TRAINITY

**PROJECT – 7**

**IMPACT OF CAR FEATURES**

Submitted By

**BOJJA JAHNAVI**

Data Analytics Trainee

Submitted On:- 10-12-2023

**IMPACT OF CAR FEATURES**

**Project Description:**

**Overview:**

The project “Impact of Car Features” is performed to analyze the price estimation of different car brands based on their features. Various Car features like make, engine fuel type, engine hp, market category, vehicle type, transmission type etc are taken to compare them with price of the cars. All these are taken as independent variables which provide value to the dependent variable price.

**Business Problem:**

There is a rapid growth in automotive industry from past 2 decades. The growth electric and hybrid vehicles increased fuel efficiency which brought demand for hydrogen and natural gas. As a data analyst we have to optimise the price based on the features on consumer demand. The price is set and decided in such a way to increase the profitability. This problem is solved by finding various car features based on consumers interests. The relationship between the car features and estimated price is found by regression analysis. We obtain clear analysis of optimizing the price by following regression method. Therefore price is set by understanding the overall car features that makes profitability to the industry.

**Description of the Dataset:**

The Dataset consists of different car features and their MSRP. There are a total of 11159 rows and 16 columns which is in CSV format. Below given are the variables of the dataset.

Make: Brand of the car

Model: Specific model

Year: Year the car is released

Engine Fuel Type: Type of fuel used (eg: Diesel)

Engine HP: Horsepower of engine

Engine Cylinders: No. of cylinders present in engine

Transmission Type: Type of transmission (eg:automotive)

Driven\_wheels: Type of wheels present (front,all)

No. of Doors: No of Doors present in the car

Marketing Type: Type of market the car belongs to

Vehicle size: Size of the car

Vehicle Style: Style of the car

Highway MPG: Estimated Miles per gallons the car gets on highway

City MPG: Estimated Miles per gallons the car gets in city

Popularity: Popularity of the car

MSRP: Manufactures suggested retail price

**Data Cleaning:**

Clean the given data by removing the dupicates, null values and blank cells to get proper results.

**Approach:**

I used Pivot Tables, Visualization Charts, Slicers to find the insights for the given tasks. Pivot tables are used to Classify the car features by adjusting them in rows, columns and values. This gives brief information of car features about their average and sum values. Slicers are used to find the particular information related to any of the car feature. This works by selecting any one option in the slicer which displays information in the pivot table and in the charts. Charts are done for visualizing the data to understand it conveniently. Visualization techniques include usage of different charts like column, Bar graph, Scatter plot, bubble chart etc. These are the techniques used to draw the solutions for the given tasks. In this project there are certain challenges encountered while doing the project. Few charts are not allowed to visualize from pivot tables. So We have to copy and paste the pivot table to get the chart. This is all about the descriptive analysis, visualization techniques used in the project.

**Tech-Stack Used:**

The Project “Impact of Car Features” is done in Microsoft Excel 2021. I used this tool as it is very simple to learn and handle for beginners. It has various functions to solve the questions and derive the answers. Therfore it is flexible to use and is free of cost.

**Insights:**

**Popularity of a car model across different market categories:**

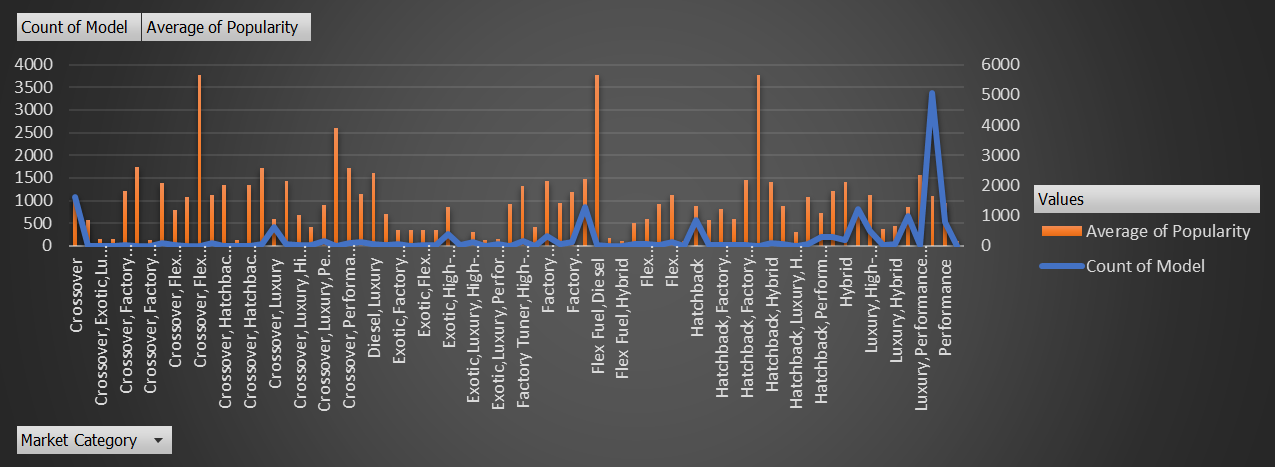
**Task 1.A**: The number of car models in each market category are found by using pivot tables. Corresponding popularity score of each market category is estimated by finding the average popularity score.

This is the pivot table I have done to find the average popularity score of models in each market category.

|  |  |  |
| --- | --- | --- |
| **Market Category** | **Count of Model** | **Average of Popularity** |
| Crossover | 1075 | 1556.168372 |
| Crossover,Diesel | 7 | 873 |
| Crossover,Exotic,Luxury,High-Performance | 1 | 238 |
| Crossover,Exotic,Luxury,Performance | 1 | 238 |
| Crossover,Factory Tuner,Luxury,High-Performance | 26 | 1823.461538 |
| Crossover,Factory Tuner,Luxury,Performance | 5 | 2607.4 |
| Crossover,Factory Tuner,Performance | 4 | 210 |
| Crossover,Flex Fuel | 64 | 2073.75 |
| Crossover,Flex Fuel,Luxury | 10 | 1173.2 |
| Crossover,Flex Fuel,Luxury,Performance | 6 | 1624 |
| Crossover,Flex Fuel,Performance | 6 | 5657 |
| Crossover,Hatchback | 72 | 1675.694444 |
| Crossover,Hatchback,Factory Tuner,Performance | 6 | 2009 |
| Crossover,Hatchback,Luxury | 7 | 204 |
| Crossover,Hatchback,Performance | 6 | 2009 |
| Crossover,Hybrid | 42 | 2563.380952 |
| Crossover,Luxury | 406 | 889.2142857 |
| Crossover,Luxury,Diesel | 34 | 2149.411765 |
| Crossover,Luxury,High-Performance | 9 | 1037.222222 |
| Crossover,Luxury,Hybrid | 24 | 630.9166667 |
| Crossover,Luxury,Performance | 112 | 1349.089286 |
| Crossover,Luxury,Performance,Hybrid | 2 | 3916 |
| Crossover,Performance | 69 | 2585.956522 |
| Diesel | 84 | 1730.904762 |
| Diesel,Luxury | 47 | 2416.106383 |
| Exotic,Factory Tuner,High-Performance | 21 | 1046.380952 |
| Exotic,Factory Tuner,Luxury,High-Performance | 51 | 523.0196078 |
| Exotic,Factory Tuner,Luxury,Performance | 3 | 520 |
| Exotic,Flex Fuel,Factory Tuner,Luxury,High-Performance | 13 | 520 |
| Exotic,Flex Fuel,Luxury,High-Performance | 11 | 520 |
| Exotic,High-Performance | 252 | 1273.678571 |
| Exotic,Luxury | 12 | 112.6666667 |
| Exotic,Luxury,High-Performance | 77 | 473.025974 |
| Exotic,Luxury,High-Performance,Hybrid | 1 | 204 |
| Exotic,Luxury,Performance | 36 | 217.0277778 |
| Exotic,Performance | 6 | 1391 |
| Factory Tuner,High-Performance | 104 | 1966.442308 |
| Factory Tuner,Luxury | 2 | 617 |
| Factory Tuner,Luxury,High-Performance | 215 | 2133.367442 |
| Factory Tuner,Luxury,Performance | 31 | 1413.419355 |
| Factory Tuner,Performance | 84 | 1774.047619 |
| Flex Fuel | 855 | 2225.71345 |
| Flex Fuel,Diesel | 16 | 5657 |
| Flex Fuel,Factory Tuner,Luxury,High-Performance | 1 | 258 |
| Flex Fuel,Hybrid | 2 | 155 |
| Flex Fuel,Luxury | 39 | 746.5384615 |
| Flex Fuel,Luxury,High-Performance | 32 | 898.3125 |
| Flex Fuel,Luxury,Performance | 28 | 1380.071429 |
| Flex Fuel,Performance | 87 | 1680.471264 |
| Flex Fuel,Performance,Hybrid | 2 | 155 |
| Hatchback | 574 | 1308.65331 |
| Hatchback,Diesel | 14 | 873 |
| Hatchback,Factory Tuner,High-Performance | 13 | 1205.153846 |
| Hatchback,Factory Tuner,Luxury,Performance | 9 | 886.8888889 |
| Hatchback,Factory Tuner,Performance | 21 | 2173.714286 |
| Hatchback,Flex Fuel | 7 | 5657 |
| Hatchback,Hybrid | 64 | 2111.15625 |
| Hatchback,Luxury | 45 | 1323.133333 |
| Hatchback,Luxury,Hybrid | 3 | 454 |
| Hatchback,Luxury,Performance | 36 | 1632.25 |
| Hatchback,Performance | 198 | 1073.661616 |
| High-Performance | 198 | 1823.378788 |
| Hybrid | 121 | 2116.586777 |
| Luxury | 819 | 1079.214896 |
| Luxury,High-Performance | 334 | 1668.017964 |
| Luxury,High-Performance,Hybrid | 12 | 568.8333333 |
| Luxury,Hybrid | 52 | 673.6346154 |
| Luxury,Performance | 659 | 1293.062215 |
| Luxury,Performance,Hybrid | 11 | 2333.181818 |
| N/A | 3376 | 1664.832938 |
| Performance | 520 | 1415.209615 |
| Performance,Hybrid | 1 | 155 |
| **Grand Total** | **11193** | **1558.449567** |

The marketing category of crossover has the highest number of models i.e 1075 with average popularity score of 1556.168372

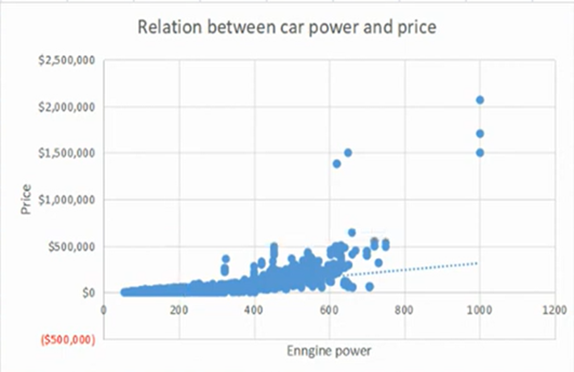
**Task 1.B:** A combo chart is visualized to analyze the relation between market category and popularity score.



Crossover has the highest no. of models. Crossover, Flex Fuel, Performance and Flex Fuel Diesel has the highest average popularity.

**Relationship between a car's engine power and its price:**

**Task 2:** A Scatter plot is drawn in in which engine power is on X – axis and Price on Y – axis.



Increase in horse power leads to increase in price of the car. So there is a positive trend line shown in the chart.

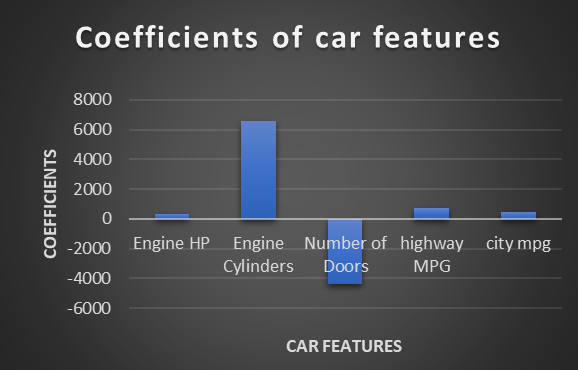
**Car features that are important to determine car price:**

**Task 3:** Regression analysis to determine the variables that have strongest relationship with cars price.

|  |  |
| --- | --- |
|  | ***Coefficients*** |
| **Intercept** | -88129.32512 |
| **Engine HP** | 315.9215914 |
| **Engine Cylinders** | 6557.490451 |
| **Number of Doors** | -4400.530194 |
| **highway MPG** | 719.6857364 |
| **city mpg** | 458.2418855 |

Engine HP, Engine Cylinders, NO. of Doors, Highway MPG, City MPG are considered as the variables responsible for optimizing cars price.

Coefficient variables to determine strongest variables of car features:

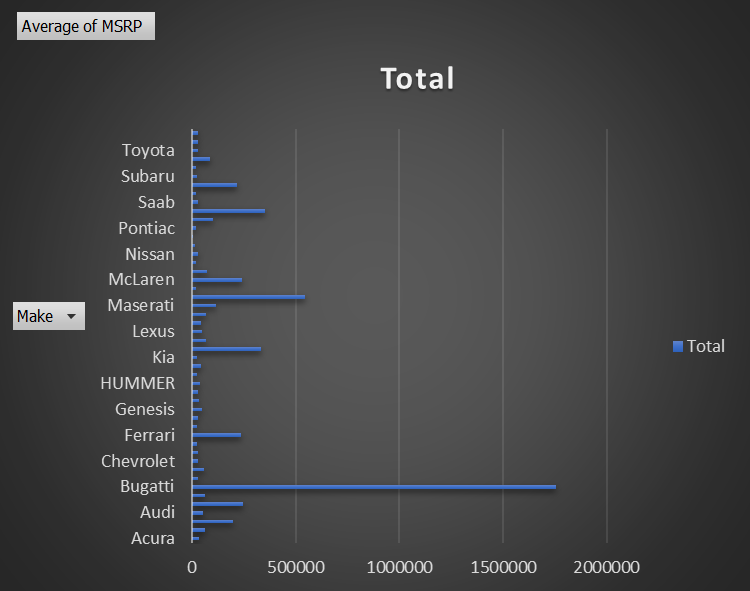


Engine cylinders has the strongest relation to cars price and No. of doors has the negative relationship. So we can conclude that if no. of cylinders are more then price of the car increases and the no. of doors decreases the price increases.

**Average price of a car across different manufacturers:**

**Task 4.A**: Used pivot tables to identify the average price of cars to identify the each markets

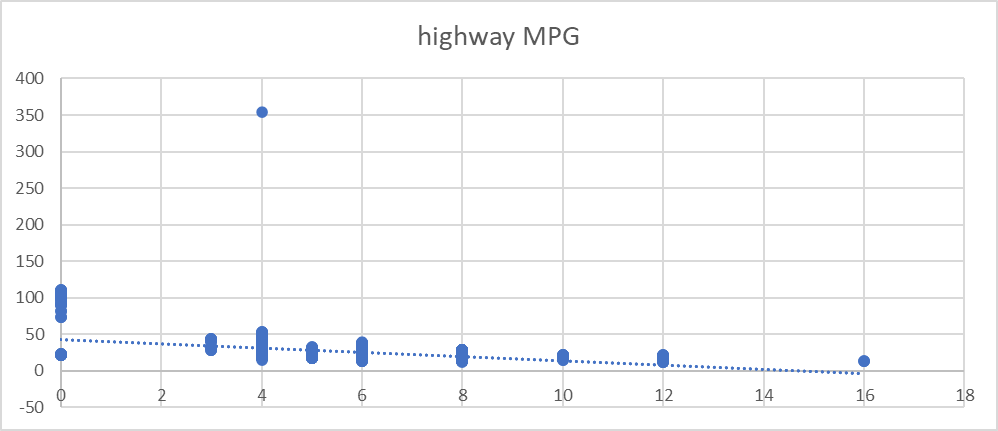
|  |  |
| --- | --- |
| **Manufacturers** | **Average of MSRP** |
| Acura | 35087.4878 |
| Alfa Romeo | 61600 |
| Aston Martin | 198123.4615 |
| Audi | 54574.1215 |
| Bentley | 247169.3243 |
| BMW | 62162.55864 |
| Bugatti | 1757223.667 |
| Buick | 29034.18947 |
| Cadillac | 56368.26515 |
| Chevrolet | 29074.72576 |
| Chrysler | 26722.96257 |
| Dodge | 24857.04537 |
| Ferrari | 237383.8235 |
| FIAT | 22670.24194 |
| Ford | 28511.30788 |
| Genesis | 46616.66667 |
| GMC | 32444.08506 |
| Honda | 26655.14781 |
| HUMMER | 36464.41176 |
| Hyundai | 24926.26255 |
| Infiniti | 42640.27134 |
| Kia | 25513.75546 |
| Lamborghini | 331567.3077 |
| Land Rover | 68067.08633 |
| Lexus | 47549.06931 |
| Lincoln | 43860.825 |
| Lotus | 68377.14286 |
| Maserati | 113684.4909 |
| Maybach | 546221.875 |
| Mazda | 20416.62379 |
| McLaren | 239805 |
| Mercedes-Benz | 72069.52786 |
| Mitsubishi | 21340.5625 |
| Nissan | 28921.15245 |
| Oldsmobile | 12843.79545 |
| Plymouth | 3296.873239 |
| Pontiac | 19800.0442 |
| Porsche | 101622.3971 |
| Rolls-Royce | 351130.6452 |
| Saab | 27879.80734 |
| Scion | 19932.5 |
| Spyker | 214990 |
| Subaru | 24240.67364 |
| Suzuki | 18026.4152 |
| Tesla | 85315.38462 |
| Toyota | 28846.5605 |
| Volkswagen | 28978.52289 |
| Volvo | 29724.68421 |
| **Grand Total** | **41884.03091** |



Bugatti has the highest average car price.

**Relationship between fuel efficiency & no. of cylinders in a car's engine:**

**Task 5.A:** Relationship between fuel efficiency and the number of cylinders in a car's engine on a scatter plot. In the chart No. of cylinders are taken on X- axis and Highway MPG is taken on Y-axis.



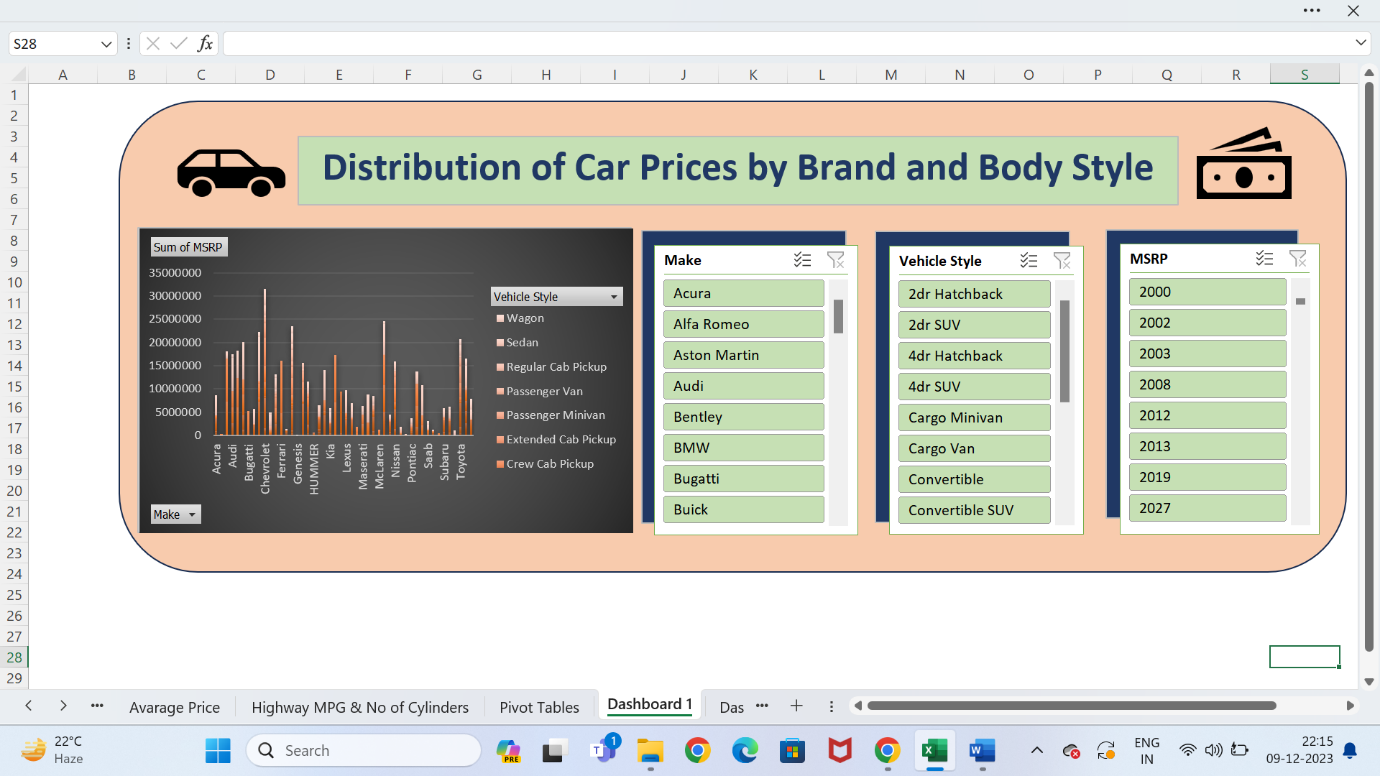
From the graph it is shown that If the number of cylinders are more then the highway MPG is also less. Increase in the cylinders has decrease in the fuel efficiency. Slope is -2.9688

**Task 5.B**: Correlation between no. of cylinders and highway MPG.

The correlation between no of cylinders and highway MPG is -0.609

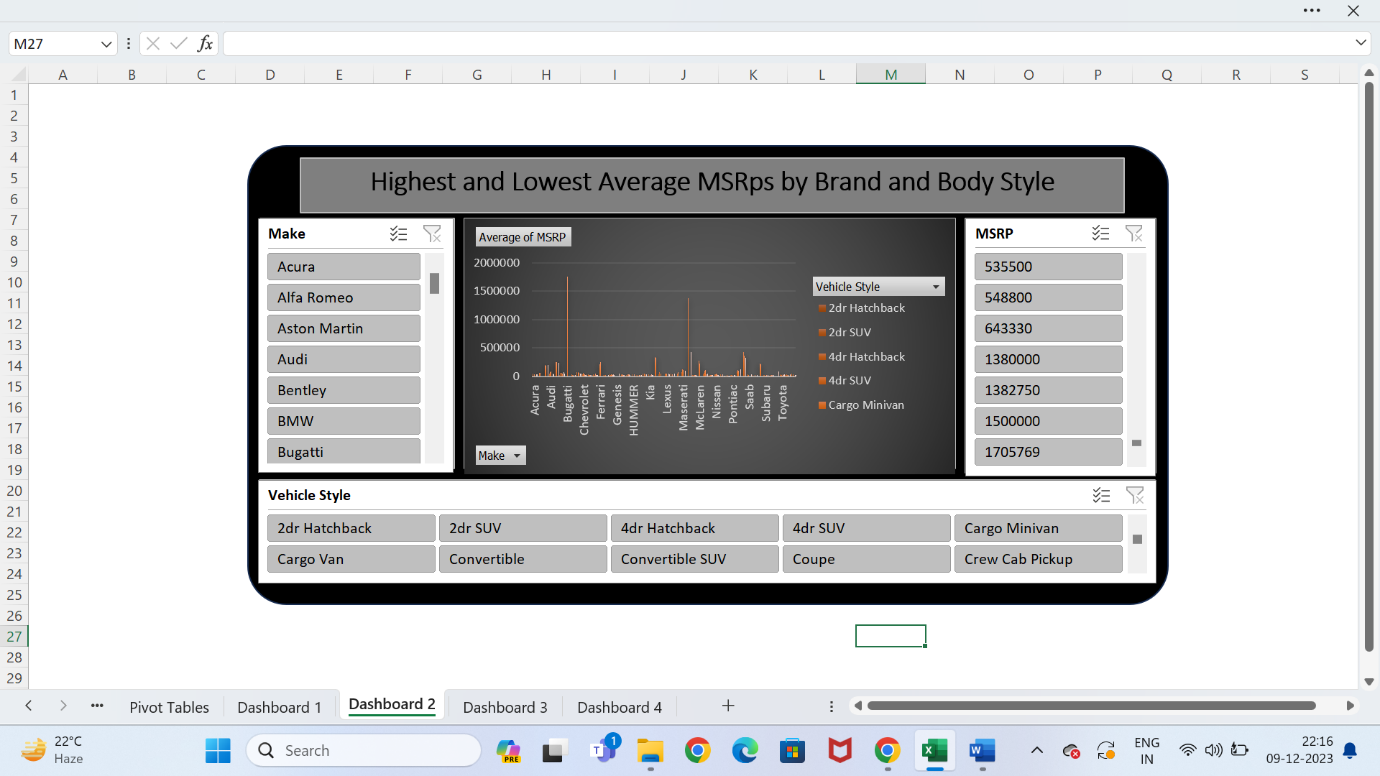
**Building Dashboard:**

**Task 1: distribution of car prices by brand and body style**

****

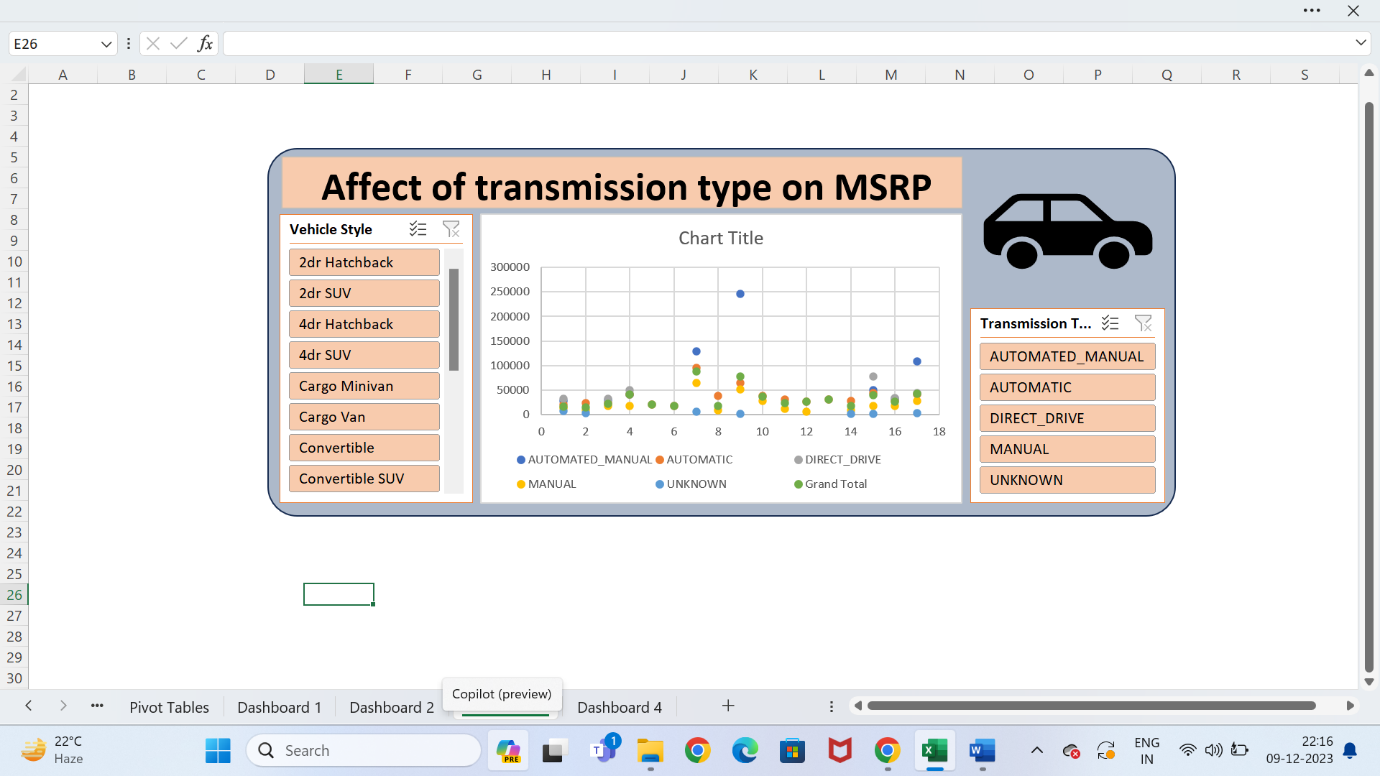
Chevloret and Mcleren has the highest car price.

**Task 2 : Highest and Lowest Average MSRps by Brand and Body Style**

****

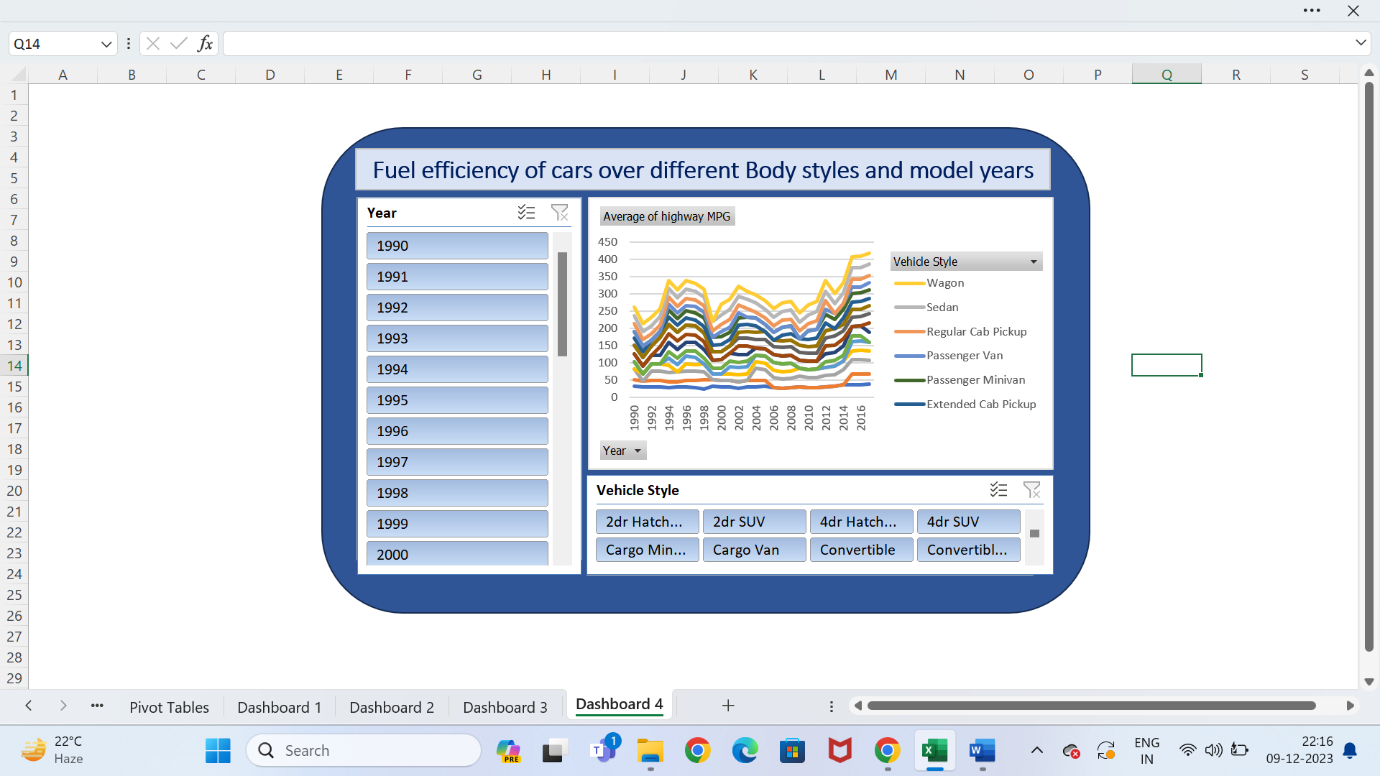
From this Dashboard Bugatti, Coupe style has the highest average MSRP and Acura, Volkswagan, Suzuki etc. has the lowest average MSRP.

