In [31]:

```
import numpy as np
import pandas as pd
from sklearn import preprocessing
from sklearn import linear_model
```

In [4]:

```
df = pd.read_csv("CarPrice_Assignment.csv")
df
```

Out[4]:

	car_ID	symboling	CarName	fueltype	aspiration	doornumber	carbody	drivewheel
0	1	3	alfa-romero giulia	gas	std	two	convertible	rwd
1	2	3	alfa-romero stelvio	gas	std	two	convertible	rwd
2	3	1	alfa-romero Quadrifoglio	gas	std	two	hatchback	rwd
3	4	2	audi 100 ls	gas	std	four	sedan	fwd
4	5	2	audi 100ls	gas	std	four	sedan	4wd
200	201	-1	volvo 145e (sw)	gas	std	four	sedan	rwd
201	202	-1	volvo 144ea	gas	turbo	four	sedan	rwd
202	203	-1	volvo 244dl	gas	std	four	sedan	rwd
203	204	-1	volvo 246	diesel	turbo	four	sedan	rwd
204	205	-1	volvo 264gl	gas	turbo	four	sedan	rwd

205 rows × 26 columns

In [7]:

df.info

Out[7]:

		ataFrame.info aspiration	of \	ca	ar_ID sy	mboling	5		С
0	1	3	`alf	a-rome	ro giuli	a	gas	std	
1	2	3			o stelvi		gas	std	
2	3				drifogli		_	std	
3	4	2	a-i Oille	•	idi 100 l		gas	std	
<i>3</i>	5	2			nudi 100 l		gas		
•				ä	1001 1001	S	gas	std	
	201	•••		7	145- /		•••	•••	
200	201	-1			145e (sw	•	gas .	std	
201	202	-1			lvo 144e		•	urbo	
202	203	-1			lvo 244d		gas	std	
203	204	-1			volvo 24			urbo	
204	205	-1		VC	lvo 264g	1	gas t	urbo	
do	ornumber	carbody	drive	wheel	enginelo	cation	wheelbase		\
0	two	convertible		rwd	•	front	88.6		
1	two	convertible		rwd		front	88.6		
2	two	hatchback		rwd		front	94.5		
3	four	sedan		fwd		front	99.8		
4	four	sedan		4wd		front	99.4		
••	•••					•••	•••		
200	four	sedan		rwd		front	109.1		
201	four	sedan		rwd		front	109.1		
202	four	sedan		rwd		front	109.1		
203	four	sedan		rwd		front	109.1		
204	four	sedan		rwd		front	109.1		
201	1001	Scaan		ı wa		110116	105.1	•••	
_				_					
	enginesize	fuelsystem	bore	ratio	stroke	compres	ssionratio	norsep	ower
\		-	bore			compres		norsep	
\ 0	130	mpfi	bore	3.47	2.68	compres	9.0	norsep	111
\ 0 1	130 130	mpfi mpfi	bore	3.47 3.47	2.68 2.68	compres	9.0 9.0	norsep	111 111
\ 0 1 2	130 130 152	mpfi mpfi mpfi	bore	3.47 3.47 2.68	2.68 2.68 3.47	compres	9.0 9.0 9.0	norsep	111 111 154
\ 0 1 2 3	130 130 152 109	mpfi mpfi mpfi mpfi	bore	3.47 3.47 2.68 3.19	2.68 2.68 3.47 3.40	compres	9.0 9.0 9.0 10.0	norsep	111 111 154 102
\ 0 1 2	130 130 152	mpfi mpfi mpfi	bore	3.47 3.47 2.68	2.68 2.68 3.47	compres	9.0 9.0 9.0	norsep	111 111 154
\ 0 1 2 3 4	130 130 152 109 136	mpfi mpfi mpfi mpfi mpfi	bore	3.47 3.47 2.68 3.19 3.19	2.68 2.68 3.47 3.40 3.40	compres	9.0 9.0 9.0 10.0 8.0	norsep	111 111 154 102 115
\ 0 1 2 3 4 200	130 130 152 109 136 	mpfi mpfi mpfi mpfi mpfi 	bore	3.47 3.47 2.68 3.19 3.19 	2.68 2.68 3.47 3.40 3.40 	compres	9.0 9.0 9.0 10.0 8.0 	norsep	111 111 154 102 115
\ 0 1 2 3 4 200 201	130 130 152 109 136 141	mpfi mpfi mpfi mpfi mpfi mpfi mpfi	bore	3.47 3.47 2.68 3.19 3.19 3.78 3.78	2.68 2.68 3.47 3.40 3.40 3.15	compres	9.0 9.0 9.0 10.0 8.0 9.5 8.7	norsep	111 111 154 102 115 114
\ 0 1 2 3 4 200 201 202	130 130 152 109 136 141 141	mpfi mpfi mpfi mpfi mpfi mpfi mpfi mpfi	bore	3.47 3.47 2.68 3.19 3.19 3.78 3.78 3.58	2.68 2.68 3.47 3.40 3.40 3.15 3.15 2.87	compres	9.0 9.0 9.0 10.0 8.0 9.5 8.7 8.8	norsep	111 111 154 102 115 114 160
\ 0 1 2 3 4 200 201 202 203	130 130 152 109 136 141 141 173 145	mpfi mpfi mpfi mpfi mpfi mpfi mpfi idi		3.47 3.47 2.68 3.19 3.19 3.78 3.78 3.58 3.01	2.68 2.68 3.47 3.40 3.40 3.15 3.15 2.87 3.40	compres	9.0 9.0 9.0 10.0 8.0 9.5 8.7 8.8 23.0	norsep	111 111 154 102 115 114 160 134 106
\ 0 1 2 3 4 200 201 202	130 130 152 109 136 141 141	mpfi mpfi mpfi mpfi mpfi mpfi mpfi mpfi		3.47 3.47 2.68 3.19 3.19 3.78 3.78 3.58	2.68 2.68 3.47 3.40 3.40 3.15 3.15 2.87	compres	9.0 9.0 9.0 10.0 8.0 9.5 8.7 8.8	norsep	111 111 154 102 115 114 160
\ 0 1 2 3 4 200 201 202 203 204	130 130 152 109 136 141 141 173 145 141	mpfi mpfi mpfi mpfi mpfi mpfi mpfi idi mpfi		3.47 3.47 2.68 3.19 3.78 3.78 3.58 3.01 3.78	2.68 2.68 3.47 3.40 3.40 3.15 3.15 2.87 3.40 3.15	compres	9.0 9.0 9.0 10.0 8.0 9.5 8.7 8.8 23.0	norsep	111 111 154 102 115 114 160 134 106
\ 0 1 2 3 4 200 201 202 203 204	130 130 152 109 136 141 173 145 141	mpfi mpfi mpfi mpfi mpfi mpfi idi mpfi	aympg	3.47 3.47 2.68 3.19 3.78 3.78 3.78 3.78 3.78 pri	2.68 2.68 3.47 3.40 3.40 3.15 3.15 2.87 3.40 3.15	compres	9.0 9.0 9.0 10.0 8.0 9.5 8.7 8.8 23.0	norsep	111 111 154 102 115 114 160 134 106
\ 0 1 2 3 4 200 201 202 203 204	130 130 152 109 136 141 173 145 141 Deakrpm cit	mpfi mpfi mpfi mpfi mpfi mpfi mpfi idi mpfi tympg highwa	aympg 27	3.47 3.47 2.68 3.19 3.78 3.78 3.78 3.78 3.78 3.78	2.68 2.68 3.47 3.40 3.40 3.15 3.15 2.87 3.40 3.15	compres	9.0 9.0 9.0 10.0 8.0 9.5 8.7 8.8 23.0	norsep	111 111 154 102 115 114 160 134 106
\ 0 1 2 3 4 200 201 202 203 204 F	130 130 152 109 136 141 173 145 141 Deakrpm cir 5000 5000	mpfi mpfi mpfi mpfi mpfi mpfi mpfi idi mpfi tympg highwa	aympg 27 27	3.47 3.47 2.68 3.19 3.78 3.78 3.78 3.58 3.01 3.78 pri 13495 16506	2.68 2.68 3.47 3.40 3.15 3.15 2.87 3.40 3.15	compres	9.0 9.0 9.0 10.0 8.0 9.5 8.7 8.8 23.0	norsep	111 111 154 102 115 114 160 134 106
\ 0 1 2 3 4 200 201 202 203 204	130 130 152 109 136 141 141 173 145 141 Deakrpm cit 5000 5000	mpfi mpfi mpfi mpfi mpfi mpfi mpfi idi mpfi tympg highwa	aympg 27 27 26	3.47 3.47 2.68 3.19 3.78 3.78 3.78 3.58 3.01 3.78 pri 13495 16506	2.68 2.68 3.47 3.40 3.15 3.15 2.87 3.40 3.15	compres	9.0 9.0 9.0 10.0 8.0 9.5 8.7 8.8 23.0	norsep	111 111 154 102 115 114 160 134 106
\ 0 1 2 3 4 200 201 202 203 204 P 0 1 2 3	130 130 152 109 136 141 141 173 145 141 Deakrpm cit 5000 5000 5000	mpfi mpfi mpfi mpfi mpfi mpfi idi mpfi itympg highwa 21 21 21	aympg 27 27 26 30	3.47 3.47 2.68 3.19 3.78 3.78 3.78 3.78 3.78 13495 16506 13956	2.68 2.68 3.47 3.40 3.40 3.15 3.15 2.87 3.40 3.15	compres	9.0 9.0 9.0 10.0 8.0 9.5 8.7 8.8 23.0	norsep	111 111 154 102 115 114 160 134 106
\ 0 1 2 3 4 200 201 202 203 204	130 130 152 109 136 141 173 145 141 Deakrpm cit 5000 5000 5000 5000 5500	mpfi mpfi mpfi mpfi mpfi mpfi mpfi idi mpfi tympg highwa 21 21 21 19 24	aympg 27 27 26 30 22	3.47 3.47 2.68 3.19 3.78 3.78 3.58 3.01 3.78 pri 13495 16506 13956 17456	2.68 2.68 3.47 3.40 3.15 3.15 2.87 3.40 3.15	compres	9.0 9.0 9.0 10.0 8.0 9.5 8.7 8.8 23.0	norsep	111 111 154 102 115 114 160 134 106
\ 0 1 2 3 4 200 201 202 203 204 F	130 130 152 109 136 141 173 145 141 Deakrpm cit 5000 5000 5000 5000	mpfi mpfi mpfi mpfi mpfi mpfi mpfi idi mpfi tympg highwa 21 21 21 19 24 18	aympg 27 27 26 30 22	3.47 3.47 2.68 3.19 3.78 3.78 3.78 3.78 3.58 3.01 3.78 pri 13495 16506 13956 17456	2.68 2.68 3.47 3.40 3.15 3.15 2.87 3.40 3.15	compres	9.0 9.0 9.0 10.0 8.0 9.5 8.7 8.8 23.0	norsep	111 111 154 102 115 114 160 134 106
\ 0 1 2 3 4 200 201 202 203 204 0 1 2 3 4 200	130 130 152 109 136 141 141 173 145 141 5000 5000 5000 5500 5500 5400	mpfi mpfi mpfi mpfi mpfi mpfi mpfi idi mpfi tympg highwa 21 21 21 21 29 24 18	aympg 27 27 26 30 22 	3.47 3.47 2.68 3.19 3.78 3.78 3.78 3.58 3.01 3.78 pri 13495 16506 13956 17456	2.68 2.68 3.47 3.40 3.15 3.15 2.87 3.40 3.15	compres	9.0 9.0 9.0 10.0 8.0 9.5 8.7 8.8 23.0	norsep	111 111 154 102 115 114 160 134 106
\ 0 1 2 3 4 200 201 202 203 204 0 1 2 3 4 200 201	130 130 152 109 136 141 141 173 145 141 5000 5000 5000 5000 5500 5500 5	mpfi mpfi mpfi mpfi mpfi mpfi mpfi idi mpfi idi mpfi 21 21 19 24 18 23 19	aympg 27 27 26 30 22 28 25	3.47 3.47 2.68 3.19 3.78 3.78 3.78 3.78 3.78 9ri 13495 16506 13956 17456	2.68 2.68 3.47 3.40 3.40 3.15 3.15 2.87 3.40 3.15	compres	9.0 9.0 9.0 10.0 8.0 9.5 8.7 8.8 23.0	norsep	111 111 154 102 115 114 160 134 106
\ 0 1 2 3 4 200 201 202 203 204 0 1 2 3 4 200 201 202	130 130 152 109 136 141 141 173 145 141 Deakrpm cit 5000 5000 5000 5000 5000 5500 5500 5300 5500	mpfi mpfi mpfi mpfi mpfi mpfi mpfi mpfi	aympg 27 27 26 30 22 28 25 23	3.47 3.47 2.68 3.19 3.78 3.78 3.78 3.78 3.78 13495 16506 13956 17456 16845 19045 21485	2.68 2.68 3.47 3.40 3.40 3.15 3.15 2.87 3.40 3.15	compres	9.0 9.0 9.0 10.0 8.0 9.5 8.7 8.8 23.0	norsep	111 111 154 102 115 114 160 134 106
\ 0 1 2 3 4 200 201 202 203 204 0 1 2 3 4 200 201	130 130 152 109 136 141 141 173 145 141 5000 5000 5000 5000 5500 5500 5	mpfi mpfi mpfi mpfi mpfi mpfi mpfi idi mpfi idi mpfi 21 21 19 24 18 23 19	aympg 27 27 26 30 22 28 25	3.47 3.47 2.68 3.19 3.78 3.78 3.78 3.78 3.78 9ri 13495 16506 13956 17456	2.68 2.68 3.47 3.40 3.15 3.15 2.87 3.40 3.15	compres	9.0 9.0 9.0 10.0 8.0 9.5 8.7 8.8 23.0	norsep	111 111 154 102 115 114 160 134 106

[205 rows x 26 columns]>

```
In [9]:
```

```
df.isna().sum()
Out[9]:
car_ID
                     0
symboling
                     0
CarName
                     0
fueltype
                     0
aspiration
                     0
doornumber
                     0
carbody
                     0
drivewheel
                     0
enginelocation
                     0
wheelbase
                     0
                     0
carlength
carwidth
                     0
carheight
                     0
                     0
curbweight
enginetype
                     0
cylindernumber
                     0
enginesize
                     0
                     0
fuelsystem
boreratio
stroke
                     0
compressionratio
                     0
horsepower
                     0
peakrpm
                     0
citympg
                     0
                     0
highwaympg
price
                     0
dtype: int64
In [ ]:
```

Data Preprocessing

```
In [13]:
```

```
lab_en = preprocessing.LabelEncoder()
lab_en.fit(df.fueltype)
df.fueltype = lab_en.transform(df.fueltype)
```

```
In [14]:
```

```
lab_en.fit(df.doornumber)
df.doornumber = lab_en.transform(df.doornumber)
```

```
In [15]:
```

```
lab_en.fit(df.aspiration)
df.aspiration = lab_en.transform(df.aspiration)
```

In [16]:

```
lab_en.fit(df.drivewheel)
df.drivewheel = lab_en.transform(df.drivewheel)
```

In [17]:

```
lab_en.fit(df.carbody)
df.carbody = lab_en.transform(df.carbody)
```

In [18]:

```
lab_en.fit(df.enginelocation)
df.enginelocation = lab_en.transform(df.enginelocation)
```

In [19]:

```
lab_en.fit(df.enginetype)
df.enginetype = lab_en.transform(df.enginetype)
```

In [20]:

```
lab_en.fit(df.cylindernumber)
df.cylindernumber = lab_en.transform(df.cylindernumber)
```

In [21]:

```
lab_en.fit(df.fuelsystem)
df.fuelsystem = lab_en.transform(df.fuelsystem)
```

```
In [39]:
```

df

Out[39]:

•	fueltype	aspiration	doornumber	carbody	drivewheel	enginelocation	wheelbase	 engine
) 3	1	0	1	0	2	0	88.6	
)	1	0	1	0	2	0	88.6	
)	1	0	1	2	2	0	94.5	
3	1	0	0	3	1	0	99.8	
3	1	0	0	3	0	0	99.4	
∋)	1	0	0	3	2	0	109.1	
Э	1	1	0	3	2	0	109.1	
П	1	0	0	3	2	0	109.1	
3	0	1	0	3	2	0	109.1	
β	1	1	0	3	2	0	109.1	

In [34]:

Out[34]:

LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=F
alse)

```
In [35]:
lreg.coef_
Out[35]:
array([-9.21677014e+03, 2.42352286e+02, -5.13930613e+02, -8.19139857e+02,
        1.16616708e+03, 1.03974427e+04, 1.10593867e+02, -2.91102195e+01,
        7.12602036e+02, 1.84232716e+02, 2.35662037e+00, 1.90948576e+02,
        1.28166071e+02, 1.01622195e+02, -1.48294905e+02, -2.74450799e+03,
       -2.82460445e+03, -5.48092812e+02, 2.06295041e+01, 2.10440991e+00,
       -1.45461967e+02, 1.35183217e+02])
In [37]:
lreg.intercept_
Out[37]:
-49313.25002503426
In [44]:
lreg.predict([[1,0,1,0,2,0,88.6,168.8,64.1,48.8,2548,0,0,130,0,3.47,2.68,9,111,5000,21,27
Out[44]:
array([13438.04815291])
In [38]:
df['price'][0]
Out[38]:
13495.0
In [ ]:
```

predicted price is similar to the actual value