In [1]:

```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [2]:

```python
# Read the data
df=pd.read_csv("C:/practice datasets/ManufactCosts.csv")
df.head(5)
```

Out[2]:

	Unnamed: 0	cost	capitalcost	laborcost	energycost	materialscost	capitalprice	laborprice
0	1	182.373	0.05107	0.24727	0.04253	0.65913	1.00000	1.00000
1	2	183.161	0.05817	0.27716	0.05127	0.61340	1.00270	1.15457
2	3	186.533	0.04602	0.25911	0.05075	0.64411	0.74371	1.15584
3	4	221.710	0.04991	0.24794	0.04606	0.65609	0.92497	1.23535
4	5	255.945	0.05039	0.25487	0.04482	0.64992	1.04877	1.33784

Check the Skewness and Kurtosis of each features

In [6]:

```python
print(df.skew(axis=0))
```

```
Unnamed: 0       0.000000
cost             0.515839
capitalcost     -0.108564
laborcost       -0.534577
energycost       0.130611
materialscost    0.882168
capitalprice    -0.237044
laborprice       0.395719
energyprice     -0.637070
materialsprice  -0.217658
dtype: float64
```

In [5]:

```
print(df.kurtosis(axis=0))
```

```
Unnamed: 0       -1.200000
cost             -0.728303
capitalcost      -0.815853
laborcost        -0.036623
energycost       -0.112566
materialscost    -0.272922
capitalprice     -0.622274
laborprice       -0.510896
energyprice       2.957819
materialsprice   -0.262818
dtype: float64
```

In [7]:

```
df.isna().sum()                    #check for missing values
```

Out[7]:

```
Unnamed: 0        0
cost              0
capitalcost       0
laborcost         0
energycost        0
materialscost     0
capitalprice      0
laborprice        0
energyprice       0
materialsprice    0
dtype: int64
```
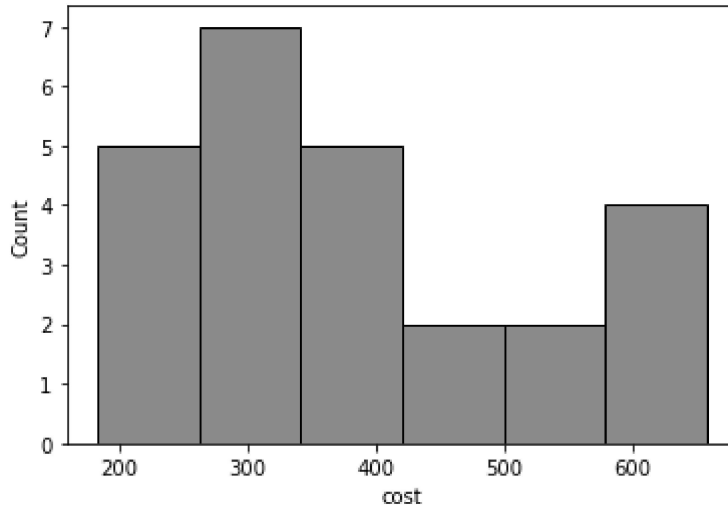
Data Visualization

In [8]:

```python
# Histogram
sns.histplot(df['cost'])
```

Out[8]:

```
<AxesSubplot:xlabel='cost', ylabel='Count'>
```


In [10]:

```python
# Dencity plot for single column
sns.distplot(df['laborcost']) #Dencity plot for single column
```

```
C:\Users\HP\anaconda3\lib\site-packages\seaborn\distributions.py:2551: Futur
eWarning: `distplot` is a deprecated function and will be removed in a futur
e version. Please adapt your code to use either `displot` (a figure-level fu
nction with similar flexibility) or `histplot` (an axes-level function for h
istograms).
  warnings.warn(msg, FutureWarning)
```
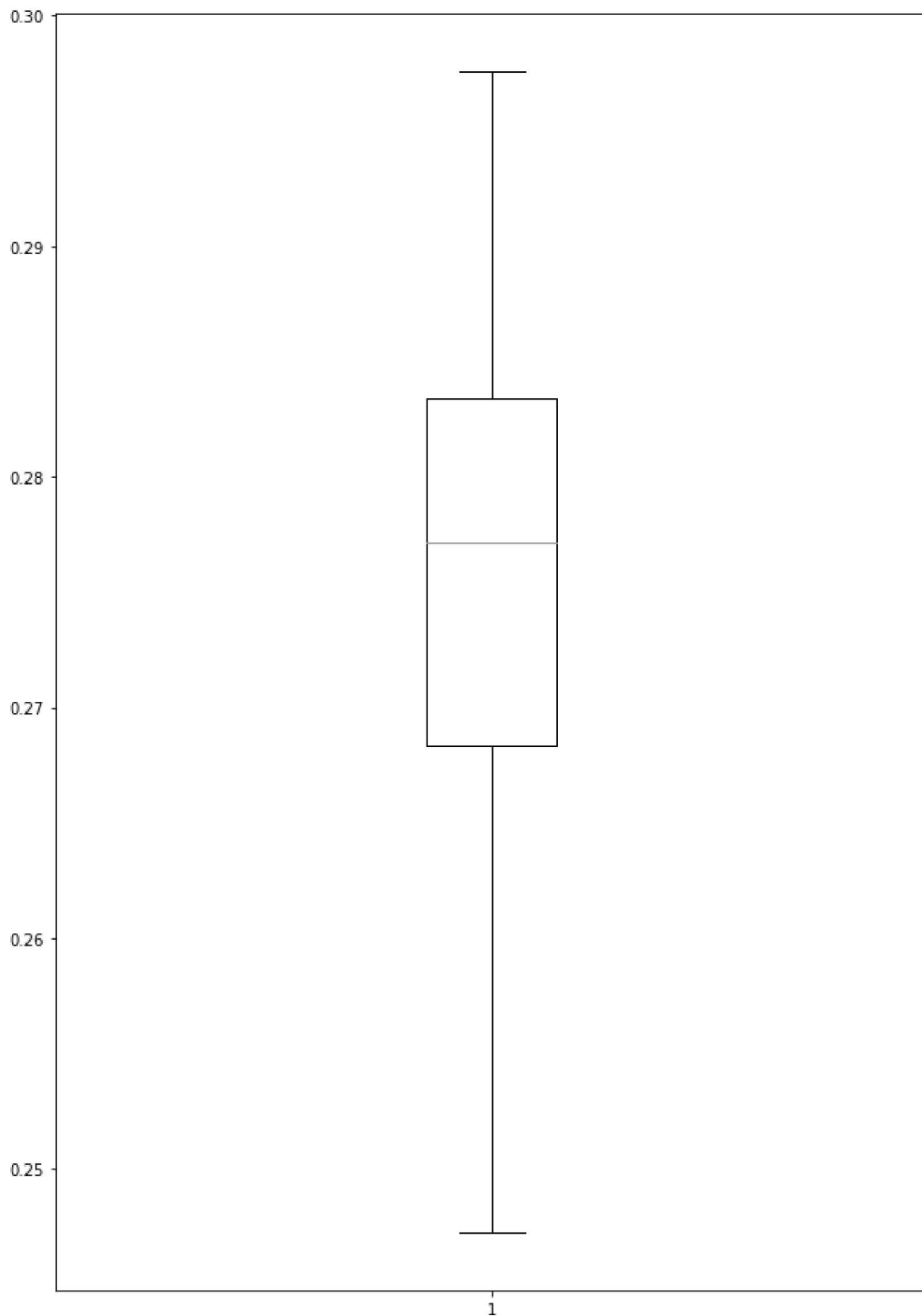
Out[10]:

```
<AxesSubplot:xlabel='laborcost', ylabel='Density'>
```
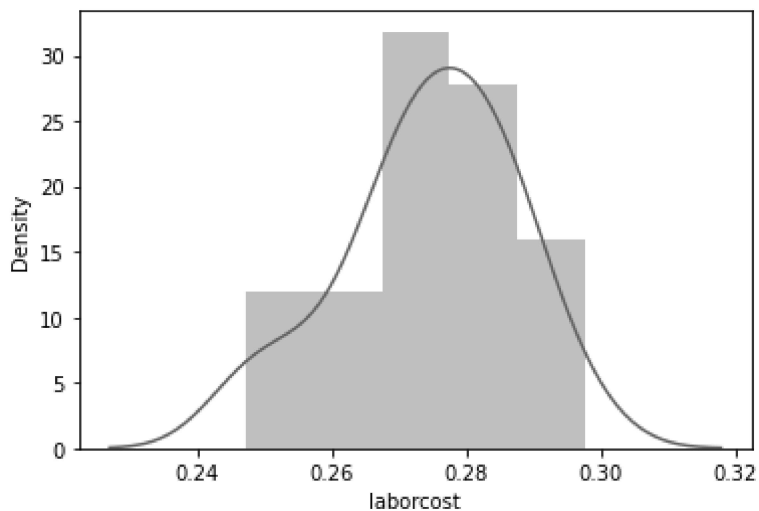
In [16]:

```python
# Boxplot
fig=plt.figure(figsize=(10,15))
plt.boxplot(df['laborcost'])
plt.show()
```

In [17]:

```
df.columns
```
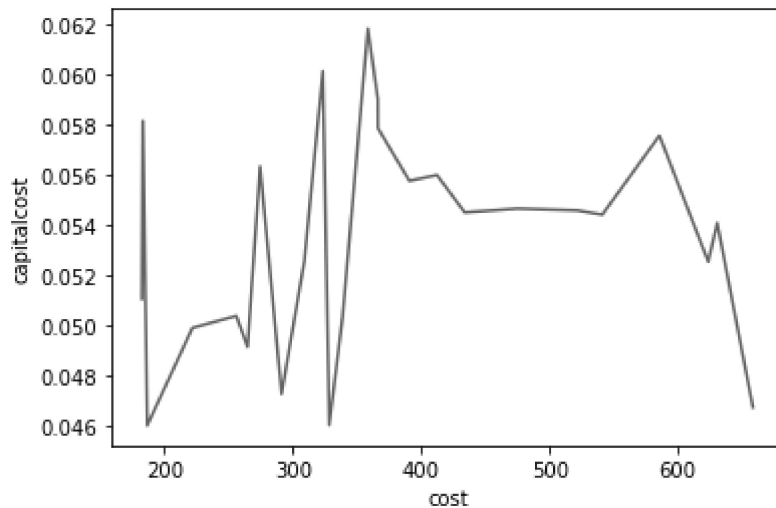
Out[17]:

```
Index(['Unnamed: 0', 'cost', 'capitalcost', 'laborcost', 'energycost',
       'materialscost', 'capitalprice', 'laborprice', 'energyprice',
       'materialsprice'],
      dtype='object')
```

In [18]:

```
# Line plot
sns.lineplot(x='cost',y='capitalcost',data=df)
```
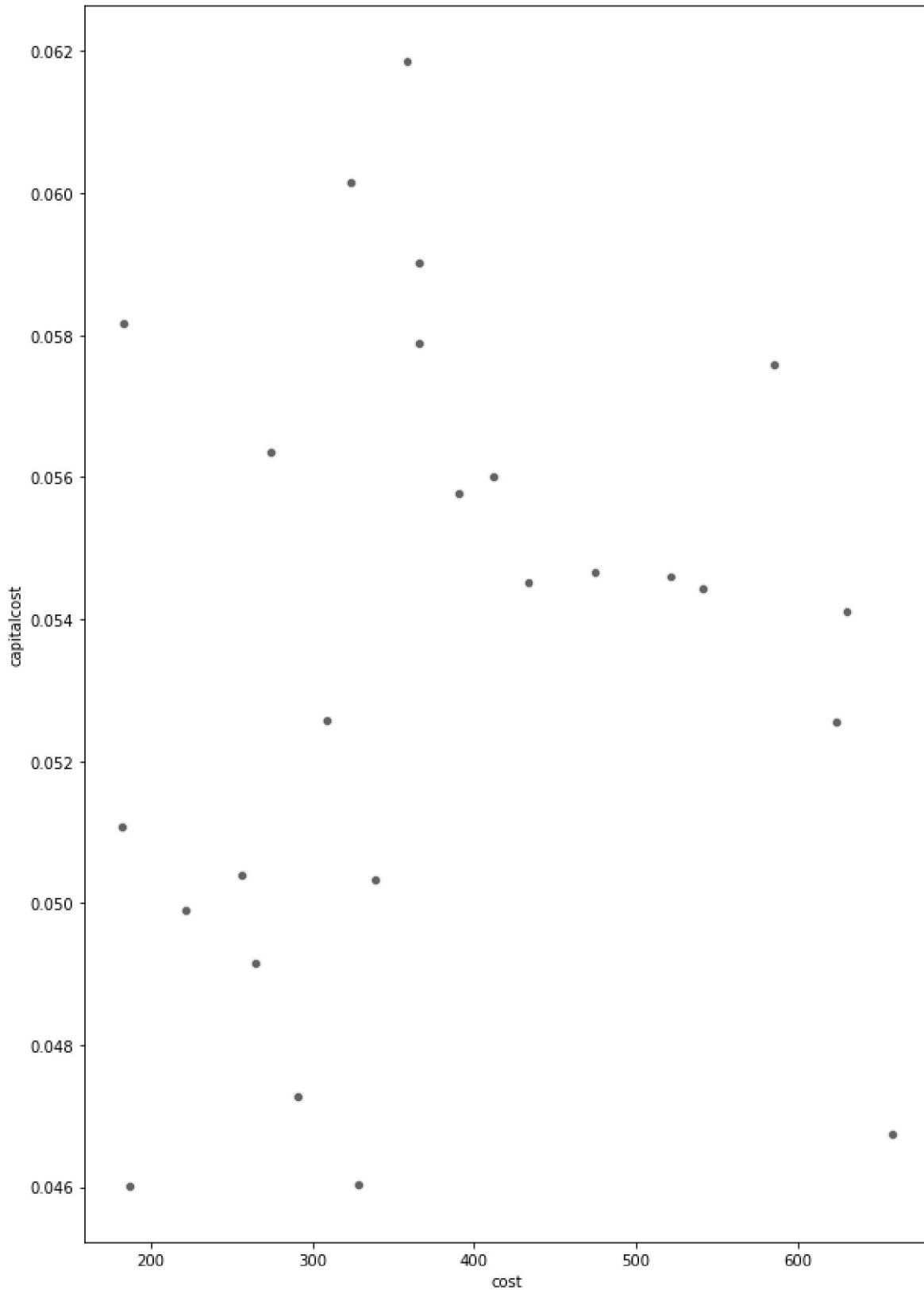
Out[18]:

```
<AxesSubplot:xlabel='cost', ylabel='capitalcost'>
```

In [20]:

```
# Scatterplot
fig=plt.figure(figsize=(10,15))
sns.scatterplot(x='cost',y='capitalcost',data=df)
```
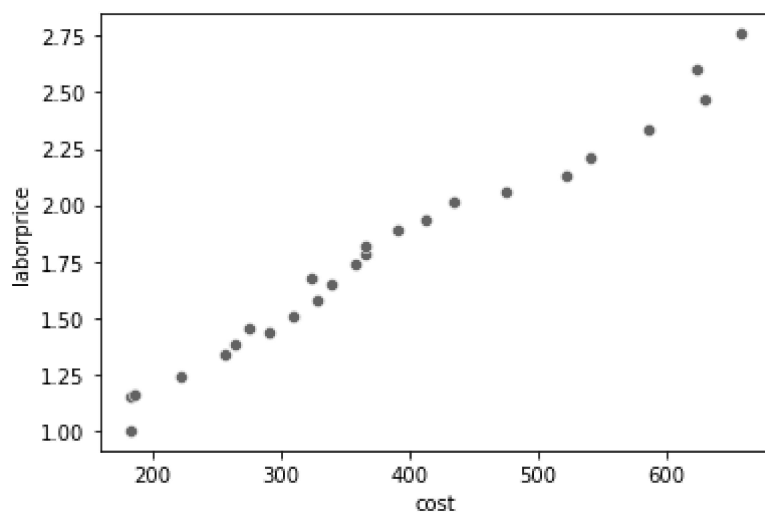
Out[20]:

```
<AxesSubplot:xlabel='cost', ylabel='capitalcost'>
```

In [22]:

```
sns.scatterplot(x='cost',y='laborprice',data=df)
```
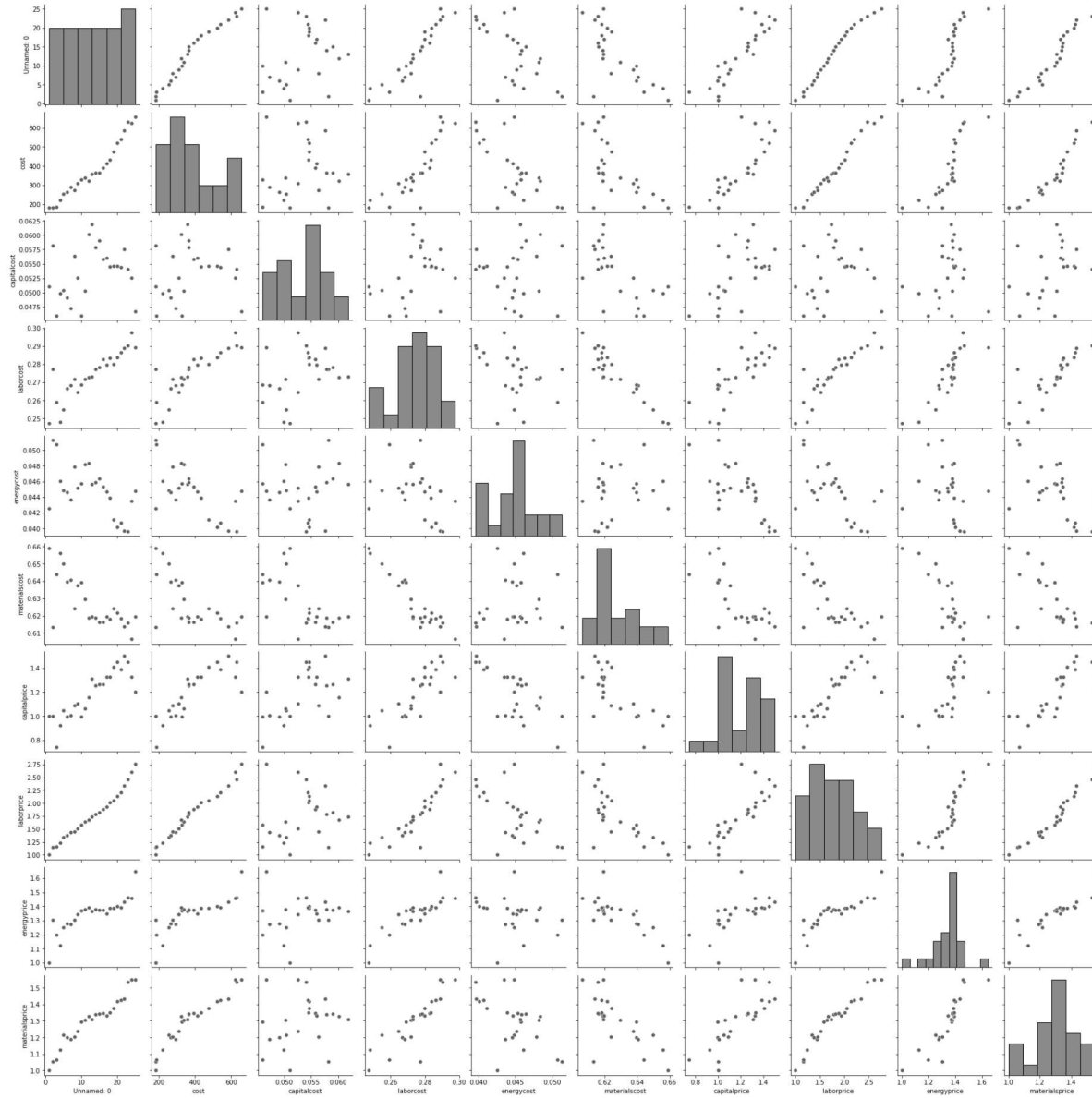
Out[22]:

```
<AxesSubplot:xlabel='cost', ylabel='laborprice'>
```

In [21]:

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

Out[21]:

```
<seaborn.axisgrid.PairGrid at 0x2cbb8bdc2e0>
```

In [ ]: