In [2]:
```
import pandas as pd
import numpy as np
from sklearn import linear_model
import matplotlib.pyplot as plt
import seaborn as sns
```

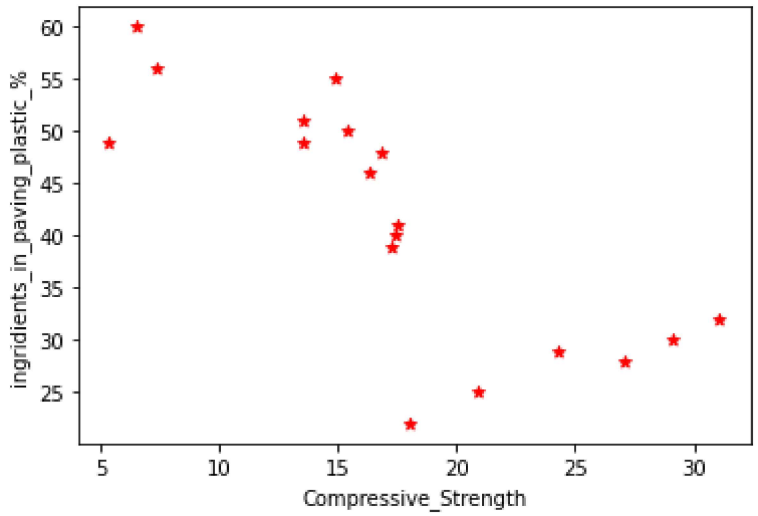
In [3]:
```
df = pd.read_csv('observation plastic.csv')
df
```

Out[3]:

	sr_no	Identification_no	Type_ of_mould	area _of _mould	ingridients_in_paving_sand_In_KG	ingridients_
0	NaN	R1	rectangular	29900.00	0.695	
1	NaN	B1	hexagon	34359.55	1.535	
2	NaN	B2	hexagon	34359.55	2.300	
3	NaN	B3	hexagon	34359.55	2.300	
4	NaN	B4	hexagon	34359.55	2.300	
5	NaN	R2	rectangular	29900.00	1.230	
6	NaN	R3	rectangular	29900.00	0.600	
7	NaN	R4	rectangular	29900.00	2.040	
8	NaN	B5	hexagon	34359.55	1.280	
9	NaN	B6	hexagon	34359.55	1.950	
10	NaN	B6	hexagon	34359.55	2.260	
11	NaN	B7	hexagon	34359.55	2.450	
12	NaN	B8	hexagon	34359.55	1.430	
13	NaN	B9	hexagon	34359.55	2.480	
14	NaN	R6	rectangular	29900.00	1.760	
15	NaN	R7	rectangular	29900.00	0.532	
16	NaN	B10	hexagon	34359.55	0.714	
17	NaN	B11	hexagon	34359.55	1.950	

In [8]:
```
%matplotlib inline
plt.xlabel('Compressive_Strength')
plt.ylabel('ingridients_in_paving_plastic_%')
plt.scatter(df.Compressive_Strength, df.ingridients_in_paving_plastic_, color=
```

Out[8]:
```
<matplotlib.collections.PathCollection at 0x221dbb159a0>
```



In [5]:
```
print(df.columns)
```

```
Index(['sr_no', 'Identification_no', 'Type_ of_mould ', 'area _of _mould',
       'ingridients_in_paving_sand_In_KG', 'ingridients_in_paving_plastic_%',
       'Load', 'Weight ', 'Compressive_Strength'],
      dtype='object')
```

In [6]:
```
# update the column name 'B' to 'C'
df = df.rename(columns={'ingridients_in_paving_plastic_%': 'ingridients_in_pav

# display the updated column names
print(df.columns)
```

```
Index(['sr_no', 'Identification_no', 'Type_ of_mould ', 'area _of _mould',
       'ingridients_in_paving_sand_In_KG', 'ingridients_in_paving_plastic_',
       'Load', 'Weight ', 'Compressive_Strength'],
      dtype='object')
```

In [7]:
```
df
```

Out[7]:

	sr_no	Identification_no	Type_ of_mould	area _of _mould	ingridients_in_paving_sand_In_KG	ingridients_
0	NaN	R1	rectangular	29900.00	0.695	
1	NaN	B1	hexagon	34359.55	1.535	
2	NaN	B2	hexagon	34359.55	2.300	
3	NaN	B3	hexagon	34359.55	2.300	
4	NaN	B4	hexagon	34359.55	2.300	
5	NaN	R2	rectangular	29900.00	1.230	
6	NaN	R3	rectangular	29900.00	0.600	
7	NaN	R4	rectangular	29900.00	2.040	
8	NaN	B5	hexagon	34359.55	1.280	
9	NaN	B6	hexagon	34359.55	1.950	
10	NaN	B6	hexagon	34359.55	2.260	
11	NaN	B7	hexagon	34359.55	2.450	
12	NaN	B8	hexagon	34359.55	1.430	
13	NaN	B9	hexagon	34359.55	2.480	
14	NaN	R6	rectangular	29900.00	1.760	
15	NaN	R7	rectangular	29900.00	0.532	
16	NaN	B10	hexagon	34359.55	0.714	
17	NaN	B11	hexagon	34359.55	1.950	

In [9]:
```
new_df = df[['Compressive_Strength', 'ingridients_in_paving_plastic_']]
```

In [10]:
```
new_df
```

Out[10]:

	Compressive_Strength	ingridients_in_paving_plastic_
0	15.410	50
1	17.462	40
2	29.070	30
3	14.930	55
4	20.890	25
5	6.534	60
6	13.528	51
7	17.290	39
8	16.862	48
9	18.062	22
10	27.050	28
11	31.090	32
12	17.532	41
13	24.248	29
14	5.350	49
15	7.358	56
16	13.560	49
17	16.300	46

In [11]:
```
n_df = new_df.drop('ingridients_in_paving_plastic_',axis='columns')
n_df
```

Out[11]:

	Compressive_Strength
0	15.410
1	17.462
2	29.070
3	14.930
4	20.890
5	6.534
6	13.528
7	17.290
8	16.862
9	18.062
10	27.050
11	31.090
12	17.532
13	24.248
14	5.350
15	7.358
16	13.560
17	16.300

In [12]:
```
ingridients_in_paving_plastic_ = df.ingridients_in_paving_plastic_
ingridients_in_paving_plastic_
```

Out[12]:
```
0     50
1     40
2     30
3     55
4     25
5     60
6     51
7     39
8     48
9     22
10    28
11    32
12    41
13    29
14    49
15    56
16    49
17    46
Name: ingridients_in_paving_plastic_, dtype: int64
```

In [13]:
```
# Create linear regression object
reg = linear_model.LinearRegression()
reg.fit(n_df,ingridients_in_paving_plastic_)
```

Out[13]:
```
LinearRegression
LinearRegression()
```

In [25]:
```
Compressive_Strength=reg.predict([[20]])
Compressive_Strength
```

```
C:\Users\ASUS\anaconda3\lib\site-packages\sklearn\base.py:420: UserWarning: X
does not have valid feature names, but LinearRegression was fitted with featu
re names
  warnings.warn(
```

Out[25]:
```
array([38.34108486])
```

In [83]:
```
reg.coef_
```

Out[83]:
```
array([-1.26091066])
```

In [84]:
```
reg.intercept_
```

Out[84]:
```
63.55929799792155
```

In [85]:
```
50*-1.26091066+63.55929799792155
```

Out[85]:
```
0.5137649979215482
```

In [31]:
```
sns.pairplot(data = df)
```

Out[31]:
```
<seaborn.axisgrid.PairGrid at 0x221dd371e50>
```