Trainity project 7

Analyzing the Impact of Car Features on Price and Profitability

Problem Statement:

The automotive industry has been rapidly evolving over the past few decades, with a growing focus on fuel efficiency, environmental sustainability, and technological innovation. With increasing competition among manufacturers and a changing consumer landscape, it has become more important than ever to understand the factors that drive consumer demand for cars.

In recent years, there has been a growing trend towards electric and hybrid vehicles and increased interest in alternative fuel sources such as hydrogen and natural gas. At the same time, traditional gasoline-powered cars remain dominant in the market, with varying fuel types and grades available to consumers.

For the given dataset, as a Data Analyst, the client has asked How can a car manufacturer optimize pricing and product development decisions to maximize profitability while meeting consumer demand?

This problem could be approached by analyzing the relationship between a car's features, market category, and 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.

Dataset Description:

The dataset contains information on various car models and their specifications, and is titled "Car Features and MSRP". It was collected and made available on Kaggle by Cooper Union, a private college located in New York City.

Here is a brief overview of the dataset:

- **Number of observations:** 11,159
- **Number of variables:** 16
- **File type:** CSV (Comma Separated Values)

The variables in the dataset are:

- **Make:** the make or brand of the car
- **Model:** the specific model of the car
- **Year:** the year the car was released
- **Engine Fuel Type**: the type of fuel used by the car (gasoline, diesel, etc.)
- **Engine HP:** the horsepower of the car's engine
- **Engine Cylinders:** the number of cylinders in the car's engine
- **Transmission Type**: the type of transmission (automatic or manual)
- **Driven_Wheels:** the type of wheels driven by the car (front, rear, all)
- **Number of Doors:** the number of doors the car has
- **Market Category:** the market category the car belongs to (Luxury, Performance, etc.)
- **Vehicle Size:** the size of the car
- **Vehicle Style:** the style of the car (Sedan, Coupe, etc.)
- **Highway MPG:** the estimated miles per gallon the car gets on the highway
- **City MPG:** the estimated miles per gallon the car gets in the city
- **Popularity:** a ranking of the popularity of the car (based on the number of times it has been viewed on Edmunds.com)
- **MSRP:** the manufacturer's suggested retail price of the car

This dataset could be useful for a variety of data analysis tasks, such as:

- Exploring trends in car features and pricing over time
- Comparing the fuel efficiency of different types of cars
- Investigating the relationship between a car's features and its popularity
- Predicting the price of a car based on its features and market category

However, it's important to note that the dataset was last updated in 2017, so it may not reflect current trends or prices in the automotive industry.

Understanding the Dataset:

The dataset contains information on over 11,000 car models and their specifications, including details on the car's make, model, year, fuel type, engine power, transmission, wheels, number of doors, market category, size, style, estimated miles per gallon, popularity, and manufacturer's suggested retail price (MSRP).

Overall, this dataset could be a valuable resource for data analysts interested in exploring various aspects of the automotive industry and could provide insights that could inform decisions related to product development, marketing, and pricing.

Tech-stack used:

I used Microsoft excel 2022 for this project for its easy to use analysis features like pivot tables etc.

Approach:

firstly I imported the data set given into excel and found the blanks in the dataset and replaced them and converted into a table and started analysing based on the given tasks

Insigths:

A data analyst could use this dataset to gain insights into various aspects of the automotive industry, such as:

Analyzing trends in car features and pricing over time: By examining the variables in the dataset, a data analyst could identify how car features and prices have changed over time, which could help manufacturers make informed decisions about product development and pricing.

Comparing the fuel efficiency of different types of cars: By looking at the MPG variables in the dataset, a data analyst could compare the fuel efficiency of different types of cars and identify which types are the most efficient. This could help consumers make informed decisions about which car to purchase.

Predicting the price of a car based on its features and market category: By using the various features and market category variables in the dataset, a data analyst could develop a model to predict the price of a car. This could help manufacturers and consumers understand how different features affect the price of a car and make informed decisions about pricing and purchasing.

Results:

Form this project we are able to find various things about the cars and its feature and how it affects the price and marketing and find what are the best feature to look for in car and analyze different models and body types etc.

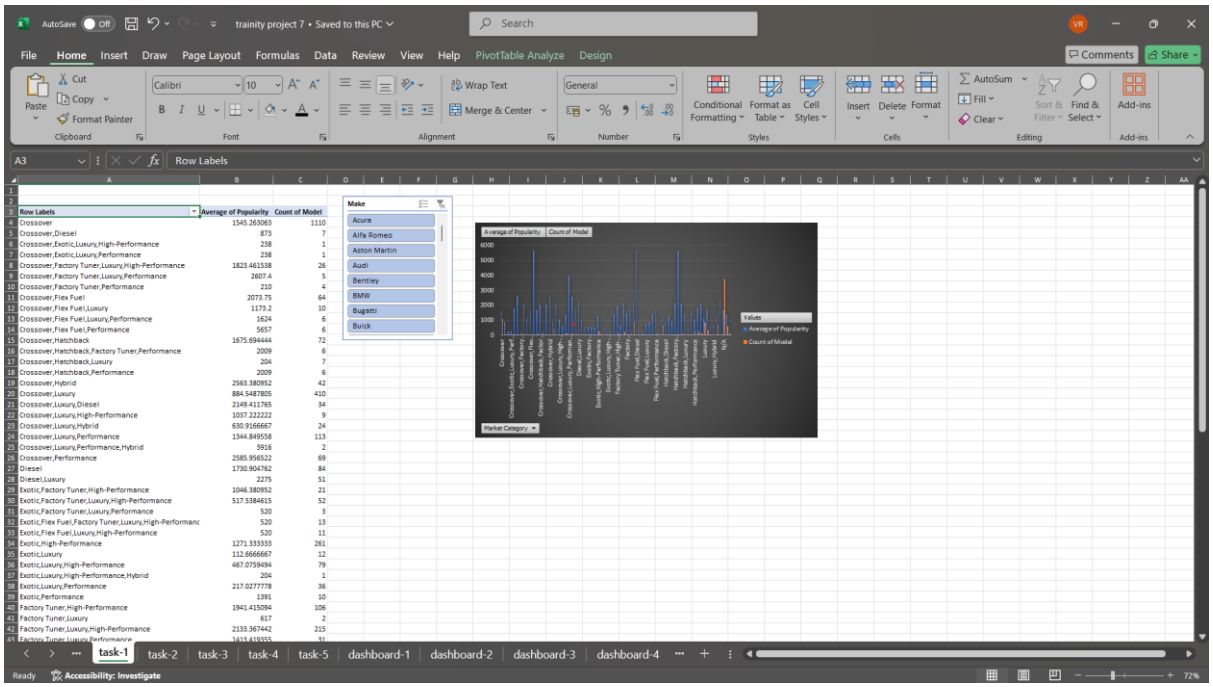
My excel workbook:
https://docs.google.com/spreadsheets/d/1lgPPp3GIXji9OZHWZZYVYSMADftCukIx/edit?usp=sharing&ouid=110428113670481159623&rtpof=true&sd=true

Tasks: Analysis

Before diving into the analysis of the given dataset, it is important to perform thorough data cleaning to ensure accurate and reliable results. Find all the blanks and replace them in the dataset.

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.

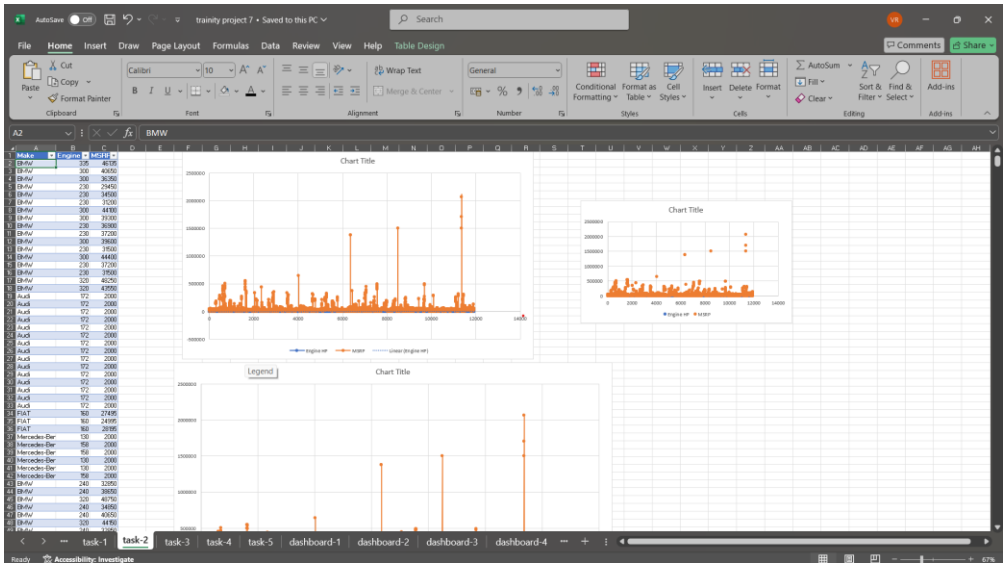


Popularity of model affects the market people go maximum for crossover model.

Insight Required: What is the relationship between a car's engine power and its price?

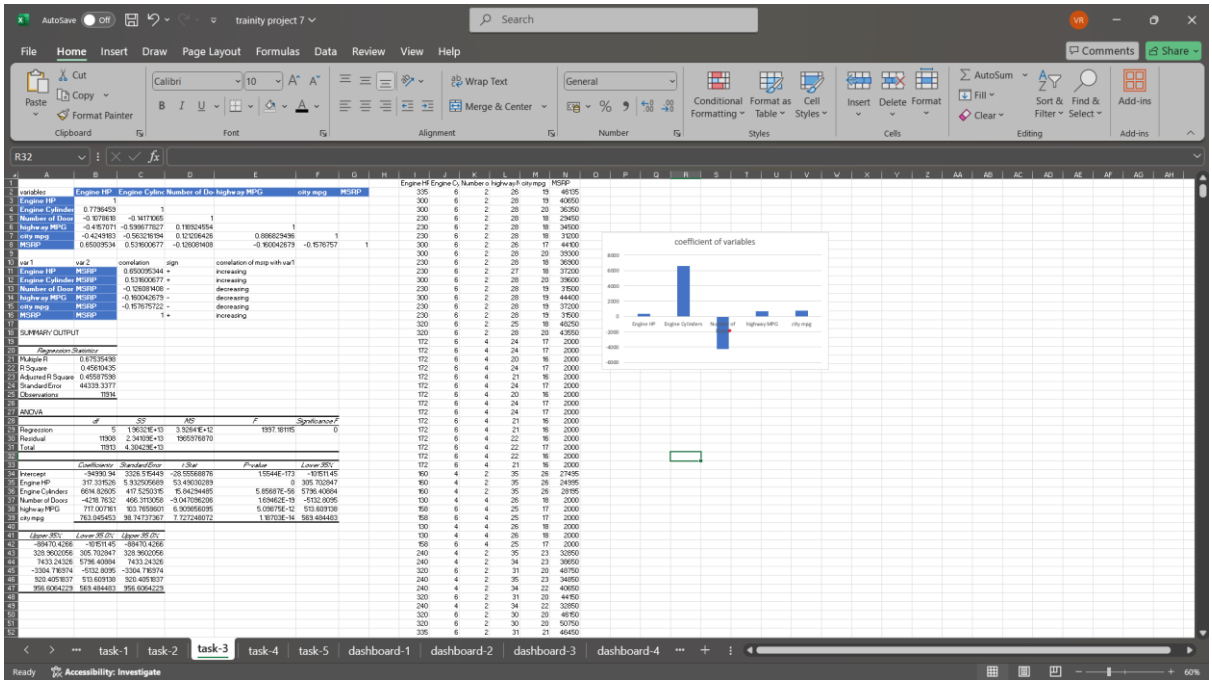
- **Task 2:** Create a scatter chart that plots engine power on the x-axis and price on the y-axis. Add a trendline to the chart to visualize the relationship between these variables.

Both are equally proptional to each as engine hp increases the msrp increases



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.

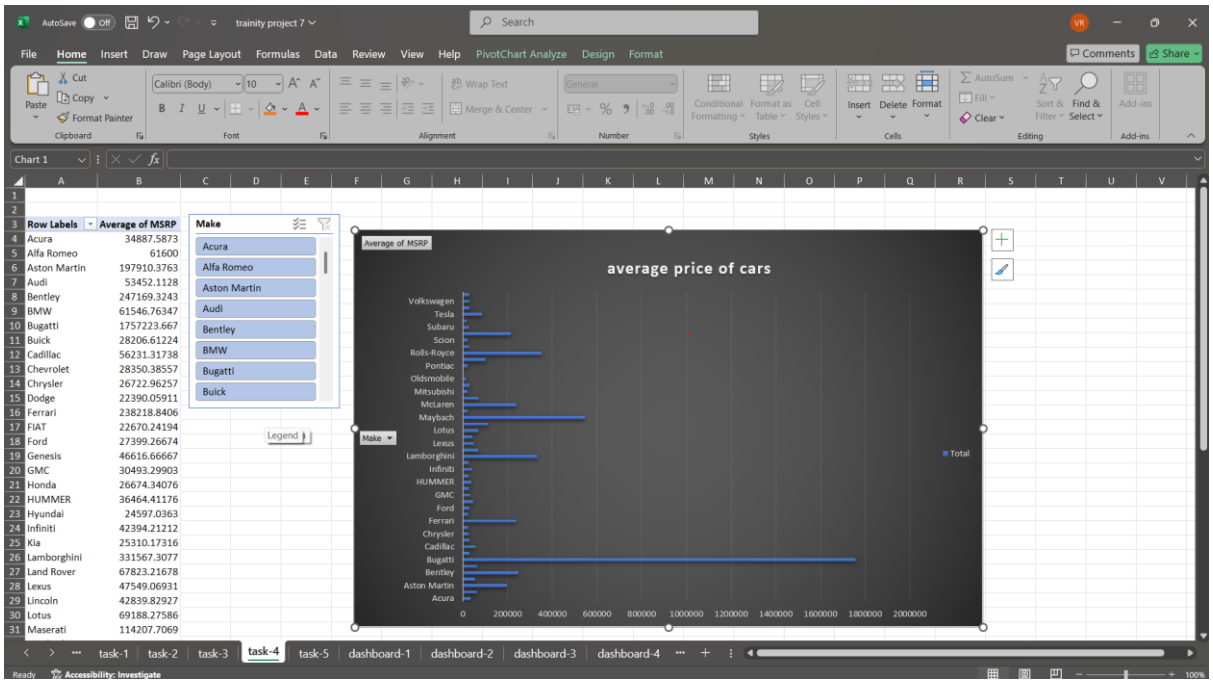


The most important features in a car are engine hp and engine cylinders with respect to the cars price.

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.

