

Impact of Car Feature

Trainity Project – 7

By Vaishnavi Patel

Project Description

- **Purpose of the project** : How can a car manufacturer optimize pricing and product development decisions to maximize profitability?
- **Business Problem** : By providing informative insights from the analyzing data, decisions on optimizing pricing and product development can be made.
- **Description of data sources** : Dataset is “Car Features and MSRP”, Dataset is collected and made available on “Kaggle” by Copper Union which is in CSV format.
- **Description of data cleaning and pre-processing steps** :
 - **Original data set** :
 - Number of observations : 11,159 , Number of variables : 16
 - **Data cleaning and pre-processing steps**:
 - **Duplicate found** : 715 data
 - Columns Engine HP, Engine Cylinders, Number of doors have few blank data

Project Description

- **Outliers:** In Column highway MPG
 - How value of outliers is replaced:
 - Filtered columns by : market category Luxury
 - Mid size vehicle size
 - Sedan Vehicle style
 - Taken understanding of overall data of same category so that outliers values can be understood
 - Data is having outliers of value 354 and replaced it by 34 (Avg. of highway MPG)

➤ After cleaning data:

- Total number of rows: 11101
- Number of variables: 16

➤ Project is in two part:

- Analysis part_1 : Perform task given and give insights on it
- Analysis part_2: Make interactive dashboard using given task

