

Analyzing the Impact of Car Features on Price and Profitability

[Link to Project Excel File](#)

Ctrl+Click the above links

Project Overview: Analyzing the Impact of Car Features on Price and Profitability

This project explores how various car features influence **pricing and profitability** in the automotive industry. With evolving consumer preferences and increasing competition, car manufacturers need data-driven strategies to **optimize pricing and product development**.

Using data analysis techniques such as **regression analysis and market segmentation**, the project examines the relationship between car features (like horsepower, fuel type, body style) and their market value. The goal is to identify the most **popular and profitable** car attributes, enabling manufacturers to **balance consumer demand with profitability**.

By leveraging insights from the analysis, manufacturers can make **informed pricing decisions, enhance product development**, and maintain a competitive edge in the market.

Task 1

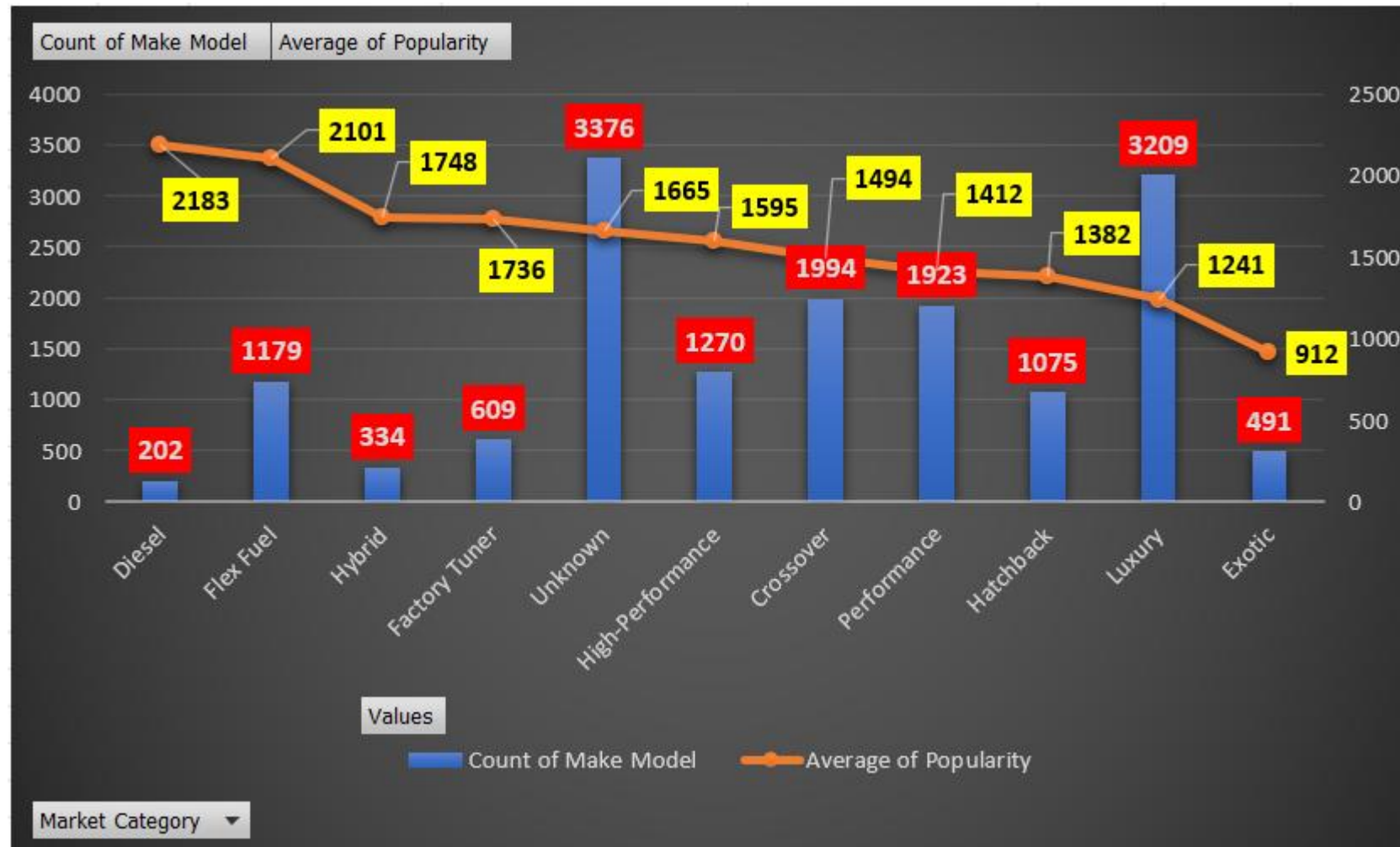
Insight Required: How does the popularity of a car model vary across different market categories?

- **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.

Task 1A - Here are the number of car models in each market category and their corresponding popularity scores.

Market Category ▼	Count of Make Model	Average of Popularity
Diesel	202	2183
Flex Fuel	1179	2101
Hybrid	334	1748
Factory Tuner	609	1736
Unknown	3376	1665
High-Performance	1270	1595
Crossover	1994	1494
Performance	1923	1412
Hatchback	1075	1382
Luxury	3209	1241
Exotic	491	912
Grand Total	15662	1521

Task 1B – Combo chart – Market Category, Number of Car models and their respective popularity



Insights

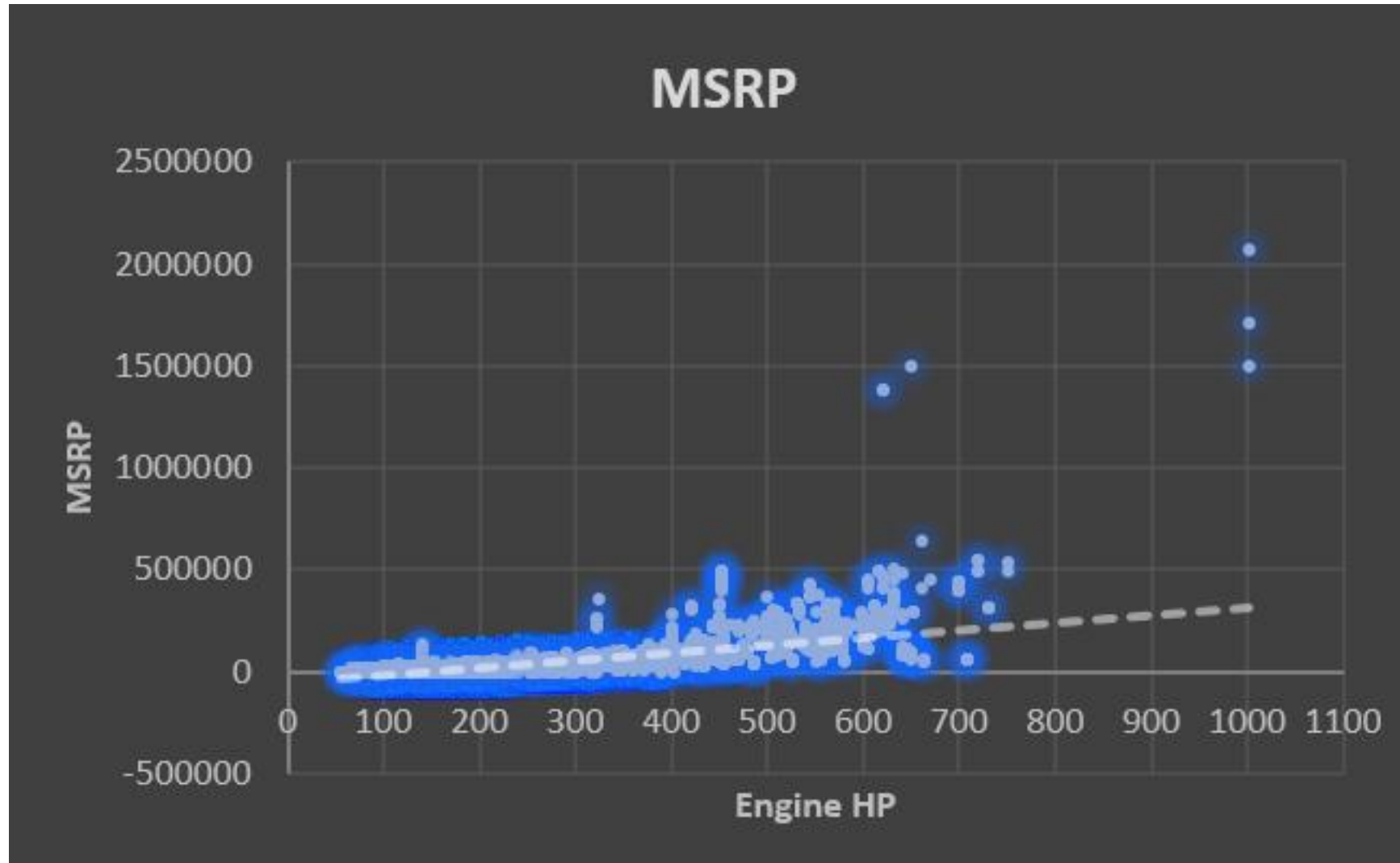
- 1."Unknown" category has the highest count (3376) but only moderate popularity (1665).
- 2.Luxury cars are widely available (3209) but have a declining popularity score (1382).
- 3.Exotic cars are rare (491) and have the lowest popularity (912).
- 4.High-Performance cars (1270) are more popular (1494) than Performance cars (1923, 1412).
- 5.Hybrid cars (334) are fewer but more popular (2101) than Flex Fuel cars (1179, 1748).

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.

Scatter chart

- Engine power on the x-axis
- MSRP on the y-axis



Insights

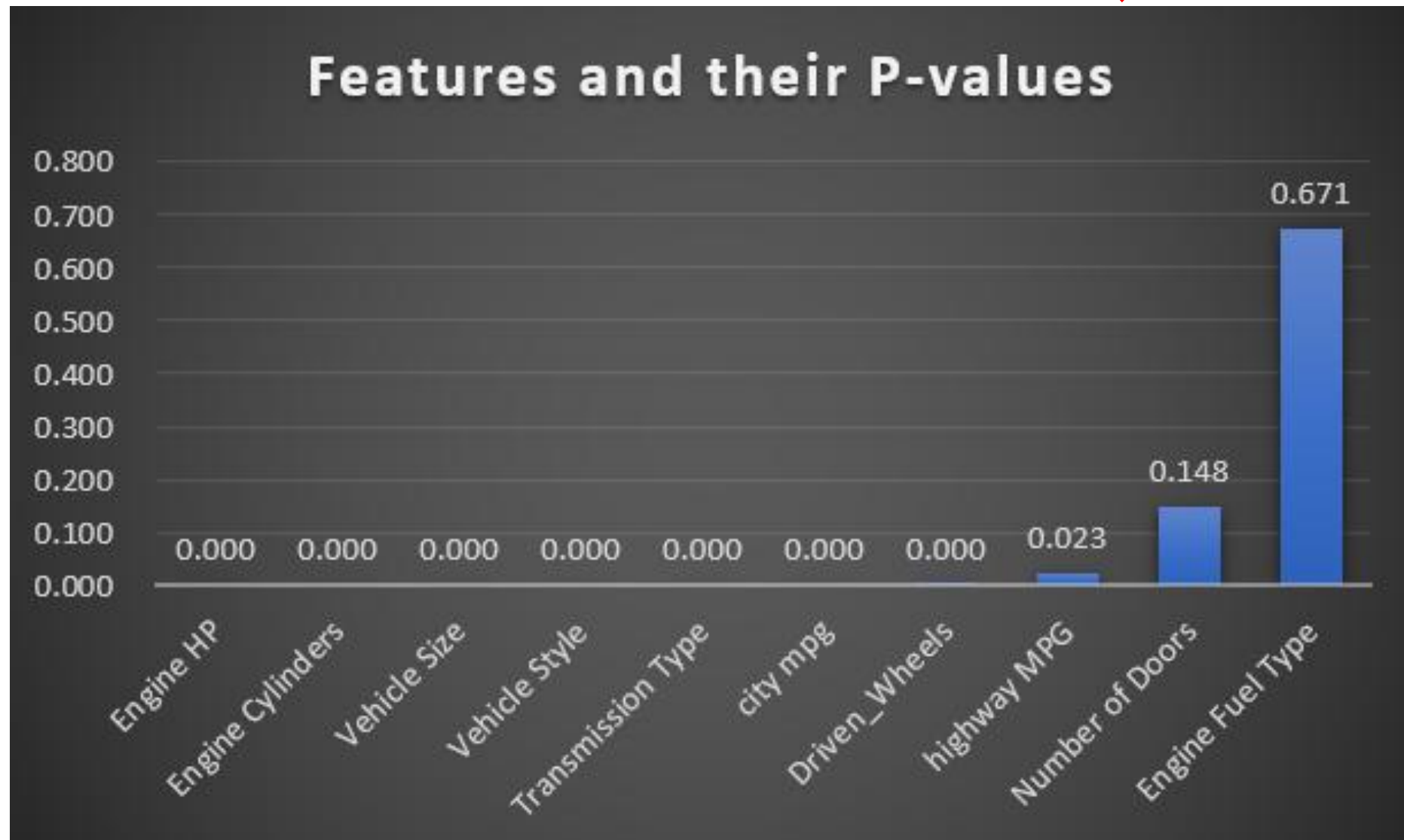
- **Positive Correlation** – Higher engine horsepower (HP) generally leads to a higher MSRP, as indicated by the upward trend in the dotted trendline.
- **Luxury & Supercars Outliers** – A few vehicles with very high HP (>800) have significantly higher MSRPs, likely representing exotic or luxury cars.
- **Majority Cluster in Lower Range** – Most cars are concentrated below 500 HP and under \$100,000, suggesting that mainstream vehicles dominate the dataset.
- **Non-Linear Pricing** – While HP increases gradually, MSRP shows sharp jumps, indicating that luxury branding, exclusivity, and features also impact pricing.
- **Some Affordable High-HP Cars** – There are instances where cars with relatively high HP (~500-600) still have moderate MSRPs, possibly due to performance-focused mainstream models.

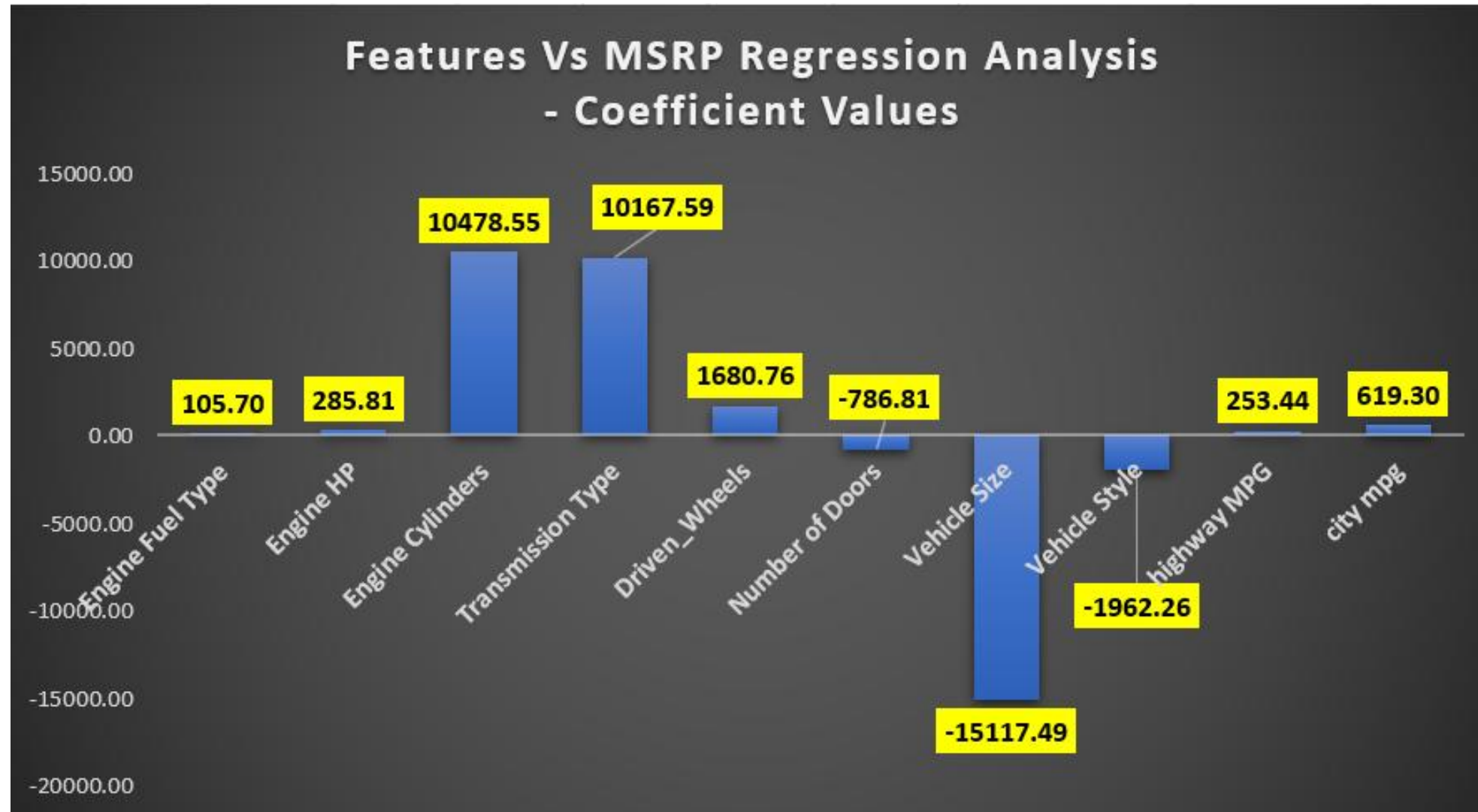
Insight Required: Which car features are most important in determining a car's price?

- **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.

Project 7 – Impact of Car Features

ni								
Regression Statistics								
Multiple R	0.704965493							
R Square	0.496976346							
Adjusted R Square	0.496526736							
Standard Error	43662.71619							
Observations	11199							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	10	2.10728E+13	2.10728E+12	1105.34988	0			
Residual	11188	2.13292E+13	1906432785					
Total	11198	4.24019E+13						
Factors	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-90283.6972	4881.954398	-18.4933512	3.0722E-75	-99853.1873	-80714.2071	-99853.1873	-80714.2071
Engine Fuel Type	105.7029477	249.2023898	0.424165064	0.67145359	-382.777607	594.1835023	-382.777607	594.1835023
Engine HP	285.80656	6.664692006	42.88368612	0	272.7425904	298.8705297	272.7425904	298.8705297
Engine Cylinders	10478.55251	457.5844774	22.89971147	1.854E-113	9581.606375	11375.49864	9581.606375	11375.49864
Transmission Type	10167.58652	878.4914764	11.57391596	8.3664E-31	8445.588572	11889.58447	8445.588572	11889.58447
Driven_Wheels	1680.76274	449.317091	3.740705113	0.00018442	800.0221417	2561.503338	800.0221417	2561.503338
Number of Doors	-786.8117167	543.2036171	-1.44846553	0.14751492	-1851.58643	277.9630008	-1851.58643	277.9630008
Vehicle Size	-15117.49134	663.5696129	-22.782073	2.427E-112	-16418.2046	-13816.7781	-16418.2046	-13816.7781
Vehicle Style	-1962.256302	142.03293	-13.8155025	4.6353E-43	-2240.66585	-1683.84675	-2240.66585	-1683.84675
highway MPG	253.4435608	111.1624762	2.279938065	0.0226301	35.54553796	471.3415836	35.54553796	471.3415836
city mpg	619.2986168	101.7405952	6.087035522	1.1879E-09	419.8691394	818.7280942	419.8691394	818.7280942





Insights

A high coefficient does not always mean a feature is important unless it has a low p-value.

Look for both high coefficients and low p-values when identifying the most important features.

Most impactful and statistically significant features:

Engine HP (Coefficient: 285.81, P-value: 0.000) → **Strong, significant impact.**

Engine Cylinders (Coefficient: 10,478.55, P-value: 1.85E-113) → **Very significant.**

Transmission Type (Coefficient: 10,167.59, P-value: 8.36E-31) → **Strong impact and highly significant.**

Vehicle Size (Coefficient: -15,117.49, P-value: 2.43E-112) → **Strong negative impact and highly significant.**

Top 4 most important features:

Vehicle Size, Engine Cylinders, Transmission Type, and Engine HP.

We need to avoid using Number of Doors & Engine Fuel Type as key predictors since their impact is not statistically significant.

City & Highway MPG affect MSRP but not as much as engine and transmission-related features.

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.

Project 7 – Impact of Car Features

Project by – Alok Joshi

Manufacturer	Avg MSRP
Bugatti	1757224
Maybach	546222
Rolls-Royce	351131
Lamborghini	331567
Bentley	247169
McLaren	239805
Ferrari	238219
Spyker	214990
Aston Martin	198123
Maserati	113684
Porsche	101622
Tesla	85256
Mercedes-Benz	72070
Lotus	68377
Land Rover	68067
BMW	62163
Alfa Romeo	61600
Cadillac	56368
Audi	54574
Lexus	47549

Manufacturer	Avg MSRP
Genesis	46617
Lincoln	43861
Infiniti	42640
HUMMER	36464
Acura	35087
GMC	32444
Volvo	29725
Chevrolet	29075
Buick	29034
Volkswagen	28979
Nissan	28921
Toyota	28847
Ford	28511
Saab	27880
Chrysler	26723
Honda	26655
Kia	25514
Hyundai	24926
Dodge	24857
Subaru	24241

Manufacturer	Avg MSRP
FIAT	22670
Mitsubishi	21341
Mazda	20417
Scion	19933
Pontiac	19800
Suzuki	18026
Oldsmobile	12844
Plymouth	3297

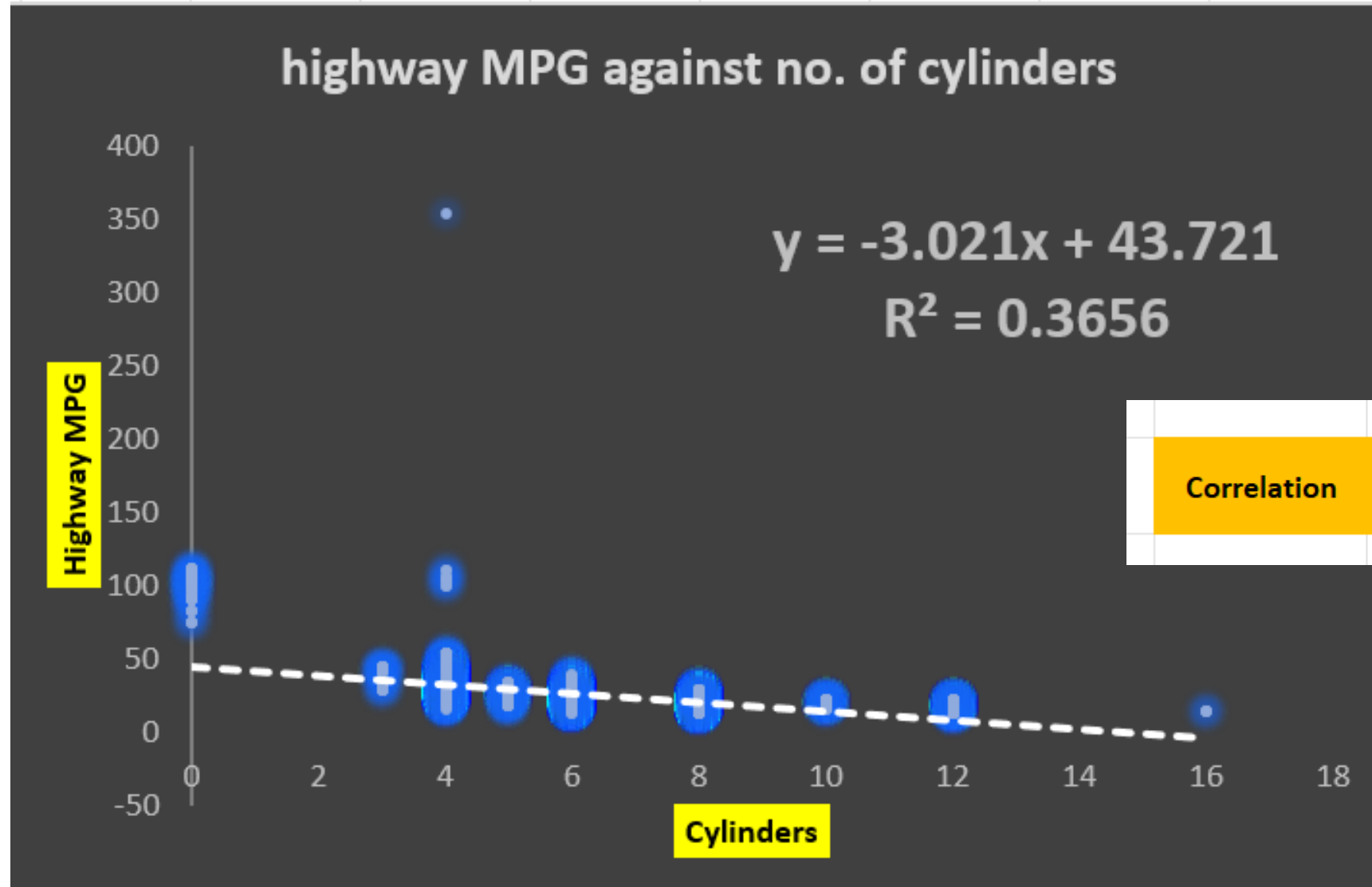
MRSP wise Top 10 Manufacturers

Manufacturer	Avg MSRP	Rank
Bugatti	1757224	1
Maybach	546222	2
Rolls-Royce	351131	3
Lamborghini	331567	4
Bentley	247169	5
McLaren	239805	6
Ferrari	238219	7
Spyker	214990	8
Aston Martin	198123	9
Maserati	113684	10
Porsche	101622	11
Tesla	85256	12
Mercedes-Benz	72070	13
Lotus	68377	14
Land Rover	68067	15



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.



Correlation	-0.60466812
-------------	-------------

Insights from the Scatter Plot

Negative Slope (-3.021) - The equation $y = -3.021x + 43.721$ suggests that for each additional cylinder, the highway MPG decreases by approximately **3.021 units**, reinforcing that higher-cylinder engines are generally less fuel-efficient.

Intercept (43.721) - When the number of cylinders is zero (hypothetically), the predicted highway MPG would be **43.721**, though this is not realistic for real-world vehicles.
(Although 0-cylinder vehicles are Electric Vehicles)

R² Value (0.3656) - Weak Correlation - The $R^2 = 0.3656$ indicates that only **36.56% of the variation in highway MPG** is explained by the number of cylinders. This means other factors (like vehicle weight, aerodynamics, fuel type) may also significantly influence MPG, beyond just the number of cylinders.

Correlation value of -0.604 indicates that as the cylinders increase in number the milage takes a backseat.

Building the Dashboard

Clipboard

Get data

Excel workbook catalog

OneLake Server

Enter data

Dataaverse data

Recent sources

Transform data

Refresh data

New visual

Text box

More visuals

Calculations

New measure

Quick measure

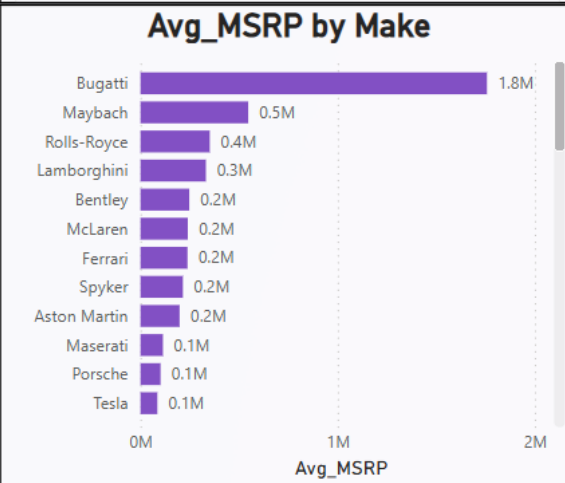
Sensitivity

Share

Copilot

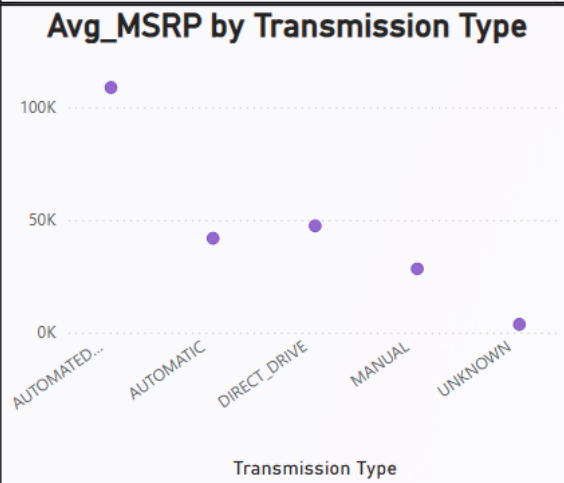
Brand

All



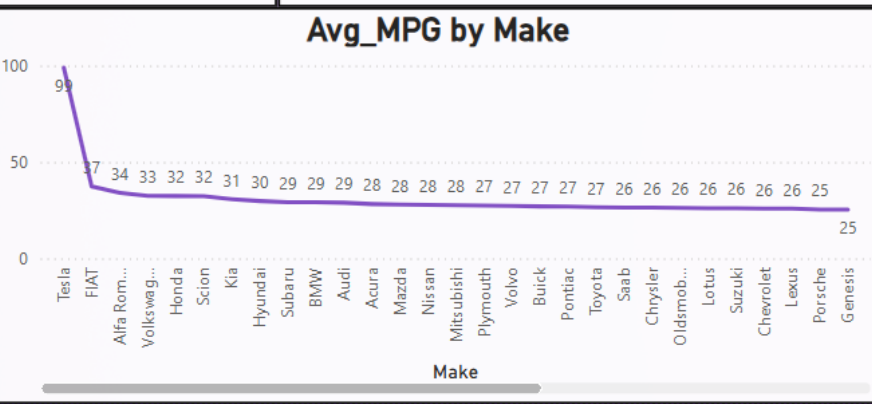
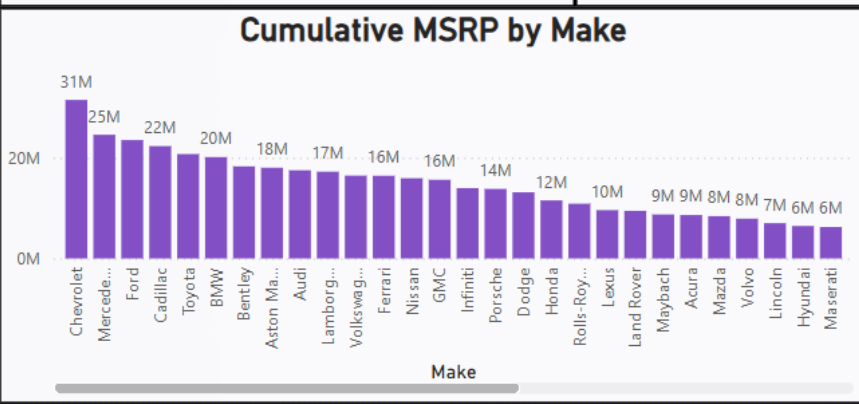
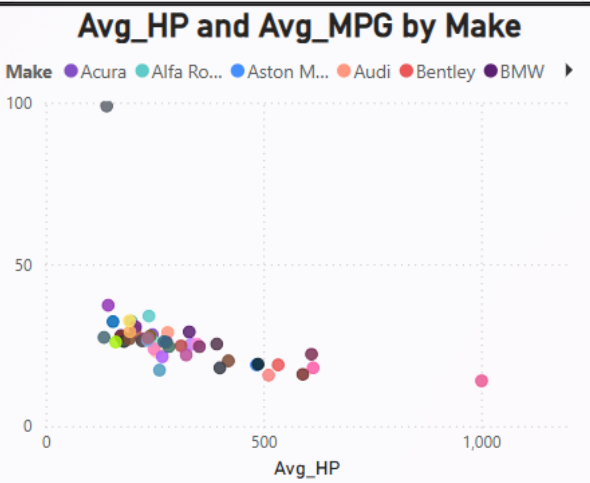
Vehicle Style

All



Year

All



Visualizations

Build visual

Filters

Data

Search

Project Data

Values

Add data fields here

Drill through

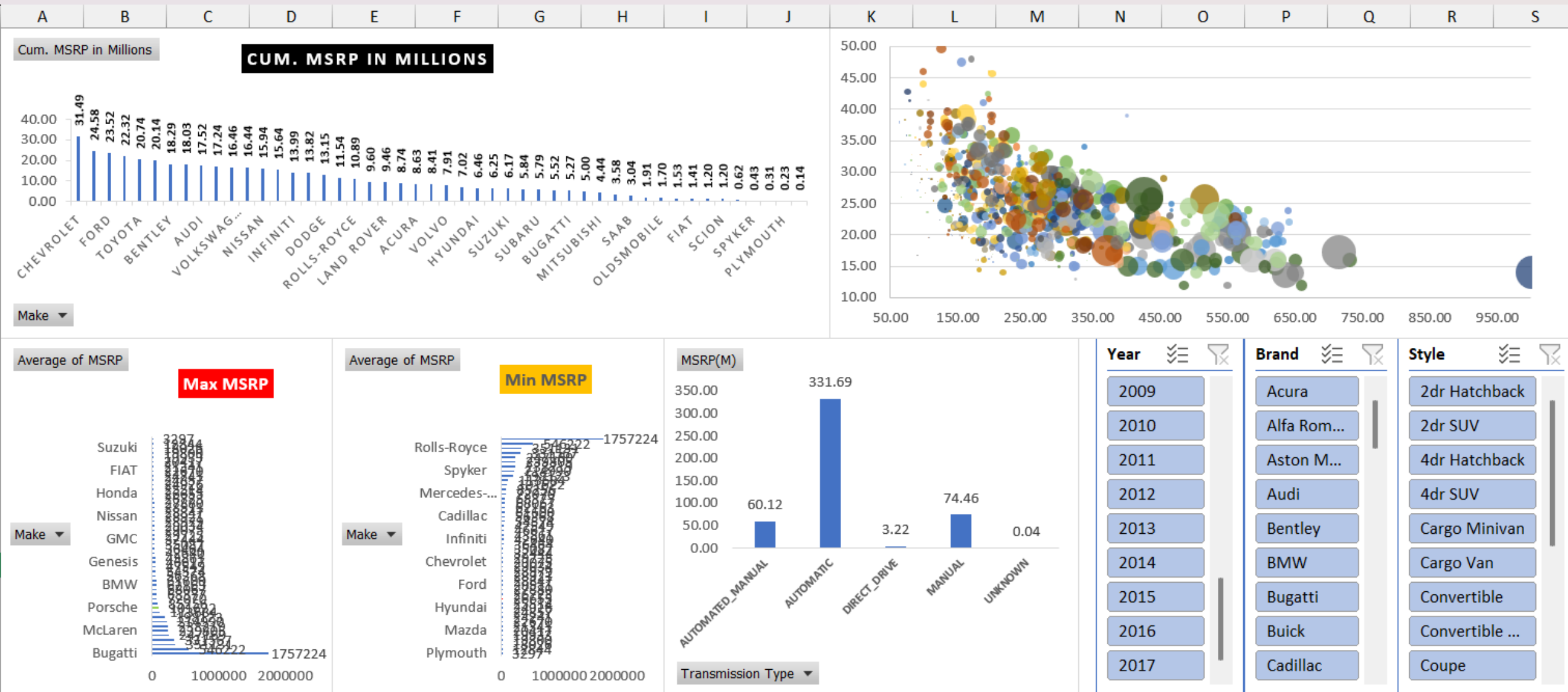
Cross-report

Keep all filters

Add drill-through fields here

Project 7 – Impact of Car Features

Project by – Alokk Joshi



5 Major Takeaways

1. Higher engine HP strongly correlates with higher MSRP, but luxury branding and exclusivity also impact pricing.
2. Vehicle Size, Engine Cylinders, Transmission Type, and Engine HP are the top four factors influencing MSRP.
3. Most cars are below 500 HP and \$100,000, while luxury and exotic cars create sharp price jumps.
4. Each additional engine cylinder decreases highway MPG by ~3.021 units, confirming lower fuel efficiency.
5. Cylinder count weakly explains MPG variation ($R^2 = 0.3656$), suggesting other factors like weight and aerodynamics matter.

5 Major Learnings

Data Cleaning is Critical – Handling missing values, renaming columns, and standardizing data are essential for accurate analysis.

Exploratory Data Analysis (EDA) is Insightful – Understanding distributions, correlations, and trends helps uncover key patterns before diving into modeling.

Statistical Thinking is Essential – Learning to interpret coefficients, p-values, and R^2 values improved your ability to derive meaningful insights.

Visualization Strengthens Storytelling – Graphs and trendlines make complex data more understandable and actionable for decision-making.

Tool Mastery Matters – Applying Excel, SQL, and Python (Pandas, Matplotlib, etc.) enhanced your ability to manipulate and analyze large datasets efficiently.