In [10]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [4]:
```
data=pd.read_csv("CarPrice_Assignment.csv")
data.head()
```

Out[4]:

	car_ID	symboling	CarName	fueltype	aspiration	doornumber	carbody	drivewheel	enginelocation	wheelbas
0	1	3	alfa-romero giulia	gas	std	two	convertible	rwd	front	88.(
1	2	3	alfa-romero stelvio	gas	std	two	convertible	rwd	front	88.(
2	3	1	alfa-romero Quadrifoglio	gas	std	two	hatchback	rwd	front	94.!
3	4	2	audi 100 ls	gas	std	four	sedan	fwd	front	99.8
4	5	2	audi 100ls	gas	std	four	sedan	4wd	front	99.4

5 rows × 26 columns

In [5]:
```
data.pop("car_ID")
```

Out[5]:
```
0        1
1        2
2        3
3        4
4        5
      ...
200    201
201    202
202    203
203    204
204    205
Name: car_ID, Length: 205, dtype: int64
```

In [7]:
```
data=data.select_dtypes(include="number")
data
```

Out[7]:

	symboling	wheelbase	carlength	carwidth	carheight	curbweight	enginesize	boreratio	stroke	compressionra
0	3	88.6	168.8	64.1	48.8	2548	130	3.47	2.68	
1	3	88.6	168.8	64.1	48.8	2548	130	3.47	2.68	
2	1	94.5	171.2	65.5	52.4	2823	152	2.68	3.47	
3	2	99.8	176.6	66.2	54.3	2337	109	3.19	3.40	1
4	2	99.4	176.6	66.4	54.3	2824	136	3.19	3.40	
...	...	...	...	...	...	...	...	...	...	
200	-1	109.1	188.8	68.9	55.5	2952	141	3.78	3.15	
201	-1	109.1	188.8	68.8	55.5	3049	141	3.78	3.15	
202	-1	109.1	188.8	68.9	55.5	3012	173	3.58	2.87	
203	-1	109.1	188.8	68.9	55.5	3217	145	3.01	3.40	2
204	-1	109.1	188.8	68.9	55.5	3062	141	3.78	3.15	

205 rows × 15 columns

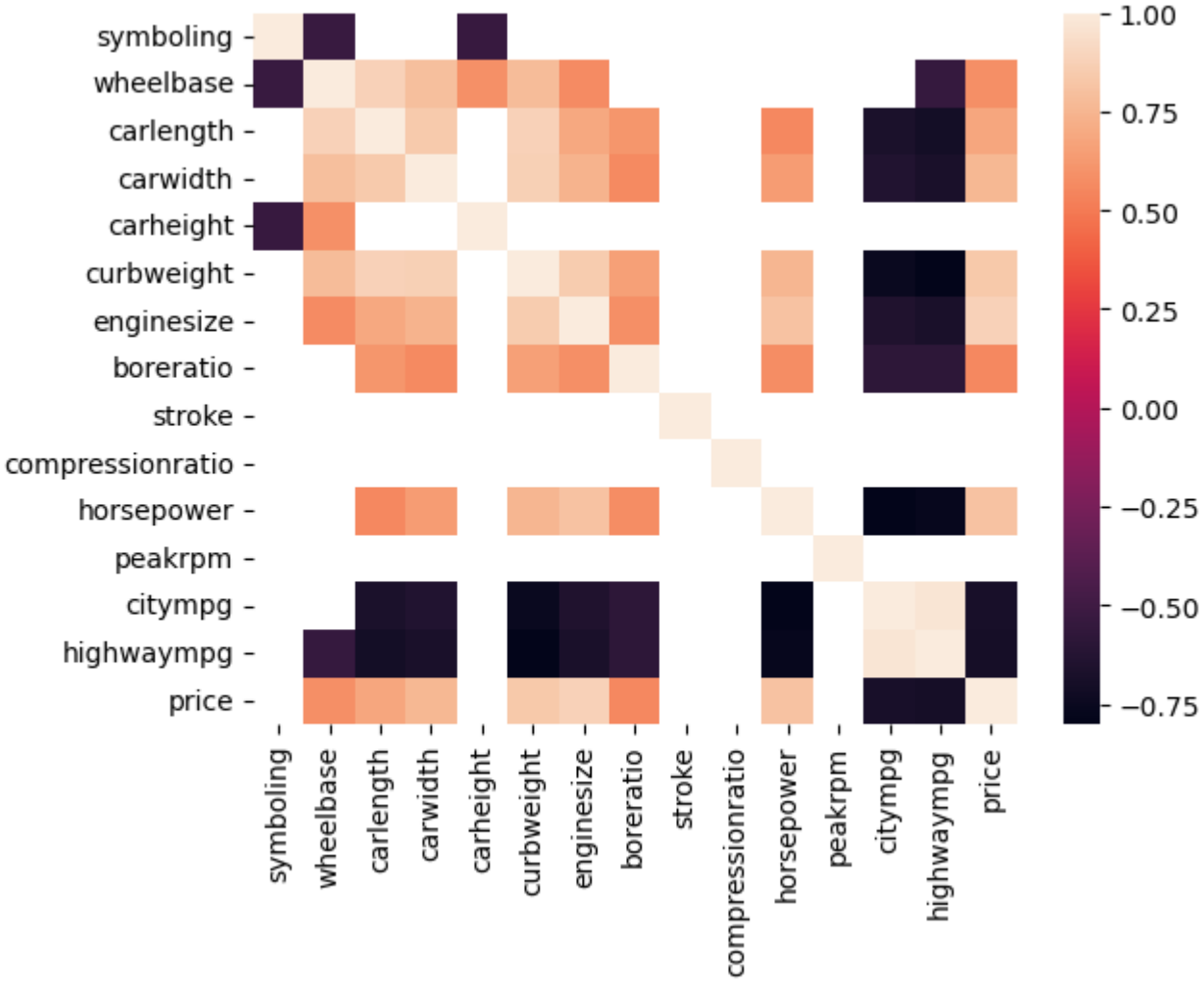
In [11]:
```
corr= data.corr()
corr=corr[(corr>0.5)|(corr<-0.5)]
corr
```

Out[11]:

	symboling	wheelbase	carlength	carwidth	carheight	curbweight	enginesize	boreratio	stroke	c
symboling	1.000000	-0.531954	NaN	NaN	-0.541038	NaN	NaN	NaN	NaN	
wheelbase	-0.531954	1.000000	0.874587	0.795144	0.589435	0.776386	0.569329	NaN	NaN	
carlength	NaN	0.874587	1.000000	0.841118	NaN	0.877728	0.683360	0.606454	NaN	
carwidth	NaN	0.795144	0.841118	1.000000	NaN	0.867032	0.735433	0.559150	NaN	
carheight	-0.541038	0.589435	NaN	NaN	1.000000	NaN	NaN	NaN	NaN	
curbweight	NaN	0.776386	0.877728	0.867032	NaN	1.000000	0.850594	0.648480	NaN	
enginesize	NaN	0.569329	0.683360	0.735433	NaN	0.850594	1.000000	0.583774	NaN	
boreratio	NaN	NaN	0.606454	0.559150	NaN	0.648480	0.583774	1.000000	NaN	
stroke	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	1.0	
compressionratio	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
horsepower	NaN	NaN	0.552623	0.640732	NaN	0.750739	0.809769	0.573677	NaN	
peakrpm	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
citympg	NaN	NaN	-0.670909	-0.642704	NaN	-0.757414	-0.653658	-0.584532	NaN	
highwaympg	NaN	-0.544082	-0.704662	-0.677218	NaN	-0.797465	-0.677470	-0.587012	NaN	
price	NaN	0.577816	0.682920	0.759325	NaN	0.835305	0.874145	0.553173	NaN	

In [12]:
```
sns.heatmap(corr)
```

Out[12]: <Axes: >



Inferences

Positive correlation

Strong

1. Car length and wheelbase, ie as the car length increases, the distance between the centers of the front and backwheel also increases.
2. Wheelbase and car width, as the distance between the front and back wheel increases, the width of the car also increases
3. Wheelbase and weight are also directly proportional
4. Car length and car width show strong correlation, ie as the length increases the width also increases
5. car width and curb weight, more width means more material and hence more weight.
6. Curb weight and engine size, more weight requires more power to pull the car forward
7. Curb weight and price, more weight means more material means more price, also means more power from the engine hence more costs for development of the engine again contributes to this strong correlation
8. Engine size and price,the bigger the engine the more power it generates and hence more development costs and material costs are associated with it. Also bigger engine size models are the higher variants hence the price goes as the engine size goes up.
9. Horsepower and price, the more power the engine generates the price goes up, this is true when we see sportscars.
10. highway mpg and city mpg, shows a very strong correlation meaning that either of as the city mileage increases or decreases, so does the other, which is very in the real world as well.

Moderate correlation

1. wheelbase and car height, enginesize, price also also show a moderate correlation
2. car width and engine size, boreratio,horsepower,price all show moderate correlation, as the width increases, the weight increases and hence more power is required is pull the car forward. And all these measures are related to generating power, hence as they go up, the car width also goes up.
3. Curb weight and boreratio, horsepower. This can be explained using the same reason as the one above.
4. Car length and engine size, boreratio, horsepower. Same explaination as the car width.
5. engine size and bore ratio. The bore is the diameter of the cyclinder in the engine. As the engine size increases, the bore usually increases but this may not always be true. Hence they show a moderate correlation and not a very strong one.

Negative Correlation

Strong

1. Horsepower and mileage(city mpg and highway mpg). As the horsepower increases,it burns more fuel in less time and hence the mileage goes down.

Moderate

1. The symboling and the wheelbase demonstrate this, hence as the wheelbase increase the symboling decreases
2. Wheelbase, Car width, car length, curb weight, engine size, and citympg , highway mpg. All the former factors contribute to generating more power from the engine, hence when they increase, it requires more fuel thus decreasing the mileage.

In [ ]: