IMPACT OF CAR FEATURES ANALYSIS

CONTENT

- PROJECT OBJECTIVE

- DOWNLOADED THE DATA

- IMPORTED THE EXCEL DATA INTO PYTHON

- ANALYSED THE IMPACT OF CAR FEATURES

- INSIGHTS

- CONCLUSION

- RECOMMENDATION

PROJECT OBJECTIVE

THE GOAL OF THIS PROJECT IS TO HELP A CAR MANUFACTURER MAKE SMARTER DECISIONS ABOUT PRICING AND PRODUCT DEVELOPMENT BY UNDERSTANDING WHAT TODAY'S CAR BUYERS REALLY WANT. USING DATA FROM THE AUTOMOTIVE MARKET, WE'LL EXPLORE HOW DIFFERENT FEATURES, FUEL TYPES, AND VEHICLE CATEGORIES IMPACT BOTH CONSUMER DEMAND AND PROFITABILITY. BY APPLYING TECHNIQUES LIKE REGRESSION ANALYSIS AND MARKET SEGMENTATION, WE AIM TO UNCOVER INSIGHTS THAT GUIDE THE CREATION OF CARS PEOPLE LOVE, WHILE ALSO BOOSTING THE COMPANY'S BOTTOM LINE. ULTIMATELY, THIS PROJECT SUPPORTS BUILDING MORE COMPETITIVE, EFFICIENT, AND FUTURE-READY VEHICLES.

DOWNLOADED THE DATA

THIS DATA IS DOWNLOADED FROM TRAINITY, AN ONLINE LEARNING PLATFORM FOR ASPIRING DATA ANALYSTS, DESIGNED TO PROVIDE HANDS-ON EXPERIENCE THROUGH LIVE PROJECTS AND VIRTUAL INTERNSHIPS.

THE LINK BELOW IS THE RAW EXCEL DATA, WHICH IS UPLOADED TO GITHUB:

HTTPS://GITHUB.COM/MANISHTOPNO/PYTHON_IMPACT-OF-CAR-FEATURES-ANALYSIS/BLOB/MAIN/IMPACT%20OF%20CAR%20FEATURES.XLSX

IMPORTED THE EXCEL DATA INTO PYTHON

AFTER DOWNLOADED THE RAW DATA FROM TRAINITY IN AN EXCEL FILE. I OPENED THE JUPYTER NOTEBOOK, IMPORTED THE NECESSARY LIBRARIES WHICH ARE NEEDED TO IMPORT AN EXCEL FILE IN PYTHON TO RUN SMOOTHLY AND DATA ANALYSIS PROCESS EASIER.

```python
[69]:   import pandas as pd
        import numpy as np
        import matplotlib.pyplot as plt
        %matplotlib inline
        import seaborn as sns

[70]:   # import Excel file
        file_path ='D:/Projects/Impact of car features/Dataset.xlsx'

[71]:   df = pd.read_excel(file_path, engine='openpyxl')
```

This is the Excel file Path

It helps to read easily in Python

ANALYSED THE IMPACT OF CAR FEATURES

- HOW DOES THE POPULARITY OF CAR MANUFACTURERS VARY ACROSS THE MARKET?

Popularity

```python
High_popularity = df.groupby(['Manufacturer'], as_index=False)['Popularity'].sum().sort_values(by='Popularity', ascending=False).head(10)
```

```python
# Top 10 Manufacturers by Popularity
plt.figure(figsize=(20,10))
sns.barplot(x = 'Manufacturer',y= 'Popularity', data = High_popularity,hue='Manufacturer')
plt.title('Manufacturer vs Popularity')
plt.show()
```

This is the Python code for bar chart
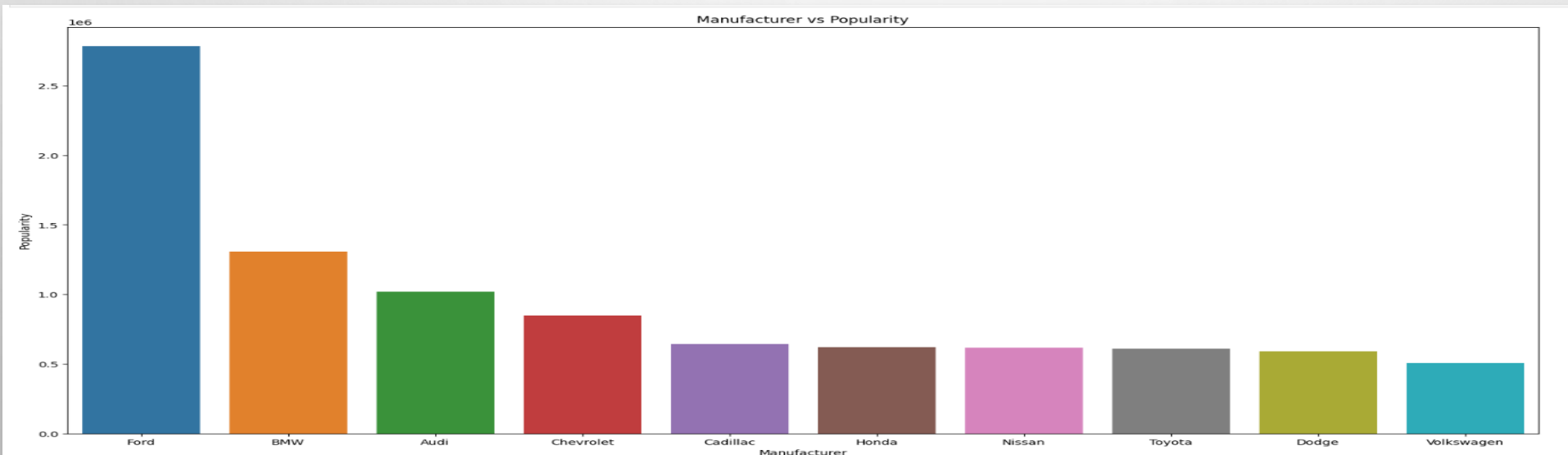
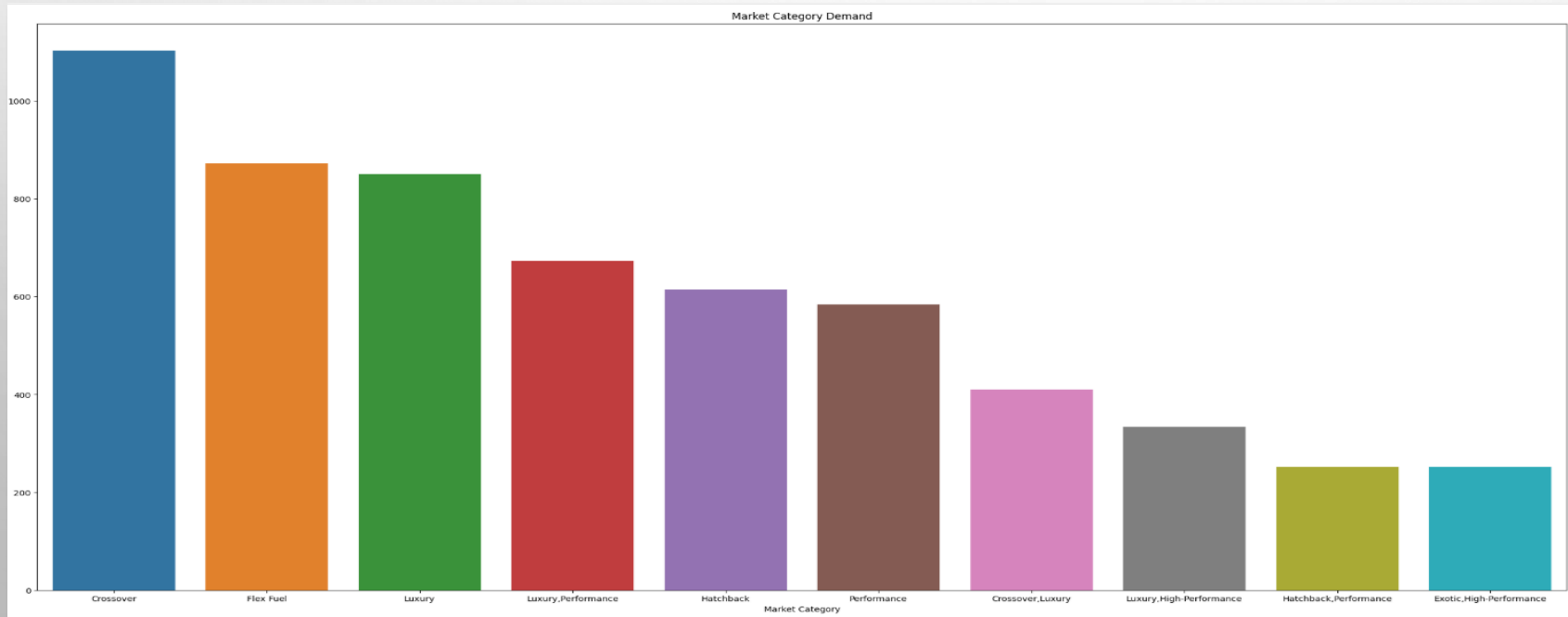**This is the output of the above code**

- HOW DOES THE DEMAND FOR THE MARKET CATEGORY VARY ACROSS DIFFERENT MANUFACTURERS?

Market Category

```
Market_demand= df.groupby(['Market Category'], as_index=False)['Manufacturer'].count().sort_values(by='Manufacturer', ascending=False).head(10)
```

```
#Top 10 highest Market Category Demand
plt.figure(figsize=(30,14))
sns.barplot(x = 'Market Category',y= 'Manufacturer', data = Market_demand,hue='Market Category')
plt.title('Market Category Demand')
plt.show()
```

This is the Python code for bar chart

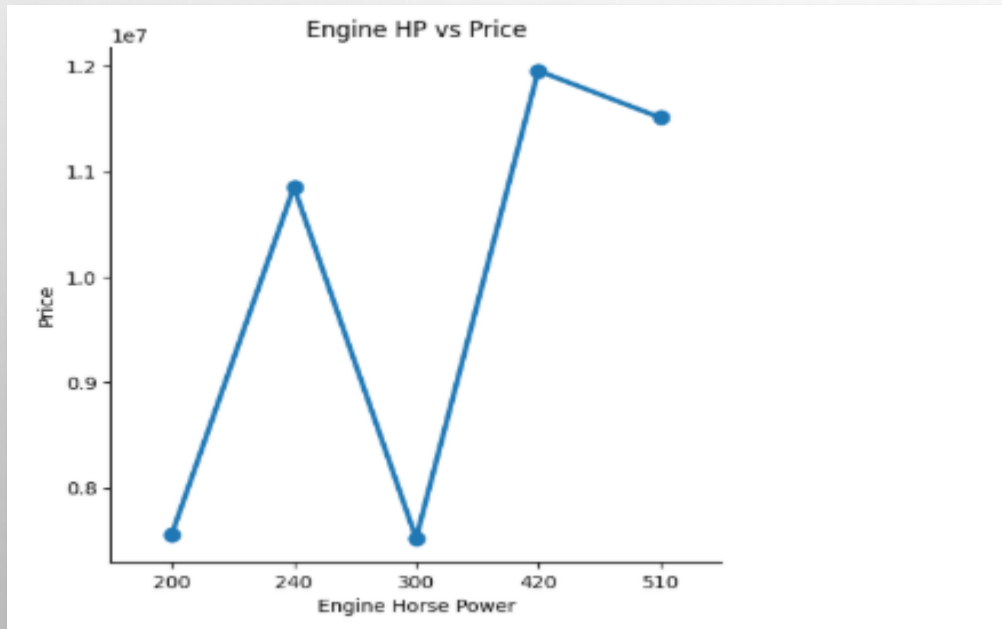**It is the output of the above code**

- WHAT IS THE RELATIONSHIP BETWEEN A CAR'S ENGINE POWER AND ITS PRICE?

Engine HP

```python
Power_price= df.groupby(['Engine HP'], as_index=False)['MSRP'].sum().sort_values(by='MSRP', ascending=False).head()
```

```python
plt.figure(figsize=(20,15))
sns.catplot(x='Engine HP', y='MSRP', data=Power_price, kind='point')
plt.xlabel('Engine Horse Power')
plt.ylabel('Price')
plt.title('Engine HP vs Price')
plt.show()
```

This is the Python code for line chart

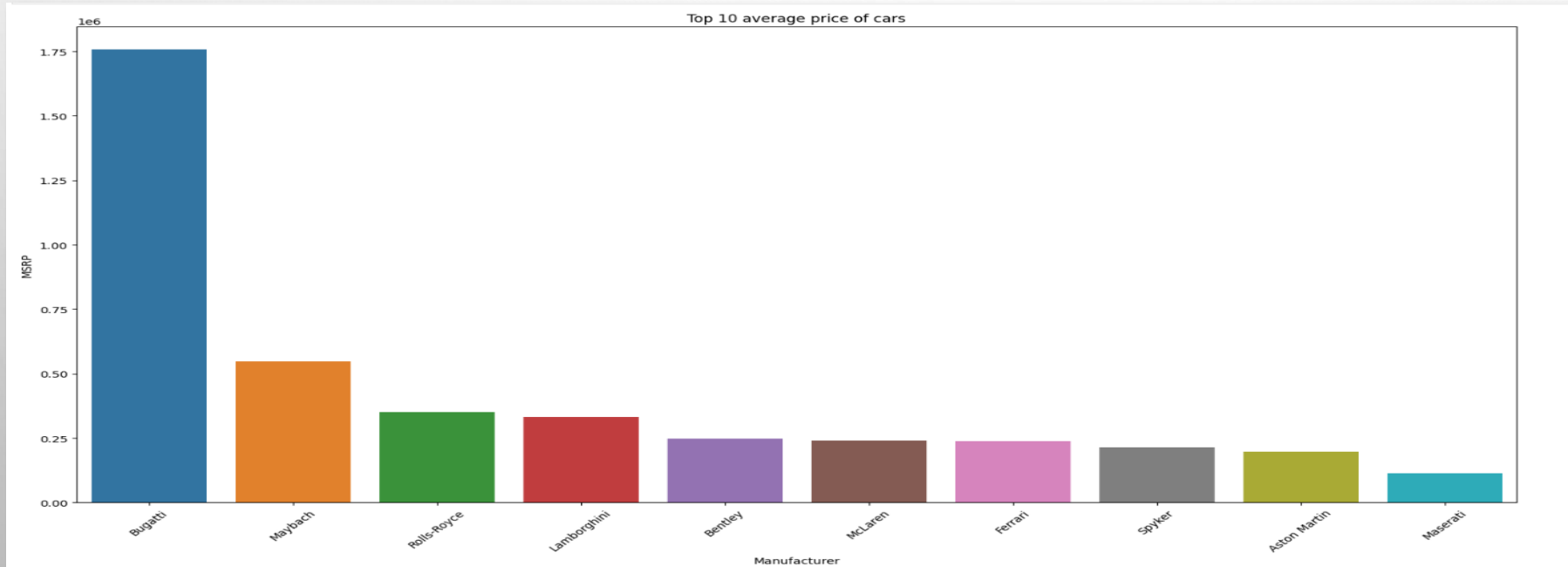**It is the output of the above code**

- WHAT IS THE AVERAGE PRICE OF THE TOP 10 CARS?

Manufacturer

```python
Average_price= df.groupby(['Manufacturer'], as_index=False)['MSRP'].mean().sort_values(by='MSRP', ascending=False).head(10)
```

```python
plt.figure(figsize=(20,10))
sns.barplot(x = 'Manufacturer',y= 'MSRP', data = Average_price,hue='Manufacturer')
plt.title('Top 10 average price of cars')
plt.xticks(rotation=45)
plt.show()
```

This is the Python code for a Bar chart

Top 10 average price of cars

It is the output of the above code

- WHAT IS THE RELATIONSHIP BETWEEN A CAR'S ENGINE CYLINDERS AND ITS PRICE?

```python
plt.figure(figsize=(20,15))
sns.relplot(data=df, x='Engine Cylinders', y='MSRP',kind='scatter',hue='Engine Cylinders')
plt.xlabel('Engine Cylinders')
plt.ylabel('Price')
plt.title('Relationship of Engine cylinders by Price')
plt.tight_layout()
plt.show()
```

This is the Python code for a scatter chart
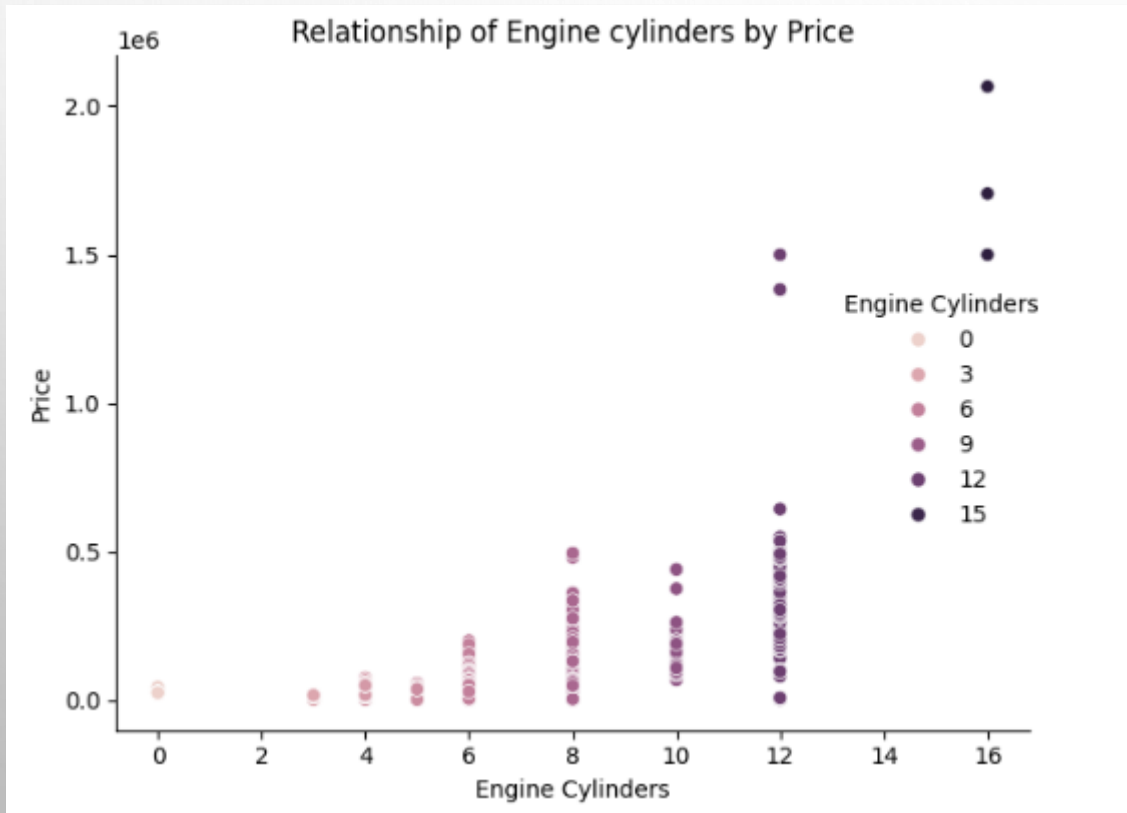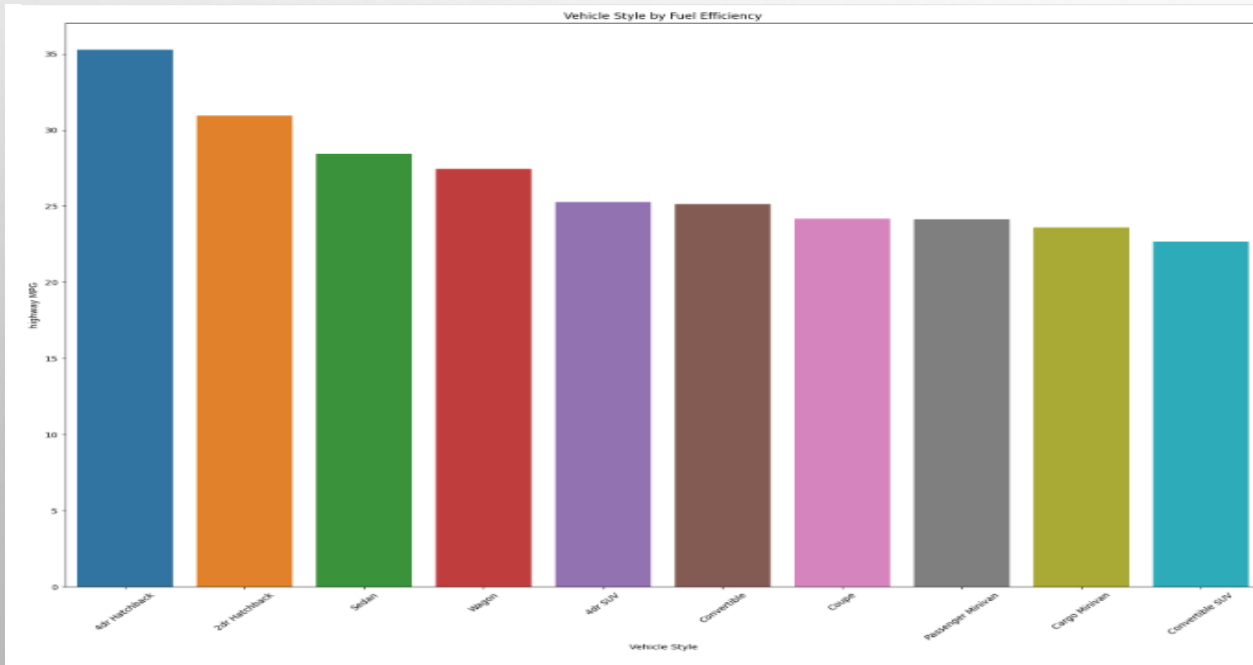

It is the output of the above code

- # HOW DOES THE FUEL EFFICIENCY OF CARS VARY ACROSS DIFFERENT BODY STYLES?

Vehicle Style

```
average_style= df.groupby(['Vehicle Style'], as_index=False)['highway MPG'].mean().sort_values(by='highway MPG', ascending=False).head(10)
```

```
plt.figure(figsize=(20,15))
sns.barplot(x = 'Vehicle Style',y= 'highway MPG', data = average_style,hue='Vehicle Style')
plt.title('Vehicle Style by Fuel Efficiency')
plt.xticks(rotation=45)
plt.show()
```

This is the Python code for a bar chart

Vehicle Style by Fuel Efficiency

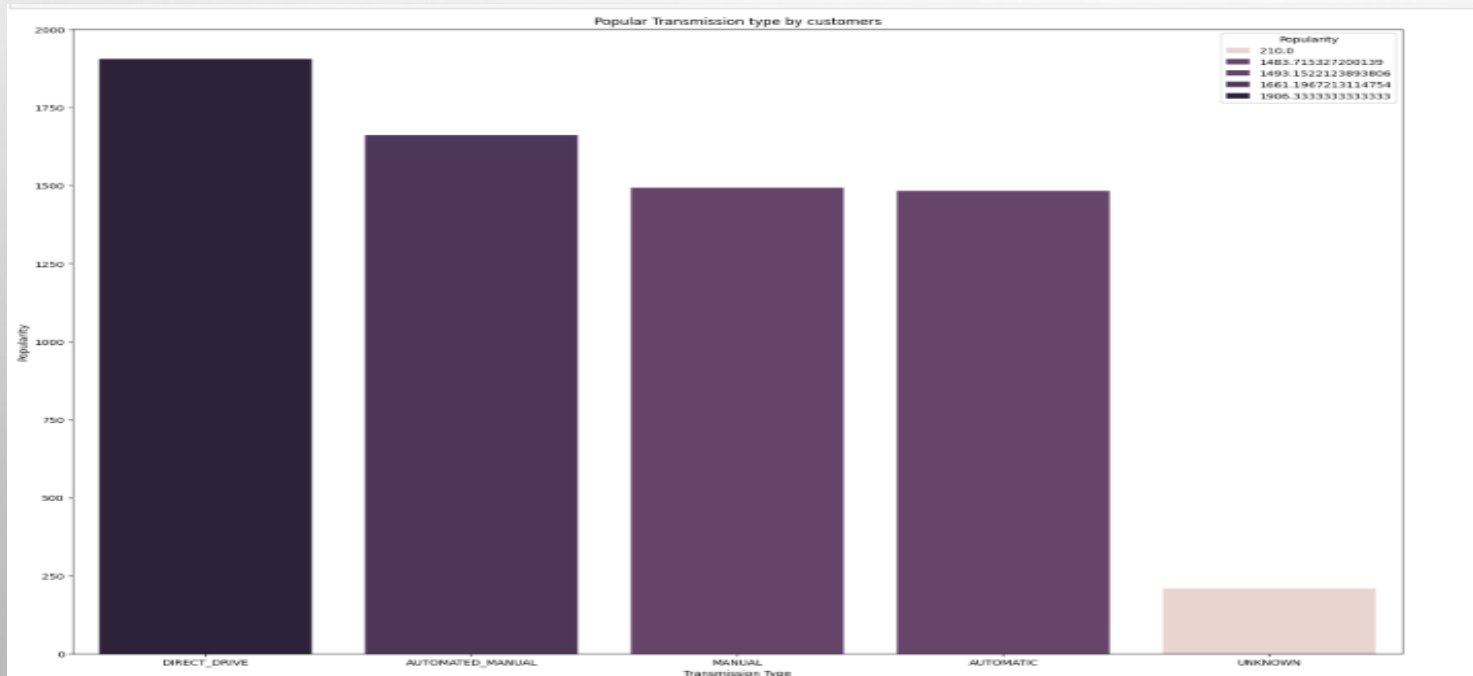**It is the output of the above code**

- HOW DO THE DIFFERENT FEATURES SUCH AS TRANSMISSION TYPE AFFECT THE MSRP?

Transmission Type

```python
transmission= df.groupby(['Transmission Type'], as_index=False)['Popularity'].mean().sort_values(by='Popularity', ascending=False).head()
```

```python
plt.figure(figsize=(20,15))
sns.barplot(x = 'Transmission Type',y= 'Popularity', data =transmission,hue='Popularity')
plt.title('Popular Transmission type by customers')
plt.show()
```

This is the Python code for a bar chart

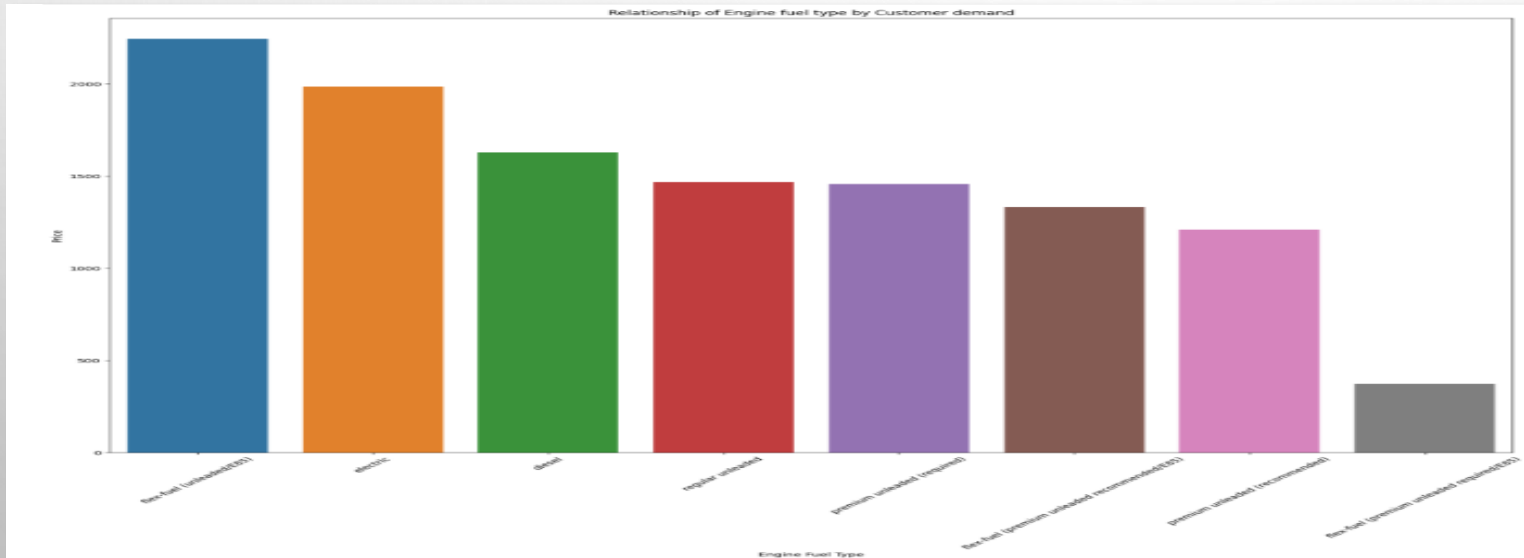**It is the output of the above code**

- WHAT IS THE RELATIONSHIP BETWEEN THE CAR ENGINE FUEL TYPE BY POPULARITY?

Engine Fuel Type

```
Fuel= df.groupby(['Engine Fuel Type'], as_index=False)['Popularity'].mean().sort_values(by='Popularity', ascending=False)
```

```python
plt.figure(figsize=(20,15))
sns.barplot(x = 'Engine Fuel Type',y= 'Popularity', data =Fuel,hue='Engine Fuel Type')
plt.xlabel('Engine Fuel Type')
plt.ylabel('Price')
plt.title('Relationship of Engine fuel type by Customer demand')
plt.xticks(rotation=45)
plt.show()
```

This is the Python code for a bar chart

Relationship of Engine fuel type by Customer demand

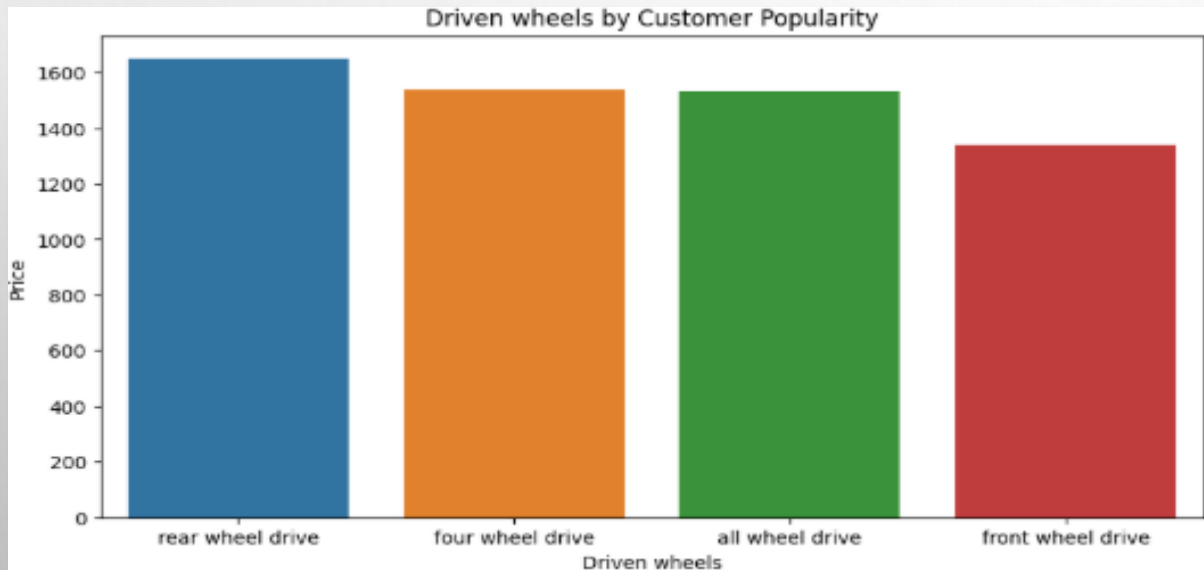**It is the output of the above code**

- WHAT ARE THE BEST DRIVEN WHEELS BY POPULARITY?

Driven Wheels

```
wheels= df.groupby(['Driven_Wheels'], as_index=False)['Popularity'].mean().sort_values(by='Popularity', ascending=False)
```

```
plt.figure(figsize=(10,5))
sns.barplot(x = 'Driven_Wheels',y= 'Popularity', data =wheels,hue='Driven_Wheels')
plt.xlabel('Driven wheels')
plt.ylabel('Price')
plt.title('Relationship of Driven wheels by Customer Popularity')
plt.show()
```

This is the Python code for a bar chart

Driven wheels by Customer Popularity

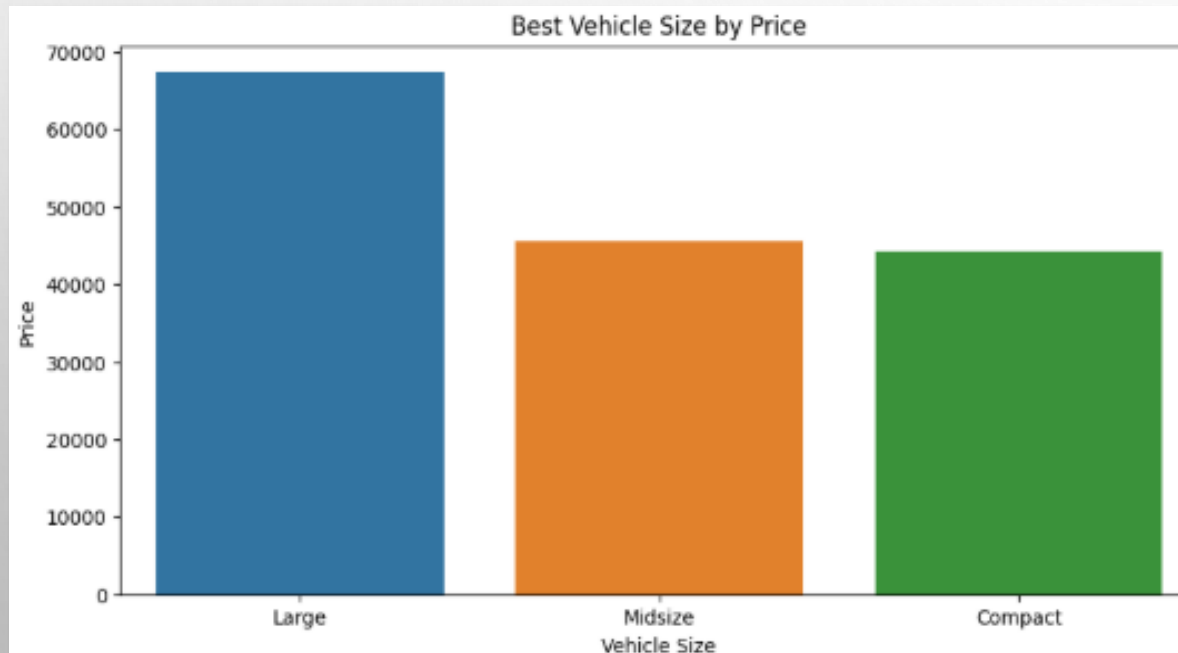**It is the output of the above code**

- WHAT ARE THE BEST VEHICLE SIZES BY PRICE?

Vehicle Size

```
size= df.groupby(['Vehicle Size'], as_index=False)['MSRP'].mean().sort_values(by='MSRP', ascending=False)
```

```python
plt.figure(figsize=(10,5))
sns.barplot(x = 'Vehicle Size',y= 'MSRP', data =size,hue="Vehicle Size")
plt.xlabel('Vehicle Size')
plt.ylabel('Price')
plt.title('Best Vehicle Size by Price')
plt.show()
```

This is the Python code for a bar chart

Best Vehicle Size by Price

It is the output of the above code

INSIGHTS

- FORD STANDS OUT AS THE MOST POPULAR AND HIGH-DEMAND CAR MANUFACTURER IN THE MARKET.

- THE CROSSOVER TYPE OF MARKET CATEGORY IS POPULAR AMONG CUSTOMERS.

- CARS WITH 420 HP ENGINES TOP THE PRICE RANGE, WHILE 200 AND 300 HP MODELS ARE THE MOST AFFORDABLE.

- THE HIGHEST AVERAGE PRICE OF THE MANUFACTURER IS BUGATTI.

- CARS WITH MORE CYLINDERS GENERALLY COST MORE DUE TO HIGHER PERFORMANCE AND ENGINE COMPLEXITY.

- THE 4-DOOR HATCHBACK IS THE MOST POPULAR STYLE WHEN IT COMES TO FUEL EFFICIENCY.

- DIRECT DRIVE TRANSMISSION IS A TOP CHOICE AMONG CUSTOMERS FOR ITS POPULARITY.

- THE FLEX-FUEL(UNLEADED/E85) IS HIGHLY DEMAND IN THE MARKET.

- REAR WHEEL DRIVE IS THE MOST POPULAR AMONG CUSTOMERS.

- COMPACT VEHICLES ARE MORE BUDGET-FRIENDLY COMPARED TO MID-SIZE AND LARGE MODELS.

CONCLUSION

UNDERSTANDING CONSUMER PREFERENCES THROUGH DATA IS KEY TO THRIVING IN TODAY'S COMPETITIVE AUTOMOTIVE INDUSTRY. BY ANALYZING HOW CAR FEATURES, PRICING, AND MARKET SEGMENTS INTERACT, MANUFACTURERS CAN MAKE SMARTER, MORE TARGETED DECISIONS. THIS APPROACH SUPPORTS THE DEVELOPMENT OF PRICING STRATEGIES AND PRODUCTS THAT REFLECT WHAT CONSUMERS TRULY VALUE, ENSURING BOTH PROFITABILITY AND RELEVANCE. IN A RAPIDLY EVOLVING MARKET, DATA-DRIVEN INSIGHTS OFFER A CLEAR PATH TO SUSTAINABLE GROWTH AND INNOVATION.

RECOMMENDATION

MANUFACTURERS SHOULD FOCUS ON PRODUCING COMPACT, FUEL-EFFICIENT VEHICLES LIKE 4-DOOR HATCHBACKS AND CROSSOVERS, WHICH ARE BOTH POPULAR AND AFFORDABLE. EMPHASIZING HIGH-DEMAND FEATURES SUCH AS DIRECT DRIVE AND REAR-WHEEL DRIVE TRANSMISSIONS CAN BOOST APPEAL. WHILE PREMIUM MODELS LIKE THOSE WITH 420 HP FROM BRANDS LIKE BUGATTI SHOW PRICING POTENTIAL, A BALANCED PORTFOLIO INCLUDING FLEX-FUEL OPTIONS AND VARYING CYLINDER COUNTS WILL HELP MEET DIVERSE CUSTOMER NEEDS. FORD'S POPULARITY HIGHLIGHTS THE VALUE OF STRONG BRAND PRESENCE, WHICH OTHERS CAN AIM TO EMULATE THROUGH STRATEGIC MARKETING AND INNOVATION.