In [186]:
```
import pandas as pd
import matplotlib.pyplot as plt
from sklearn import linear_model
import numpy as np
```

In [233]:
```
#data load
df = pd.read_csv('room.csv')
df
```

Out[233]:

| | area | bedrom | age | price |
|---|---|---|---|---|
| **0** | 129 | 2 | 10 | NaN |
| **1** | 159 | 4 | 5 | 2500.0 |
| **2** | 200 | 3 | 20 | 3000.0 |
| **3** | 131 | 5 | 10 | 2300.0 |
| **4** | 300 | 2 | 15 | 4000.0 |
| **5** | 174 | 2 | 22 | 2900.0 |

In [234]:
```
df['price']= df['price'].fillna(df['price'].median())
df
```

Out[234]:

| | area | bedrom | age | price |
|---|---|---|---|---|
| **0** | 129 | 2 | 10 | 2900.0 |
| **1** | 159 | 4 | 5 | 2500.0 |
| **2** | 200 | 3 | 20 | 3000.0 |
| **3** | 131 | 5 | 10 | 2300.0 |
| **4** | 300 | 2 | 15 | 4000.0 |
| **5** | 174 | 2 | 22 | 2900.0 |

In [235]:
```
df['price'].median()
```

Out[235]: 2900.0

In [236]:
```
reg = linear_model.LinearRegression()
```

In [237]:
```
reg.fit(df.drop('price', axis= 'columns'),df['price'])
```

Out[237]:
```
▾ LinearRegression
LinearRegression()
```

In [238]:
```
reg.fit(df.drop('age', axis= 'columns'),df['age'])
```

Out[238]:
```
LinearRegression
LinearRegression()
```

In [239]:
```
reg.predict([[129,2,2000]])
```

```
C:\Users\Administrator\AppData\Local\Programs\Python\Python311\Lib\site-packa
ges\sklearn\base.py:465: UserWarning: X does not have valid feature names, bu
t LinearRegression was fitted with feature names
  warnings.warn(
```

Out[239]:
```
array([31.86907377])
```

In [240]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6 entries, 0 to 5
Data columns (total 4 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   area    6 non-null      int64
 1   bedrom  6 non-null      int64
 2   age     6 non-null      int64
 3   price   6 non-null      float64
dtypes: float64(1), int64(3)
memory usage: 324.0 bytes
```

In [241]:
```
df.describe()
```

Out[241]:

| | area | bedrom | age | price |
|---|---|---|---|---|
| **count** | 6.000000 | 6.000000 | 6.000000 | 6.000000 |
| **mean** | 182.166667 | 3.000000 | 13.666667 | 2933.333333 |
| **std** | 63.640920 | 1.264911 | 6.531973 | 588.784058 |
| **min** | 129.000000 | 2.000000 | 5.000000 | 2300.000000 |
| **25%** | 138.000000 | 2.000000 | 10.000000 | 2600.000000 |
| **50%** | 166.500000 | 2.500000 | 12.500000 | 2900.000000 |
| **75%** | 193.500000 | 3.750000 | 18.750000 | 2975.000000 |
| **max** | 300.000000 | 5.000000 | 22.000000 | 4000.000000 |

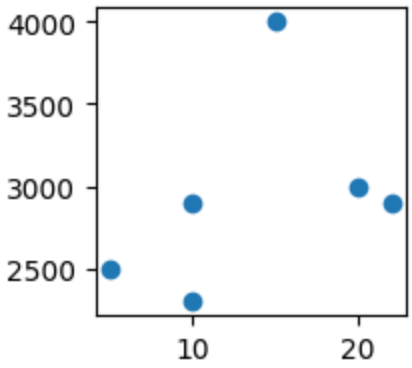
In [249]:
```
reg.fit(df[['age']], df['price'])
```

Out[249]:
```
LinearRegression
LinearRegression()
```

In [248]:
```
# ntercept (c)= y when x=0
reg.intercept_
```

Out[248]: 2435.78125

In [247]:
```
plt.figure(figsize=(2,2))
plt.scatter(df['age'], df['price'])
plt.show()
```




In [245]:
```
# Model Building
size= df.drop('price', axis=1)
size
```

Out[245]:

| | area | bedrom | age |
|---|---|---|---|
| **0** | 129 | 2 | 10 |
| **1** | 159 | 4 | 5 |
| **2** | 200 | 3 | 20 |
| **3** | 131 | 5 | 10 |
| **4** | 300 | 2 | 15 |
| **5** | 174 | 2 | 22 |

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: