In [39]:
```
import pandas as pd
import numpy as np
import seaborn as sns
```

In [2]:
```
datase=pd.read_csv("C:\\Users\\Admin\\Downloads\\assignment 3\\delivery_time.csv")
dataset
```

Out[2]:

	Delivery Time	Sorting Time
0	21.00	10
1	13.50	4
2	19.75	6
3	24.00	9
4	29.00	10
5	15.35	6
6	19.00	7
7	9.50	3
8	17.90	10
9	18.75	9
10	19.83	8
11	10.75	4
12	16.68	7
13	11.50	3
14	12.03	3
15	14.88	4
16	13.75	6
17	18.11	7
18	8.00	2
19	17.83	7
20	21.50	5

In [4]:
```
dataset.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 21 entries, 0 to 20
Data columns (total 2 columns):
 #   Column         Non-Null Count  Dtype  
---  ------         --------------  -----  
 0   Delivery Time  21 non-null     float64
 1   Sorting Time   21 non-null     int64  
```

```
dtypes: float64(1), int64(1)
memory usage: 464.0 bytes
```

In [41]:
```
dataset.describe()
```

Out[41]:

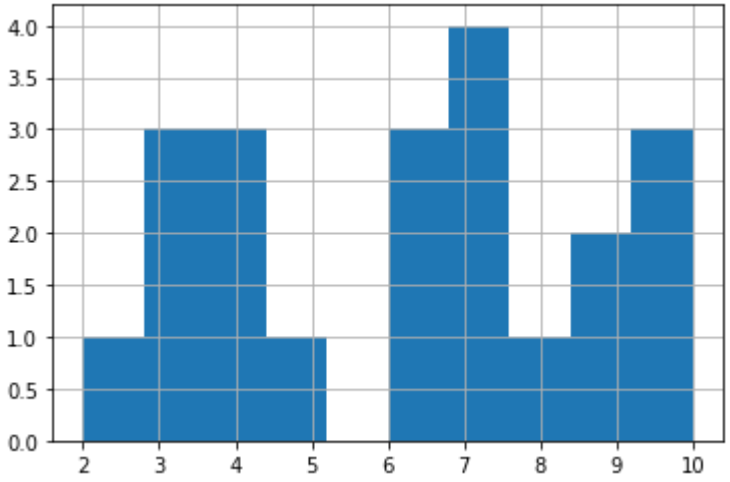
	Delivery Time	Sorting Time
count	21.000000	21.000000
mean	16.790952	6.190476
std	5.074901	2.542028
min	8.000000	2.000000
25%	13.500000	4.000000
50%	17.830000	6.000000
75%	19.750000	8.000000
max	29.000000	10.000000

In [43]:
```
dataset[dataset.duplicated()].shape
```

Out[43]:
```
(0, 2)
```

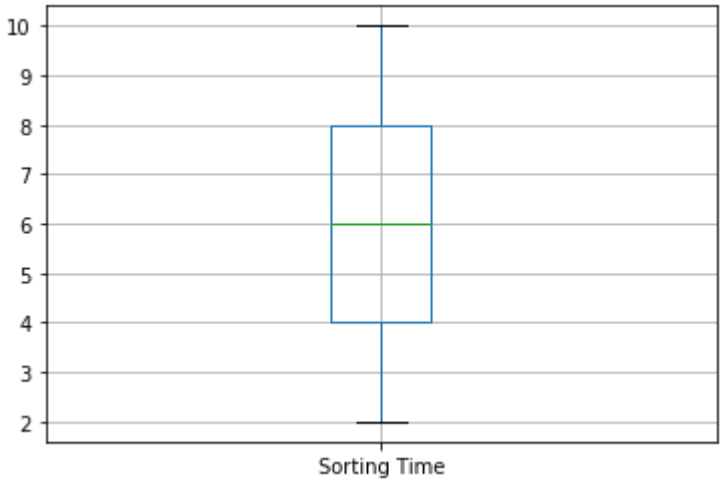
In [44]:
```
dataset['Sorting Time'].hist()
```

Out[44]:
```
<AxesSubplot:>
```



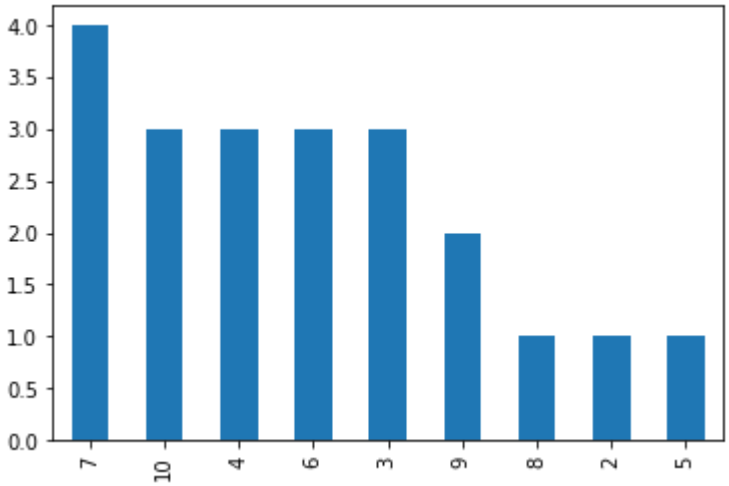
In [45]:
```
dataset.boxplot(column=['Sorting Time'])
```

Out[45]:
```
<AxesSubplot:>
```



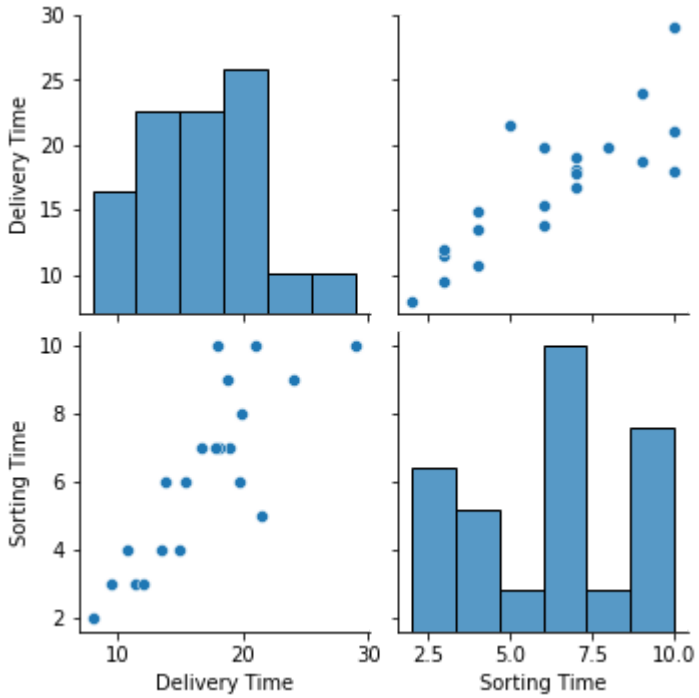
In [47]:
```
dataset['Sorting Time'].value_counts().plot.bar()
```

Out[47]:
```
<AxesSubplot:>
```



In [48]:
```
import seaborn as sns
sns.pairplot(dataset)
```

Out[48]:
```
<seaborn.axisgrid.PairGrid at 0x2ae0e3c18e0>
```



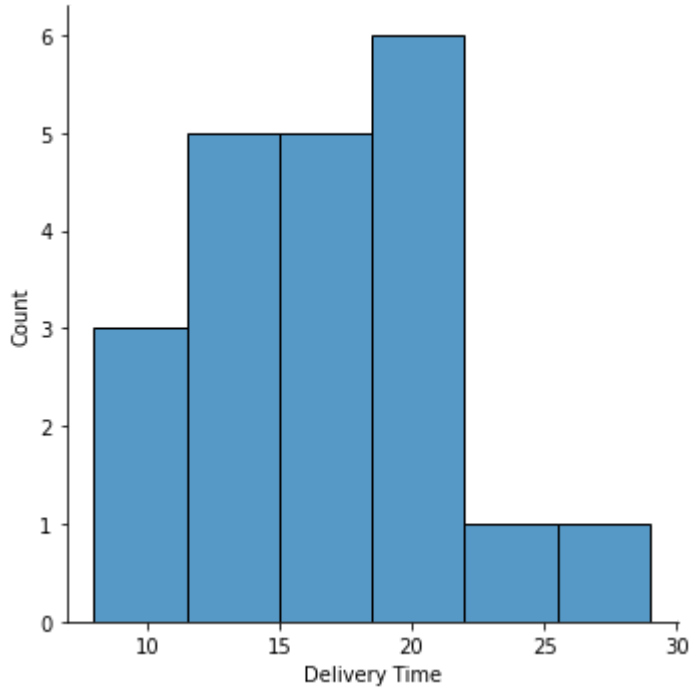
In [49]:
```
dataset.corr()
```

Out[49]:

	Delivery Time	Sorting Time
Delivery Time	1.000000	0.825997
Sorting Time	0.825997	1.000000

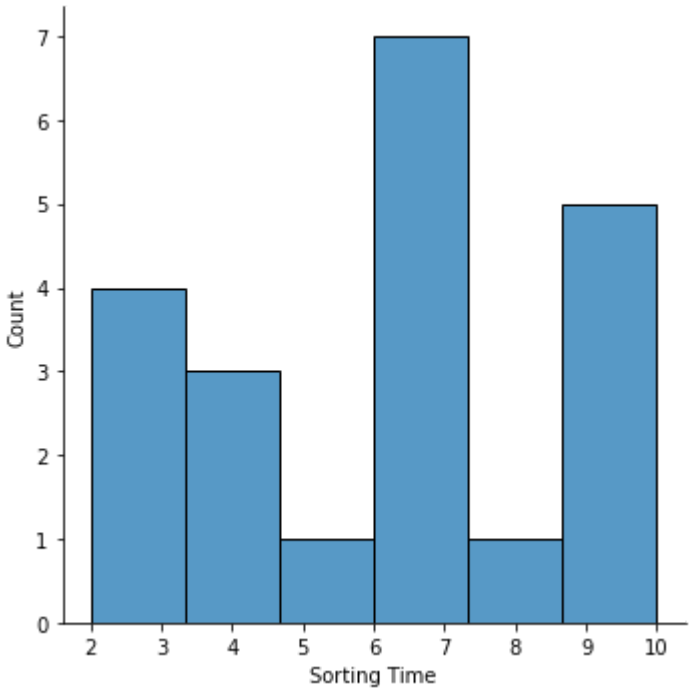
In [50]:
```
sns.displot(dataset['Delivery Time'])
```

Out[50]:
```
<seaborn.axisgrid.FacetGrid at 0x2ae0e36ff40>
```



In [51]:
```
sns.displot(dataset['Sorting Time'])
```

Out[51]:
```
<seaborn.axisgrid.FacetGrid at 0x2ae0e0df160>
```



In [55]:
```
dataset1= dataset1.rename(columns={'Delivery Time': 'DT', 'Sorting Time':'ST'})
```

In [56]:
```
dataset1
```

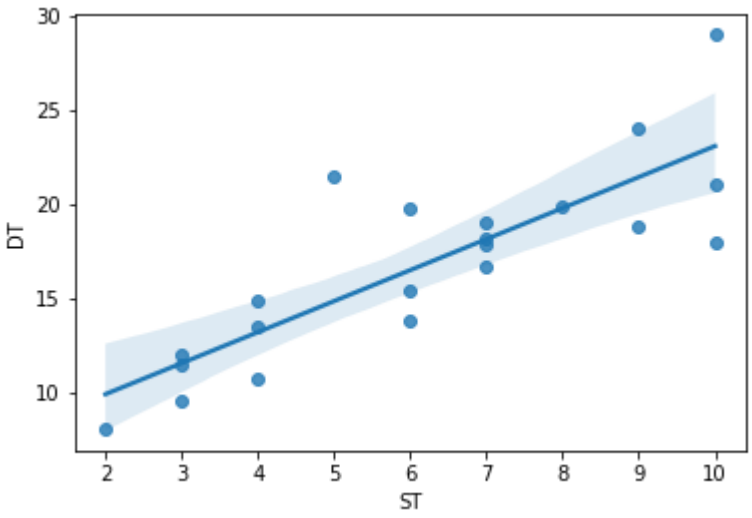
Out[56]:

	DT	ST
0	21.00	10
1	13.50	4
2	19.75	6
3	24.00	9
4	29.00	10
5	15.35	6
6	19.00	7
7	9.50	3
8	17.90	10
9	18.75	9
10	19.83	8
11	10.75	4
12	16.68	7
13	11.50	3

	DT	ST
14	12.03	3
15	14.88	4
16	13.75	6
17	18.11	7
18	8.00	2
19	17.83	7
20	21.50	5

In [63]:
```
import statsmodels.formula.api as smf
model=smf.ols("DT~ST",data=dataset1).fit()
```

In [64]:
```
sns.regplot(x="ST",y="DT",data=dataset1);
```



In [65]:
```
model.params
```

Out[65]:
```
Intercept    6.582734
ST           1.649020
dtype: float64
```

In [66]:
```
print(model.tvalues, '\n',model.pvalues)
```

```
Intercept    3.823349
ST           6.387447
dtype: float64
 Intercept    0.001147
ST           0.000004
dtype: float64
```

In [69]:
```
(model.rsquared,model.rsquared_adj)
```

```
(0.6822714748417231, 0.6655489208860244)
```

Out[69]:

In [ ]: