehicle Style	highway MPG	city mpg	Popu
BMW	1 Series M	2011	premium unleaded (required)	335	6 MANUAL	rear wheel drive	2	Factory Tuner,Luxur	Compact	Coupe	26	19		
BMW	1 Series	2011	premium unleaded (required)	300	6 MANUAL	rear wheel drive	2	Luxury,Performance	Compact	Convertible	28	19		
BMW	1 Series	2011	premium unleaded (required)	300	6 MANUAL	rear wheel drive	2	Luxury,High-Perform	Compact	Coupe	28	20		
BMW	1 Series	2011	premium unleaded (required)	230	6 MANUAL	rear wheel drive	2	Luxury,Performance	Compact	Coupe	28	18		
BMW	1 Series	2011	premium unleaded (required)	230	6 MANUAL	rear wheel drive	2	Luxury	Compact	Convertible	28	18		
BMW	1 Series	2012	premium unleaded (required)	230	6 MANUAL	rear wheel drive	2	Luxury,Performance	Compact	Coupe	28	18		
BMW	1 Series	2012	premium unleaded (required)	300	6 MANUAL	rear wheel drive	2	Luxury,Performance	Compact	Convertible	26	17		
BMW	1 Series	2012	premium unleaded (required)	300	6 MANUAL	rear wheel drive	2	Luxury,High-Perform	Compact	Coupe	28	20		
BMW	1 Series	2012	premium unleaded (required)	230	6 MANUAL	rear wheel drive	2	Luxury	Compact	Convertible	28	18		
BMW	1 Series	2013	premium unleaded (required)	230	6 MANUAL	rear wheel drive	2	Luxury	Compact	Convertible	27	18		
BMW	1 Series	2013	premium unleaded (required)	300	6 MANUAL	rear wheel drive	2	Luxury,High-Perform	Compact	Coupe	28	20		
BMW	1 Series	2013	premium unleaded (required)	230	6 MANUAL	rear wheel drive	2	Luxury,Performance	Compact	Coupe	28	19		
BMW	1 Series	2013	premium unleaded (required)	300	6 MANUAL	rear wheel drive	2	Luxury,Performance	Compact	Convertible	28	19		
BMW	1 Series	2013	premium unleaded (required)	230	6 MANUAL	rear wheel drive	2	Luxury	Compact	Convertible	28	19		
BMW	1 Series	2013	premium unleaded (required)	320	6 MANUAL	rear wheel drive	2	Luxury,High-Perform	Compact	Convertible	25	18		
BMW	1 Series	2013	premium unleaded (required)	320	6 MANUAL	rear wheel drive	2	Luxury,High-Perform	Compact	Coupe	28	20		
Audi	100	1992	regular unleaded	172	6 MANUAL	front wheel drive	4	Luxury	Midsiz	Sedan	24	17		
Audi	100	1992	regular unleaded	172	6 AUTOMATIC	all wheel drive	4	Luxury	Midsiz	Wagon	20	16		
Audi	100	1992	regular unleaded	172	6 MANUAL	all wheel drive	4	Luxury	Midsiz	Sedan	21	16		
Audi	100	1993	regular unleaded	172	6 MANUAL	front wheel drive	4	Luxury	Midsiz	Sedan	24	17		
Audi	100	1993	regular unleaded	172	6 AUTOMATIC	all wheel drive	4	Luxury	Midsiz	Wagon	20	16		
Audi	100	1993	regular unleaded	172	6 MANUAL	all wheel drive	4	Luxury	Midsiz	Sedan	21	16		
Audi	100	1994	regular unleaded	172	6 AUTOMATIC	front wheel drive	4	Luxury	Midsiz	Wagon	21	16		
Audi	100	1994	regular unleaded	172	6 MANUAL	all wheel drive	4	Luxury	Midsiz	Sedan	22	16		
Audi	100	1994	regular unleaded	172	6 MANUAL	front wheel drive	4	Luxury	Midsiz	Sedan	22	17		
Audi	100	1994	regular unleaded	172	6 AUTOMATIC	front wheel drive	4	Luxury	Midsiz	Sedan	22	16		
Audi	100	1994	regular unleaded	172	6 AUTOMATIC	all wheel drive	4	Luxury	Midsiz	Wagon	21	16		
FIAT	124 Spider	2017	premium unleaded (recommended)	160	4 MANUAL	rear wheel drive	2	Performance	Compact	Convertible	35	26		
FIAT	124 Spider	2017	premium unleaded (recommended)	160	4 MANUAL	rear wheel drive	2	Performance	Compact	Convertible	35	26		
FIAT	124 Spider	2017	premium unleaded (recommended)	160	4 MANUAL	rear wheel drive	2	Performance	Compact	Convertible	35	26		
Mercedes	190-Class	1991	regular unleaded	130	4 MANUAL	rear wheel drive	4	Luxury	Compact	Sedan	26	18		
Mercedes	190-Class	1991	regular unleaded	158	6 MANUAL	rear wheel drive	4	Luxury	Compact	Sedan	25	17		

The Dataset

Screenshot of Microsoft Excel showing a PivotTable named "Table1" in "Table Design" mode. The PivotTable has columns for Make, Model, Year, Engine Fuel Type, Engine HP, Engine Cylinders, Transmission Type, Driven\_Wheels, Number of Doors, Market Category, Vehicle Size, Vehicle Style, highway MPG, and city mpg. The PivotTable displays data for various vehicles, including Chevrolet Bolt EV, Volkswagen e-Golf, Toyota RAV4 EV, and Mazda RX-8. The "Table Styles" section shows several color-coded styles for the table.

Make	Model	Year	Engine Fuel Type	Engine HP	Engine Cylinders	Transmission Type	Driven_Wheels	Number of Doors	Market Category	Vehicle Size	Vehicle Style	highway MPG	city mpg
382	Chevrolet	Bolt EV	2017 electric		200	DIRECT_DRIVE	front wheel drive	4	Hatchback	Compact	4dr Hatchback	110	
383	Chevrolet	Bolt EV	2017 electric		200	DIRECT_DRIVE	front wheel drive	4	Hatchback	Compact	4dr Hatchback	110	
711	Volkswagen	e-Golf	2015 electric		115	DIRECT_DRIVE	front wheel drive	4	Hatchback	Compact	4dr Hatchback	105	
712	Volkswagen	e-Golf	2015 electric		115	DIRECT_DRIVE	front wheel drive	4	Hatchback	Compact	4dr Hatchback	105	
713	Volkswagen	e-Golf	2016 electric		115	DIRECT_DRIVE	front wheel drive	4	Hatchback	Compact	4dr Hatchback	105	
714	Volkswagen	e-Golf	2016 electric		115	DIRECT_DRIVE	front wheel drive	4	Hatchback	Compact	4dr Hatchback	105	
758	Mitsubishi	i-MiEV	2016 electric		66	DIRECT_DRIVE	rear wheel drive	4	Hatchback	Compact	4dr Hatchback	99	
759	Mitsubishi	i-MiEV	2017 electric		66	DIRECT_DRIVE	rear wheel drive	4	Hatchback	Compact	4dr Hatchback	102	
313	Toyota	RAV4 EV	2012 electric		154	DIRECT_DRIVE	front wheel drive	4	Crossover	Midsized	4dr SUV	74	
533	Mazda	RX-7	1993 regular unleaded		255	MANUAL	rear wheel drive	2	Factory Tuner	Perfomance	Coupe	23	
534	Mazda	RX-7	1994 regular unleaded		255	MANUAL	rear wheel drive	2	Factory Tuner	Perfomance	Coupe	23	
535	Mazda	RX-7	1995 regular unleaded		255	MANUAL	rear wheel drive	2	Factory Tuner	Perfomance	Coupe	23	
536	Mazda	RX-8	2009 premium unleaded (required)		232	MANUAL	rear wheel drive	4	Performance	Compact	Coupe	22	
537	Mazda	RX-8	2009 premium unleaded (required)		212	AUTOMATIC	rear wheel drive	4	Performance	Compact	Coupe	23	
538	Mazda	RX-8	2009 premium unleaded (required)		232	MANUAL	rear wheel drive	4	Performance	Compact	Coupe	22	
539	Mazda	RX-8	2009 premium unleaded (required)		232	MANUAL	rear wheel drive	4	Performance	Compact	Coupe	22	
540	Mazda	RX-8	2009 premium unleaded (required)		232	MANUAL	rear wheel drive	4	Performance	Compact	Coupe	22	
541	Mazda	RX-8	2009 premium unleaded (required)		212	AUTOMATIC	rear wheel drive	4	Performance	Compact	Coupe	23	
542	Mazda	RX-8	2009 premium unleaded (required)		212	AUTOMATIC	rear wheel drive	4	Performance	Compact	Coupe	23	
543	Mazda	RX-8	2010 premium unleaded (required)		232	MANUAL	rear wheel drive	4	Performance	Compact	Coupe	22	
544	Mazda	RX-8	2010 premium unleaded (required)		212	AUTOMATIC	rear wheel drive	4	Performance	Compact	Coupe	23	
545	Mazda	RX-8	2010 premium unleaded (required)		212	AUTOMATIC	rear wheel drive	4	Performance	Compact	Coupe	23	
546	Mazda	RX-8	2010 premium unleaded (required)		232	MANUAL	rear wheel drive	4	Performance	Compact	Coupe	22	
547	Mazda	RX-8	2010 premium unleaded (required)		232	MANUAL	rear wheel drive	4	Performance	Compact	Coupe	22	
548	Mazda	RX-8	2011 premium unleaded (required)		212	AUTOMATIC	rear wheel drive	4	Performance	Compact	Coupe	23	
549	Mazda	RX-8	2011 premium unleaded (required)		232	MANUAL	rear wheel drive	4	Performance	Compact	Coupe	22	
550	Mazda	RX-8	2011 premium unleaded (required)		232	MANUAL	rear wheel drive	4	Performance	Compact	Coupe	22	
551	Mazda	RX-8	2011 premium unleaded (required)		232	MANUAL	rear wheel drive	4	Performance	Compact	Coupe	22	

## Blanks in Engine Cylinders

Screenshot of Microsoft Excel showing a PivotTable in "Home" mode. The PivotTable has columns for Make, Model, Year, Engine Fuel Type, Engine HP, Engine Cylinders, Transmission Type, Driven\_Wheels, Number of Doors, Market Category, Vehicle Size, Vehicle Style, highway MPG, city mpg, and Popularity. The PivotTable displays data for various vehicles, including Suzuki Verona. The PivotTable interface includes the Home tab, Font, Alignment, Number, Styles, Cells, and Editing sections.

Make	Model	Year	Engine Fuel Type	Engine HP	Engine Cylinders	Transmission Type	Driven_Wheels	Number of Doors	Market Category	Vehicle Size	Vehicle Style	highway MPG	city mpg	Popularity
323	Suzuki	Verona	2004		155	6 AUTOMATIC	front wheel drive	4	N/A	Midsize	Sedan	25	17	48
324	Suzuki	Verona	2004		155	6 AUTOMATIC	front wheel drive	4	N/A	Midsize	Sedan	25	17	48
325	Suzuki	Verona	2004		155	6 AUTOMATIC	front wheel drive	4	N/A	Midsize	Sedan	25	17	48
916														
917														
918														
919														
920														
921														
922														
923														
924														
925														
926														
927														
928														
929														
930														
931														
932														
933														
934														
935														
936														

## Blanks in engine Fuel Type

Suzuki

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Make	Model	Year	Engine Fuel Type	Engine HP	Engine Cylinders	Transmission Type	Driven_Wheels	Number of Doors	Market Category	Vehicle Size	Vehicle Style	highway MPG	city mpg
541	FIAT	500e	2015	electric		0	DIRECT_DRIVE	front wheel drive	2	Hatchback	Compact	2dr Hatchback	108	
542	FIAT	500e	2016	electric		0	DIRECT_DRIVE	front wheel drive	2	Hatchback	Compact	2dr Hatchback	103	
543	FIAT	500e	2017	electric		0	DIRECT_DRIVE	front wheel drive	2	Hatchback	Compact	2dr Hatchback	103	
2907	Lincoln	Continent	2017	premium unleaded (recommended)		6	AUTOMATIC	all wheel drive	4	Luxury	Large	Sedan	25	
2908	Lincoln	Continent	2017	premium unleaded (recommended)		6	AUTOMATIC	front wheel drive	4	Luxury	Large	Sedan	27	
2909	Lincoln	Continent	2017	premium unleaded (recommended)		6	AUTOMATIC	front wheel drive	4	Luxury	Large	Sedan	27	
2910	Lincoln	Continent	2017	premium unleaded (recommended)		6	AUTOMATIC	all wheel drive	4	Luxury	Large	Sedan	25	
4205	Ford	Escape	2017	regular unleaded		4	AUTOMATIC	front wheel drive	4	Crossover	Compact	4dr SUV	30	
4206	Ford	Escape	2017	regular unleaded		4	AUTOMATIC	all wheel drive	4	Crossover	Compact	4dr SUV	28	
4207	Ford	Escape	2017	regular unleaded		4	AUTOMATIC	all wheel drive	4	Crossover	Compact	4dr SUV	28	
4208	Ford	Escape	2017	regular unleaded		4	AUTOMATIC	front wheel drive	4	Crossover	Compact	4dr SUV	30	
4707	Honda	Fit EV	2013	electric		0	DIRECT_DRIVE	front wheel drive	4	Hatchback	Compact	4dr Hatchback	105	
4708	Honda	Fit EV	2014	electric		0	DIRECT_DRIVE	front wheel drive	4	Hatchback	Compact	4dr Hatchback	105	
4787	Ford	Focus	2015	electric		0	DIRECT_DRIVE	front wheel drive	4	Hatchback	Compact	4dr Hatchback	99	
4791	Ford	Focus	2016	electric		0	DIRECT_DRIVE	front wheel drive	4	Hatchback	Compact	4dr Hatchback	99	
4800	Ford	Focus	2017	electric		0	DIRECT_DRIVE	front wheel drive	4	Hatchback	Compact	4dr Hatchback	99	
4916	Ford	Freestar	2005	regular unleaded		6	AUTOMATIC	front wheel drive	4	N/A	Midsized	Pasenger Miniv	22	
4917	Ford	Freestar	2005	regular unleaded		6	AUTOMATIC	front wheel drive	4	N/A	Midsized	Pasenger Miniv	22	
4918	Ford	Freestar	2005	regular unleaded		6	AUTOMATIC	front wheel drive	4	N/A	Midsized	Cargo Minivan	22	
4919	Ford	Freestar	2005	regular unleaded		6	AUTOMATIC	front wheel drive	4	N/A	Midsized	Pasenger Miniv	22	
4920	Ford	Freestar	2005	regular unleaded		6	AUTOMATIC	front wheel drive	4	N/A	Midsized	Pasenger Miniv	21	
4921	Ford	Freestar	2005	regular unleaded		6	AUTOMATIC	front wheel drive	4	N/A	Midsized	Pasenger Miniv	21	
5780	Mitsubishi	i-MiEV	2014	electric		0	DIRECT_DRIVE	rear wheel drive	4	Hatchback	Compact	4dr Hatchback	99	
5827	Chevrolet	Impala	2015	flex-fuel (unleaded/natural gas)		6	AUTOMATIC	front wheel drive	4	Flex Fuel,Perfomance	Large	Sedan	25	
5832	Chevrolet	Impala	2015	flex-fuel (unleaded/natural gas)		6	AUTOMATIC	front wheel drive	4	Flex Fuel,Perfomance	Large	Sedan	25	
5833	Chevrolet	Impala	2016	flex-fuel (unleaded/natural gas)		6	AUTOMATIC	front wheel drive	4	Flex Fuel,Perfomance	Large	Sedan	25	
5835	Chevrolet	Impala	2016	flex-fuel (unleaded/natural gas)		6	AUTOMATIC	front wheel drive	4	Flex Fuel,Perfomance	Large	Sedan	25	
5841	Chevrolet	Impala	2017	flex-fuel (unleaded/natural gas)		6	AUTOMATIC	front wheel drive	4	Flex Fuel,Perfomance	Large	Sedan	25	

## Blanks in engine hp

Pontiac

	A	B	C	D	E	F	G	H	I	J	K	L	M	
1	Make	Model	Year	Engine Fuel Type	Engine HP	Engine Cylinders	Transmission Type	Driven_Wheels	Number of Doors	Market Category	Vehicle Size	Vehicle Style	highway MPG	city mpg
4651	Ferrari	FF	2013	premium unleaded (required)		651	12	AUTOMATED_MANUAL	all wheel drive	Exotic,High-Perform-Large		Coupe		
11815														
11816														
11817														
11818														
11819														
11820														
11821														
11822														
11823														
11824														
11825														
11826														
11827														
11828														
11829														
11830														
11831														
11832														
11833														
11834														
11835														
11836														
11837														

## Blanks in no of Doors

The screenshot shows a Microsoft Excel spreadsheet titled "Market Category". The table has columns for Engine HP, Engine Cylinders, Transmission Type, Driven\_Wheels, Number of Doors, Market Category, Vehicle Size, Vehicle Style, highway MPG, city mpg, Popularity, and MSRP. The data includes rows for different vehicle models like BMW and Audi across various years and configurations.

	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Engine HP	Engine Cylinders	Transmission Type	Driven_Wheels	Number of Doors	Market Category	Vehicle Size	Vehicle Style	highway MPG	city mpg	Popularity	MSRP	Yr
114	620	12 AUTOMATIC	rear wheel drive	4	Exotic,Luxury,High-P Large	Convertible	16	10	67	1380000			
115	620	12 AUTOMATIC	rear wheel drive	4	Exotic,Luxury,Perfor Large	Convertible	16	10	67	1382750			
114	650	12 AUTOMATED_MANU/all wheel drive		2	Exotic,High-Perform Compact	Coupe	14	9	1158	1500000			
262	1001	16 AUTOMATED_MANU/all wheel drive		2	Exotic,High-Perform Compact	Coupe	14	8	820	2065902			
263	1001	16 AUTOMATED_MANU/all wheel drive		2	Exotic,High-Perform Compact	Coupe	14	8	820	1500000			
264	1001	16 AUTOMATED_MANU/all wheel drive		2	Exotic,High-Perform Compact	Coupe	14	8	820	1705769			
814													
815													
816													
817													
818													
819													
820													
821													
822													
823													
824													
825													
826													
827													
828													
829													
830													
831													
832													
833													
834													

## Anomaly in MSRP

The screenshot shows a Microsoft Excel spreadsheet with a "Remove Duplicates" dialog box open. The dialog box states: "715 duplicate values found and removed; 11091 unique values remain. Note that counts may include empty cells, spaces, etc." The OK button is visible at the bottom of the dialog. The main table below contains data for various car models like BMW and Audi across different years and configurations.

A	B	C	D	E	F	G	H	I	J	K	L	M	
1	Make	Model	Year	Engine Fuel Type	Engine HP	Engine Cylinders	Transmission Type	Driven_Wheels	Number of Doors	Market Category	Vehicle Size	Vehicle Style	high
2	BMW	1 Series M	2011	premium unleaded (required)	335	6 MANUAL	rear wheel drive		2	Factory Tuner,Luxury,Compact	Coupe		
3	BMW	1 Series	2011	premium unleaded (required)	300	6 MANUAL	rear wheel drive		2	Luxury,Performance Compact	Convertible		
4	BMW	1 Series	2011	premium unleaded (required)	300	6 MANUAL	rear wheel drive		2	Luxury,High-Perform Compact	Coupe		
5	BMW	1 Series	2011	premium unleaded (required)	230	6 MANUAL	rear wheel drive		2	Luxury,Performance Compact	Coupe		
5	BMW	1 Series	2011	premium unleaded (required)	230	6 MANUAL	rear wheel drive		2	Luxury	Compact	Convertible	
7	BMW	1 Series	2012	premium unleaded (required)						xury,Performance Compact	Coupe		
3	BMW	1 Series	2012	premium unleaded (required)						xury,Performance Compact	Convertible		
9	BMW	1 Series	2012	premium unleaded (required)						xury,High-Perform Compact	Coupe		
0	BMW	1 Series	2012	premium unleaded (required)						xury	Compact	Convertible	
1	BMW	1 Series	2013	premium unleaded (required)						xury	Compact	Convertible	
2	BMW	1 Series	2013	premium unleaded (required)						xury,High-Perform Compact	Coupe		
3	BMW	1 Series	2013	premium unleaded (required)	230	6 MANUAL	rear wheel drive		2	Luxury,Performance Compact	Coupe		
4	BMW	1 Series	2013	premium unleaded (required)	300	6 MANUAL	rear wheel drive		2	Luxury,Performance Compact	Convertible		
5	BMW	1 Series	2013	premium unleaded (required)	230	6 MANUAL	rear wheel drive		2	Luxury	Compact	Convertible	
6	BMW	1 Series	2013	premium unleaded (required)	320	6 MANUAL	rear wheel drive		2	Luxury,High-Perform Compact	Convertible		
7	BMW	1 Series	2013	premium unleaded (required)	320	6 MANUAL	rear wheel drive		2	Luxury,High-Perform Compact	Coupe		
8	Audi	100	1992	regular unleaded	172	6 MANUAL	front wheel drive		4	Luxury	Midsized	Sedan	
9	Audi	100	1992	regular unleaded	172	6 AUTOMATIC	all wheel drive		4	Luxury	Midsized	Wagon	
10	Audi	100	1992	regular unleaded	172	6 MANUAL	all wheel drive		4	Luxury	Sedan		
11	Audi	100	1993	regular unleaded	172	6 MANUAL	front wheel drive		4	Luxury	Midsized	Sedan	
12	Audi	100	1993	regular unleaded	172	6 AUTOMATIC	all wheel drive		4	Luxury	Midsized	Wagon	
13	Audi	100	1993	regular unleaded	172	6 MANUAL	all wheel drive		4	Luxury	Midsized	Sedan	
14	Audi	100	1994	regular unleaded	172	6 AUTOMATIC	front wheel drive		4	Luxury	Midsized	Wagon	

## Removing Duplicates

## 1. How does the popularity of a car model vary across different market categories

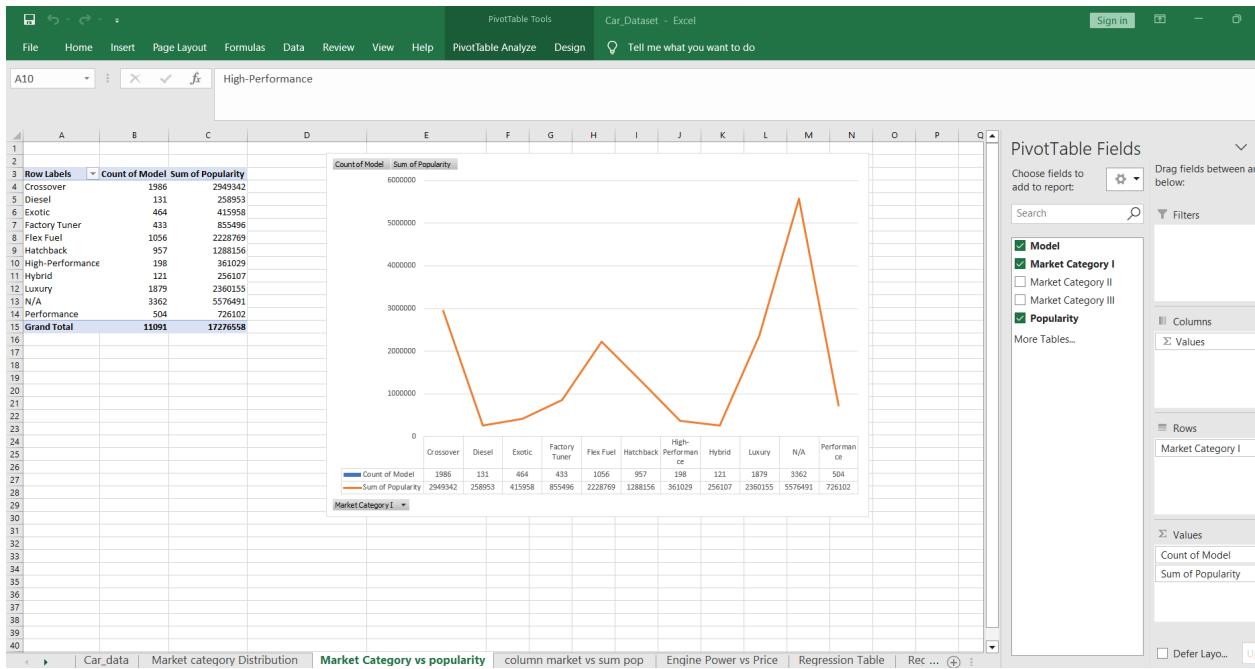
By Create a pivot table that shows the number of car models in each market category and their corresponding popularity scores and Create a combo chart that visualizes the relationship between market category and popularity

The screenshot shows an Excel spreadsheet with two main tables. The first table, titled 'Market category Distribution', contains data for various car models categorized by market type (Performance, High-Performance, Luxury) and popularity score (3916, 3105, 819). The second table, titled 'Market Category - No of Models', lists market categories with their respective counts: Crossover (1986), Diesel (201), Exotic (466), Factory Tuner (606), Flex Fuel (1173), Hatchback (1048), High-Performance (1359), Hybrid (333), Luxury (3211), Performance (3275), and N/A (3362). The bottom of the screen shows a ribbon with tabs like 'Car\_data', 'Market category Distribution' (which is selected), 'Market Category vs popularity', etc.

Model	Market Category	Market Category II	Market Category III	Popularity	F	G	H	I	J	K	L
3 1 Series	Luxury,Performance	Performance		3916							
4 1 Series	Luxury,High-Performance	High-Performance		3916							
5 1 Series	Luxury,Performance	Performance		3916							
6 1 Series	Luxury			3916							
7 1 Series	Luxury,Performance	Performance		3916							
8 1 Series	Luxury,Performance	Performance		3916							
9 1 Series	Luxury,High-Performance	High-Performance		3916							
10 1 Series	Luxury			3916							
11 1 Series	Luxury			3916							
12 1 Series	Luxury,High-Performance	High-Performance		3916							
13 1 Series	Luxury,Performance	Performance		3916							
14 1 Series	Luxury,Performance	Performance		3916							
15 1 Series	Luxury			3916							
16 1 Series	Luxury,High-Performance	High-Performance		3916							
17 1 Series	Luxury,High-Performance	High-Performance		3916							
18 100	Luxury			3105							
19 100	Luxury			3105							
20 100	Luxury			3105							
21 100	Luxury			3105							
22 100	Luxury			3105							
23 100	Luxury			3105							
24 100	Luxury			3105							
25 100	Luxury			3105							
26 100	Luxury			3105							
27 100	Luxury			3105							
28 100	Luxury			3105							
29 124 Spider	Performance			819							
30 124 Spider	Performance			819							
31 124 Spider	Performance			819							
32 190-Class	Luxury			617							
33 190-Class	Luxury			617							
34 190-Class	Luxury			617							

Market Category	No of Models
Crossover	1986
Diesel	201
Exotic	466
Factory Tuner	606
Flex Fuel	1173
Hatchback	1048
High-Performance	1359
Hybrid	333
Luxury	3211
Performance	3275
N/A	3362

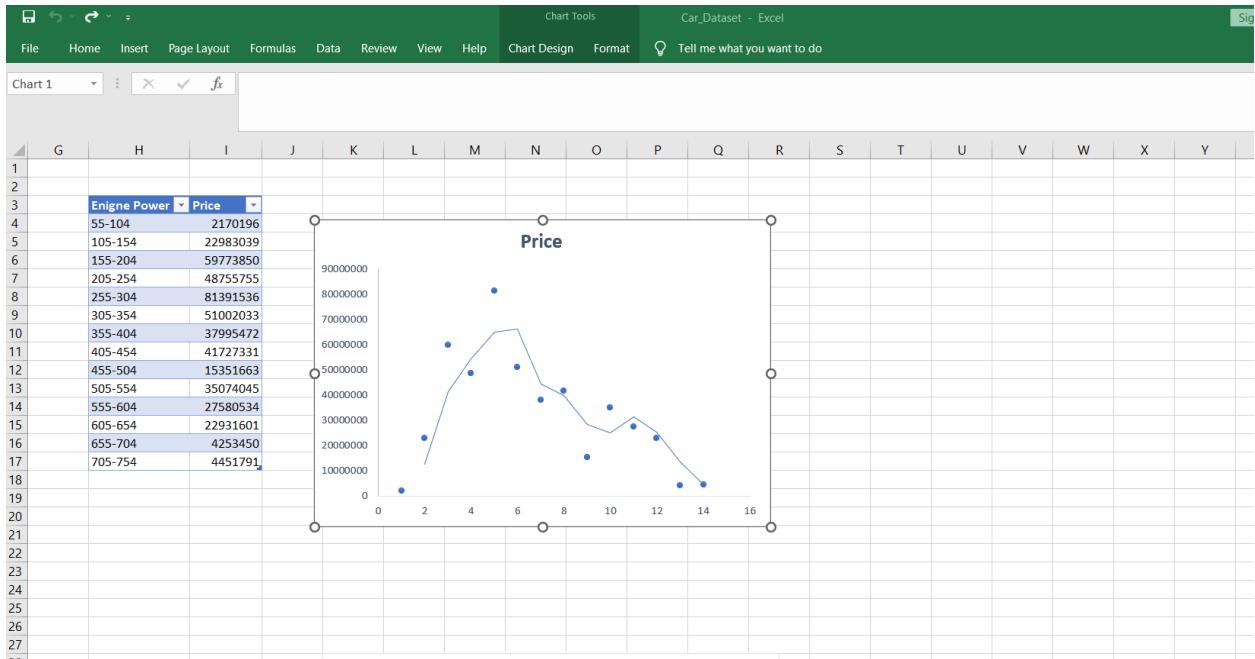
Separating the market category in different columns in a new table, and importing model and popularity, we can count the number of models using count if in combination of wildcard or by using pivot table



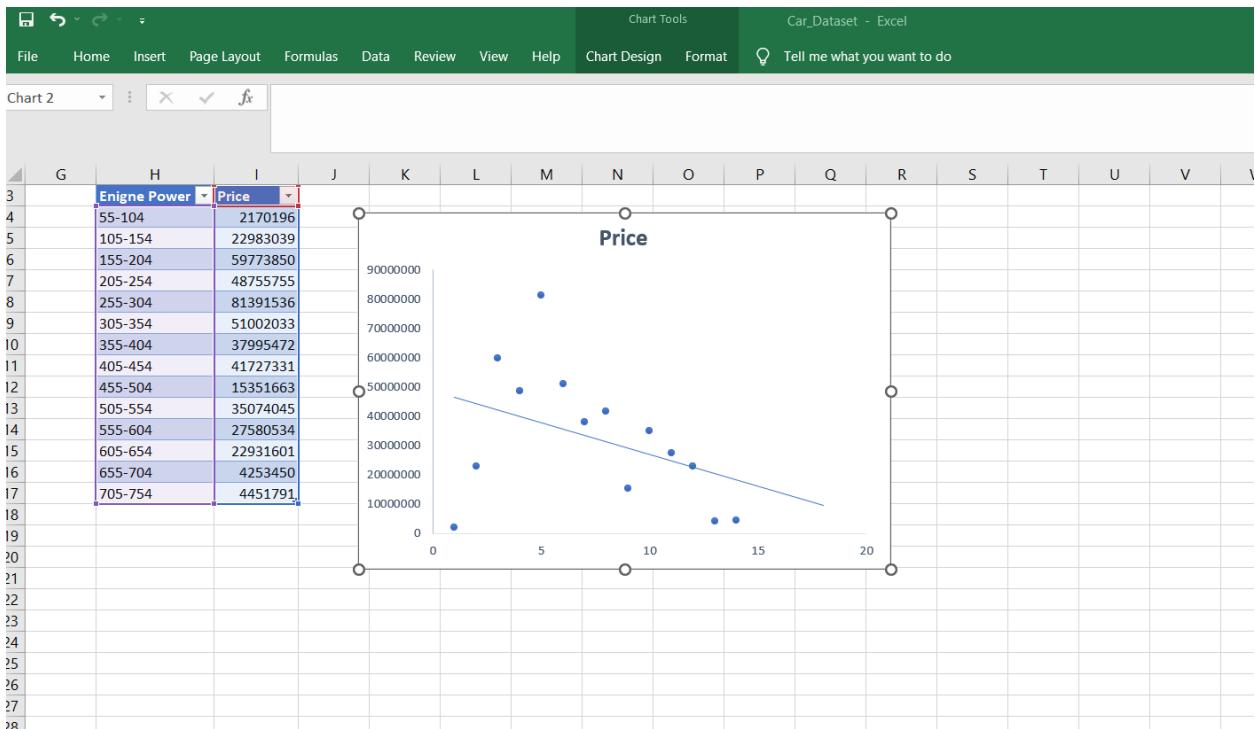
Based on pivot table creating a combo chart

## 2. the relationship between a car's engine power and its price

By creating 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 trendline of moving average to give overall on how the data is varies with engine power and MSRP



A predictive trendline based on linear forecast

### 3. car features are most important in determining a car's price

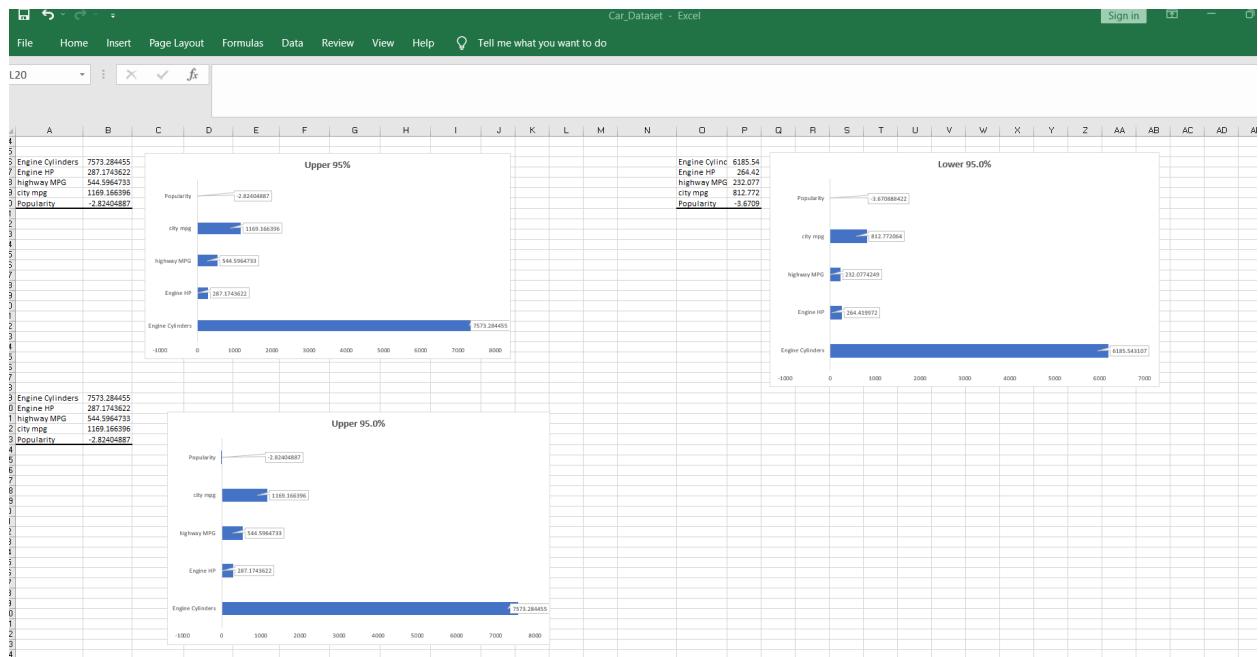
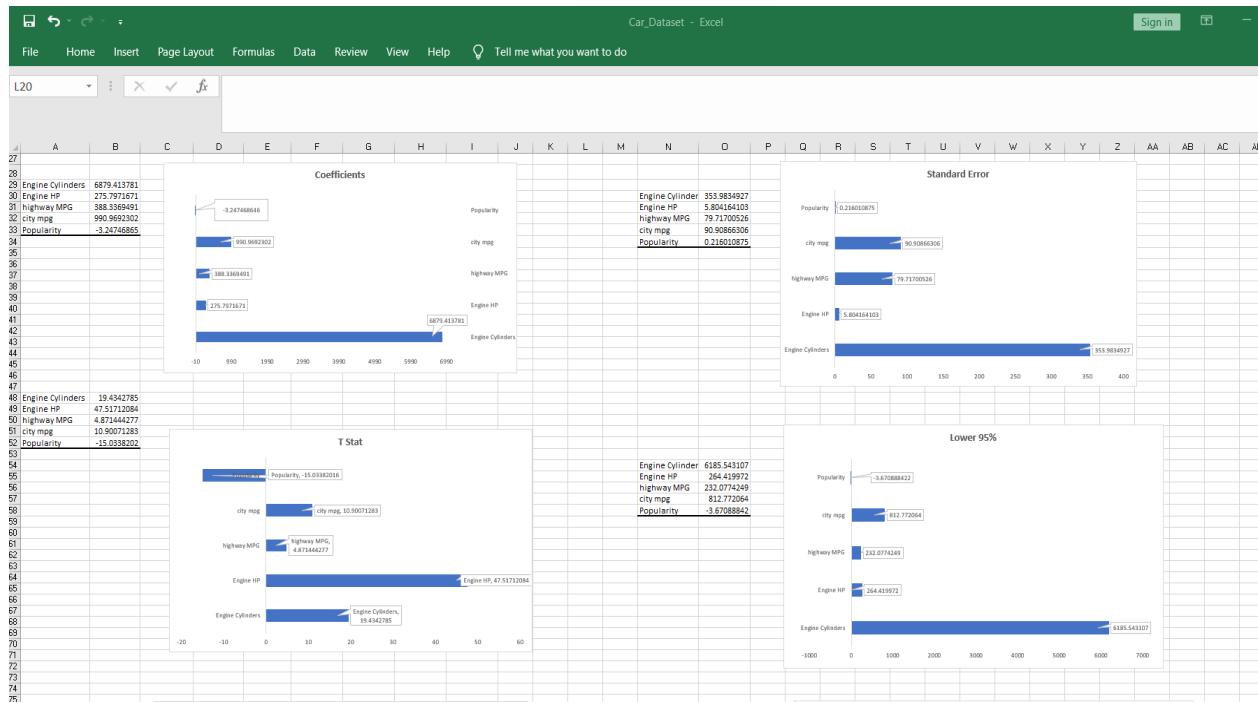
By using regression analysis to identify the variables that have the strongest relationship with a car's price. Then create a bar chart that shows the coefficient values for each variable to visualize their relative importance.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Year	Number of Doors	Engine Cylinders	Engine HP	highway MPG	city mpg	Popularity	MSRP											
2	2011	2	6	335	26	19	3916	46135											
3	2011	2	6	300	28	19	3916	40650											
4	2011	2	6	300	28	20	3916	36350											
5	2011	2	6	230	28	18	3916	29450											
6	2011	2	6	230	28	18	3916	34500											
7	2012	2	6	230	28	18	3916	31200											
8	2012	2	6	300	26	17	3916	44100											
9	2012	2	6	300	28	20	3916	39300											
10	2012	2	6	230	28	18	3916	36900											
11	2013	2	6	230	27	18	3916	37200											
12	2013	2	6	300	28	20	3916	39600											
13	2013	2	6	230	28	19	3916	31500											
14	2013	2	6	300	28	19	3916	44400											
15	2013	2	6	230	28	19	3916	37200											
16	2013	2	6	320	25	18	3916	48250											
17	2013	2	6	320	28	20	3916	43550											
18	1992	4	6	172	24	17	3105	2000											
19	1992	4	6	172	20	16	3105	2000											
20	1992	4	6	172	21	16	3105	2000											
21	1993	4	6	172	24	17	3105	2000											
22	1993	4	6	172	20	16	3105	2000											
23	1993	4	6	172	21	16	3105	2000											
24	1994	4	6	172	21	16	3105	2000											
25	1994	4	6	172	22	16	3105	2000											
26	1994	4	6	172	22	17	3105	2000											
27	1994	4	6	172	22	16	3105	2000											
28	1994	4	6	172	21	16	3105	2000											
29	2017	2	4	160	35	26	819	27495											
30	2017	2	4	160	35	26	819	24995											
31	2017	2	4	160	35	26	819	28195											
32	1991	4	4	120	26	19	647	2000											

Extracting the all the variables with number datatypes

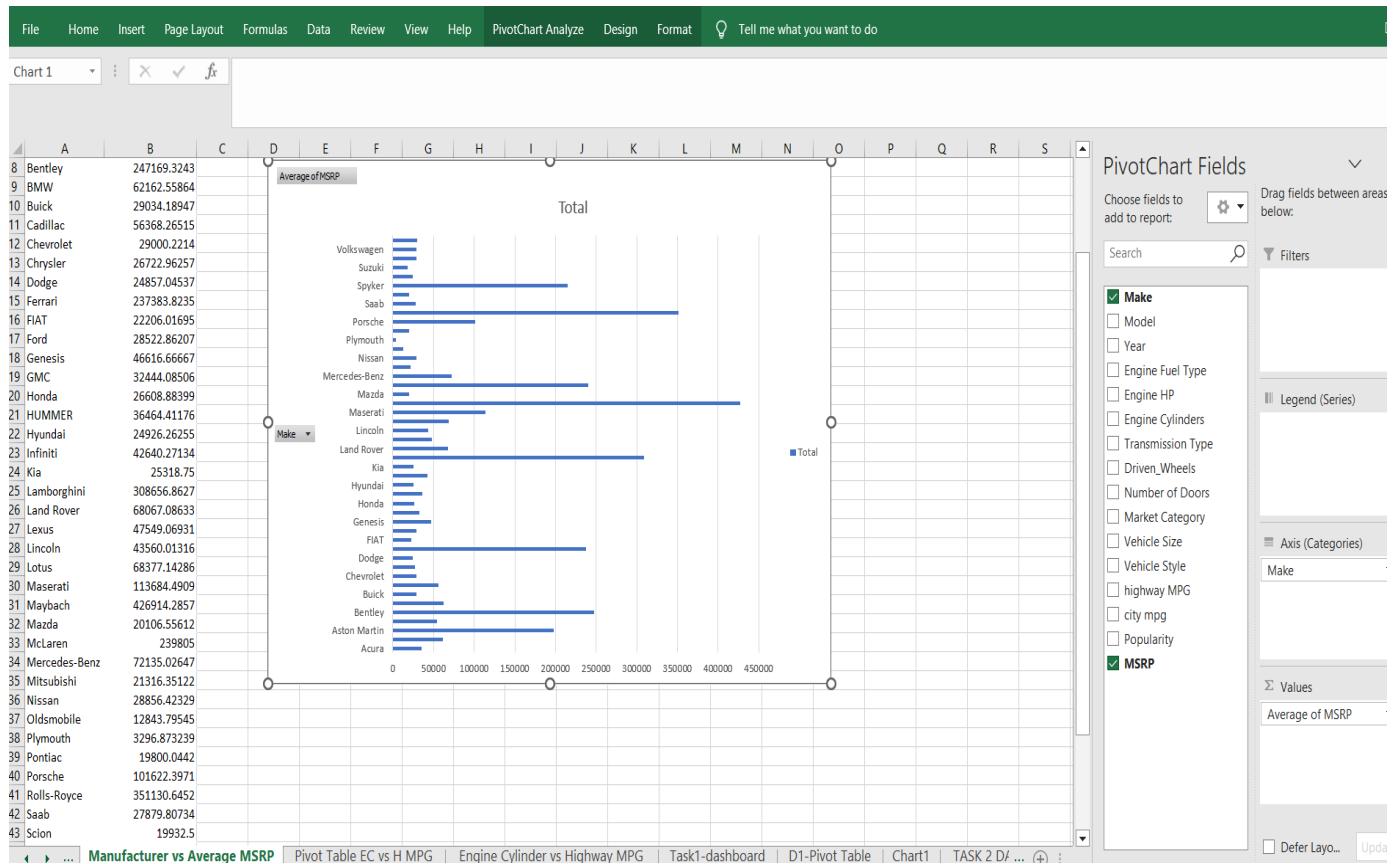
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	
SUMMARY OUTPUT																										
Regression Statistics																										
Multiple R																										
R Square																										
Adjusted R Square																										
Standard Error																										
Observations																										
ANOVA																										
11				df	SS	MS	F	Significance F																		
12	Regression			7	1.26728E+13	2.26754E+12	2148.45133	0																		
13	Residual			11083	1.16971E+13	1055410628																				
14	Total			11090	2.75699E+13																					
16				Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%															
17	Intercept	-768124.2432	117119.3085	-6.558476592	5.678E-11	-997698.9414	-538549.5451																			
18	Year	345.438494	58.8093457	5.874710947	4.3576E-09	230.1781925	460.6987954	230.1781925	460.6987954																	
19	Number of Doors	-53.794116	31.340416	-1.67578779	14.24578779	1.2031E-09	-6.4648E-09	613.138831	4464.146029																	
20	Engine Cylinders	6879.41377	43.3943927	157.712785	9.3628E-09	6185.543107	284.4555	6185.543107	284.4555																	
21	Engine HP	375.7971671	5.804164103	47.5171227	0	354.419972	237.1743622	264.419972	287.1743622																	
22	highway MPG	388.3369491	79.71700526	4.87144277	1.1232E-06	232.0774249	544.5964733	232.0774249	544.5964733																	
23	city mpg	990.9693236	90.90866306	10.90071283	1.5773E-27	812.7722064	1169.166396	812.7722064	1169.166396																	
24	Popularity	-3.247468646	0.216010875	-15.0382016	1.3888E-50	-3.67088442	-2.82404887	-3.67088442	-2.82404887																	

Regression Analysis



#### 4. the average price of a car vary across different manufacturers

By creating a pivot table that shows the average price of cars for each manufacturer and create a bar chart or a horizontal stacked bar chart that visualizes the relationship between manufacturer and average price



#### 5. The relationship between fuel efficiency and the number of cylinders in a car's engine

By creating 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. And calculate the correlation coefficient between the number of cylinders and highway MPG to quantify the strength and direction of the relationship.

Screenshot of Microsoft Excel showing a PivotTable Fields pane on the right. The PivotTable Fields pane lists various vehicle attributes: Make, Model, Year, Engine Fuel Type, Engine HP, Engine Cylinders (selected), Transmission Type, Driven\_Wheels, Number of Doors, Market Category, Vehicle Size, Vehicle Style, highway MPG (selected), city mpg, Popularity, and MSRP. The main table shows the sum of highway MPG for different engine cylinder counts.

Engine Cylinders	Sum of highway MPG
0	1339
3	1160
4	137104
5	4405
6	102530
8	39648
10	1300
12	3982
<b>Grand Total</b>	<b>291468</b>

Pivot table of Engine Cylinders ang Highway MPG

Screenshot of Microsoft Excel showing a scatter plot titled "Highway MPG" with a linear regression line. The x-axis represents Engine Cylinders (0 to 16) and the y-axis represents Highway MPG (0 to 160,000). The data points show a weak negative correlation. The regression equation is  $y = -1714.5x + 46721$  with  $R^2 = 0.0153$ .

Engine Cylinder	Highway MPG
0	1339
3	1160
4	137104
5	4405
6	102530
8	39648
10	1300
12	3982

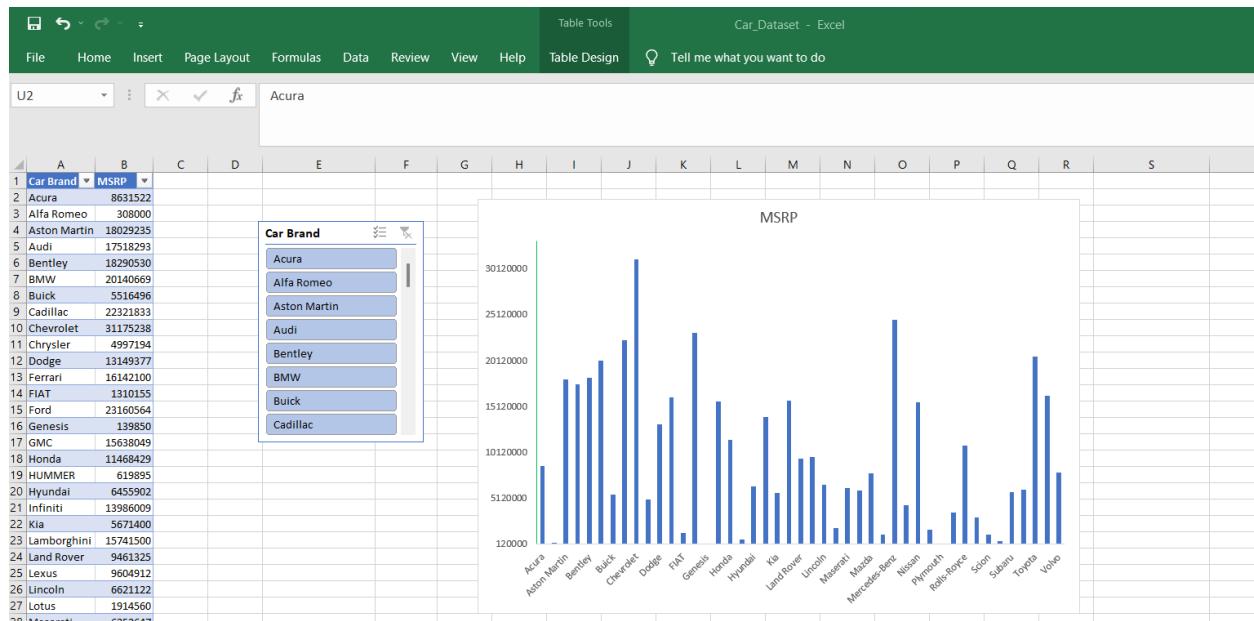
Correlation Coefficient

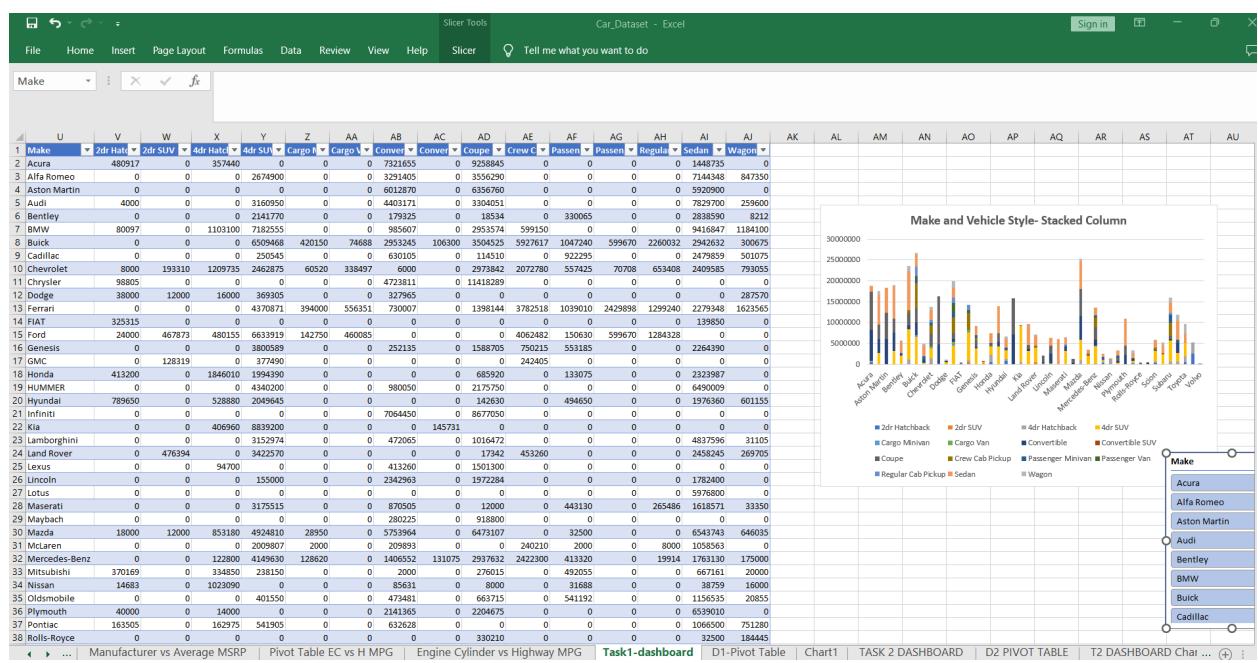
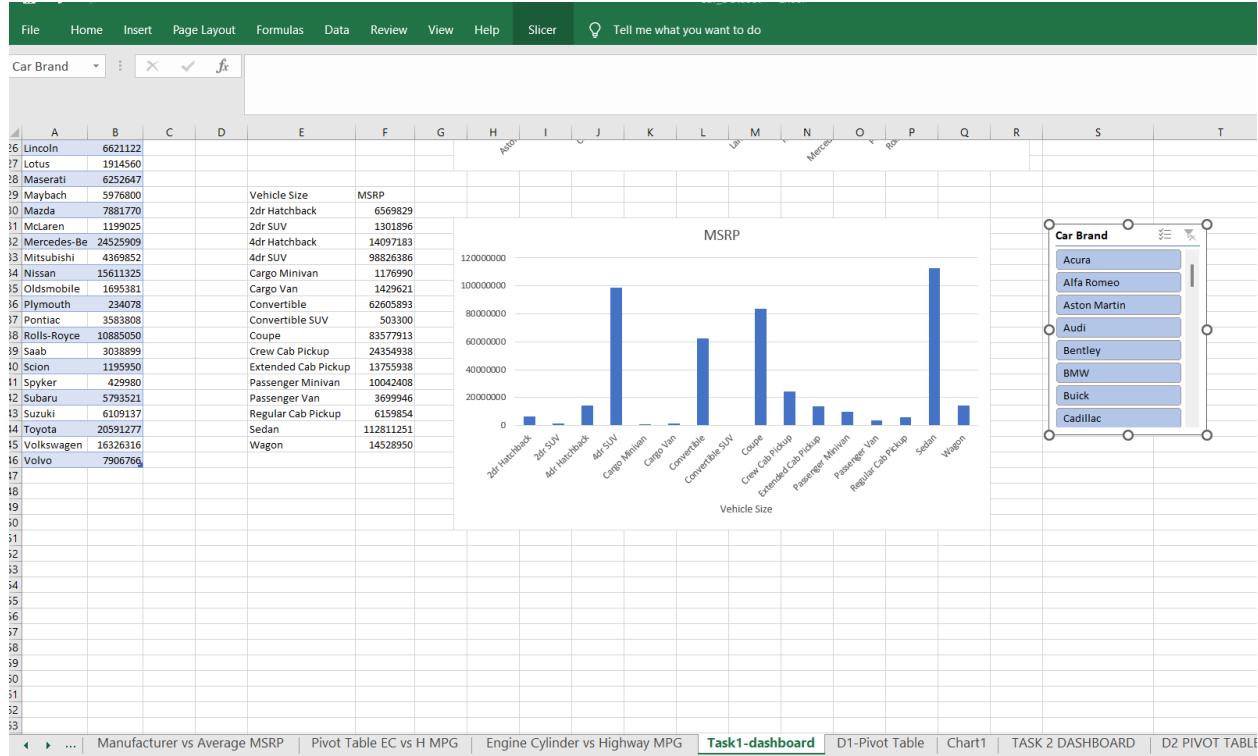
# Creating Dashboard

Creating an Interactive Dashboard Using Slicers, pivot table and filters

1. The distribution of car prices vary by brand and body style?

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 SUMIF or Pivot Tables.



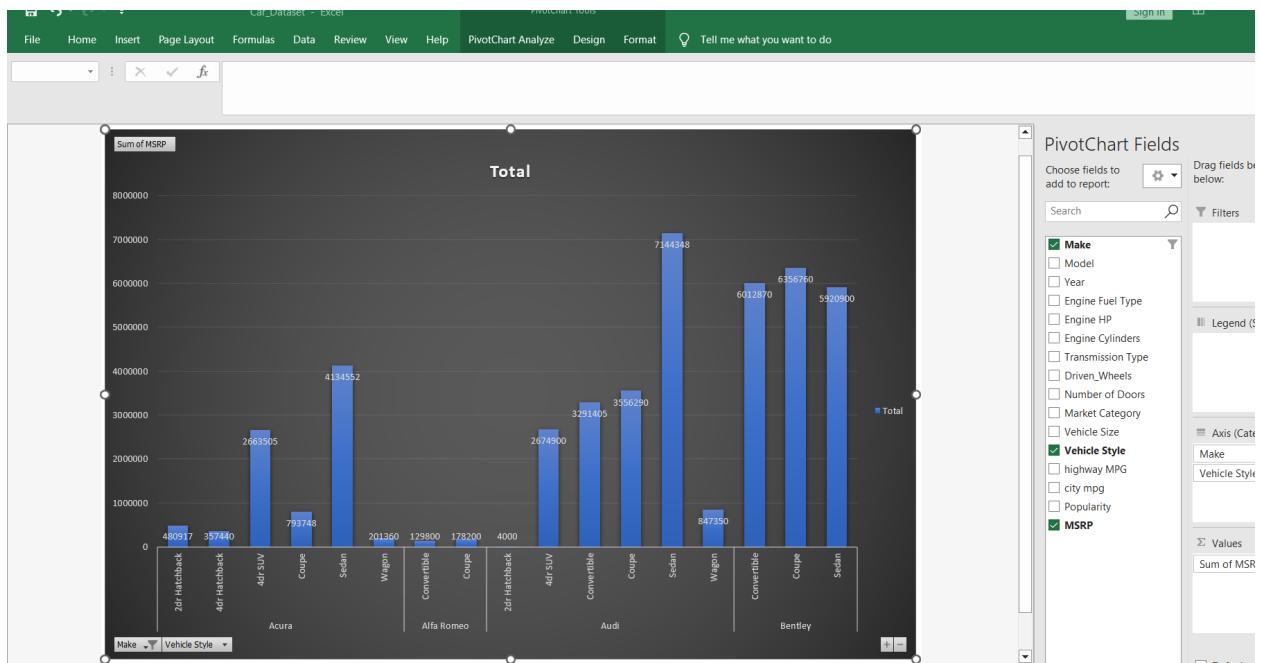


File Home Insert Page Layout Formulas Data Review View Help PivotTable Analyze Design Tell me what you want to do

A3 Row Labels Sum of MSRP

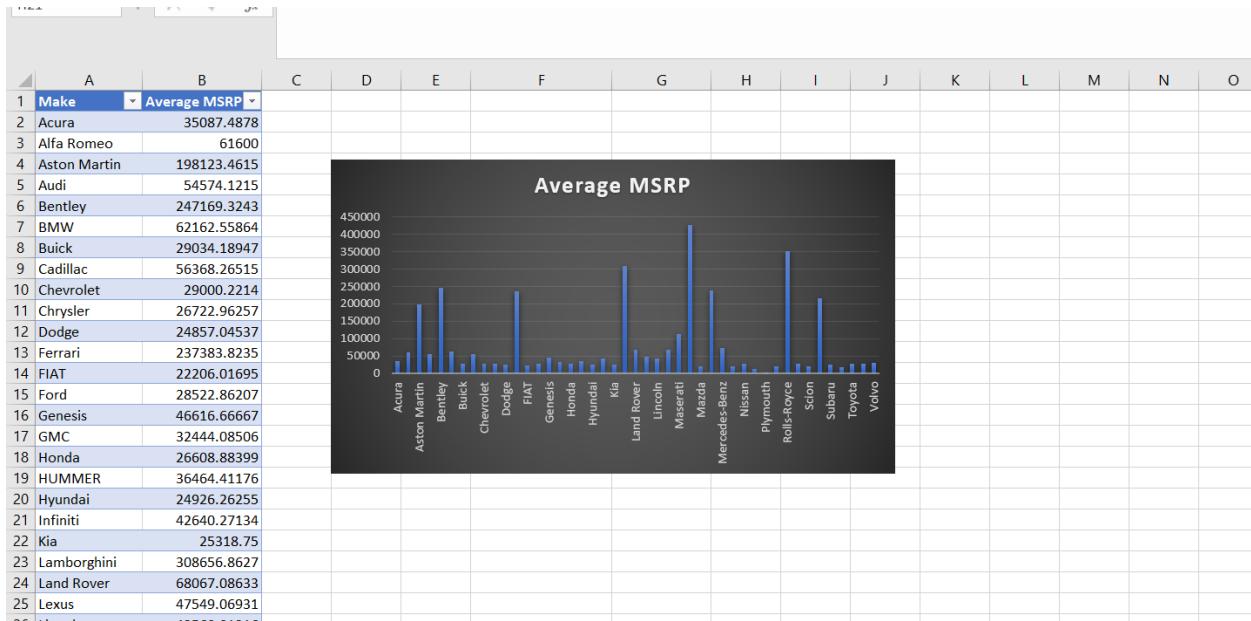
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1																	
2																	
3	<b>Row Labels</b>		<b>Sum of MSRP</b>														
4	<b>Acura</b>		<b>8631522</b>														
5	2dr Hatchback		480917														
6	4dr Hatchback		357440														
7	4dr SUV		2663505														
8	Coupe		793748														
9	Sedan		4134552														
10	Wagon		201360														
11	<b>Alfa Romeo</b>		<b>308000</b>														
12	Convertible		129800														
13	Coupe		178200														
14	<b>Audi</b>		<b>17518293</b>														
15	2dr Hatchback		4000														
16	4dr SUV		2674900														
17	Convertible		3291405														
18	Coupe		3556290														
19	Sedan		7144348														
20	Wagon		847350														
21	<b>Bentley</b>		<b>18290530</b>														
22	Convertible		6012870														
23	Coupe		6356760														
24	Sedan		5920900														
25	<b>Grand Total</b>		<b>44748345</b>														
26																	
27																	
28																	
29																	
30																	
31																	
32																	

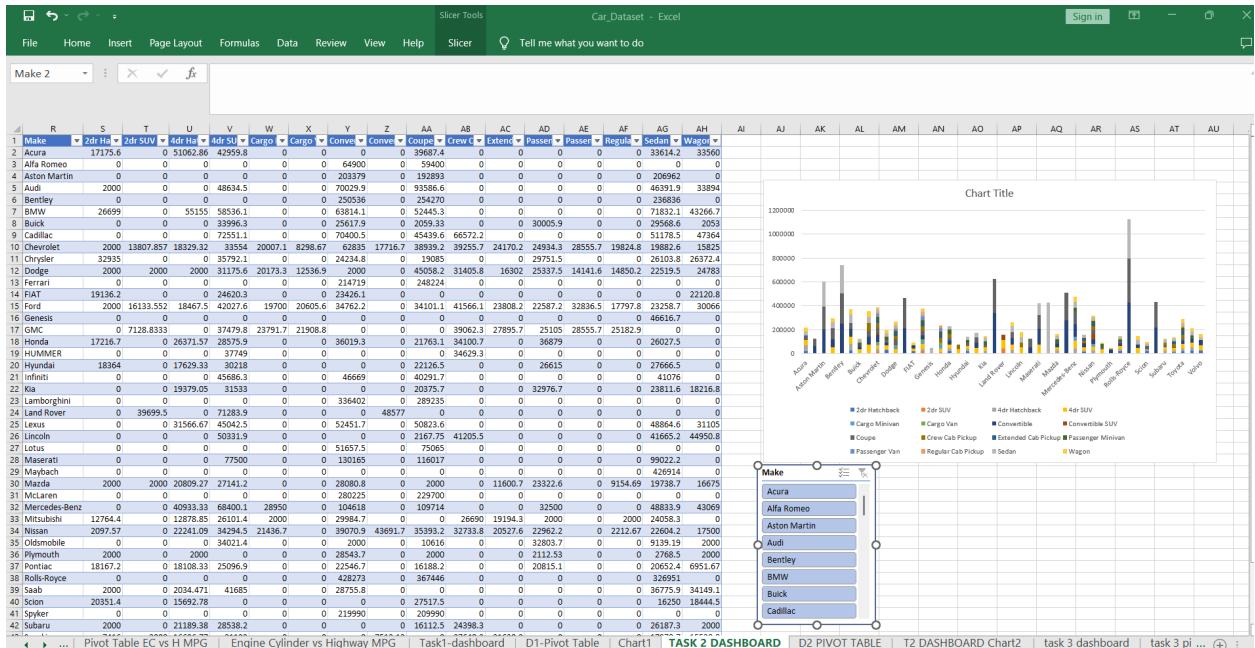
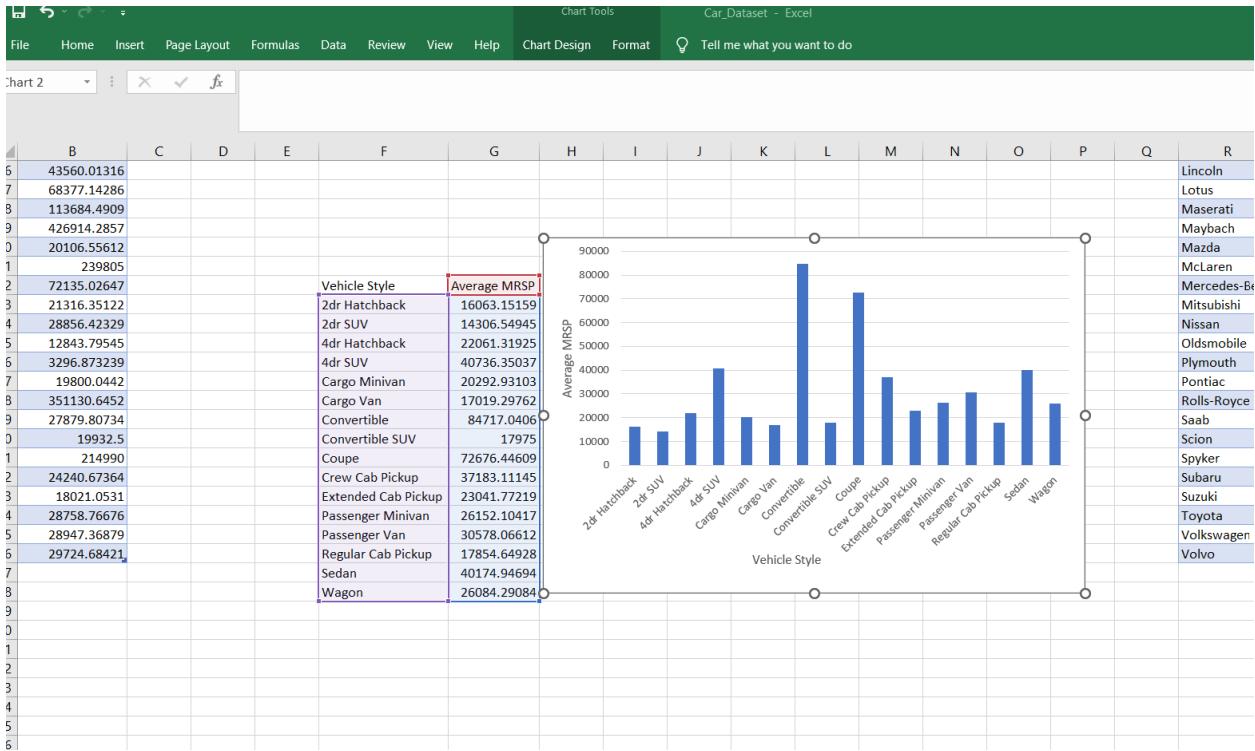
Manufacturer vs Average MSRP | Pivot Table EC vs H MPG | Engine Cylinder vs Highway MPG | Task1-dashboard | D1-Pivot Table | Chart1 | TASK 2 D/ ... | + | Def



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

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 AVERAGEIF or Pivot Tables.





Car\_Dataset - Excel

PivotTable Tools

Average of MSRP

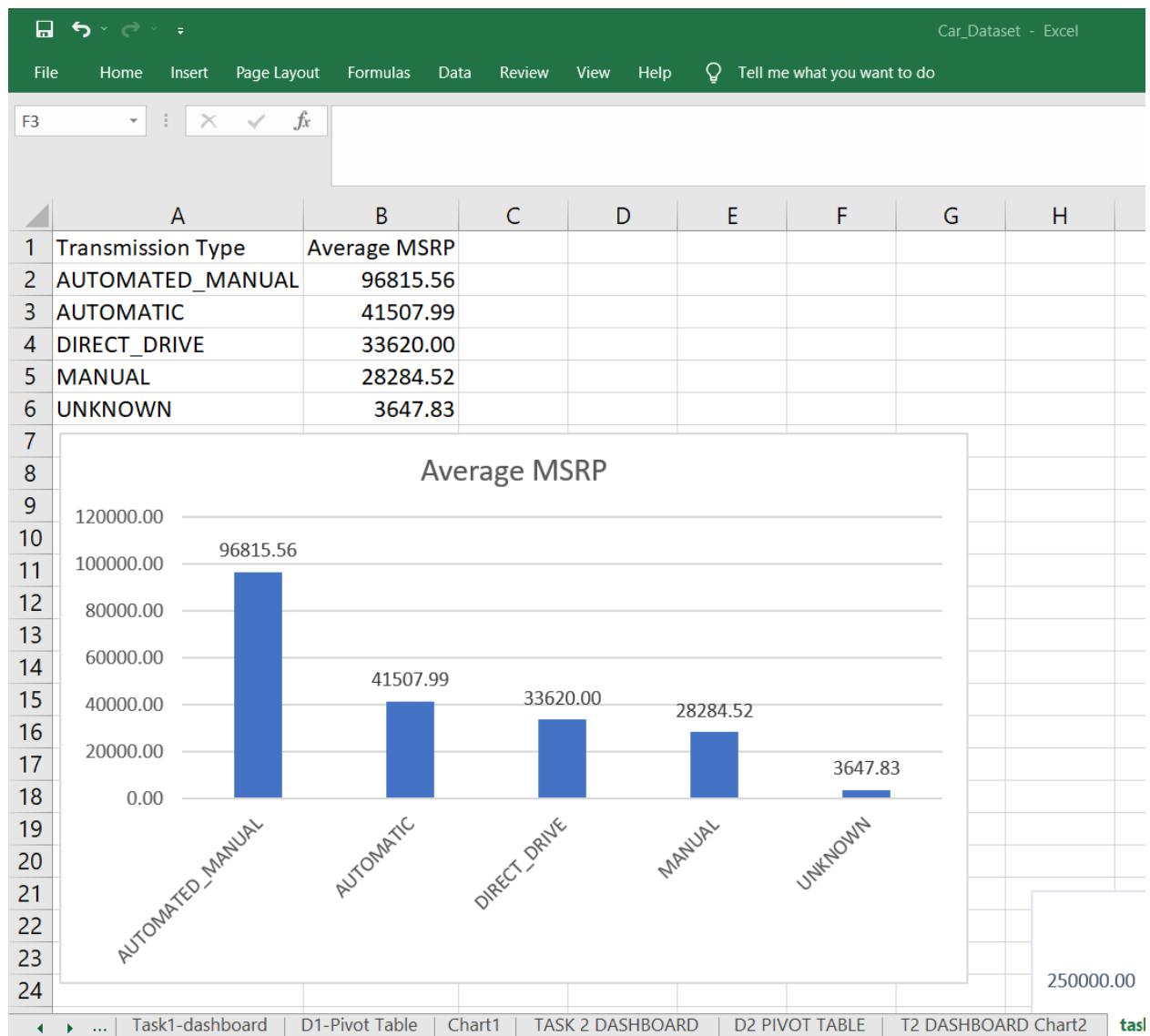
2dr Hatchback

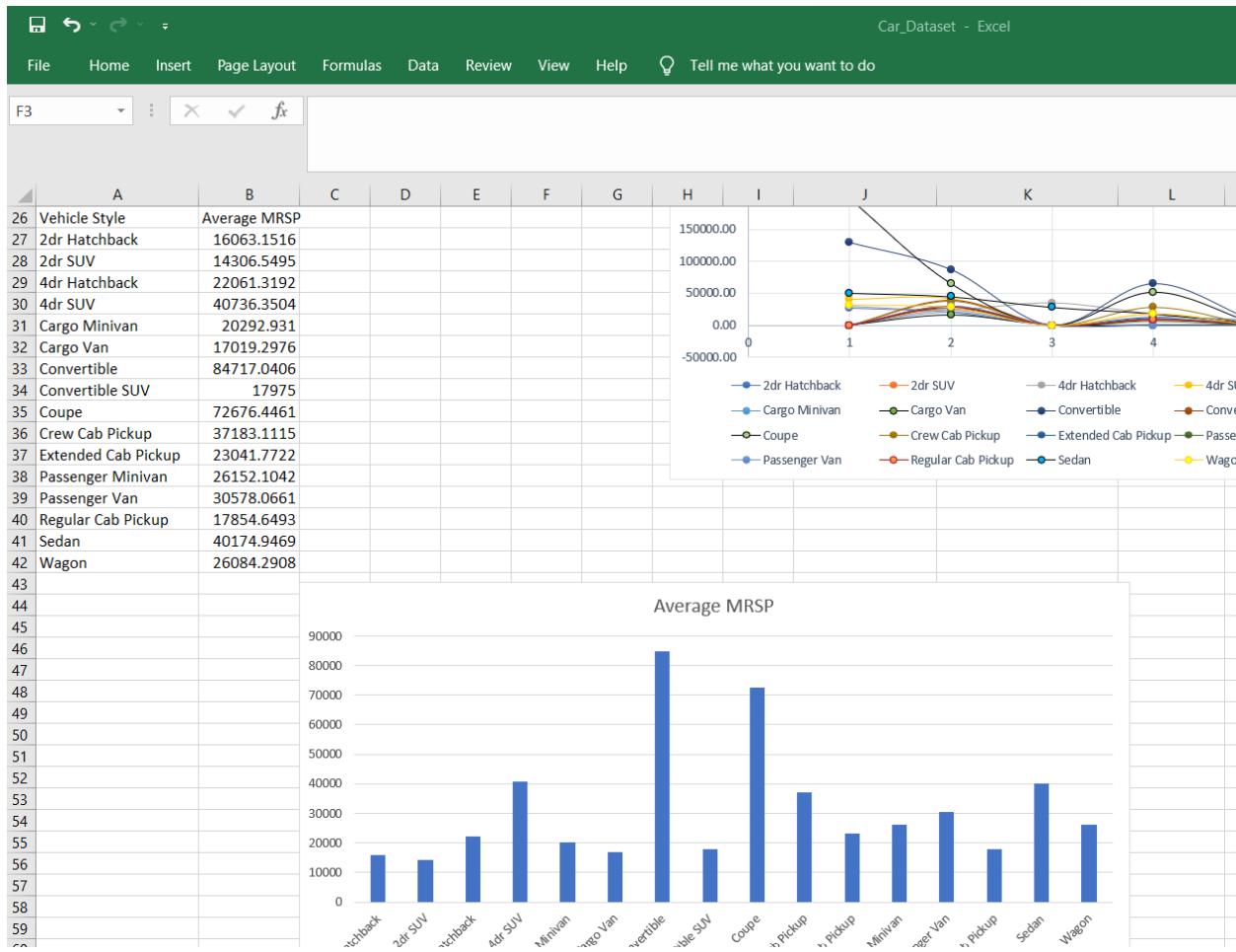
			C	D	E	F	G	H	I	J	K	L	M	N	O
1	A	B													
3	Row Labels	Average of MSRP													
4	2dr Hatchback	16163.90													
5	Acura	17175.61													
6	Audi	2000.00													
7	4dr Hatchback	51062.86													
8	Acura	51062.86													
9	4dr SUV	45627.39													
10	Acura	42959.76													
11	Audi	48634.55													
12	Convertible	153722.29													
13	Alfa Romeo	64900.00													
14	Aston Martin	203379.31													
15	Audi	70029.89													
16	Bentley	250536.25													
17	Coupe	150327.19													
18	Acura	39687.40													
19	Alfa Romeo	59400.00													
20	Aston Martin	192892.60													
21	Audi	9386.58													
22	Bentley	254270.40													
23	Sedan	60351.25													
24	Acura	33614.24													
25	Aston Martin	206962.14													
26	Audi	46391.87													
27	Bentley	236836.00													
28	Wagon	33829.35													
29	Acura	33560.00													
30	Audi	33894.00													
31	Grand Total	85179.89													

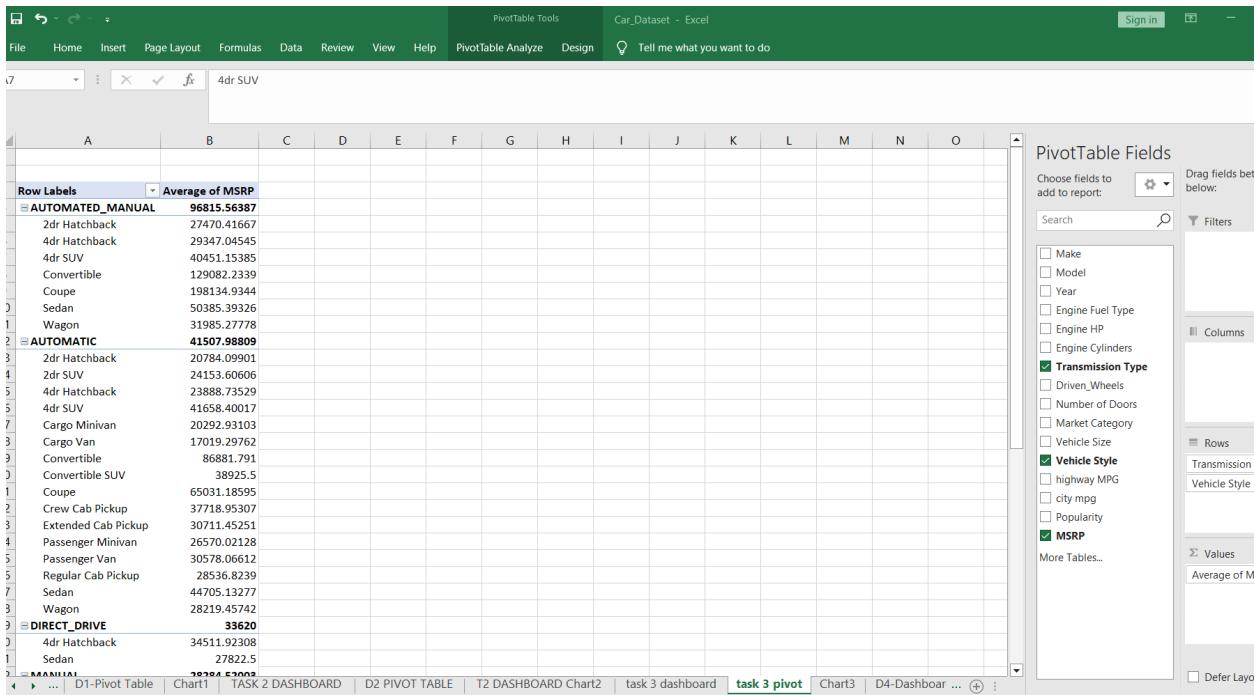
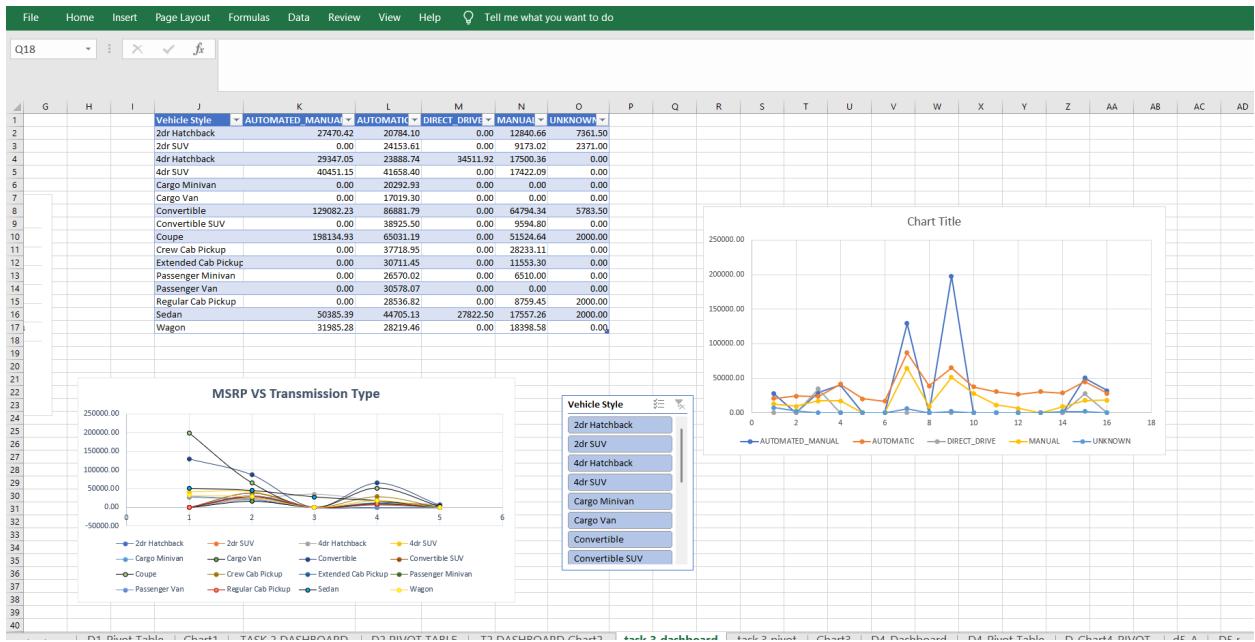


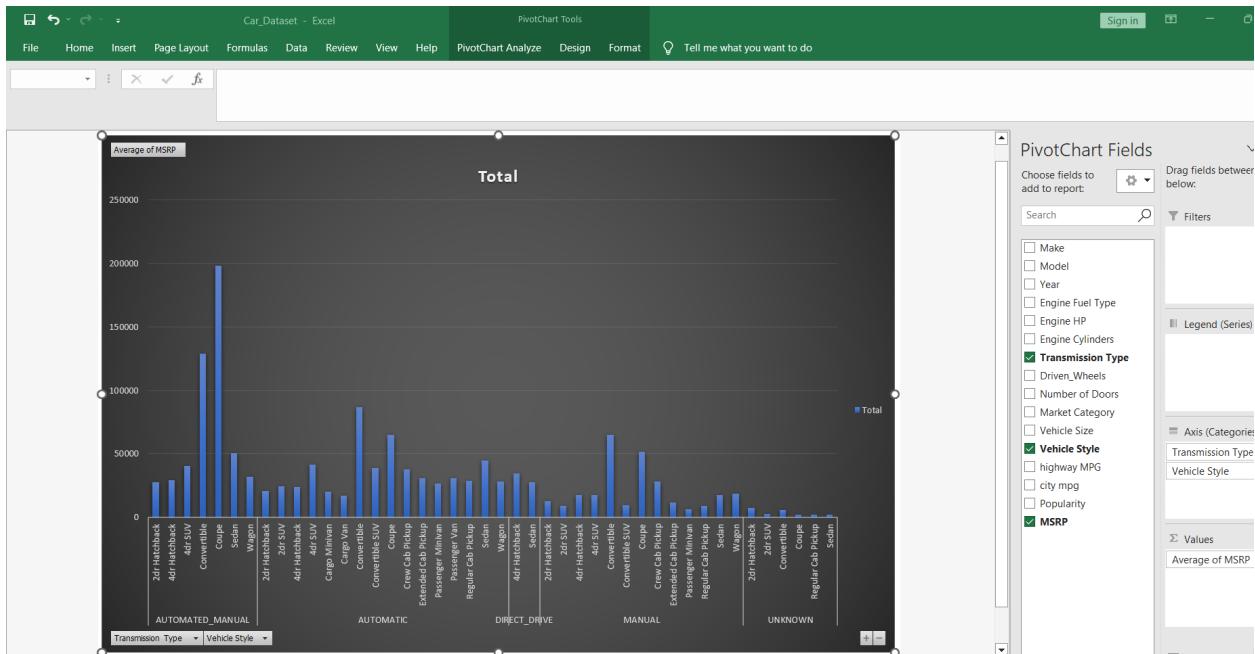
3. The different feature such as transmission type affect the MSRP, and how does this vary by body style

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 AVERAGEIFS or Pivot Tables



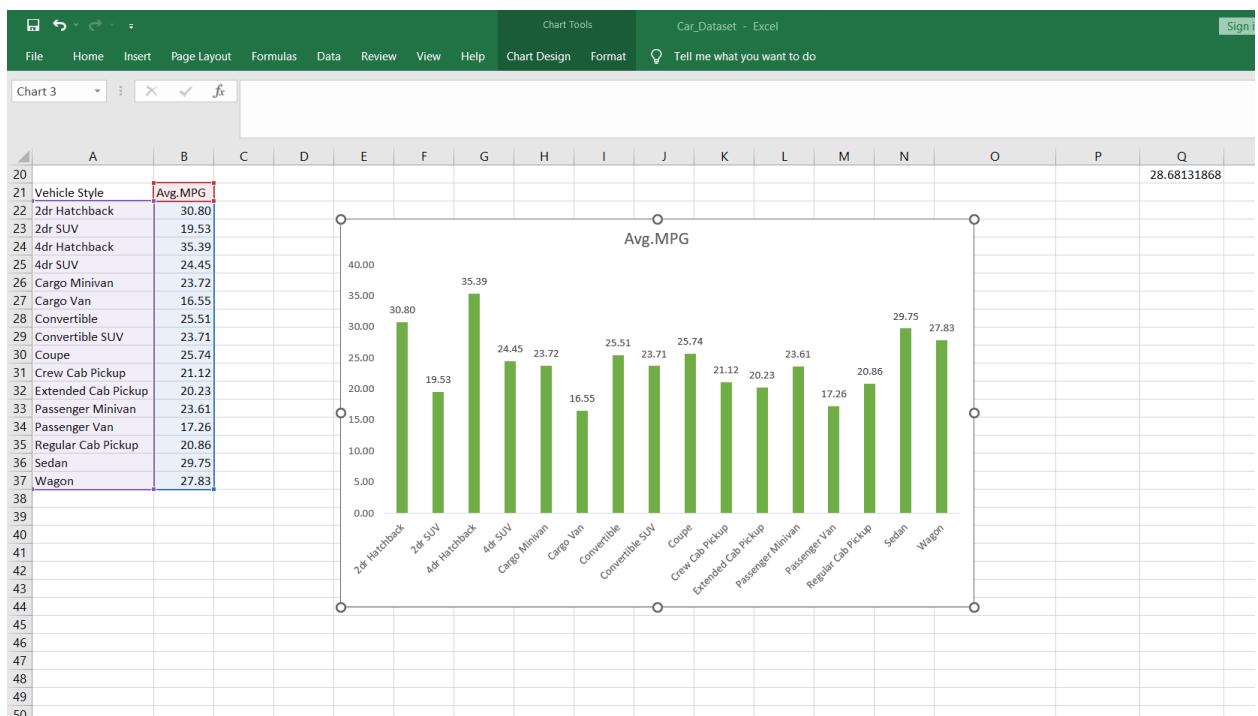
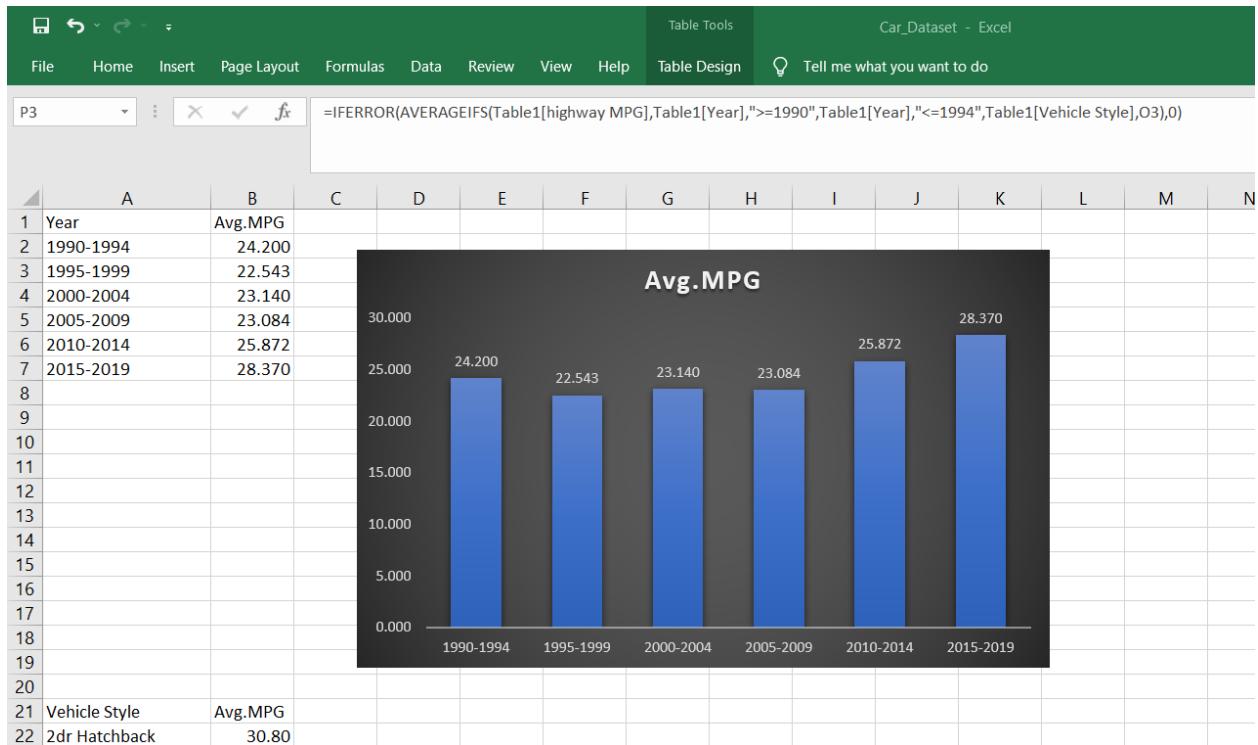


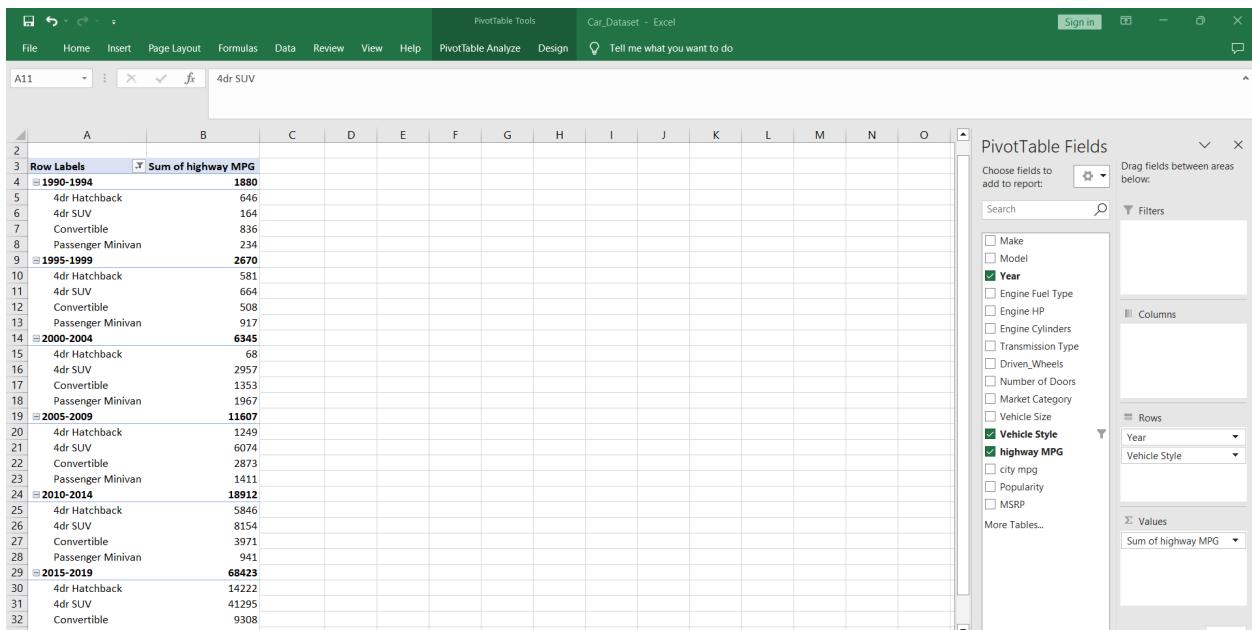
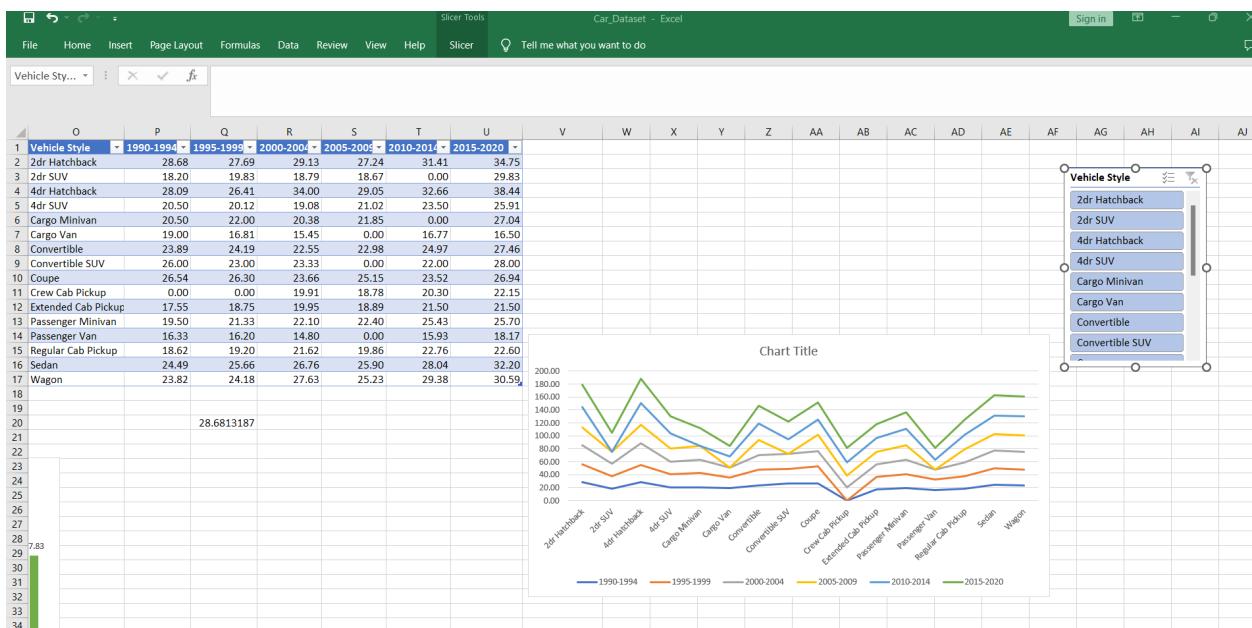


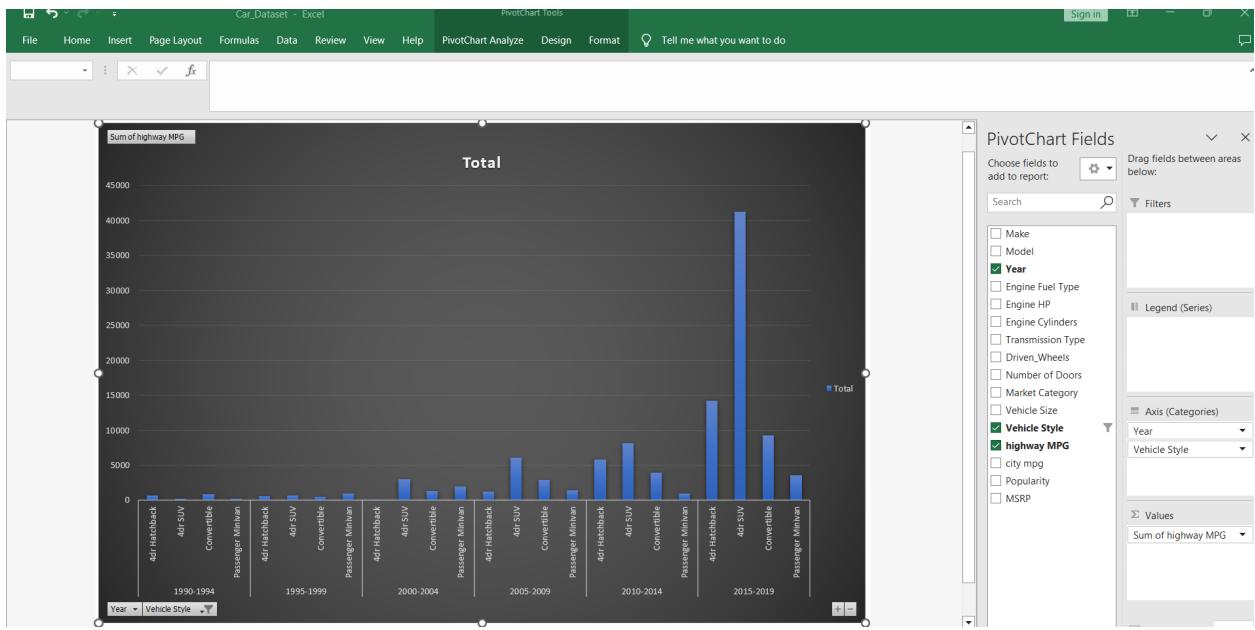


#### 4. The fuel efficiency of cars vary across different body styles and model years

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.

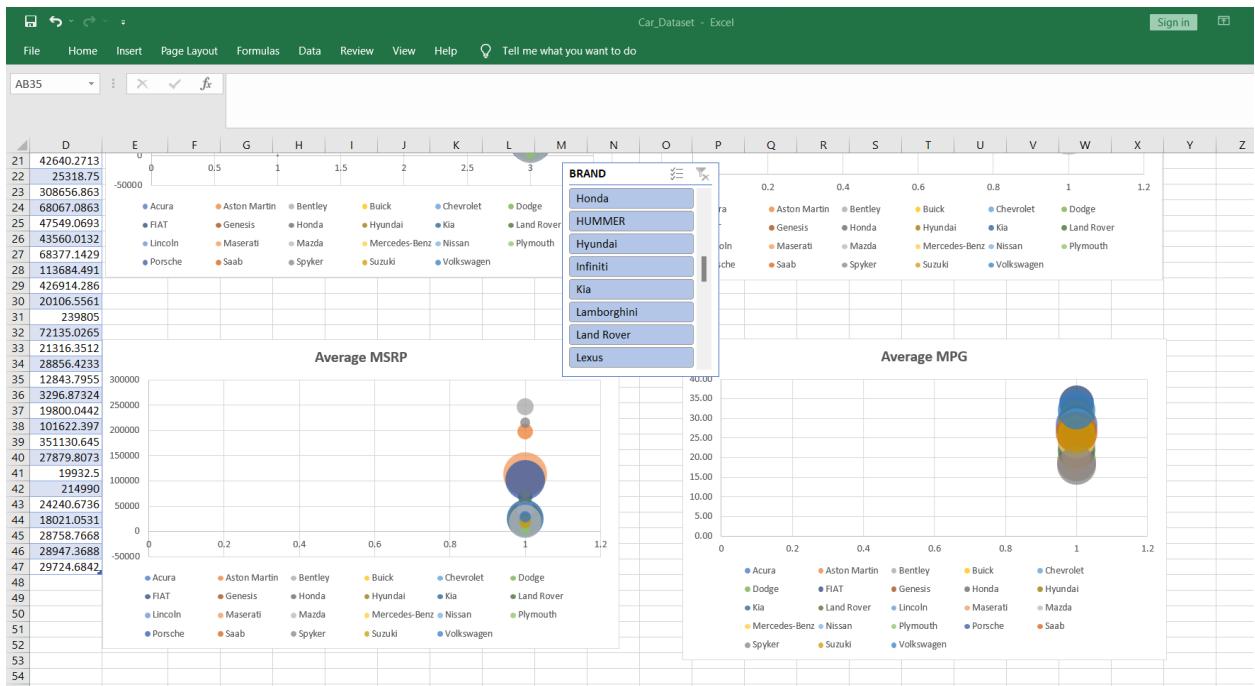
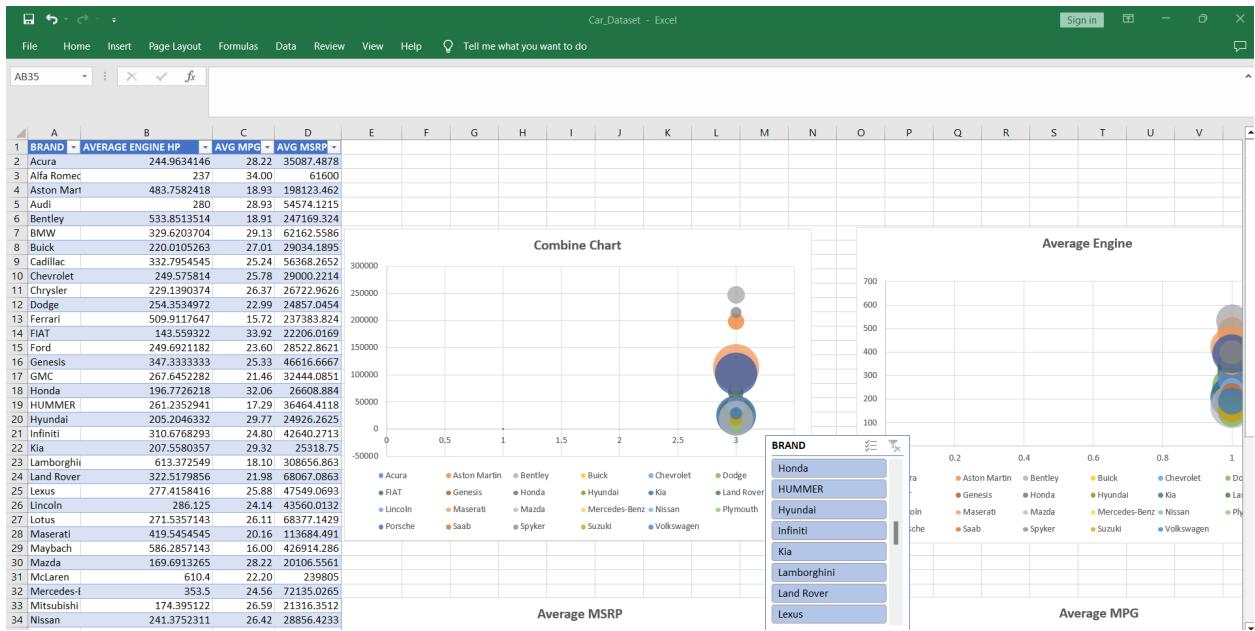






The car's horsepower, MPG, and price vary across different Brands

By using 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.

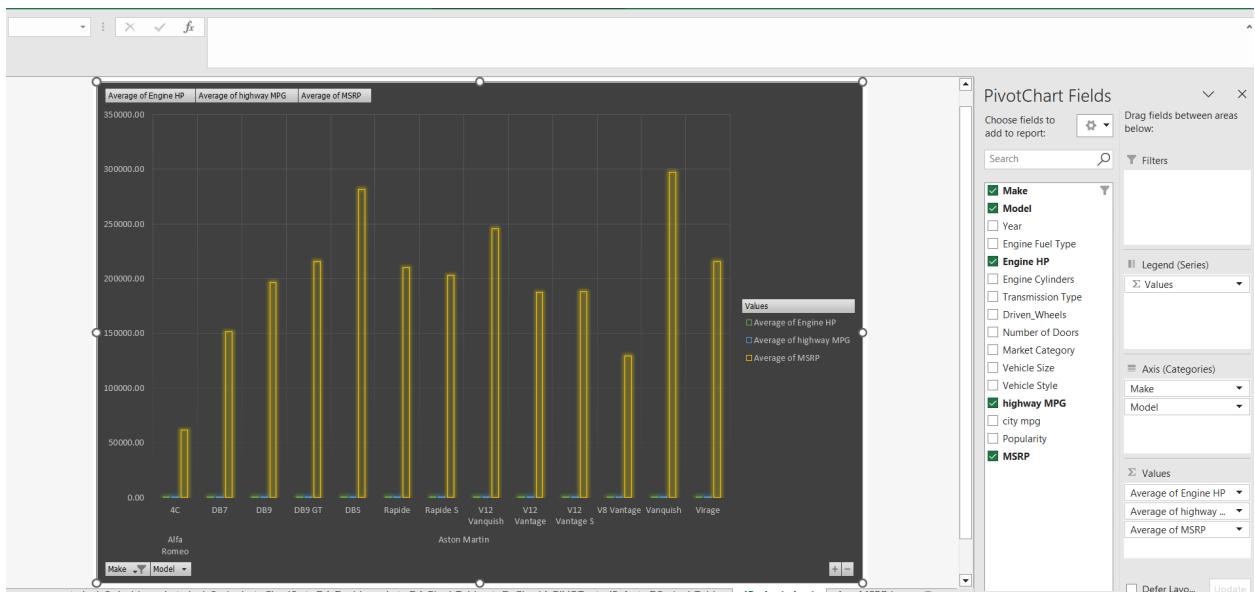


Excel screenshot showing a PivotTable report titled "Car\_Dataset - Excel". The PivotTable Fields pane on the right shows fields like Make, Model, Engine HP, highway MPG, and MSRP.

**PivotTable Fields**

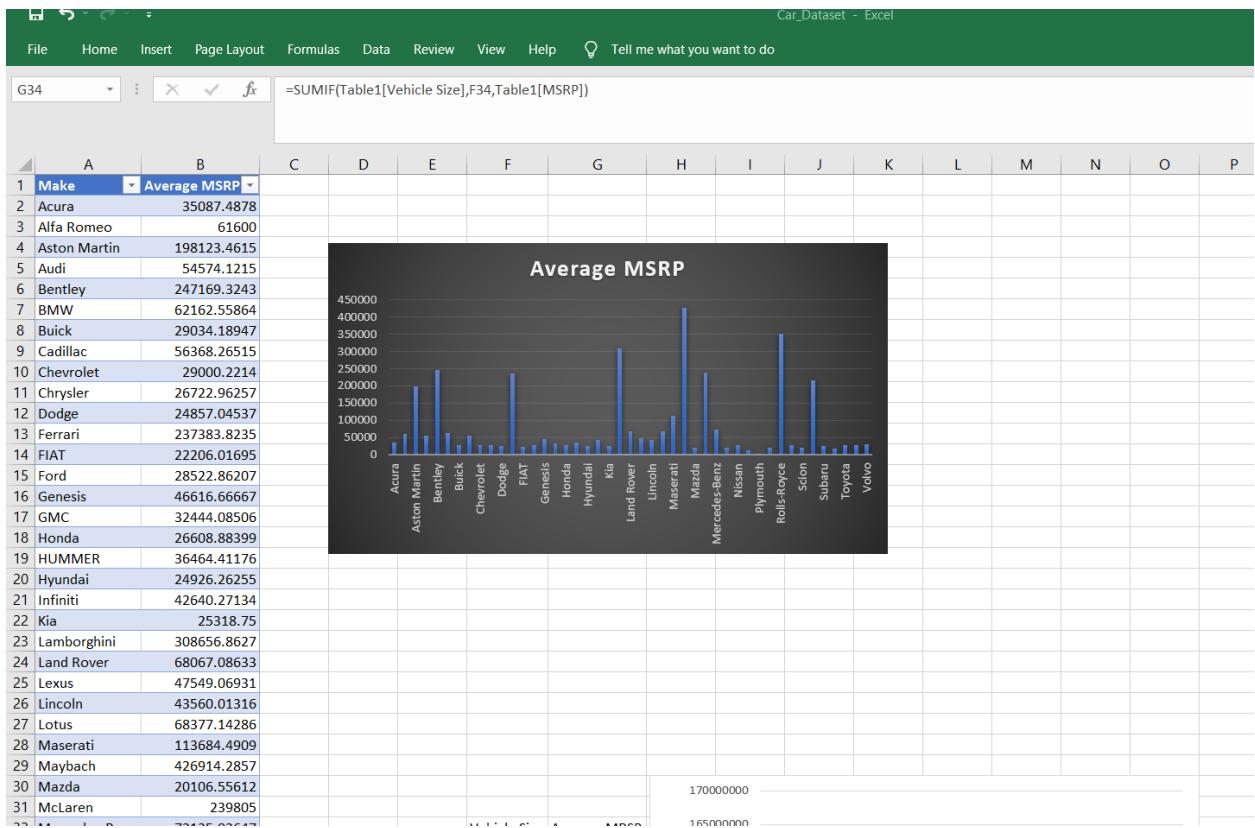
- Choose fields to add to report:
- Filters:
  - Columns: Values
  - Rows: Make, Model
  - Values: Average of Engine HP, Average of highway MPG, Average of MSRP