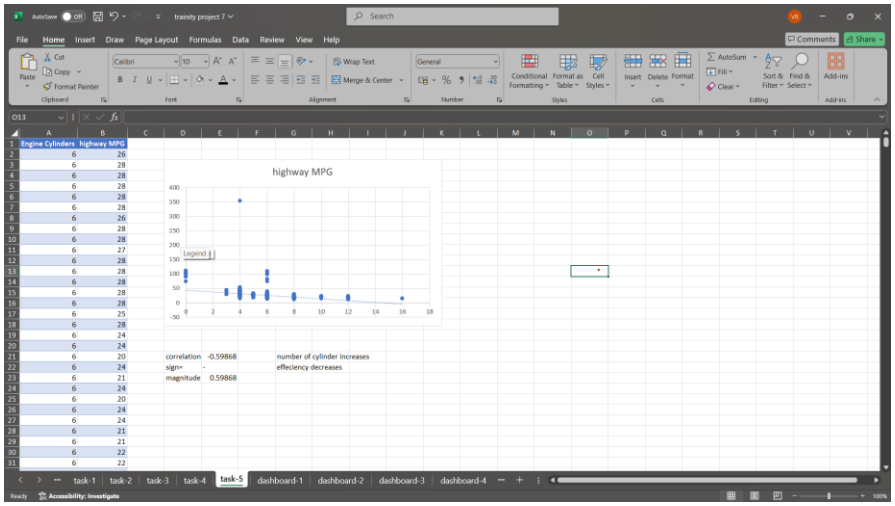


Row Labels	Average of MSRP
Acura	34887.5873
Alfa Romeo	61600
Aston Martin	197910.3763
Audi	53452.1128
Bentley	247169.3243
BMW	61546.76347
Bugatti	1757223.667
Buick	28206.61224
Cadillac	56231.31738
Chevrolet	28350.38557
Chrysler	26722.96257
Dodge	22390.05911
Ferrari	238218.8406
FIAT	22670.24194
Ford	27399.26674
Genesis	46616.66667
GMC	30493.29903
Honda	26674.34076
HUMMER	36464.41176
Hyundai	24597.0363
Infiniti	42394.21212
Kia	25310.17316
Lamborghini	331567.3077
Land Rover	67823.21678
Lexus	47549.06931
Lincoln	42839.82927
Lotus	69188.27586
Maserati	114207.7069

Average price of a car company varies widely according the company buggati is a luxury hypercar brand so it has a highest average car price.

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.



As the number of cylinders increases fuel efficiency decreases.

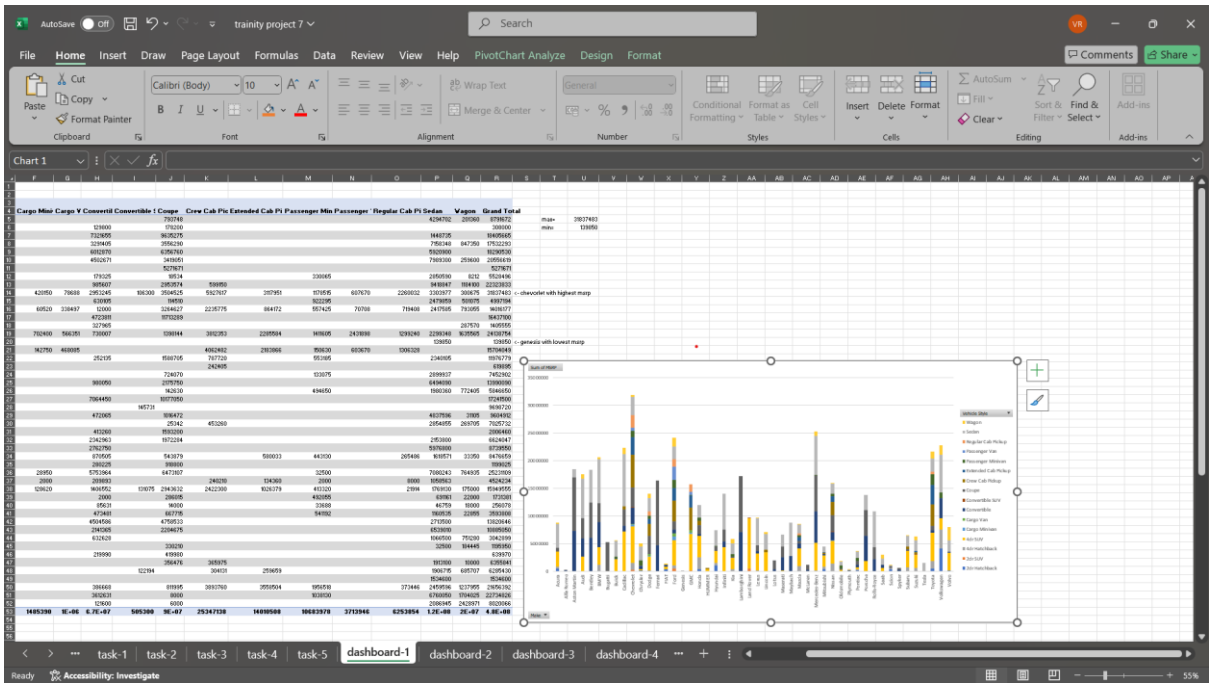
Building the Dashboard:

Now for the Next portion of the Project, you need to create the Interactive Dashboard.

Use filters and slicers to make the chart interactive. The client has requested these questions given below:

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

- **Hints:** 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.

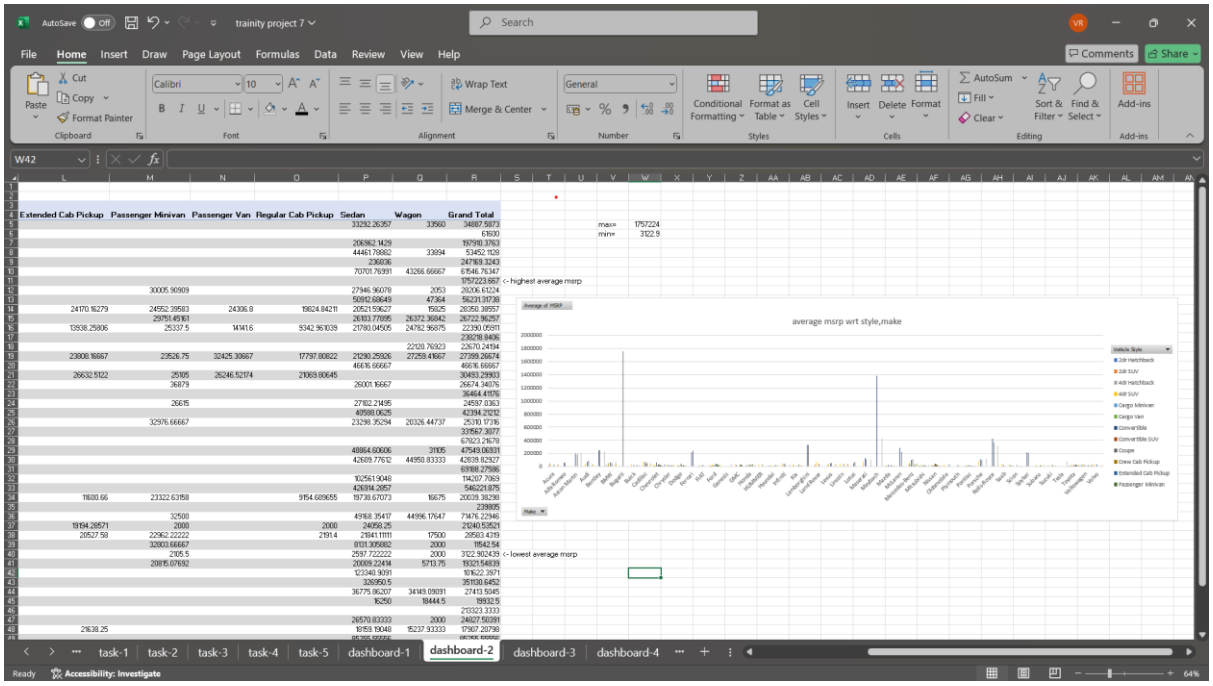


The car prices vary widely with respect to the brand and with each body style with chevorlt and genesis with highest and lowest prices of cars.