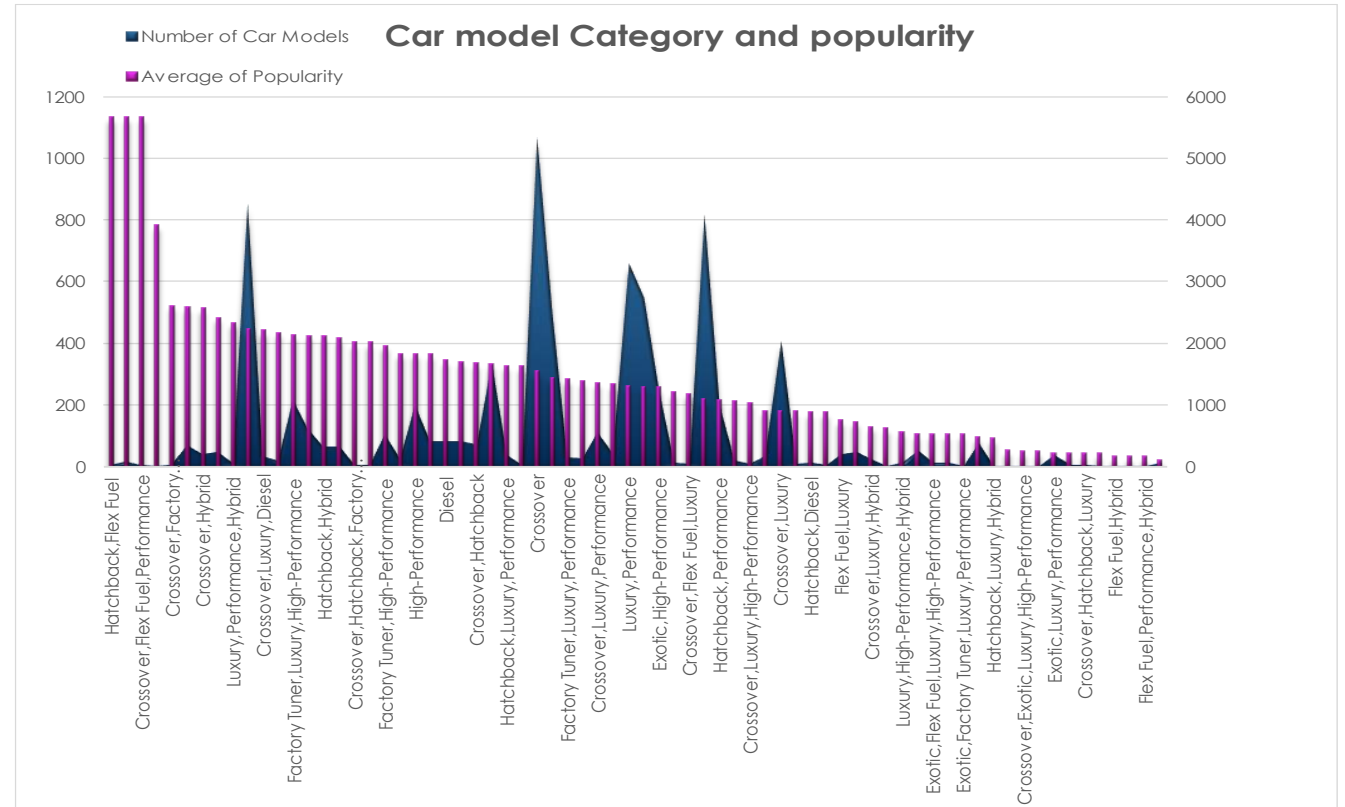
Tech-stack Used

- Software used : **Excel**
- Excel is easy to use, with easy accessible tools, formulas to perform task
- Pivot table , Data analysis (add-ins) for regression analysis, description statistics to replace outliers, and charts for visualization

Approach for Tasks : Analysis Part_1

- Task:1** : How does the popularity of a car model vary across different market categories? Create a pivot table that shows the number of car models in each market category and their corresponding popularity scores. Create a combo chart that visualizes the relationship between market category and popularity.

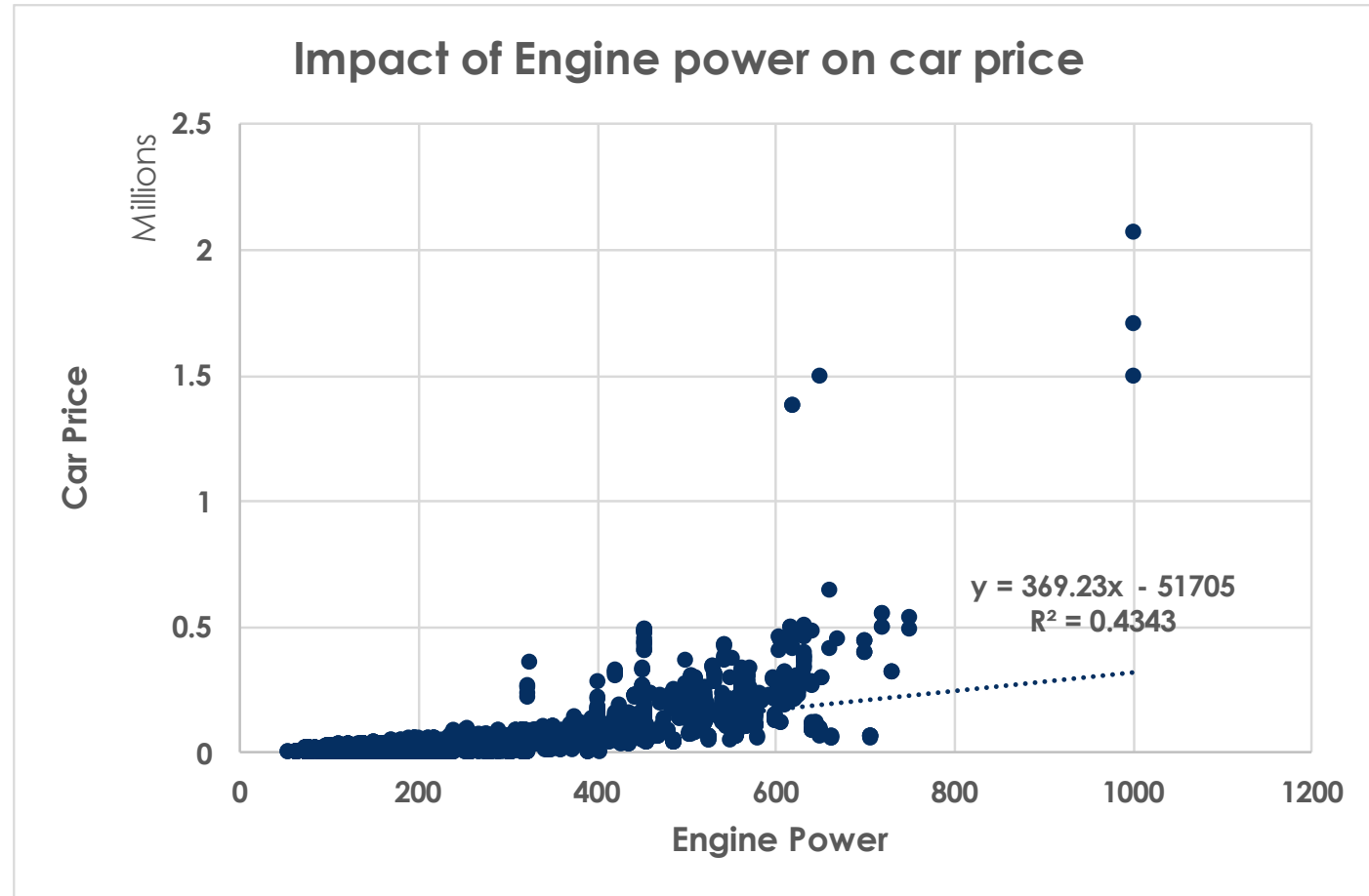
Market Category	Number of Car Models	Average of Popularity
Hatchback, Flex Fuel	7	5657
Flex Fuel, Diesel	16	5657
Crossover, Flex Fuel, Performance	6	5657
Crossover, Luxury, Performance, Hybrid	2	3916
Crossover, Factory Tuner, Luxury, Performance	5	2607
Crossover, Performance	69	2586
Crossover, Hybrid	42	2563
Diesel, Luxury	47	2416
Luxury, Performance, Hybrid	11	2333
Flex Fuel	855	2226
Crossover, Luxury, Diesel	33	2196
Hatchback, Factory Tuner, Performance	21	2174
Factory Tuner, Luxury, High-Performance	215	2133
Hybrid	121	2117
Hatchback, Hybrid	64	2111
Crossover, Flex Fuel	64	2074
Crossover, Hatchback, Factory Tuner, Performance	6	2009
Crossover, Hatchback, Performance	6	2009
Factory Tuner, High-Performance	104	1966
Crossover, Factory Tuner, Luxury, High-Performance	26	1823
High-Performance	198	1823
Factory Tuner, Performance	81	1818
Diesel	84	1731
Flex Fuel, Performance	81	1702
Crossover, Hatchback	72	1676
Luxury, High-Performance	334	1668
Hatchback, Luxury, Performance	36	1632
Crossover, Flex Fuel, Luxury, Performance	6	1624
Crossover	1068	1539
Performance	503	1443
Factory Tuner, Luxury, Performance	31	1413
Flex Fuel, Luxury, Performance	28	1380
Crossover, Luxury, Performance	112	1349
Hatchback, Luxury	45	1323
Luxury, Performance	659	1293
Hatchback	547	1279



- Key findings** : The popularity of a car model vary across different market categories: As seen from the combo chart, average popularity is more for hatchback, flex fuel type of car models and for that number of car models are less.

Approach for Tasks : Analysis Part_1

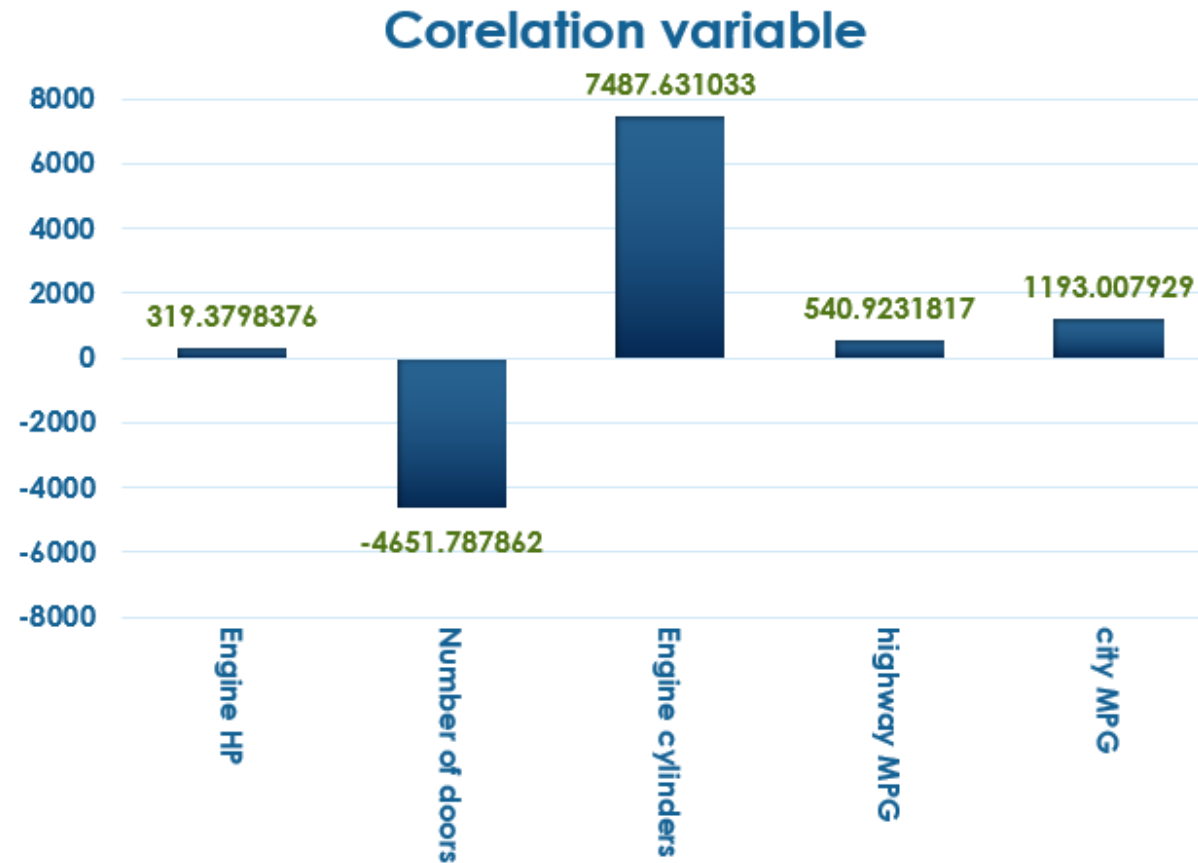
Task:2 : What is the relationship between a car's engine power and its price? Create a scatter chart that plots engine power on the x-axis and price on the y-axis. Add a trendline to the chart to visualize the relationship between these variables.



- **Key findings** : Scatter plot is more for engine power between 200-600 HP with less car price. It is linearly dependent. As engine power increases it will increase the price of car.

Approach for Tasks : Analysis Part_1

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

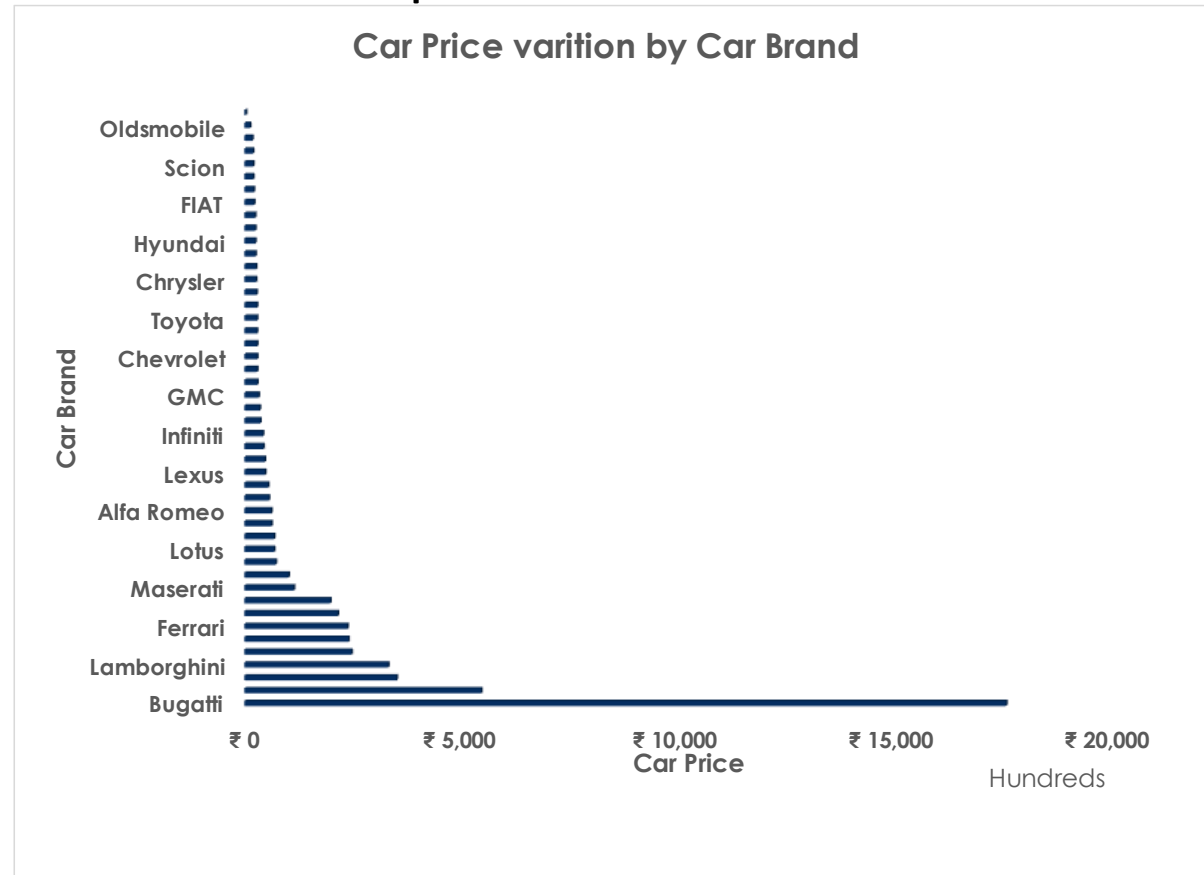


- **Key findings :** The visualized chart shows correlation of Car price with 5-car feature variable: Engine HP, Number of doors, Number of engine cylinders, Highway MPG, city MPG. As seen from these, most important car feature is engine cylinders. It is highly correlated.

Approach for Tasks : Analysis Part_1

Task: 4: How does the average price of a car vary across different manufacturers? Create a pivot table that shows the average price of cars for each manufacturer. Create a bar chart or a horizontal stacked bar chart that visualizes the relationship between manufacturer and average price.

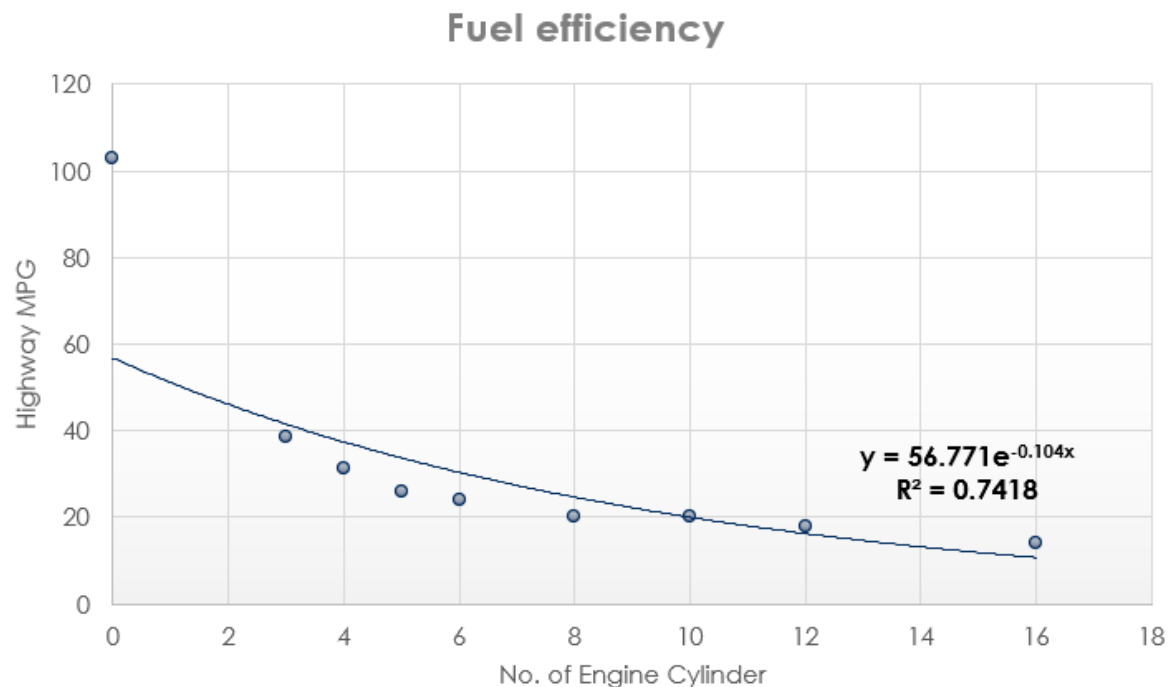
Manufacture	Average of MSRP
Bugatti	₹ 17,57,224
Maybach	₹ 5,46,222
Rolls-Royce	₹ 3,51,131
Lamborghini	₹ 3,31,567
Bentley	₹ 2,47,169
McLaren	₹ 2,39,805
Ferrari	₹ 2,37,384
Spyker	₹ 2,14,990
Aston Martin	₹ 1,98,123
Maserati	₹ 1,13,684
Porsche	₹ 1,01,622
Mercedes-Benz	₹ 72,135
Lotus	₹ 68,377
Land Rover	₹ 68,067
BMW	₹ 62,163
Alfa Romeo	₹ 61,600
Cadillac	₹ 56,368
Audi	₹ 54,574
Lexus	₹ 47,549
Genesis	₹ 46,617
Lincoln	₹ 43,560
Infiniti	₹ 42,640
HUMMER	₹ 36,464
Acura	₹ 35,087
GMC	₹ 32,444
Volvo	₹ 29,725
Buick	₹ 29,034



- Key findings :** The average price of cars for each manufacturer is seen using horizontal bar chart. The brand Bugatti has maximum and Plymouth with minimum avg. car price.

Approach for Tasks : Analysis Part_1

Task: 5: What is the relationship between fuel efficiency and the number of cylinders in a car's engine? 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. Calculate the correlation coefficient between the number of cylinders and highway MPG to quantify the strength and direction of the relationship



```
=CORREL(Table5[Engine Cylinders],Table5[highway MPG])
```

-0.67185 correlation coefficient

- Key findings :** The correlation coefficient between the number of cylinders and highway MPG is 67 % negatively impacted. The lesser the no. of cylinders more will be highway MPG and good fuel efficiency. Because cars with more cylinder are heavier , need more fuel to operate.

Approach for Dashboard : Analysis Part_2

Task 1: How does the distribution of car prices vary by brand and body style? (Stacked column chart to show the distribution of car prices by brand and body style. Use filters and slicers to make the chart interactive. Calculate the total MSRP for each brand and body style using SUMIF or Pivot Tables.)

