Analysed File URL: CLICK HERE
Video Explanation URL: CLICK HERE

Analysing the Impact of Car Features on Price and Profitability

Shakti prasad Nayak

Description

This project aims to find insights of car features data which was collected from Kaggle.com

Also it aims to find specific statistical business questionaries which will help businesses to identify problems in them and use strategical approach to solve their business problems

For this project we have used data cleaning methodologies to clean our data for achieving precise results

We also performed imputation method to heal blank cells with most accurate data

Approach

Data Cleaning

In this step we have analysed blank cells and performed imputation using mean, median, mode method and finally we have removed duplicates

First Step

Second Step

Detecting Outliers

In this step we have detected outliers of some specific columns as the outliers affects analysis

Analysing Data

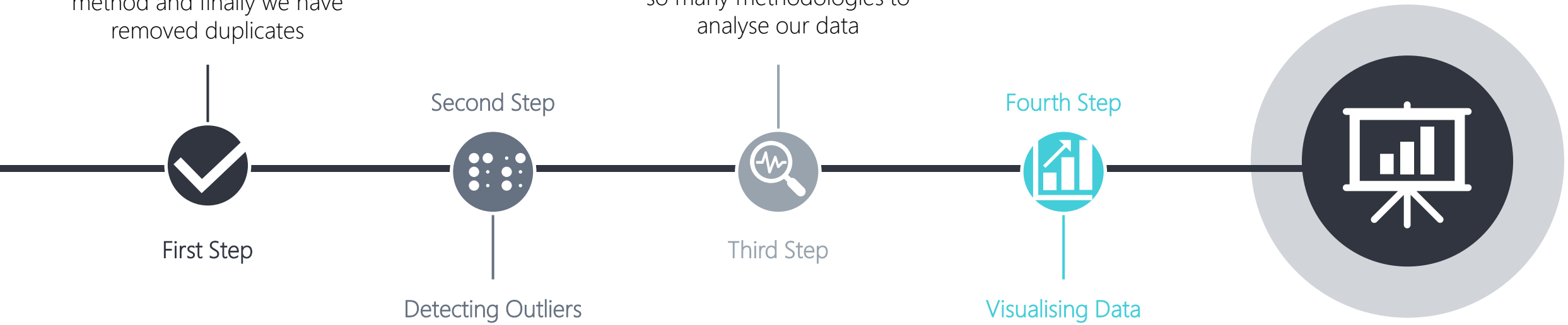
We have used Excel functions, pivot tables and so many methodologies to analyse our data

Third Step

Fourth Step

Visualising Data

We have used different charts like scatter plot, box plot, bubble chart to analyse our data. We also used slicers to filter our data for desired results.



Tech Stack Used-MS EXCEL

Formulas and functions

You can use formulas and functions to perform calculations and operations on your data. Formulas are expressions that you type in a cell, while functions are predefined formulas that you can choose from a list

Pivot tables and pivot charts

You can use pivot tables and pivot charts to summarize and analyse large amounts of data in a dynamic and interactive way. You can arrange your data into rows, columns, filters and values, and then change the layout and format of your pivot table or pivot chart. You can also use slicers and timelines to filter your data and see different views.

Charts and graphs

You can use charts and graphs to visualize your data and show trends, patterns and relationships. You can choose from different types of charts and graphs, such as pie, bar, line, scatter and more.

Data analysis tools

You can use data analysis tools to perform advanced statistical and analytical tasks on your data. Some of the data analysis tools available in MS Excel are: Solver, which can help you find optimal solutions for problems that involve constraints; Data Table, which can help you perform what-if analysis by changing one or more variables and seeing how they affect the outcome

Insights

- ✓ The average popularity is highest for market categories Crossover, Flex Fuel, Performance; Flex Fuel, Diesel; Hatchback, Flex Fuel.
- ✓ For each increment of number of cylinders the predicted MSRP increases by 18446.
- ✓ For each increment of Engine HP predicted MSRP increases by 369.09.
- ✓ For each increment of number of Cylinders predicted Highway MPG decreases by 3.0129.
- ✓ Sedan vehicle style has the highest sum of MSRP's across all vehicle styles.
- ✓ Genesis and plymouth has the lowest sum of MSRPs across all brands.
- ✓ Convertible SUV has the lowest sum of MSRP's across all vehicle styles.
- ✓ For Coupe vehicle style Ferrari has the highest sum of MSRPs across all brands.
- ✓ Chevrolet brand has the highest sum of MSRP's across all brands.
- ✓ For Sedan vehicle style Cardillac brand has the highest sum of MSRP's across all brands.
- ✓ For Cardillac brand sedan vehicle style contriibutes higher.

Insights

- ✓ For Chevrolet brand 4dr SUV vehicle style contributes higher.
- ✓ For 4dr SUV vehicle style Land Rover contributes higher in sum of MSRP's.
- ✓ The Bugatti brand has the highest average MSRP across all the brands.
- ✓ Convertible vehicle style has the highest average MSRP across all vehicle styles.
- ✓ Bugatti brand has only one vehicle style that is Coupe.
- ✓ For Convertible vehicle style, Maybach brand contributes higher average MSRP across all brands.
- ✓ For Coupe vehicle style, Bugatti brand contributes highest average MSRP across all brands.
- ✓ For Sedan vehicle style Maybach contributes highest average MSRP across all brands.
- ✓ For Maybach brand, Convertible vehicle style contributes highest average MSRP across all vehicle style.

Insights

- ✓ Coupe vehicle style has the highest average MSRP for Automated Manual Transmission type.
- ✓ For Automatic transmission type Convertible vehicle style has the highest average MSRP.
- ✓ For Direct Drive transmission type Sedan vehicle style has the highest average MSRP.
- ✓ For Manual transmission type Convertible vehicle style has the highest average MSRP.
- ✓ The overall decrement of efficiency is seen on 1991,1995,1999,2006,2009.
- ✓ Overall increment of efficiency is seen on 2017.
- ✓ For Sedan vehicle style highest efficiency is seen on 2016.
- ✓ For Convertible SUV highest efficiency is seen on 2017.
- ✓ For 4dr SUV vehicle style highest efficiency is seen on 2016.
- ✓ For Convertible vehicle style highest efficiency is seen on 2017.
- ✓ For Coupe vehicle style average MPG on 1992 is nearly same as average MPG on 2017.

Insights

- ✓ For Chevrolet brand most of it's models, average MPG falls below 40.
- ✓ For Bugatti brand average MPG is at 14 while average HP is at 1001 and average MSRP is 1757223 which is the highest average MSRP.
- ✓ For Genesis brand average MPG is at 25 while average HP is at 347 and average MSRP is 46616.

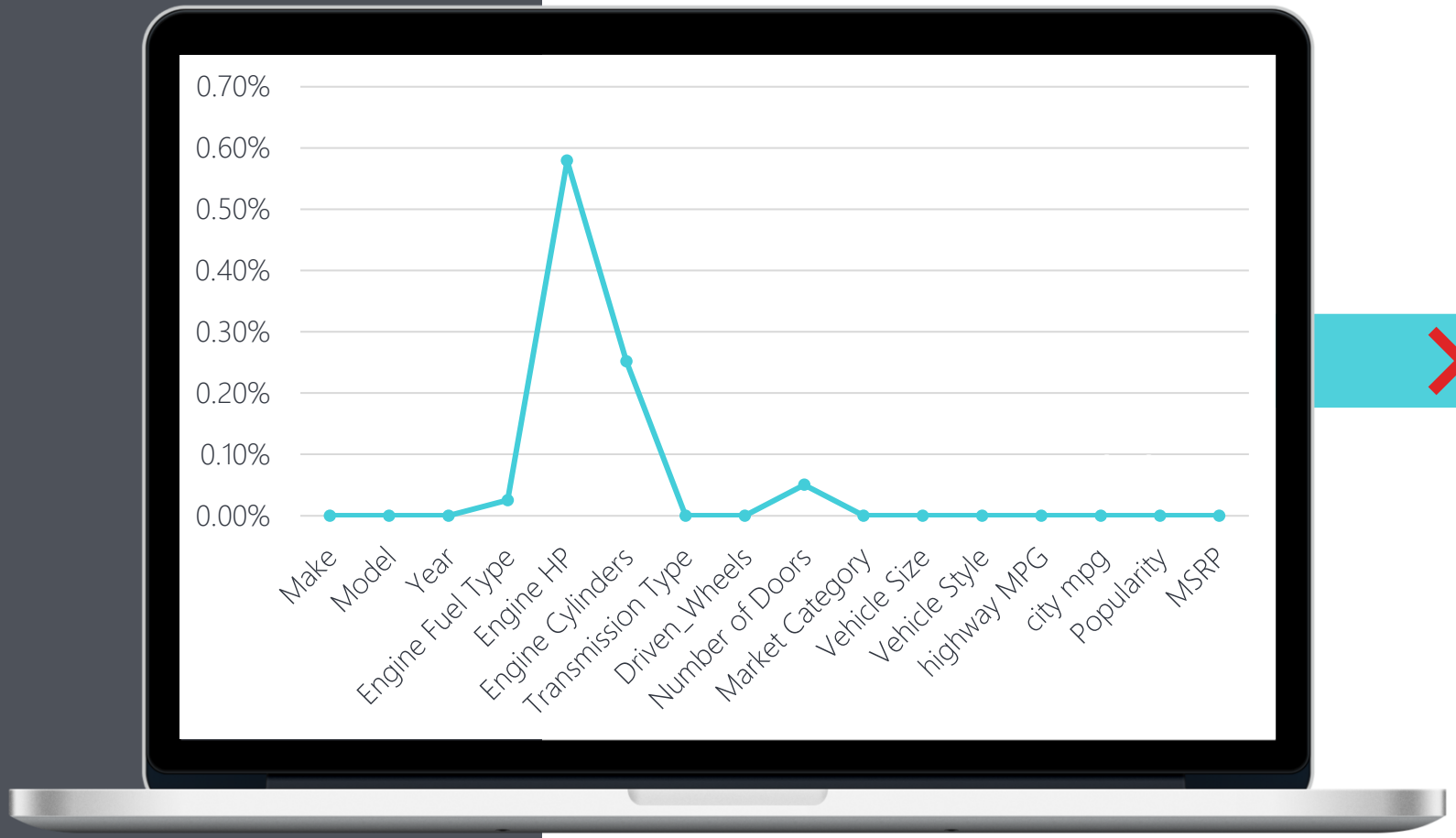
Results

Blank Cells

The graph shows percentage of blank cells in the corresponding columns.

Since the percentage of blank cells are less than 15% we can impute them with mode or median values.

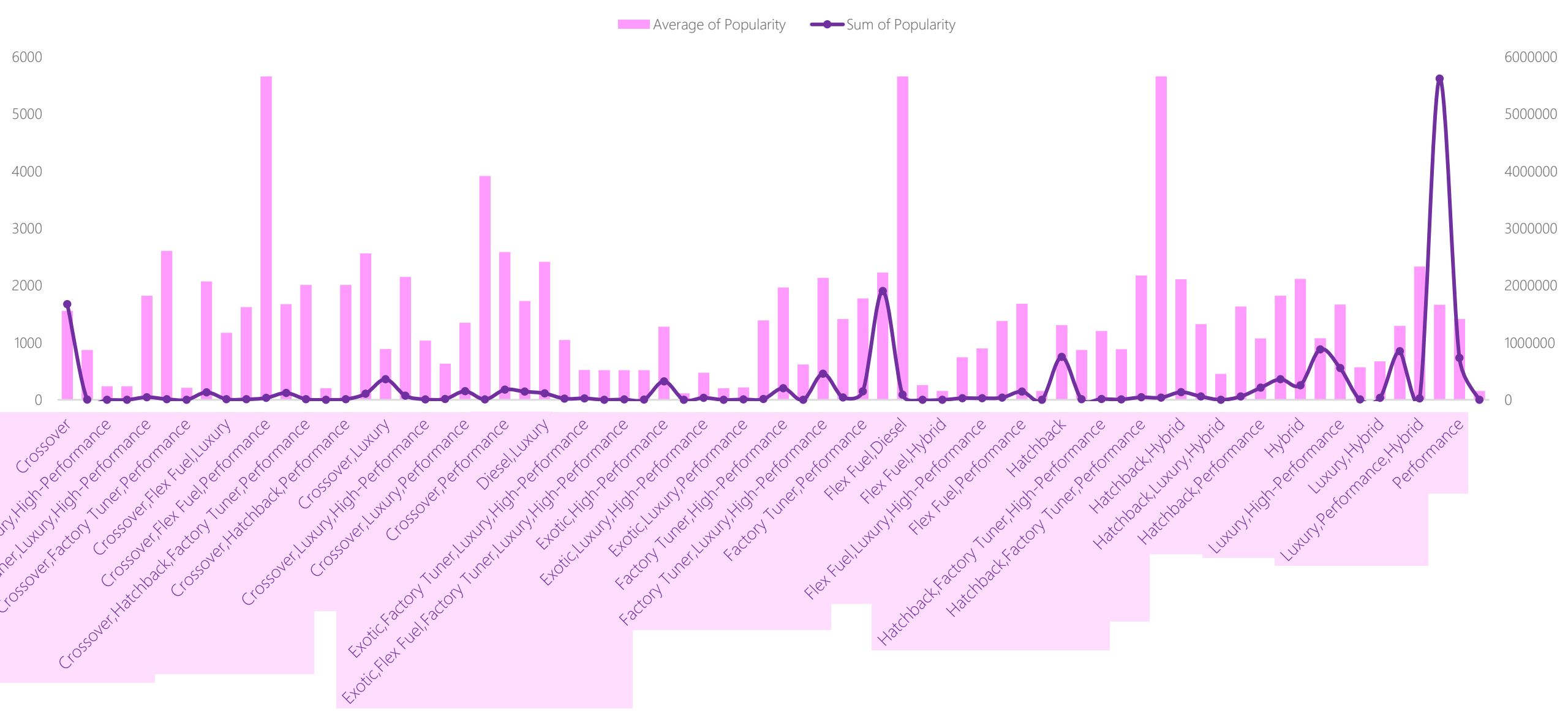
Engine Fuel Type: Regular Unleaded
Engine HP: 239
Engine Cylinders: 4
Number of doors: 4