**Task 3: Affect of transmission type on MSRP**

****

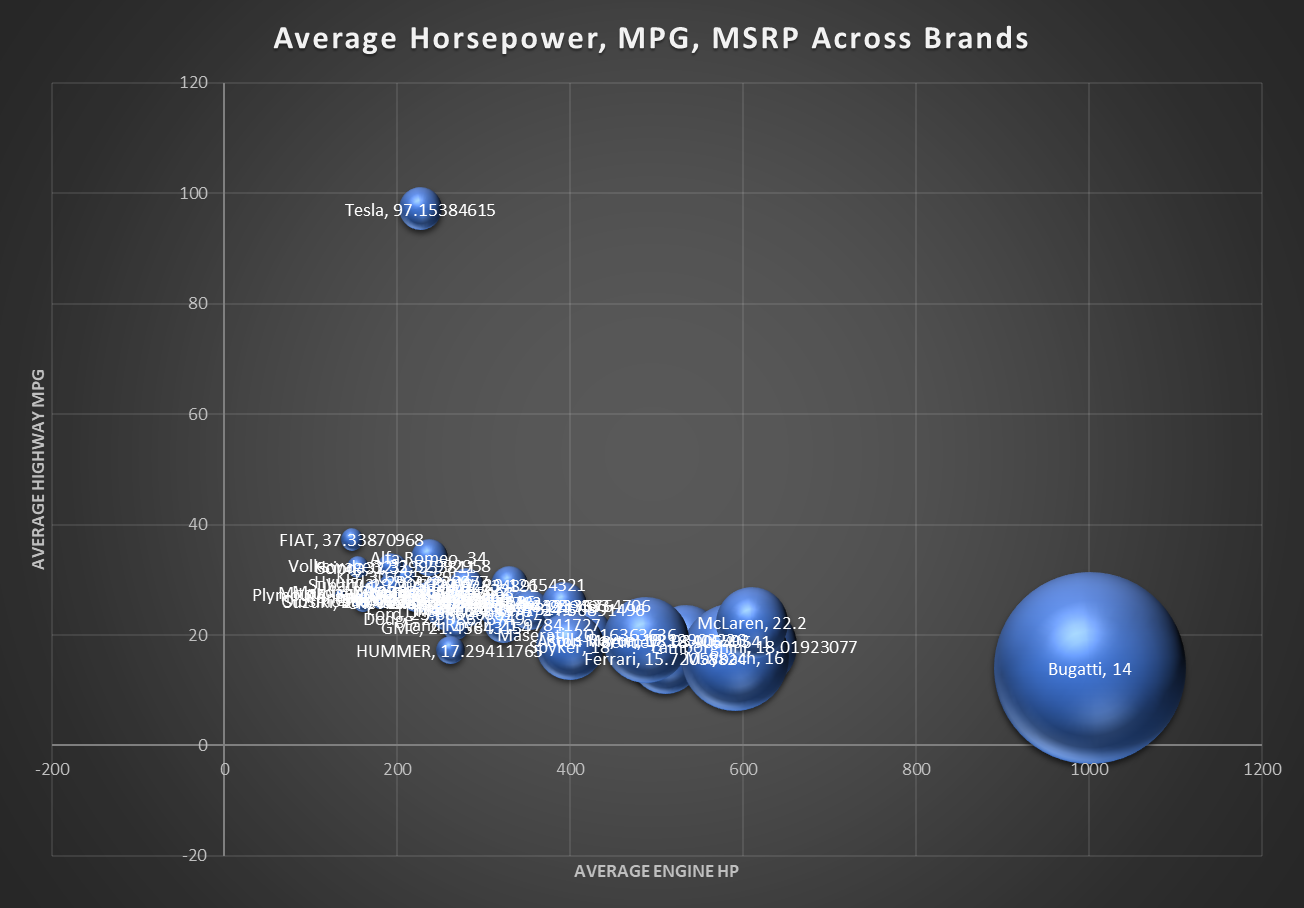
Automated manual transmission type has highest MSRP and it varies a lot in the vehicle styles.

**Task 4: Fuel efficiency of cars over different Body styles and model years**



The fuel efficiency of different body styles has increased from the year 1990 to 2017.

**Task 5: Horsepower, MPG and Price across different Brands**

****

|  |  |  |  |
| --- | --- | --- | --- |
| **Row Labels** | **Average of Engine HP** | **Average of highway MPG** | **Average of MSRP** |
| Acura | 244.9634146 | 28.2195122 | 35087.4878 |
| Alfa Romeo | 237 | 34 | 61600 |
| Aston Martin | 483.7582418 | 18.93406593 | 198123.4615 |
| Audi | 280 | 28.92834891 | 54574.1215 |
| Bentley | 533.8513514 | 18.90540541 | 247169.3243 |
| BMW | 329.6203704 | 29.12654321 | 62162.55864 |
| Bugatti | 1001 | 14 | 1757223.667 |
| Buick | 220.0105263 | 27.01052632 | 29034.18947 |
| Cadillac | 332.7954545 | 25.24494949 | 56368.26515 |
| Chevrolet | 249.3591874 | 25.92705448 | 29074.72576 |
| Chrysler | 229.1390374 | 26.36898396 | 26722.96257 |
| Dodge | 254.3534972 | 22.98676749 | 24857.04537 |
| Ferrari | 509.9117647 | 15.72058824 | 237383.8235 |
| FIAT | 147.5967742 | 37.33870968 | 22670.24194 |
| Ford | 249.3345455 | 23.88606061 | 28511.30788 |
| Genesis | 347.3333333 | 25.33333333 | 46616.66667 |
| GMC | 267.6452282 | 21.45643154 | 32444.08506 |
| Honda | 196.9122402 | 32.39953811 | 26655.14781 |
| HUMMER | 261.2352941 | 17.29411765 | 36464.41176 |
| Hyundai | 205.2046332 | 29.77220077 | 24926.26255 |
| Infiniti | 310.6768293 | 24.79573171 | 42640.27134 |
| Kia | 207.9825328 | 30.68558952 | 25513.75546 |
| Lamborghini | 614.0769231 | 18.01923077 | 331567.3077 |
| Land Rover | 322.5179856 | 21.97841727 | 68067.08633 |
| Lexus | 277.4158416 | 25.87623762 | 47549.06931 |
| Lincoln | 283.16875 | 24.5375 | 43860.825 |
| Lotus | 271.5357143 | 26.10714286 | 68377.14286 |
| Maserati | 419.5454545 | 20.16363636 | 113684.4909 |
| Maybach | 590.5 | 16 | 546221.875 |
| Mazda | 172.5436893 | 27.93932039 | 20416.62379 |
| McLaren | 610.4 | 22.2 | 239805 |
| Mercedes-Benz | 353.1290323 | 24.56891496 | 72069.52786 |
| Mitsubishi | 173.6057692 | 27.64423077 | 21340.5625 |
| Nissan | 241.1143376 | 27.77495463 | 28921.15245 |
| Oldsmobile | 179.7348485 | 26.18939394 | 12843.79545 |
| Plymouth | 133.7464789 | 27.4084507 | 3296.873239 |
| Pontiac | 192.3370166 | 26.96132597 | 19800.0442 |
| Porsche | 392.7941176 | 25.36764706 | 101622.3971 |
| Rolls-Royce | 487.5483871 | 19.12903226 | 351130.6452 |
| Saab | 221.1743119 | 26.37614679 | 27879.80734 |
| Scion | 154.4333333 | 32.3 | 19932.5 |
| Spyker | 400 | 18 | 214990 |
| Subaru | 193.2887029 | 29.20502092 | 24240.67364 |
| Suzuki | 160.4853801 | 26.00292398 | 18026.4152 |
| Tesla | 227 | 97.15384615 | 85315.38462 |
| Toyota | 234.2002782 | 26.62447844 | 28846.5605 |
| Volkswagen | 192.1267606 | 32.52992958 | 28978.52289 |
| Volvo | 234.5601504 | 27.26315789 | 29724.68421 |
| **Grand Total** | **253.202448** | **26.57714643** | **41884.03091** |

Bugatti has the highest average Engine HP. Various brand have different Highway MPG and price. This Variations may be due to fuel efficiency and other car features.

**Results:** From the project “Impact of Car Features” I have learnt the advanced excel functions. Firstly there is a huge data which must be cleaned by removing the null values. Then the tasks are derived by using pivot tables and slicers which are converted into meaningful interactive dashboards. Bugatti has the highest average price and Acura has the lowest price. Different brands have differed fuel efficiency based on their no. of cylinders, Engine HP. The more the no. of cylinders in a cars engine the less is the fuel efficiency. Therfore this project gives insights on optimizing the cars price based on the car’s features.

**My Excel Worksheet:**

[**https://docs.google.com/spreadsheets/d/1R1bKNqOHtezpq4UklFOHWfcX\_QcRLb5d/edit?usp=sharing&ouid=101880124803050791429&rtpof=true&sd=true**](https://docs.google.com/spreadsheets/d/1R1bKNqOHtezpq4UklFOHWfcX_QcRLb5d/edit?usp=sharing&ouid=101880124803050791429&rtpof=true&sd=true)

**My Presentation:**

[**https://drive.google.com/file/d/1vcK7OYWDG1glSVf-3a88H5yZzANST-G8/view?usp=sharing**](https://drive.google.com/file/d/1vcK7OYWDG1glSVf-3a88H5yZzANST-G8/view?usp=sharing)