```python
In [1]: import numpy as np
        import pandas as pd
        import seaborn as sns
        import matplotlib.pyplot as plt
        from sklearn import preprocessing,svm
        from sklearn.model_selection import train_test_split
        from sklearn.linear_model import LinearRegression
        import warnings
        warnings.simplefilter(action='ignore')
```

```python
In [2]: df=pd.read_csv(r"C:\Users\pappu\Downloads\used_cars_data.csv")
        df
```

Out[2]:

	S.No.	Name	Location	Year	Kilometers_Driven	Fuel_Type	Transmission	Owner_
0	0	Maruti Wagon R LXI CNG	Mumbai	2010	72000	CNG	Manual	
1	1	Hyundai Creta 1.6 CRDi SX Option	Pune	2015	41000	Diesel	Manual	
2	2	Honda Jazz V	Chennai	2011	46000	Petrol	Manual	
3	3	Maruti Ertiga VDI	Chennai	2012	87000	Diesel	Manual	
4	4	Audi A4 New 2.0 TDI Multitronic	Coimbatore	2013	40670	Diesel	Automatic	Se
...	...	...	...	...	...	...	...	
7248	7248	Volkswagen Vento Diesel Trendline	Hyderabad	2011	89411	Diesel	Manual	
7249	7249	Volkswagen Polo GT TSI	Mumbai	2015	59000	Petrol	Automatic	
7250	7250	Nissan Micra Diesel XV	Kolkata	2012	28000	Diesel	Manual	
7251	7251	Volkswagen Polo GT TSI	Pune	2013	52262	Petrol	Automatic	
7252	7252	Mercedes-Benz E-Class 2009-2013 E 220 CDI Avan...	Kochi	2014	72443	Diesel	Automatic	

7253 rows × 14 columns

In [3]: ```python
df=df[['Seats','Price']]
```

In [4]: ```python
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 7253 entries, 0 to 7252
Data columns (total 2 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   Seats   7200 non-null   float64
 1   Price   6019 non-null   float64
dtypes: float64(2)
memory usage: 113.5 KB
```

In [5]: ```python
df.dropna(inplace=True)
```

In [6]: ```python
df.isna().any()
```

Out[6]:
```
Seats    False
Price    False
dtype: bool
```

In [7]: ```python
df.head()
```

Out[7]:

|   | Seats | Price |
|---|-------|-------|
| 0 | 5.0   | 1.75  |
| 1 | 5.0   | 12.50 |
| 2 | 5.0   | 4.50  |
| 3 | 7.0   | 6.00  |
| 4 | 5.0   | 17.74 |

```
In [8]:  #sns.lmplot(x='kil',y='price',data=df)
         plt.scatter(df['Seats'],df['Price'])
```

Out[8]: <matplotlib.collections.PathCollection at 0x204d1db5fc0>


```
In [9]:  x=df[['Seats']]
         y=df['Price']
```

```
In [10]: x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.25)
```

```
In [11]: from sklearn.linear_model import LinearRegression
         lr=LinearRegression()
```

```
In [12]: lr.fit(x_train,y_train)
```

Out[12]: ▾ LinearRegression

          LinearRegression()

```
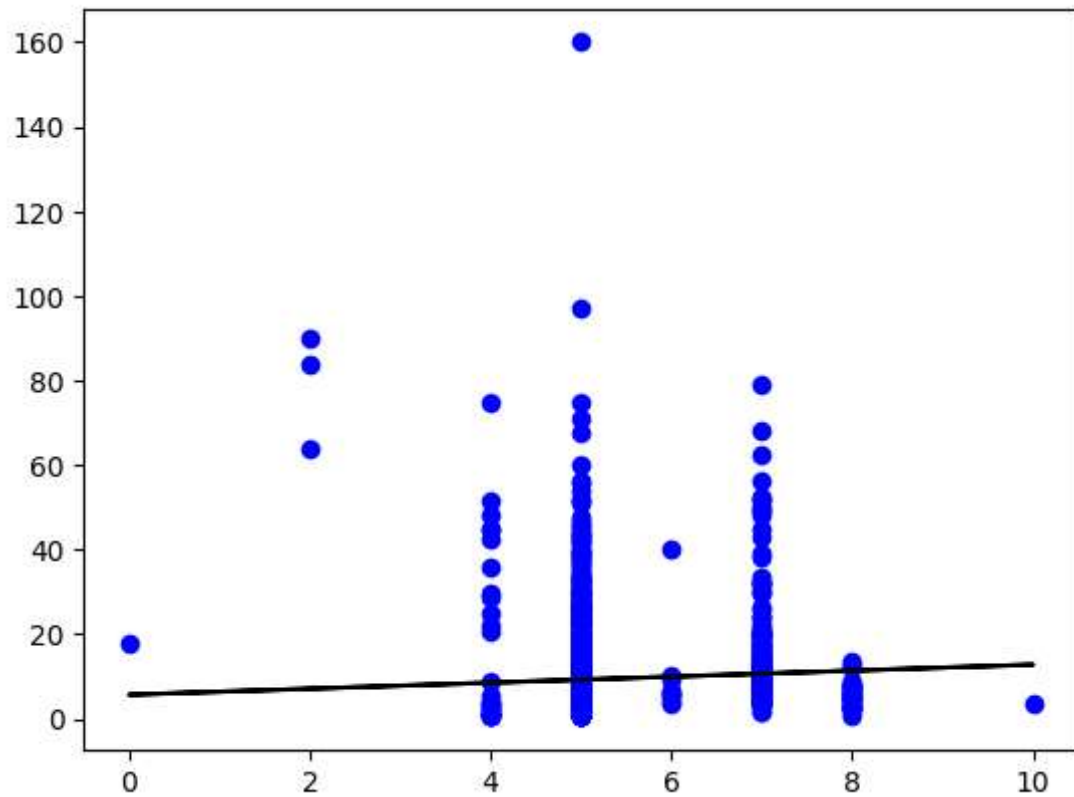In [13]: lr.predict(x_test)
```

Out[13]: array([9.23098781, 9.23098781, 9.23098781, ..., 9.23098781, 9.23098781,
                9.23098781])

In [14]:
```python
print(lr.score(x_test,y_test))
```
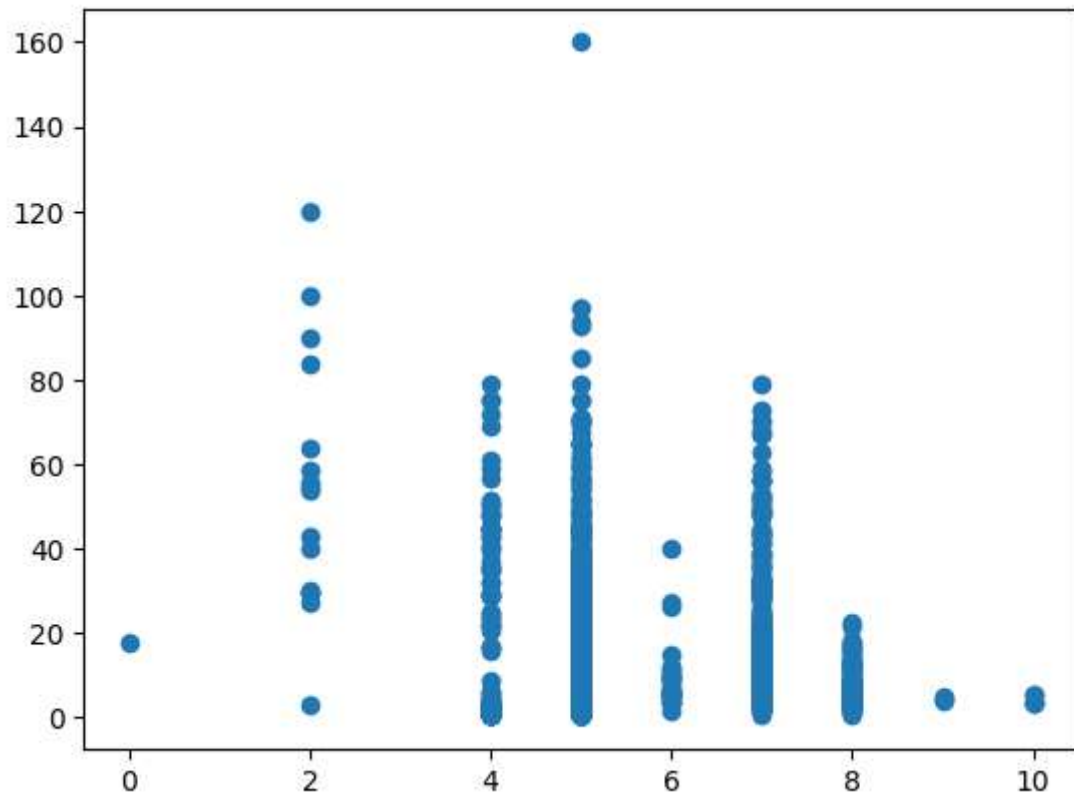
0.0020763086361255123

In [15]:
```python
y_pred=lr.predict(x_test)
plt.scatter(x_test,y_test,color='b')
plt.plot(x_test,y_pred,color='k')
plt.show()
```


In [ ]: