```
In [1]:  import pandas as pd
         import numpy as np
         from sklearn import linear_model
```

```
In [3]:  df = pd.read_csv('car data.csv')
```

```
In [4]:  df
```

Out[4]:
	speed	car_age	experien	risk
0	200	15	5.0	85
1	90	17	13.0	20
2	165	12	4.0	93
3	110	20	NaN	60
4	140	5	3.0	82
5	115	2	8.0	10

```
In [5]:  df.experien
```

```
Out[5]:  0     5.0
         1    13.0
         2     4.0
         3     NaN
         4     3.0
         5     8.0
         Name: experien, dtype: float64
```

```
In [15]:  exm_fit = df.experien.median()
```

```
In [16]:  exm_fit
```

Out[16]: 5.0

```
In [17]:  df.experien = df.experien.fillna(exm_fit)
```

```
In [18]:  df.experien
```

```
Out[18]:  0     5.0
          1    13.0
          2     4.0
          3     5.0
          4     3.0
          5     8.0
          Name: experien, dtype: float64
```

```
In [19]:  df
```

Out[19]:
	speed	car_age	experien	risk
0	200	15	5.0	85
1	90	17	13.0	20
2	165	12	4.0	93
3	110	20	5.0	60
4	140	5	3.0	82
5	115	2	8.0	10

```
In [20]:  reg = linear_model.LinearRegression()
```

```
In [22]:  reg.fit(df[['speed','car_age','experien']],df.risk)
```

Out[22]: LinearRegression()

```
In [23]:  reg.predict([[180,20,10]])
```

Out[23]: array([63.05003302])

```
In [24]:  reg.coef_
```

Out[24]: array([0.33059217, 1.61053246, -6.20772074])

```
In [25]:  reg.intercept_
```

Out[25]: 33.410000910435905

```
In [26]:  0.33059217*180 + 1.61053246*20 + 10*-6.20772074 + 33.410000910435905
```

Out[26]: 63.05003331043591

```
In [ ]:
```