



# IMPACT OF CAR FEATURES ON PRICE AND PROFITABILITY

Data Analysis Project 7

Ritika Chaudhary

# Contents

- Project Overview
- Tech-Stack Used
- Tasks: Analysis
- Building the Dashboard
- Excel Sheet Drive Link
- Result

# PROJECT OVERVIEW

The automotive industry, characterized by rapid innovation and evolving consumer preferences, faces the ongoing challenge of optimizing pricing and product development strategies. As the industry undergoes shifts towards eco-friendly and technologically advanced vehicles, understanding consumer demand, market dynamics and the impact of car features on pricing has become paramount. In light of this, our data analysis project aims to address the following business problems:

How can a car manufacturer optimize pricing and product development decisions to maximise profitability while meeting consumer demand?

To tackle this question, we will leverage data driven insights and analytical techniques to explore the relationship between car features, market categories and pricing. Our goal is to identify the most popular and profitable product attributes, enabling manufacturers to make informed decisions that enhance competitiveness and profitability. Through this project, we aim to provide actionable recommendations for the automotive industry's pricing and product development strategies.

# APPROACHES

- **DATA CLEANING:** Removal of duplicates, rows containing blank cells.
- **DESCRIPTIVE STATISTICS:** Descriptive statistics served as our starting point, providing a comprehensive overview of the dataset. Measures such as means, median and standard deviations allowed us to grasp the central tendencies and dispersions of key variables.
- **DATA VISUALIZATION:** Visualisation techniques including bar charts, scatter plots and pivot tables, were instrumental in presenting complex data in an interpretable format. These visualizations enhanced our ability to uncover patterns, trends and relationships within the data.
- **REGRESSION ANALYSIS:** Regression analysis emerged as a powerful tool to explore the relationship between variables. By conducting regression analysis we identified the factors most strongly associated with the car prices, offering invaluable insights into pricing determinants.

# TECH-STACK USED



## **SOFTWARE USED:** MICROSOFT EXCEL

Our primary tool for data analysis and visualisation throughout this project was Microsoft Excel. Excel's versatility and user-friendly interface made it the ideal choice for handling, exploring and visualising the data set. Excel's PivotTables, Pivot Charts, charting capabilities were particularly instrumental in generating meaningful insights.

TASKS: ANALYSIS

A red toy car is positioned on top of two stacks of gold coins. The car is a small-scale model, possibly a sports car, with a sleek design. The coins are stacked in two separate piles, one on the left and one on the right, with the car's front wheel resting on the left stack. The background is a soft, out-of-focus gradient of light green and yellow, suggesting an outdoor setting. The overall image conveys a sense of value and investment in the automotive market.

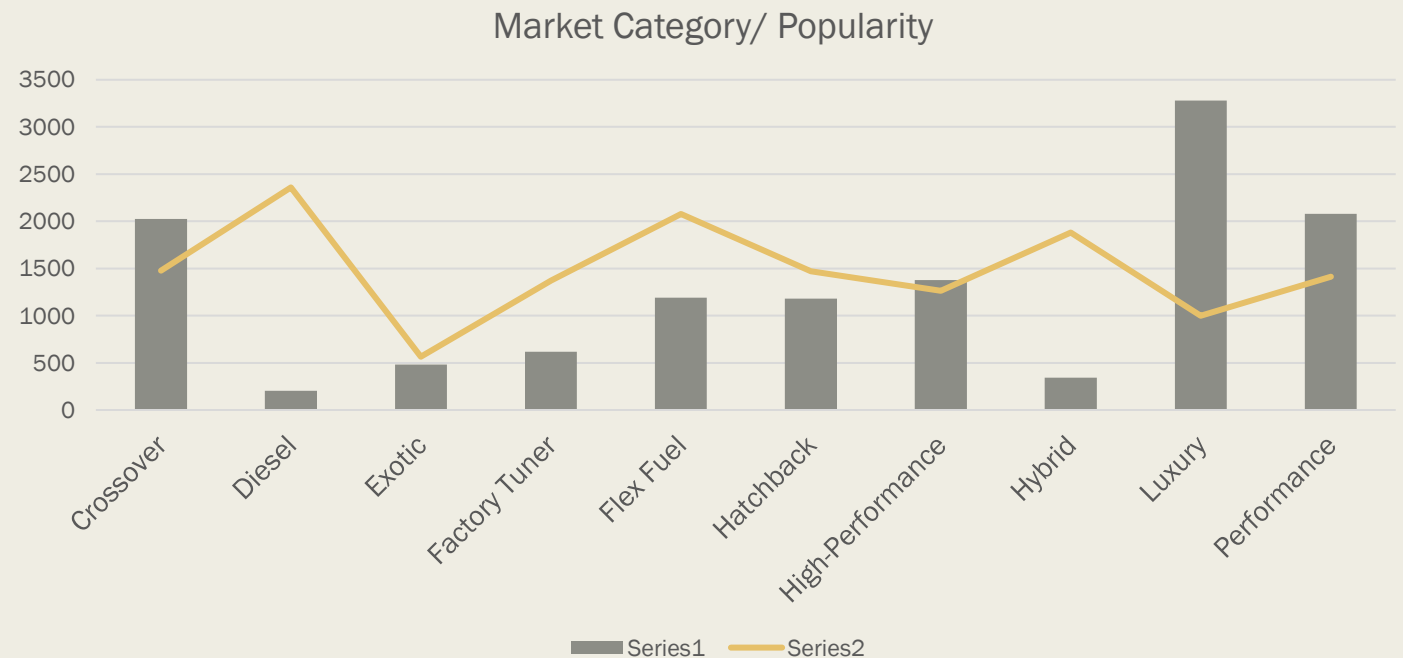
**HOW DOES THE POPULARITY OF A  
CAR MODULE VARY ACROSS THE  
MARKET CATEGORIES?**

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.

Multiple categories in a single column so after splitting got the following table and chart.

Market Category	Count	Popularity
Crossover	2026	1477.90
Diesel	205	2358.02
Exotic	482	563.91
Factory Tuner	618	1371.33
Flex Fuel	1191	2078.49
Hatchback	1181	1471.07
High-Performance	1378	1262.51
Hybrid	343	1881.04
Luxury	3279	997.50
Performance	2078	1413.01



