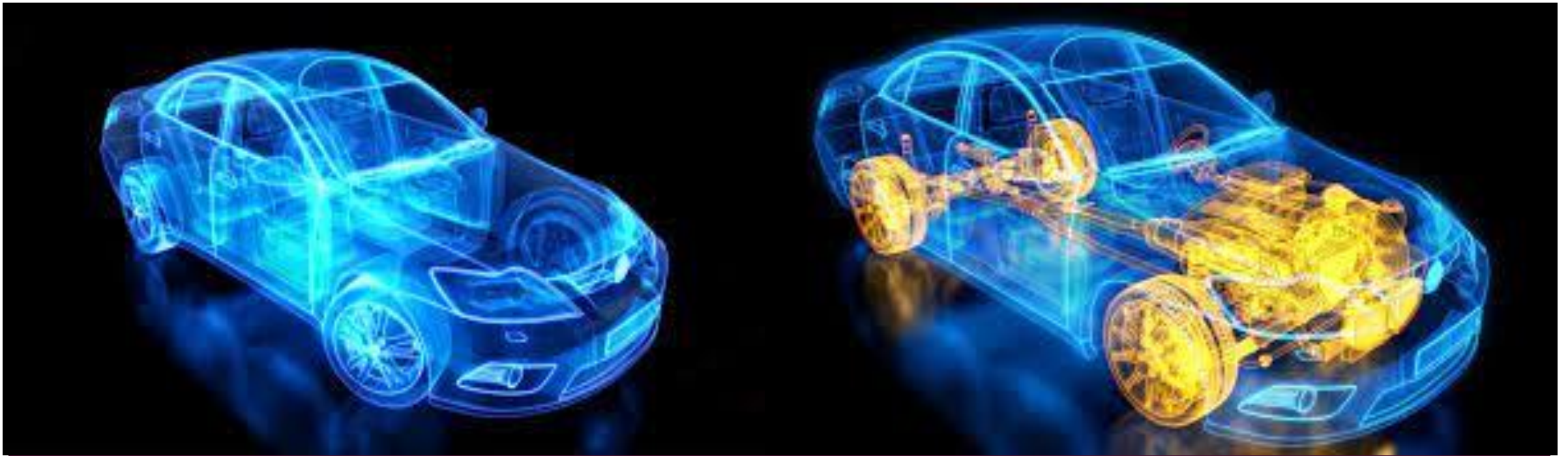

Analyzing the Impact of Car Features on Price and Profitability



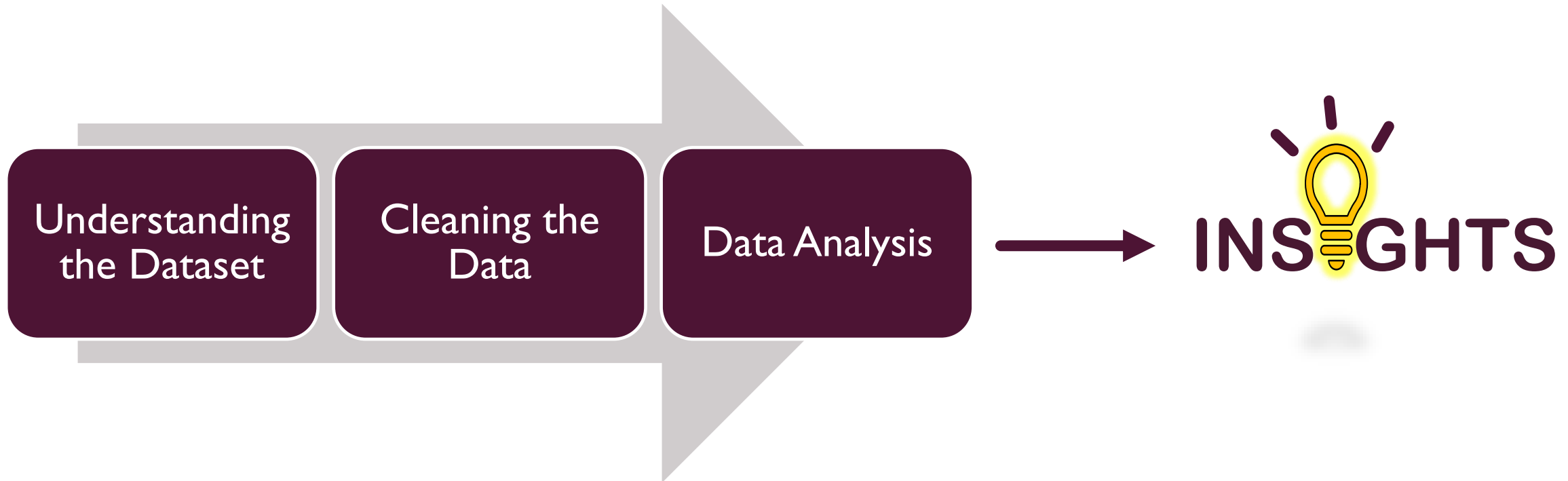
PROJECT DESCRIPTION:

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.

For the given dataset, as a Data Analyst, the client has asked 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, using data analysis techniques such as **regression analysis** and **market segmentation** to identify which features and categories are most popular among consumers and most profitable for the manufacturer. This could help the manufacturer improve its competitiveness in the market and increase its profitability over time.

APPROACH:



TECH-STACK USED:



- ❖ We have used Microsoft® Excel® 2021 MSO (Version 2310 Build 16.0.16924.20054) 64-bit for our data analysis as Excel is jam-packed with features and functions that can be used to clean, aggregate, pivot, and graph data. Also, Excel has a user-friendly visual interface that allows individuals at any level of expertise to easily learn and utilize its capabilities.
- ❖ Add-in, such as Analysis ToolPak was used for Regression Analysis.

[Excel File](#)

[Video Presentation](#)

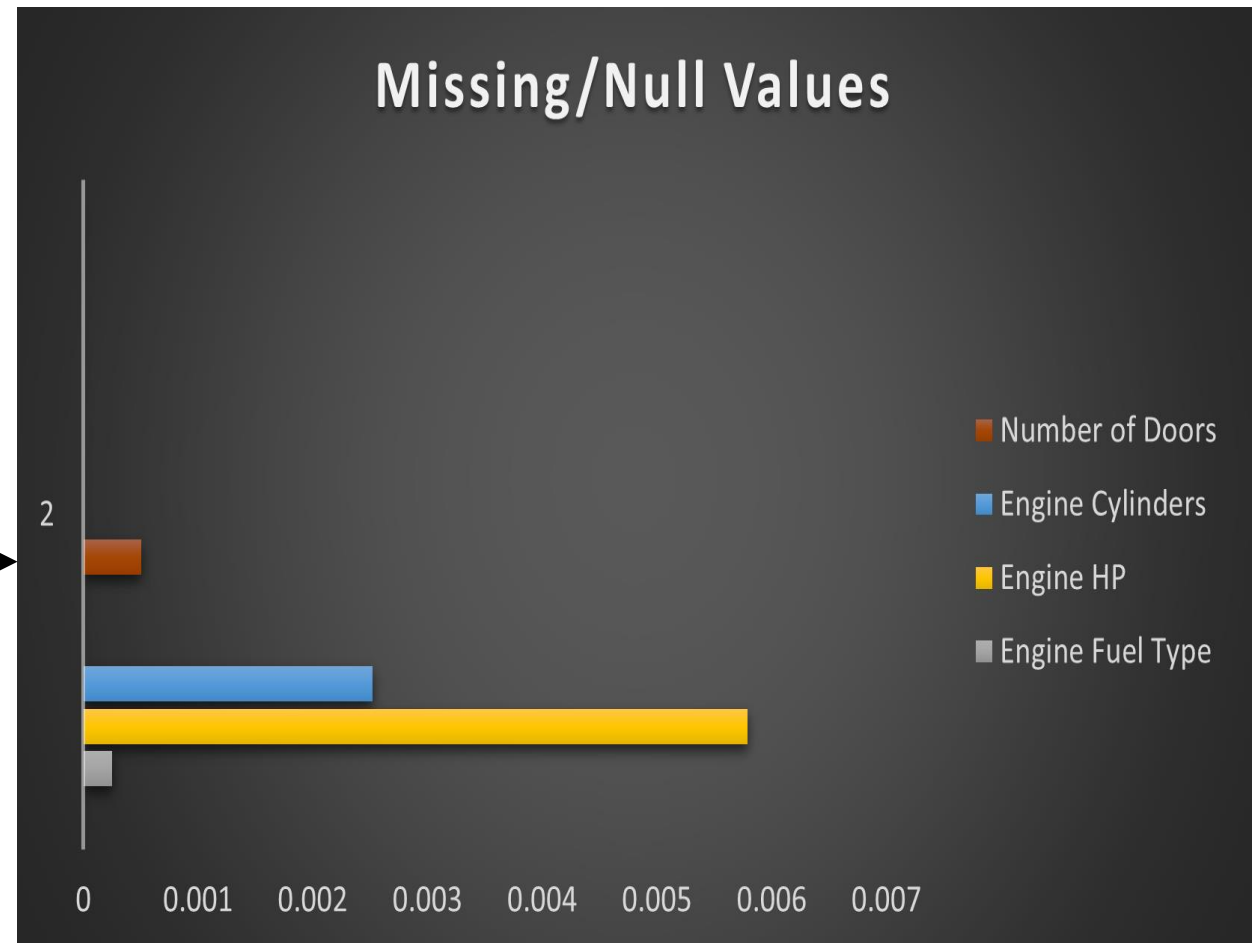
THE DATASET:

ROWS = 11914

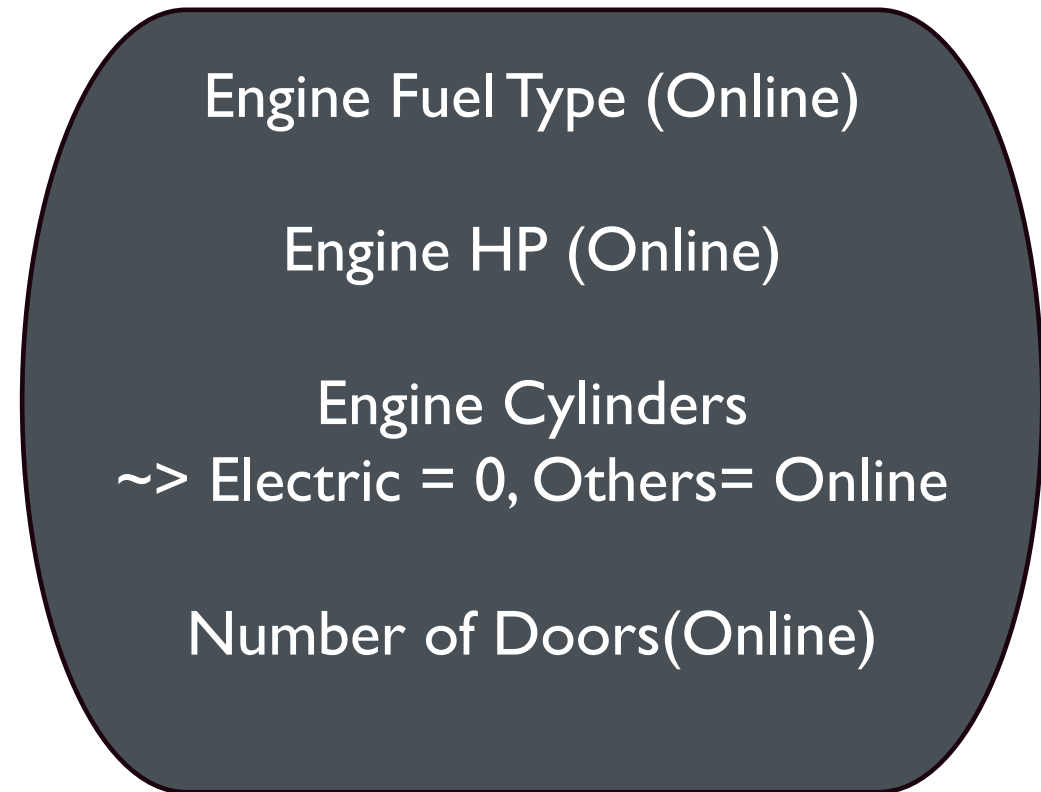
COLUMNS = 16

Duplicates = 715

Missing/Null
Values



DATA CLEANING:



OUTLIERS:

Engine HP



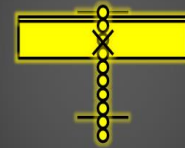
Engine Cylinders



Number of Doors



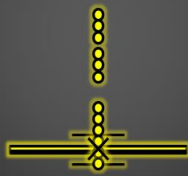
Year



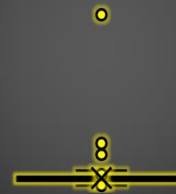
MSRP



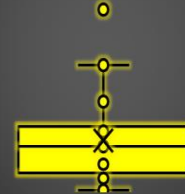
city mpg



highway MPG



Popularity



All the outliers are valid, so no data removed.

CLEANED DATA

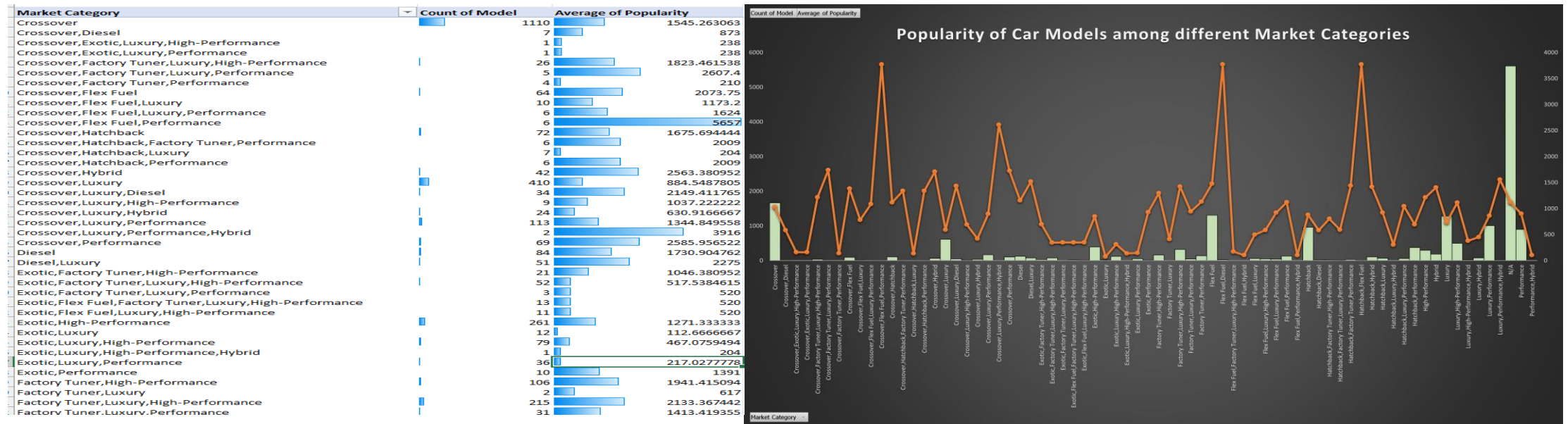


ROWS = 11199

COLUMNS = 16

INSIGHTS:

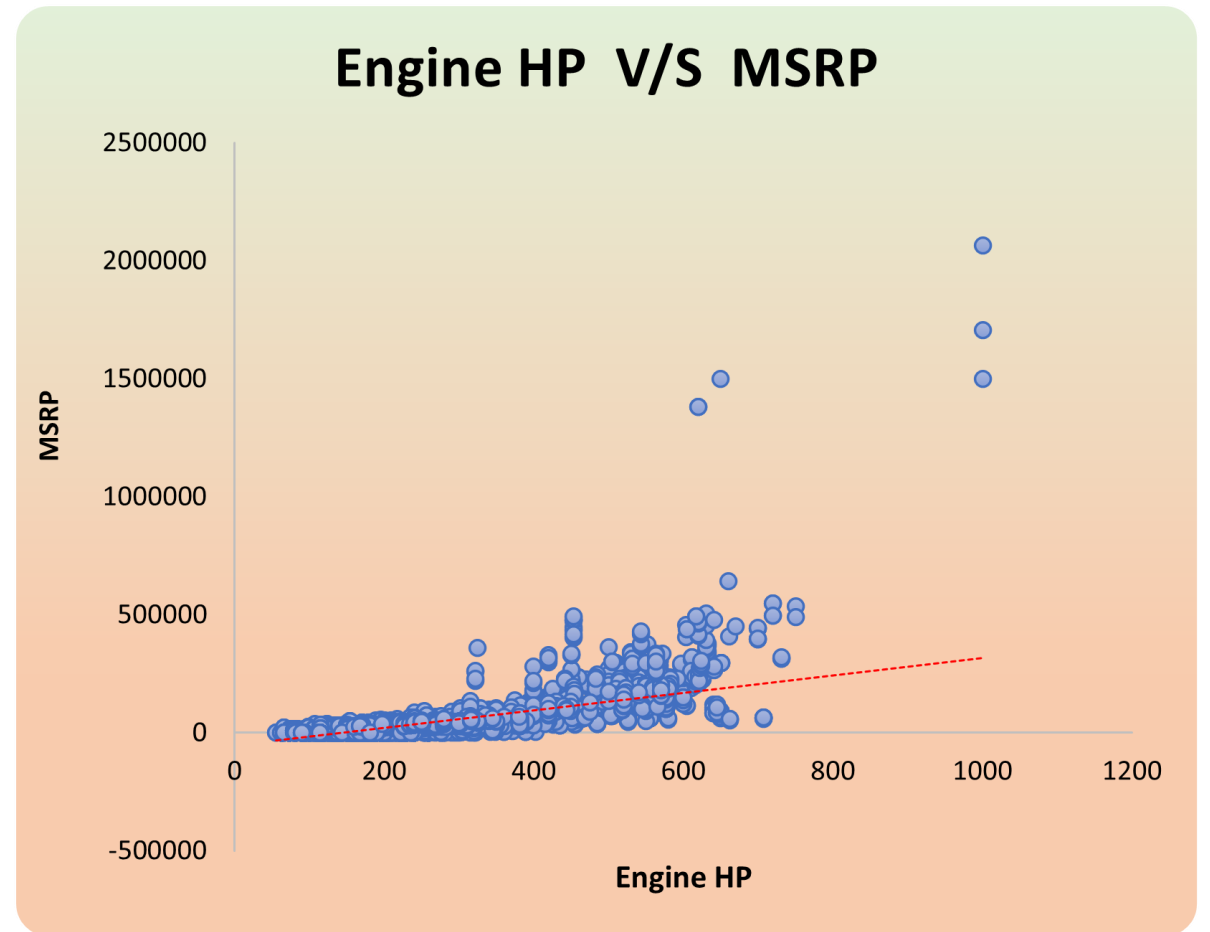
❖ How does the popularity of a car model vary across different market categories?



- Market Category N/A has the most number of car models followed by Crossover, Flex Fuel and Luxury.
- Crossover, Flex Fuel, Performance, Flex Fuel,Diesel, Hatchback,Flex Fuel are the most popular Market Category, followed by Crossover,Luxury,Performance,Hybrid.

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

- We can check from the scatter plot that Car's Engine Power and its Price have a positive relationship i.e., as we increase the HP, MSRP will also increase.



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

SUMMARY OUTPUT

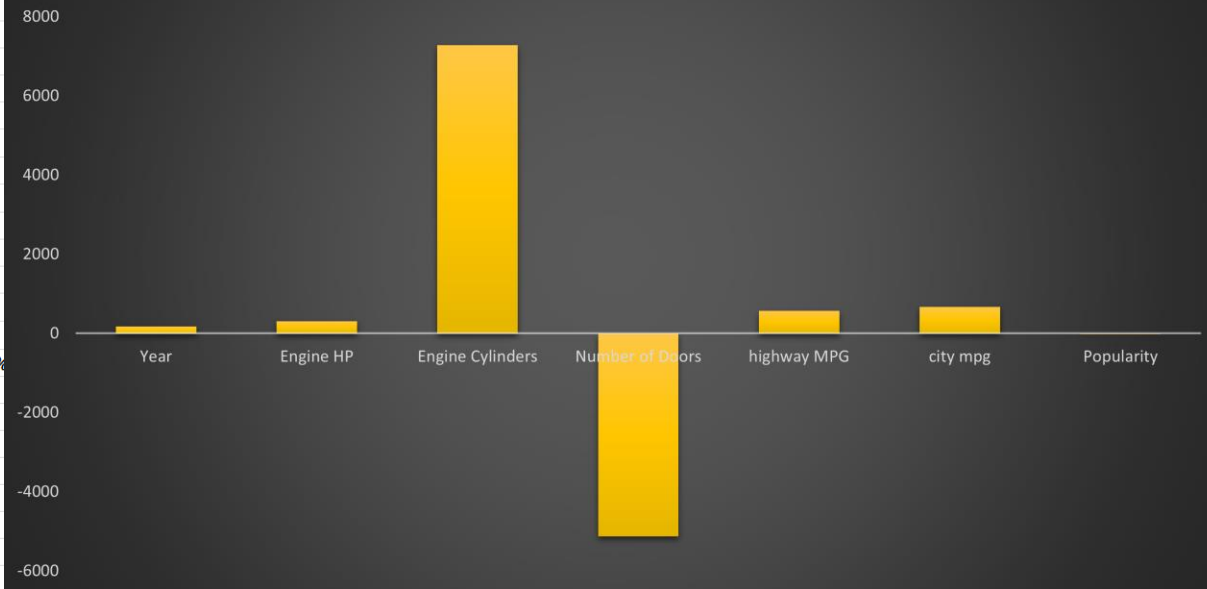
Regression Statistics	
Multiple R	0.681461415
R Square	0.46438966
Adjusted R Square	0.464054634
Standard Error	45048.75598
Observations	11199

ANOVA

	df	SS	MS	F	Significance F
Regression	7	1.9691E+13	2.813E+12	1386.131536	0
Residual	11191	2.27109E+13	2029390415		
Total	11198	4.24019E+13			

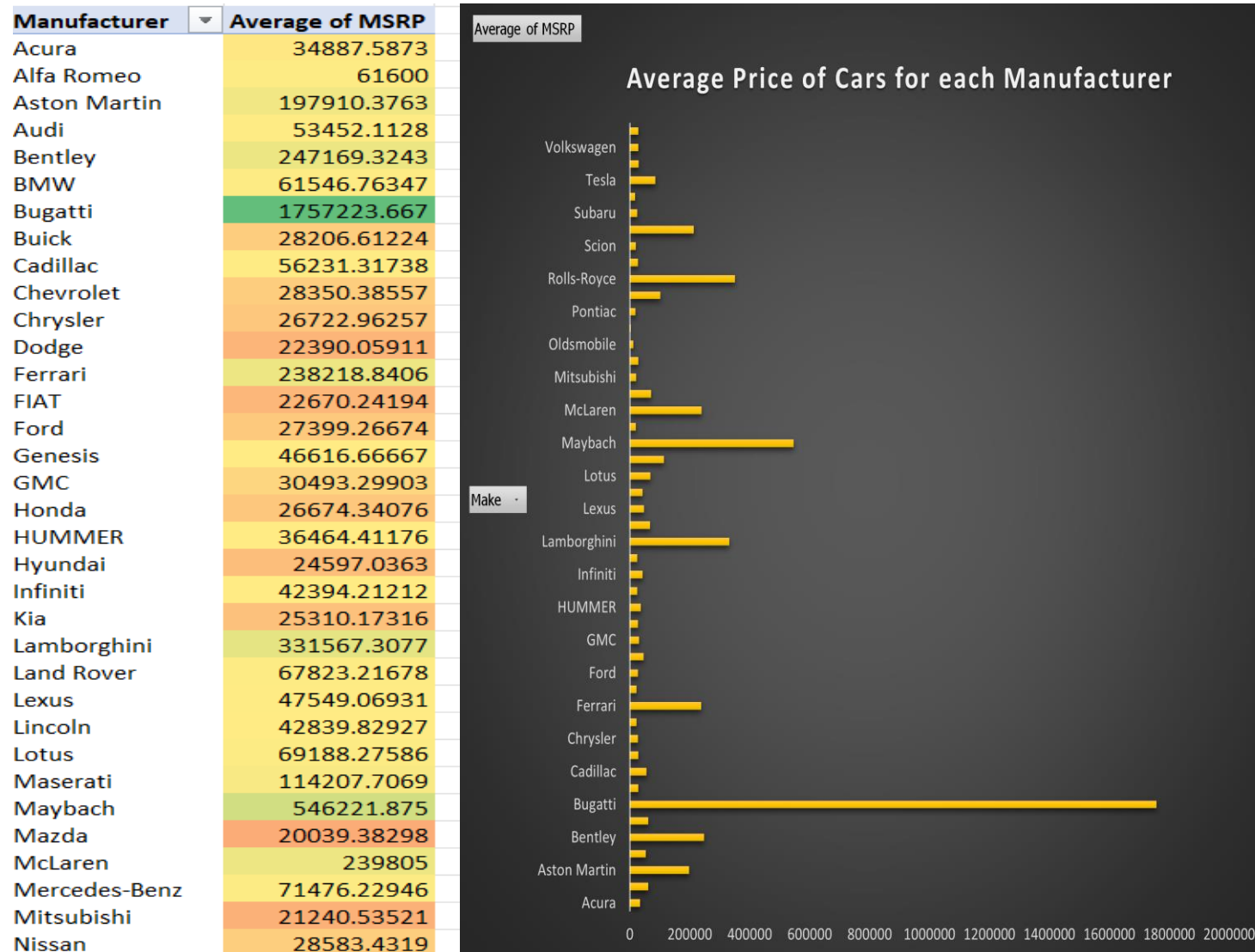
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-417900.1822	155576.4038	-2.686141163	0.007238923	-722857.3132	-112943.0513	-722857.3132	-112943.0513
Year	167.0242063	77.93346492	2.143164127	0.032121418	14.2608998	319.7875129	14.2608998	319.788
Engine HP	306.6558484	7.620932685	40.23862446	0	291.7174791	321.5942176	291.7174791	321.594
Engine Cylinders	7277.983696	478.7334256	15.20258103	1.11445E-51	6339.581931	8216.385461	6339.581931	8216.39
Number of Doors	-5130.993068	522.4266686	-9.821460842	1.12389E-22	-6155.041279	-4106.944857	-6155.041279	-4106.9
highway MPG	573.6524568	108.1958572	5.301981717	1.16733E-07	361.5695356	785.735378	361.5695356	785.735
city mpg	668.4024188	101.2509907	6.601440777	4.25599E-11	469.9326581	866.8721795	469.9326581	866.872
Popularity	-3.545184762	0.297678286	-11.90945035	1.66044E-32	-4.128686589	-2.961682935	-4.128686589	-2.9617

Coefficient of Variables



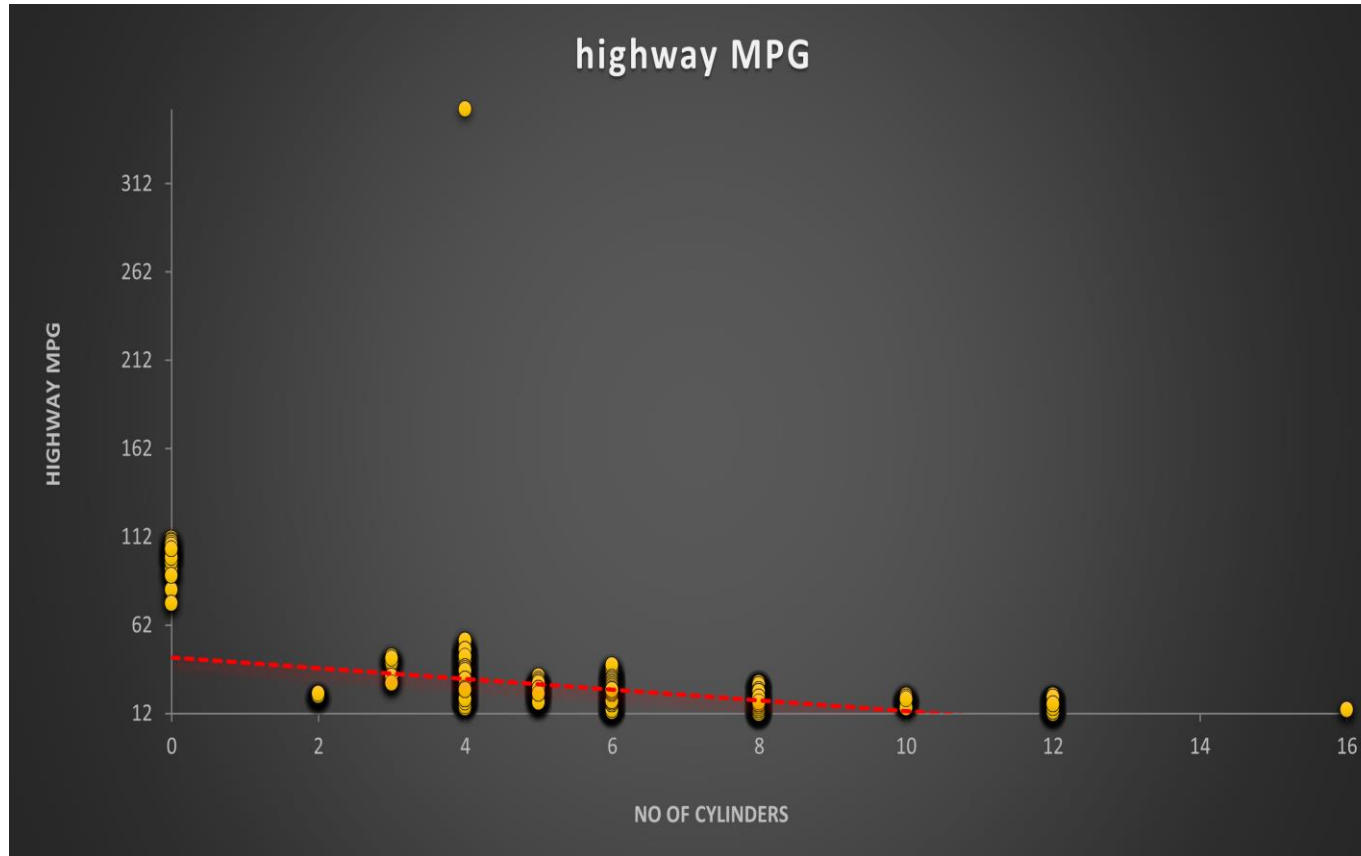
- Engine Cylinders is most important while determining a car's price, followed by city MPG, highway MPG and Engine HP.

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



- Bugatti has the highest average price, followed by Maybach, Rolls-Royce and Lamborghini

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



- We can see from the scatter-plot that the fuel efficiency decreases as the number of cylinders in the car's engine increases ie., MPG and No of Cylinders has a negative relationship.
- Correlation Coefficient also says that both have a negative correlation.

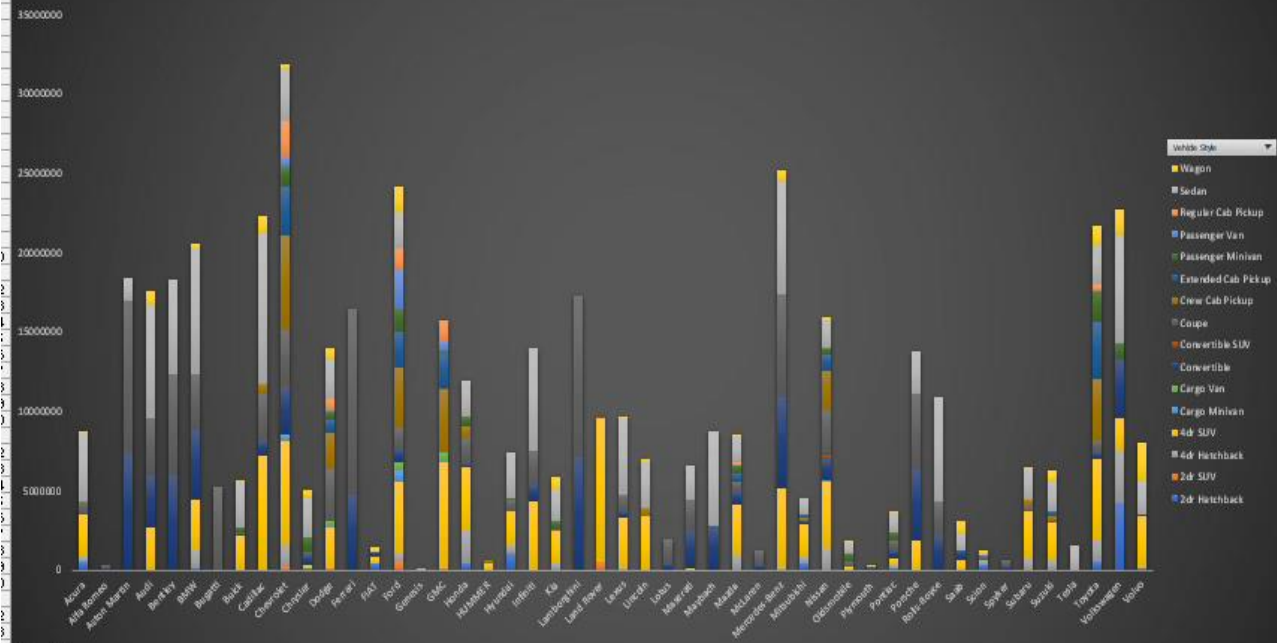
CORRELATION COEFFICIENT = -0.61433

DASHBOARD

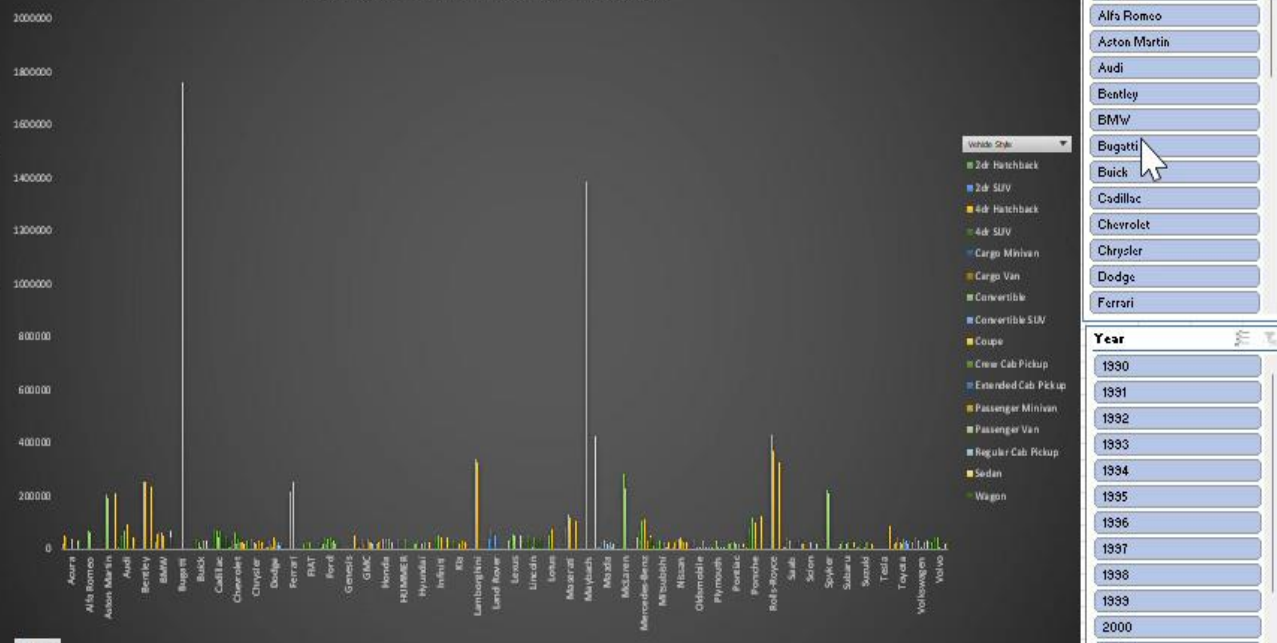
TASKS

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

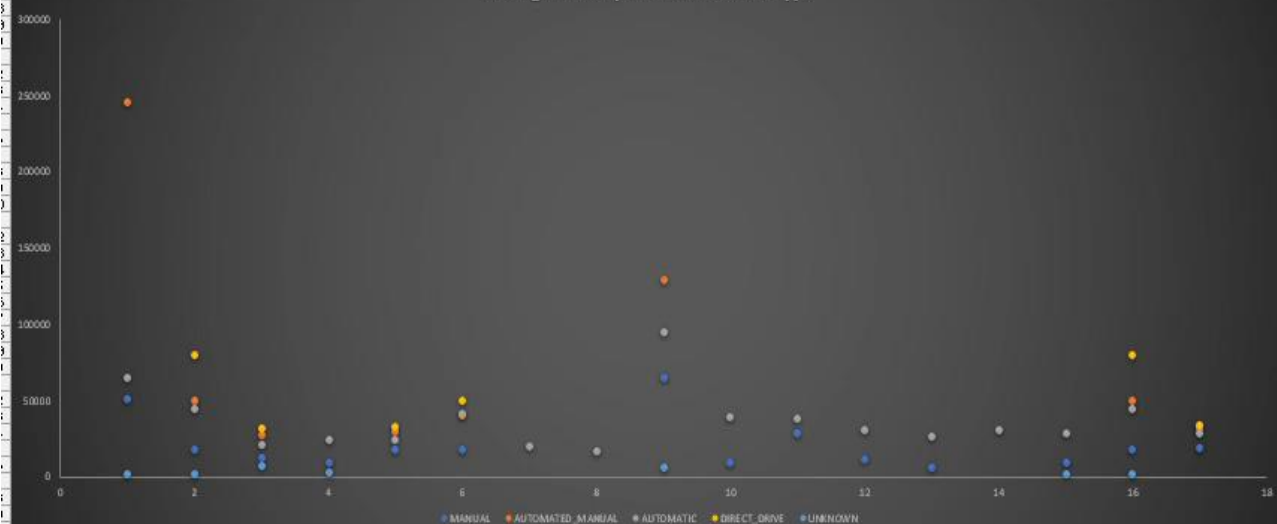
Total Price per Brand and Body Style



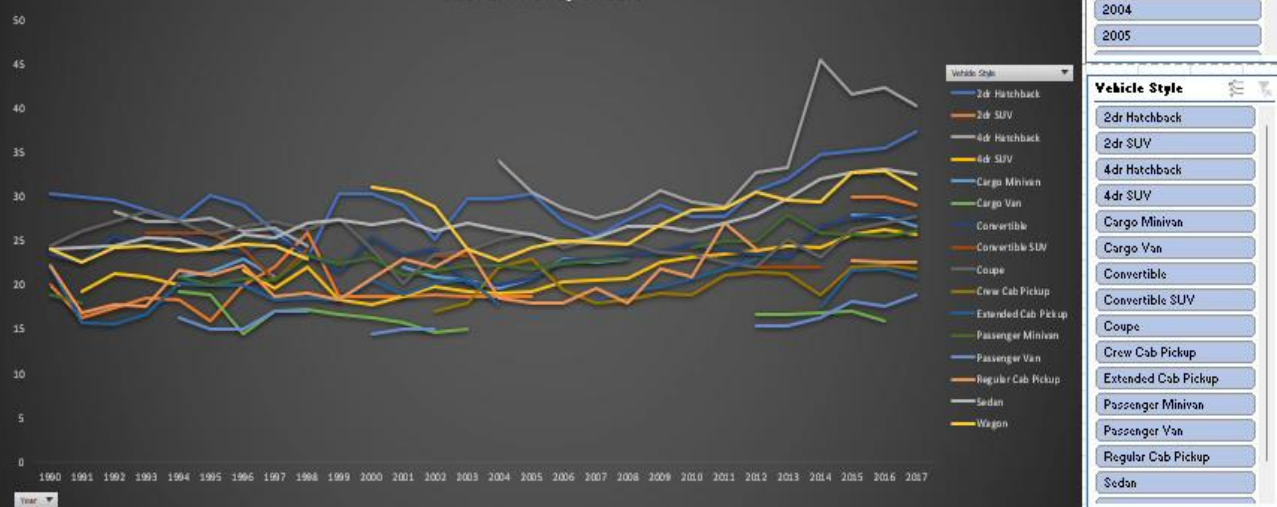
Average Price per Brand and Body Style



Average MSRP per Transmission Type



Fuel Efficiency of Cars



Make

- Acura
- Alfa Romeo
- Auton Martin
- Audi
- Bentley
- BMW
- Bugatti
- Buick
- Cadillac
- Chevrolet
- Chrysler
- Dodge
- Ferrari

Year

- 1990
- 1991
- 1992
- 1993
- 1994
- 1995
- 1996
- 1997
- 1998
- 1999
- 2000
- 2001
- 2002
- 2003
- 2004
- 2005

Vehicle Style

- 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

KEY INSIGHTS & SUGGESTIONS:

- In the Market Category, **N/A** has the most number of car models followed by **Crossover, Flex Fuel and Luxury**. So, we can research more on this category and find out in which category it falls. Also, **Crossover, Flex Fuel, Performance, Flex Fuel, Diesel, Hatchback, Flex Fuel** are the most popular Market Category, followed by **Crossover, Luxury, Performance, Hybrid**.
Hence, a strategy can be made based on this and company can make more cars production for better profit.
- Scatter Plot shows that Car's Engine Power and its Price have a **positive** relationship and fuel efficiency and the no of cylinders has a **negative correlation**.
- **Engine Cylinders** is most important while determining a car's price, followed by **city MPG, highway MPG and Engine HP**. Thus, the company can keep this when deciding the car's price.
- **Bugatti** has the highest average price , followed by **Maybach, Rolls-Royce and Lamborghini**.
- From the **Dashboard**, we see that **Bugatti** has the **strongest** relationship among HP, MPG and Price.

RESULT:

Completing the project, helped me understand how to preprocess the data, clean it and handle missing values. I learned how to perform Regression Analysis and Market Segmentation using add-ins like Analysis Toolpak. The project required extensive use of EXCEL, its formulas and functions, pivot tables, scatter plots, etc. Also, I got a better understanding on the use of slicers and filters to make interactive dashboards.

Thank
you