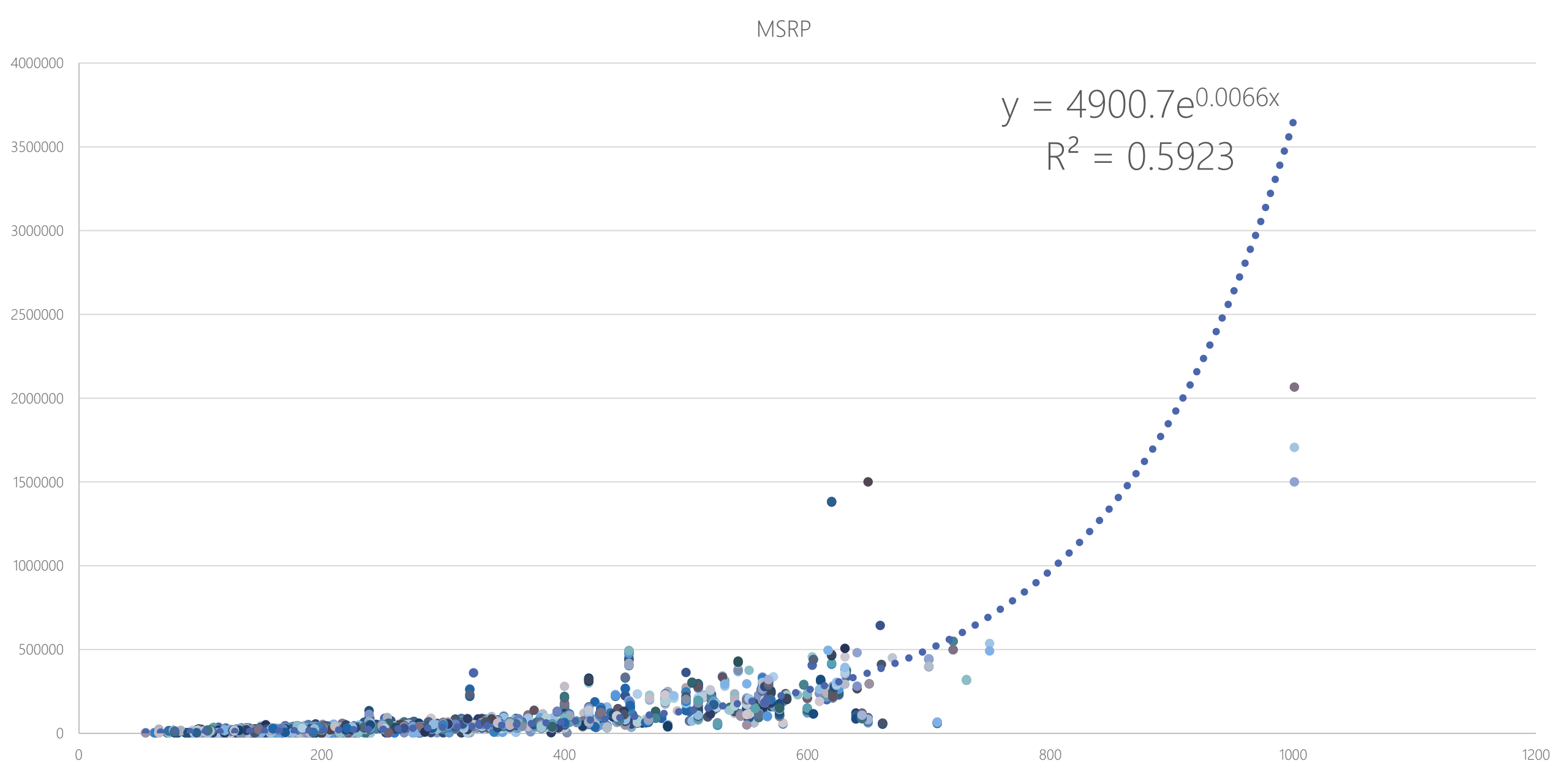
Identifying Outliers

The two graphs shows box plots of the numerical columns Engine HP and MSRP. The outliers are shown in the dots and the mean value is shown as cross mark. It's a good idea to remove the outliers from our dataset but we are keeping the outliers for effective and precise results.

Your logo



The Above Graph Shows the Relationship between Market Category and Population Score



Relationship between car's engine power and it's price:

This shows the positive correlation between two variables i.e. engine power and price. The R^2 value and exponential equation is labelled above.

Car's Price Regression Analysis

Correlation	Number of Doors	Engine Cylinders	Engine HP	highway MPG	city mpg	Popularity	MSRP
Number of Doors	1	-0.150814073	-0.129216667	0.115080984	0.121012512	-0.057212828	-0.144353121
Engine Cylinders	-0.150814073	1	0.774299381	-0.602959305	-0.572117031	0.040031449	0.538580448
Engine HP	-0.129216667	0.774299381	1	-0.362951802	-0.355036126	0.041946841	0.658692783
highway MPG	0.115080984	-0.602959305	-0.362951802	1	0.886298533	-0.01715925	-0.166631495
city mpg	0.121012512	-0.572117031	-0.355036126	0.886298533	1	-0.000548917	-0.162342905
Popularity	-0.057212828	0.040031449	0.041946841	-0.01715925	-0.000548917	1	-0.048370622
MSRP	-0.144353121	0.538580448	0.658692783	-0.166631495	-0.162342905	-0.048370622	1

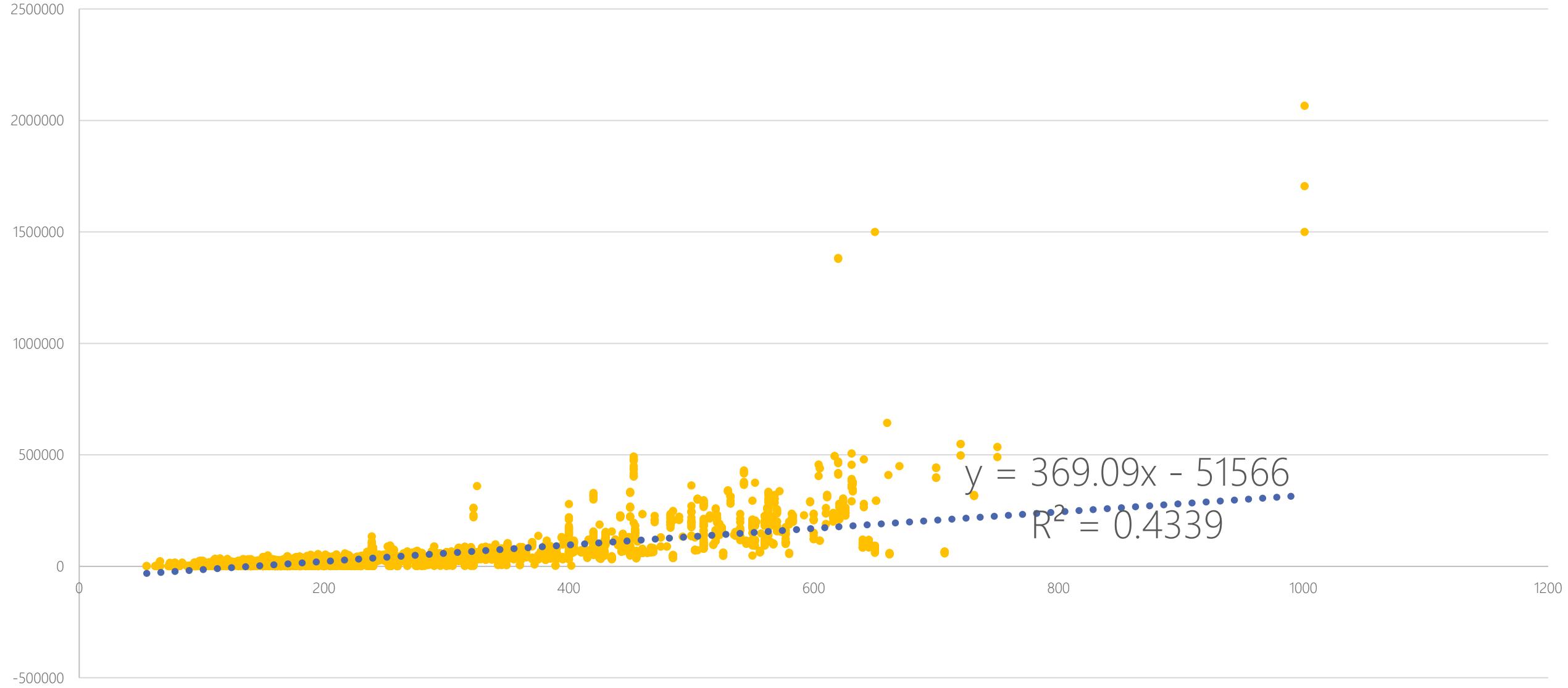
The above table shows correlation coefficient between different numerical columns from our dataset using CORREL() function.

As we can see from the above table that Engine Cylinders and Engine HP has the highest correlation with MSRP i.e. price.

Now, we can find Regression line and equation by plotting a scatterplot between MSRP on Y-axis and other column on X-axis.

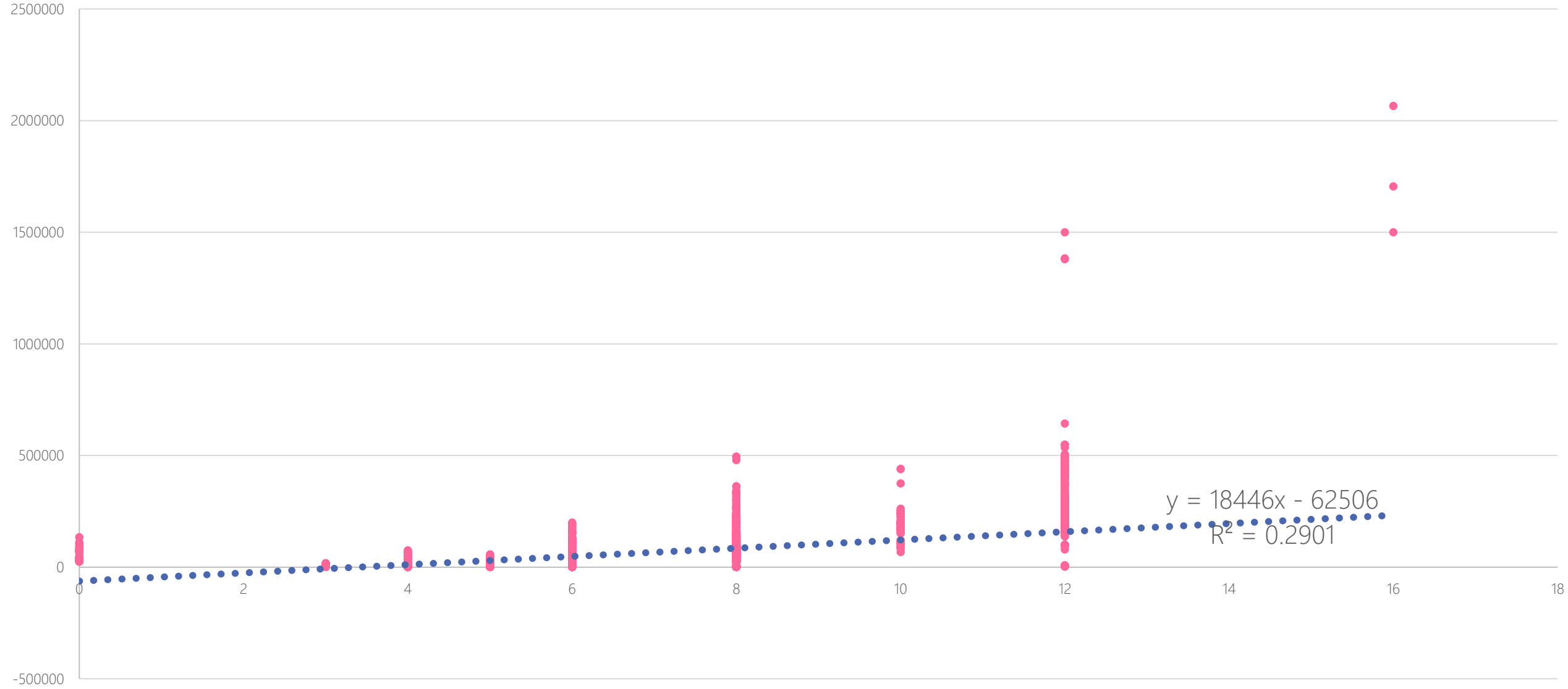
This graph shows the correlation of MSRP with different variables.
From this graph we can conclude that only Engine HP and Engine Cylinders are positively correlated with MSRP and other variables are negatively correlated with MSRP.

MSRP Vs Engine HP



The above graph shows relationship between price of the car and engine horse power

MSRP Vs Engine Cylinders



The above graph shows relationship between price of the car and engine cylinders

Multiple Regression Output

We have formulated an regression equation from the summary of regression analysis based on following formula

Predicted MSRP="Intercept"+("Coefficient"*"Number of Doors")+("Coefficient"*"Engine Cylinders")+("Coefficient"*"Engine HP")+("Coefficient"*"highway MPG")+("Coefficient"*"city mpg")+("Coefficient"*"Popularity")

The equation is-

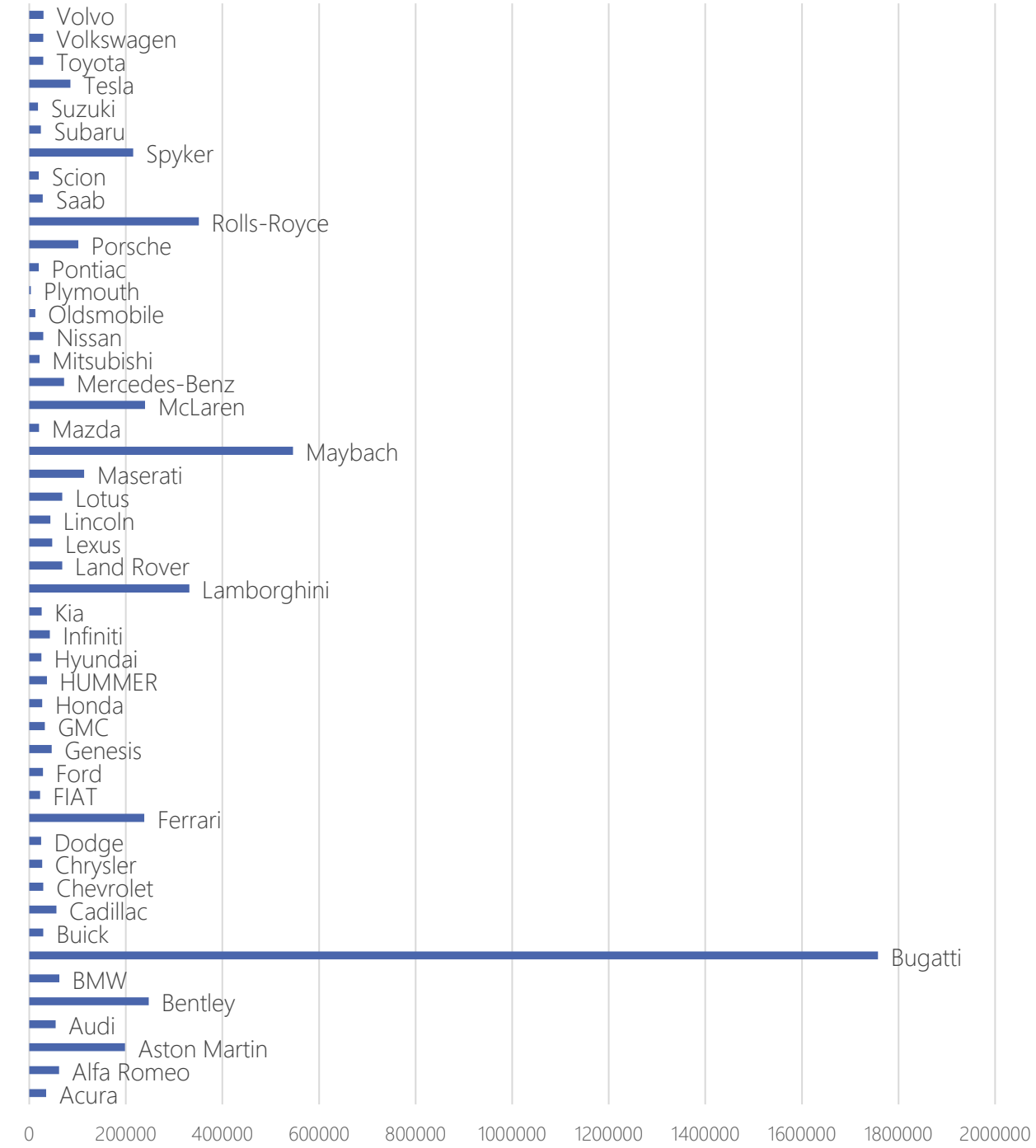
Predicted MSRP=-82094.8628617037+(-4697.29639835653*"Number of Doors")+(6743.09485652704*"Engine Cylinders")+(315.015624509003*"Engine HP")+(716.959632335862*"highway MPG")+(437.180753030329*"city mpg")+(-3.48246979940141*"Popularity")

	Coefficients
Intercept	-82094.8629
Number of Doors	-4697.2964
Engine Cylinders	6743.094857
Engine HP	315.0156245
highway MPG	716.9596323
city mpg	437.180753
Popularity	-3.4824698

The adjusted R^2 value is 0.46 i.e. 46% of values fits into this model

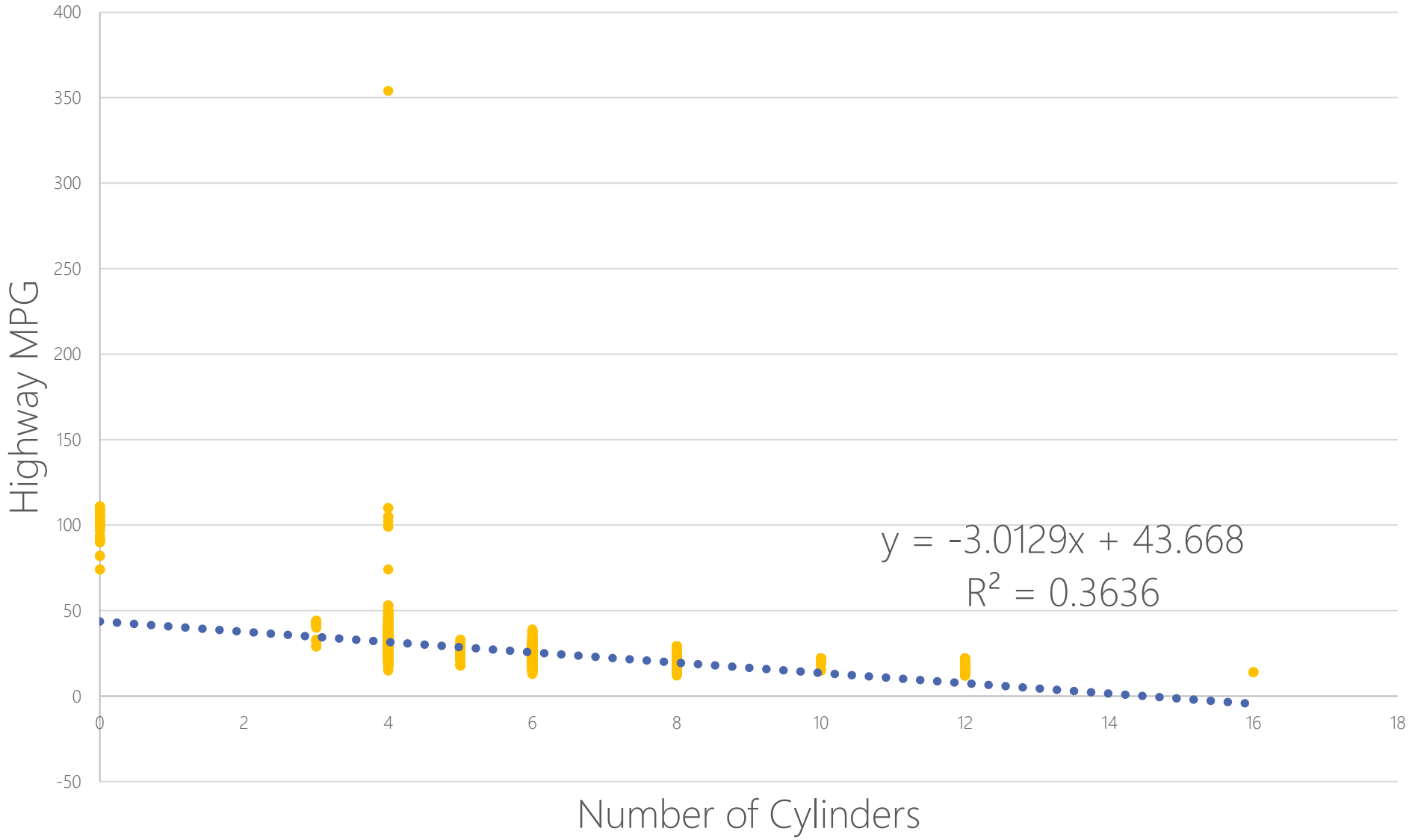
Average price of cars for different Manufactures

Row Labels	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	238218.8406
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	85255.55556
Toyota	28846.5605
Volkswagen	28978.52289
Volvo	29724.68421



The Graph Shows Average price of cars for different Manufactures

Relationship between fuel efficiency and the number of cylinders in a car's engine



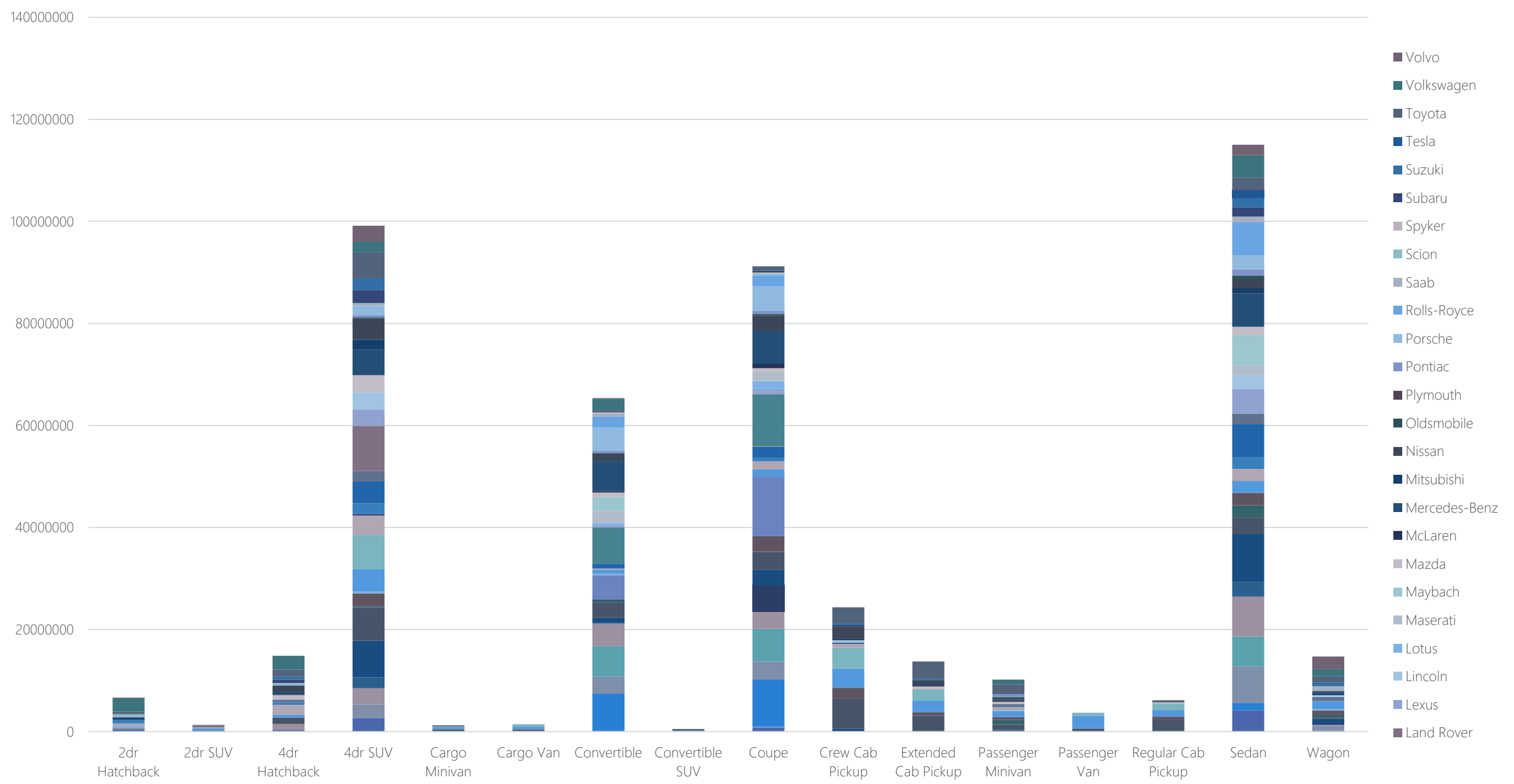
Relationship between fuel efficiency and the number of cylinders in a car's engine

The slope between Highway MPG and Engine Cylinders is -3.0128681379476

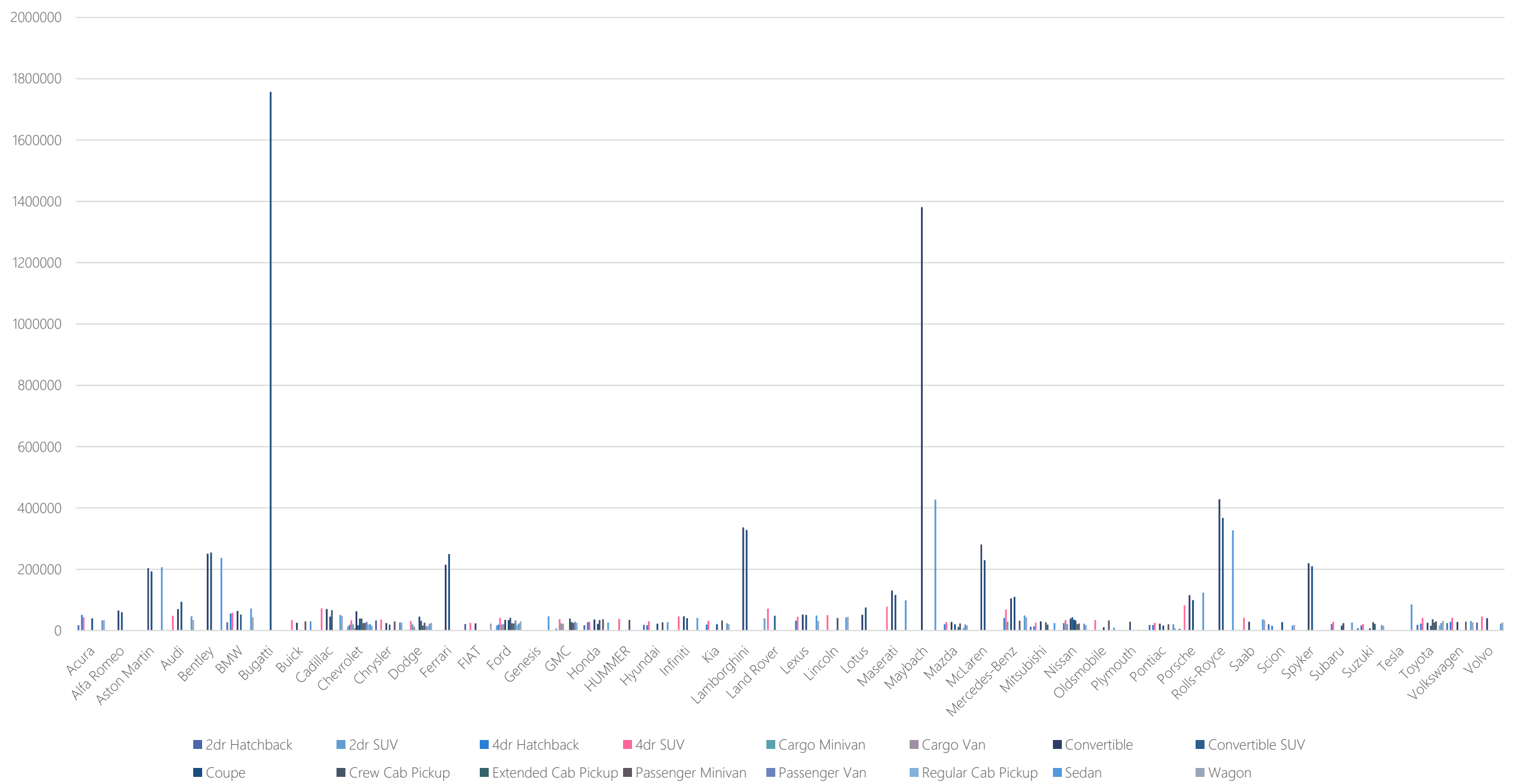
The correlation coefficient between two variables is -0.602959304615793