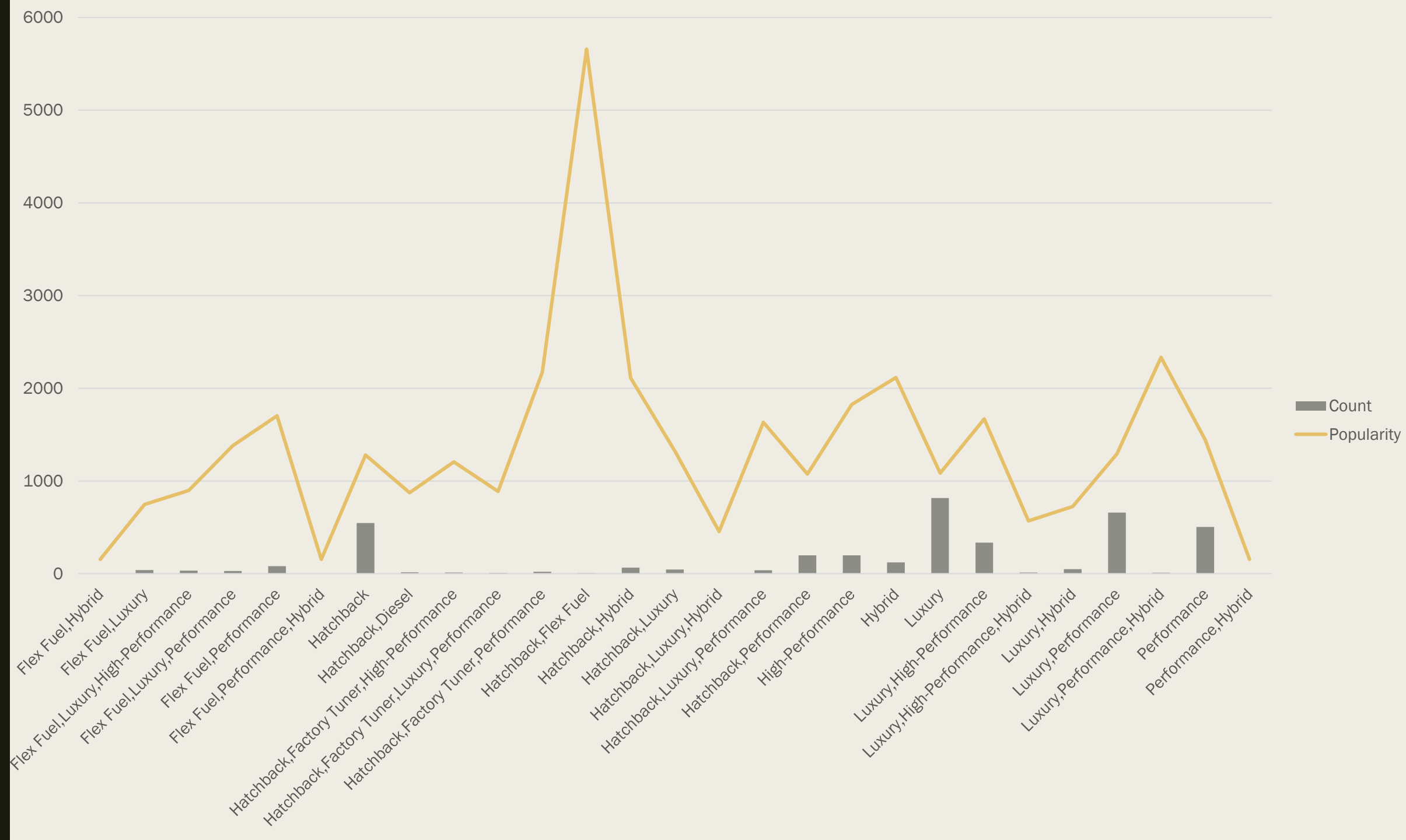


# Pivot table with multiple categories in a single column.

Market Category	Count	Popularity
Crossover	1068	1539.48
Crossover, Diesel	7	873.00
Crossover, Exotic, Luxury, High-Performance	1	238.00
Crossover,Exotic,Luxury,Performance	1	238.00
Crossover,Factory Tuner,Luxury,High-Performance	26	1823.46
Crossover,Factory Tuner,Luxury,Performance	5	2607.40
Crossover, Factory Tuner, Performance	4	210.00
Crossover,Flex Fuel	64	2073.75
Crossover,Flex Fuel,Luxury	10	1173.20
Crossover,Flex Fuel,Luxury,Performance	6	1624.00
Crossover,Flex Fuel,Performance	6	5657.00
Crossover,Hatchback	72	1675.69
Crossover,Hatchback,Factory Tuner,Performance	6	2009.00
Crossover,Hatchback,Luxury	7	204.00
Crossover,Hatchback,Performance	6	2009.00
Crossover,Hybrid	42	2563.38
Crossover,Luxury	406	889.21
Crossover,Luxury,Diesel	33	2195.85
Crossover,Luxury,High-Performance	9	1037.22
Crossover,Luxury,Hybrid	24	630.92
Crossover,Luxury,Performance	112	1349.09

Market Category	Count	Popularity
Crossover, Luxury, Performance, Hybrid	2	3916.00
Crossover,Performance	69	2585.96
Diesel	84	1730.90
Diesel,Luxury	47	2416.11
Exotic, Factory Tuner, High-Performance	21	1046.38
Exotic, Factory Tuner, Luxury, High-Performance	51	523.02
Exotic,Factory Tuner,Luxury,Performance	3	520.00
Exotic,Flex Fuel,Factory Tuner,Luxury,High-Performance	13	520.00
Exotic,Flex Fuel,Luxury,High-Performance	11	520.00
Exotic,High-Performance	245	1270.33
Exotic,Luxury	12	112.67
Exotic,Luxury,High-Performance	77	473.03
Exotic,Luxury,High-Performance,Hybrid	1	204.00
Exotic,Luxury,Performance	36	217.03
Factory Tuner,High-Performance	104	1966.44
Factory Tuner,Luxury	2	617.00
Factory Tuner,Luxury,High-Performance	215	2133.37
Factory Tuner,Luxury,Performance	31	1413.42
Factory Tuner,Performance	81	1818.05
Flex Fuel	855	2225.71
Flex Fuel,Diesel	16	5657.00
Flex Fuel,Factory Tuner,Luxury,High-Performance	1	258.00

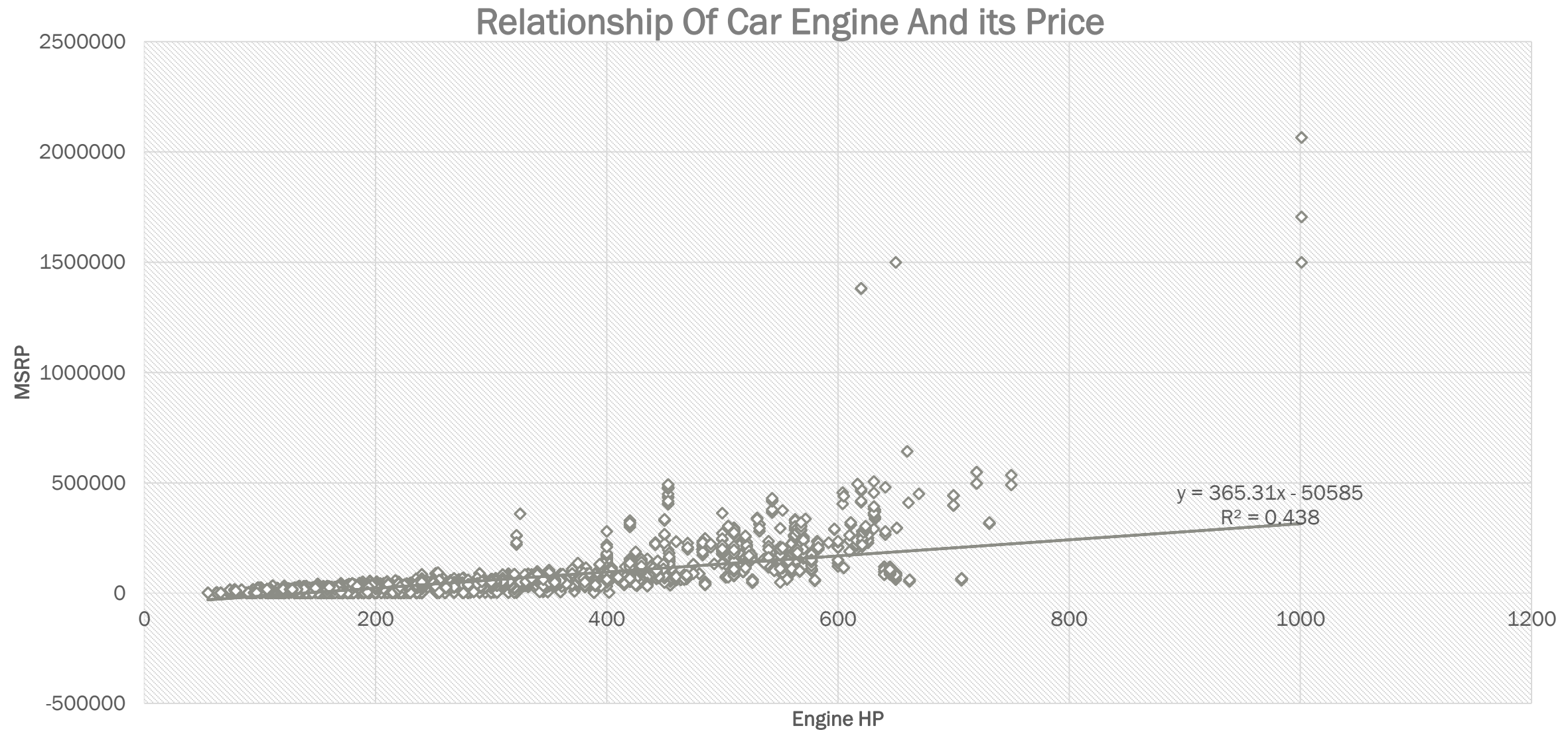
Market Category	Count	Popularity
Flex Fuel, Hybrid	2	155.00
Flex Fuel,Luxury	39	746.54
Flex Fuel,Luxury,High-Performance	32	898.31
Flex Fuel,Luxury,Performance	28	1380.07
Flex Fuel,Performance	81	1702.36
Flex Fuel,Performance,Hybrid	2	155.00
Hatchback	547	1279.11
Hatchback,Diesel	14	873.00
Hatchback,Factory Tuner,High-Performance	13	1205.15
Hatchback,Factory Tuner,Luxury,Performance	9	886.89
Hatchback, Factory Tuner, Performance	21	2173.71
Hatchback,Flex Fuel	7	5657.00
Hatchback,Hybrid	64	2111.16
Hatchback,Luxury	45	1323.13
Hatchback,Luxury,Hybrid	3	454.00
Hatchback,Luxury,Performance	36	1632.25
Hatchback,Performance	198	1073.66
High-Performance	198	1823.38
Hybrid	121	2116.59
Luxury	815	1084.21
Luxury,High-Performance	334	1668.02
Luxury,High-Performance,Hybrid	12	568.83
Luxury,Hybrid	48	724.69
Luxury,Performance	659	1293.06
Luxury,Performance,Hybrid	11	2333.18
Performance	503	1443.23
Performance, Hybrid	1	155.00



A red toy car is positioned on top of a stack of gold coins. The car is a small-scale model, possibly a sports car, with a sleek design. The coins are stacked in two columns, one on the left and one on the right, with the car resting on the left column. The background is a soft, out-of-focus gradient of light green and yellow. The text is overlaid on the image, centered horizontally and slightly below the middle vertically.

**WHAT IS A RELATIONSHIP BETWEEN  
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.**

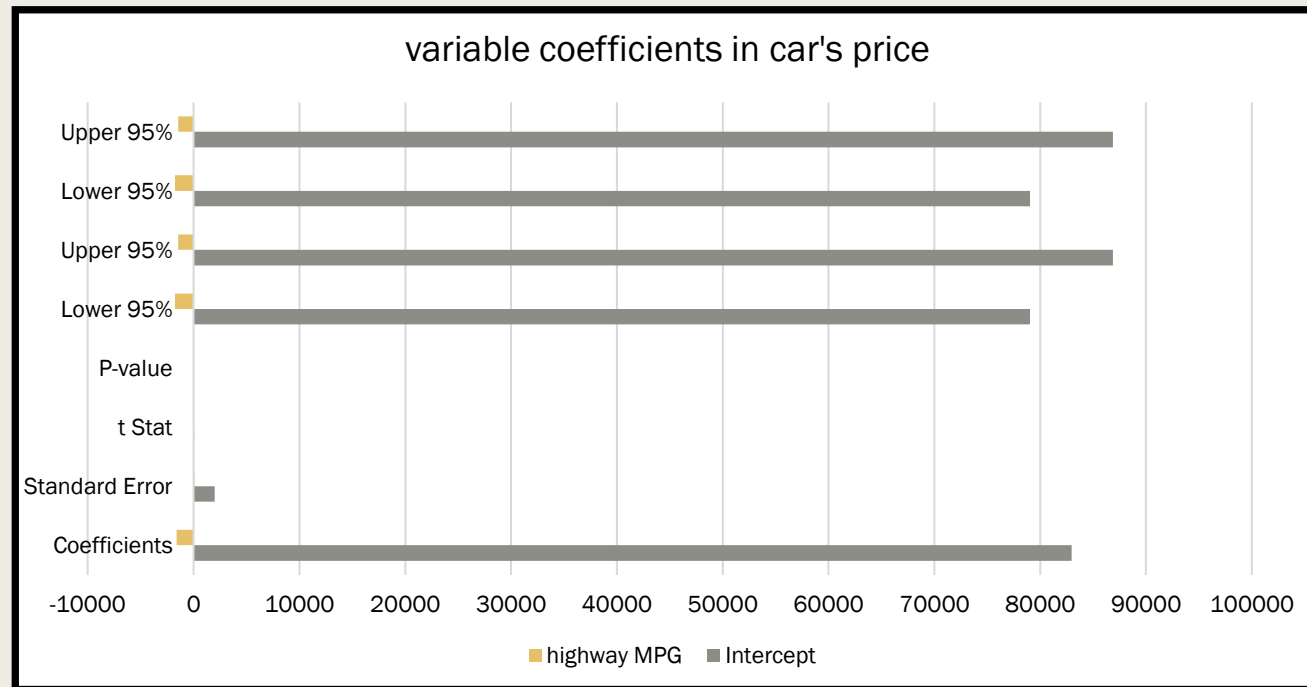
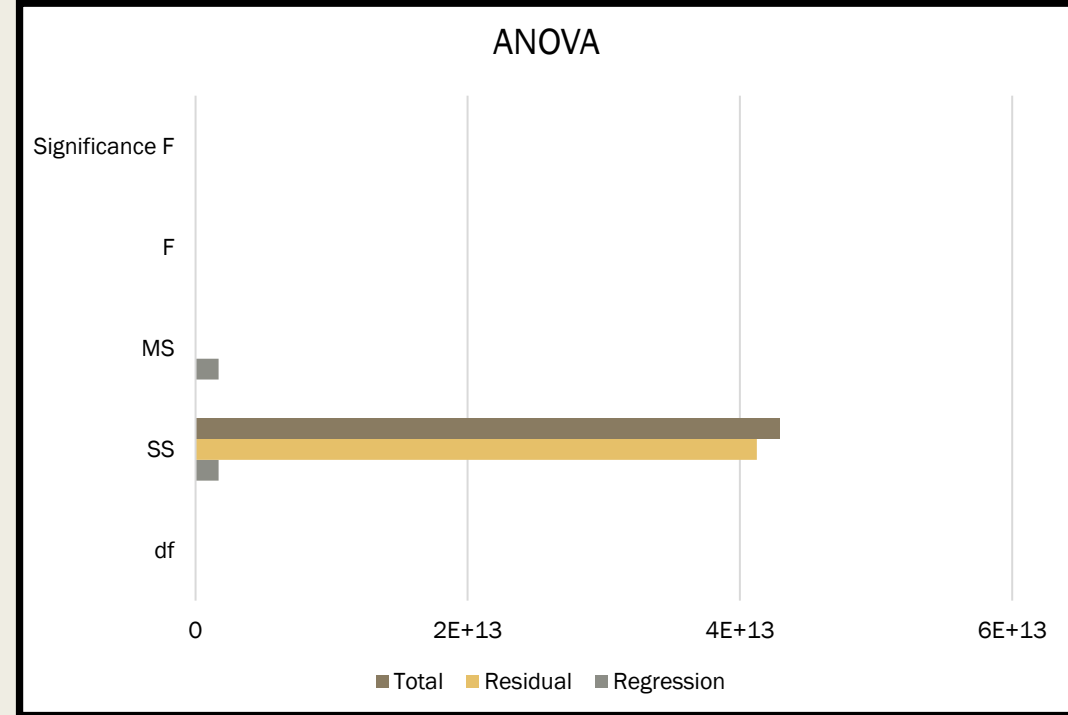
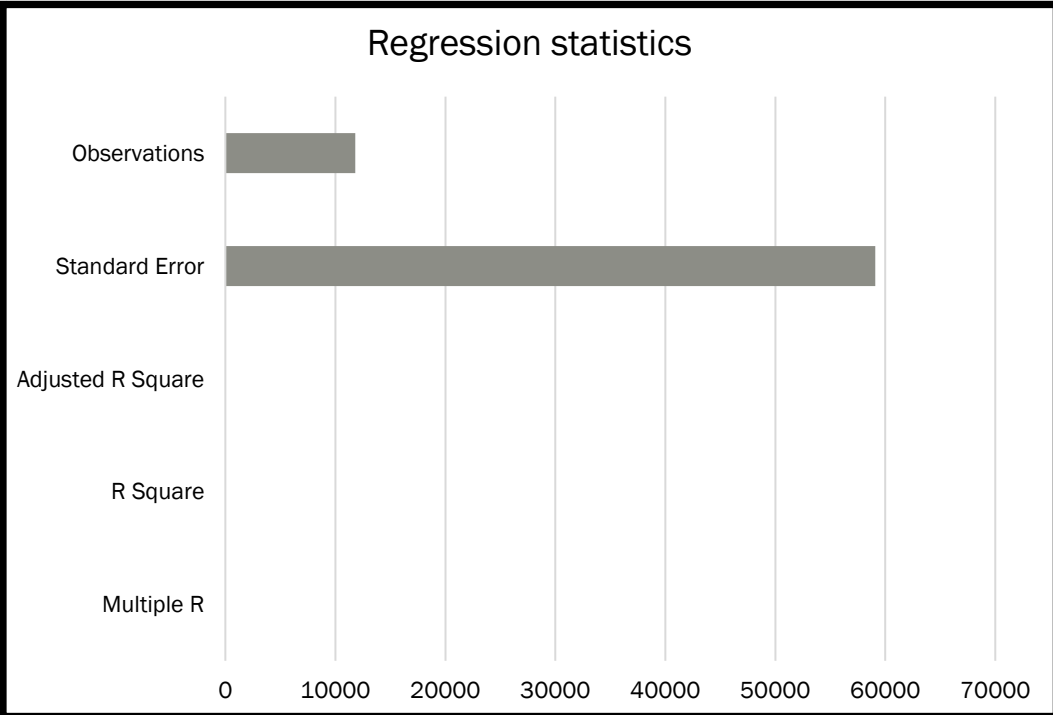


A red toy car is positioned on top of two stacks of gold coins. The car is a small-scale model, possibly a sports car, with a sleek design. The coins are stacked in two separate piles, one on the left and one on the right, with the car's front wheel resting on the left stack. The background is a soft, out-of-focus gradient of light colors, suggesting an indoor setting with natural light. The overall composition is centered, with the car and coins being the primary focus.

**WHICH CAR FEATURES ARE MOST  
IMPORTANT IN DETEREMING THE  
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.**

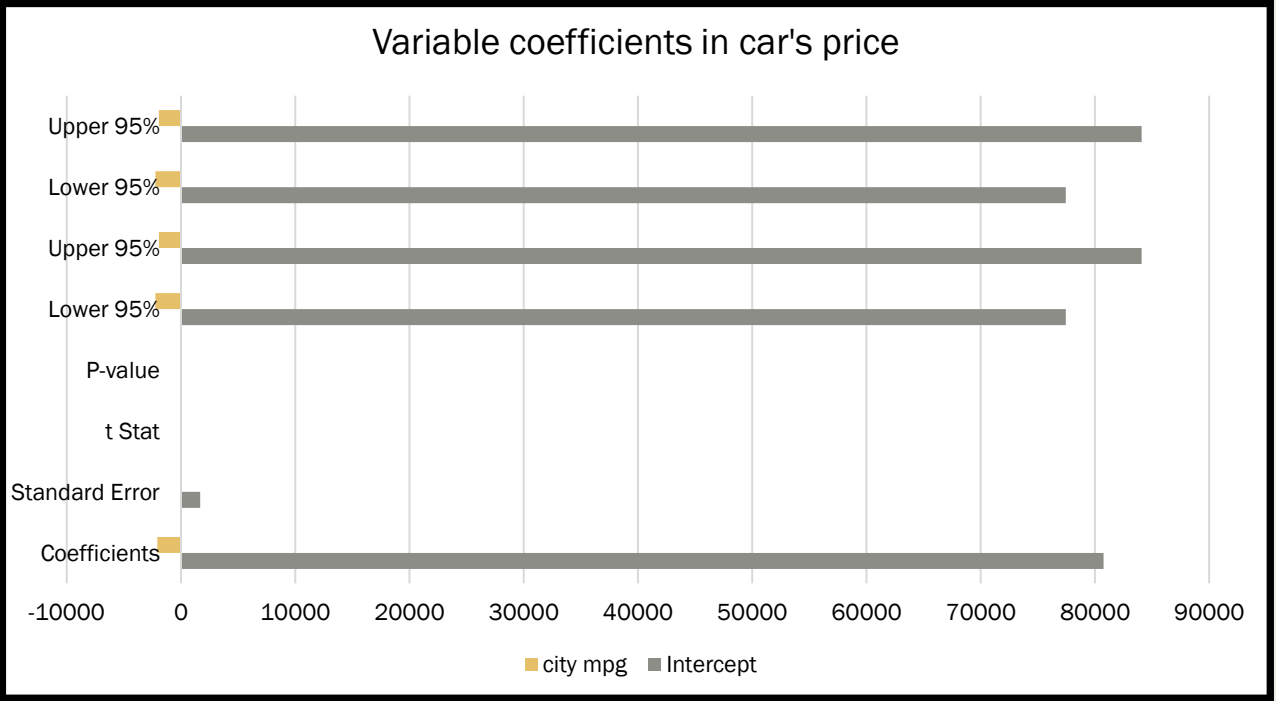
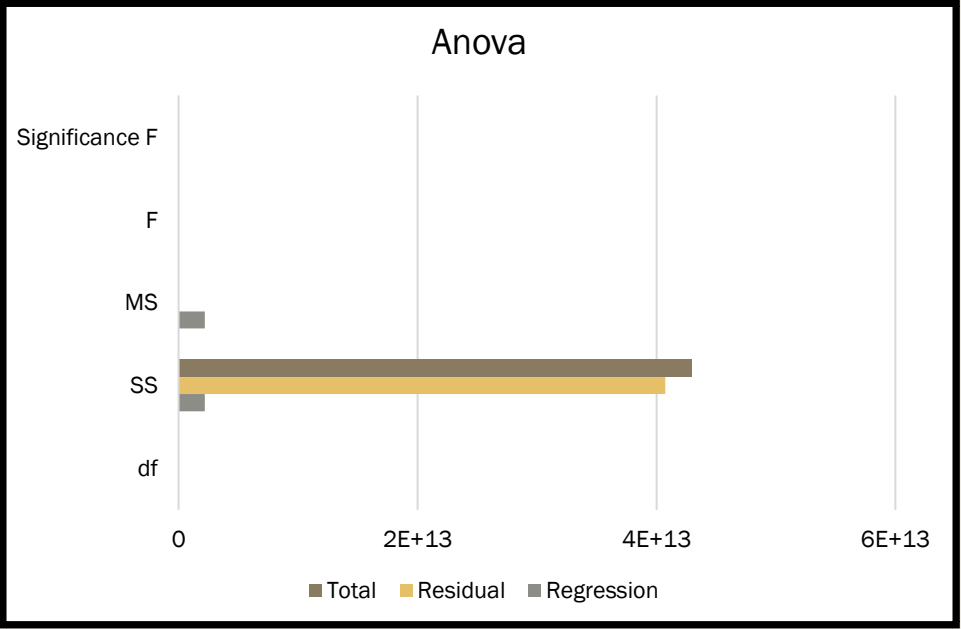
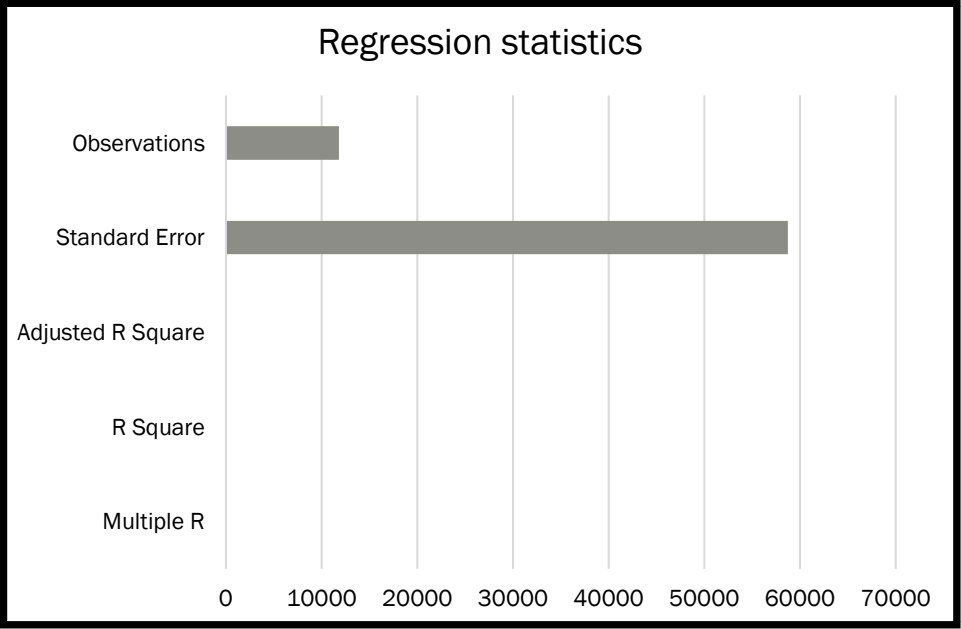
HIGHWAY MPG								
Regression Statistics								
Multiple R	0.198941632							
R Square	0.039577773							
Adjusted R Square	0.039496471							
Standard Error	59075.12953							
Observations	11815							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	1.69886E+12	1.69886E+12	486.7986353	0			
Residual	11813	4.12258E+13	3489870929					
Total	11814	4.29247E+13						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95%	Upper 95%
Intercept	82962.37196	1997.447559	41.53419278	0	79047.04563	86877.69829	79047.04563	86877.69829
highway MPG	-1611.209062	73.02595074	-22.06351367	9.531E-106	-1754.351958	-1468.066166	-1754.351958	-1468.066166



**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.**

CITY MPG								
Regression Statistics								
Multiple R	0.225276521							
R Square	0.050749511							
Adjusted R Square	0.050669154							
Standard Error	58730.54029							
Observations	11815							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	2.17841E+12	2.17841E+12	631.555083	0			
Residual	11813	4.07463E+13	3449276363					
Total	11814	4.29247E+13						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95%	Upper 95%
Intercept	80764.32799	1688.796997	47.82358575	0	77454.00762	84074.64837	77454.00762	84074.64837
city mpg	-2080.447909	82.78491912	-25.1307597	8.17E-136	-2242.719991	-1918.175827	-2242.719991	-1918.175827





A conceptual image showing a small red car model placed on a stack of gold coins, symbolizing the cost of a vehicle. The background is a soft-focus gradient of green and yellow.

**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.

Manufacturer	Price
Acura	34887.59
Alfa Romeo	61600.00
Aston Martin	197910.38
Audi	53452.11
Bentley	247169.32
BMW	61546.76
Bugatti	1757223.67
Buick	28206.61
Cadillac	56231.32
Chevrolet	28273.36
Chrysler	26722.96
Dodge	22390.06
Ferrari	237383.82
FIAT	22206.02
Ford	27393.42
Genesis	46616.67
GMC	30493.30
Honda	26629.82
HUMMER	36464.41
Hyundai	24597.04
Infiniti	42394.21
Kia	25112.39
Lamborghini	331567.31

Manufacturer	Price
Land Rover	67823.22
Lexus	47549.07
Lincoln	42494.37
Lotus	69188.28
Maserati	114207.71
Maybach	546221.88
Mazda	19719.06
McLaren	239805.00
Mercedes-Benz	71537.81
Mitsubishi	21215.47
Nissan	28513.37
Oldsmobile	11542.54
Plymouth	3122.90
Pontiac	19321.55
Porsche	101622.40
Rolls-Royce	351130.65
Saab	27413.50
Scion	19932.50
Spyker	213323.33
Subaru	24827.50
Suzuki	17907.21
Toyota	28946.15
Volkswagen	28076.20

**Task 4.B:** Create a bar chart or a horizontal stacked bar chart that visualizes the relationship between manufacturer and average price.



A red toy car is positioned on top of a stack of gold coins. The car is a small, detailed model with black wheels and a white interior. The coins are stacked in two columns, with the car resting on the left column. The background is a soft, out-of-focus green and yellow gradient. The text is overlaid on the right side of the image, partially obscuring the car and the coins.

**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.

correlation coefficient between the number of cylinders and highway MPG
-0.620315935

By using the CORREL function correlation coefficient is calculated.

This coefficient quantifies with the strength and direction of the relationship between these two variables. As the number of cylinders increases, the highway MPG tends to decrease. This findings suggest that the car with fewer cylinders tend to be more fuel efficient on the highway, while those with the more cylinders typically consume more fuel.

# BUILDING THE DASHBOARD



(P)

P



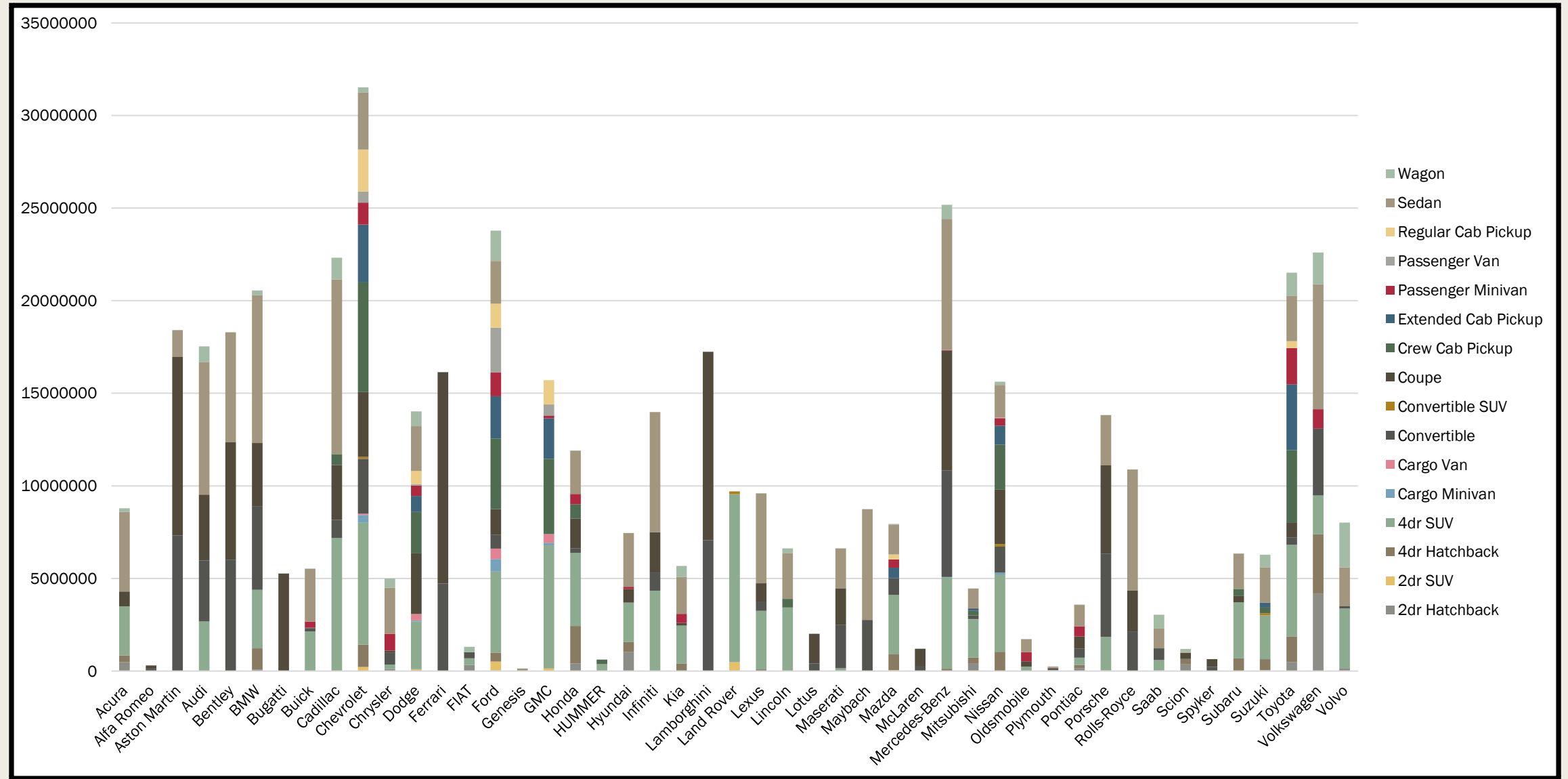
148 km  
95.4 km



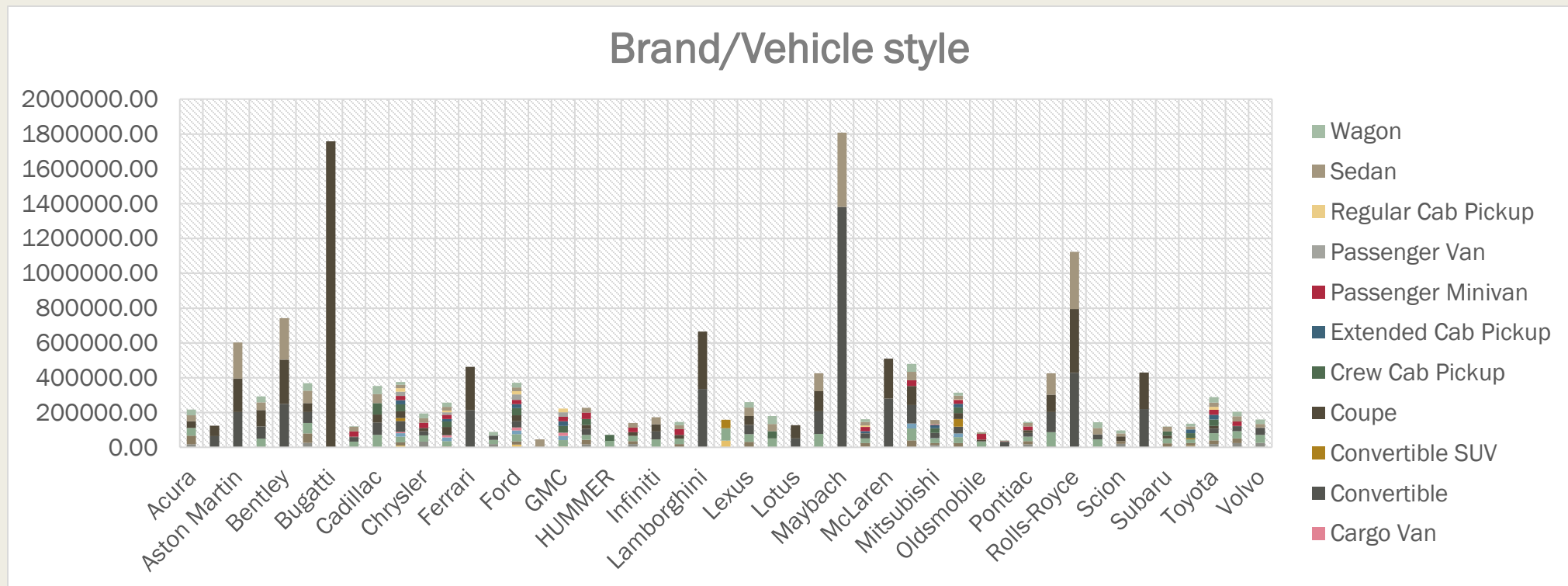
→ 99  
km



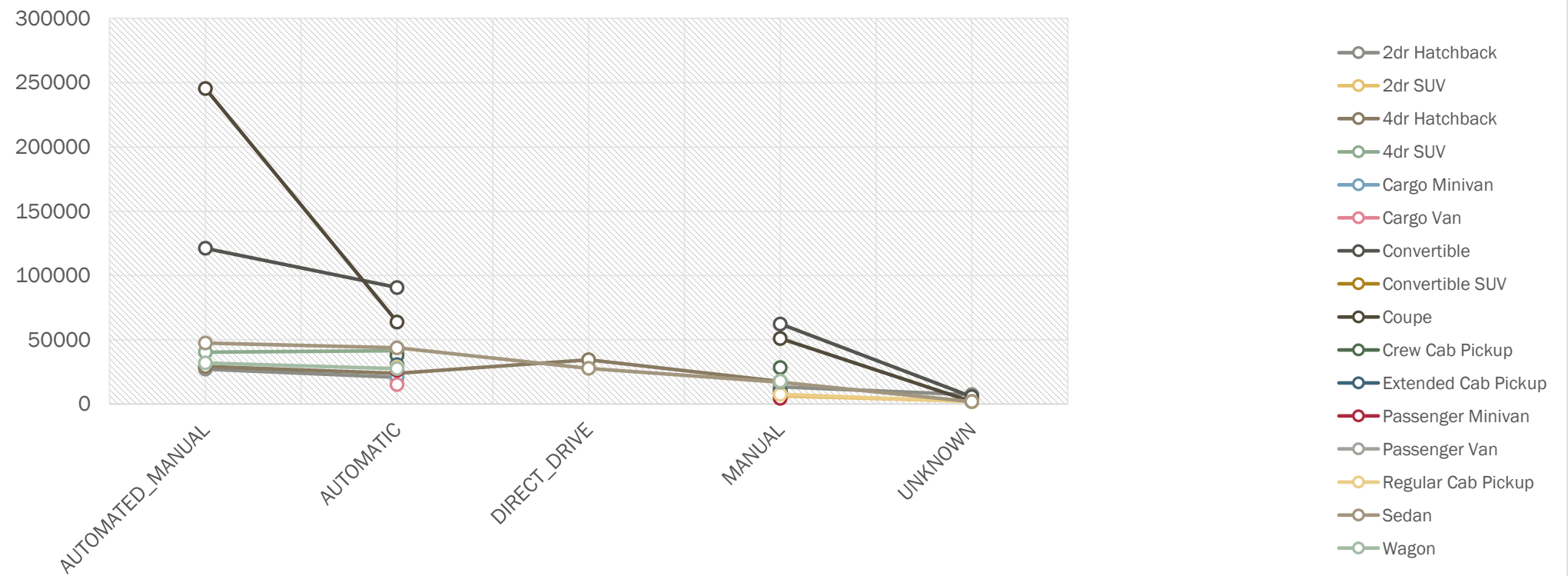
## Task 1: How does the distribution of car prices vary by brand and body style?



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

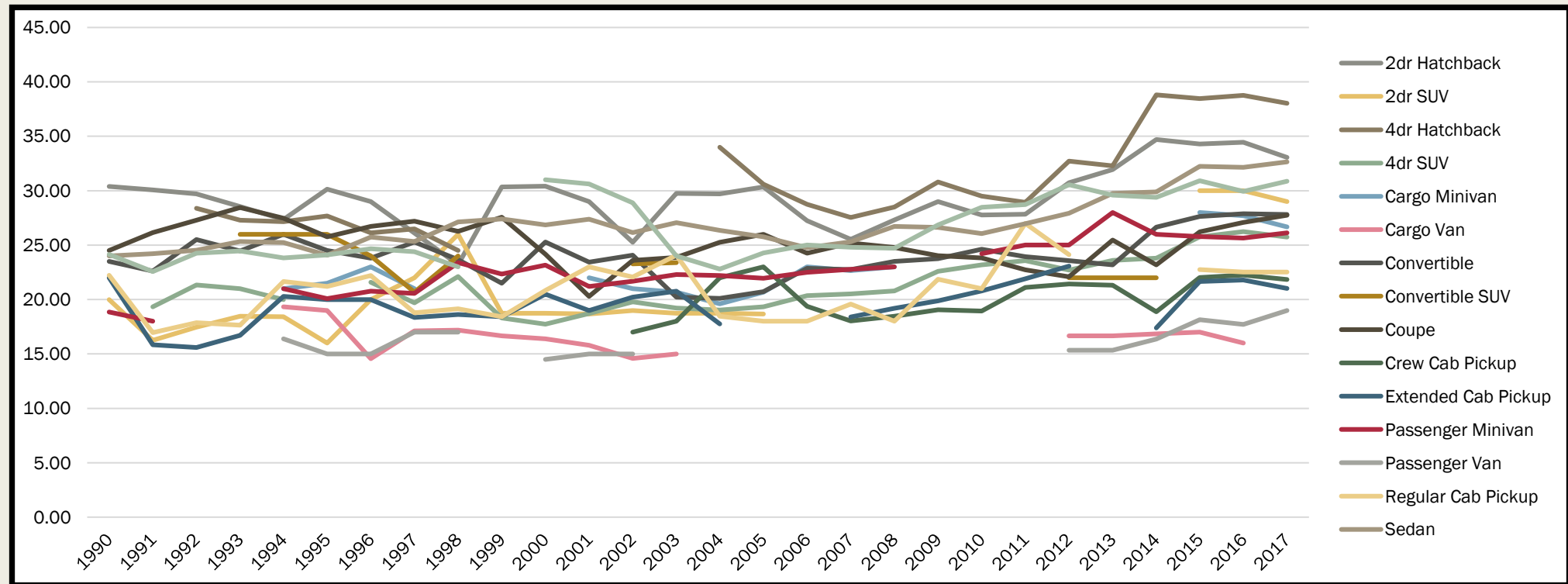


### Task 3: How do the different features such as transmission type affect the MSRP, and how does this vary by body style?



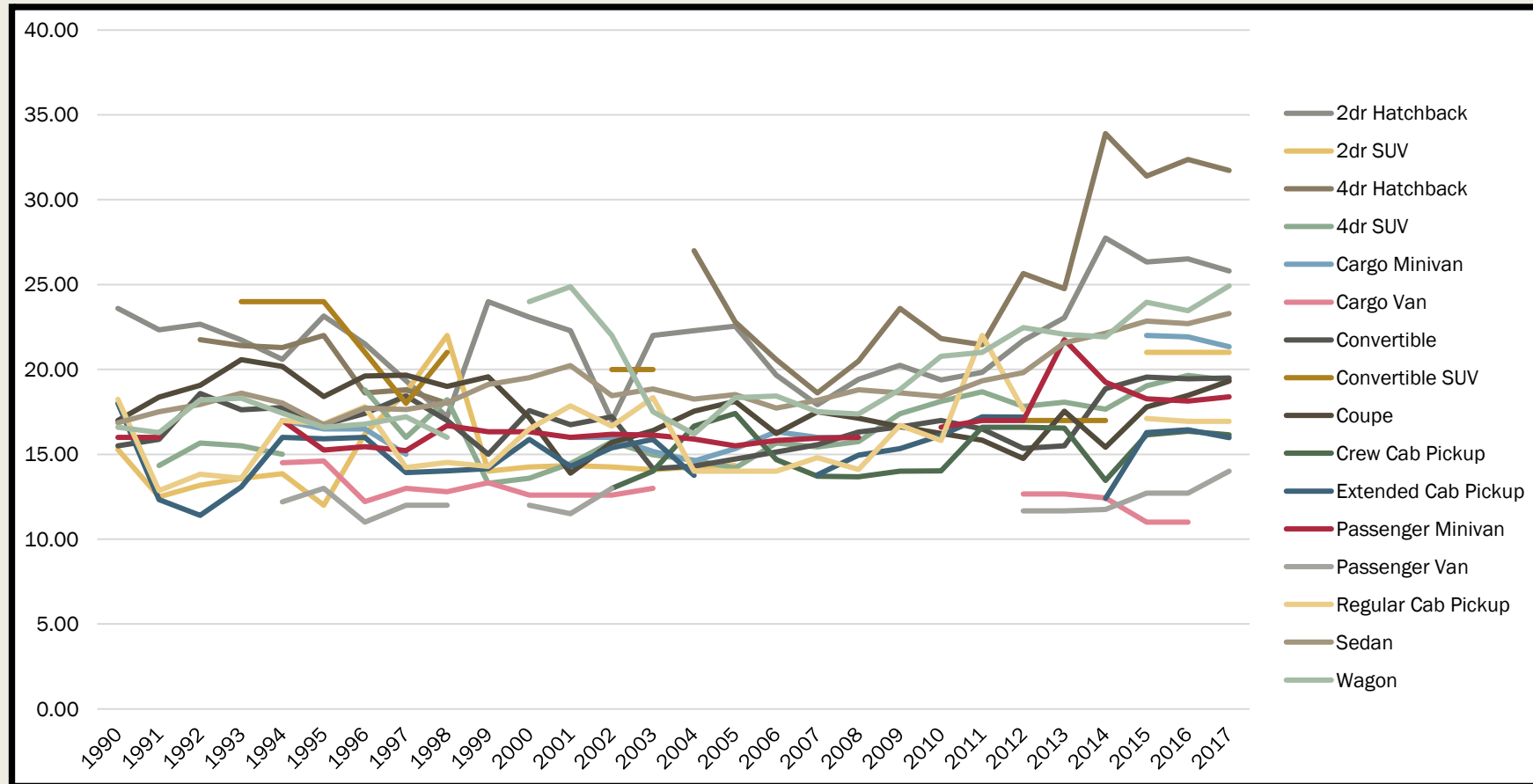
## Task 4: How does the fuel efficiency of cars vary across different body styles and model years?

### HIGHWAY MPG

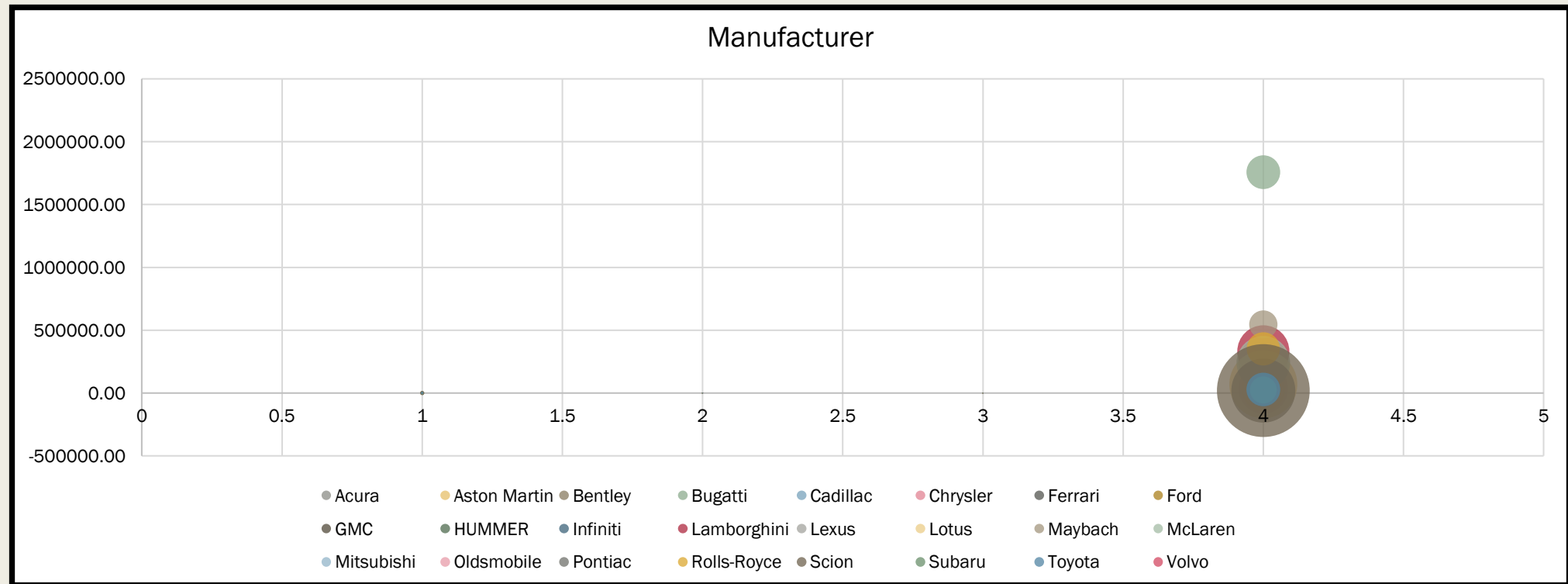


## Task 4: How does the fuel efficiency of cars vary across different body styles and model years?

### CITY MPG



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





# Excel Sheet Drive Link

<https://docs.google.com/spreadsheets/d/1vckHBNXj71w2aFlpbVE3xWcE99hhIr3e/edit?usp=sharing&oid=102610263604488941402&rtpof=true&sd=true>

# RESULT

- This project has greatly benefited me by providing practical hands on experience in data analysis. It equipped me with essential skills in data processing, regression analysis and data visualisation using tools like Excel. Moreover, the project enhanced problem-solving and communication skills as to present actionable insight. These acquired skills and the successful completion of a real world analysis project have not only bolstered my confidence but also improved the employability and prepared me for the future data-related roles.
- For the business this project has been instrumental in informed decision making it allowed the business to optimise its pricing and product development strategies based on data driven insights by understanding the factors influencing car prices the business can make adjustment that enhance the competitiveness reduce cost and increase the profitability. The project demonstrates the value of data analysis in guiding market efforts, risk medication, and improving customer satisfaction, reinforcing the importance of data-driven approaches in the business's future endeavors.



Thank You

Ritika Chaudhary  
ameliarc2105@gmail.com

