



Analyzing the Impact of Car Features on Price and Profitability

By Somella Rohith

CONTENT

Introduction

Project Description

Approach

Tech Stack Used

Insights

Result



INTRODUCTION

UNDERSTANDING THE PROJECT CONCEPT

INTRODUCTION

- In this Project analysing the Trends of the Car features with cost.
- Analysing the Impact of Car features on Price of the Car.
- Understanding the models and body styles of the Car with different Brands.
- Performing Regressive Analysis to understand, which features plays important role in deciding the Car price.
- Creating dashboard using Power Bi to for further analysis.

PROJECT DESCRIPTION

PURPOSE OF THE PROJECT AND DATASET DETAILS

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.

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, 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.

DATASET DESCRIPTION

- 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.
- Here is a brief overview of the dataset:
 - **Number of observations:** 11,159
 - **Number of variables:** 16
 - **File type:** CSV (Comma Separated Values)

DATASET DESCRIPTION: Variable Details

- **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
- **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
- **Market Category:** the market category the car belongs to (Luxury, Performance, etc...)

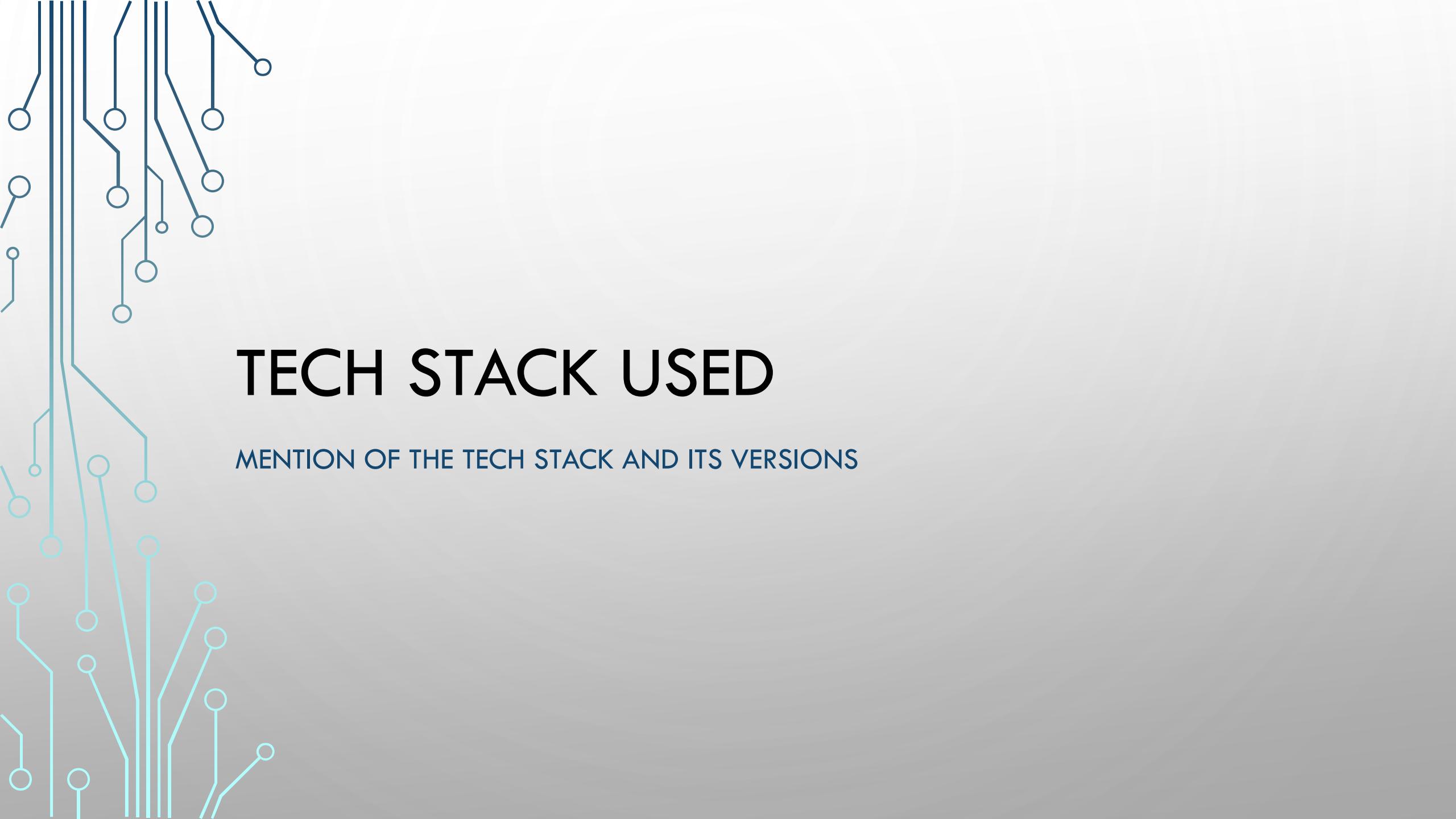
The background features a complex network of vertical and diagonal lines in various colors (dark blue, light blue, teal, yellow, orange) forming a branching, tree-like structure.

APPROACH

UNDERSTANDING OF MY APPROACH TO THIS PROJECT

APPROACH

- Used python to clean the data by removing the missing values w.r.t rows
- Used Excel to Analyse and visualise the given Tasks.
- Used Pivot tables to understand the Trends inside the dataset.
- Used regression analysis to understand the important features of the data and find out the proportionality of the Independent variables with the Dependent variables.
We can improve the accuracy (r^2 score) by optimising the model.
- Used Power Bi to create a Dashboard Report to Present the Insights.



TECH STACK USED

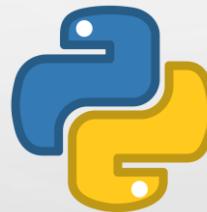
MENTION OF THE TECH STACK AND ITS VERSIONS

TECH STACK USED

1. Python 3.8
 - I. Google Colab
 - II. Data Preprocessing

2. Excel 2019
 - I. Pivot tables
 - II. Data Visualisation

3. Power Bi Desktop
 - I. Dashboard Report



The background features a complex, abstract graphic composed of numerous thin, colored lines (dark blue, teal, light blue) forming a network of nodes (small circles). These lines and circles are concentrated on the left side of the slide, creating a visual metaphor for data flow or connectivity.

INSIGHTS

TASKS, INSIGHTS, DATA VISUALS AND DASHBOARD REPORT

TASKS

1. Insight Required: 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 combo chart that visualizes the relationship between market category and popularity.

2. Insight Required: 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.

3. Insight Required: Which car features are most important in determining a car's price?

- **Task 3:** Use regression analysis to identify the variables that have the strongest relationship with a car's price. Then create a bar chart that shows the coefficient values for each variable to visualize their relative importance.

TASKS

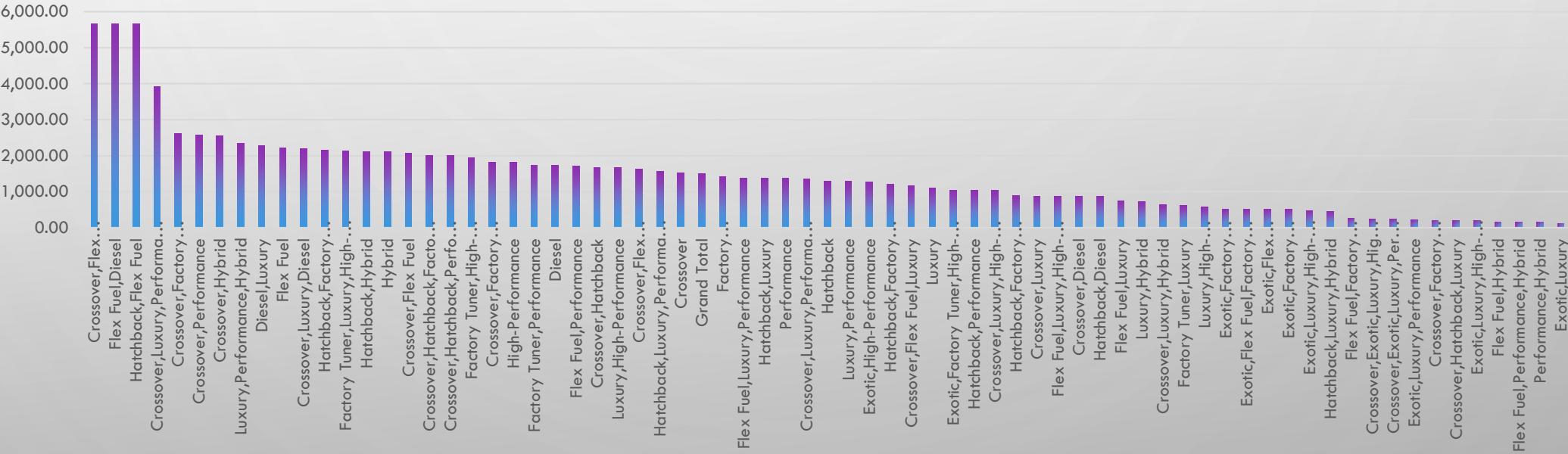
4. Insight Required: How does the average price of a car vary across different manufacturers?

- **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 manufacturer and average price.

5. Insight Required: 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.

TASK 1. POPULARITY OF CAR MODEL

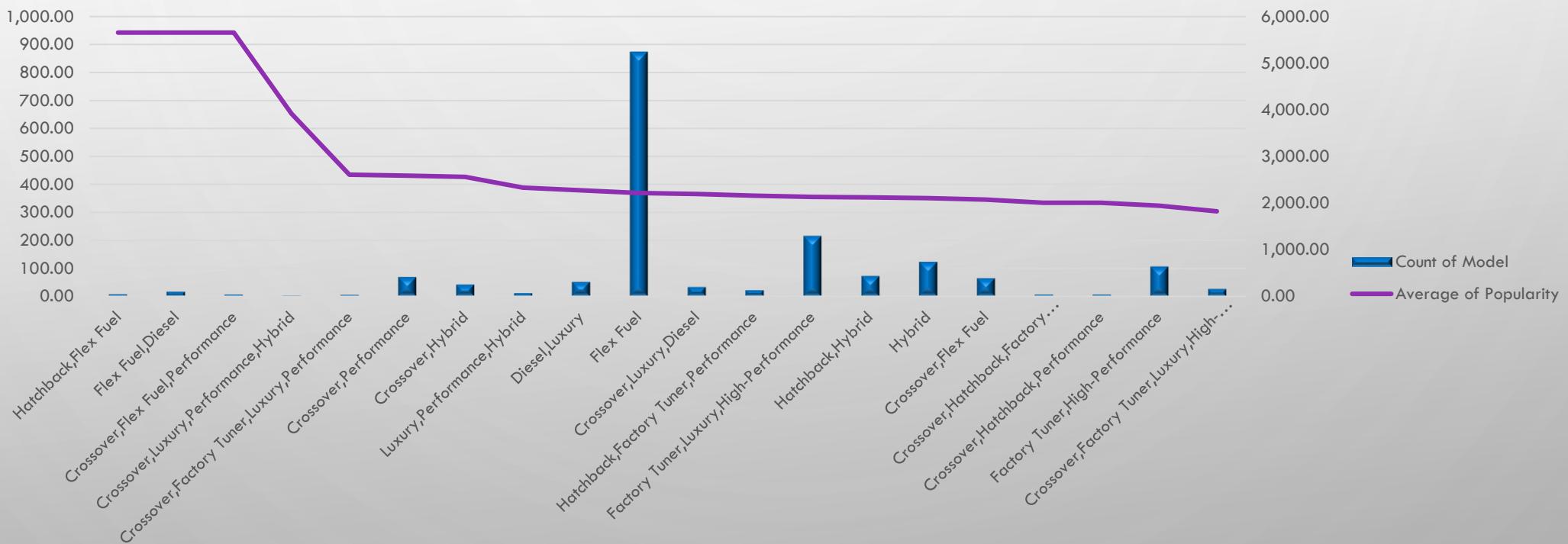


TASK 1A. PIVOT TABLE: Car Models In Each Market Category And Popularity.