Since, the variables are negatively correlated we can say that when the number of cylinders increases in car Highway MPG goes down.

The slope indicates that for each increment of number of cylinders how much Highway MPG goes down.



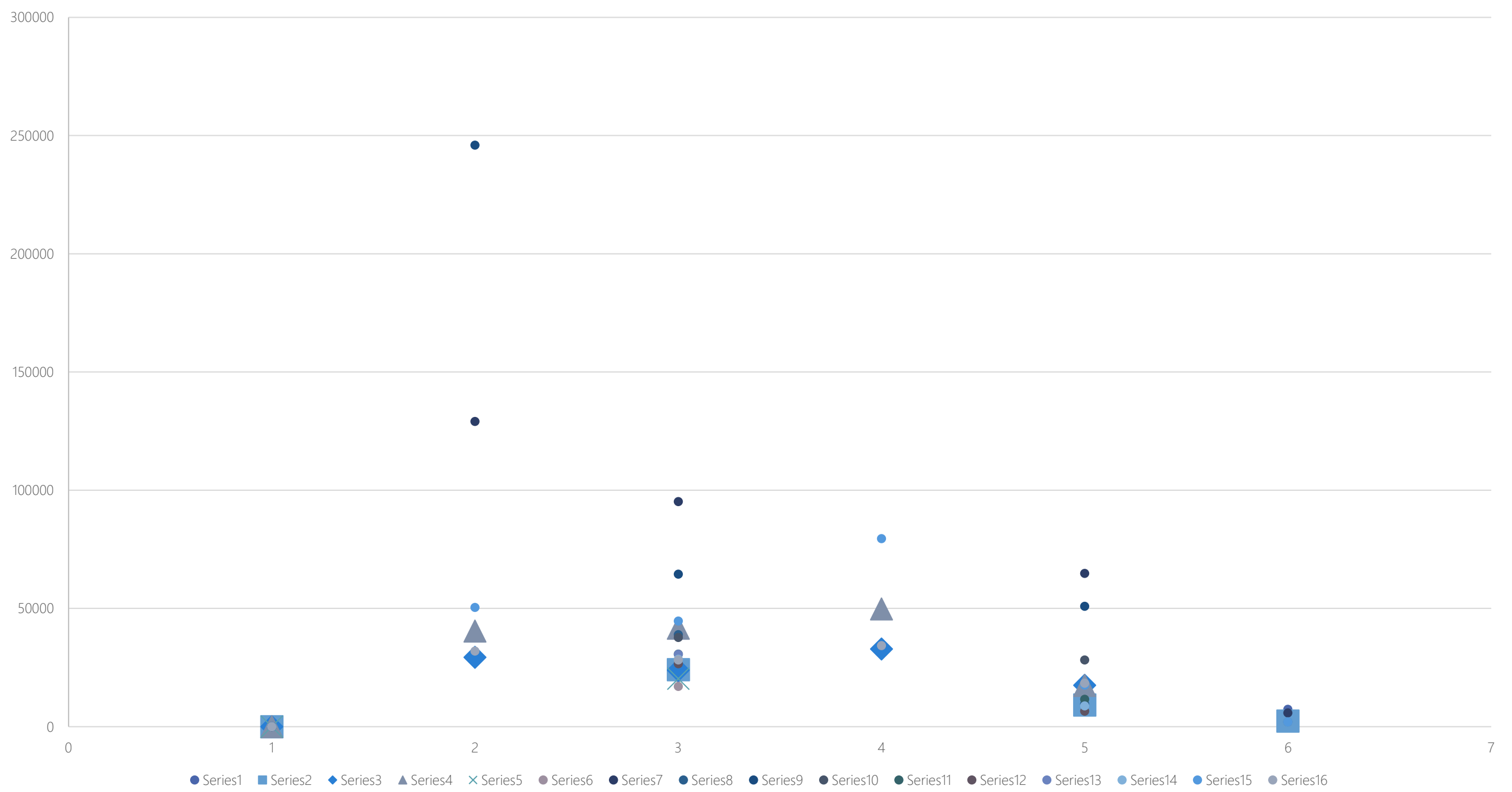
Distribution of car prices vary by Brand and vehicle style



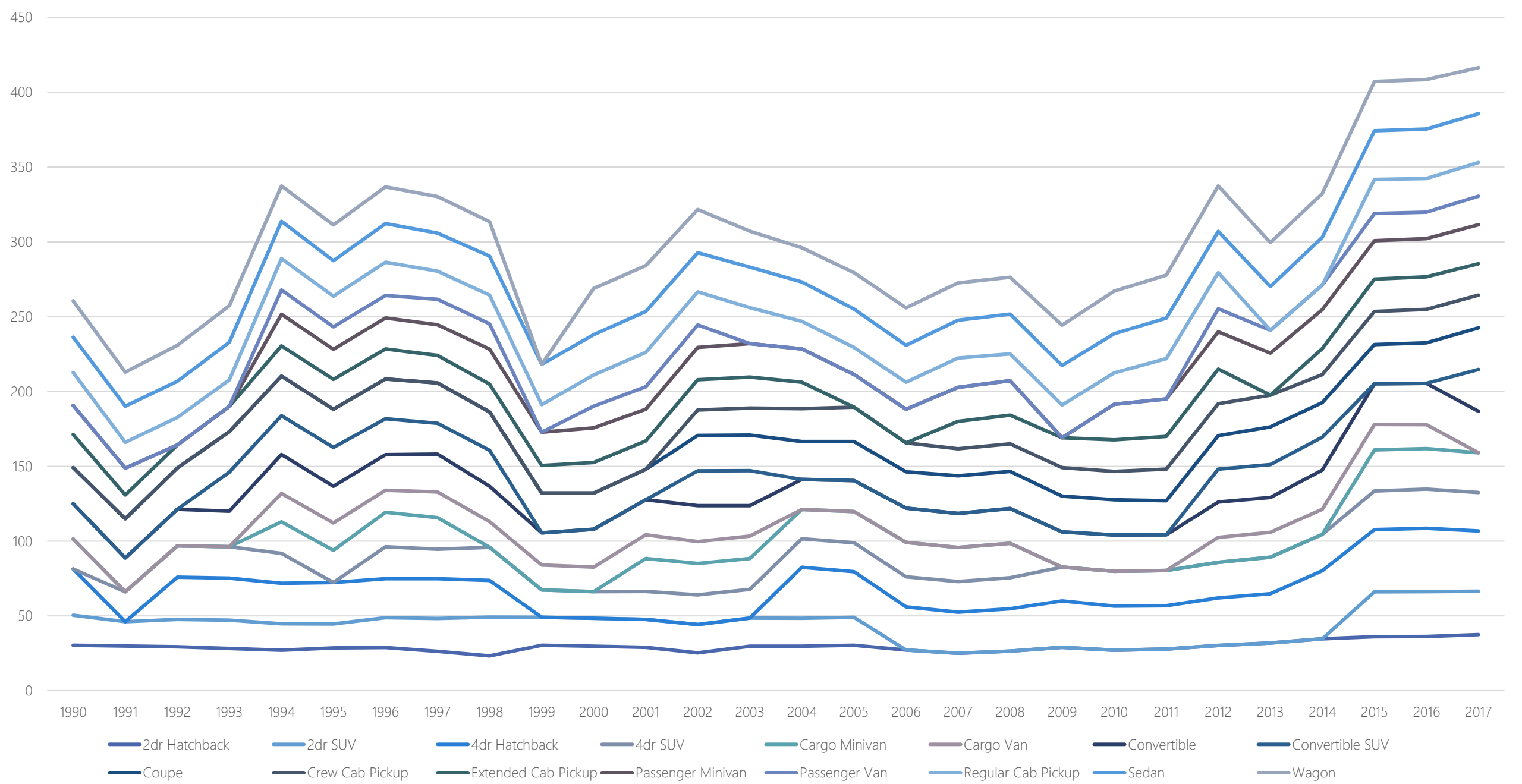
Average MSRPs across different car brands and vehicle styles

Row Labels	AUTOMATED_MANUAL	AUTOMATIC	DIRECT_DRIVE	MANUAL	UNKNOWN
2dr Hatchback	27470.41667	20784.09901	31800	12840.65556	7361.5
2dr SUV		24153.60606		9173.018519	2371
4dr Hatchback	29347.04545	23888.73529	32799.72973	17500.36364	
4dr SUV	40451.15385	41638.26534	49800	17422.08791	
Cargo Minivan		20315.59322			
Cargo Van		17019.29762			
Convertible	129082.2339	95153.3131		64794.34437	5783.5
Convertible SUV		38925.5		9594.8	
Coupe	245977.4252	64523.41955		50901.4973	2000
Crew Cab Pickup		37718.95307		28233.10811	
Extended Cab Pickup		30711.45251		11553.29707	
Passenger Minivan		26589.50919		6510	
Passenger Van		30578.06612			
Regular Cab Pickup		28536.8239		8759.454054	2000
Sedan	50385.39326	44671.35638	79512.25	17557.26441	2000
Wagon	31985.27778	28219.45742	34250	18398.57813	

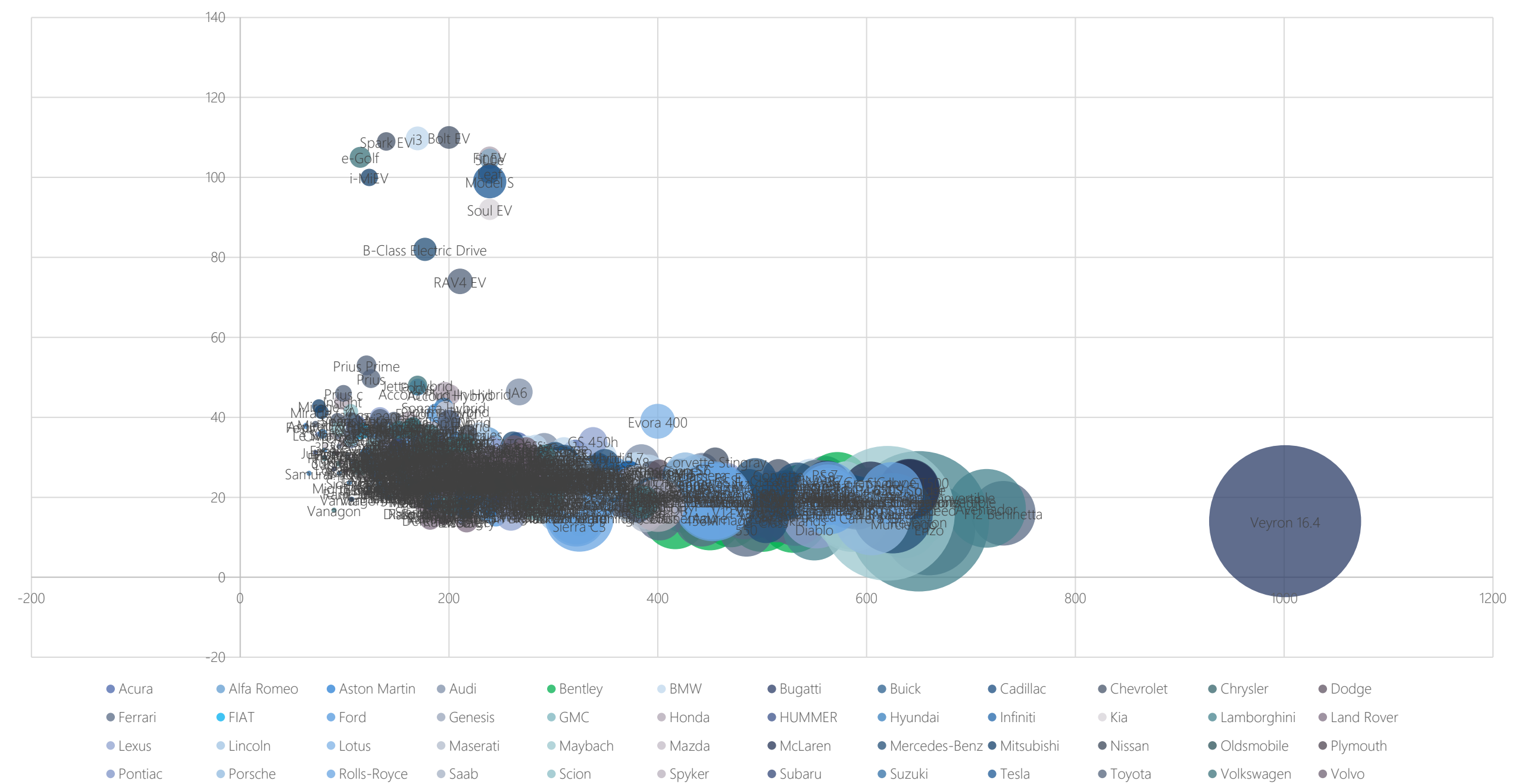
Relationship between MSRP and transmission type for each body style



Relationship between MSRP and transmission type for each body style; 1,2,3.. In x axis represents corresponding transmission type



Fuel efficiency of cars vary across different body styles and model years



Relationship between horsepower, MPG, and price across different car brands


THANK YOU