

Analyzing the Impact of Car Features on Price and Profitability

~TANUSHRI ASTHANA

PROJECT DESCRIPTION:

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: *"Car Features and MSRP"*

Number of observations: 11,159

Number of variables: 16

File type: CSV (Comma Separated Values)

COLUMN

- **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

TECH-STACK USED:

- **MS Excel 2019 for Data Analysis.**
- **MS PowerPoint for Presentation.**

APPROACH:

- **Data Cleaning: Removing duplicates, Null Values, etc.**
- **Understanding the relationships of columns.**
- **Performing descriptive analysis for the given tasks.**
- **Plotting the required graphs.**
- **Established and generated Insights through Data Analysis.**

Data Cleaning

Book1 - Excel

Table Tools Query Tools

Tanushri Asthana

File Home Insert Page Layout Formulas Data Review View Add-ins Help Analytic Solver Table Design Query Tell me what you want to do

BMW

	A	B	C	D	E	F	G	H	I	J
1	Make	Model	Year	Engine Fuel Type	Engine HP	Engine Cylinders	Transmission Type	Driven_Wheels	Number of Doors	Market Category
2	BMW	1 Series M	2011	premium unleaded (required)	335	6	MANUAL	rear wheel drive	2	Factory Tuner,Luxury,High-Performance
3	BMW	1 Series	2011	premium unleaded (required)	300	6	MANUAL	rear wheel drive	2	Luxury,Performance
4	BMW	1 Series	2011	premium unleaded (required)	300	6	MANUAL	rear wheel drive	2	Luxury,High-Performance
5	BMW	1 Series	2011	premium unleaded (required)	230	6	MANUAL	rear wheel drive	2	Luxury,Performance
6	BMW	1 Series	2011	premium unleaded (required)	230	6	MANUAL	rear wheel drive	2	Luxury
7	BMW	1 Series	2012	premium unleaded (required)	230	6	MANUAL	rear wheel drive	2	Luxury,Performance
8	BMW	1 Series	2012	premium unleaded (required)	300	6	MANUAL	rear wheel drive	2	Luxury,Performance
9	BMW	1 Series	2012	premium unleaded (required)	300	6	MANUAL	rear wheel drive	2	Luxury,High-Performance
10	BMW	1 Series	2012	premium unleaded (required)	230	6	MANUAL	rear wheel drive	2	Luxury
11	BMW	1 Series	2013	premium				el drive	2	Luxury
12	BMW	1 Series	2013	premium				el drive	2	Luxury,High-Performance
13	BMW	1 Series	2013	premium				el drive	2	Luxury,Performance
14	BMW	1 Series	2013	premium				el drive	2	Luxury,Performance
15	BMW	1 Series	2013	premium				el drive	2	Luxury
16	BMW	1 Series	2013	premium				el drive	2	Luxury,High-Performance
17	BMW	1 Series	2013	premium unleaded (required)	320	6	MANUAL	rear wheel drive	2	Luxury,High-Performance
18	Audi	100	1992	regular unleaded	172	6	MANUAL	front wheel drive	4	Luxury
19	Audi	100	1992	regular unleaded	172	6	AUTOMATIC	all wheel drive	4	Luxury
20	Audi	100	1992	regular unleaded	172	6	MANUAL	all wheel drive	4	Luxury
21	Audi	100	1993	regular unleaded	172	6	MANUAL	front wheel drive	4	Luxury
22	Audi	100	1993	regular unleaded	172	6	AUTOMATIC	all wheel drive	4	Luxury
23	Audi	100	1993	regular unleaded	172	6	MANUAL	all wheel drive	4	Luxury
24	Audi	100	1994	regular unleaded	172	6	AUTOMATIC	front wheel drive	4	Luxury
25	Audi	100	1994	regular unleaded	172	6	MANUAL	all wheel drive	4	Luxury
26	Audi	100	1994	regular unleaded	172	6	MANUAL	front wheel drive	4	Luxury
27	Audi	100	1994	regular unleaded	172	6	AUTOMATIC	front wheel drive	4	Luxury
28	Audi	100	1994	regular unleaded	172	6	AUTOMATIC	all wheel drive	4	Luxury
29	FIAT	124 Spider	2017	premium unleaded (recommended)	160	4	MANUAL	rear wheel drive	2	Performance
30	FIAT	124 Spider	2017	premium unleaded (recommended)	160	4	MANUAL	rear wheel drive	2	Performance

715 duplicate values found and removed; 11199 unique values remain. Note that counts may include empty cells, spaces, etc.

OK

Car_data Car_data (2) Sheet1

Ready Accessibility: Investigate

Average: 5564.053872 Count: 190519 Sum: 529736877

26°C Mostly cloudy

Search

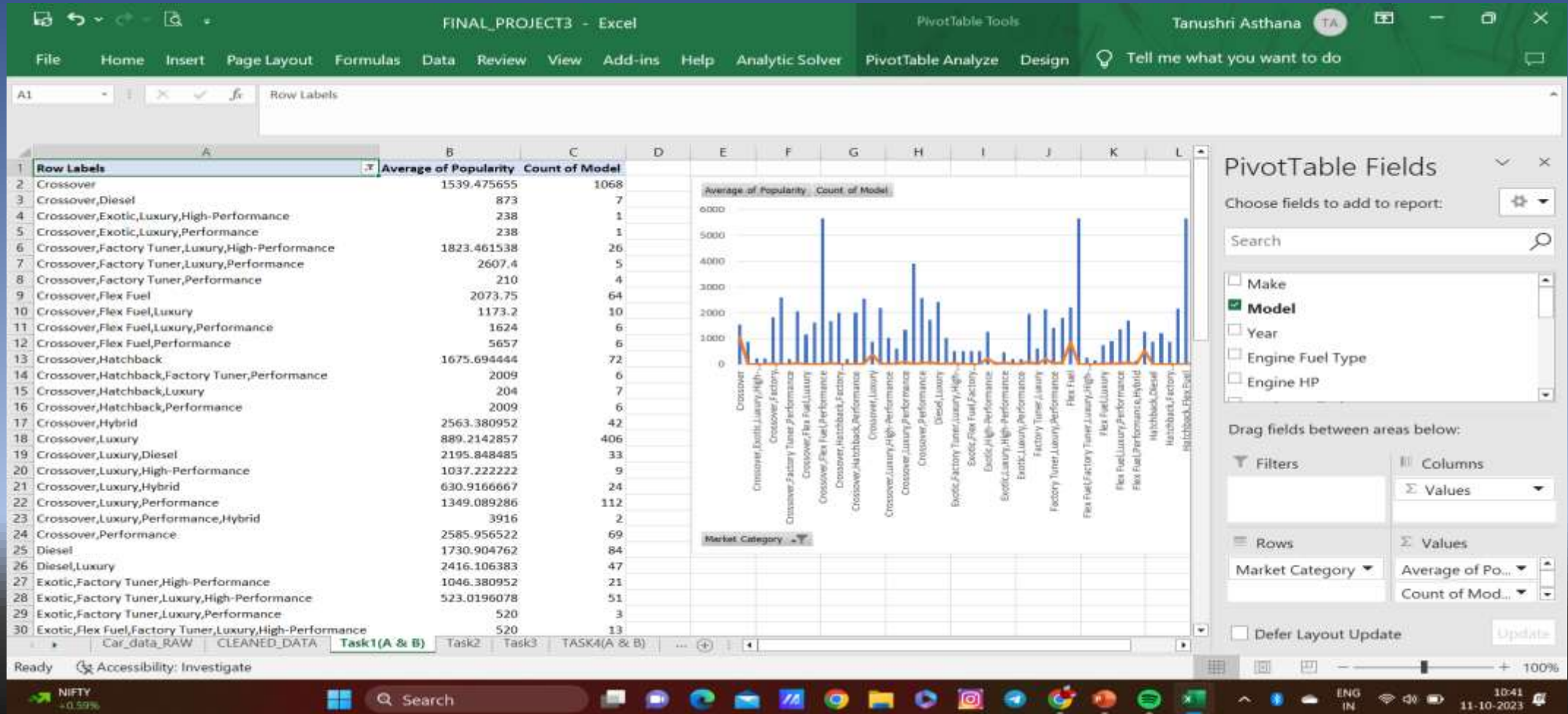
ENG

21:03 04-10-2023

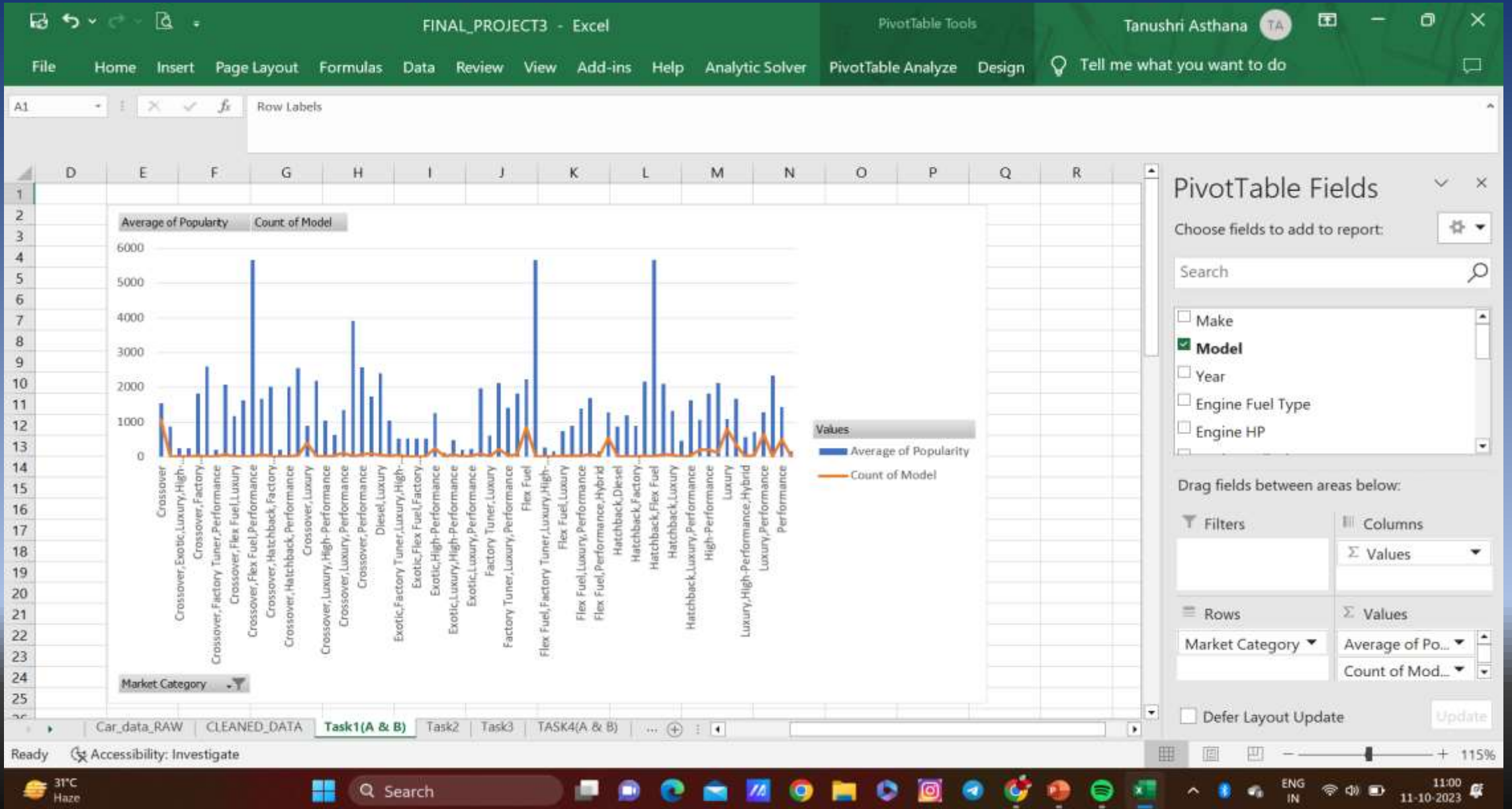
TASK (1)

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

A. Create a pivot table that shows the number of car models in each market category and their corresponding popularity scores.



(B) Create a combo chart that visualizes the relationship between market category and popularity.





Desired Insight:

Market Categories that are most popular:

- ❖ Hatchback vehicles with Flex fuel
- ❖ Diesel vehicles with flex fuel
- ❖ Performance, Crossover, and Flex Fuel vehicle

Market Categories that are not popular:

- ❖ Luxury
- ❖ Exotic

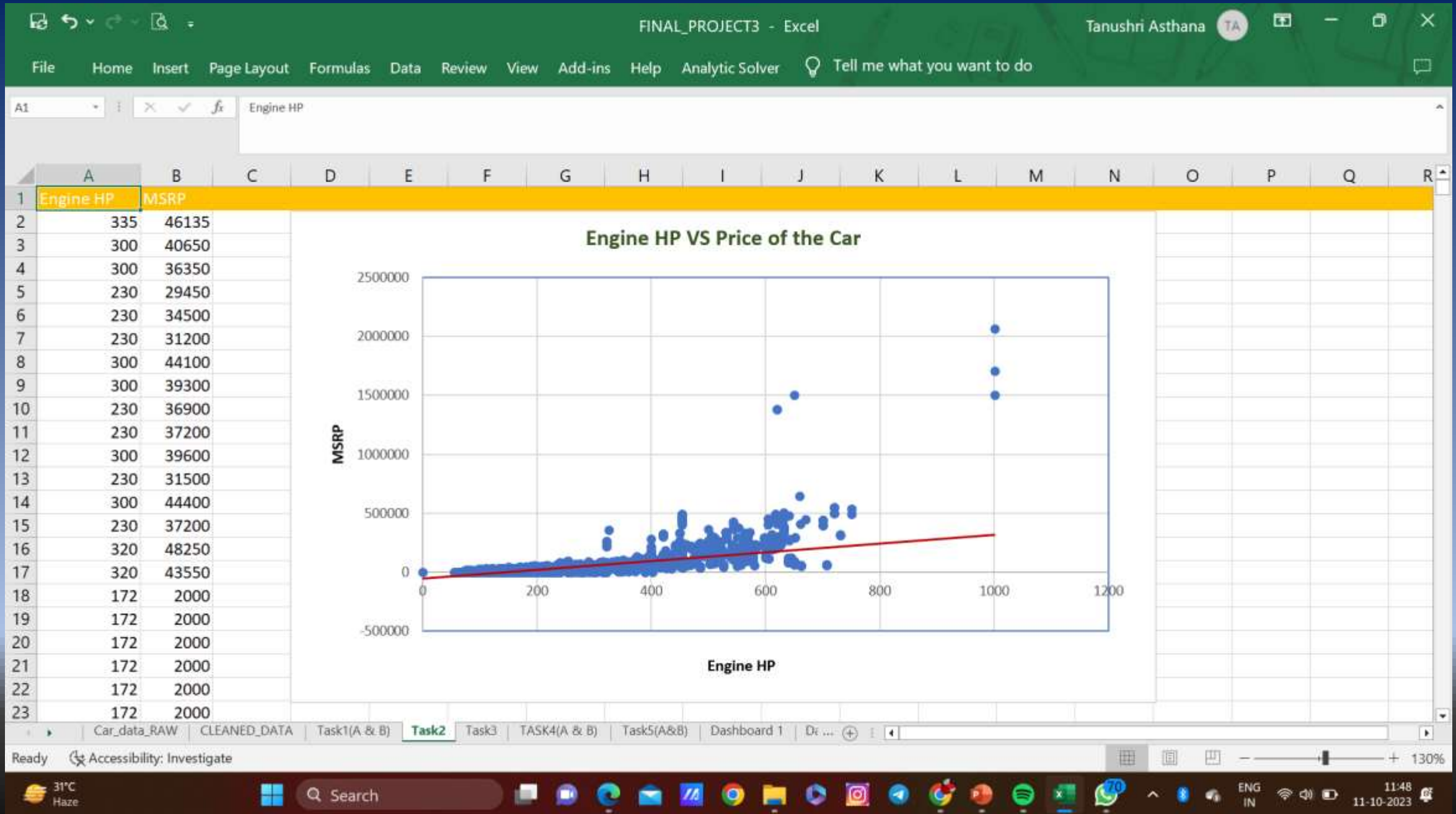
Task 2

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

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.

Desired Insight:

- ❖ There is a presence of very strong linear relationship.
- ❖ With respect to linear trendline we can eventually notice with the increase in Horsepower of the engine, there is an increase in the Price.



Task 3

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

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.

Desired Insight:

- ❖ The positive regression coefficient of engine cylinders shows that increasing the number of engine cylinders will lead to an automatic increase in the price of the car.
- ❖ The negative correlation signifies that when one of these variables increases, the others tend to decrease. There is a negative correlation between the MSRP and the variables of highway mileage, city MPG, and the number of doors.

T54

SUMMARY OUTPUT

Regression Statistics

Multiple R	0.683387437
R Square	0.46701839
Adjusted R Square	0.466730032
Standard Error	45078.99614
Observations	11097

ANOVA

	df	SS	MS	F	Significance F
Regression	6	1.9747E+13	3.29117E+12	1619.578688	0
Residual	11090	2.25362E+13	2032115893		
Total	11096	4.22832E+13			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-97167.87274	3898.078612	-24.92711985	1.7136E-133	-104808.8004	-89527	-104808.8	-89526.945
Engine HP	320.4942139	6.3774589	50.25421863	0	307.9932598	332.995	307.99326	332.99517
Engine Cylinders	7578.79133	461.2602827	16.43061762	5.88653E-60	6674.639109	8482.94	6674.63911	8482.9436
Number of Doors	-4980.209981	496.4047724	-10.03255863	1.38198E-23	-5953.251655	-4007.2	-5953.25165	-4007.1683
highway MPG	503.5834871	109.2773107	4.608307836	4.10488E-06	289.3805157	717.786	289.380516	717.78646
city mpg	1253.468123	125.6629389	9.974843293	2.46287E-23	1007.146405	1499.79	1007.14641	1499.7898
Popularity	-3.553387511	0.297352947	-11.95006653	1.02989E-32	-4.136252193	-2.9705	-4.13625219	-2.9705228

CORRELATION

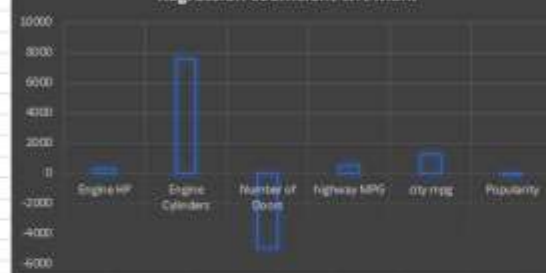
	Engine HP	Engine Cylinders	Number of Doors	highway MPG	city mpg	Popularity
Engine HP	1					
Engine Cylinders	0.788318554	1				
Number of Doors	-0.128492859	-0.147591136	1			
highway MPG	-0.42373423	-0.614703148	0.116919706	1		
city mpg	-0.474810111	-0.63445975	0.137386114	0.842834493	1	
Popularity	0.041418279	0.042705254	-0.058978607	-0.024926937	-0.0072	1

Engine HP Line Fit Plot

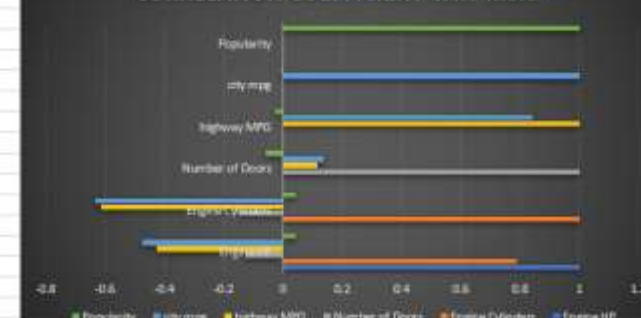
\$29,00,000.00

\$30,00,000.00

Regression coefficient wrt MSRP



CORRELATION COEFFICIENT WRT MSRP



Car_data_RAW

CLEANED_DATA

Task1(A & B)

Task2

Task3

TASK4(A & B)

Task5(A&B)

Dashboard 1

Di ...

Ready Accessibility: Investigate

25°C
Haze

Search

ENG
IN22:10
11-10-2023

Task 4

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

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

Desired Insight:

- ❖ According to the graph, the manufacturer "Bugatti" has the highest average price, followed by "Maybach," "Rolls-Royce," and "Lamborghini."

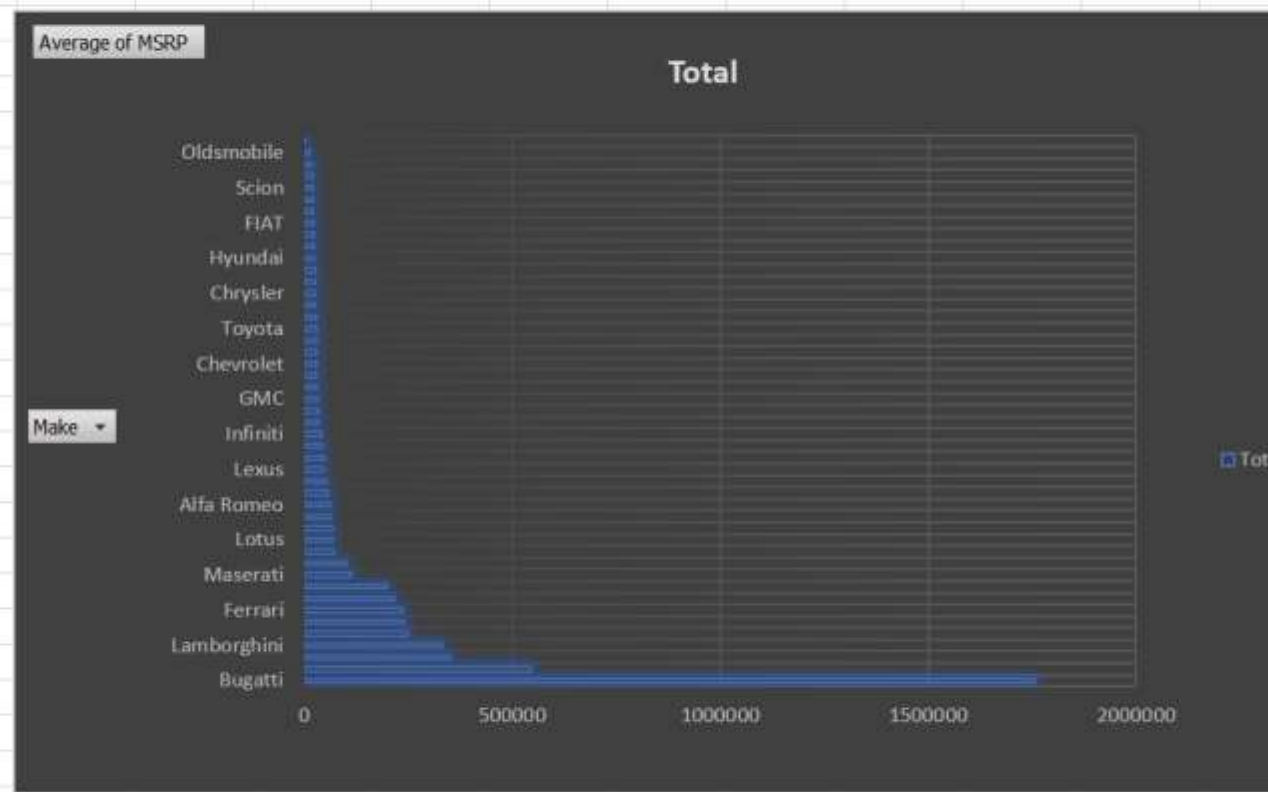
Bugatti	1757223.667
Maybach	546221.875
Rolls-Royce	351130.6452

- ❖ "Plymouth" has the lowest average price.

Suzuki	18021.0531
Oldsmobile	12843.79545
Plymouth	3296.873239

A1 Row Labels

1	Row Labels	Average of MSRP
2	Bugatti	1757223.667
3	Maybach	546221.875
4	Rolls-Royce	351130.6452
5	Lamborghini	331567.3077
6	Bentley	247169.3243
7	McLaren	239805
8	Ferrari	237383.8235
9	Spyker	214990
10	Aston Martin	198123.4615
11	Maserati	113684.4909
12	Porsche	101622.3971
13	Mercedes-Benz	72135.02647
14	Lotus	68377.14286
15	Land Rover	68067.08633
16	BMW	62162.55864
17	Alfa Romeo	61600
18	Cadillac	56368.26515
19	Audi	54574.1215
20	Lexus	47549.06931
21	Genesis	46616.66667
22	Lincoln	43560.01316
23	Infiniti	42640.27134
24	HUMMER	36464.41176
25	Acura	35027.10279



PivotTable Fields

Choose fields to add to report:

Search

- ☒ Make
- ☐ Model
- ☐ Year
- ☐ Engine Fuel Type
- ☐ Engine HP

Drag fields between areas below:

Filters	Columns

Rows	Values
Make	Average of MSRP

☐ Defer Layout Update

Update

Task 5

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

(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.

(B) Calculate the correlation coefficient between the number of cylinders and highway MPG to quantify the strength and direction of the relationship.

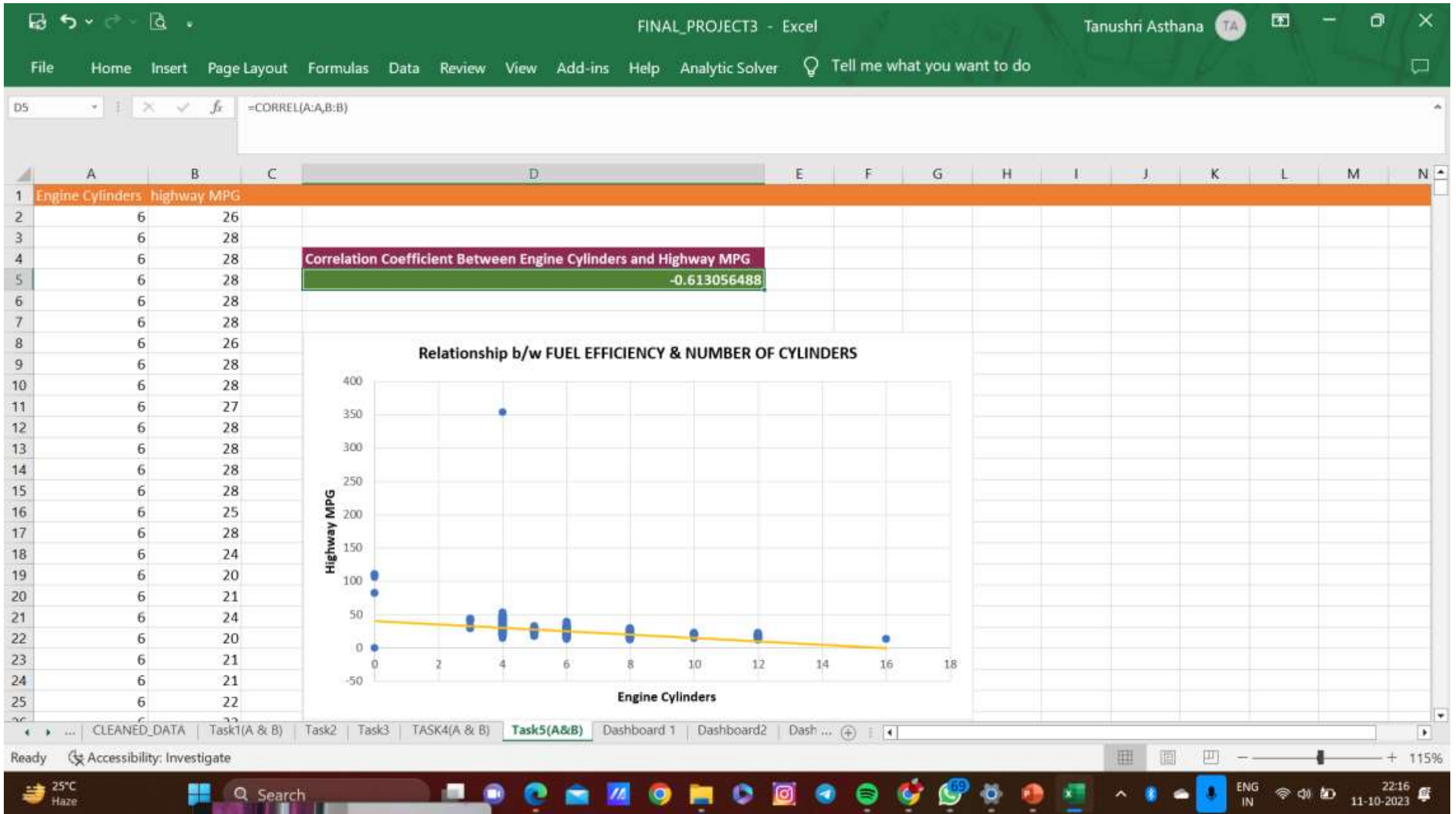
Desired Insight

❖ Negative trendline that states there is a negation relation between the number of cylinders and fuel efficiency, which means decrease in number of cylinders increases the fuel efficiency.

❖ The correlation coefficient turns out to be negative.

Correlation Coefficient Between Engine Cylinders and Highway MPG

-0.613056488



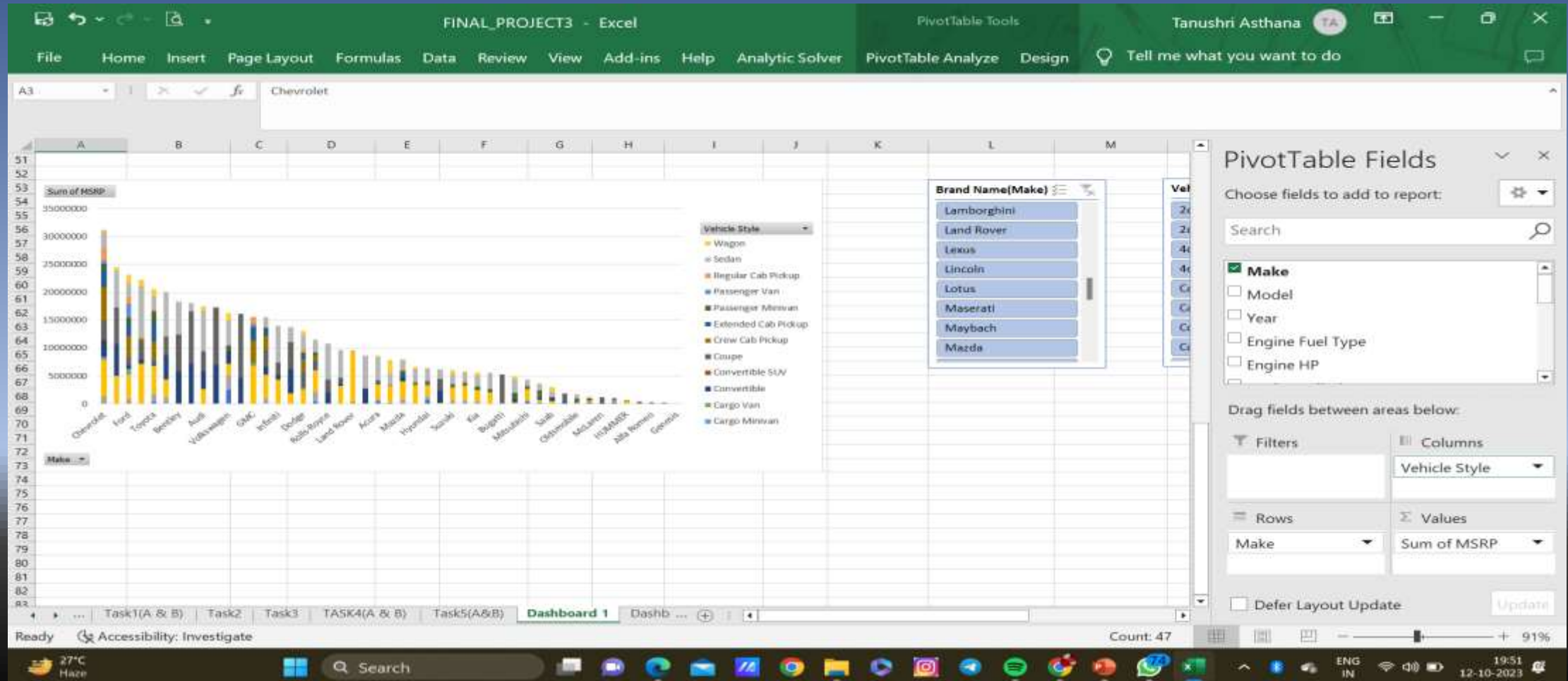
DASHBOARD:



Dashboard 1:

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

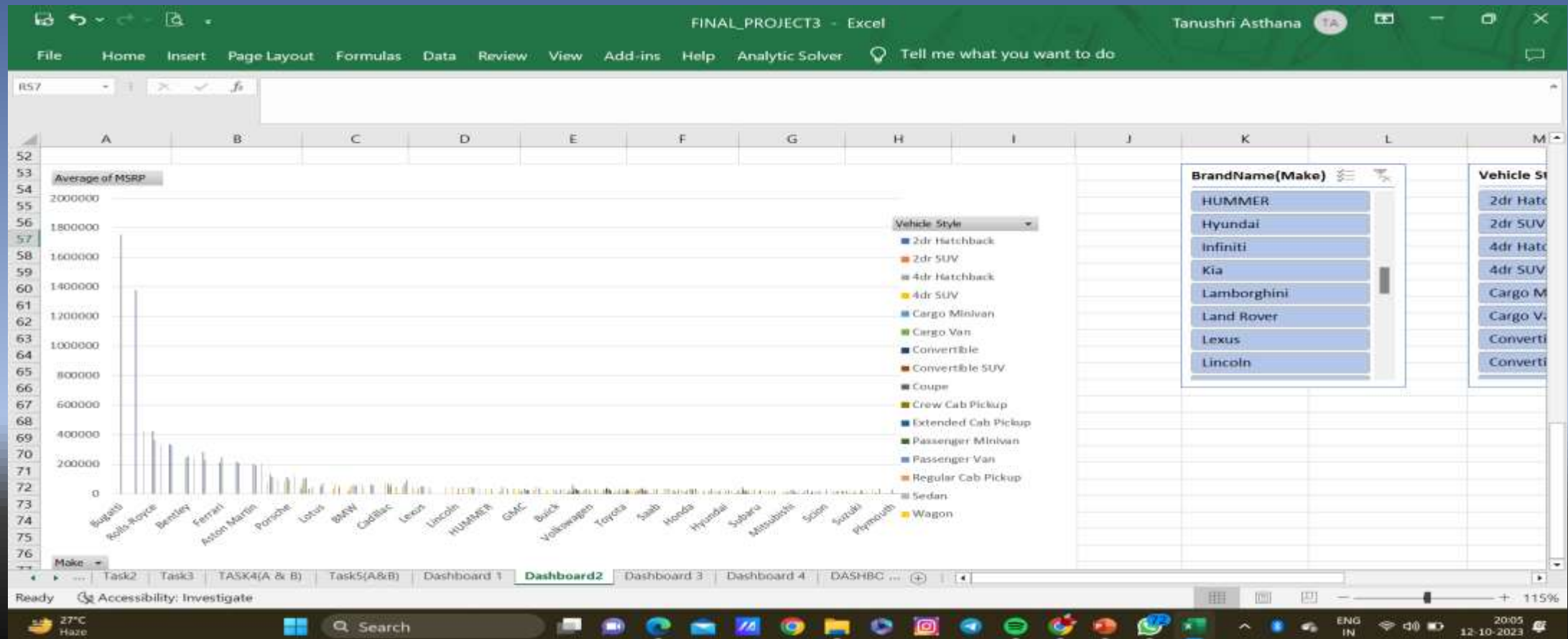
- ❖ **Desired Insight:** "Chevrolet" has the highest Maximum Selling Retail Price (MSRP) among all manufacturers, followed by "Mercedes-Benz" in the second position. Among various vehicle styles, the "Sedan" category exhibits the highest maximum selling retail price.



DASHBOARD 2

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

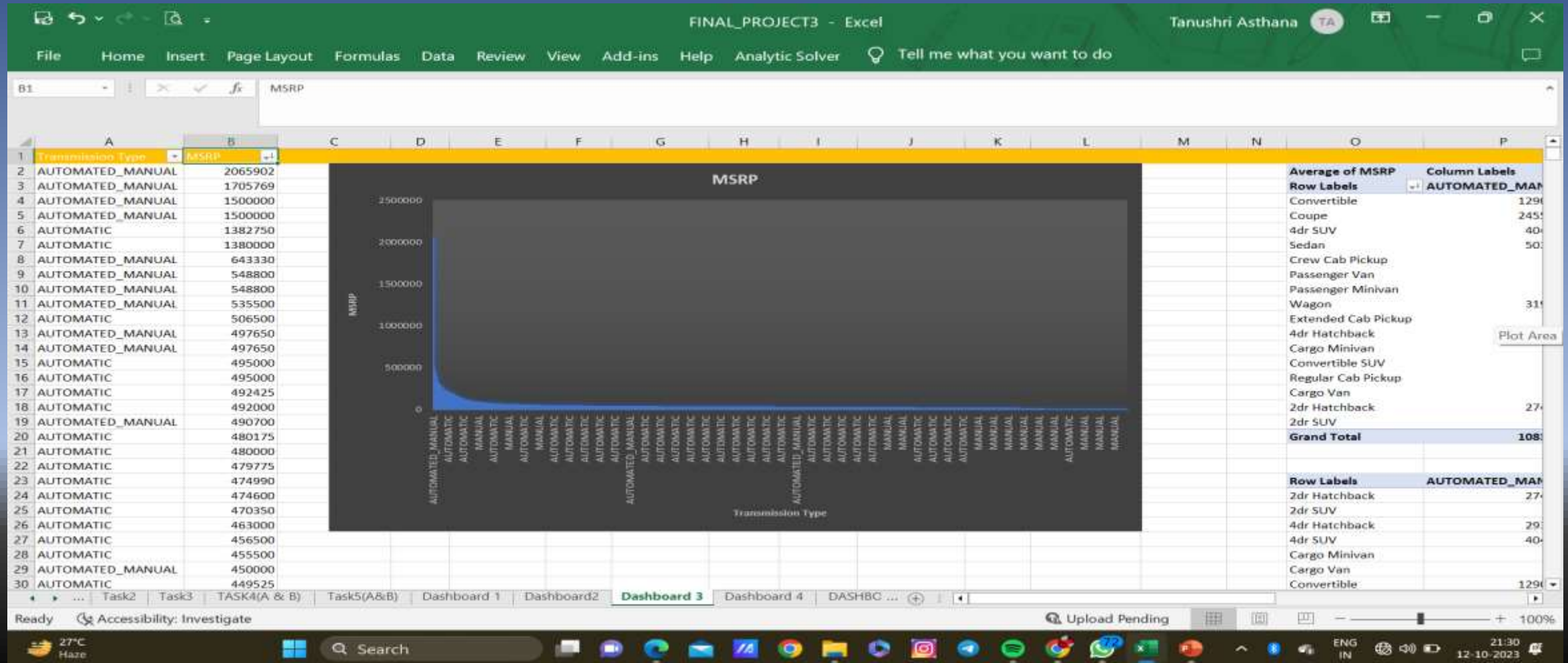
- ❖ **Desired Insight:** Bugatti has the highest MSRP while Coupe has the most popular vehicle style. Maybach is the second most popular brand.
- ❖ Plymouth has the lowest price for 2dr and 4dr hatchback models.



Dashboard 3

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

Desired Insight: "Automated_Manual" has the highest MSRP while "Manual" transmissions have the lowest MSRP.



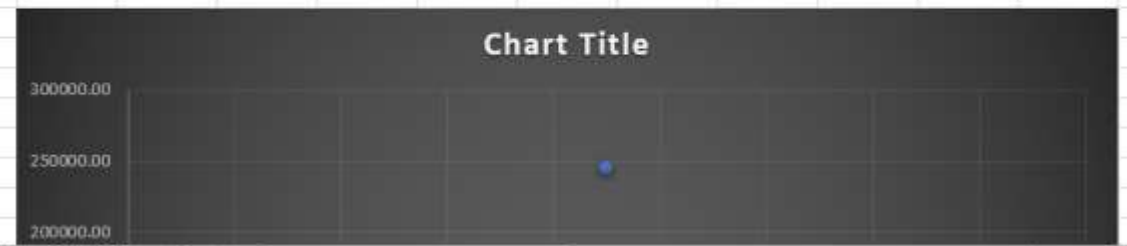
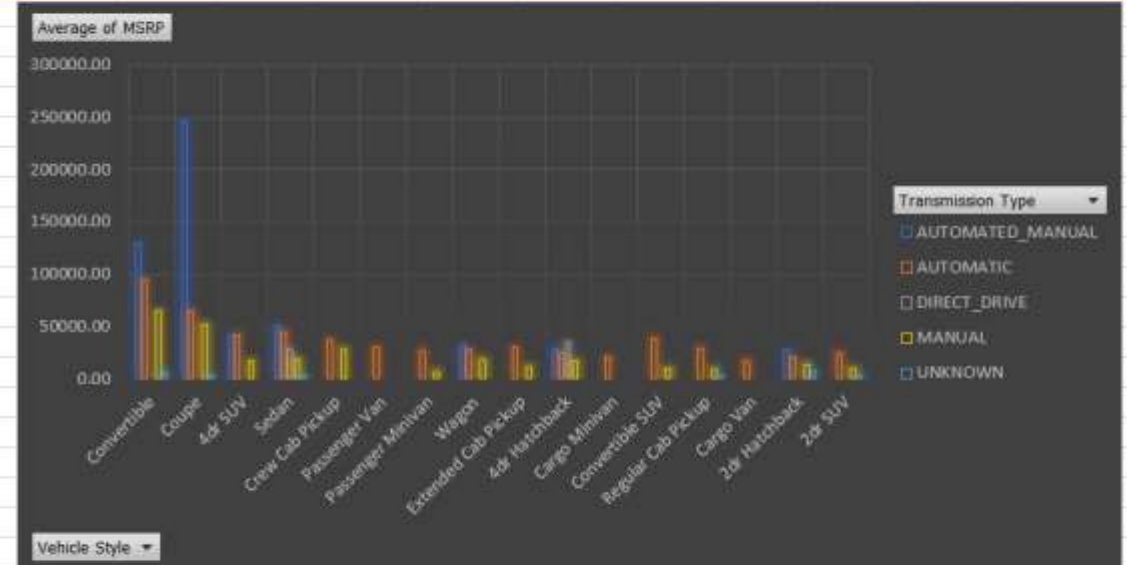
B1