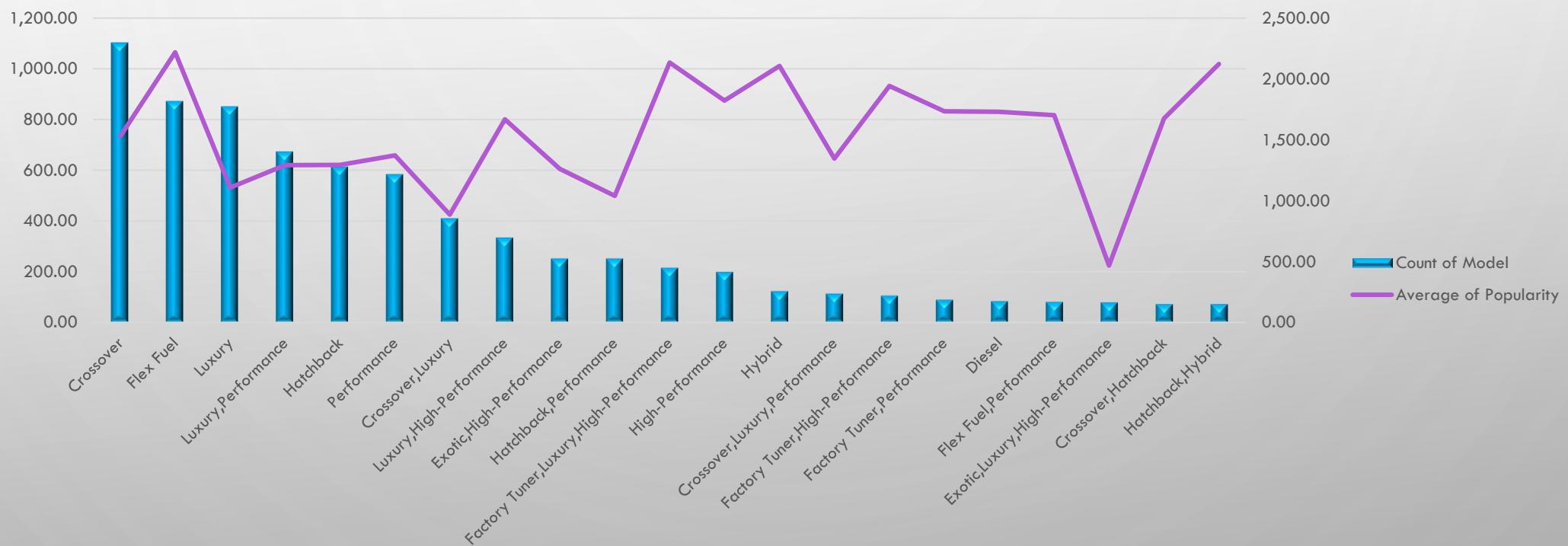
Market Category	Count of Model	Average of Popularity
Crossover	1,103.00	1,529.03
Crossover,Diesel	7.00	873.00
Crossover,Exotic,Luxury,High-Performance	1.00	238.00
Crossover,Exotic,Luxury,Performance	1.00	238.00
Crossover,Factory Tuner,Luxury,High-Performance	26.00	1,823.46
Crossover,Factory Tuner,Luxury,Performance	5.00	2,607.40
Crossover,Factory Tuner,Performance	4.00	210.00
Crossover,Flex Fuel	64.00	2,073.75
Crossover,Flex Fuel,Luxury	10.00	1,173.20
Crossover,Flex Fuel,Luxury,Performance	6.00	1,624.00
Crossover,Flex Fuel,Performance	6.00	5,657.00
Crossover,Hatchback	72.00	1,675.69
Crossover,Hatchback,Factory Tuner,Performance	6.00	2,009.00
Crossover,Hatchback,Luxury	7.00	204.00
Crossover,Hatchback,Performance	6.00	2,009.00
Crossover,Hybrid	42.00	2,563.38
Crossover,Luxury	410.00	884.55
Crossover,Luxury,Diesel	33.00	2,195.85
Crossover,Luxury,High-Performance	9.00	1,037.22
Crossover,Luxury,Hybrid	24.00	630.92
Crossover,Luxury,Performance	113.00	1,344.85
Crossover,Luxury,Performance,Hybrid	2.00	3,916.00
Crossover,Performance	69.00	2,585.96

Diesel	84.00	1,730.90
Diesel,Luxury	51.00	2,275.00
Exotic,Factory Tuner,High-Performance	21.00	1,046.38
Exotic,Factory Tuner,Luxury,High-Performance	52.00	517.54
Exotic,Factory Tuner,Luxury,Performance	3.00	520.00
Exotic,Flex Fuel,Factory Tuner,Luxury,High-Performance	13.00	520.00
Exotic,Flex Fuel,Luxury,High-Performance	11.00	520.00
Exotic,High-Performance	252.00	1,261.57
Exotic,Luxury	12.00	112.67
Exotic,Luxury,High-Performance	79.00	467.08
Exotic,Luxury,High-Performance,Hybrid	1.00	204.00
Exotic,Luxury,Performance	36.00	217.03
Factory Tuner,High-Performance	106.00	1,941.42
Factory Tuner,Luxury	2.00	617.00
Factory Tuner,Luxury,High-Performance	21.50	2,133.37
Factory Tuner,Luxury,Performance	31.00	1,413.42
Factory Tuner,Performance	89.00	1,733.10
Flex Fuel	872.00	2,217.30
Flex Fuel,Diesel	16.00	5,657.00
Flex Fuel,Factory Tuner,Luxury,High-Performance	1.00	258.00
Flex Fuel,Hybrid	2.00	155.00
Flex Fuel,Luxury	39.00	746.54
Flex Fuel,Luxury,High-Performance	33.00	878.91
Flex Fuel,Luxury,Performance	28.00	1,380.07
Flex Fuel,Performance	81.00	1,702.36
Flex Fuel,Performance,Hybrid	2.00	155.00
Hatchback	614.00	1,293.00
Hatchback,Diesel	14.00	873.00
Hatchback,Factory Tuner,High-Performance	13.00	1,205.15
Hatchback,Factory Tuner,Luxury,Performance	9.00	886.89
Hatchback,Factory Tuner,Performance	22.00	2,159.05
Hatchback,Flex Fuel	7.00	5,657.00
Hatchback,Hybrid	72.00	2,121.25
Hatchback,Luxury	46.00	1,379.50
Hatchback,Luxury,Hybrid	3.00	454.00
Hatchback,Luxury,Performance	38.00	1,566.13

TASK 1B: RELATIONSHIP BETWEEN MARKET CATEGORY AND POPULARITY.



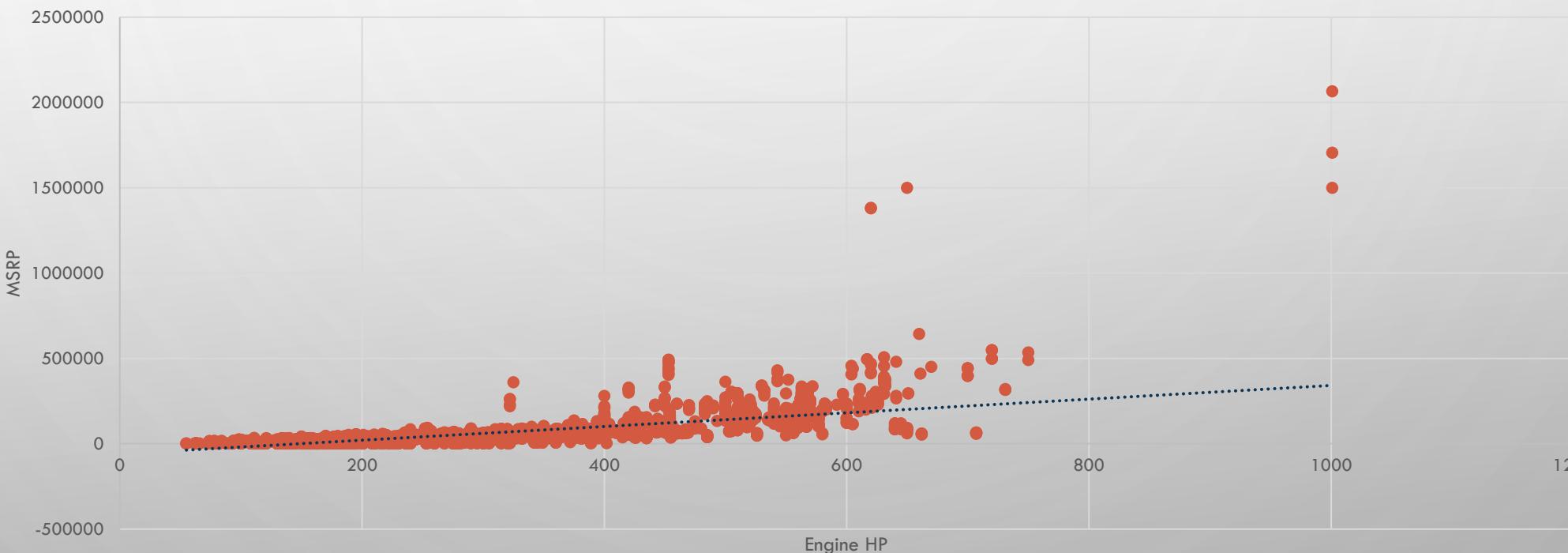
TASK 1B: HIGHEST MODEL COUNT WITH HIGHEST AVERAGE POPULARITY OF MARKET CATEGORY



TASK 1: INSIGHTS

- popularity of a car model vary across different market categories.
- 1st visual wont give the best understanding of the relationship between market category and popularity coz it contains low count of models category with High average popularity but in the 2nd visuals we can easily understand the relationship between market category and the popularity which contains highest count of models with highest popularity.
- Crossover,Flex Fuel,Performance | Flex Fuel,Diesel | Hatchback,Flex Fuel are have the Higher average Population of 5,657.00 lowest count of models
- Crossover, Flex Fuel, Luxury these categories have highest average population with 800 to 1100 count of models
- in which flex fuel score top with max of 2,217 average population with 872 count of models.

TASK 2: RELATIONSHIP BETWEEN A CAR'S ENGINE POWER AND ITS PRICE



TASK 2: INSIGHTS

- relationship between a car's engine power and its price.
- Here we can see with the help of trend line there is a slightly positive correlation of Price(MSRP) with the Engine HP.
- Minimum of Engine Horse Power is 55HP and Maximum was 1001HP
- Most of the Engine HP values are from 55HP to 600HP
- There is high positive correlation from 600HP of Engine Horse Power with Price(MSRP).

TASK 3: REGRESSION ANALYSIS

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.682034332							
R Square	0.465170831							
Adjusted R Square	0.464681892							
Standard Error	52432.95163							
Observations	7665							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	7	1.8309E+13	2.61557E+12	951.3887143	0			
Residual	7657	2.10507E+13	2749214416					
Total	7664	3.93597E+13						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-278380.6734	240778.0302	-1.156171405	0.247647126	-750371.5496	193610.2028	-750371.5496	193610.2028
Year	82.97568507	120.4905391	0.688648965	0.491065064	-153.218768	319.1701381	-153.218768	319.1701381
Engine HP	318.0617235	10.4227743	30.51603289	4.0788E-193	297.6302315	338.4932154	297.6302315	338.4932154
Engine Cylinders	10348.7952	633.3362415	16.34012791	5.06043E-59	9107.282724	11590.30767	9107.282724	11590.30767
Number of Doors	-4942.55206	714.8364828	-6.914241478	5.08224E-12	-6343.827324	-3541.276795	-6343.827324	-3541.276795
highway MPG	393.3588139	138.2144462	2.846003617	0.004438933	122.4206493	664.2969785	122.4206493	664.2969785
city mpg	1381.104907	150.0601499	9.203675378	4.38875E-20	1086.945919	1675.263895	1086.945919	1675.263895
Popularity	-4.34701955	0.423762494	-10.25815076	1.56693E-24	-5.177710085	-3.516329014	-5.177710085	-3.516329014

TASK 3. CAR FEATURES ARE MOST IMPORTANT IN DETERMINING A CAR'S PRICE

COEFFICIENT VALUES FOR EACH VARIABLE
AND THEIR RELATIVE IMPORTANCE.

■ Year ■ Engine HP ■ Engine Cylinders ■ Number of Doors ■ highway MPG ■ city mpg ■ Popularity



TASK 3: INSIGHTS

- R2 score = evaluation value of the prediction of regression model
- Standard error = MSE (Mean Squared Error) = $\text{sum}(\text{Actual} - \text{Predicted})^2$
- Intercept = (starting point of slope (intersection point of y axis)) = -278380.6
- positive correlated variables = Engine HP, Engine Cylinders and MPG
- Negatively correlated variables = number of doors
- No correlation variables = Years and popularity

TASK 4A: PIVOT TABLE THAT SHOWS THE AVERAGE PRICE OF CARS FOR EACH MANUFACTURERS.

Brands (Make)	Average of MSRP
Plymouth	4,076.82
Mitsubishi	20,221.88
Scion	20,395.94
Suzuki	21,153.05
FIAT	22,370.66
Mazda	22,983.35
Pontiac	24,728.13
Subaru	26,407.22
Honda	26,956.89
Hyundai	26,985.96
Saab	27,413.50
Volvo	28,541.16
Volkswagen	29,908.67
Chrysler	29,978.87
Kia	30,149.31
Dodge	30,995.38
Toyota	31,044.22
Nissan	32,908.42
Ford	33,245.37
Buick	33,770.40
Oldsmobile	34,868.00
Acura	34,887.59
Chevrolet	35,833.28
HUMMER	36,464.41

GMC	37,385.75
Infiniti	42,394.21
Lincoln	42,494.37
Genesis	46,616.67
Lexus	47,549.07
Audi	53,452.11
Cadillac	56,231.32
BMW	61,546.76
Alfa Romeo	61,600.00
Land Rover	67,823.22
Lotus	69,188.28
Mercedes-Benz	71,537.81
Porsche	1,01,622.40
Maserati	1,14,207.71
Aston Martin	1,97,910.38
Spyker	2,13,323.33
Ferrari	2,37,383.82
McLaren	2,39,805.00
Bentley	2,47,169.32
Lamborghini	3,31,567.31
Rolls-Royce	3,51,130.65
Maybach	5,46,221.88
Bugatti	17,57,223.67
Grand Total	50,072.67

TASK 4B: RELATIONSHIP BETWEEN MANUFACTURER AND AVERAGE PRICE.

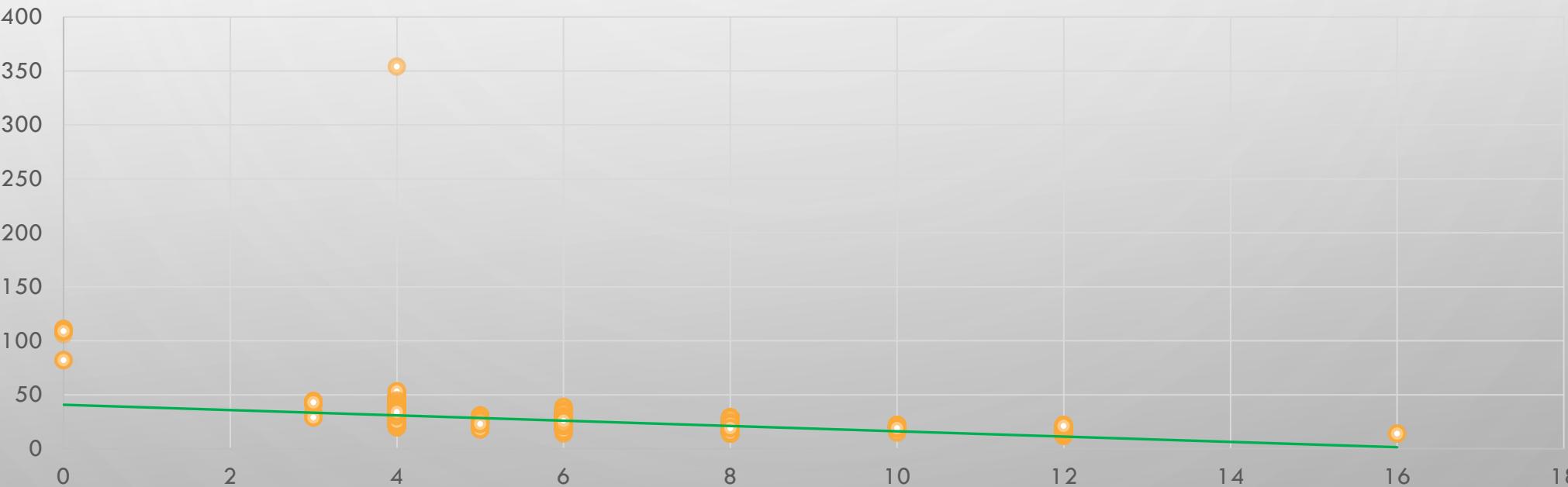


TASK 4: INSIGHTS

- Bugatti have the highest average price among all the other makers with \$17,57,223 and followed by Maybach and Rolls Royce.
- Previous visual represent the top 10 makes who have highest average Price of their Cars.
- But we also need to consider the number of the units manufactured for total revenue generated but yeah in terms of costliest Cars Bugatti was the top among all other makers.