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

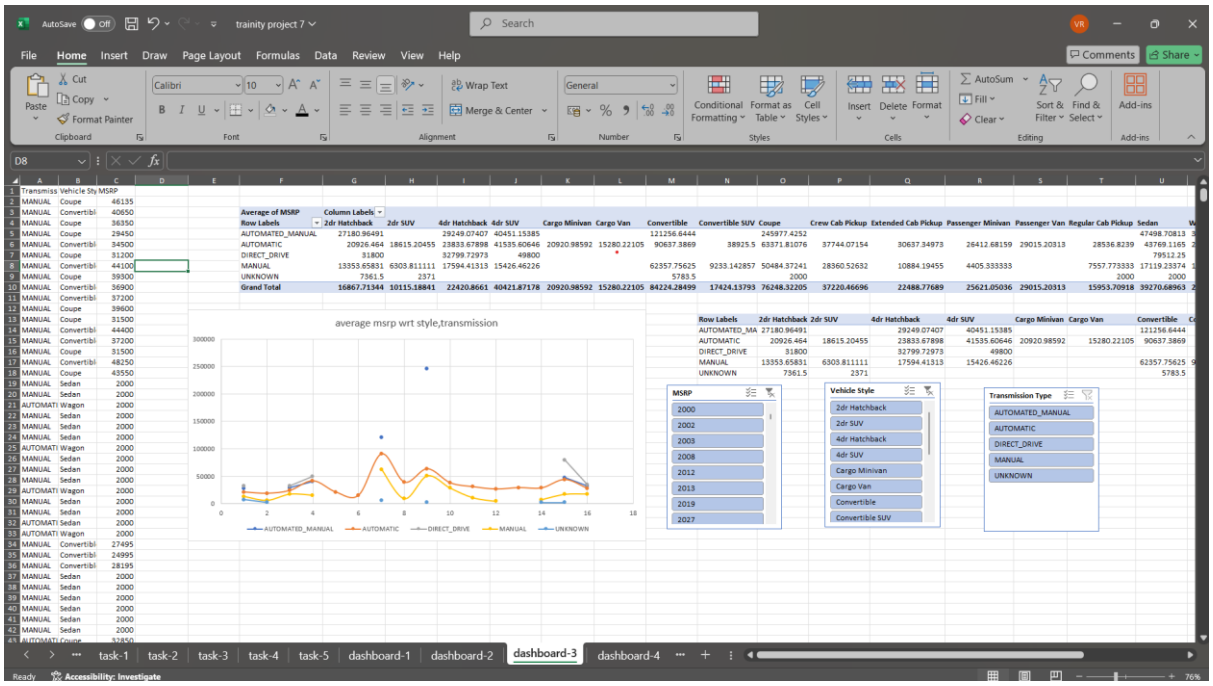
- **Hints:** Clustered column chart to compare the average MSRPs across different car brands and body styles. Calculate the average MSRP for each brand and body style using AVERAGEIF or Pivot Tables.

Bugatti and Plymouth have the highest and lowest average msrps because of their main body style manufacturing and cost of their cars.



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

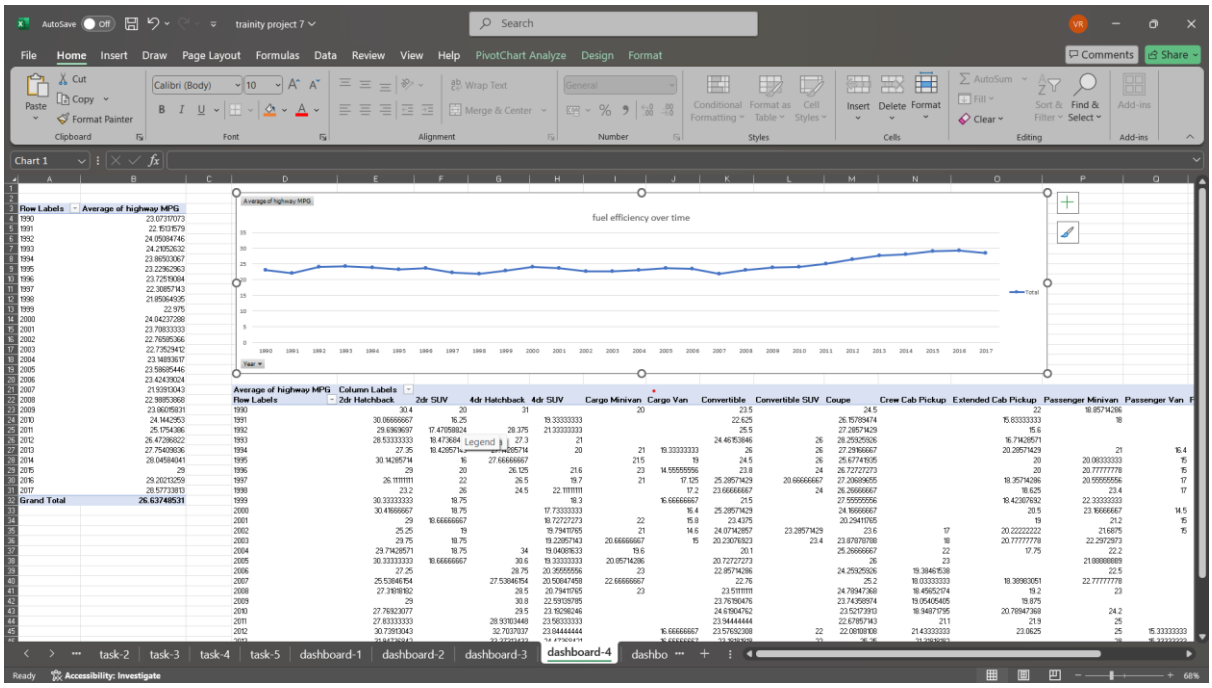
- **Hints:** Scatter plot chart to visualize the relationship between MSRP and transmission type, with different symbols for each body style. Calculate the average MSRP for each combination of transmission type and body style using AVERAGEIFS or Pivot Tables.



Msrp is largely affected by the transmission type and the car body types with automatic transmission type being the highest msrp.

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

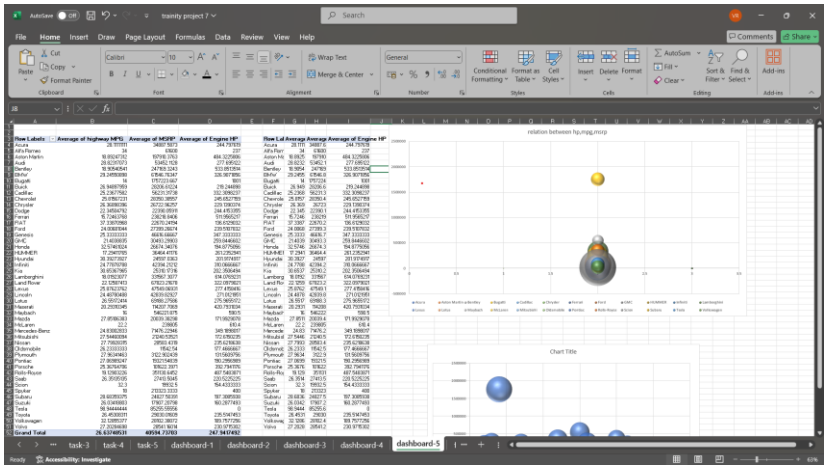
- **Hints:** Line chart to show the trend of fuel efficiency (MPG) over time for each body style. Calculate the average MPG for each combination of body style and model year using AVERAGEIFS or Pivot Tables.



Fuel efficiency does not vary much with respect to model and body type and slowly improving fuel efficiency with each model.

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

Hints: 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.



The cars hp and mpg and price varies across different brands with respect to their manufacturing type and the cost to manufacture the car and the type of car.

The final dashboard with all the slicers: