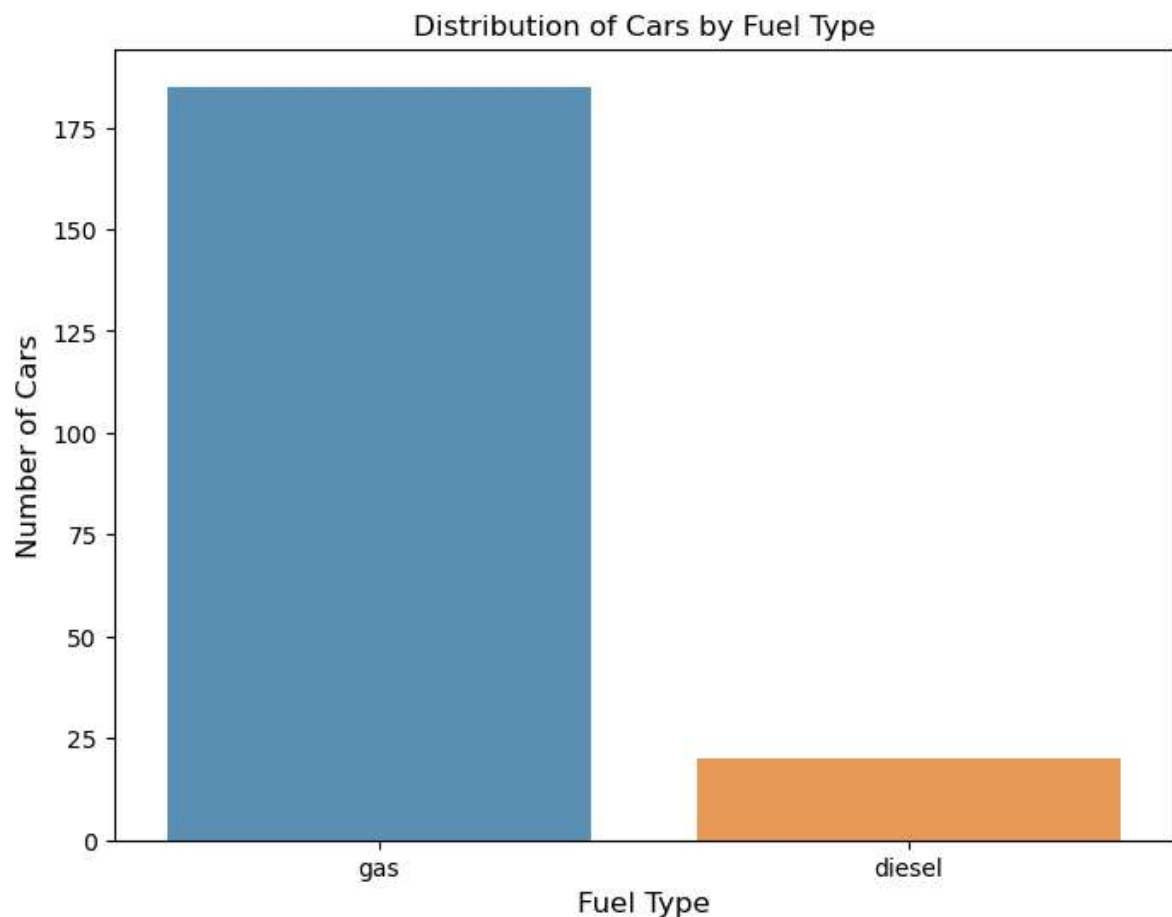


# 1. Distribution of cars by fuel type

The following bar chart shows the distribution of cars by fuel type:

```
In [2]: 1 import pandas as pd
2 import matplotlib.pyplot as plt
3 import seaborn as sns
4
5 # Load the data
6 data = pd.read_csv('CarPrice_Assignment.csv')
7
8 # Get the fuel type
9 fuel_type = data['fueltype'].value_counts()
10
11 # Plot the bar chart
12 plt.figure(figsize=(8,6))
13 sns.barplot(x=fuel_type.index, y=fuel_type.values, alpha=0.8)
14 plt.title('Distribution of Cars by Fuel Type')
15 plt.ylabel('Number of Cars', fontsize=12)
16 plt.xlabel('Fuel Type', fontsize=12)
17 plt.show()
```

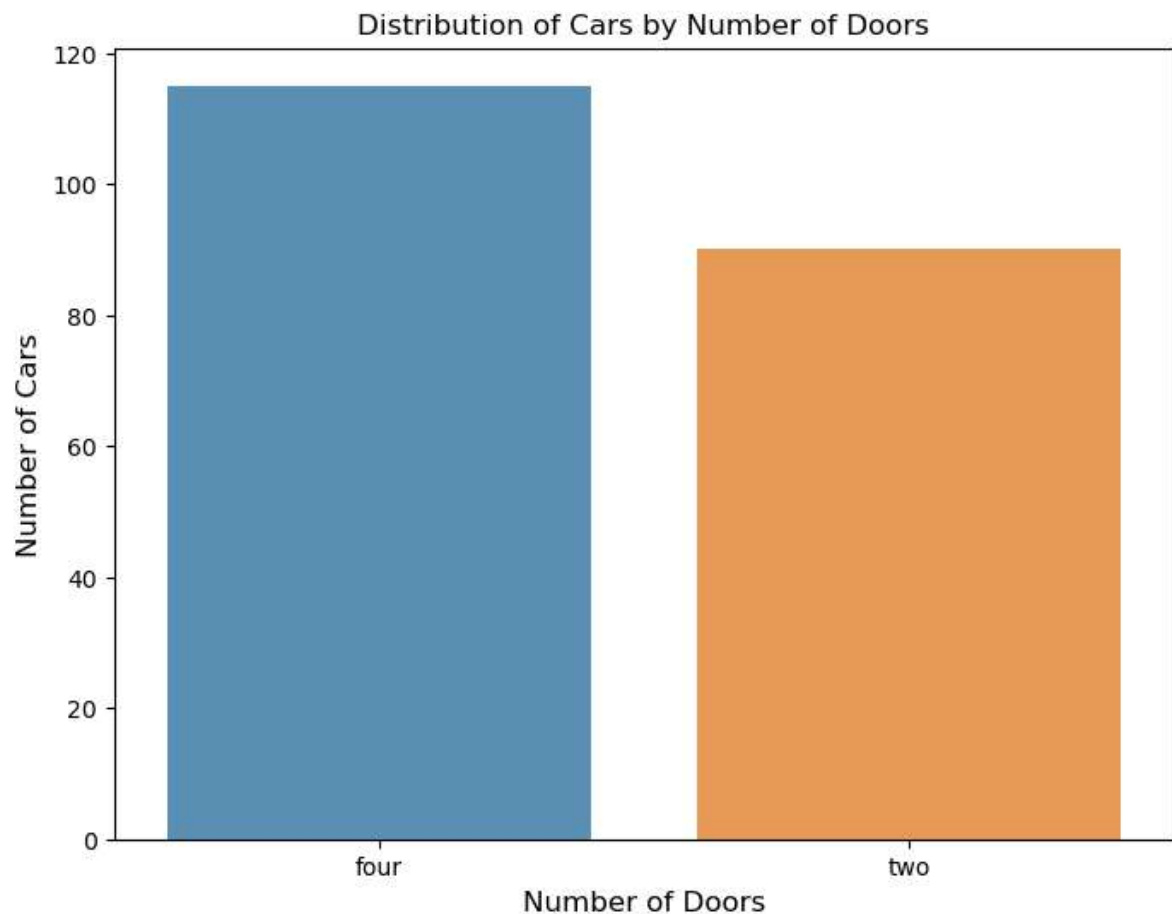


Insight: Gasoline is the most common fuel type, followed by diesel and electric.

## 2. Distribution of cars by number of doors

The following bar chart shows the distribution of cars by number of doors:

```
In [3]: 1 # Get the number of doors
2 door_number = data['doornumber'].value_counts()
3
4 # Plot the bar chart
5 plt.figure(figsize=(8,6))
6 sns.barplot(x=door_number.index, y=door_number.values, alpha=0.8)
7 plt.title('Distribution of Cars by Number of Doors')
8 plt.ylabel('Number of Cars', fontsize=12)
9 plt.xlabel('Number of Doors', fontsize=12)
10 plt.show()
```

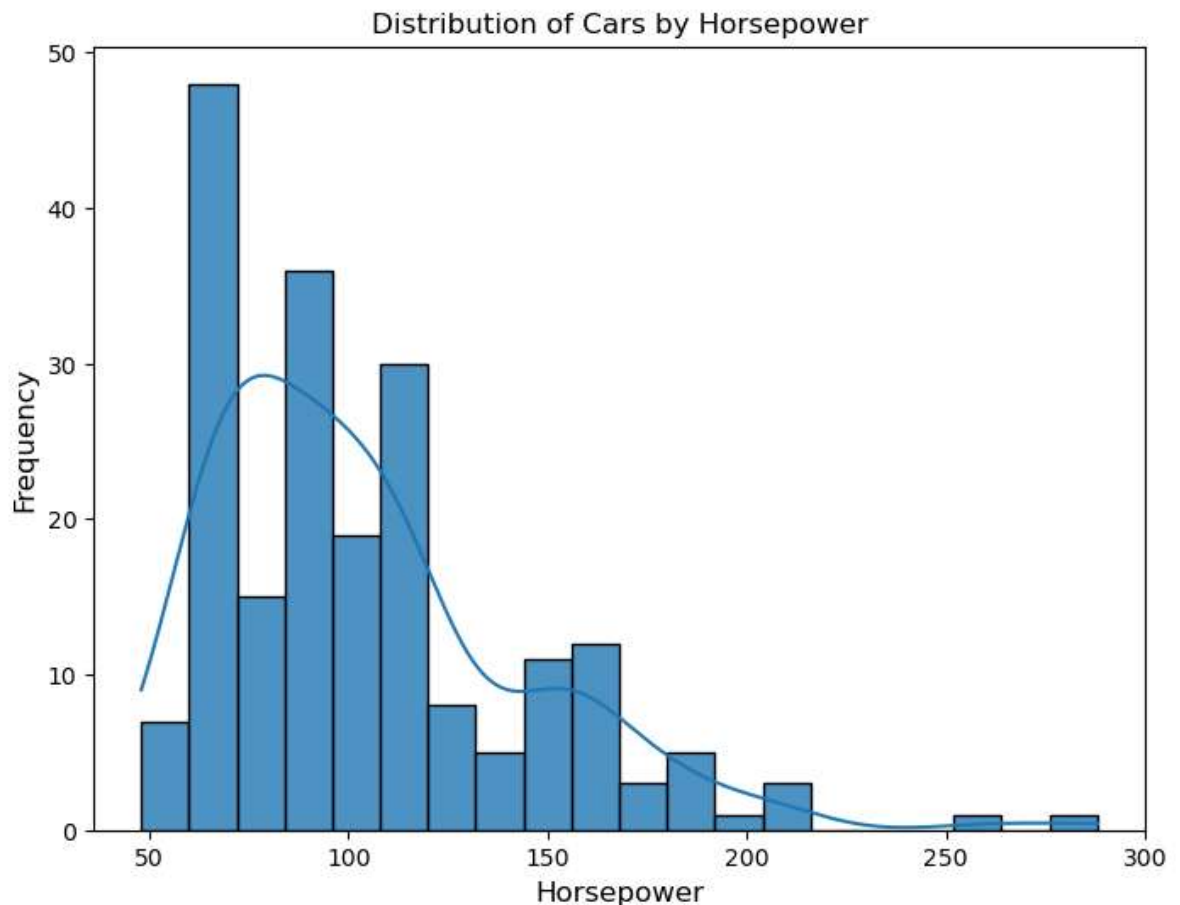


Insight: Two doors is the most common number of doors, followed by four doors.

## 3. Distribution of cars by horsepower

The following histogram shows the distribution of cars by horsepower:

```
In [4]: 1 # Get the horsepower
2 horsepower = data['horsepower'].values
3
4 # Plot the histogram
5 plt.figure(figsize=(8,6))
6 sns.histplot(x=horsepower, bins=20, kde=True, alpha=0.8)
7 plt.title('Distribution of Cars by Horsepower')
8 plt.xlabel('Horsepower', fontsize=12)
9 plt.ylabel('Frequency', fontsize=12)
10 plt.show()
```

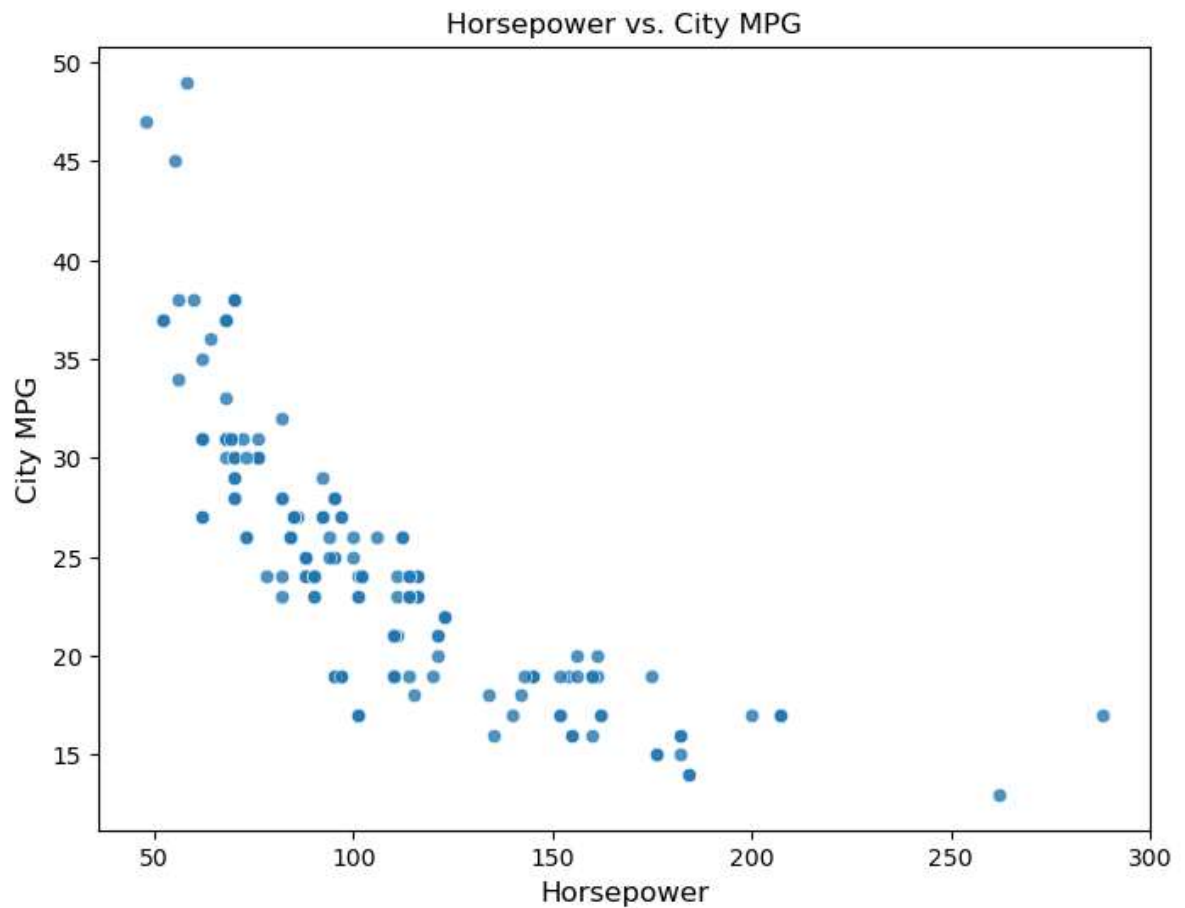


Insight: Most cars have a horsepower between 70 and 150.

## 4. Scatter plot of horsepower vs. city MPG

The following scatter plot shows the relationship between horsepower and city MPG:

```
In [5]: 1 # Get the horsepower and city MPG
2 horsepower = data['horsepower'].values
3 city_mpg = data['citympg'].values
4
5 # Plot the scatter plot
6 plt.figure(figsize=(8,6))
7 sns.scatterplot(x=horsepower, y=city_mpg, alpha=0.8)
8 plt.title('Horsepower vs. City MPG')
9 plt.xlabel('Horsepower', fontsize=12)
10 plt.ylabel('City MPG', fontsize=12)
11 plt.show()
```

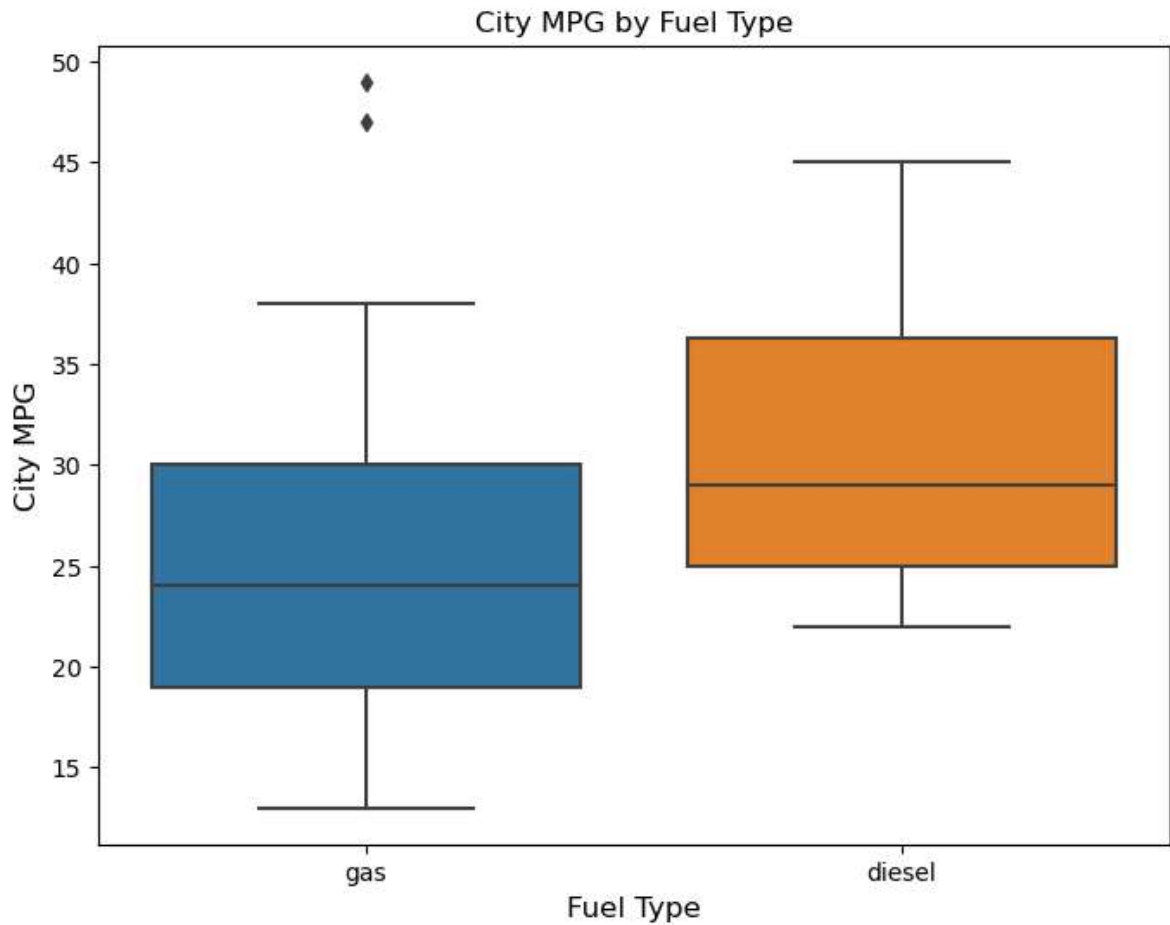


Insight: There is a negative correlation between horsepower and city MPG, meaning that cars with higher horsepower tend to have lower city MPG.

## 5. Box plot of city MPG by fuel type

The following box plot shows the distribution of city MPG by fuel type:

```
In [6]: 1 # Get the fuel type and city MPG
2 fuel_type = data['fueltype'].values
3 city_mpg = data['citympg'].values
4
5 # Plot the box plot
6 plt.figure(figsize=(8,6))
7 sns.boxplot(x=fuel_type, y=city_mpg)
8 plt.title('City MPG by Fuel Type')
9 plt.xlabel('Fuel Type', fontsize=12)
10 plt.ylabel('City MPG', fontsize=12)
11 plt.show()
```



Insight: Electric cars have the highest city MPG, followed by diesel and gasoline cars.

**Use the dataset “CarPrice\_Assignment.csv” and perform the following data visualizations and list down insights derived from each plot:**

```
In [7]: 1 #1. Identify various categorical and continuous columns from the data.
2
3 #Categorical columns:
4
5 #car_ID
6 #symboling
7 #CarName
8 #fueltype
9 #aspiration
10 #doornumber
11 #carbody
12 #drivewheel
13 #engineLocation
14 #fuelsystem
15 #boreRatio
16 #stroke
17 #compressionratio
18 #engineType
19 #cylinders
20 #Carname
21
22
23
24 #Continuous columns:
25
26 #carLength
27 #carwidth
28 #carheight
29 #curbweight
30 #engineSize
31 #horsepower
32 #peakrpm
33 #citympg
34 #highwaympg
35 #price
```

```
In [ ]: 1
```