TASK 5A: RELATIONSHIP BETWEEN FUEL EFFICIENCY AND THE NUMBER OF CYLINDERS IN A CAR'S ENGINE

highway MPG



TASK 5B: CORRELATION COEFFICIENT BETWEEN THE NUMBER OF CYLINDERS AND HIGHWAY MPG

Engine Cylinders	highway MPG	Corelation Coefficient
6	26	-0.597863785
6	28	-0.597863785
6	28	-0.597863785
6	28	-0.597863785
6	28	-0.597863785
6	28	-0.597863785
6	26	-0.597863785
6	28	-0.597863785
6	28	-0.597863785
6	27	-0.597863785
6	28	-0.597863785
6	28	-0.597863785
6	28	-0.597863785
6	25	-0.597863785
6	28	-0.597863785
6	24	-0.597863785
6	20	-0.597863785
6	21	-0.597863785

6	24	-0.597863785
6	20	-0.597863785
6	21	-0.597863785
6	21	-0.597863785
6	22	-0.597863785
6	22	-0.597863785
6	22	-0.597863785
6	21	-0.597863785
4	35	-0.597863785
4	35	-0.597863785
4	35	-0.597863785
4	26	-0.597863785
6	25	-0.597863785
6	25	-0.597863785
4	26	-0.597863785
4	26	-0.597863785
6	25	-0.597863785
4	35	-0.597863785
4	34	-0.597863785
6	31	-0.597863785

TASK 5: INSIGHTS

- There is a Negative correlation of Highway MPG with Number of Engine Cylinders which is -0.597
- So, as the numbers of cylinders of Engine is Increasing the MPG is decreasing, Same vice versa
- As the trend line suggest the negative correlation with fall of MPG as the number of cylinders Increasing. So, the number of cylinders are directly proportional to the loss of fuel efficiency.



DASHBOARD REPORT

POWER BI DASHBOARD REPORT

DASHBOARD TASKS

- **Task 1:** How does the distribution of car prices vary 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?
- **Task 3:** How do the different feature such as transmission type affect the MSRP, and how does this vary by body style?
- **Task 4:** How does the fuel efficiency of cars vary across different body styles and model years?
- **Task 5:** How does the car's horsepower, MPG, and price vary across different Brands?

Analyzing the Impact of Car Features on Price and Profitability

[Clear Filter](#)

Filters

2dr Hatchback

2dr SUV

4dr Hatchback

4dr SUV

Cargo Minivan

Cargo Van

Convertible

Convertible SUV

Coupe

Crew Cab Pickup

Extended Cab ...

Passenger Mini...

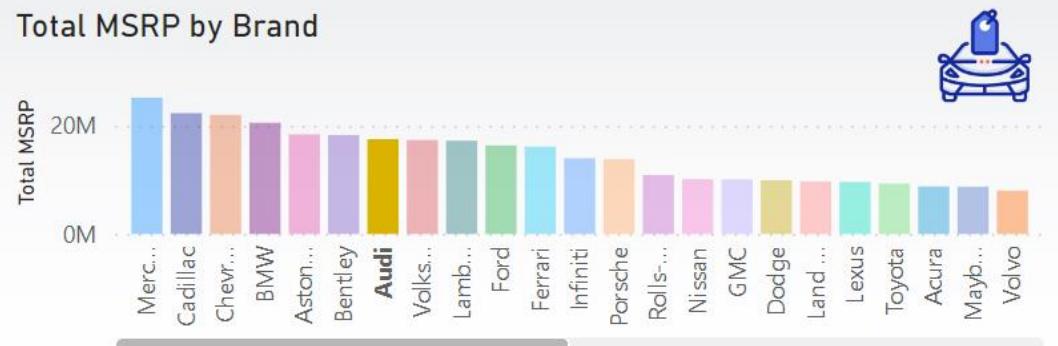
Passenger Van

Regular Cab Pi...

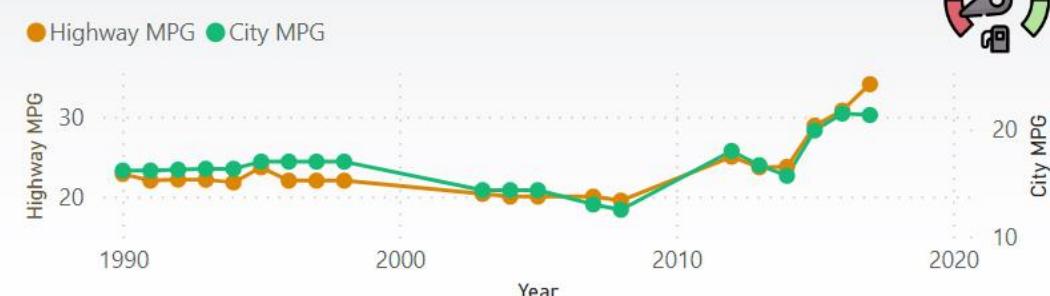
Sedan

Wagon

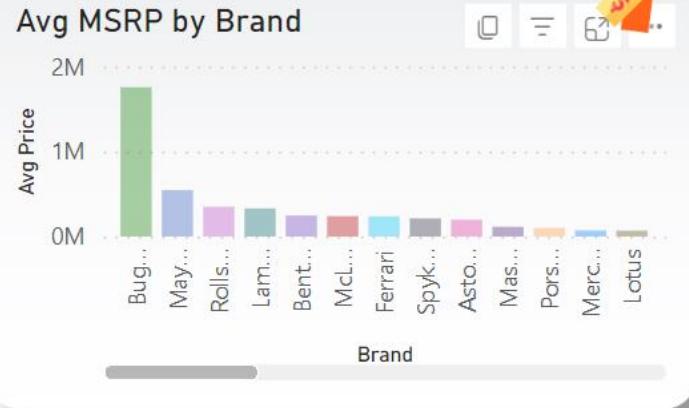
Total MSRP by Brand



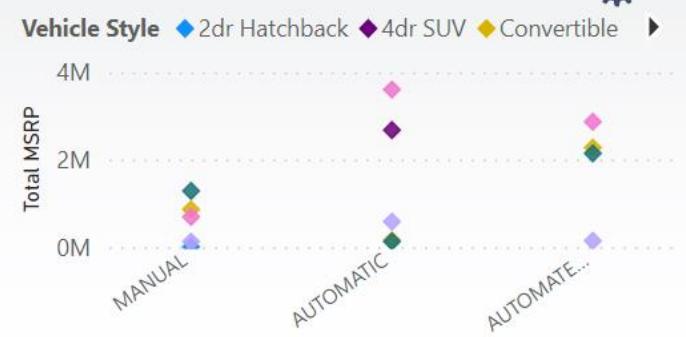
Fuel Efficiency by Model Year



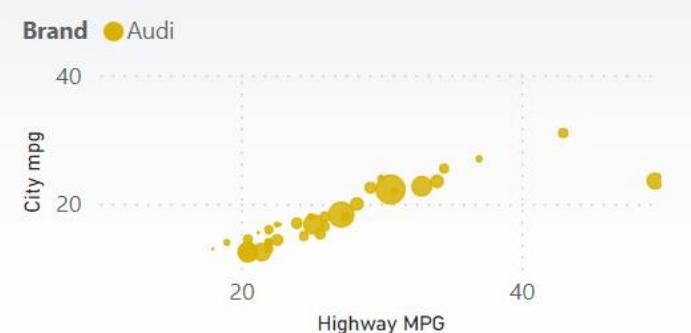
Avg MSRP by Brand



Style and Transmission by MSRP



HP, MPG and Price of Diff Brands



DASHBOARD: TASK 1 – INSIGHTS

DISTRIBUTION OF CAR PRICES VARY BY BRAND AND BODY STYLE

- There is a huge difference of body style and brand in terms of pricing.
- Highest was Mercedes Benz with maximum total MSRP for all models as an average.

1. Mercedes-Benz	2. Cadillac	3. Chevrolet
2,51,81,309	2,23,23,833	2,20,01,631

- Lowest was Oldsmobile with 1,04,604 of Minimum total MSRP for all models of their average.

1. Oldsmobile	2. Genesis	3. Plymouth
1,04,604	1,39,850	1,58,996

DASHBOARD: TASK 2 – INSIGHTS

HIGHEST FOR EACH MARKET CATEGORY

Market category	Brand (make)
2dr Hatchback	Chrysler
2dr SUV	Land Rover
4dr Hatchback	BMW
4dr SUV	Porsche
Cargo Minivan	Mercedes-Benz
Cargo Van	GMC
Convertible	Maybach
Convertible SUV	Land Rover

Market category	Brand (make)
Coupe	Bugatti
Crew Cab Pickup	Cadillac
Extended Cab Pickup	Ford
Passenger Minivan	Volkswagen
Passenger Van	Ford
Regular Cab Pickup	Toyota
Sedan	Maybach
Wagon	Cadillac

DASHBOARD: TASK 2 – INSIGHTS

LOWEST FOR EACH MARKET CATEGORY

Market category	Brand (make)
2dr Hatchback	Audi, Chevrolet, Dodge, Ford, Mazda, Plymouth, Saab, Subaru with 2000
2dr SUV	GMC
4dr Hatchback	Dodge, Plymouth
4dr SUV	FIAT
Cargo Minivan	Dodge
Cargo Van	Ford
Convertible	Buick, Dodge
Convertible SUV	Nissan

Market category	Brand (make)
Coupe	Mazda, Volvo
Crew Cab Pickup	Subaru
Extended Cab Pickup	Mazda
Passenger Minivan	Plymouth
Passenger Van	Chevrolet, GMC
Regular Cab Pickup	Mazda
Sedan	Volvo
Wagon	Chevrolet

DASHBOARD: TASK 3 – INSIGHTS

DIFFERENT FEATURE SUCH AS TRANSMISSION TYPE AFFECT THE MSRP, AND HOW DOES THIS VARY BY BODY STYLE

- Mercedes-Benzes, BMW, Cadillac, Audi, Lexus, Lincoln, Chevrolet, Acura, Ford, Toyota, Dodge, Chrysler, Volvo, Saab. There are the premium brands which offers Automatic Transmission and apart from these other brands are offering only manual or direct drive.
- Maybach has the highest Average price among all other Makers with \$ 5,46,221. which offers Automatic Transmission.
- Automatic Transmission have the highest MSRP as compared with other Transmission types.
- 2 door SUV, 4 door SUV, sedan, and wagon. Vehicle body style have the highest MSRP (80M to 90M+) than in Automatic or Automatic Manual.

DASHBOARD: TASK 4 – INSIGHTS

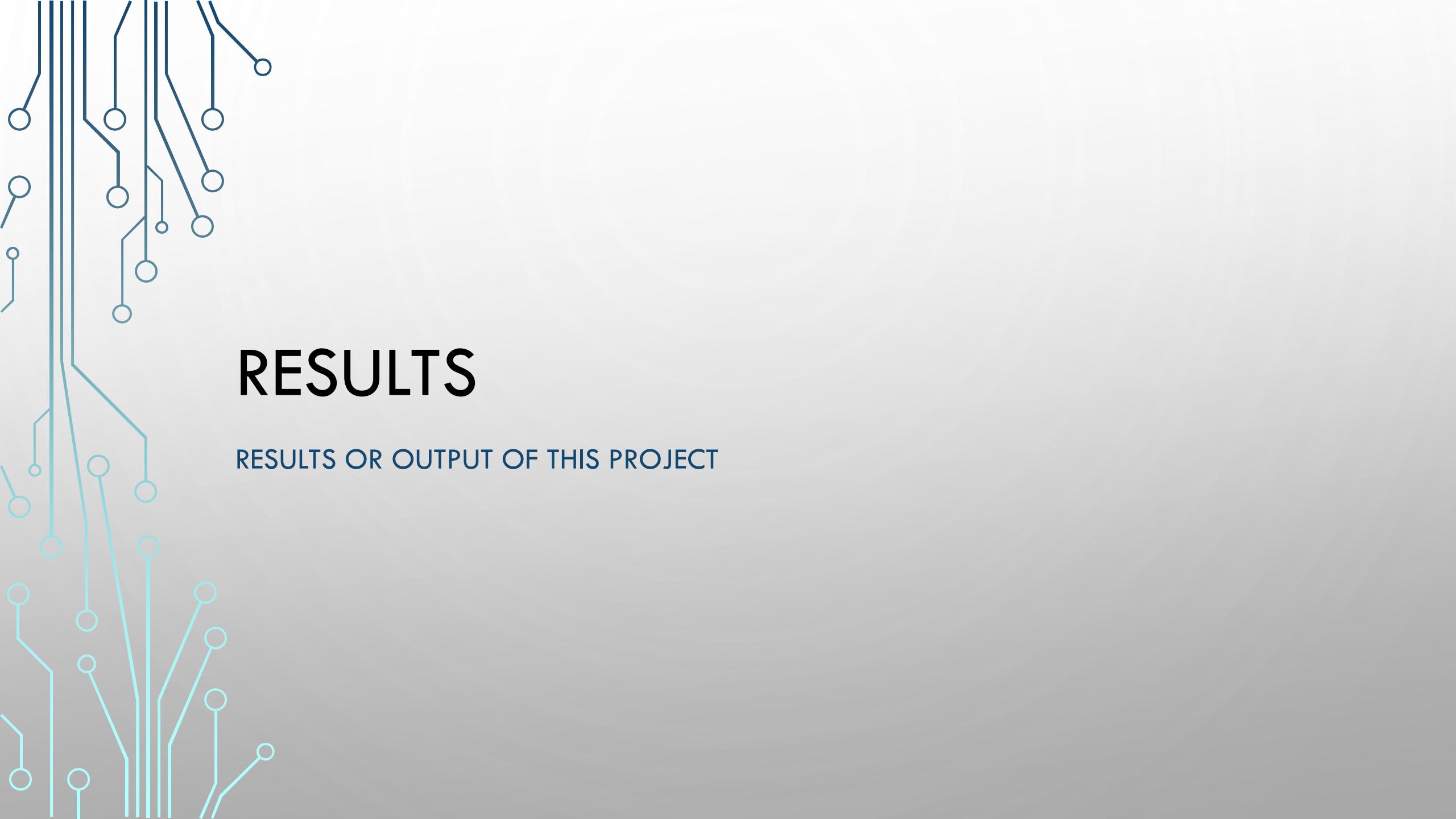
THE FUEL EFFICIENCY OF CARS VARY ACROSS DIFFERENT BODY STYLES AND MODEL YEARS

- Each and every body style have the rise in the fuel efficiency except cargo vans from 1990 to 2017
- But most of the body styles have a downfall in the fuel efficiency between 2000 and 2008
- In every body type both Highway MPG and City MPG have similar growth rate over the years.
- Hatchbacks and Passenger vehicles have better fuel efficiency in Highway over city rides.

DASHBOARD: TASK 5 – INSIGHTS

DOES THE CAR'S HORSEPOWER, MPG, AND PRICE VARY ACROSS DIFFERENT BRANDS

- There is a Negative corelation of Engine HP and Both Highway MPG, City MPG. As the Engine HP Increases the Milage decreases and as the Engine HP decrease the Fuel Efficiency Increases.
- there is no influence of any brand in this negative corelation, for every car there is same negative corelation value -0.46
- But there is a positive correlation of both Highway MPG and City MPG and perform according to the brand and body style.



RESULTS

RESULTS OR OUTPUT OF THIS PROJECT

CAR FEATURES' IMPACT ON PRICE AND PROFITABILITY BASED ON 5 TASKS

Analysing car features' influence on pricing and profitability reveals vital insights. Market category popularity indicates a clear connection with specific categories like Crossover, Flex Fuel, and Performance. Notably, higher engine power correlates with increased pricing, especially beyond 600HP. Features like Engine HP, Cylinders, and MPG play pivotal roles in determining car prices, while the number of doors has a negative impact. Bugatti stands out as a premium-priced manufacturer, offering valuable features. Fuel efficiency considerations in relation to engine cylinders emphasize the necessity for optimization while maintaining performance standards. These insights provide actionable strategies for pricing, feature emphasis, and overall competitiveness in the ever-evolving automotive industry.

DASHBOARD INSIGHTS:

AUTOMOTIVE ANALYSIS FOR INFORMED DECISION-MAKING

Our Power BI dashboard reveals crucial insights for the automotive industry. Examining car prices across brands and body styles, we identify stark variations. Mercedes Benz stands out with the highest average MSRP, while Oldsmobile records the lowest.

Transmission type influences pricing, with premium brands favoring automatic transmission. Maybach leads in average price at \$546,221. Fuel efficiency trends highlight improvements in most body styles, except cargo vans, with a dip between 2000 and 2008. Engine horsepower negatively correlates with MPG, and we find consistent positive correlations between MPG and brand or body style. These insights empower strategic decisions in pricing, product development, and efficiency optimization, enhancing industry competitiveness.

LINKS

- E mail: rohithsomella@gmail.com
- Project Folder: [Analyzing the Impact of Car Features on Price and Profitability](#)
- Excel File: [Car Features and their Impact](#)
- Power Bi File: [Car Features Analysis](#)
- Loom Video : [Analysis of Car Features and their Impact](#)

THANK YOU