

# Box Plot Interpretation

boxplot.png

```
In [3]: import pandas as pd
import seaborn as sns

auto = pd.read_csv('Automobile_data.csv', na_values=['?'])
```

```
In [4]: auto.head()
```

```
Out [4]:
```

	symboling	normalized-losses	make	fuel-type	aspiration	num-of-doors	body-style	drive-wheels	engine-location	wheel-base	...	engine-size	fuel-system	bore	stroke	compression-ratio
0	3	NaN	alfa-romero	gas	std	two	convertible	rwd	front	88.6	...	130	mpfi	3.47	2.68	9.0
1	3	NaN	alfa-romero	gas	std	two	convertible	rwd	front	88.6	...	130	mpfi	3.47	2.68	9.0
2	1	NaN	alfa-romero	gas	std	two	hatchback	rwd	front	94.5	...	152	mpfi	2.68	3.47	9.0
3	2	164.0	audi	gas	std	four	sedan	fwd	front	99.8	...	109	mpfi	3.19	3.40	10.0
4	2	164.0	audi	gas	std	four	sedan	4wd	front	99.4	...	136	mpfi	3.19	3.40	8.0

5 rows × 26 columns

```
In [5]: auto.describe()
```

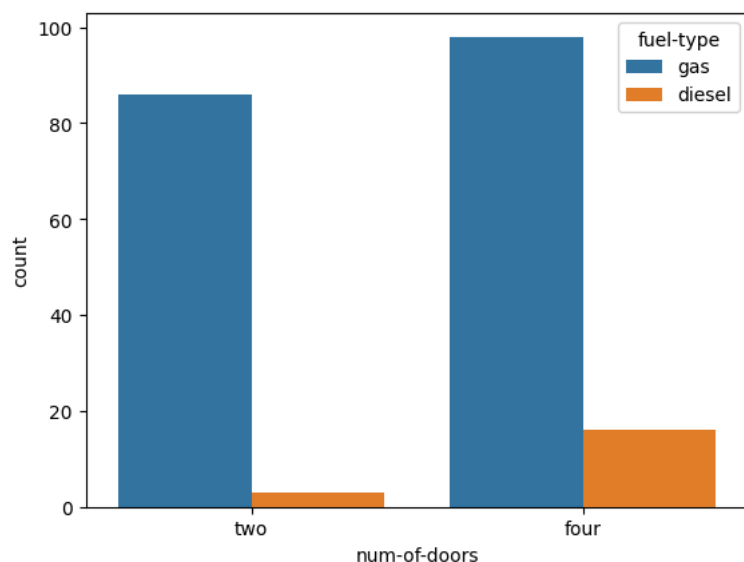
```
Out [5]:
```

	symboling	normalized-losses	wheel-base	length	width	height	curb-weight	engine-size	bore	stroke	compression-ratio
count	205.000000	164.000000	205.000000	205.000000	205.000000	205.000000	205.000000	205.000000	201.000000	201.000000	205.000000
mean	0.834146	122.000000	98.756585	174.049268	65.907805	53.724878	2555.565854	126.907317	3.329751	3.255423	10.142531
std	1.245307	35.442168	6.021776	12.337289	2.145204	2.443522	520.680204	41.642693	0.273539	0.316717	3.972040
min	-2.000000	65.000000	86.600000	141.100000	60.300000	47.800000	1488.000000	61.000000	2.540000	2.070000	7.000000
25%	0.000000	94.000000	94.500000	166.300000	64.100000	52.000000	2145.000000	97.000000	3.150000	3.110000	8.600000
50%	1.000000	115.000000	97.000000	173.200000	65.500000	54.100000	2414.000000	120.000000	3.310000	3.290000	9.000000
75%	2.000000	150.000000	102.400000	183.100000	66.900000	55.500000	2935.000000	141.000000	3.590000	3.410000	9.400000
max	3.000000	256.000000	120.900000	208.100000	72.300000	59.800000	4066.000000	326.000000	3.940000	4.170000	23.000000

```
In [6]: #distribution of the the two and four door cars with respect to the type of fuel they use.

sns.countplot(x= auto['num-of-doors'], hue= auto['fuel-type']);

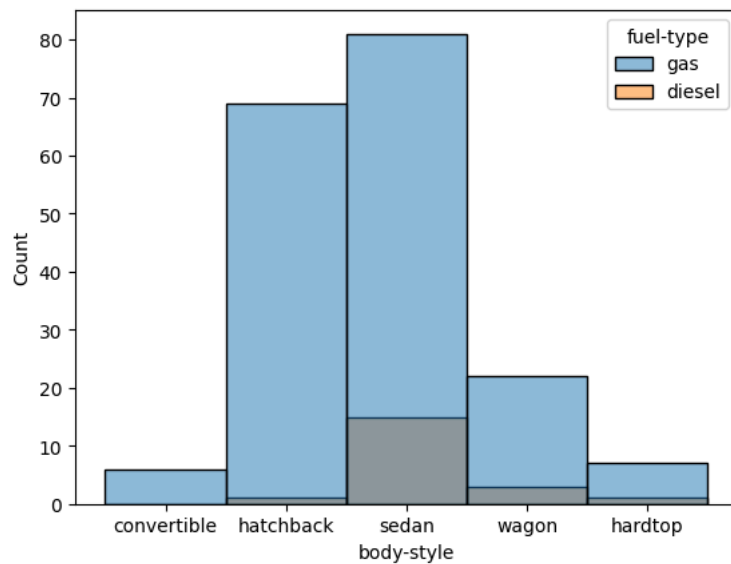
#it seems there are fewer four door diesel cars that gasoline cars and same stands for the two door cars as well
```



```
In [13]: #Distribution of cars of different body styles with respect to the type of fuel they use

sns.histplot(x = auto['body-style'], hue=auto['fuel-type']);
```

```
# it can be seen for some body-style the diesel car doesn't even exist(convertible)
```



```
In [ ]: #Total number of each type of body style cars categorized by fuel type
sns.countplot(x = auto['body-style'], hue= auto['fuel-type']);

#it can be observed that sedan cars of gas fuel type are more. There are no diesel convertible cars
```

```
In [ ]: #Horsepower of each of the fuel type with reference to the type of drive wheel present in cars

sns.catplot(x="fuel-type",
            y= "horsepower",
            hue = "num-of-doors",
            col= "drive-wheels",
            data = auto,
            kind = "box"
            );

# with reference to the type of drive wheel the car is. From data we can see there are many rear wheel drive cars.
# horsepower as well in gasoline category under two or four door cars.
```

```
In [ ]: #or can plot box plot, but insights obtained is less compared to categorical plot
sns.boxplot(x = auto['fuel-type'], y = auto['horsepower'], hue=auto['drive-wheels']);
```

```
In [ ]: !pip install seaborn
```