Task 2: Which car brands have the highest and lowest average MSRPs, and how does this vary by body style? (Clustered column chart to compare the average MSRPs across different car brands and body styles. Calculate the average MSRP for each brand and body style using AVERAGEIF or Pivot Tables)

Task 3: How do the different feature such as transmission type affect the MSRP, and how does this vary by body style?(Scatter plot chart to visualize the relationship between MSRP and transmission type, with different symbols for each body style. Calculate the average MSRP for each combination of transmission type and body style using AVERAGEIFS or Pivot Tables.)

Task 4: How does the fuel efficiency of cars vary across different body styles and model years? (Line chart to show the trend of fuel efficiency (MPG) over time for each body style. Calculate the average MPG for each combination of body style and model year using AVERAGEIFS or Pivot Tables.)

Task 5: How does the car's horsepower, MPG, and price vary across different Brands? (Bubble chart to visualize the relationship between horsepower, MPG, and price across different car brands. Assign different colors to each brand and label the bubbles with the car model name. Calculate the average horsepower, MPG, and MSRP for each car brand using AVERAGEIFS or Pivot Tables.)

Approach for Dashboard : Analysis Part_2

- ✓ All the task have performed and each graph with required slicer are kept in dashboard.
- ✓ Here is the Excel File link :

<https://docs.google.com/spreadsheets/d/1ZvaLFld6KlvghLRAeZseCWotDAIV3eiQ/edit?usp=sharing&oid=114706903831395447522&rtpof=true&sd=true>

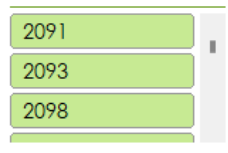
(download the excel file : Excel version : 2019)

- ✓ Excel file include :
 1. Task 1 to Task 5 : Tables, graphs
 2. Graphs for dashboard as DB_1 to DB_5
 3. Clicking on link will show only interactive dashboard. All the other sheets are hidden. (Right click on db sheet : Unhide option will show each sheet).
 4. Next slide shows screenshot of dashboard created in Excel.

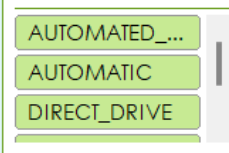
Approach for Dashboard : Analysis Part_2

IMPACT OF CAR FEATURE DASHBOARD

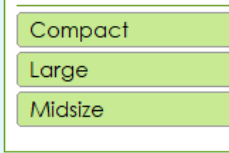
MSRP



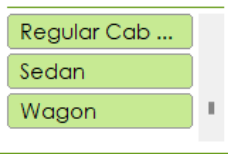
Transmisio...



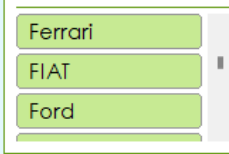
Vehicle Size



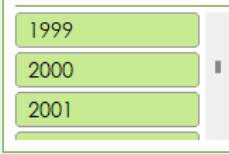
Vehicle Style



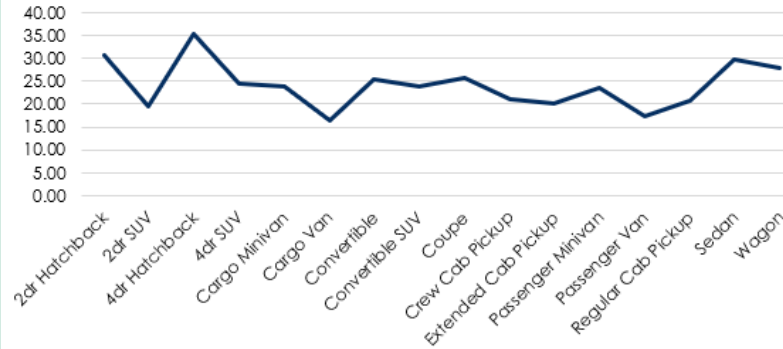
Make



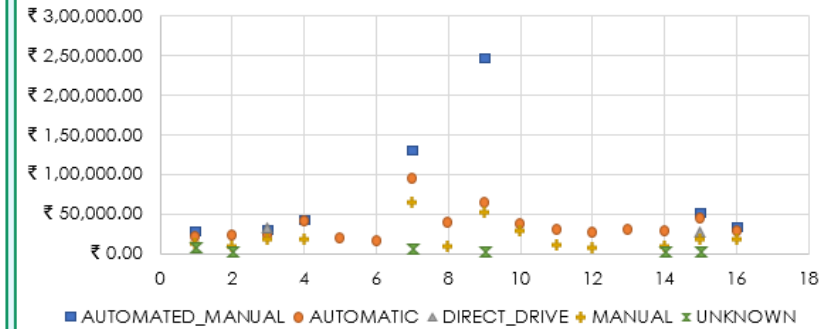
Year



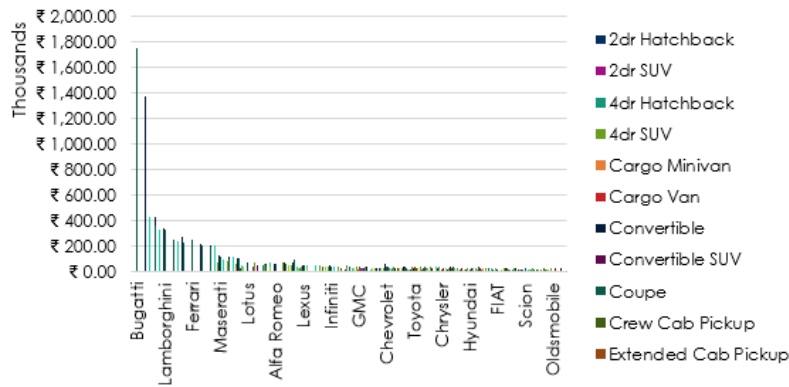
Fuel efficiency of cars across model year and body style



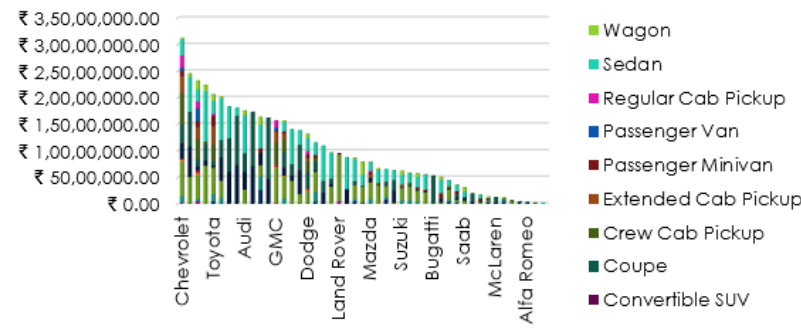
Avg. Price distribution by Transmission type



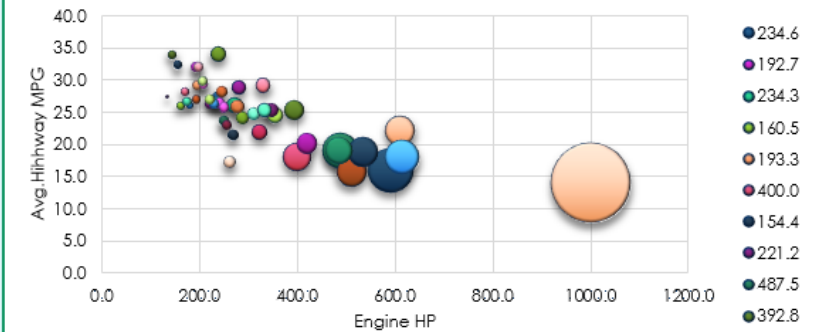
Avg. Car Price distribution by Brand and Body style



Car Price Distribution by Brand & Body style



Engine HP, highway MPG and Avg. Price of car distribution by Car brand



Results

- Conclusions from the analysis:
 - Car price mainly depends upon the no. of cylinders car has, because it highly effects the fuel efficiency. Less cylinders, more fuel efficiency, more popularity, less price of car, more sell and more profit.
 - Demand of electric vehicle is more.

Thank you