MSRP

O P Q R S T U V W X Y Z AA AB AC AD AE AF AG

1						
2	Average of MSRP	Column Labels				
3	Row Labels	AUTOMATED_MANUAL	AUTOMATIC	DIRECT_DRIVE	MANUAL	UNKNOWN
4	Convertible	129082.23	95153.31		64794.34	5783.50
5	Coupe	245588.36	65031.19		51524.64	2000.00
6	4dr SUV	40451.15	41658.40		17422.09	
7	Sedan	50385.39	44705.13	27822.50	17557.26	2000.00
8	Crew Cab Pickup		37718.95		28233.11	
9	Passenger Van		30578.07			
10	Passenger Minivan		26570.02		6510.00	
11	Wagon	31985.28	28219.46		18398.58	
12	Extended Cab Pickup		30711.45		11553.30	
13	4dr Hatchback	29347.05	23888.74	34511.92	17500.36	
14	Cargo Minivan		20292.93			
15	Convertible SUV		38925.50		9594.80	
16	Regular Cab Pickup		28536.82		8759.45	2000.00
17	Cargo Van		17019.30			
18	2dr Hatchback	27470.42	20784.10		12840.66	7361.50
19	2dr SUV		24153.61		9173.02	2371.00
20	Grand Total	108381.52	41847.32	33620.00	28284.52	3647.83

23	Row Labels	AUTOMATED_MANUAL	AUTOMATIC	DIRECT_DRIVE	MANUAL	UNKNOWN
24	2dr Hatchback	27470.42	20784.10		12840.66	7361.50
25	2dr SUV		24153.61		9173.02	2371.00
26	4dr Hatchback	29347.05	23888.74	34511.92	17500.36	
27	4dr SUV	40451.15	41658.40		17422.09	
28	Cargo Minivan		20292.93			
29	Cargo Van		17019.30			
30	Convertible	129082.23	95153.31		64794.34	5783.50

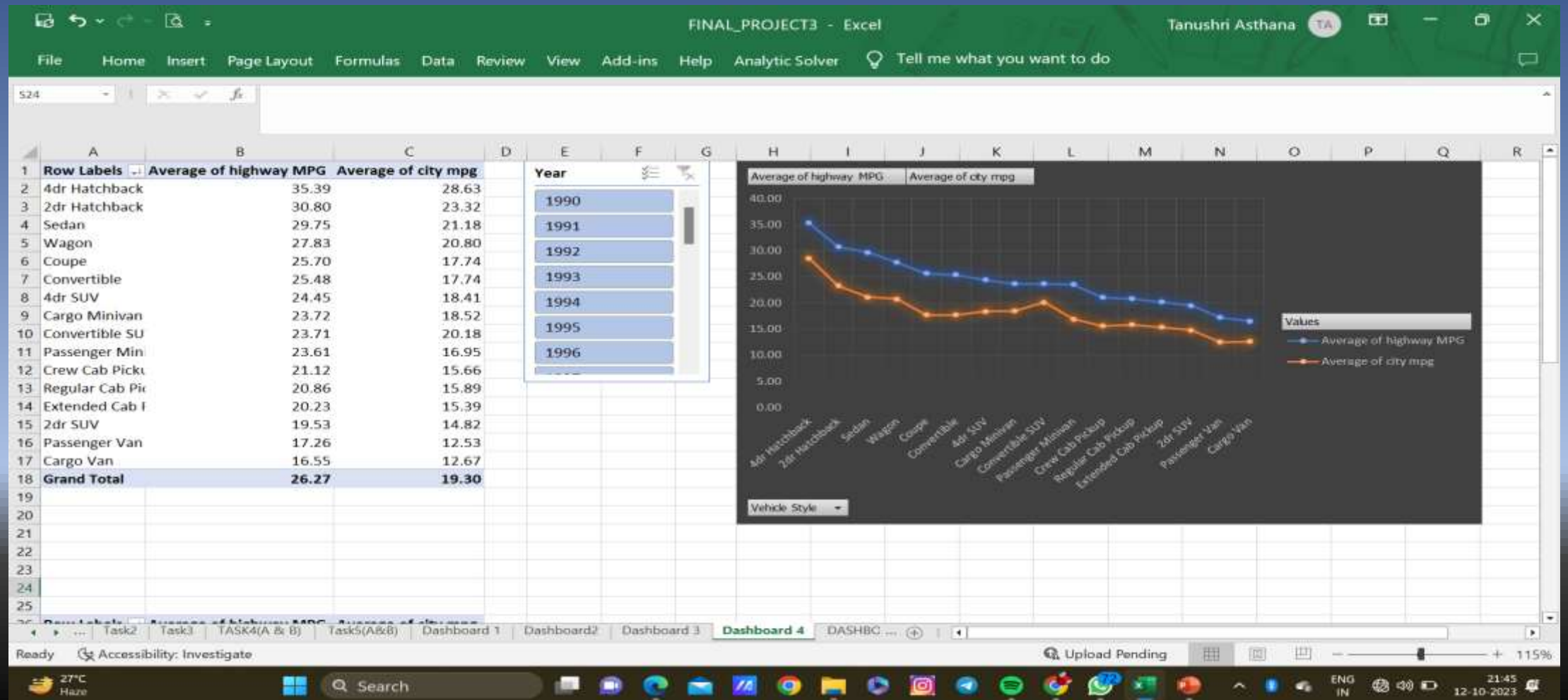


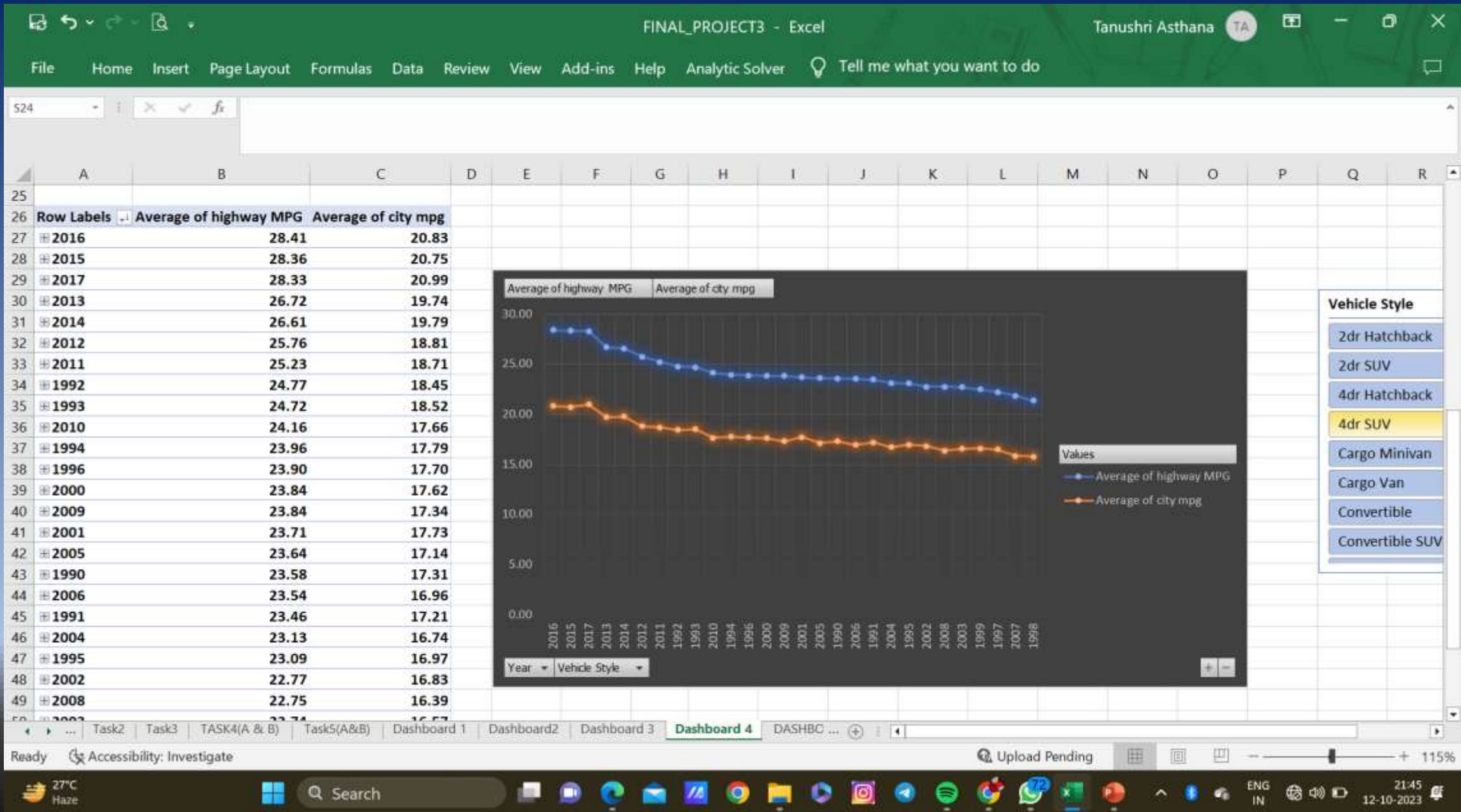
Task2 Task3 TASK4(A & B) Task5(A&B) Dashboard 1 Dashboard2 Dashboard 3 Dashboard 4 DASHBOC ...

