


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

```
df= pd.read_csv('/content/cars.csv')
```

```
df.head()
```

	year_bought	km_driven	transmission	owner	selling_price	
0	2007	70000	Manual	First Owner	60000	
1	2007	50000	Manual	First Owner	135000	
2	2012	100000	Manual	First Owner	600000	
3	2017	46000	Manual	First Owner	250000	
4	2014	141000	Manual	Second Owner	450000	

```
df.describe()
```

	year_bought	km_driven	selling_price	
count	4340.000000	4340.000000	4.340000e+03	
mean	2013.090783	66215.777419	5.041273e+05	
std	4.215344	46644.102194	5.785487e+05	
min	1992.000000	1.000000	2.000000e+04	
25%	2011.000000	35000.000000	2.087498e+05	
50%	2014.000000	60000.000000	3.500000e+05	
75%	2016.000000	90000.000000	6.000000e+05	
max	2020.000000	806599.000000	8.900000e+06	

```
df.corr()['selling_price']
```

```
year_bought    0.413922
km_driven      -0.192289
selling_price   1.000000
Name: selling_price, dtype: float64
```

```
df.isnull().sum()
```

```
year_bought    0
km_driven      0
transmission    0
owner          0
selling_price   0
dtype: int64
```

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

```
le=LabelEncoder()  
df.owner=le.fit_transform(df.owner)
```

```
df.head()
```

	year_bought	km_driven	transmission	owner	selling_price
0	2007	70000	Manual	0	60000
1	2007	50000	Manual	0	135000
2	2012	100000	Manual	0	600000
3	2017	46000	Manual	0	250000
4	2014	141000	Manual	2	450000

```
df['owner'].unique()
```

```
array([0, 2, 1, 4, 3])
```

```
df=df.drop(["transmission"],axis="columns")
```

```
df.insert(0, "X0" , 1)
```

```
df.head()
```

	X0	year_bought	km_driven	owner	selling_price
0	1	2007	70000	0	60000
1	1	2007	50000	0	135000
2	1	2012	100000	0	600000
3	1	2017	46000	0	250000
4	1	2014	141000	2	450000

```
y=df["selling_price"]  
x=df[["X0", "year_bought", "km_driven", "owner"]]
```

```
x=x.to_numpy()  
y=y.to_numpy()
```

x

```
array([[ 1, 2007, 70000, 0],
       [ 1, 2007, 50000, 0],
       [ 1, 2012, 100000, 0],
       ...,
       [ 1, 2009, 83000, 2],
       [ 1, 2016, 90000, 0],
       [ 1, 2016, 40000, 0]])
```

y

```
array([ 60000, 135000, 600000, ..., 110000, 865000, 225000])
```

```
x_t = x.transpose()
print(x_t)
```

```
[[ 1 1 1 ... 1 1 1]
 [ 2007 2007 2012 ... 2009 2016 2016]
 [ 70000 50000 100000 ... 83000 90000 40000]
 [ 0 0 0 ... 2 0 0]]
```

```
x_m_t = np.matmul(x_t,x)
```

```
inv = np.linalg.inv(x_m_t)
```

```
inv_x_t = np.matmul(inv, x_t)
```

```
answer= np.matmul(inv_x_t, y)
```

```
print(answer)
```

```
[-1.07203019e+08  5.35179407e+04 -2.03126104e-01 -1.93553204e+04]
```

```
x1= int(input())
```

```
x2=int(input())
```

```
x3= int(input())
```

```
y_pred= -1.07203019e+08 + 5.35179407e+04*x1 + (-2.03126104e-01)*x2 + (-1.93553204e+04)*x3
print(y_pred)
```

```
☞ 2007
40000
2
160652.29993999773
```

[Colab paid products](#) - [Cancel contracts here](#)

✓ 13s completed at 4:39 PM ● ✕