

Analyzing the Impact of Car Features on Price and Profitability

Project Description:

The project is about how the Car Features can impact on its Price & Profitability. Where as a Data Analyst we need to find relations between car's features, market category, 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.

Approach:

My first approach is to extract & transform methodology on the dataset provided and follow the below steps.

- The data needs to be cleaned, formatted like removing duplicate values and to check if there are any missing values in the dataset. If there are, decide on the best strategy to handle them.
- After cleaning Perform relevant descriptive statistic calculations to gain a general understanding of dataset.
- This could involve calculating averages, medians, or other statistical measures. It could also involve creating visualizations to better understand the data.

Tech-Stack Used:

I used Microsoft Excel version(Microsoft Office Home and Student 2021) . It provides various tools for organising data, including sorting, filtering, and searching. It is easier to find and analyse specific data points quickly.

Also it has various mathematical functions and formulas that can perform complex calculations and comments on data sets. These functions can summarise data, perform statistical analysis, and more.

Insights:

Data preprocessing/Cleaning

- Starting with Data preprocessing, I found there are 715 duplicate records present dataset. So I removed them accordingly.
- After that I went for Missing Treatment on dataset in which I found 31% of Null (N/A) values were in column “Market Category” as well as below 10 % data were missing in “Engine Fuel Type”, “Engine HP”, “Engine Cylinders” and “Number of Doors” columns.
- I used Mode & Median imputation method for missing data as per Numerical & Categorical variables per below.

Columns	Missing Values	Missing Treatment	Values
Engine Fuel Type	3	Mode	regular unleaded
Engine HP	66	Median	255
Engine Cylinders	30	Median	6
Number of Doors	6	Median	4
Market Category	3376(N/A)	Mode	Crossover

Analysis:

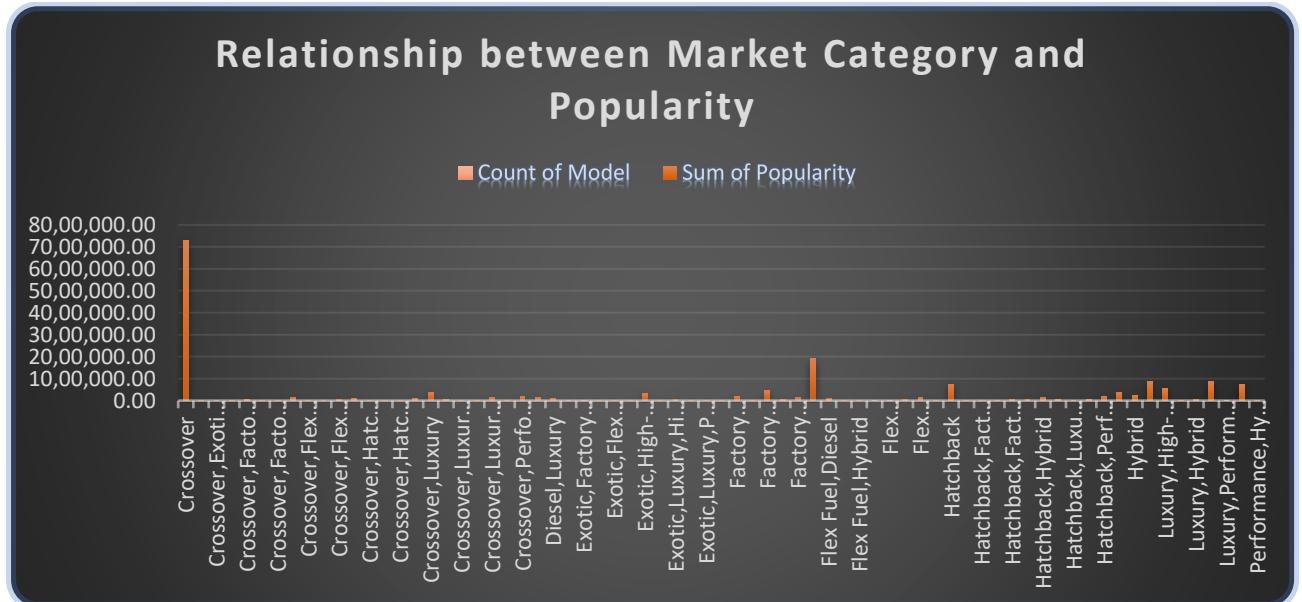
Insight Required: How does the popularity of a car model vary across different market categories?

- **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.

The relationship between Market Category and Popularity Scores shows the “Crossover” has the highest popularity with 7293357 popularity score in Market Category. The Top 5 Market categories with high popularities below.

Market Category	Popularity Score
Crossover	7293357
Flex Fuel	1902985

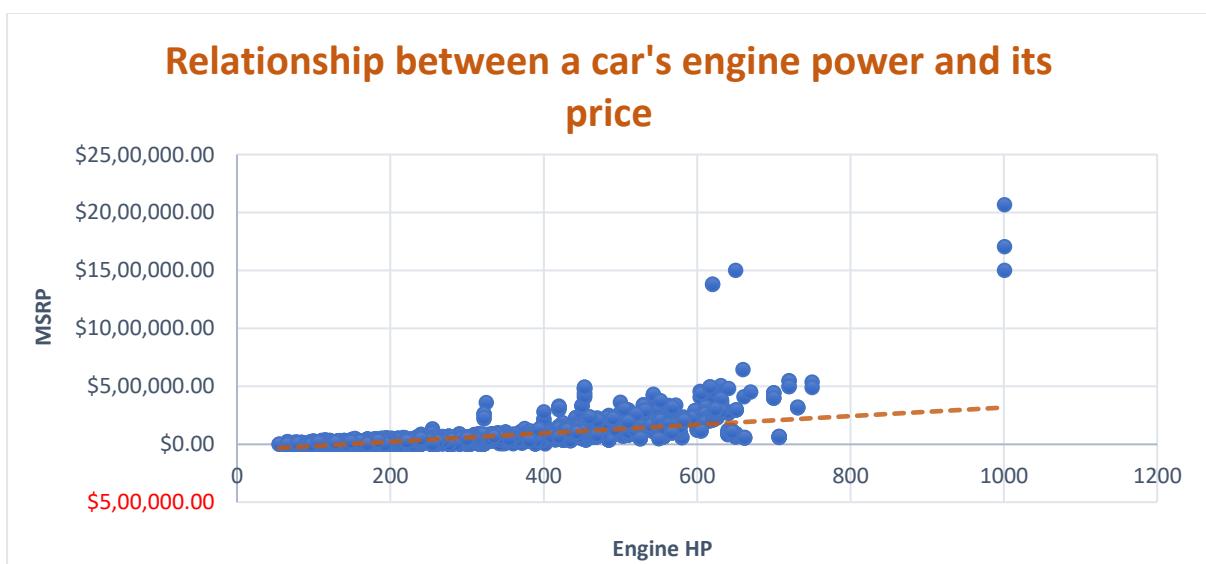
Luxury	883877
Luxury, Performance	852128
Hatchback	751167



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.

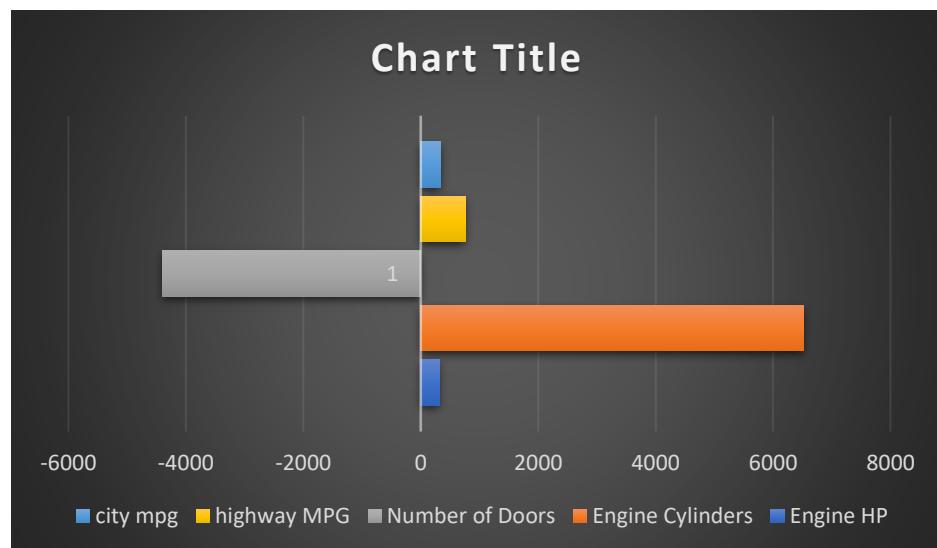
We found the Positive correlation as the **Engine Horsepower** is increasing then Price is also increasing.



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.

In Regression analysis we can see **Number of Doors** has Negative correlation with price while the **Engine Cylinders** have high Positive correlation with Price.
Which shows the [Engine Cylinders](#) are the most important car feature in determining a car's price.



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

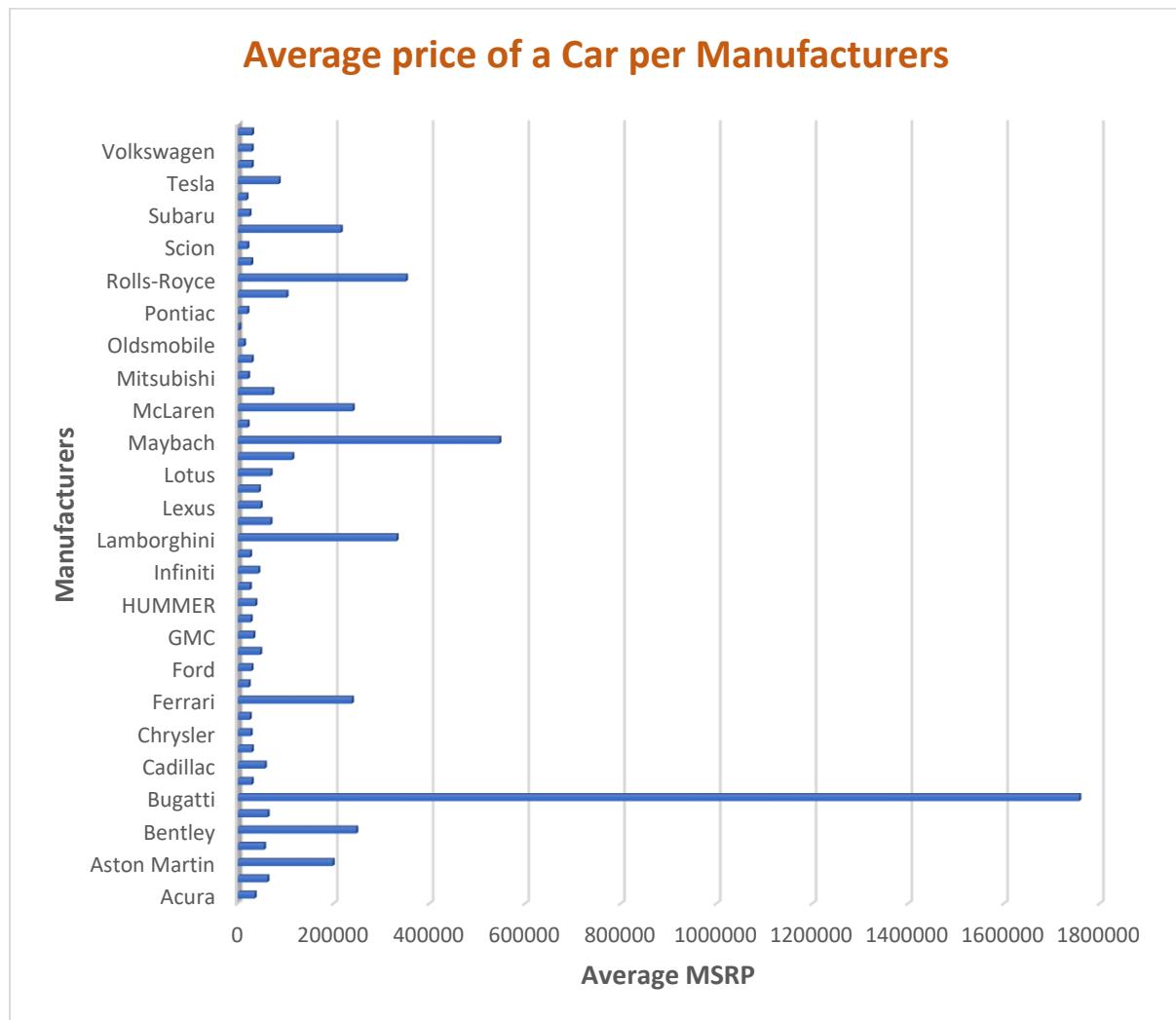
- **Task 4.A:** Create a pivot table that shows the average price of cars for each manufacturer.
- **Task 4.B:** Create a bar chart or a horizontal stacked bar chart that visualizes the relationship between manufacturer and average price.

In terms of average price of a car vary across different manufacturers we see Bugatti has the highest average car price of \$ 17,57,223.67 followed by Maybach, Rolls-Royce, Lamborghini & Bently.

Top 5 Manufacturers:

Manufacturers	Average MSRP
Bugatti	\$17,57,223.67
Maybach	\$5,46,221.88
Rolls-Royce	\$3,51,130.65

Lamborghini	\$3,31,567.31
Bentley	\$2,47,169.32

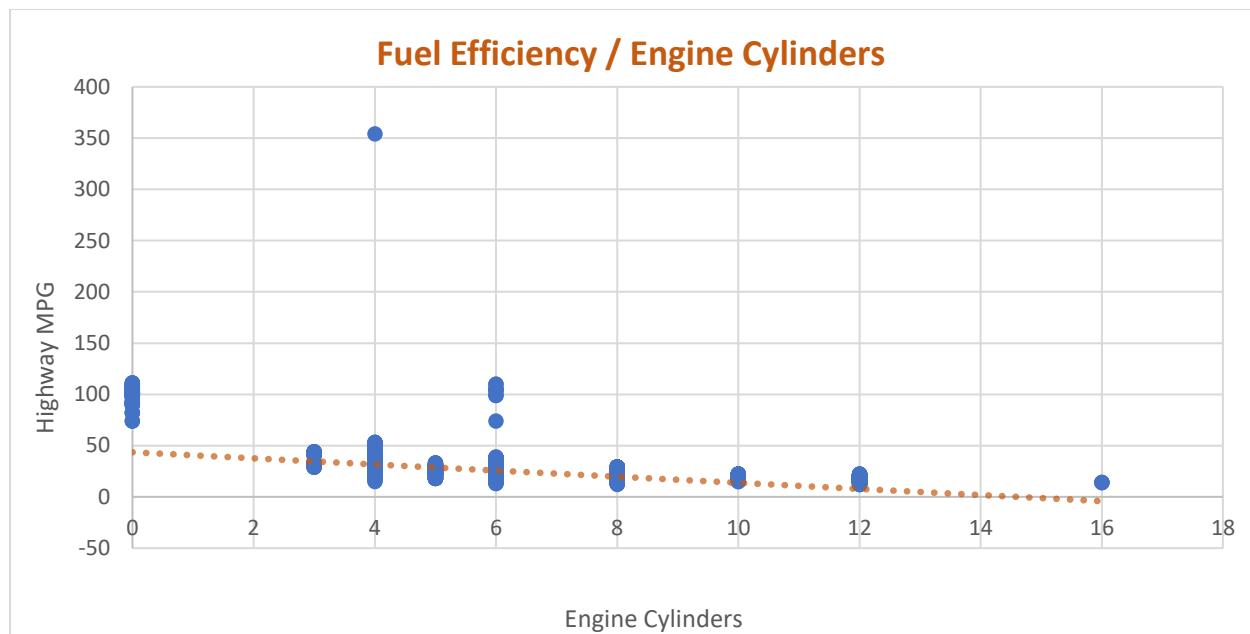


Insight Required: What is the relationship between fuel efficiency and the number of cylinders in a car's engine?

- **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.

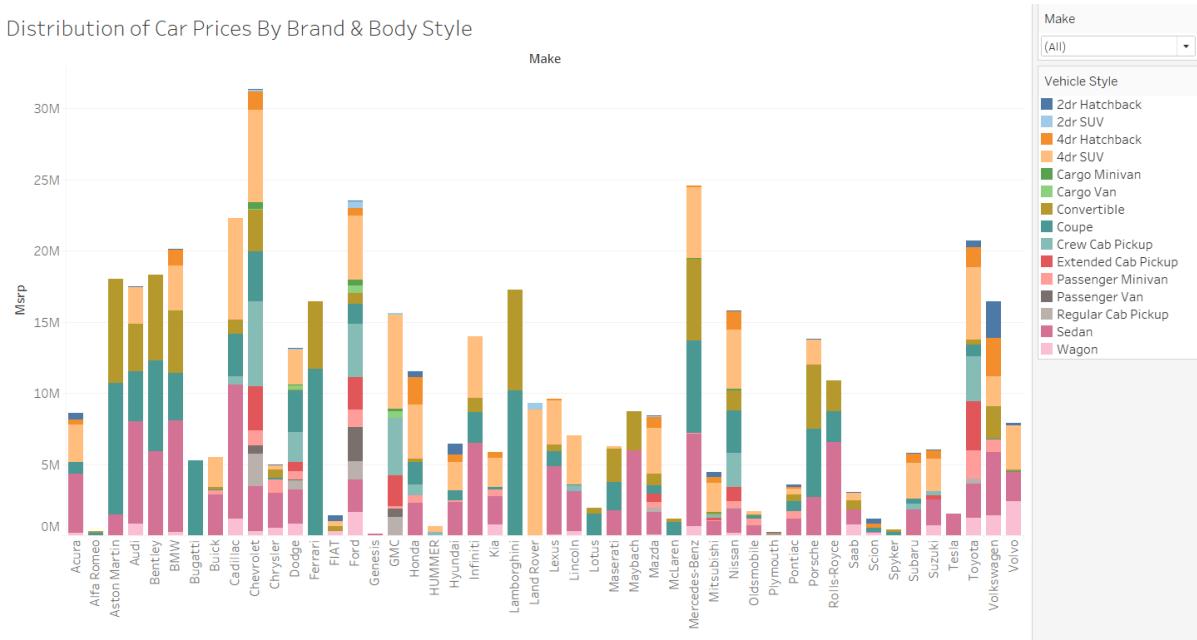
In the relationship between fuel efficiency and the number of cylinders in a car's engine. We see the Negative Correlation between them. As the No. of Engine Cylinders increases the Fuel Efficiency Decreases.

Correlation Coefficient	-0.596246019
-------------------------	--------------



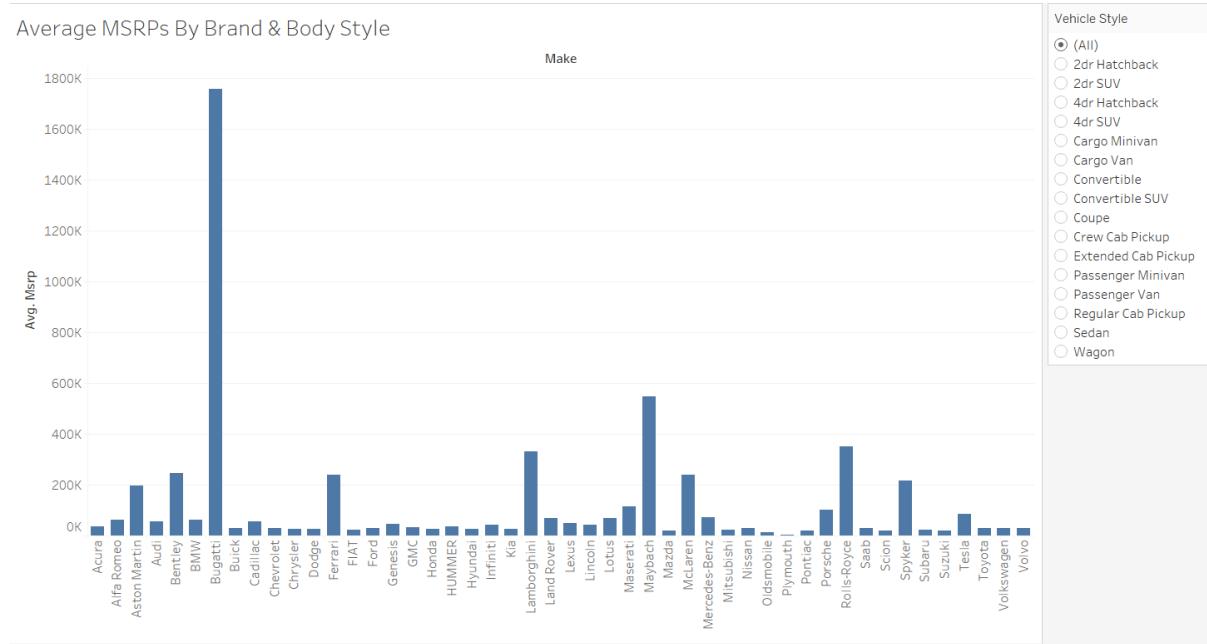
Building the Dashboard:

Task 1: How does the distribution of car prices vary by brand and body style?



Result- As we can see in above chart Chevrolet has the highest price distribution by body style.

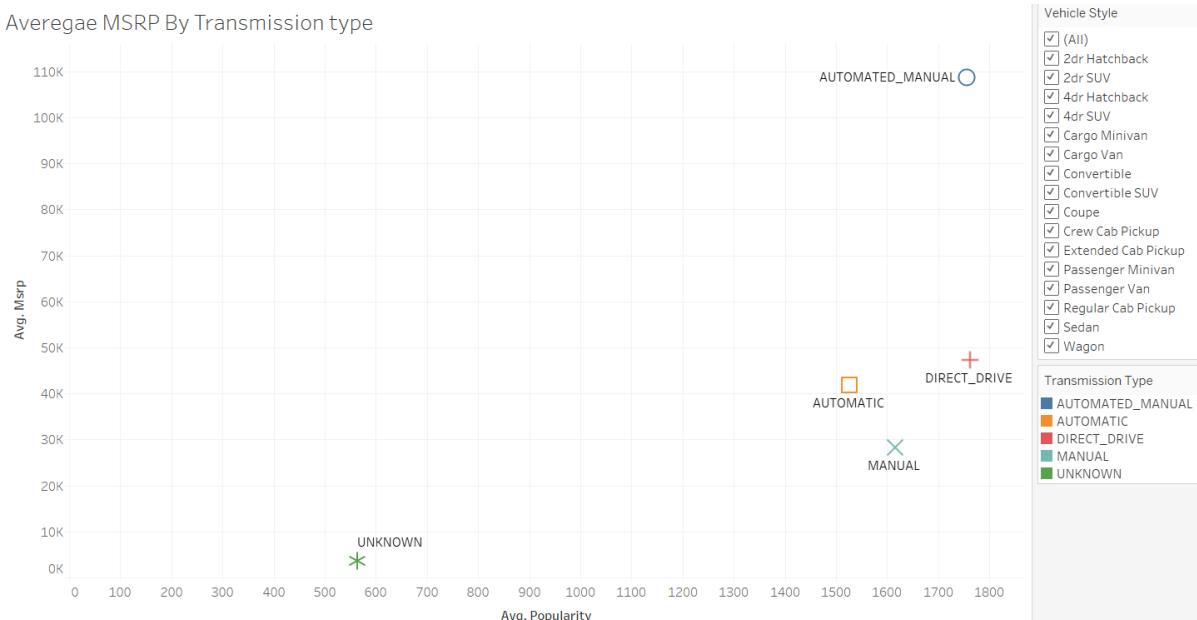
Task 2: Which car brands have the highest and lowest average MSRPs, and how does this vary by body style?



Result- Bugatti has the highest average MSRP and Plymouth has the lowest average MSRP.

Task 3: How do the different feature such as transmission type affect the MSRP, and how does this vary by body style?

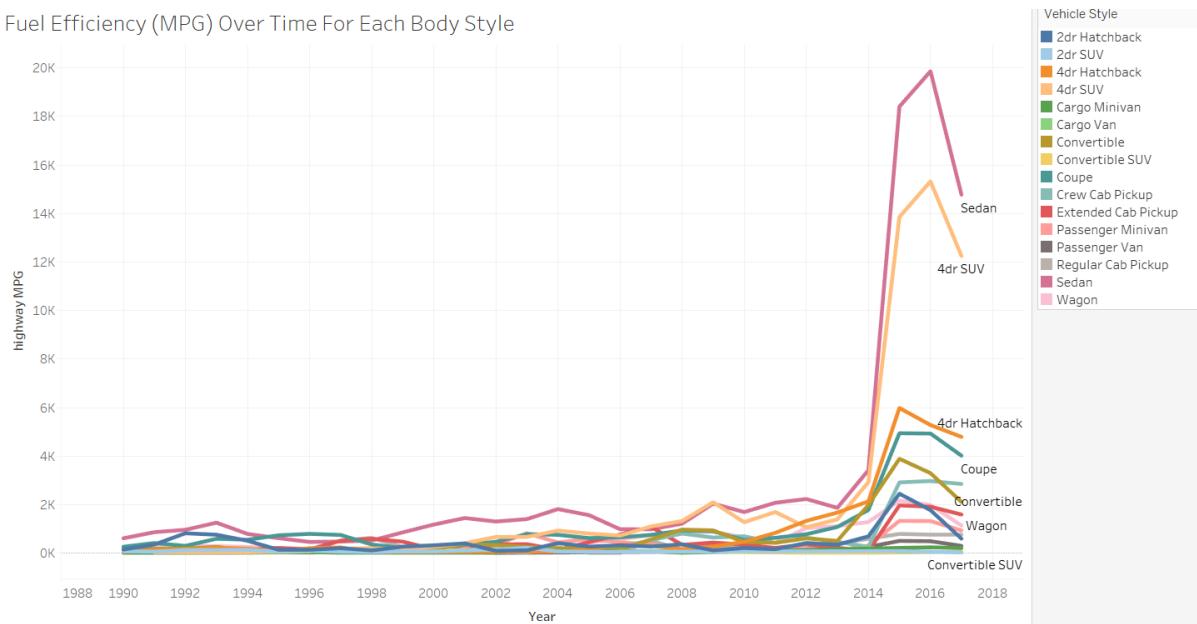
Averegae MSRP By Transmission type



Result- Automated_Manual & Automatic are the highest average MSRP transmission type in Convertible vehicle type. While the Manual is second most popular transmission type.

Task 4: How does the fuel efficiency of cars vary across different body styles and model years?

Fuel Efficiency (MPG) Over Time For Each Body Style



Result- . So in above both the graph shows slight difference in Highway & City MPG's. Also the trend line shows there is continuous increase in MPG in both after 2008 followed by each vehicle type.

Task 5: How does the car's horsepower, MPG, and price vary across different Brands?