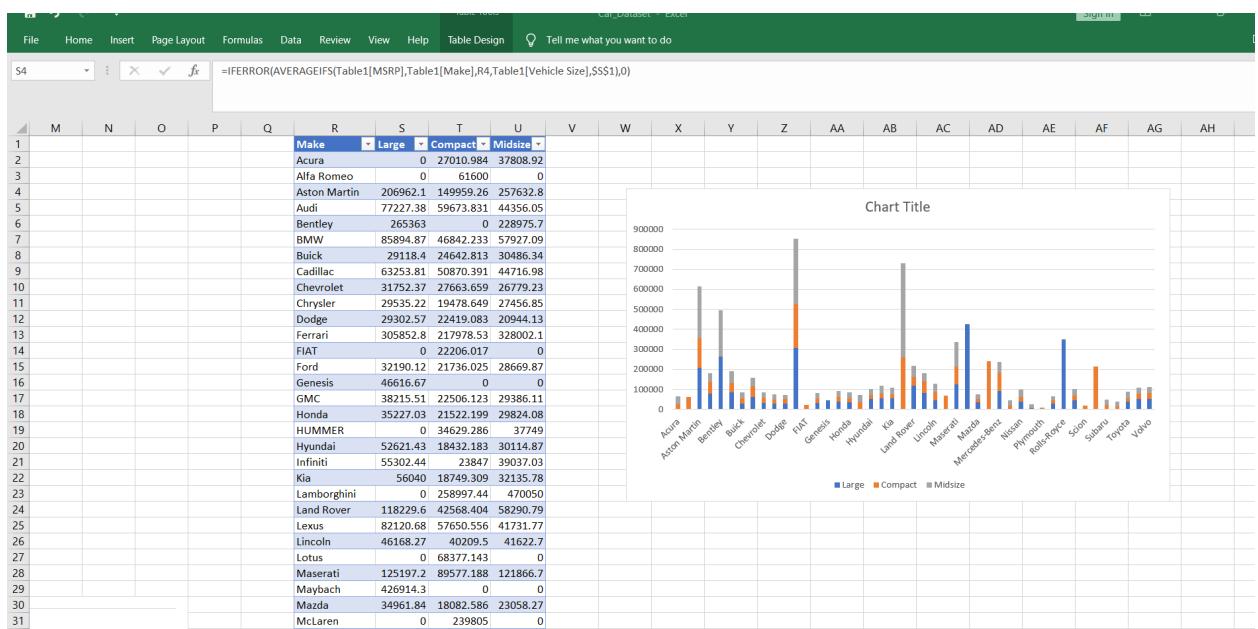
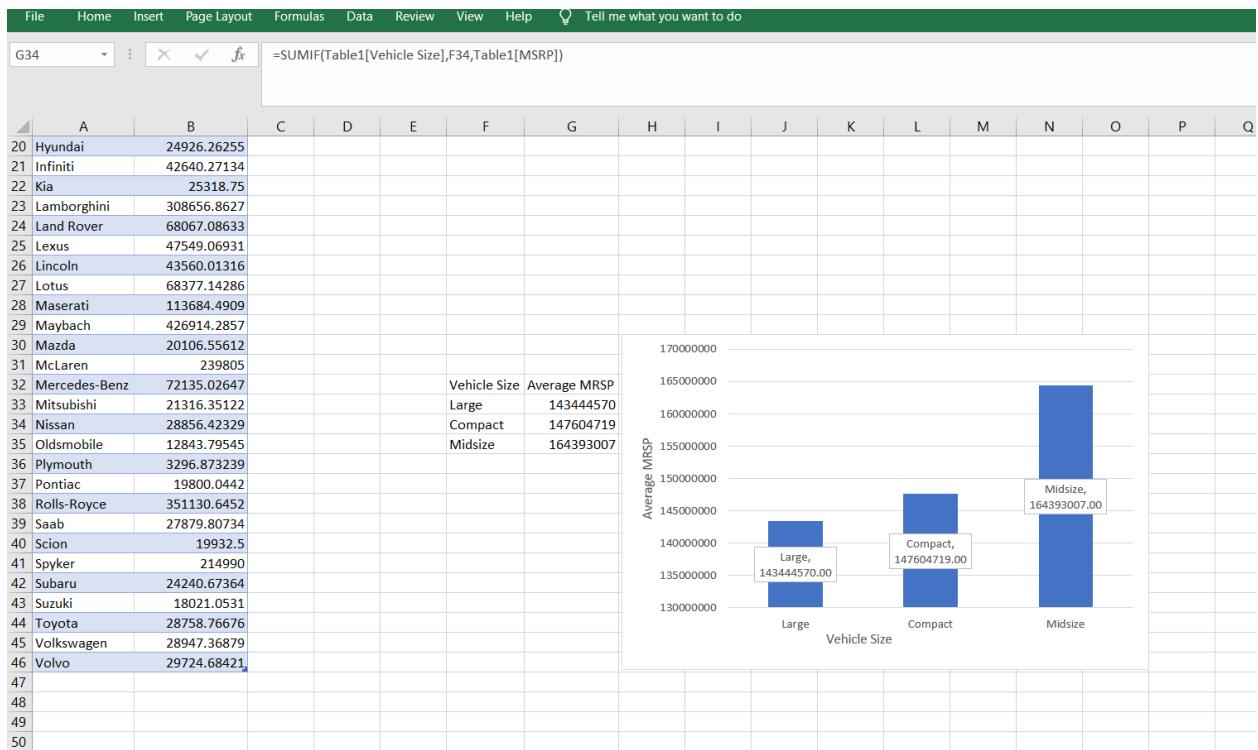
Row Labels	Average of Engine HP	Average of highway MPG	Average of MSRP
Alfa Romeo	237.00	34.00	61600.00
4C	237.00	34.00	61600.00
Aston Martin	483.76	18.93	198123.46
DB7	423.75	16.75	151550.00
DB9	510.00	19.00	196522.50
DB9 GT	540.00	19.00	215602.33
DBS	510.00	17.50	281573.00
Rapide	470.00	19.00	210122.50
Rapide S	551.33	21.00	202748.33
V12 Vanquish	490.00	16.50	245880.00
V12 Vantage	510.00	17.00	187240.00
V12 Vantage S	565.00	18.00	188095.00
V8 Vantage	428.75	20.00	129535.00
Vanquish	567.25	20.50	297083.75
Virage	490.00	18.00	215795.00
<b>Grand Total</b>	<b>470.91</b>	<b>19.72</b>	<b>191012.86</b>



6 car brands have the highest and lowest average MSRPs, and how does this vary by Vehicle Size

By using clustered column chart to compare the average MSRPs across different car brands and Vehicle Size. Calculate the average MSRP for each brand and Vehicle size using AVERAGEIF or Pivot Tables.

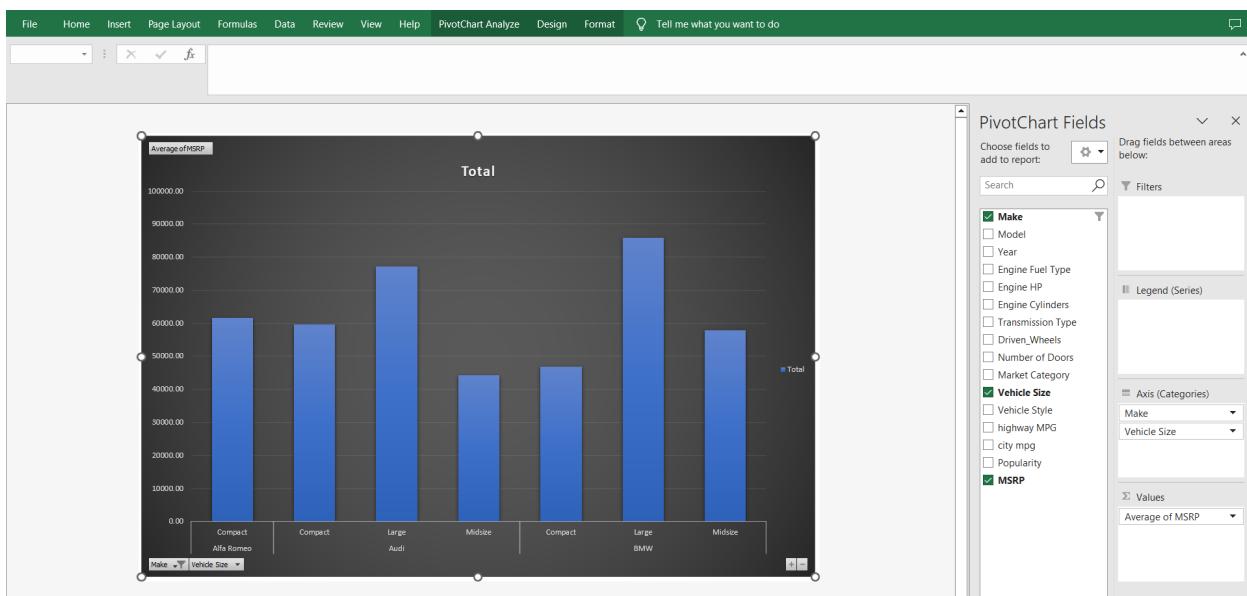




File Home Insert Page Layout Formulas Data Review View Help PivotTable Analyze Design Tell me what you want to do

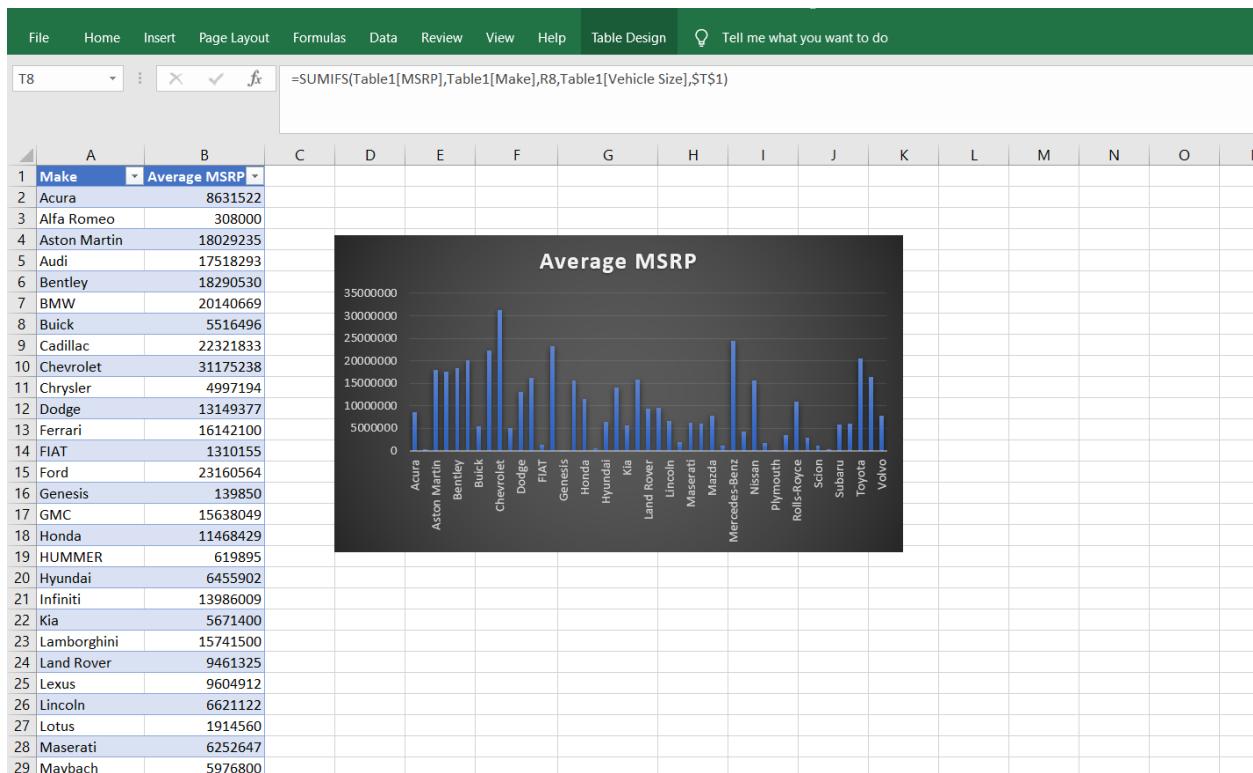
B9 44356.0516129032

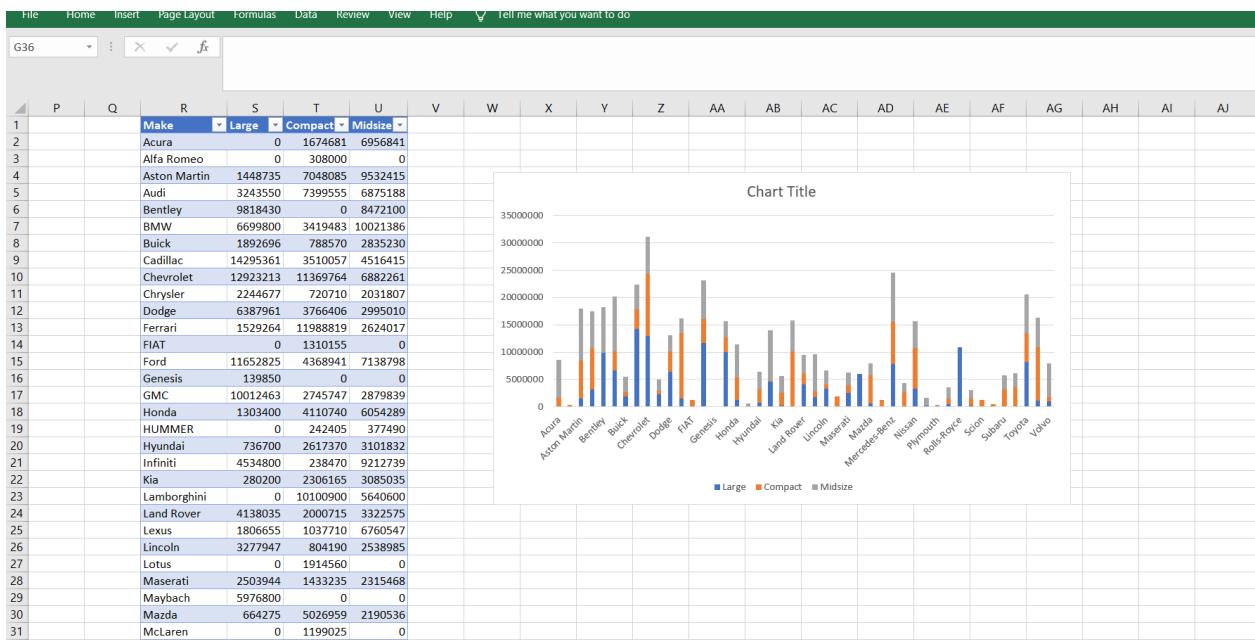
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1																
2																
3	Row Labels	Average of MSRP														
4	Alfa Romeo	61600.00														
5	Compact	61600.00														
6	Audi	54574.12														
7	Compact	59673.83														
8	Large	77227.38														
9	Midsize	44356.05														
10	BMW	62162.56														
11	Compact	46842.23														
12	Large	85894.87														
13	Midsize	57927.09														
14	Grand Total	58410.71														
15																
16																
17																
18																
19																
20																
21																
22																
23																
24																
25																
26																
27																
28																

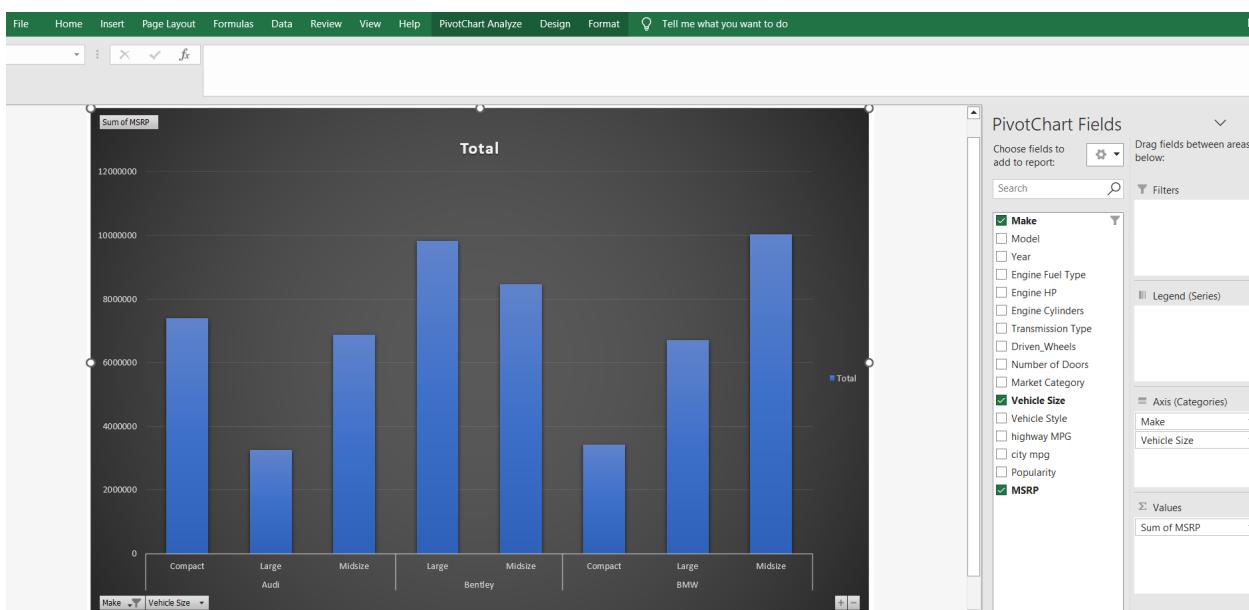


7. car brands have the highest and lowest MSRPs, and how does this vary by Vehicle Size

By clustered column chart to compare the MSRPs across different car brands and Vehicle Size. Calculate the sum of MSRP for each brand and Vehicle size using SUMIF or Pivot Tables.







In terms of market category, we should be focused on Crossover as they have values in terms of model count and popularity

by observing linear forecast we can see that as we increase the no of cylinder our price will decline with it so the manufacturers should focus on models with no of cylinders in range of 225-304

the manufacturer as for to increase its sales should focus on brands such as Maybach, Rolls-Royce, Lamborghini, Spyker car models to maximize its profitability

the manufacturers should focus of automobiles that have 4 cylinders as we cannot correlate between manufactures and numbers of cylinder due to low correlation coefficient

Based on msrp of the make of the Vehicle the brands the highest is Chevrolet, followed by Mercedes-benz, Ford, Cadillac, Toyota

Sedan, 4dr Suv, coupe are more preferred in terms of Vehicle style

Automated Manual are the most popular in terms of Transmission Type

Based on the Vehicle style over the years 4dr Hatchback is overall poplar throughout the time frame followed by sedan, coupe, convertible, passenger minivan with the least and drop through the time period is crew cab Pickup so the manufacturer should avoid crew cab Pickup

In terms of vehicle style and make the most bought among them are wagon and sedan of Buick

In terms of transmission type and vehicle style the most well liked in automated manuals are coupe and convertible

Based on vehicle size the most brought are the midsize vehicle

The most brought among the midsize vehicles are Ferrari, Chevrolet, Lamborghini, Aston Martin