IMPACT OF CAR FEATURES

DATA ANALYTICS

BY TUSHAR BHAGWAT

PROJECT DESCRIPTION

The automotive industry has been rapidly changing in recent years, with a strong emphasis on fuel efficiency, environmental sustainability, and technological advancements. This has led to a shift in consumer preferences towards electric and hybrid vehicles, as well as an interest in alternative fuel options like hydrogen and natural gas. However, traditional gasoline-powered cars still dominate the market, and there are different types and grades of fuel available.

The client wants to know how a car manufacturer can optimize their pricing and product development decisions to maximize profitability while meeting consumer demand. To tackle this problem, we can analyze the relationship between a car's features, market category, and pricing. By examining which features and categories are popular among consumers and most profitable for the manufacturer, we can develop a pricing strategy that balances both consumer demand and profitability.

To achieve this, we can utilize data analysis techniques such as regression analysis and market segmentation. Regression analysis helps us understand how different features and categories affect the pricing of cars. By identifying the features and categories that have the most significant impact on consumer demand and profitability, the manufacturer can make informed decisions on pricing.

Market segmentation allows us to group consumers based on their preferences, needs, and buying behaviour. By identifying different segments within the market, the manufacturer can tailor their pricing and product development strategies to target specific consumer groups effectively. This helps in optimizing profitability by catering to the demands of different customer segments.

By combining the findings from regression analysis and market segmentation, the car manufacturer can develop a pricing strategy that maximizes profitability while ensuring consumer satisfaction. Additionally, the analysis can provide insights into which product features should be emphasized in future product development efforts. This information helps the manufacturer stay competitive in the market and improve profitability over time

> **Data Details**

The dataset "Car Features and MSRP" contains information on various car models and their specifications. It was provided by Cooper Union, a private college in New York City, and is available on Kaggle. The dataset consists of 11,915 observations and 16 variables, and it is in CSV format.

This dataset can be utilized for several data analysis tasks. It allows for the exploration of trends in car features and pricing over time. Additionally, it enables comparisons of fuel efficiency among different car types. The dataset can also be used to investigate the relationship between a car's features and its popularity.

Furthermore, it can support the development of models to predict car prices based on their features and market category.

However, it's important to note that the dataset's last update was in 2017. Consequently, it may not reflect the most current trends or prices in the automotive industry.

> **Data Understanding**

The dataset contains detailed information on over 11,000 car models, including their make, model, year, fuel type, engine power, transmission, wheels, number of doors, market category, size, style, estimated miles per gallon (MPG), popularity, and manufacturer's suggested retail price (MSRP).

This dataset is valuable for data analysts who want to gain insights into different aspects of the automotive industry. They can:

- Analyze trends in car features and pricing over time: By examining the variables in the dataset, analysts can identify how car features and prices have evolved, aiding manufacturers in making informed decisions about product development and pricing.
- Compare the fuel efficiency of different car types: The MPG variables allow analysts to compare the fuel efficiency of various cars, enabling consumers to make informed purchasing decisions.
- Investigate the relationship between a car's features and its popularity: By analyzing the popularity variable, analysts can determine which features are popular among consumers and how they impact a car's popularity. This helps manufacturers with product development and marketing strategies.
- Predict a car's price based on its features and market category: By leveraging the dataset's features and market category variables, analysts can develop models to predict a car's price. This aids manufacturers and consumers in understanding how different features influence pricing decisions.

In summary, this dataset provides a valuable resource for data analysts exploring the automotive industry, enabling informed decisions related to product development, marketing, and pricing.

➢ **Problems and Planning**

The following tasks were provided to follow while doing the analysis

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.
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.
6. **Building Dashboard:** For the next portion of the Project, you need to create the Interactive Dashboard. Use filters and slicers to make the chart interactive.
   - **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?

➤ Data Cleaning

At first, to evade data corruption, I made a copy of the raw data so that the original data won't get effected. Next, I followed the following steps to clean and manipulate the data whenever needed:

- At first, I did an inspection and found out the following columns having null values Engine fuel type, Engine HP, Engine cylinders, Number of doors and Market Category
- Engine Fuel Type had only 3 blank cells corresponding to verona model. Therefore, after doing some research, I filled up the null cells with "regular unleaded".
- Removing the null value rows from Engine HP and Engine Cylinders and No. of Doors columns.
- Converting the values of MSRP into $ currency.
- After doing the data cleaning, I've been left with 11816 data from 11915 data.

Approach

Now that the data cleaning is complete, I'll be moving towards analysis process. I'll be processing the tasks provided, one by one and then end up creating a dashboard as suggested.

  ➢ **Analysis**

  ▪ **Task: 1.A:** To create a pivot table that shows the number of car models in each market category and their corresponding popularity scores.
  Once the pivot table is created, it should display the market categories as rows and the popularity scores as values along with the number of car models that falls under the categories.

Popularity of Car Model		
Market Categories	**Average of Popularity**	**Count of Model**
Hatchback,Flex Fuel	5657	7
Flex Fuel,Diesel	5657	16
Crossover,Flex Fuel,Performance	5657	6
Crossover,Luxury,Performance,Hybrid	3916	2
Crossover,Factory Tuner,Luxury,Performance	2607.4	5
Crossover,Performance	2585.956522	69
Crossover,Hybrid	2563.380952	42
Luxury,Performance,Hybrid	2333.181818	11
Diesel,Luxury	2275	51
Flex Fuel	2217.302752	872
Crossover,Luxury,Diesel	2195.848485	33
Hatchback,Factory Tuner,Performance	2159.045455	22
Factory Tuner,Luxury,High-Performance	2133.367442	215
Hatchback,Hybrid	2121.25	72
Hybrid	2105.569106	123
Crossover,Flex Fuel	2073.75	64
Crossover,Hatchback,Factory Tuner,Performance	2009	6
Crossover,Hatchback,Performance	2009	6
Factory Tuner,High-Performance	1941.415094	106
Crossover,Factory Tuner,Luxury,High-Performance	1823.461538	26
High-Performance	1821.447236	199
Factory Tuner,Performance	1733.101124	89
Diesel	1730.904762	84
Flex Fuel,Performance	1702.358025	81
Crossover,Hatchback	1675.694444	72
Luxury,High-Performance	1668.017964	334
Crossover,Flex Fuel,Luxury,Performance	1624	6
Hatchback,Luxury,Performance	1566.131579	38
Crossover	1529.030825	1103
Factory Tuner,Luxury,Performance	1413.419355	31
Flex Fuel,Luxury,Performance	1380.071429	28
Hatchback,Luxury	1379.5	46
Performance	1371.080479	584
Crossover,Luxury,Performance	1344.849558	113
Hatchback	1292.998371	614
Luxury,Performance	1292.615156	673
Exotic,High-Performance	1261.571429	252
Hatchback,Factory Tuner,High-Performance	1205.153846	13
Crossover,Flex Fuel,Luxury	1173.2	10
Luxury	1107.553467	851
Exotic,Factory Tuner,High-Performance	1046.380952	21
Hatchback,Performance	1039.646825	252
Crossover,Luxury,High-Performance	1037.222222	9
Hatchback,Factory Tuner,Luxury,Performance	886.8888889	9
Crossover,Luxury	884.5487805	410
Flex Fuel,Luxury,High-Performance	878.9090909	33
Hatchback,Diesel	873	14
Crossover,Diesel	873	7
Flex Fuel,Luxury	746.5384615	39
Luxury,Hybrid	724.6875	48
Crossover,Luxury,Hybrid	630.9166667	24
Factory Tuner,Luxury	617	2
Luxury,High-Performance,Hybrid	568.8333333	12
Exotic,Flex Fuel,Factory Tuner,Luxury,High-Performance	520	13
Exotic,Flex Fuel,Luxury,High-Performance	520	11
Exotic,Factory Tuner,Luxury,Performance	520	3
Exotic,Factory Tuner,Luxury,High-Performance	517.5384615	52
Exotic,Luxury,High-Performance	467.0759494	79
Hatchback,Luxury,Hybrid	454	3
Flex Fuel,Factory Tuner,Luxury,High-Performance	258	1
Crossover,Exotic,Luxury,High-Performance	238	1
Crossover,Exotic,Luxury,Performance	238	1
Exotic,Luxury,Performance	217.0277778	36
Crossover,Factory Tuner,Performance	210	4
Crossover,Hatchback,Luxury	204	7
Exotic,Luxury,High-Performance,Hybrid	204	1
Flex Fuel,Hybrid	155	2
Performance,Hybrid	155	1
Flex Fuel,Performance,Hybrid	155	2
Exotic,Luxury	112.6666667	12

The above table is being created by dragging the "Market Category" column to the "Row Labels" section and dragging the "Popularity Score" and "Model" column to the "Values" section. Also, later changing the field settings of values of popularity score from SUM to AVERAGE and Model from SUM to COUNT.

- **Task: 1.B:** To create a combo chart, I'll be selecting the Market categories and Average of popularity as the initial data and the chart as combo chart. Then, I'll be selecting Count of Model column that'll act as the line to the average of popularity.
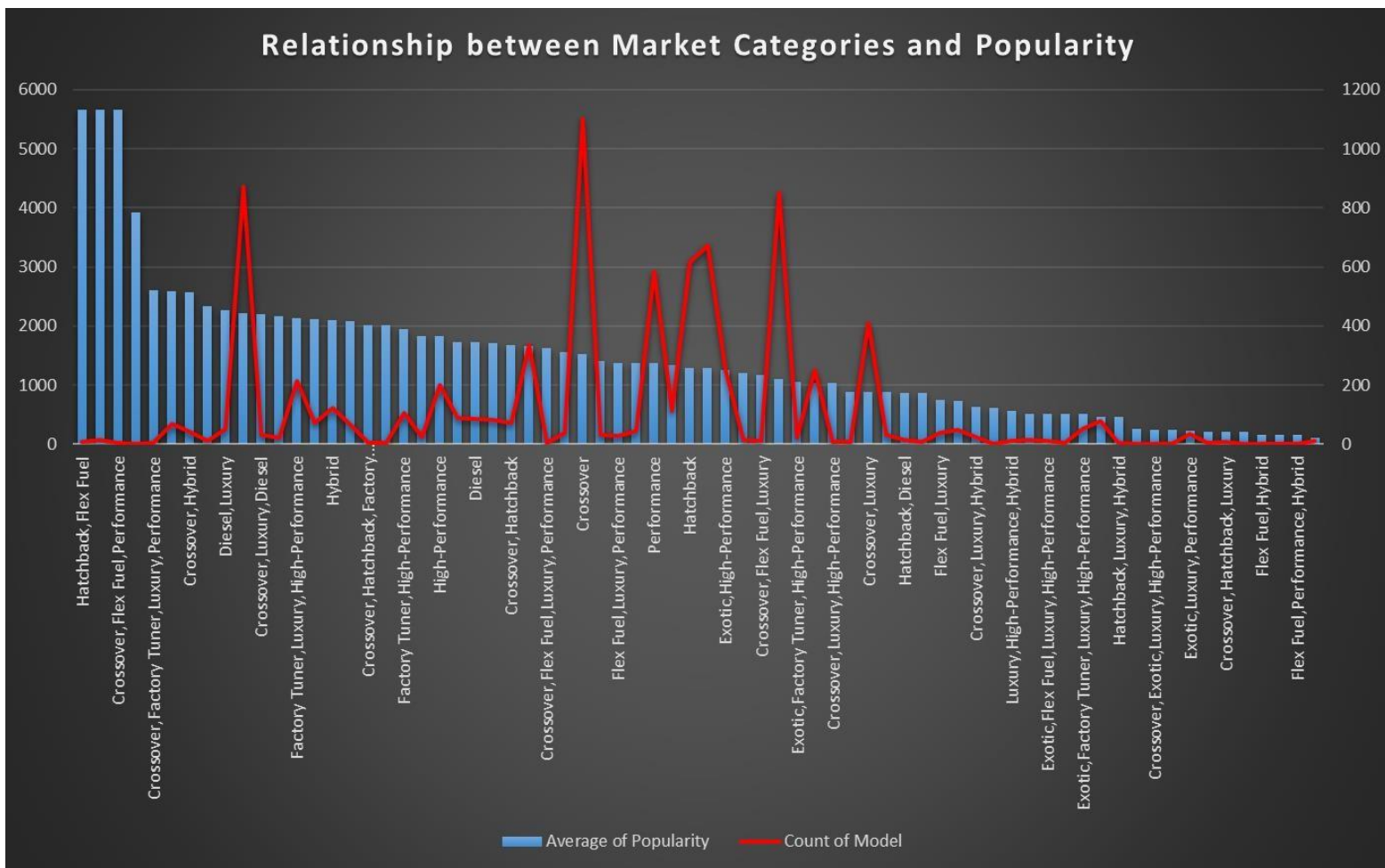

The combo chart will display the market categories along the x-axis, and the popularity values will be represented along the y-axis. This visualization will provide a clear and concise representation of the relationship between market categories and their corresponding popularity scores, allowing for easy interpretation and analysis of the data.

Insight Required: How does the popularity of a car model vary across different market categories?

Result: The popularity of a car model can vary across different market categories based on various factors. Generally, the popularity of a car model may be influenced by factors such as consumer preferences, market demand, price range, and target audience within each market category. Such insights can guide manufacturers and marketers in developing and positioning their car models to effectively cater to the specific preferences and demands of each market segment.

- **Task 2:** To create a scatter chart that plots engine power on the x-axis and price on the y-axis while adding a trendline
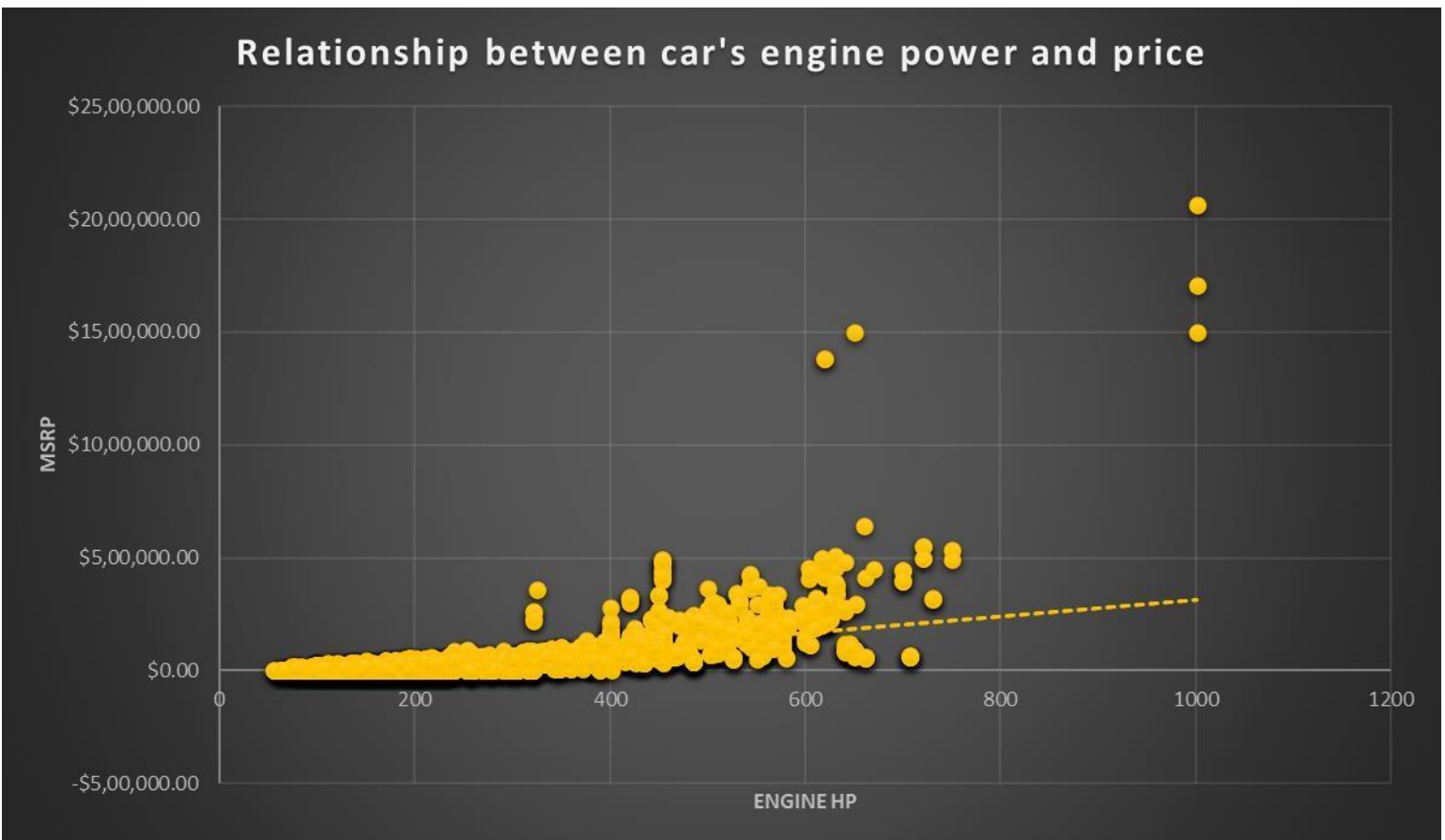
To create the above plot, I created a new sheet and added a scatter plot from the charts menu. Then, I added the data according to x-axis and y-axis. Also, I added a trend line from the option to depict the relationship between the data.

The scatter chart will display data points with their engine power values on the x-axis and corresponding prices on the y-axis. The trendline will provide a visual representation of the overall relationship between engine power and price, helping to identify any patterns or correlations. This visualization enables a quick understanding of how changes in engine power might relate to price variations in the car models being analyzed.

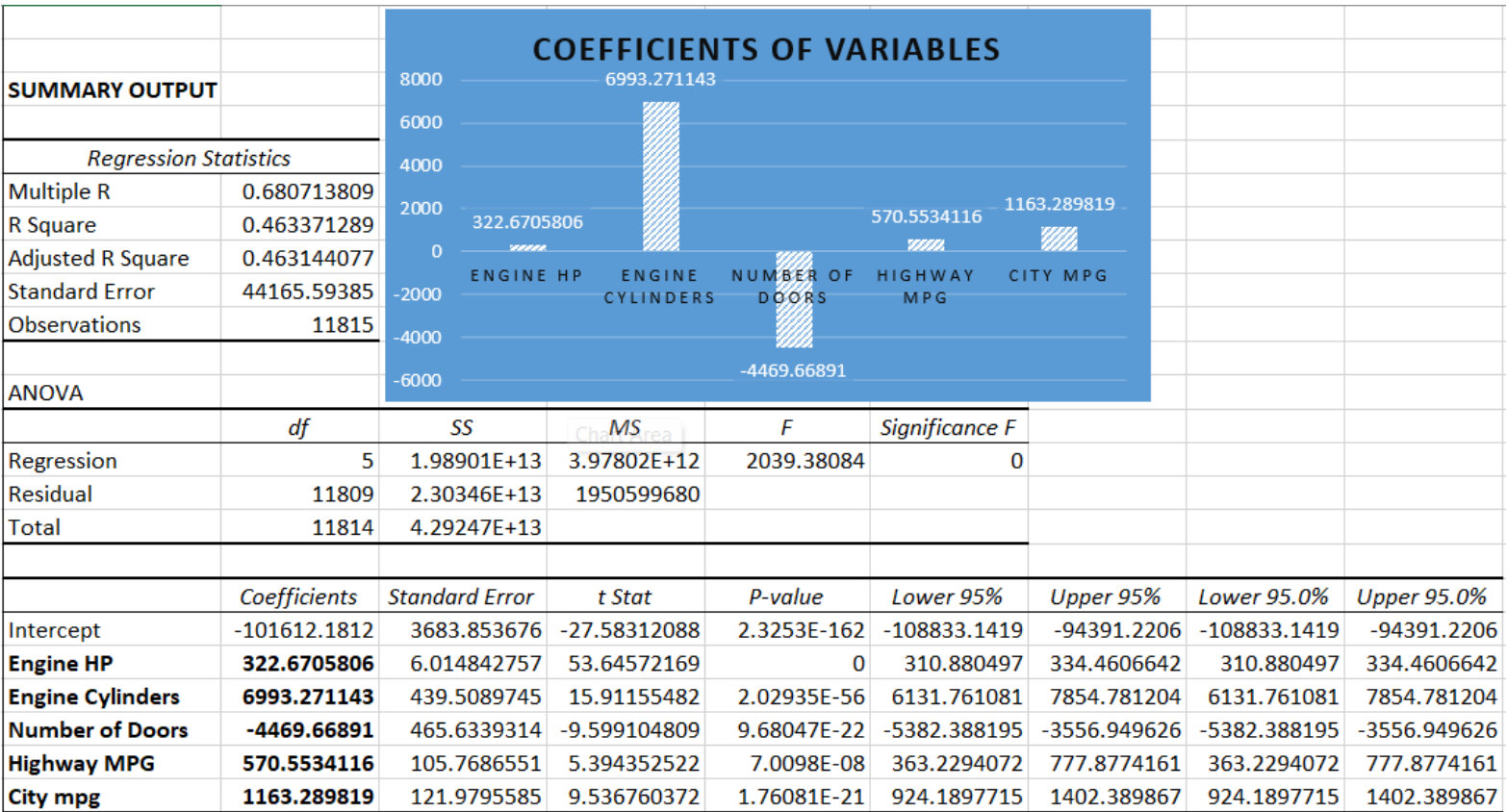**Insight Required:** What is the relationship between a car's engine power and its price?

Result: The relationship between a car's engine power and its price can vary depending on several factors. In general, there is often a positive correlation between engine power and price, meaning that as engine power increases, the price of the car tends to increase as well. This is because higher engine power often

requires more advanced technology, engineering, and components, which can drive up production costs and, subsequently, the price of the vehicle.

- **Task 3:** To identify the variables that have the strongest relationship with a car's price, regression analysis can be employed. This statistical technique analyzes the relationship between a dependent variable and independent variables. By conducting regression analysis, the coefficient values for each variable can be obtained, indicating the strength and direction of their impact on the car's price.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.680713809
R Square	0.463371289
Adjusted R Square	0.463144077
Standard Error	44165.59385
Observations	11815

COEFFICIENTS OF VARIABLES

ANOVA

	df	SS	MS	F	Significance F
Regression	5	1.98901E+13	3.97802E+12	2039.38084	0
Residual	11809	2.30346E+13	1950599680		
Total	11814	4.29247E+13			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-101612.1812	3683.853676	-27.58312088	2.3253E-162	-108833.1419	-94391.2206	-108833.1419	-94391.2206
Engine HP	322.6705806	6.014842757	53.64572169	0	310.880497	334.4606642	310.880497	334.4606642
Engine Cylinders	6993.271143	439.5089745	15.91155482	2.02935E-56	6131.761081	7854.781204	6131.761081	7854.781204
Number of Doors	-4469.66891	465.6339314	-9.599104809	9.68047E-22	-5382.388195	-3556.949626	-5382.388195	-3556.949626
Highway MPG	570.5534116	105.7686551	5.394352522	7.0098E-08	363.2294072	777.8774161	363.2294072	777.8774161
City mpg	1163.289819	121.9795585	9.536760372	1.76081E-21	924.1897715	1402.389867	924.1897715	1402.389867

For the process of regression analysis, I copied and pasted the 6 important columns that determines the price of a car i.e., Engine HP, Engine Cylinder, Number of Doors, Highway MPG, City MPG & MSRP to a new sheet. Then I selected Regression from Data Analysis tab. Upon revelation I added MSRP to Y-axis and the rest of the columns to X-axis.

Once the regression analysis is performed and the coefficient values are determined, a bar chart can be created to visualize the relative importance of each variable. In the bar chart, each variable will be represented by a bar, and the height of the bar will correspond to its coefficient value. This visual representation allows for a quick

and intuitive understanding of which variables have the most significant influence on the car's price, based on the magnitude of their coefficient values.

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

Result: From the above analysis, I can be seen that Engine Cylinders followed by City MPG and Highway MPG are the most important points that determine the price of a car.

- ▪ **Task 4.A:** A pivot table was generated to calculate and display the average price of cars for each manufacturer.

Manufacturer	Average Price
Bugatti	1757223.667
Maybach	546221.875
Rolls-Royce	351130.6452
Lamborghini	331567.3077
Bentley	247169.3243
McLaren	239805
Ferrari	237383.8235
Spyker	213323.3333
Aston Martin	197910.3763
Maserati	114207.7069
Porsche	101622.3971
Mercedes-Benz	71537.80966
Lotus	69188.27586
Land Rover	67823.21678
Alfa Romeo	61600
BMW	61546.76347
Cadillac	56231.31738
Audi	53452.1128
Lexus	47549.06931
Genesis	46616.66667
Lincoln	42494.37179
Infiniti	42394.21212
HUMMER	36464.41176
Acura	34887.5873
GMC	30493.29903
Toyota	28946.15343
Volvo	28541.16014
Nissan	28513.36679
Chevrolet	28273.35695
Buick	28206.61224
Volkswagen	28076.2
Saab	27413.5045
Ford	27393.42051
Chrysler	26722.96257
Honda	26629.81879
Kia	25112.38938
Subaru	24827.50391
Hyundai	24597.0363
Dodge	22390.05911
FIAT	22206.01695
Mitsubishi	21215.47143
Scion	19932.5
Mazda	19719.05707
Pontiac	19321.54839
Suzuki	17907.20798
Oldsmobile	11542.54
Plymouth	3122.902439

The table presents the manufacturer names in rows and the corresponding average prices in the cells. This analysis provides a clear overview of the pricing patterns across different car manufacturers, allowing for easy comparison of their average prices. By utilizing this pivot table, manufacturers can gain valuable insights into their

pricing strategies and make informed decisions regarding pricing adjustments, market positioning, and competitive analysis.

- ▪ Task 4.B: To create a bar chart or a horizontal stacked bar chart that visualizes the relationship between manufacturer and average price.

AVERAGE PRICE OF CARS ACROSS DIFFERENT MANUFACTURERS
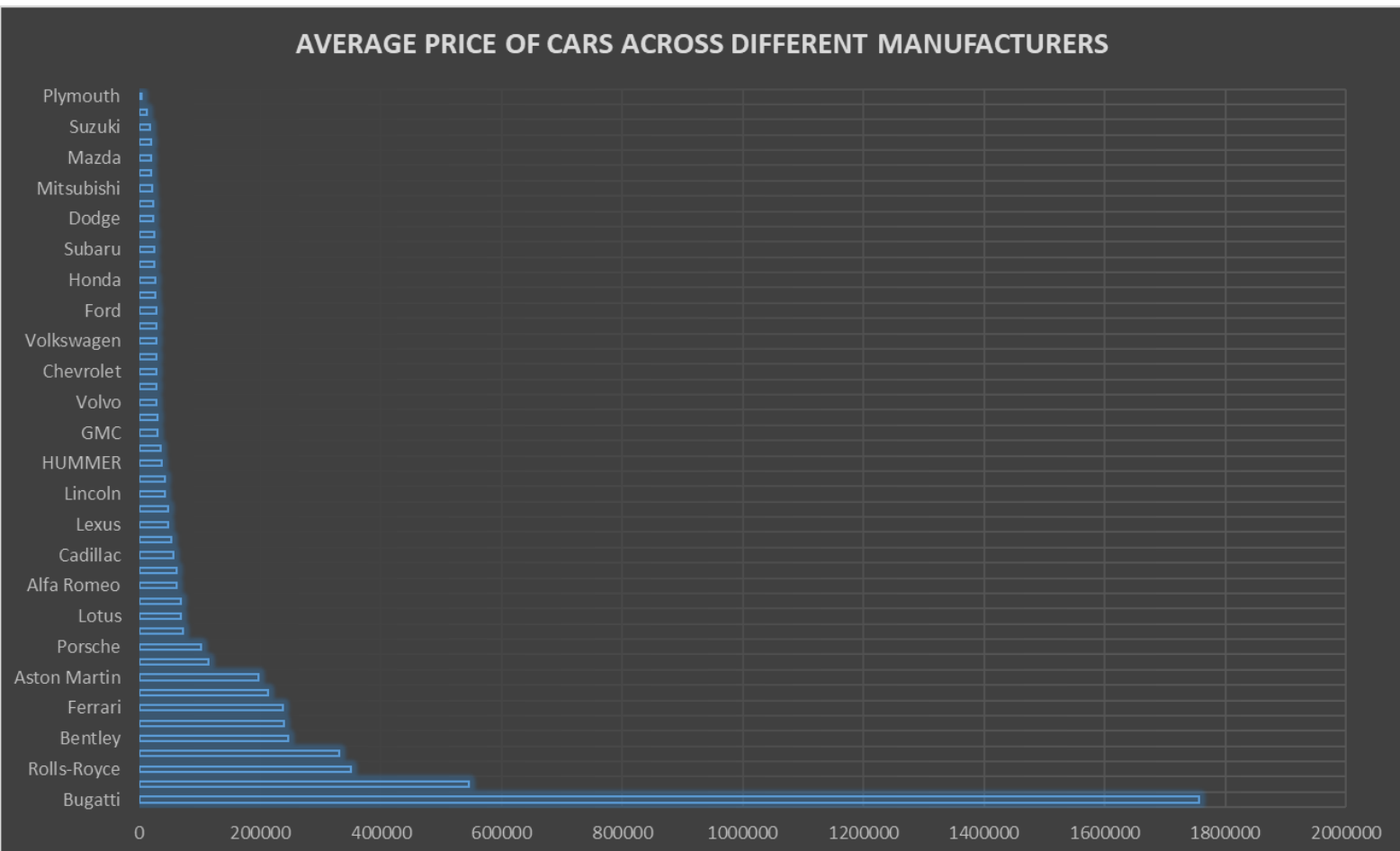
A bar chart was created to visually represent the relationship between manufacturers and their average prices. In this chart, each manufacturer is represented by a bar, and the height or width of the bar corresponds to the average price of cars associated with that manufacturer. This visualization allows for easy comparison of average prices across different manufacturers, providing insights into the pricing variations within the automotive industry.

This graphical representation facilitates data-driven decision-making regarding pricing strategies, market positioning, and competitive analysis, aiding in effective pricing and business planning.

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

Result: The average price of a car can vary significantly across different manufacturers. Each manufacturer has its own pricing strategy, target market, brand positioning, and product offerings, which contribute to the variations in average prices.

Luxury and high-end manufacturers often have higher average prices due to their focus on premium features, advanced technology, and superior craftsmanship.
These brands cater to consumers seeking exclusivity, prestige, and top-tier performance, which comes with a higher price tag.

On the other hand, mainstream manufacturers typically offer a wider range of vehicles at various price points to target different segments of the market. Their average prices can vary depending on factors such as vehicle size, model type, features, and market competition. Economy or budget-oriented brands often have lower average prices as they prioritize affordability and cost-effectiveness.

- ▪ **Task 5.A:** To create a scatter plot with the number of cylinders on the x-axis and highway MPG on the y-axis. Also, provide a trendline on the scatter plot to visually estimate the slope of the relationship and assess its significance.
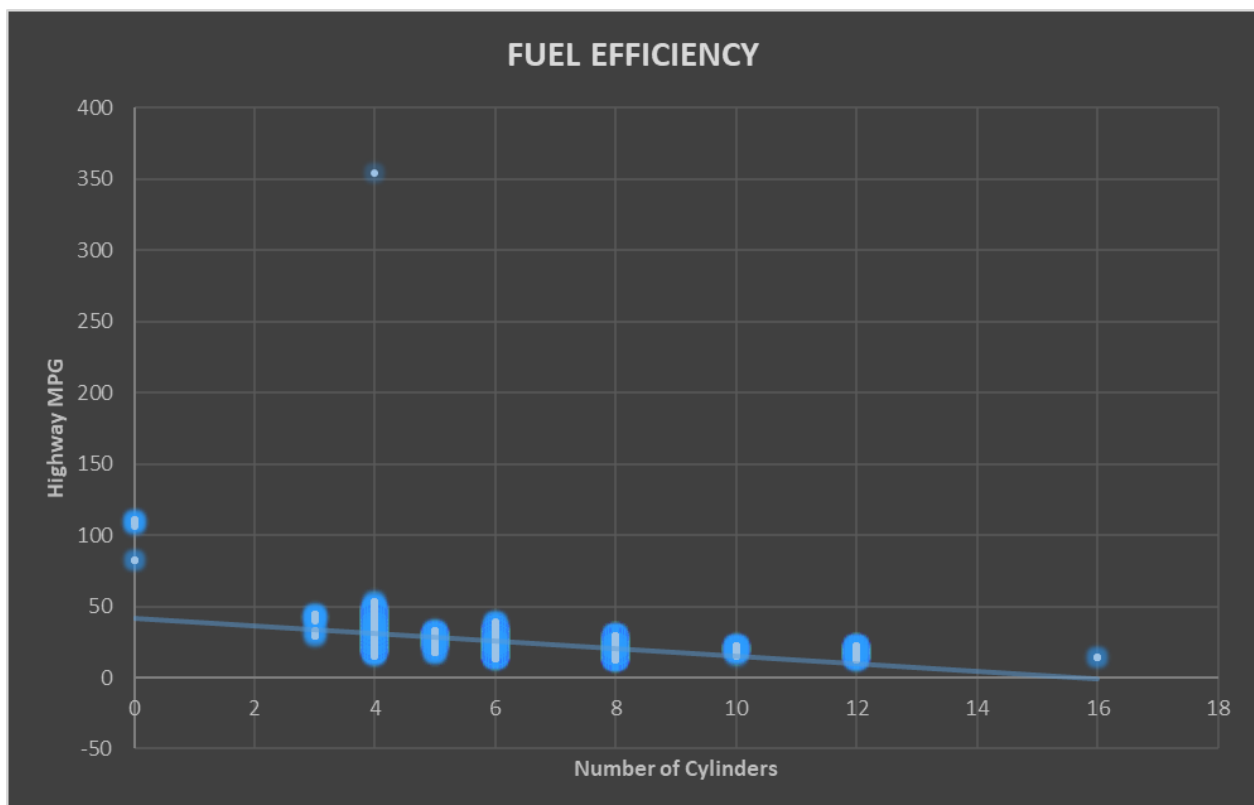
The above plot is being created by assigning the "Number of Cylinders" as independent variable (x-axis) and "Highway MPG" as dependent variable (y-axis). Once the scatter plot is created, a trendline can be added to visually estimate the slope of the relationship between the variables.

- **Task 5.B:** To calculate the correlation coefficient between the number of cylinders and highway MPG to quantify the strength and direction of the relationship. The result will be a single value between -1 and 1, representing the strength and direction of the linear relationship between the two variables. A positive value indicates a positive correlation, a negative value indicates a negative correlation, and a value close to zero indicates little to no correlation.

Correlation Coefficient	-0.62032	

The correlation coefficient is being calculated by using the following formula
=CORREL(A2:A11816, B2:B11816)

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

Result: The relationship between the number of cylinders in a car's engine and fuel efficiency, which is frequently expressed in terms of miles per gallon or MPG, can be complicated and influenced by a variety of factors. Moreover, a higher cylinder count typically leads to a worse fuel efficiency. This is true because more cylinders often result in greater power and performance, but they also consume more gasoline to run. Adding more cylinders frequently results in higher fuel usage, which lowers a vehicle's overall MPG.

In comparison to larger engines with more cylinders, such 6-cylinder or 8-cylinder engines, smaller engines with fewer cylinders, like 4-cylinder engines, frequently have a higher fuel efficiency. This is so because smaller engines use less fuel to produce the same amount of power due to their reduced displacement.

The next portion of the project demands creating an interactive dashboard using necessary filters and slicers.

▪ **Task 6.1:** How does the distribution of car prices vary by brand and body style?

Sum of MSRP	Body Style															
Row Labels	2dr Hatchback	2dr SUV	4dr Hatchback	4dr SUV	Cargo Minivan	Cargo Van	Convertible	Convertible SUV	Coupe	Crew Cab Pickup	Extended Cab Pickup	Passenger Minivan	Passenger Van	Regular Cab Pickup	Sedan	Wagon	Grand Total
Chevrolet	$8,000.00	$2,13,310.00	$12,09,735.00	$65,69,568.00	$4,20,150.00	$78,688.00	$29,53,245.00	$1,06,300.00	$35,04,525.00	$59,27,617.00	$31,17,951.00	$11,78,515.00	$6,07,670.00	$22,60,032.00	$30,68,812.00	$3,00,675.00	$3,15,24,793.00
Mercedes-Benz		$1,22,800.00	$49,24,810.00	$28,950.00			$57,53,964.00		$64,73,107.00			$32,500.00			$70,80,243.00	$7,64,935.00	$2,51,81,309.00
Ford	$36,000.00	$4,79,873.00	$4,80,155.00	$43,70,871.00	$6,80,770.00	$5,66,351.00	$7,30,007.00		$13,98,144.00	$38,12,353.00	$22,85,584.00	$12,71,330.00	$24,31,898.00	$12,99,240.00	$22,99,348.00	$16,35,565.00	$2,37,77,489.00
Volkswagen	$41,71,275.00		$32,22,275.00	$20,84,955.00			$36,12,631.00		$8,000.00			$10,38,130.00			$67,60,050.00	$17,04,025.00	$2,26,01,341.00
Cadillac				$71,82,555.00			$9,85,607.00		$29,53,574.00	$5,99,150.00					$94,18,847.00	$11,84,100.00	$2,23,23,833.00
Toyota	$4,73,750.00		$13,97,750.00	$49,57,050.00			$3,86,668.00		$8,11,995.00	$38,93,760.00	$35,58,504.00	$19,56,518.00		$3,73,446.00	$24,59,596.00	$12,37,955.00	$2,15,06,992.00
BMW	$80,097.00		$11,44,950.00	$31,60,950.00			$45,02,671.00		$34,19,051.00						$79,89,300.00	$2,59,600.00	$2,05,56,619.00
Aston Martin							$73,21,655.00		$96,35,275.00						$14,48,735.00		$1,84,05,665.00
Bentley							$60,12,870.00		$63,56,760.00						$59,20,900.00		$1,82,90,530.00
Audi	$4,000.00			$26,74,900.00			$32,91,405.00		$35,56,290.00						$71,58,348.00	$8,47,350.00	$1,75,32,293.00
Lamborghini							$70,64,450.00		$1,01,77,050.00								$1,72,41,500.00
Ferrari							$47,23,811.00		$1,14,18,289.00								$1,61,42,100.00
GMC		$1,44,319.00		$66,41,919.00	$1,42,750.00	$4,68,085.00				$40,62,482.00	$21,83,866.00	$1,50,630.00	$6,03,670.00	$13,06,328.00			$1,57,04,049.00
Nissan	$14,683.00		$10,23,090.00	$41,49,630.00	$1,28,620.00		$14,06,552.00	$1,31,075.00	$29,43,632.00	$24,22,300.00	$10,26,379.00	$4,13,320.00		$21,914.00	$17,69,130.00	$1,75,000.00	$1,56,25,325.00
Dodge	$48,000.00	$44,000.00	$18,000.00	$25,72,405.00	$60,520.00	$3,38,497.00	$12,000.00		$32,64,627.00	$22,35,775.00	$8,64,172.00	$5,57,425.00	$70,708.00	$7,19,408.00	$24,17,585.00	$7,93,055.00	$1,40,16,177.00
Infiniti				$43,40,200.00			$9,80,050.00		$21,75,750.00						$64,94,090.00		$1,39,90,090.00
Porsche	$28,827.00			$18,15,200.00			$45,04,586.00		$47,58,533.00						$27,13,500.00		$1,38,20,646.00
Honda	$4,13,200.00		$20,15,270.00	$39,53,209.00			$2,52,135.00		$15,88,705.00	$7,87,720.00		$5,53,185.00			$23,40,105.00		$1,19,03,529.00
Rolls-Royce							$21,41,365.00		$22,04,675.00						$65,39,010.00		$1,08,85,050.00
Land Rover		$4,76,394.00		$90,76,595.00				$1,45,731.00									$96,98,720.00
Lexus			$94,700.00	$31,52,974.00			$4,72,065.00		$10,16,472.00						$48,37,596.00	$31,105.00	$96,04,912.00

(Only for Visual Representation. For full table data please refer to the excel file)
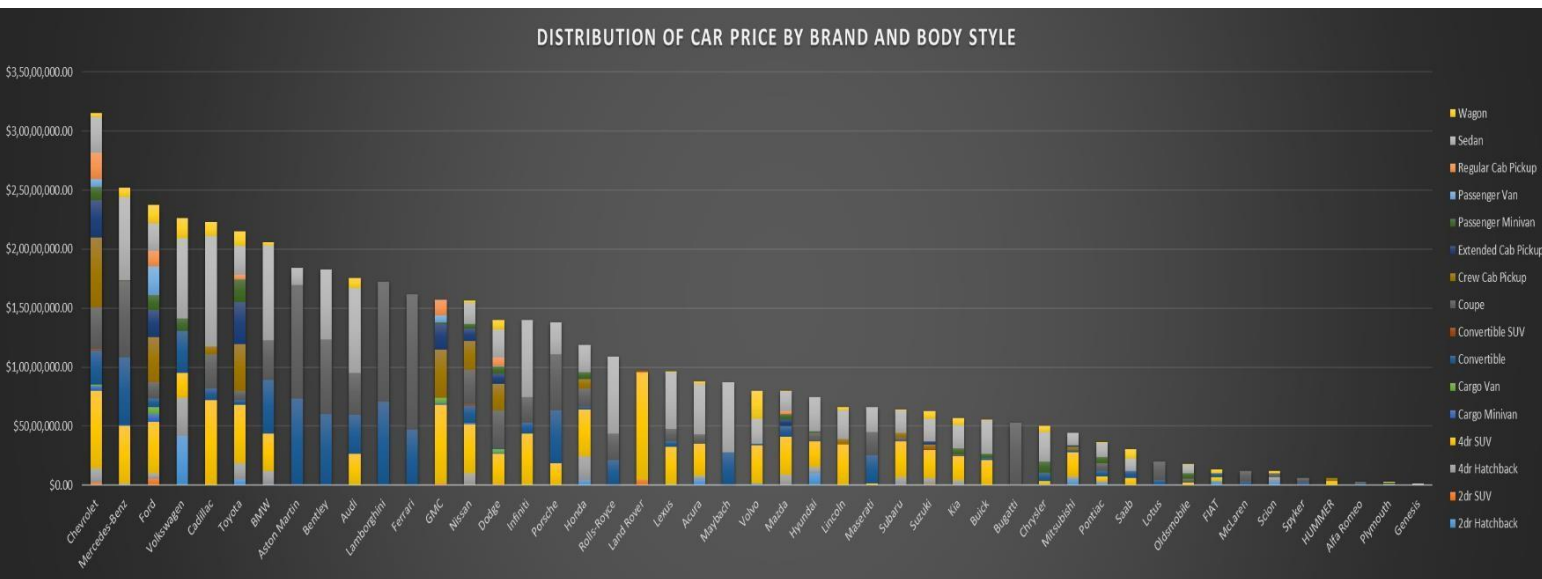
Result: Due to a number of factors, including the brand's location in the market, target market, features and specifications given, perceived quality, and market demand, the distribution of automobile pricing might differ by brand and body design. Here is a general description of how the distribution of automobile pricing by brand and body design could change depending on the market and trends today:

✓ The target audiences and brand positioning of various auto brands can affect their pricing tactics. High-end or luxury brands often come with premium features, cutting-edge technology, and

better materials, commanding higher price points. On the other hand, cost-effective or economy brands could emphasise affordability and provide more affordable solutions.

✓ Pricing may vary depending on the body type of the vehicle, including sedans, SUVs, coupes, hatchbacks, and convertibles. Due to their increased size, extra equipment, and perceived prestige, larger and more spacious body types, such as SUVs and luxury sedans, typically have higher price ranges. Price ranges are frequently lower for smaller, more compact body types like hatchbacks or entry-level sedans.

✓ Pricing is heavily influenced by perceptions of a brand and its reputation. Due to the perceived value connected with the brand, well-known brands with a reputation for quality, dependability, and prestige frequently attract higher costs.

▪ **Task 6.2:** Which car brands have the highest and lowest average MSRPs, and how does this vary by body style?

	AVERAGE OF CAR PRICE BY BRAND AND BODY STYLE															
Average of MSRP	Body Style															
Row Labels	2dr Hatchback	2dr SUV	4dr Hatchback	4dr SUV	Cargo Minivan	Cargo Van	Convertible	Convertible SUV	Coupe	Crew Cab Pickup	Extended Cab Pickup	Passenger Minivan	Passenger Van	Regular Cab Pickup	Sedan	Wagon	Grand Total
Bugatti									$17,57,223.67								$17,57,223.67
Maybach							$13,81,375.00								$4,26,914.29		$5,46,221.88
Rolls-Royce							$4,28,273.00		$3,67,445.83						$3,26,950.50		$3,51,130.65
Lamborghini							$3,36,402.38		$3,28,291.94								$3,31,567.31
Bentley							$2,50,536.25		$2,54,270.40						$2,36,836.00		$2,47,169.32
McLaren							$2,80,225.00		$2,29,700.00								$2,39,805.00
Ferrari							$2,14,718.68		$2,48,223.67								$2,37,383.82
Spyker							$2,19,990.00		$2,09,990.00								$2,13,323.33
Aston Martin							$2,03,379.31		$1,92,705.50						$2,06,962.14		$1,97,910.38
Maserati			$77,500.00				$1,30,164.61		$1,16,016.71						$1,02,561.90		$1,14,207.71
Porsche	$5,765.40		$82,509.09				$1,15,502.21		$99,136.10						$1,23,340.91		$1,01,622.40
Mercedes-Benz			$40,933.33	$68,400.14	$28,950.00		$1,04,617.53		$1,09,713.68			$32,500.00			$49,168.35	$44,996.18	$71,537.81
Lotus							$51,657.50		$75,866.67								$69,188.28
Land Rover		$39,699.50		$70,910.90				$48,577.00									$67,823.22
Alfa Romeo							$64,900.00		$59,400.00								$61,600.00
BMW	$26,699.00		$54,521.43	$58,536.11			$63,417.90		$51,803.80						$70,701.77	$43,266.67	$61,546.76
Cadillac				$72,551.06			$70,400.50		$45,439.60	$66,572.22					$50,912.69	$47,364.00	$56,231.32
Audi	$2,000.00			$48,634.55			$70,029.89		$93,586.58						$44,461.79	$33,894.00	$53,452.11
Lexus			$31,566.67	$45,042.49			$52,451.67		$50,823.60						$48,864.61	$31,105.00	$47,549.07
Genesis															$46,616.67		$46,616.67
Lincoln				$50,331.91					$2,111.83	$41,205.45					$41,665.17	$44,950.83	$42,494.37
Infiniti				$45,686.32			$46,669.05		$40,291.67						$40,588.06		$42,394.21
HUMMER				$37,749.00							$34,629.29						$36,464.41
Acura	$17,175.61		$51,062.86	$42,959.76					$39,687.40						$33,292.26	$33,560.00	$34,887.59
GMC		$5,550.73		$36,695.69	$23,791.67	$18,723.40				$39,062.33	$26,632.51	$25,105.00	$26,246.52	$21,069.81			$30,493.30

(Only for Visual Representation. For full table data please refer to the excel file)

AVERAGE PRICE OF CAR BY BRAND AND BODY STYLE
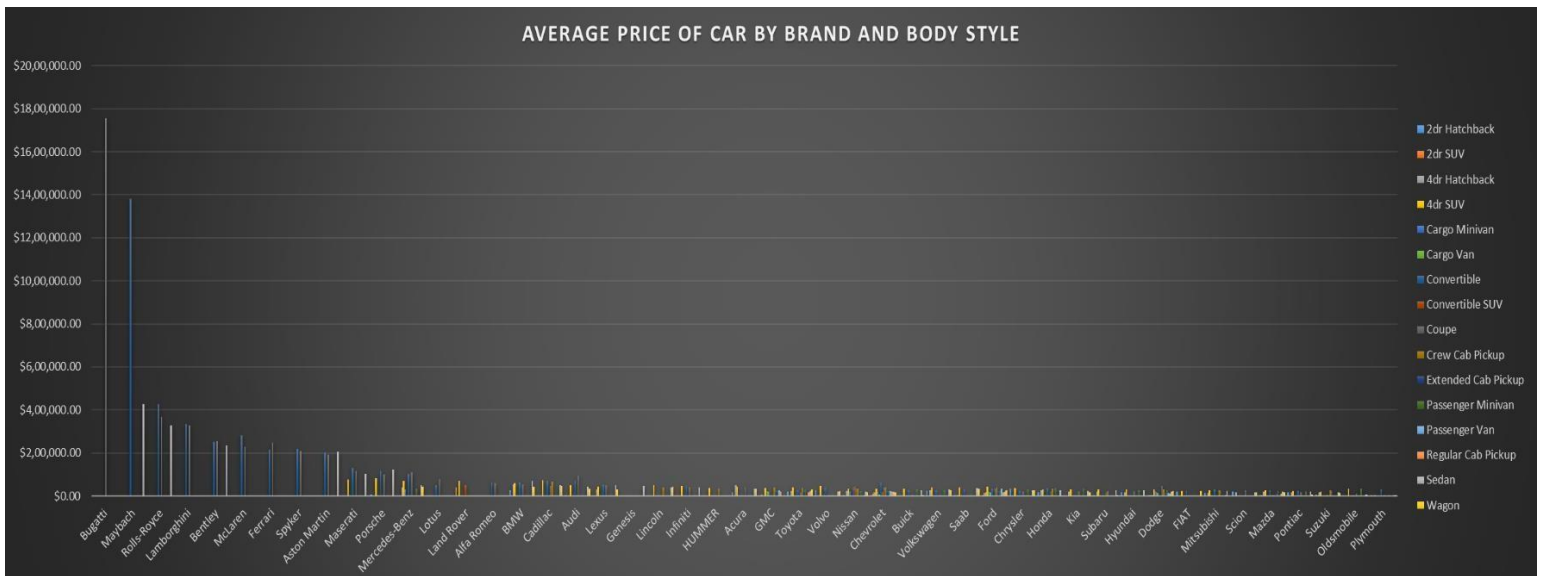
Result: Luxury brands like Bugatti and Maybach often have the highest average MSRPs for automobiles. These companies are renowned for producing high-end, upscale cars that come equipped with top-notch materials, cutting-edge technology, and opulent amenities. Their automobiles' typical MSRPs might be in the hundreds of thousands or even millions of dollars.

Conversely, the average MSRP of popular and budget brands is typically lower. Brands often provide a selection of reasonably priced cars at reasonable prices. Depending on the model, trim level, and body type, the average MSRPs for the vehicles from these manufacturers might vary greatly, but they are often more affordable for a wider spectrum of buyers. Plymouth on the other hand have the lowest MSRP.

It's vital to keep in mind that certain models and trims within a brand can greatly affect the average price when discussing the range in typical MSRPs per body shape. However, the following broad tendencies may be seen:
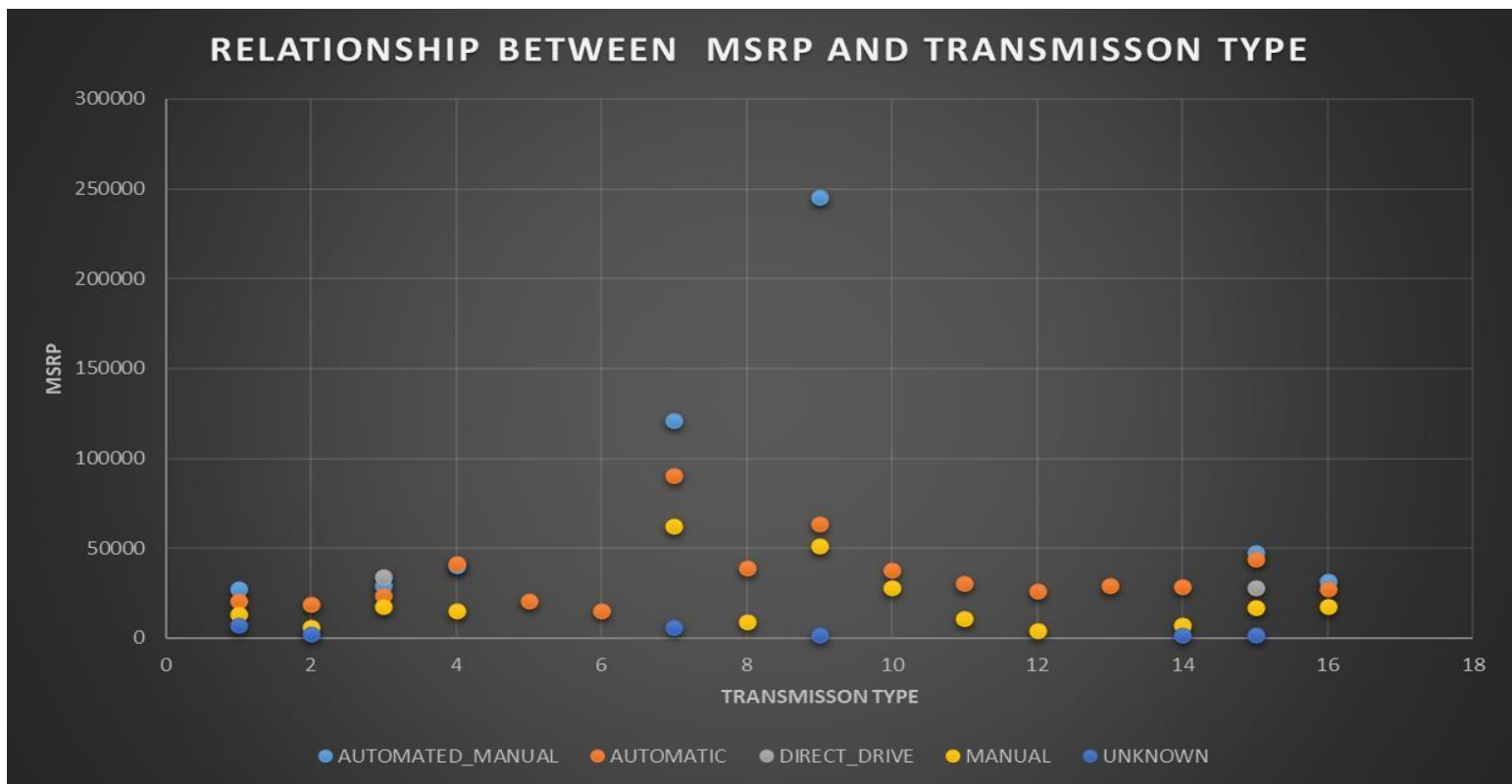
- ✓ The average MSRP for sedan and SUV models from luxury companies is often expensive. These cars frequently cost more because they have more expensive materials, cutting-edge technology, and powerful engines.
- ✓ With reasonable average MSRPs, mainstream companies provide a variety of compact and midsize sedans and hatchbacks. These cars are intended to be practical, economical, and economical solutions for daily travel.

▪ **Task 6.3:** How do the different feature such as transmission type affect the MSRP, and how does this vary by body style?

Average of MSRP	Transmisson Type					
Body Style	AUTOMATED_MANUAL	AUTOMATIC	DIRECT_DRIVE	MANUAL	UNKNOWN	Grand Total
2dr Hatchback	27180.96491	20926.464		13353.658	7361.5	16778.65408
2dr SUV		18615.20455		6303.8111	2371	10115.18841
4dr Hatchback	29249.07407	23833.67898	34511.92308	17594.413		22086.30236
4dr SUV	40451.15385	41555.18825		15426.462		40426.82137
Cargo Minivan		20910.85714				20910.85714
Cargo Van		15280.22105				15280.22105
Convertible	121256.6444	90637.3869		62357.756	5783.5	84224.28499
Convertible SUV		38925.5		9233.1429		17424.13793
Coupe	245588.3571	63852.00808		51070.48	2000	76900.70504
Crew Cab Pickup		37744.07154		28360.526		37220.46696
Extended Cab Pickup		30637.34973		10884.195		22488.77689
Passenger Minivan		26391.99748		4405.3333		25591.51214
Passenger Van		29015.20313				29015.20313
Regular Cab Pickup		28536.8239		7557.7733	2000	15953.70918
Sedan	47498.70813	43760.61208	27822.5	17119.234	2000	38969.06068
Wagon	31985.27778	27613.19169		17844.14		25483.90119
Grand Total	99195.584	41129.06449	33620	26671.397	3040.736842	40554.36758

For creating the table, I copied and pasted the three columns necessary for the task, i.e., Transmission Type, MSRP and Vehicle Style. Next, I created a pivot table to get the data about average value. Then, I copied and pasted the average data to a new

sheet. And later from that data I created a scatter plot that shows the relationship between transmission type and MSRP.


RELATIONSHIP BETWEEN MSRP AND TRANSMISSON TYPE

Result: The different features of a car, such as the transmission type, can have an impact on the Manufacturer's Suggested Retail Price (MSRP), and this can vary based on the body style of the vehicle. The cost of automated transmissions is typically higher than that of manual transmissions. When opposed to manual transmissions, automatic transmissions provide comfort and ease of use, but their cost can be higher overall.

The impact of transmission type on MSRP can vary depending on the body style of the vehicle. The effect of transmission type on MSRP can be particularly pronounced in the sedan sector. For sedans, especially those with higher trim levels, automatic transmissions are frequently the preferred option since they offer a more comfortable driving experience. As a result, sedans with automatic transmissions could have more expensive MSRPs than equivalent models with manual transmissions.

- **Task 6.4:** How does the fuel efficiency of cars vary across different body styles and model years?

Years	Average of highway MPG
⊞ 1990	23.1
⊞ 1991	22.2
⊞ 1992	24.1
⊞ 1993	24.2
⊞ 1994	23.9
⊞ 1995	23.2
⊞ 1996	23.7
⊞ 1997	22.3
⊞ 1998	21.9
⊞ 1999	23.0
⊞ 2000	24.0
⊞ 2001	23.7
⊞ 2002	22.8
⊞ 2003	22.7
⊞ 2004	23.1
⊞ 2005	23.6
⊞ 2006	23.4
⊞ 2007	21.9
⊞ 2008	23.0
⊞ 2009	23.9
⊞ 2010	24.2
⊞ 2011	25.2
⊞ 2012	26.3
⊞ 2013	27.4
⊞ 2014	26.9
⊞ 2015	28.5
⊞ 2016	28.5
⊞ 2017	28.3
Grand Total	26.3

For the table, I used pivot table to generate the data for the analysis. I assembled years Years and Vehicle style to Rows section and Highway MPG to Values section. The table represents the average mileage corresponding to the years from 1990 to 2017.

The line graph provides an insight for the fuel efficiency over each and every year. Also the trend line indicates the increment in efficiency.

FUEL EFFICIENCY OVERTIME
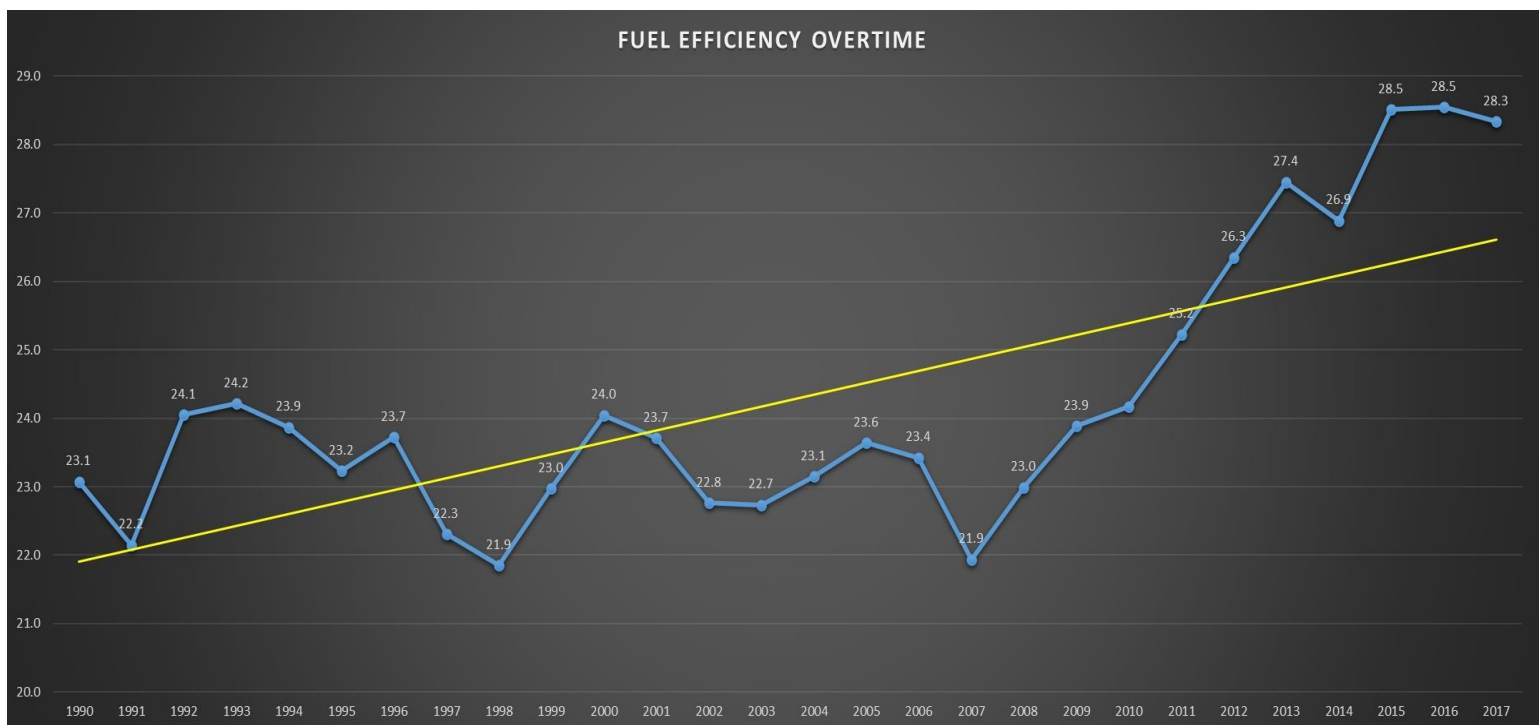
Average of highway MPG	Body Style																
AVERAGE HIGHWAY MPG FOR EACH COMBNATION OF BODY STYLE CORRESPONDING TO YEAR																	
Years	2dr Hatchback	2dr SUV	4dr Hatchback	4dr SUV	Cargo Minivan	Cargo Van	Convertible	Convertible SUV	Coupe	Crew Cab Pickup	Extended Cab Pickup	Passenger Minivan	Passenger Van	Regular Cab Pickup	Sedan	Wagon	Grand Total
1990	30.4	20.0	31.0		20.0		23.5		24.5			22.0	18.9	22.2	24.0	24.1	23.1
1991	30.1	16.3		19.3			22.6		26.2			15.8	18.0	17.0	24.2	22.6	22.2
1992	29.7	17.5	28.4	21.3			25.5		27.3				15.6	17.9	24.5	24.3	24.1
1993	28.5	18.5	27.3	21.0			24.5	26.0	28.5				16.7	17.6	25.3	24.5	24.2
1994	27.4	18.4	27.1	20.0	21.0	19.3	26.0	26.0	27.5		20.3	21.0	16.4	21.7	25.2	23.8	23.9
1995	30.1	16.0	27.7		21.5	19.0	24.5	26.0	25.8		20.0	20.1	15.0	21.2	24.1	24.1	23.2
1996	29.0	20.0	26.1	21.6	23.0	14.6	23.8	24.0	26.7		20.0	20.8	15.0	22.2	25.7	24.7	23.7
1997	26.1	22.0	26.5	19.7	21.0	17.1	25.3	20.7	27.2		18.4	20.6	17.0	18.8	25.3	24.4	22.3
1998	23.2	26.0	24.5	22.1		17.2	23.7	24.0	26.3		18.6	23.4	17.0	19.2	27.1	23.0	21.9
1999	30.3	18.8		18.3		16.7	21.5		27.6			18.4	22.3	18.4	27.4		23.0
2000	30.4	18.8		17.7		16.4	25.3		24.2		20.5	23.2	14.5	20.8	26.8	31.0	24.0
2001	29.0	18.7		18.7	22.0	15.8	23.4		20.3		19.0	21.2	15.0	23.0	27.4	30.6	23.7
2002	25.3	19.0		19.8	21.0	14.6	24.1	23.3	23.6	17.0	20.2	21.7	15.0	22.1	26.1	28.9	22.8
2003	29.8	18.8		19.2	20.7	15.0	20.2	23.4	23.9	18.0	20.8	22.3		24.1	27.1	24.0	22.7
2004	29.7	18.8	34.0	19.0	19.6		20.1		25.3	22.0	17.8	22.2	18.5		26.4	22.8	23.1
2005	30.3	18.7	30.6	19.3	20.7		20.7		26.0		23.0	22.0	18.0		25.8	24.3	23.6
2006	27.3		28.8	20.4	23.0		22.9		24.3	19.4		22.5		18.0	24.8	25.0	23.4
2007	25.5		27.5	20.5	22.7		22.8		25.2	18.0	18.4	22.8		19.6	25.3	24.8	21.9
2008	27.3		28.5	20.8	23.0		23.5		24.8	18.5	19.2	23.0		18.0	26.7	24.7	23.0
2009	29.0		30.8	22.6			23.8		24.0	19.1	19.9			21.9	26.6	26.8	23.9
2010	27.8		29.5	23.2			24.6		23.8	18.9	20.8	24.2		21.0	26.1	28.5	24.2
2011	27.8		28.9	23.6			23.9		22.7	21.1	21.9	25.0		27.0	27.0	28.7	25.2
2012	30.7		32.7	22.7		16.7	23.6	22.0	22.1	21.4	23.1	25.0	15.3	24.1	27.9	30.6	26.3
2013	31.9		32.3	23.6		16.7	23.2	22.0	25.5	21.3		28.0	15.3		29.8	29.6	27.4
2014	34.7		38.8	23.8		16.9	26.6	22.0	23.2	18.9	17.4	26.0	16.4		29.9	29.4	26.9
2015	34.3	30.0	38.5	25.8	28.0	17.0	27.6		26.2	22.0	21.7	25.8	18.1	22.7	32.2	30.9	28.5
2016	34.5	30.0	38.8	26.2	27.7	16.0	27.9			22.3	21.8	25.6	17.7	22.5	32.1	29.9	28.5
2017	33.1	29.0	38.0	25.7	26.7		27.8	28.0	27.7	21.8	21.0	26.1	19.0	22.5	32.6	30.9	28.3
Grand Total	30.9	19.1	35.3	24.4	24.5	16.6	25.8	23.7	25.8	21.1	20.1	23.6	17.2	20.6	29.8	27.8	26.3

Result: A car's fuel efficiency might range between body types and model years due to technological developments, modifications in the way automobiles are designed, and changing consumer tastes. Different body styles have varying aerodynamics, weights, and sizes, which can impact fuel efficiency. Also, because of developments in engine technology, materials, and vehicle design, fuel economy has increased over time. More fuel-efficient engines have been developed by automakers to improve fuel economy by streamlining power delivery and consuming less gasoline. To lighten vehicles, automakers have started using lightweight materials like aluminium and carbon fibre. Vehicles that are lighter require less energy to move, which increases fuel efficiency.

As manufacturers improve their designs and use new technology, each model year may see slight increases in fuel economy. Newer models frequently gain from enhanced aerodynamics, powertrain combinations, and engine economy.

- **Task 6.5:** How does the car's horsepower, MPG, and price vary across different Brands?

For the table creation, I selected the main data and opted for a pivot table. Then, I assembled Engine HP, Highway MPG and MSRP to values section and converted the field to average. Then, I assembled Make column to Rows section.

AVERAGE OF ENGINE HP, HIGHWAY MPG & MSRP

Brand	Average of Engine HP	Average of highway MPG	Average of MSRP
Acura	244.8	28.1	34887.6
Alfa Romeo	237.0	34.0	61600.0
Aston Martin	484.3	18.9	197910.4
Audi	277.7	28.8	53452.1
Bentley	533.9	18.9	247169.3
BMW	326.9	29.2	61546.8
Bugatti	1001.0	14.0	1757223.7
Buick	219.2	26.9	28206.6
Cadillac	332.3	25.2	56231.3
Chevrolet	247.1	25.7	28273.4
Chrysler	229.1	26.4	26723.0
Dodge	244.4	22.3	22390.1
Ferrari	509.9	15.7	237383.8
FIAT	143.6	33.9	22206.0
Ford	243.1	23.7	27393.4
Genesis	347.3	25.3	46616.7
GMC	259.8	21.4	30493.3
Honda	195.7	32.3	26629.8
HUMMER	261.2	17.3	36464.4
Hyundai	201.9	30.4	24597.0
Infiniti	310.1	24.8	42394.2
Kia	206.8	29.3	25112.4
Lamborghini	614.1	18.0	331567.3
Land Rover	322.1	22.1	67823.2
Lexus	277.4	25.9	47549.1
Lincoln	284.9	24.1	42494.4
Lotus	276.0	26.6	69188.3
Maserati	420.8	20.3	114207.7

Maybach	590.5	16.0	546221.9
Mazda	169.2	28.1	19719.1
McLaren	610.4	22.2	239805.0
Mercedes-Benz	350.2	24.8	71537.8
Mitsubishi	174.5	26.5	21215.5
Nissan	239.9	26.5	28513.4
Oldsmobile	177.5	26.2	11542.5
Plymouth	131.6	28.0	3122.9
Pontiac	190.3	27.1	19321.5
Porsche	392.8	25.4	101622.4
Rolls-Royce	487.5	19.1	351130.6
Saab	220.5	26.4	27413.5
Scion	154.4	32.3	19932.5
Spyker	400.0	18.0	213323.3
Subaru	197.3	28.7	24827.5
Suzuki	160.3	26.0	17907.2
Toyota	236.3	26.3	28946.2
Volkswagen	190.1	31.8	28076.2
Volvo	231.0	27.2	28541.2
Grand Total	**249.5**	**26.3**	**40554.4**

For the chart preparation, I copied the data from pivot table and pasted it in a new worksheet. Then, I opted for bubble chart from chats tab and selected Engine HP for x-axis, Highway MPG for y-axis and MSRP as size of the bubbles.
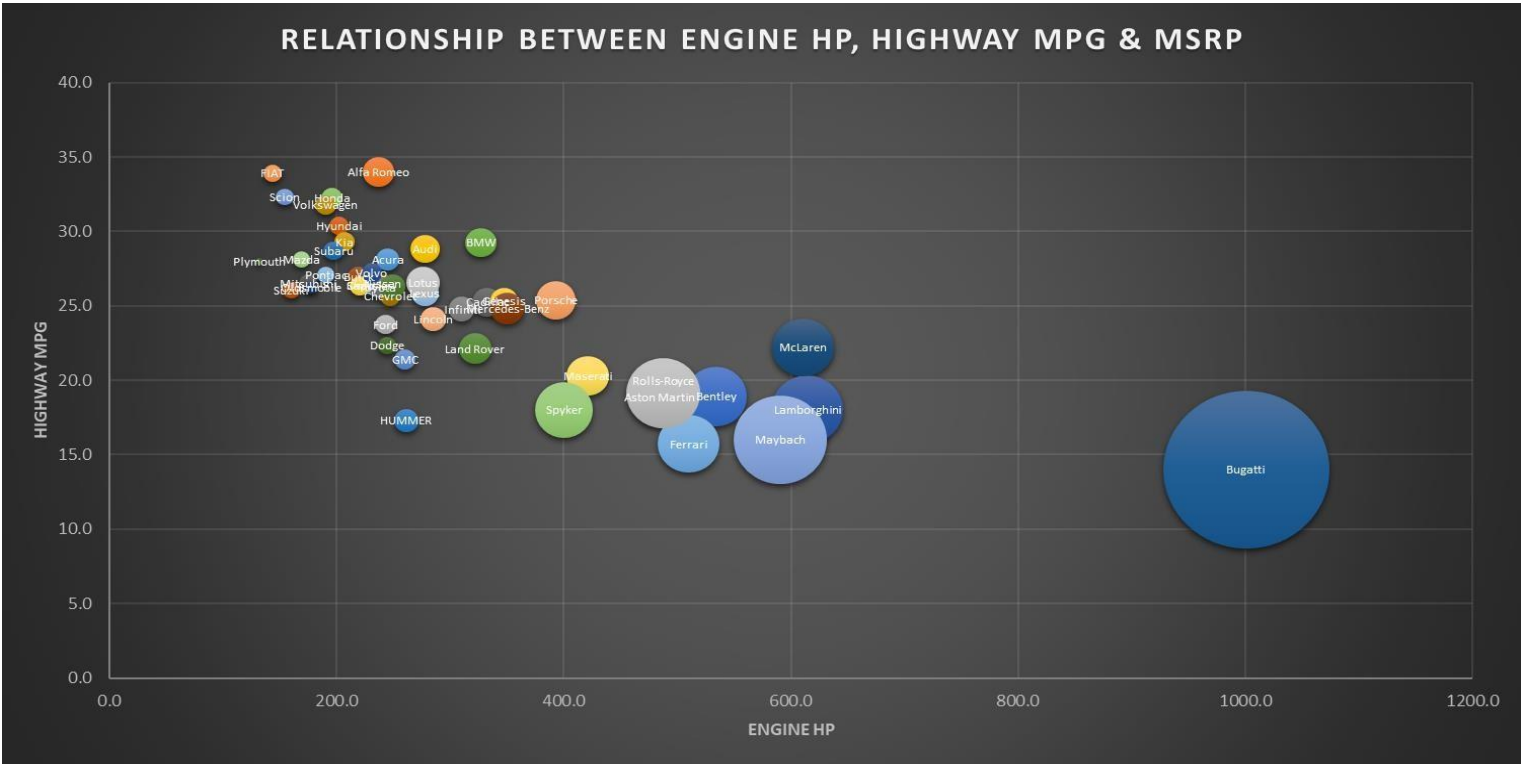
Result: Due to several factors including brand positioning, target market, technology improvements, and production decisions, the horsepower, MPG, and price of a car can change dramatically between different automobile manufacturers.

✓ High-performance automobiles with incredibly high horsepower are a hallmark of performance-oriented manufacturers like Bugatti, Maybach, Lamborghini. These automakers frequently place a high priority on acceleration, speed, and

track performance, providing vehicles with potent motors that produce outstanding horsepower numbers. The variety of automobiles that luxury manufacturers like Rolls Royce, Bentley, BMW, Mercedes Benz, Porche etc., normally provide has varied horsepower outputs. They provide a balance between power and elegance for a variety of clients, including luxury- and performance-focused ones. A large selection of automobiles aimed at the mass market are available from mainstream manufacturers including Toyota, Honda, Ford, and Chevrolet. Although they may provide high-performance variations, the bulk of their vehicles are focused on utility, fuel economy, and affordability, and have moderate to low horsepower.

✓ The body shape of a vehicle can also affect its fuel economy. In general, compared to bigger SUVs, smaller, more aerodynamic cars like compact cars, hatchbacks, and sedans tend to deliver greater MPG.

✓ Luxury automakers like Rolls-Royce, Bentley, Bugatti etc., create expensive, special automobiles with cutting-edge features and top-notch construction. Due to their upscale facilities, premium amenities, and restricted manufacturing, these brands can fetch high rates. Mainstream brands like Toyota, Honda, Ford, Chevrolet etc., offer a wide range of vehicles targeting different market segments. They provide affordable options as well as higher- priced models with advanced features, resulting in a varied price range.

✓ Some companies prioritise affordability by providing cars at more affordable pricing points. To entice budget-conscious shoppers, brands like Kia, Hyundai, Fiat, Mitsubishi etc., frequently provide reasonable pricing and feature-packed vehicles.

Tech Stack Used

➢ **Microsoft Excel:** In data analysis, Excel's PivotTable and PivotChart capabilities are highly regarded. They let users to swiftly reorganise and aggregate data depending on many dimensions, which enables users to summarise and analyse enormous datasets. Users may dynamically slice, filter, and dig down into data using pivot tables, which offers insights and promotes interactive data exploration. Excel provides users with a number of graphing and charting options that let them design aesthetically appealing and educational data visualisations. Data patterns, trends, and correlations may be better understood and communicated with the use of visual representations like charts and graphs.

➢ **Microsoft Word:** The formatting choices in Word are extensive and include fonts, styles, colours, headers, footers, and page layout options. With the help of these tools, users may produce reports that have a professional aesthetic and uniform layout. It is helpful for organising and presenting information in a report format since users can simply alter margins, include tables and photos, and apply formatting to headers and paragraphs. A built-in function in Microsoft Word enables users to save documents in the Portable Document Format (PDF) format. Since PDF maintains the formatting and style across many devices and operating systems, it is frequently used for document sharing. Users may make sure that reports preserve their intended design and can be seen regularly by others by turning them into PDFs.

Insights

The main motive of this project is to provide an insight on the creation of dashboards. The client was interested in learning how a vehicle manufacturer may make the most profitable pricing and product development choices while also satisfying consumer demand. It was possible examine the connection between a car's attributes, market segment, and price to solve this issue.

A data analyst may evaluate the fuel economy of different automobile kinds by looking at the MPG factors in the dataset, which enables customers to make wise vehicle purchase decisions. This research makes it possible to identify the different automobile categories that are the most energy-efficient, giving important information about the fuel-efficient and environmentally friendly car models. Consumers may prioritise fuel economy based on this information and select vehicles that suit their preferences for sustainability and long-term cost savings. Such research enables people to make wise decisions, which helps to create a more sustainable and cost- effective transportation environment.

The properties of the dataset and the market category variables may be used by a data analyst to build a model to forecast the cost of a car. Manufacturers and customers may better understand how different features affect automobile costs with the help of our predictive study. It helps customers understand the variables influencing automobile prices and enables manufacturers to make well-informed decisions about pricing strategies. By utilising this model, producers and customers may set prices and make better informed purchases in the automobile industry.

For data analysts interested in examining many parts of the automobile sector, the dataset is a useful resource. Analysts may acquire important insights that guide important decisions on product development, marketing strategies, and pricing strategies thanks to its extensive information on automotive characteristics, market categories, and price. Analysts may learn more about the automobile market by investigating this information for patterns, trends, and correlations. This information enables key players in the sector to execute well-informed, data-driven decisions that result in better goods, focused advertising campaigns, and aggressive pricing tactics.

It's important to note that market trends can vary regionally and are subject to change over time. Automakers continuously introduce new models and adapt their offerings to meet evolving consumer demands. Additionally, factors such as pricing, performance, brand reputation, and marketing efforts also influence the popularity of car models within specific market categories. Therefore, it is essential to consider the specific dynamics of each market segment and consult up-to-date market research and sales data to understand the current popularity of car models across different market categories.

<u>Results</u>

Regression analysis, pivot tables, sensitivity analysis, optimisation, and time series analysis are just a few of the data analysis techniques that must be mastered in order to complete the specified tasks that are connected to this business challenge. A data analyst may provide a vehicle manufacturer with priceless information if these duties are completed effectively along with the creation of an interactive dashboard. These insights support decisions about product development and price strategies, maximising revenue while successfully satisfying customer demand.

The data analyst is essential to assisting the decision-making process at the vehicle manufacturing and promoting overall business success due to their expertise in Excel and data analysis methodologies.

The following points provide the necessary insight to all the tasks given by the client:

- ✓ A variety of variables might cause the popularity of a certain automobile model to change throughout various market segments. These variables include target markets within each market sector, price range, and customer preferences. Manufacturers and marketers must understand these dynamics in order to create and promote automobile models that appeal to the unique tastes and needs of each market group. Manufacturers and marketers may increase their chances of success and gain a competitive edge in the automobile sector by customising their solutions to fit the distinct demands of various segments.

- ✓ The price of an automobile generally has a positive connection with the engine's output, meaning that as output grows, so does the price. Higher engine power-related production costs and, as a result, higher car prices are caused by advanced engineering, technology, and componentry.

- ✓ To illustrate the relative relevance of each variable, a bar chart has been made after completing regression analysis and collecting coefficient values. The height of each bar correlates to the coefficient value, making it easy to see how the many factors affect the cost of the automobile.

- ✓ Due to variations in pricing strategies, target markets, brand positioning, and product offers, automobile prices vary widely amongst manufacturers. While mainstream manufacturers provide a variety of vehicles at various price points to target different market groups, luxury brands charge greater costs.

✓ Higher cylinder counts typically result in worse fuel economy. The link between the number of cylinders in a car's engine and fuel efficiency is frequently complicated. Due to their smaller displacement and lower fuel consumption, 4-cylinder engines and other smaller engines with fewer cylinders often have superior fuel economy.

✓ Luxury automakers with high average MSRPs, such Bugatti and Maybach, are recognised for their luxury vehicles with high-end materials and cutting- edge technology. While Plymouth normally has the lowest MSRP, popular and low-cost brands frequently have lower average MSRPs, offering a variety of economical alternatives for a wider spectrum of purchasers.

✓ The MSRP of an automobile can be impacted by the gearbox type, with automatic transmissions often costing more than manual transmissions. varied body types have varied effects on this, particularly in the sedan market, where automatic transmissions are frequently favoured for their comfort, cars with automatic transmissions tend to have higher MSRPs than those with manual gearboxes.

✓ Due to variables including technology developments, modifications in vehicle design, and changing consumer tastes, fuel economy in automobiles might differ between body types and model years. Different body types have different aerodynamics and weights, and as engine technology and lightweight materials progress, this leads to an overall rise in fuel economy.

Drive Link

Excel Sheet Link:-

https://docs.google.com/spreadsheets/d/1DBRjoQkV1brc3kx0TvNGNL1h2vPk
LDgU/edit?usp=sharing&ouid=116704808750149683637&rtpof=true&sd=true

Video Presentation Link:-

https://drive.google.com/file/d/1a3QN4aPn7hKhuqca1k1_JUSINY1l0Fy2/view?
usp=sharing