Dashboard 4

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

- ❖ **Desired Insight:** The "4dr Hatchback" vehicle model is identified as the most fuel-efficient, with an average of 35 miles per gallon on the highway and 28.6 miles per gallon in the city.
- ❖ The "cargo van" is the least fuel-efficient vehicle type, with an average of 12 mpg in cities and 16 mpg on the highway.





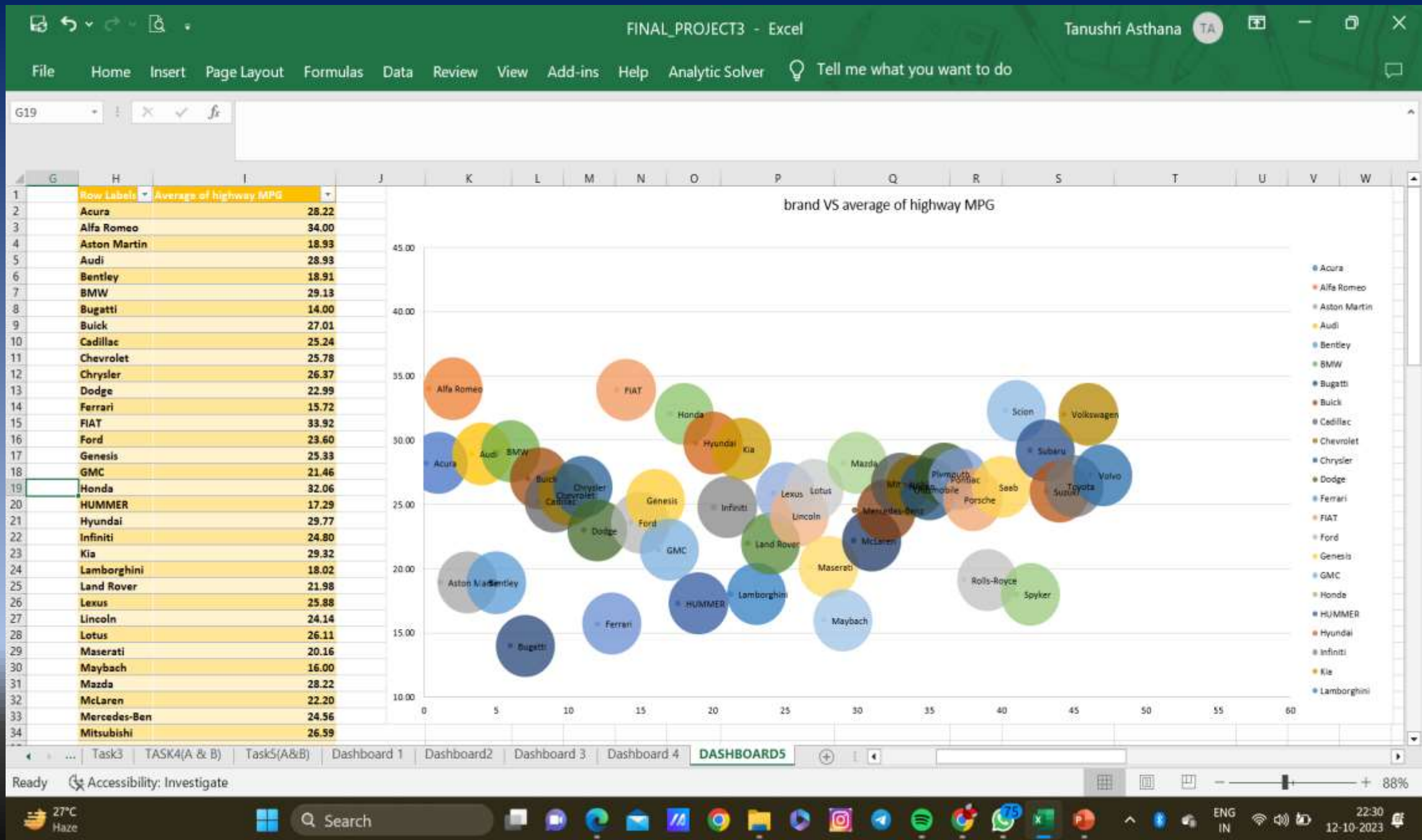
Dashboard 5

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

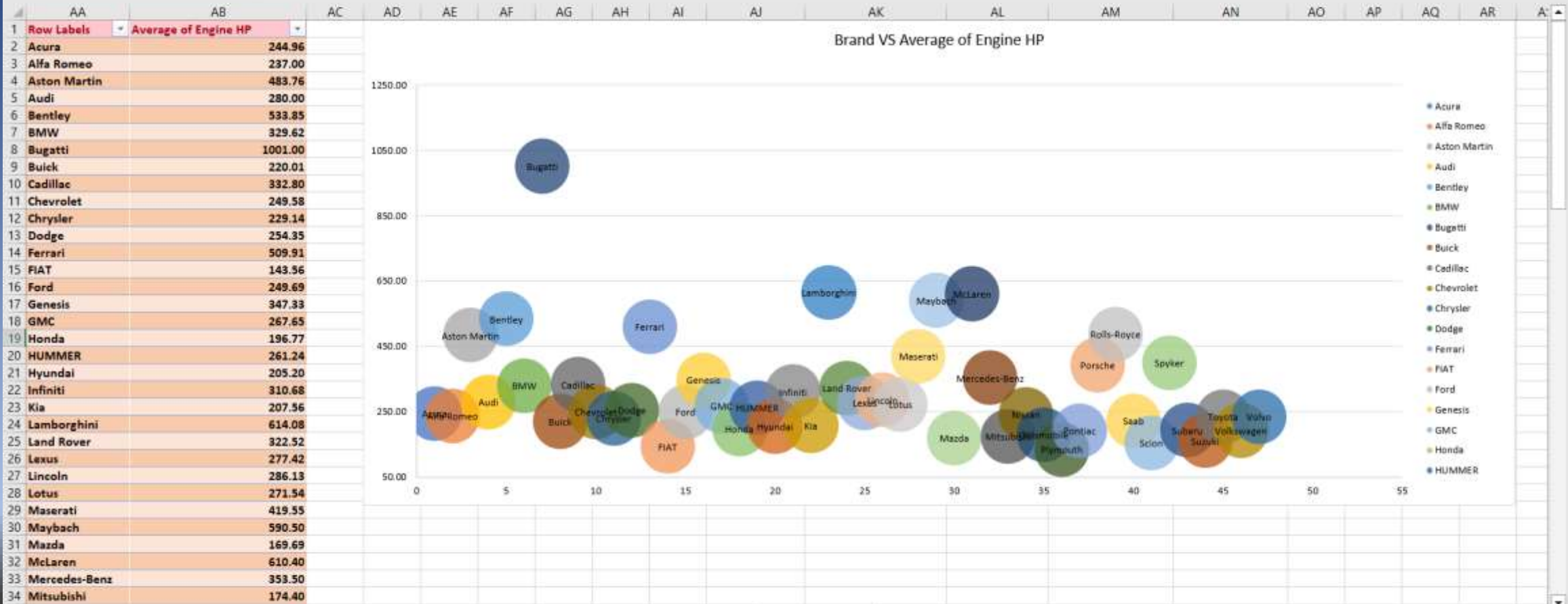
Desired Insight:

- ❖ “Bugatti” has the highest MSRP.
- ❖ “Lamborghini” has the second most position with respect to horsepower i.e. 614.
- ❖ Vehicles under the “Plymouth” brand are the least expensive.
- ❖ The efficiency of the engine cylinders decreases as the car becomes more expensive.

Row Labels	Average of highway MPG	Average of Engine HP	Average of city mpg	Average of MSRP
Bugatti	14.00	1001.00	8.00	1757223.67
Maybach	16.00	590.50	10.00	546221.88
Rolls-Royce	19.13	487.55	11.84	351130.65
Lamborghini	18.02	614.08	11.52	331567.31
Bentley	18.91	533.85	11.55	247169.32

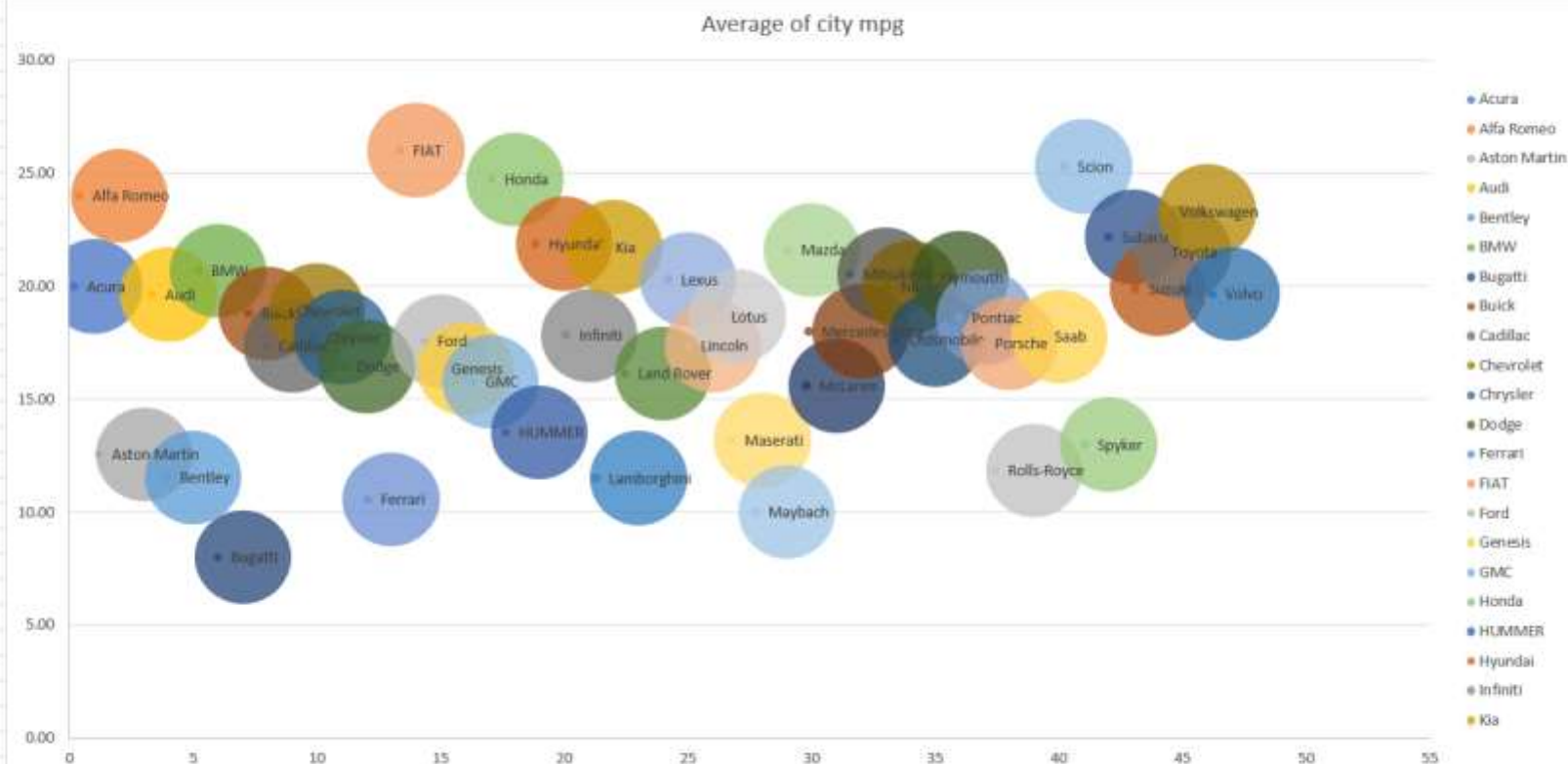


G19

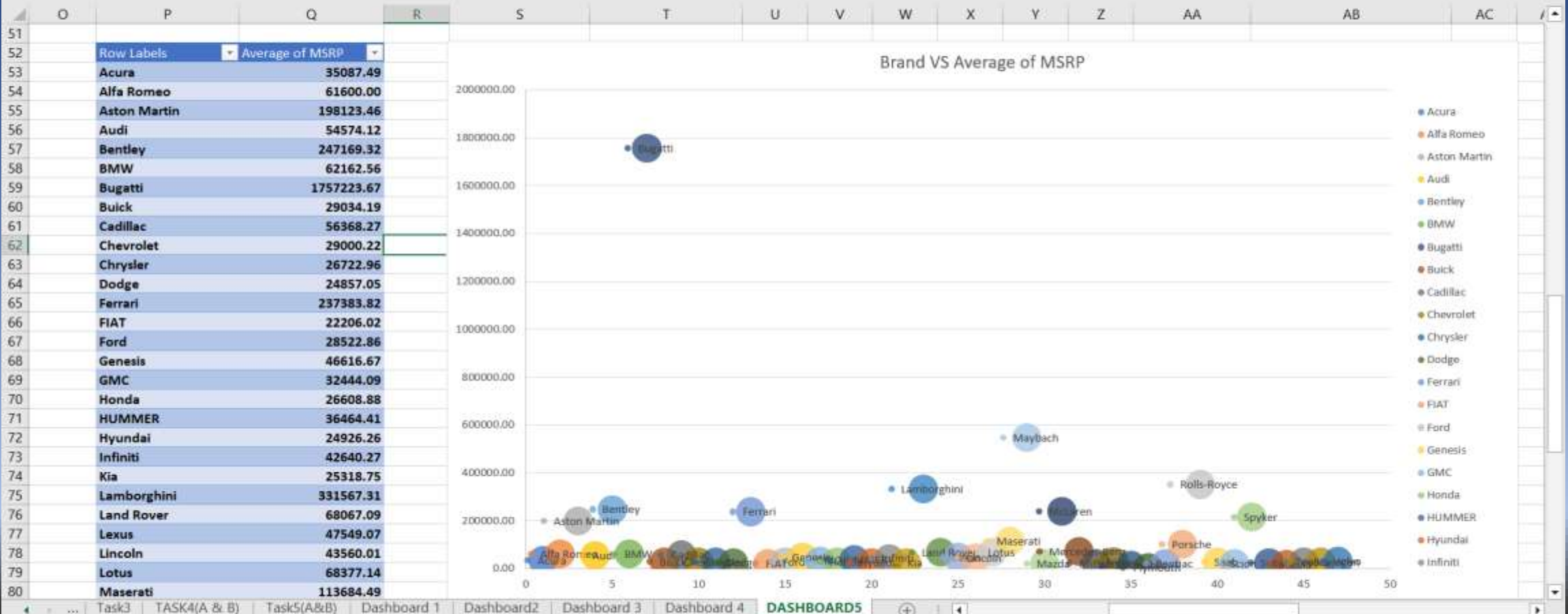


G19

	A	B
52	Row Labels	Average of city mpg
53	Acura	20.00
54	Alfa Romeo	24.00
55	Aston Martin	12.56
56	Audi	19.64
57	Bentley	11.55
58	BMW	20.70
59	Bugatti	8.00
60	Buick	18.78
61	Cadillac	17.36
62	Chevrolet	18.92
63	Chrysler	17.76
64	Dodge	16.45
65	Ferrari	10.56
66	FIAT	26.03
67	Ford	17.56
68	Genesis	16.33
69	GMC	15.79
70	Honda	24.72
71	HUMMER	13.53
72	Hyundai	21.88
73	Infiniti	17.84
74	Kia	21.73
75	Lamborghini	11.52
76	Land Rover	16.14
77	Lexus	20.31
78	Lincoln	17.36
79	Lotus	18.68
80	Maserati	13.20
81	Maybach	10.00
82	Mazda	21.59
83	McLaren	15.60
84	Mercedes-Benz	18.01



R62



Ready Accessibility: Investigate

27°C
Haze

Search

ENG
IN22:31
12-10-2023

Conclusion:

- ❖ *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*

DRIVE LINK EXCEL:

https://docs.google.com/spreadsheets/d/1haox82RnTA8LGEdV55h95eu9zQWrXKLW/edit?usp=drive_link&ouid=116264600853168329913&rtpof=true&sd=true

THANK YOU!