

```
import numpy as np
import pandas as pd
import sklearn
import matplotlib.pyplot as plt
import matplotlib.ticker as mtick
import seaborn as sns
```

▼ New Section

```
df = pd.read_csv('energy_efficiency_data.csv')
df.head()
```

	Relative_Compactness	Surface_Area	Wall_Area	Roof_Area	Overall_Height	Orientation
0	0.98	514.5	294.0	110.25	7.0	
1	0.98	514.5	294.0	110.25	7.0	
2	0.98	514.5	294.0	110.25	7.0	
3	0.98	514.5	294.0	110.25	7.0	
4	0.90	563.5	318.5	122.50	7.0	



```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 768 entries, 0 to 767
Data columns (total 10 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Relative_Compactness                  768 non-null    float64
1   Surface_Area                         768 non-null    float64
2   Wall_Area                           768 non-null    float64
3   Roof_Area                           768 non-null    float64
4   Overall_Height                       768 non-null    float64
5   Orientation                          768 non-null    int64
6   Glazing_Area                        768 non-null    float64
7   Glazing_Area_Distribution            768 non-null    int64
8   Heating_Load                        768 non-null    float64
9   Cooling_Load                        768 non-null    float64
dtypes: float64(8), int64(2)
memory usage: 60.1 KB
```

```
df.describe()
```

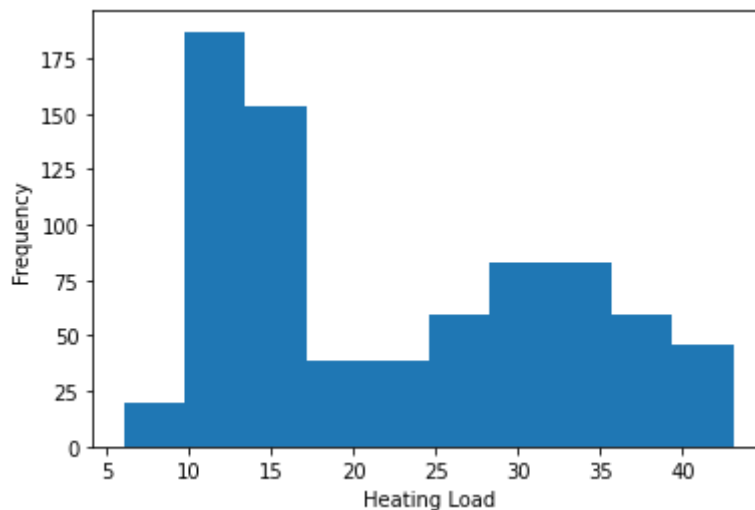
	Relative_Compactness	Surface_Area	Wall_Area	Roof_Area	Overall_Height	Or
count	768.000000	768.000000	768.000000	768.000000	768.000000	
mean	0.764167	671.708333	318.500000	176.604167	5.250000	
std	0.105777	88.086116	43.626481	45.165950	1.75114	
min	0.620000	514.500000	245.000000	110.250000	3.500000	
25%	0.682500	606.375000	294.000000	140.875000	3.500000	
50%	0.750000	673.750000	318.500000	183.750000	5.250000	
75%	0.830000	741.125000	343.000000	220.500000	7.000000	
max	0.980000	808.500000	416.500000	220.500000	7.000000	



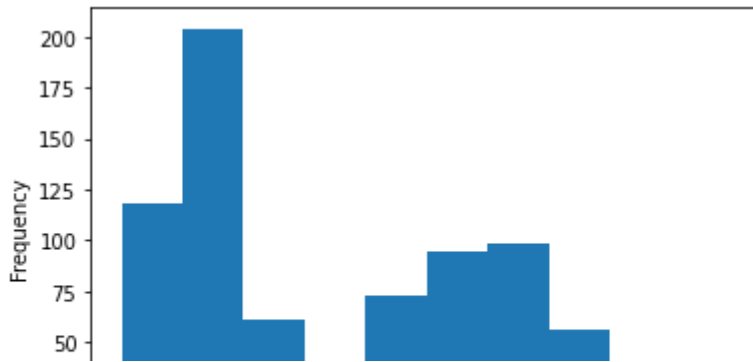
```
df.columns
```

```
Index(['Relative_Compactness', 'Surface_Area', 'Wall_Area', 'Roof_Area',
      'Overall_Height', 'Orientation', 'Glazing_Area',
      'Glazing_Area_Distribution', 'Heating_Load', 'Cooling_Load'],
      dtype='object')
```

```
plt.hist(df["Heating_Load"])
plt.xlabel("Heating Load")
plt.ylabel('Frequency')
plt.show()
```



```
plt.hist(df["Cooling_Load"])
plt.xlabel("Cooling Load")
plt.ylabel('Frequency')
plt.show()
```

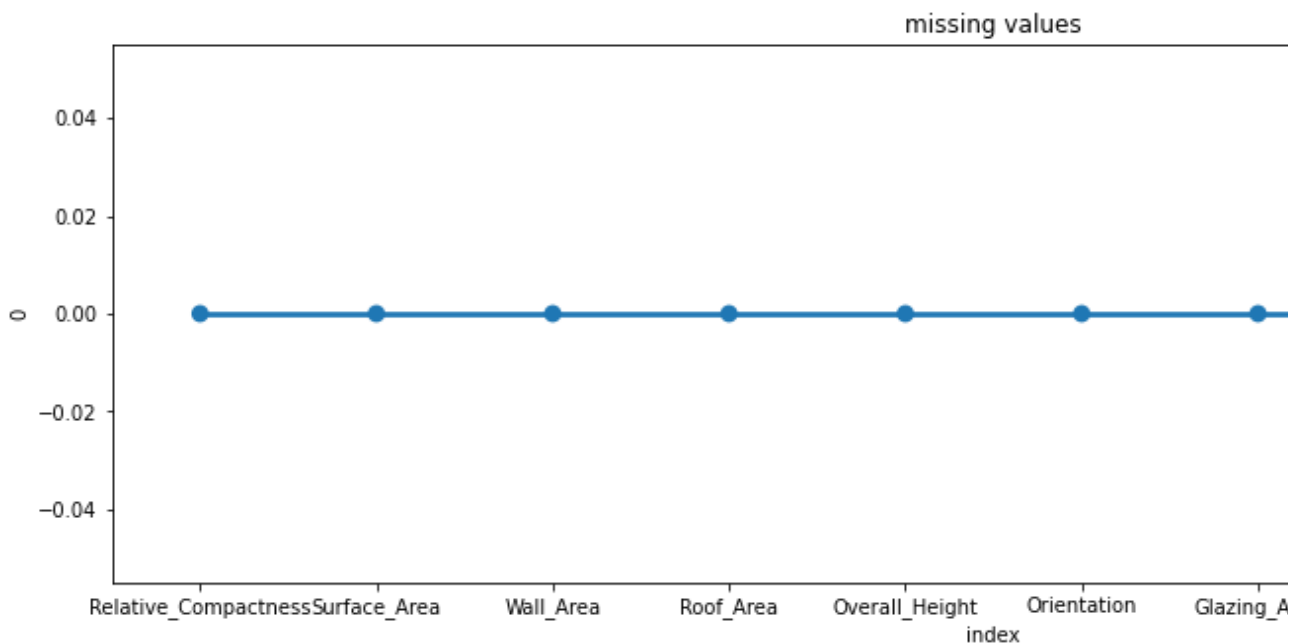


```
df.isnull().sum()
```

```
Relative_Compactness    0
Surface_Area            0
Wall_Area               0
Roof_Area              0
Overall_Height         0
Orientation             0
Glazing_Area           0
Glazing_Area_Distribution 0
Heating_Load           0
Cooling_Load           0
dtype: int64
```

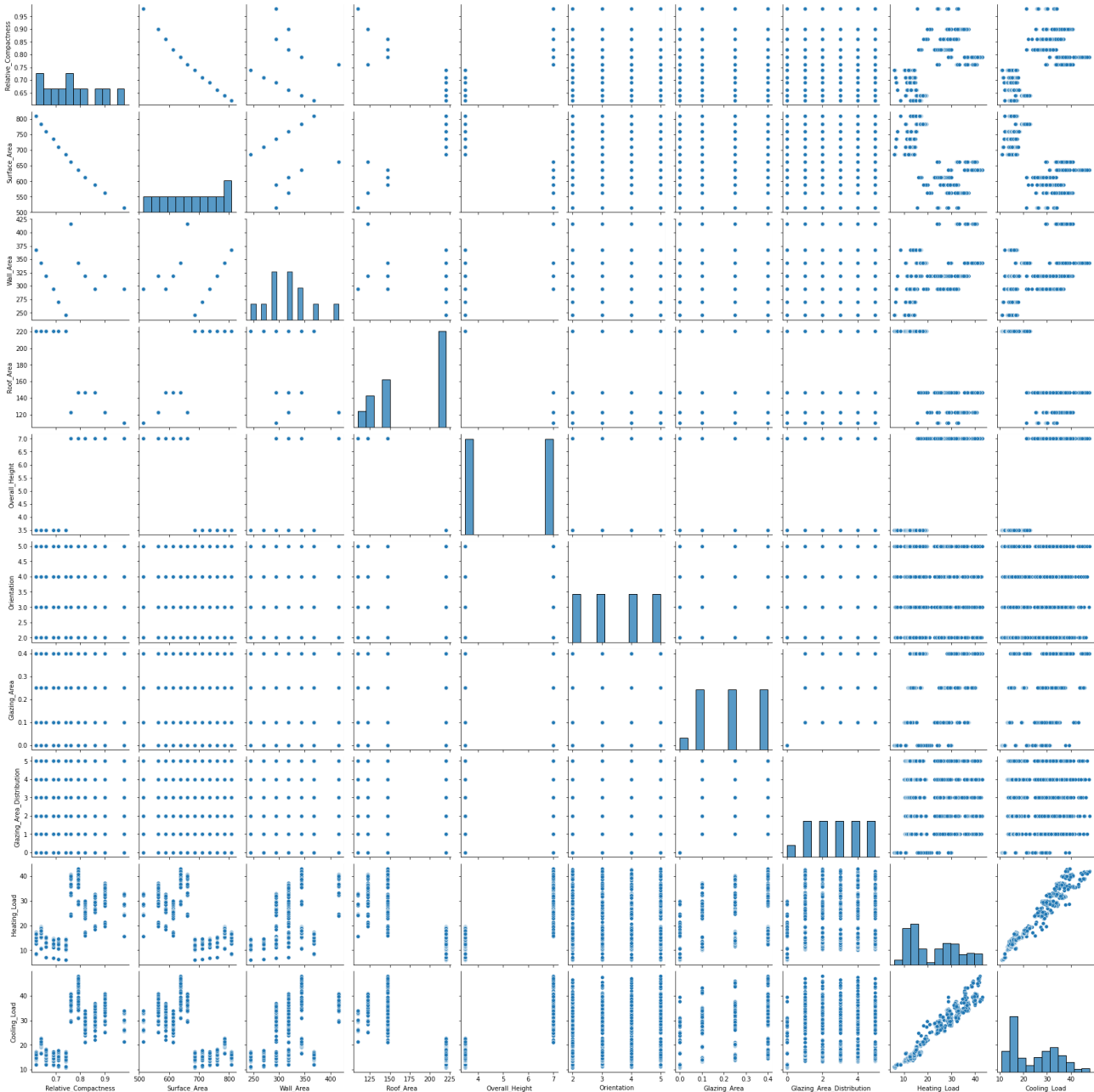
```
missing = pd.DataFrame((df.isnull().sum()*100/df.shape[0]).reset_index())
plt.figure(figsize=(16,5))
ax = sns.pointplot('index', 0 , data=missing)
plt.title("missing values")
plt.show()
```

```
/usr/local/lib/python3.7/dist-packages/seaborn/_decorators.py:43: FutureWarning: Pas
FutureWarning
```



```
sns.pairplot(df)
```

<seaborn.axisgrid.PairGrid at 0x7f3c722993d0>



✓ 46s completed at 9:33 AM ● ✕