

```
import numpy as np
import pandas as pd
```

```
df=pd.read_csv("/content/cars - cars.....csv")
df.head(5)
```

	name	year	selling_price	km_driven	fuel	seller_type	transmission	owner	grid
0	Maruti 800 AC BSII	1992	50000	100000	Petrol	Individual	Manual	Fourth & Above Owner	
1	Maruti Gypsy E MG410W ST	1995	95000	100000	Petrol	Individual	Manual	Second Owner	
2	Mahindra Jeep CL 500 MDI	1996	250000	35000	Diesel	Individual	Manual	Second Owner	
3	Mahindra Jeep MM 540	1996	200000	60000	Diesel	Individual	Manual	First Owner	
4	Mahindra Jeep CL 500 MDI	1997	150000	120000	Diesel	Individual	Manual	Third Owner	

Next steps: [Generate code with df](#) [New interactive sheet](#)

```
df=df.drop(columns=['name'])
```

```
df.head(2)
```

	year	selling_price	km_driven	fuel	seller_type	transmission	owner	grid
0	1992	50000	100000	Petrol	Individual	Manual	Fourth & Above Owner	
1	1995	95000	100000	Petrol	Individual	Manual	Second Owner	

Next steps: [Generate code with df](#) [New interactive sheet](#)

```
x=df.drop(columns=['owner'])
y=df['owner']
```

```
from sklearn.preprocessing import LabelEncoder
```

```
lb=LabelEncoder()
```

```
x['fuel']=lb.fit_transform(x['fuel'])
```

```
x['transmission']=lb.fit_transform(x['transmission'])
x['seller_type']=lb.fit_transform(x['seller_type'])
```

```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.2,random_state=2)
```

```
np.round(x_train.describe(),2)
```

	year	selling_price	km_driven	fuel	seller_type	transmission	grid
count	3472.00	3472.00	3472.00	3472.00	3472.00	3472.0	
mean	2013.13	499976.03	66102.83	2.50	0.79	0.9	
std	4.19	554884.90	46094.87	1.51	0.46	0.3	
min	1992.00	20000.00	1.00	0.00	0.00	0.0	
25%	2011.00	210000.00	35000.00	1.00	1.00	1.0	
50%	2014.00	350000.00	60000.00	3.00	1.00	1.0	
75%	2016.00	600000.00	90000.00	4.00	1.00	1.0	
max	2020.00	8150000.00	560000.00	4.00	2.00	1.0	

```
from sklearn.preprocessing import MinMaxScaler
```

```
MM=MinMaxScaler()
```

```
mn=MM.fit_transform(x_train)
```

```
new=pd.DataFrame(mn,columns=x_train.columns)
```

```
np.round(new.describe(),2)
```

	year	selling_price	km_driven	fuel	seller_type	transmission	grid
count	3472.00	3472.00	3472.00	3472.00	3472.00	3472.0	
mean	0.75	0.06	0.12	0.62	0.39	0.9	
std	0.15	0.07	0.08	0.38	0.23	0.3	
min	0.00	0.00	0.00	0.00	0.00	0.0	
25%	0.68	0.02	0.06	0.25	0.50	1.0	
50%	0.79	0.04	0.11	0.75	0.50	1.0	
75%	0.86	0.07	0.16	1.00	0.50	1.0	
max	1.00	1.00	1.00	1.00	1.00	1.0	

Start coding or [generate](#) with AI.