

# Ford Price Analysis

*10/06/2023  
Data Analysis*

*Prakhar Tripathi*

## Problem Statement

*In the context of the automotive industry, the problem at hand is to understand and quantify the key factors that influence the resale value of Ford car models.*

*This problem is critical for various stakeholders, including Ford Motor Company, car buyers, and the broader automotive market. The overarching goal is to provide actionable insights that can inform pricing strategies, consumer decisions, and future product development efforts.*

*To address this problem effectively, we need to analyze a dataset containing attributes such as production year, initial price, transmission type, mileage, fuel type, annual tax, miles per gallon (MPG), and engine size, and determine how these variables affect the resale value of Ford vehicles.*

*Now we must analyze the price according to different factor they are as follows...*

*First, we must do the data processing to clean and check the data and maintain the values.*

*Next, we do the EDA process to perform different tasks like drop, cleaning, value evaluation, renaming and modification.*

*Last we do visualization of data to analyze them.*



# Task to Perform in Data Analysis

## Data Preprocessing

- The dataset is first cleaned to handle missing values and outliers.
- Categorical variables like "Transmission" and "Fuel Type" are encoded for analysis.

## Exploratory Data Analysis (EDA)

- Descriptive statistics are used to summarize key attributes.
- Visualizations such as histograms, scatter plots, and correlation matrices are generated to understand the relationships between variables.

## Factors Affecting Resale Value

- Resale value is the dependent variable, and a regression analysis is performed to identify significant predictors.
- Factors like "Year," "Mileage," "Fuel Type," "MPG," and "Engine Size" are expected to have a significant impact on resale value.

## Engine Size and Resale Value

- The effect of engine size on resale value is investigated.
- Larger engine sizes may lead to higher resale values, especially for performance-oriented models.

## Year and Resale Value

- A time-series analysis is conducted to assess how the resale value changes with the production year.
- It is expected that newer models tend to have higher resale values.

## Mileage and Resale Value

- The relationship between mileage and resale value is examined.
- A negative correlation is expected, as lower mileage typically results in a higher resale value.

## Fuel Type and Resale Value

- The impact of different fuel types on resale value is analyzed.
- Electric and hybrid cars may have a higher resale value due to their eco-friendliness.

## MPG and Resale Value

- The relationship between fuel efficiency (MPG) and resale value is explored.
- Higher MPG may positively influence resale value.

## Conclusion

- The findings are summarized, and the key factors influencing the resale value of Ford car models are identified.
- Recommendations for Ford and potential buyers are provided based on the analysis.

# EDA (EXPLORATORY DATA ANALYSIS)

First, we find the information about dataset.

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"

Exploratory Data Analysis

Information About Dataset
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 17966 entries, 0 to 17965
Data columns (total 9 columns):
#   Column          Non-Null Count  Dtype
---  -
0   model           17966 non-null  object
1   year            17966 non-null  int64
2   price           17966 non-null  int64
3   transmission    17966 non-null  object
4   mileage         17966 non-null  int64
5   fuelType       17966 non-null  object
6   tax             17966 non-null  int64
7   mpg            17966 non-null  float64
8   engineSize     17966 non-null  float64
dtypes: float64(2), int64(4), object(3)
memory usage: 1.2+ MB
None

Process finished with exit code 0
```

By Applying info() method we get the information about all the column, like their datatype, memory usages, null values.

This is used to identify the data information.

# Top-10 rows of the dataset

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"
```

Exploratory Data Analysis

Top 10 rows of dataset

	model	year	price	transmission	mileage	fuelType	tax	mpg	engineSize
0	Fiesta	2017	12000	Automatic	15944	Petrol	150	57.7	1.0
1	Focus	2018	14000	Manual	9883	Petrol	150	57.7	1.0
2	Focus	2017	13000	Manual	12456	Petrol	150	57.7	1.0
3	Fiesta	2019	17500	Manual	10460	Petrol	145	40.3	1.5
4	Fiesta	2019	16500	Automatic	1482	Petrol	145	48.7	1.0
5	Fiesta	2015	10500	Manual	35432	Petrol	145	47.9	1.6
6	Puma	2019	22500	Manual	2029	Petrol	145	50.4	1.0
7	Fiesta	2017	9000	Manual	13054	Petrol	145	54.3	1.2
8	Kuga	2019	25500	Automatic	6894	Diesel	145	42.2	2.0
9	Focus	2018	10000	Manual	48141	Petrol	145	61.4	1.0

Process finished with exit code 0

# Last-10 rows of the dataset

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"
```

Exploratory Data Analysis

Last 10 rows of dataset

	model	year	price	transmission	mileage	fuelType	tax	mpg	engineSize
17956	Grand C-MAX	2018	14750	Manual	4850	Petrol	145	47.1	1.0
17957	EcoSport	2015	7650	Manual	46123	Petrol	125	53.3	1.0
17958	C-MAX	2019	13250	Manual	13359	Petrol	145	48.7	1.0
17959	C-MAX	2016	9299	Manual	35637	Diesel	20	68.9	1.5
17960	Fiesta	2016	7999	Manual	31348	Petrol	125	54.3	1.2
17961	B-MAX	2017	8999	Manual	16700	Petrol	150	47.1	1.4
17962	B-MAX	2014	7499	Manual	40700	Petrol	30	57.7	1.0
17963	Focus	2015	9999	Manual	7010	Diesel	20	67.3	1.6
17964	KA	2018	8299	Manual	5007	Petrol	145	57.7	1.2
17965	Focus	2015	8299	Manual	5007	Petrol	22	57.7	1.0

Process finished with exit code 0

# Columns of the dataset

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"
```

```
Exploratory Data Analysis
```

```
columns of dataset
```

```
Index(['model', 'year', 'price', 'transmission', 'mileage', 'fuelType', 'tax',  
      'mpg', 'engineSize'],  
      dtype='object')
```

# No. of rows and Columns in dataset

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"
```

```
Exploratory Data Analysis
```

```
Number of rows and columns of dataset
```

```
Number of rows: 17966
```

```
Number of columns: 9
```

```
Process finished with exit code 0
```

# Value Count of model

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"
```

```
Exploratory Data Analysis
```

```
Count values of model
```

```
model
Fiesta          6557
Focus           4588
Kuga            2225
EcoSport        1143
C-MAX           543
Ka+             531
Mondeo          526
B-MAX           355
S-MAX           296
Grand C-MAX     247
Galaxy          228
Edge            208
KA              199
Puma            80
Tourneo Custom  69
Grand Tourneo Connect 59
Mustang         57
Tourneo Connect 33
Fusion          16
Streetka        2
Ranger          1
Escort          1
Transit Tourneo 1
Focus           1
```

```
Name: count, dtype: int64
```

```
Process finished with exit code 0
```

In the value count we have a unwanted value that is Focus.

# Remove unwanted value from Model column.

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"
```

```
Exploratory Data Analysis
```

```
Remove values of model
```

```
model
Fiesta          6557
Focus           4588
Kuga            2225
EcoSport        1143
C-MAX           543
Ka+             531
Mondeo          526
B-MAX           355
S-MAX           296
Grand C-MAX     247
Galaxy          228
Edge            208
KA              199
Puma            80
Tourneo Custom  69
Grand Tourneo Connect 59
Mustang         57
Tourneo Connect 33
Fusion          16
Streetka        2
Ranger          1
Escort          1
Transit Tourneo 1
```

```
Name: count, dtype: int64
```

```
Process finished with exit code 0
```

**By using the drop function, we remove the value from column.**



# Value count of year

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"
```

```
Exploratory Data Analysis
```

```
Count values of year
```

```
year
```

```
2017    4888
```

```
2018    4014
```

```
2019    3194
```

```
2016    2331
```

```
2015    1368
```

```
2014     805
```

```
2013     609
```

```
2020     258
```

```
2012     115
```

```
2011      94
```

```
2009      91
```

```
2010      67
```

```
2008      57
```

```
2007      32
```

```
2005      16
```

```
2006      13
```

```
2004       4
```

```
2002       3
```

```
2003       3
```

```
1998       1
```

```
1996       1
```

```
2000       1
```

```
2060       1
```

```
Name: count, dtype: int64
```

```
Process finished with exit code 0
```

In the value count we have a unwanted value that is 2060. It is a faulty value for that data set.

# Remove unwanted value from Year column.

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"
```

```
Exploratory Data Analysis
```

```
Remove values of year
```

```
year
```

```
2017    4888
```

```
2018    4014
```

```
2019    3194
```

```
2016    2331
```

```
2015    1368
```

```
2014     805
```

```
2013     609
```

```
2020     258
```

```
2012     115
```

```
2011      94
```

```
2009      91
```

```
2010      67
```

```
2008      57
```

```
2007      32
```

```
2005      16
```

```
2006      13
```

```
2004       4
```

```
2002       3
```

```
2003       3
```

```
1998       1
```

```
1996       1
```

```
2000       1
```

```
Name: count, dtype: int64
```

```
Process finished with exit code 0
```

**By using the drop function, we remove the value from column.**

# Value count of Price

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"
```

```
Exploratory Data Analysis
```

```
Count values of price
```

```
price
```

```
10000    164
```

```
11000    153
```

```
10500    148
```

```
12000    126
```

```
9000     118
```

```
...
```

```
9410      1
```

```
19360      1
```

```
9287       1
```

```
10570      1
```

```
18122      1
```

```
Name: count, Length: 3511, dtype: int64
```

```
Process finished with exit code 0
```

# Value count of Transmission

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"
```

```
Exploratory Data Analysis
```

```
Count values of transmission
```

```
transmission
```

```
Manual    15518
```

```
Automatic  1361
```

```
Semi-Auto  1087
```

```
Name: count, dtype: int64
```

```
Process finished with exit code 0
```

# Value count of Mileage

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"
```

```
Exploratory Data Analysis
```

```
Count values of mileage
```

```
mileage
```

```
10      40
```

```
15000   38
```

```
25000   34
```

```
10000   32
```

```
9000    31
```

```
..
```

```
80427   1
```

```
1376    1
```

```
24376   1
```

```
23496   1
```

```
10057   1
```

```
Name: count, Length: 13528, dtype: int64
```

```
Process finished with exit code 0
```

# Value count of Fuel Type

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"
```

```
Exploratory Data Analysis
```

```
Count values of fuelType
```

```
fuelType
```

```
Petrol    12179
```

```
Diesel     5762
```

```
Hybrid      22
```

```
Electric     2
```

```
Other        1
```

```
Name: count, dtype: int64
```

```
Process finished with exit code 0
```

# Value count of MPG

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"
```

```
Exploratory Data Analysis
```

```
Count values of mpg
```

```
mpg
```

```
65.7    2528
```

```
54.3    1774
```

```
60.1    1674
```

```
64.2     995
```

```
57.7     962
```

```
...
```

```
37.1      1
```

```
28.0      1
```

```
54.2      1
```

```
28.5      1
```

```
23.9      1
```

```
Name: count, Length: 90, dtype: int64
```

```
Process finished with exit code 0
```

# Value count of engine size

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"
```

```
Exploratory Data Analysis
```

```
Count values of engineSize
```

```
engineSize
```

```
1.0    7765
```

```
1.5    3418
```

```
2.0    3311
```

```
1.2    1626
```

```
1.6     923
```

```
1.1     559
```

```
1.4     112
```

```
2.3      80
```

```
0.0      51
```

```
5.0      45
```

```
1.8      35
```

```
2.2      13
```

```
2.5      13
```

```
1.3      13
```

```
3.2       1
```

```
1.7       1
```

```
Name: count, dtype: int64
```

```
Process finished with exit code 0
```

# DATA VISUALIZATION AND ANALYSIS

## Price analysis based on Engine size

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"
```

```
Data Visualization and Analysis
```

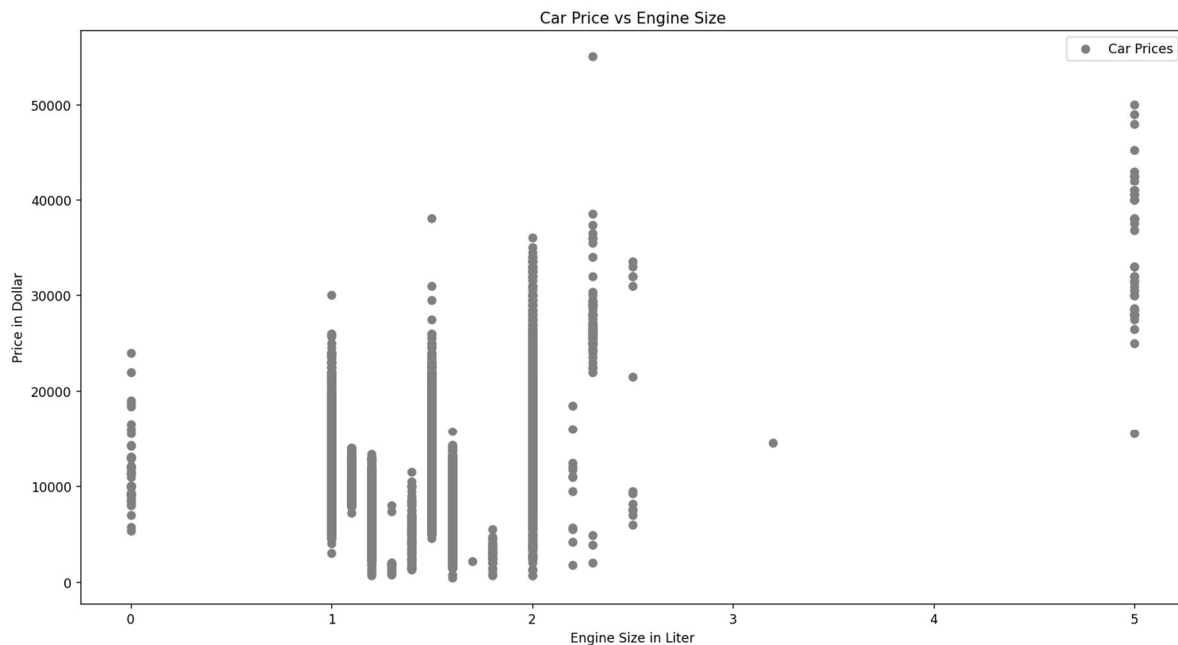
```
Price analysis based on EngineSize
```

```
Price of cars based on their EngineSize
```

```
engineSize  price
0           1.0  12000
1           1.0  14000
2           1.0  13000
3           1.5  17500
4           1.0  16500
...         ...   ...
17961       1.4   8999
17962       1.0   7499
17963       1.6   9999
17964       1.2   8299
17965       1.0   8299
```

```
[17966 rows x 2 columns]
```

## Scatter-graph representation



# Price analysis based on MPG

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"
```

Data Visualization and Analysis

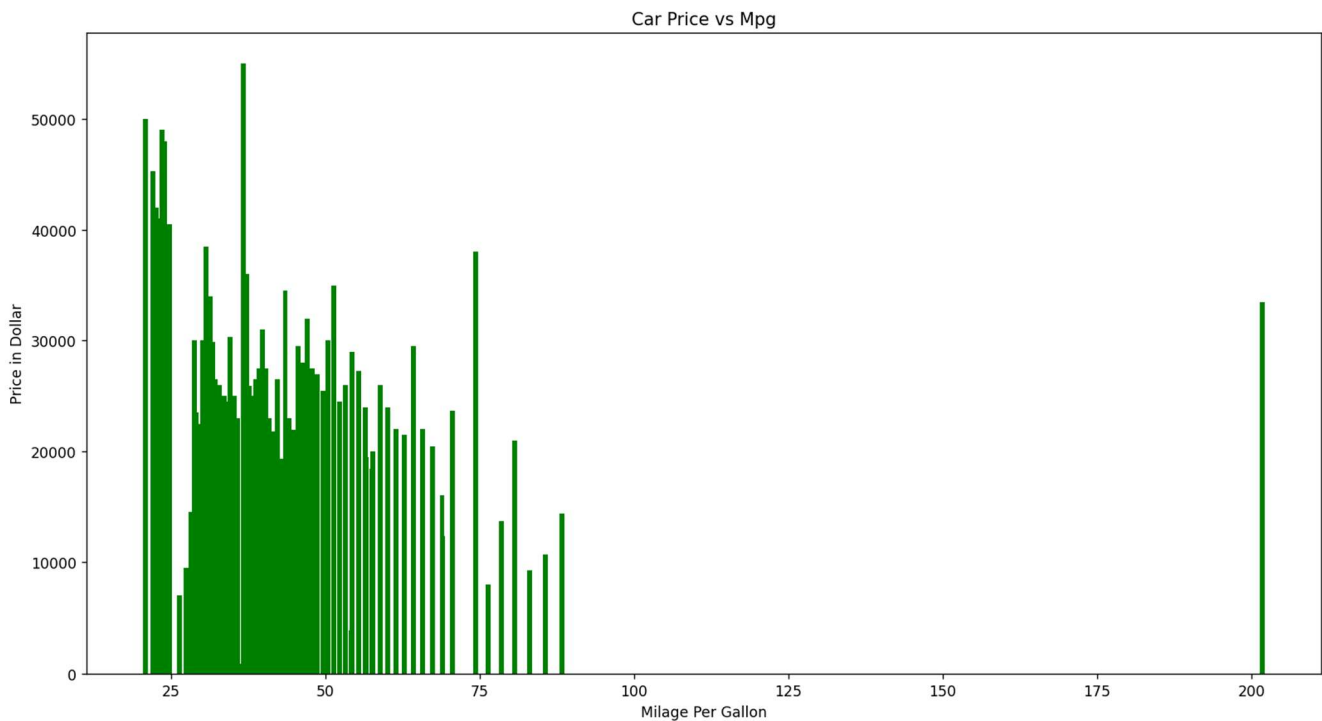
Price analysis based on mpg

Price of cars based on their MPG

	mpg	price
0	57.7	12000
1	57.7	14000
2	57.7	13000
3	40.3	17500
4	48.7	16500
...	...	...
17961	47.1	8999
17962	57.7	7499
17963	67.3	9999
17964	57.7	8299
17965	57.7	8299

[17966 rows x 2 columns]

## Bar-graph representation



# Price analysis based on Year

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"
```

Data Visualization and Analysis

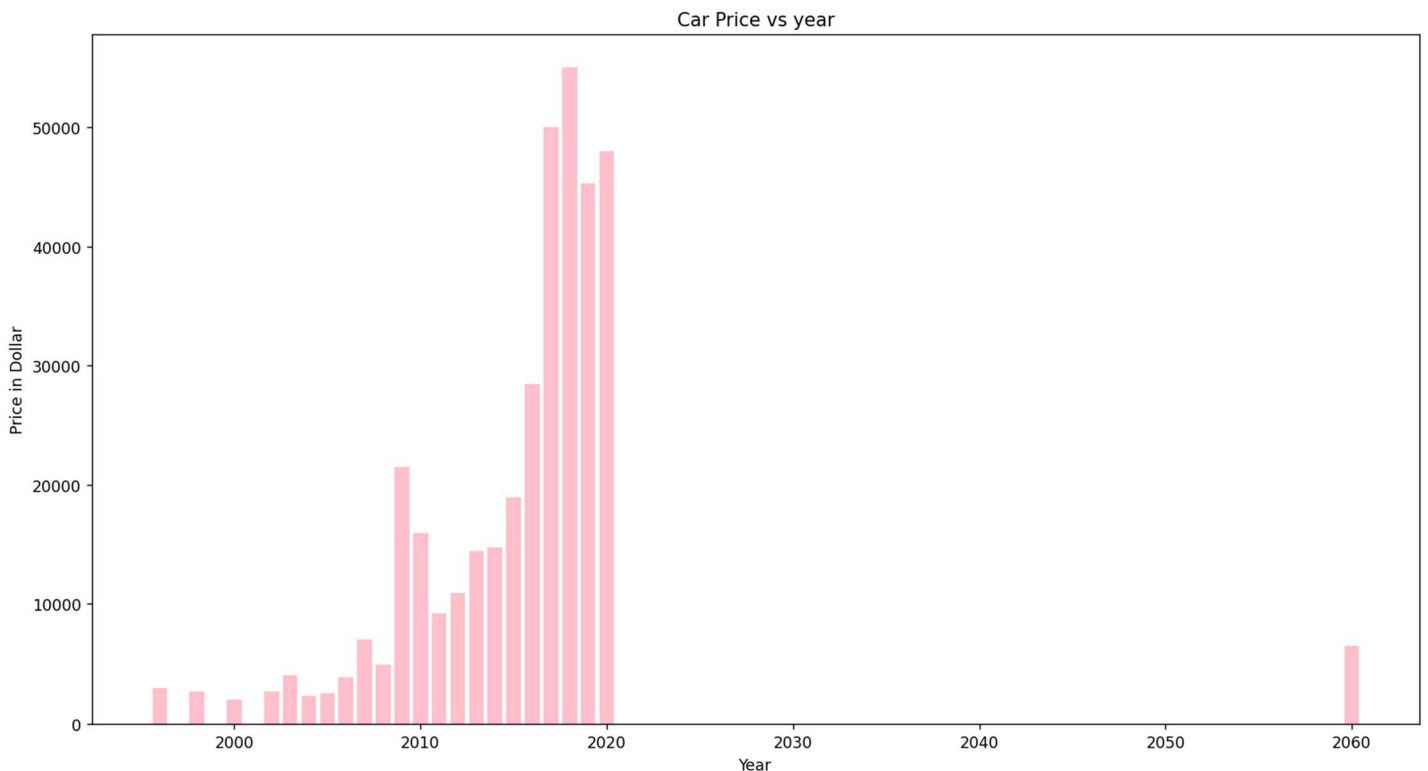
Price analysis based on Year

Price of cars based on year

	year	price
0	2017	12000
1	2018	14000
2	2017	13000
3	2019	17500
4	2019	16500
...	...	...
17961	2017	8999
17962	2014	7499
17963	2015	9999
17964	2018	8299
17965	2015	8299

[17966 rows x 2 columns]

## Bar-graph representation





# Price analysis based on transmission

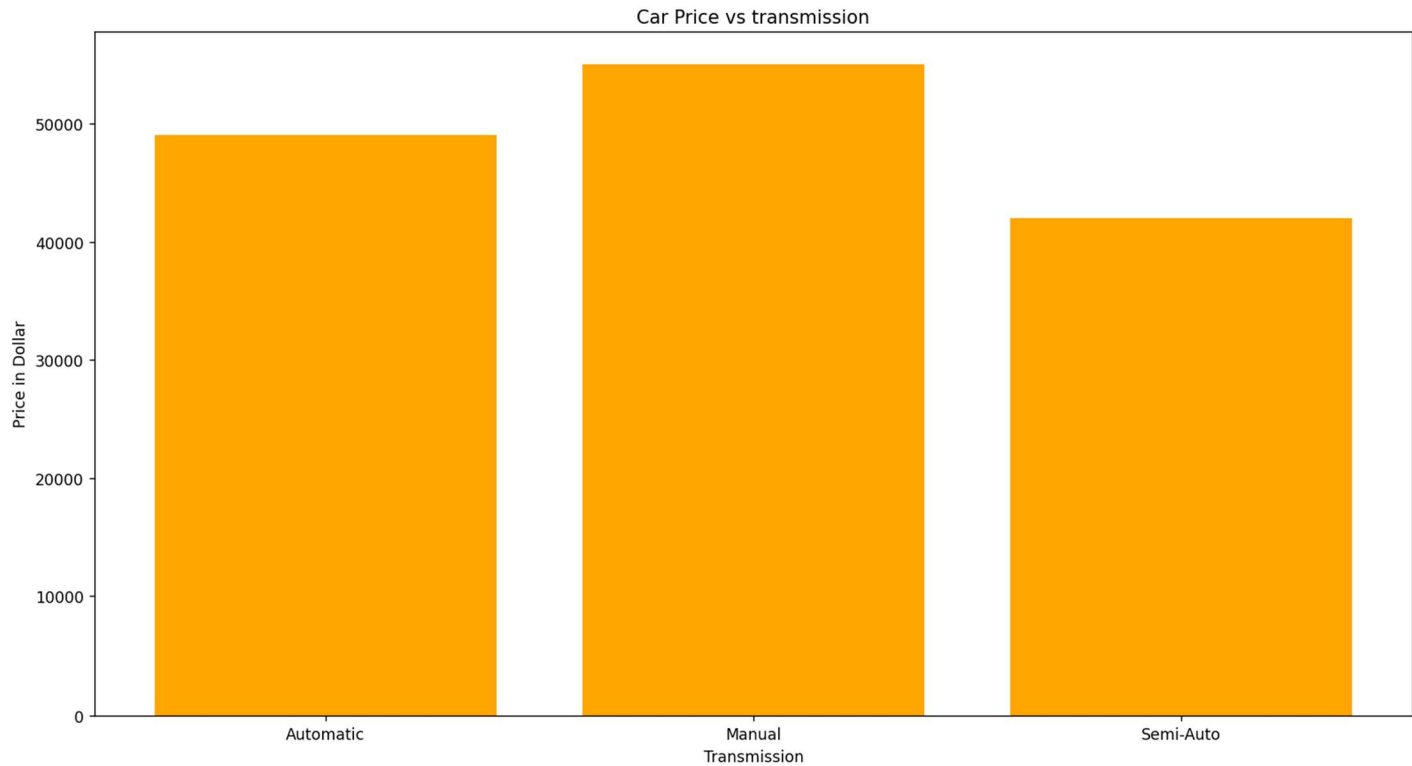
```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"

Data Visualization and Analysis

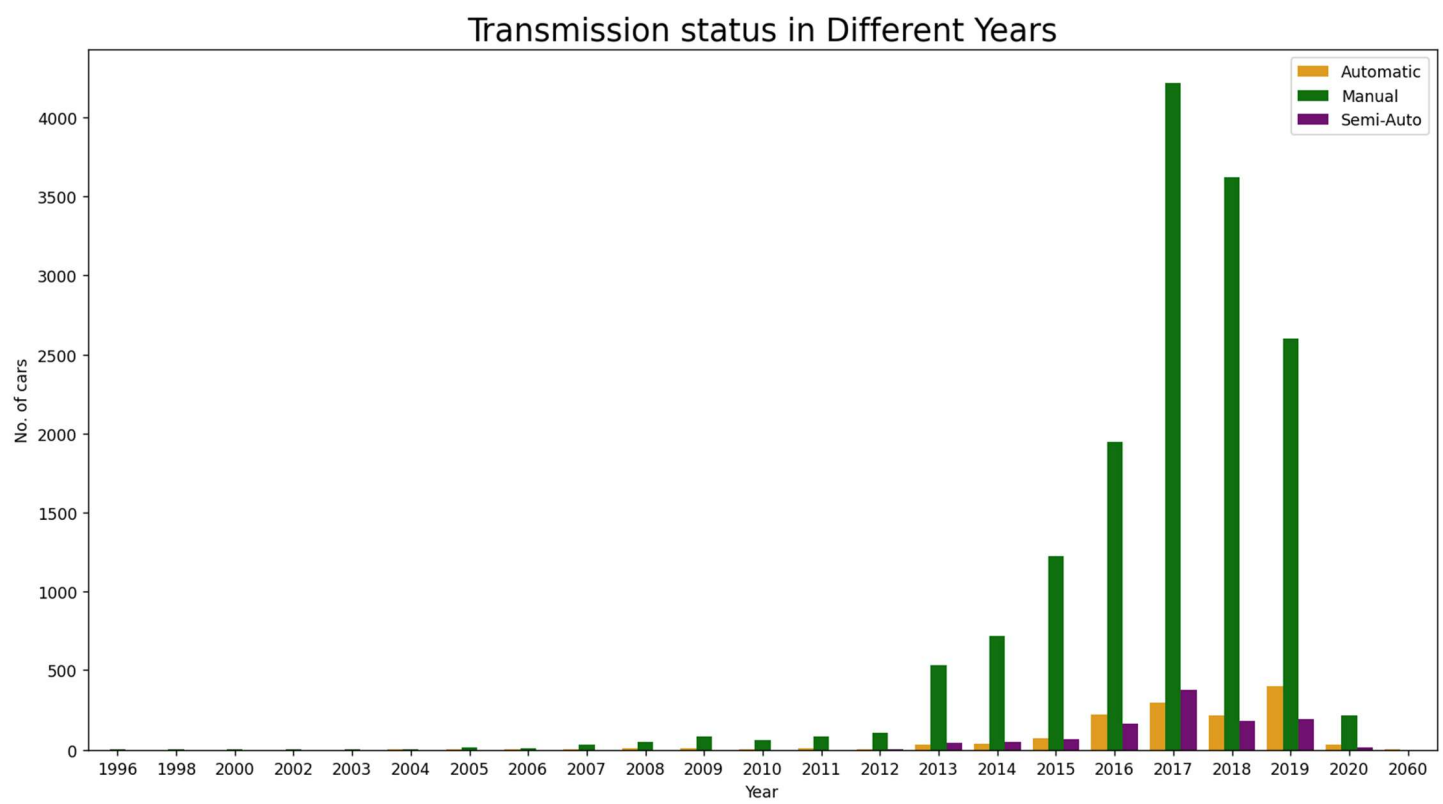
Price analysis based on transmission
Price of cars based on transmission
  transmission  price
0      Automatic 12000
1         Manual 14000
2         Manual 13000
3         Manual 17500
4      Automatic 16500
...          ...   ...
17961      Manual  8999
17962      Manual  7499
17963      Manual  9999
17964      Manual  8299
17965      Manual  8299

[17966 rows x 2 columns]
```

## Bar-graph representation



# No. of cars of different transmission in Different Years



# Top 5 models with automatic Transmission status

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"
```

Data Visualization and Analysis

Top 5 models with automatic Transmission status

model

Focus 415

Kuga 265

Fiesta 186

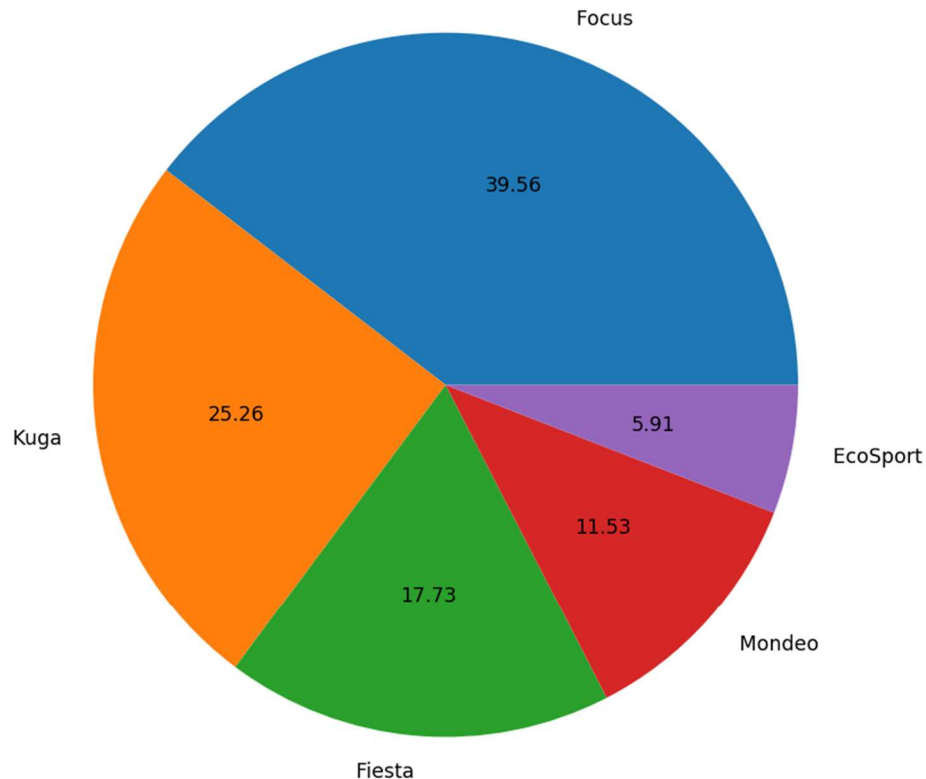
Mondeo 121

EcoSport 62

Name: count, dtype: int64

## Pi-chart Representation

Top 5 models with automatic Transmission status



# Top 5 models with manual Transmission status

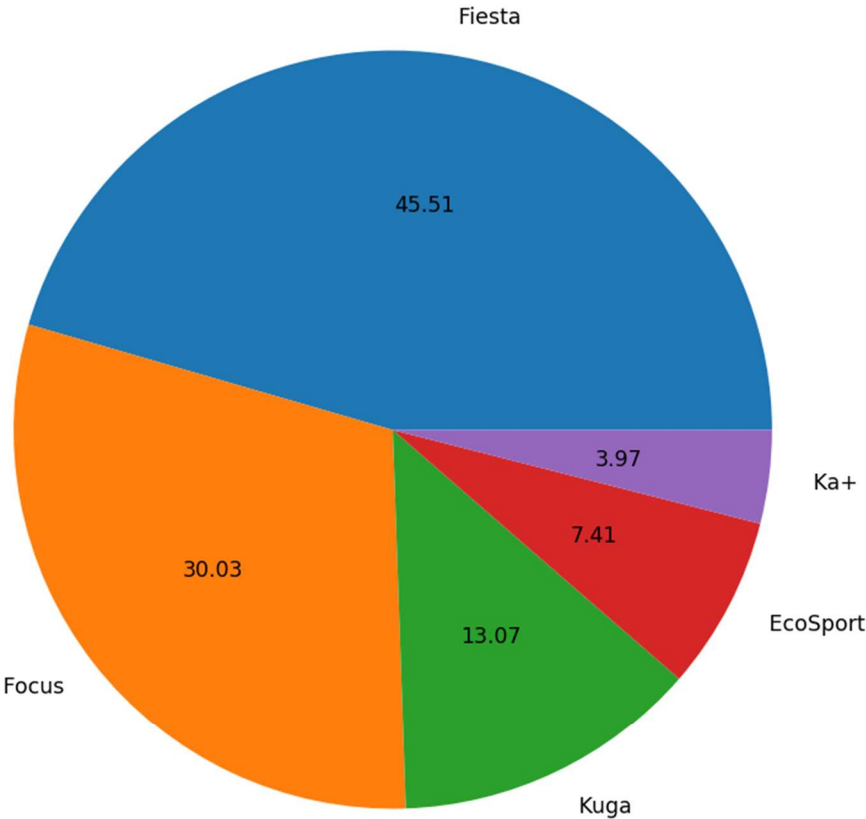
```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"

Data Visualization and Analysis

Top 5 models with manual Transmission status
model
Fiesta      6083
Focus       4014
Kuga        1747
EcoSport     990
Ka+         531
Name: count, dtype: int64
```

## Pi-chart Representation

Top 5 models with manual Transmission status



# Price analysis based on tax

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"
```

## Data Visualization and Analysis

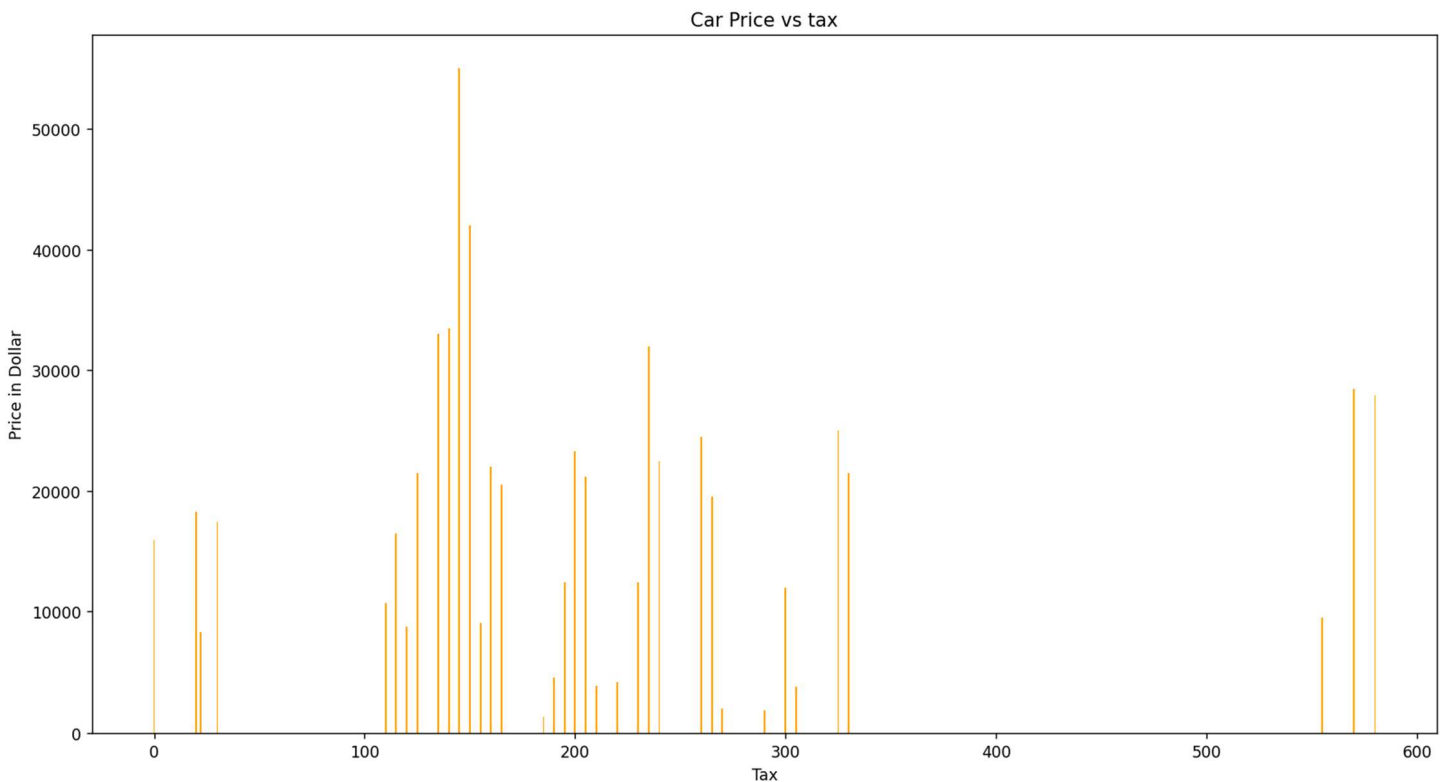
### Price analysis based on tax

#### Price of cars based on tax

	tax	price
0	150	12000
1	150	14000
2	150	13000
3	145	17500
4	145	16500
...	...	...
17961	150	8999
17962	30	7499
17963	20	9999
17964	145	8299
17965	22	8299

```
[17966 rows x 2 columns]
```

## Bar-Graph Representation



# Price analysis based on Fuel type

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"
```

```
Data Visualization and Analysis
```

```
Price analysis based on Fuel Type
```

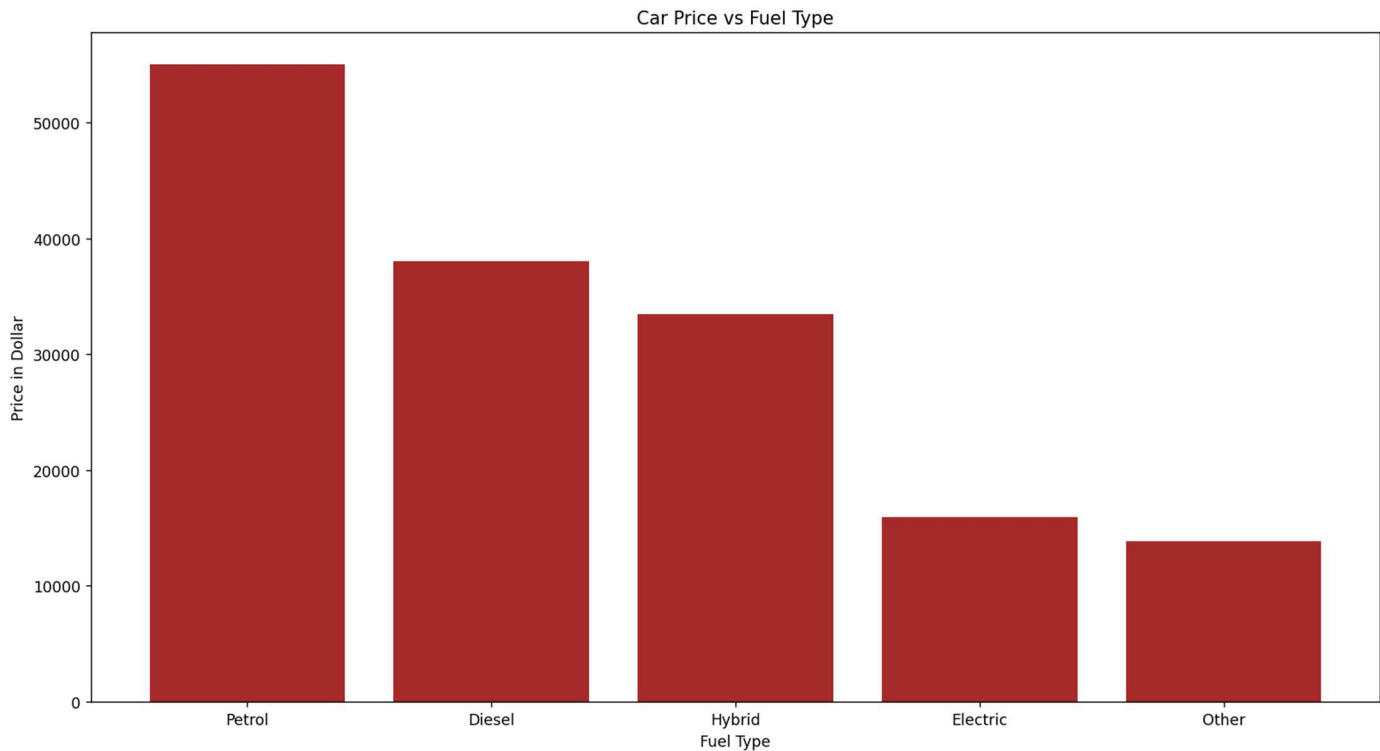
```
Price of cars based on Fuel Type
```

```
   fuelType  price
0    Petrol  12000
1    Petrol  14000
2    Petrol  13000
3    Petrol  17500
4    Petrol  16500
...      ...    ...
17961  Petrol   8999
17962  Petrol   7499
17963  Diesel   9999
17964  Petrol   8299
17965  Petrol   8299
```

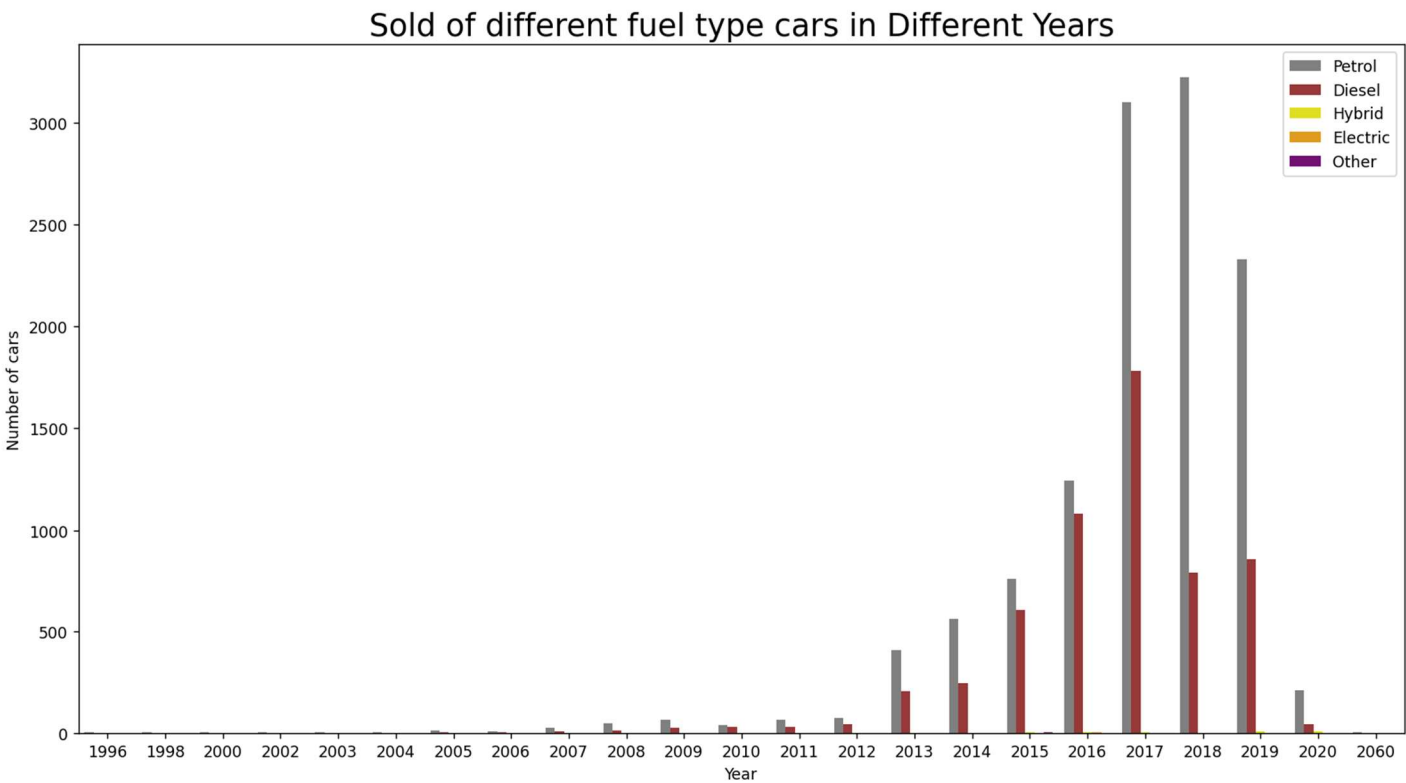
```
[17966 rows x 2 columns]
```

## Bar-Graph Representation

Sold of different fuel type cars in Different Years



# Sold of different fuel type cars in Different Years



# Price analysis based on mileage

```
"C:\Python Examples-Accenture\venv\Scripts\python.exe" "C:\Python Examples-Accenture\Accenture-case-study-3\Ford_analysis.py"
```

```
Data Visualization and Analysis
```

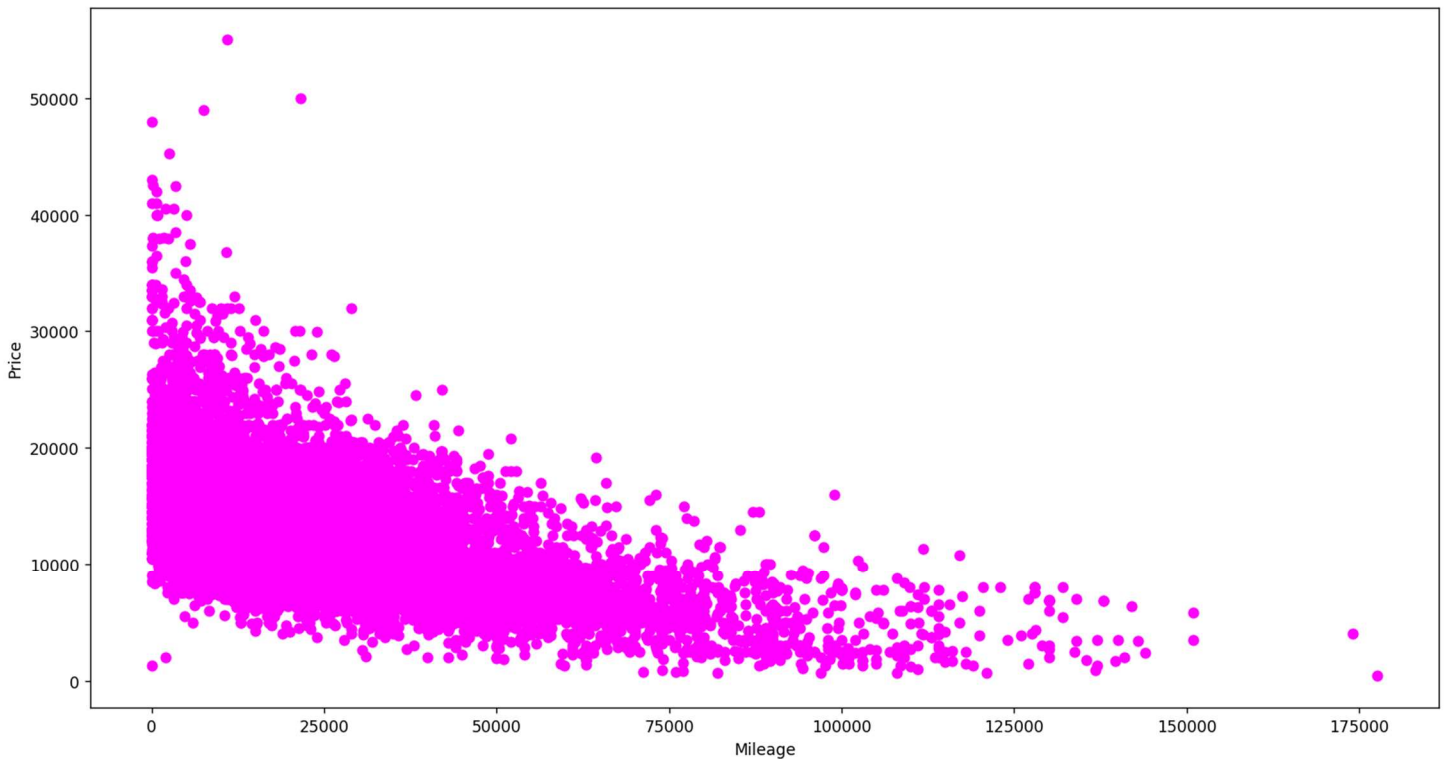
```
Price analysis based on mileage
```

```
Price of cars based on Fuel Type
```

	mileage	price
0	15944	12000
1	9083	14000
2	12456	13000
3	10460	17500
4	1482	16500
...	...	...
17961	16700	8999
17962	40700	7499
17963	7010	9999
17964	5007	8299
17965	5007	8299

```
[17966 rows x 2 columns]
```

## Scatter-graph Representation





# Finding and Insights

## Points of findings

- a) Price value counts based on the size of engine size.
- b) Price value counts based on the size of MPG.
- c) Price value counts based on the size of transmissions.
- d) Price value counts based on the size of engine year.
- e) Price value counts based on the size of tax.
- f) Price value counts based on the size of MPG.
- g) Price value counts based on the size of fuel type.
- h) Top 5 models with manual Transmission status.
- i) Top 5 models with automatic Transmission status
- j) Transmission status in Different Years

## Points of insights

- a) According to the visualization the engine size of 2 to 2.5 are maximum in price.
- b) According to the visualization the MPG OF 20 and between 27 to 35 are maximum in price.
- c) According to the visualization in 2018 car's price is maximum.
- d) According to the visualization manual transmission car's price is maximum.
- e) According to the visualization the tax of 150 dollars cars is high in price.
- f) According to the visualization patrol cars are high in price.
- g) According to the visualization the mileage of cars from 0 to 15000 are maximum in price.
- h) According to visualization patrol cars in 2018 are sold very much.
- i) According to visualization manual cars in 2017 are sold very much.
- j) According to visualization Fiesta is the top model car as manual car.
- k) According to visualization Focus is the top model car as automatic car.

So, according to this analysis, if we say that we must decide the price of the car according to their mileage, transmission, and fuel type.