In [32]:

```python
import pandas as pd
import matplotlib.pyplot as plt
from numpy.polynomial.polynomial import polyfit
import statsmodels.api as sm
import numpy as np
```

executed in 17ms, finished 08:56:33 2021-11-26

In [13]:

```python
data = pd.read_csv("delivery_time.csv")
```

executed in 12ms, finished 08:45:44 2021-11-26

In [14]:

```python
data.head()
```

executed in 37ms, finished 08:45:46 2021-11-26

Out[14]:

	Delivery Time	Sorting Time
0	21.00	10
1	13.50	4
2	19.75	6
3	24.00	9
4	29.00	10

In [15]:

```python
data.info()
```

executed in 41ms, finished 08:45:49 2021-11-26

```
<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 [6]:

```python
data.corr()
```

executed in 38ms, finished 08:39:36 2021-11-26

Out[6]:

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

In [7]:

```python
import seaborn as sns
sns.distplot(data['Delivery Time'])
```
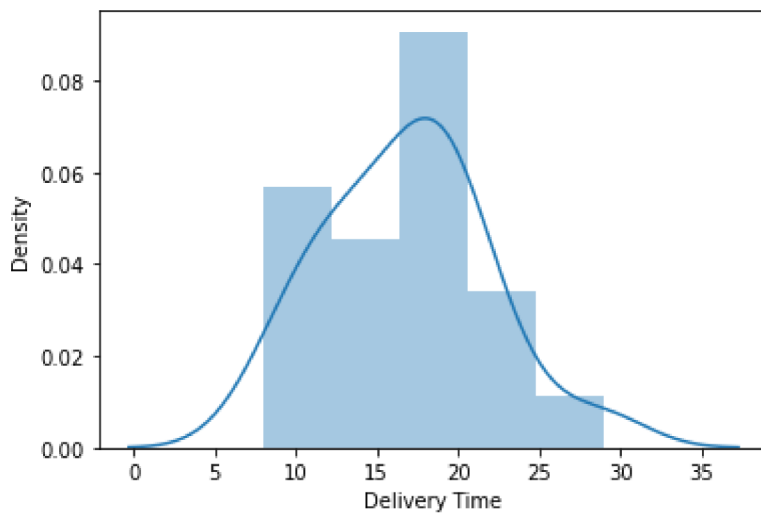
executed in 1.80s, finished 08:41:00 2021-11-26

C:\Users\win\anaconda3\lib\site-packages\seaborn\distributions.py:2557: Futu
reWarning: `distplot` is a deprecated function and will be removed in a futu
re version. Please adapt your code to use either `displot` (a figure-level f
unction with similar flexibility) or `histplot` (an axes-level function for
histograms).
  warnings.warn(msg, FutureWarning)

Out[7]:

```
<AxesSubplot:xlabel='Delivery Time', ylabel='Density'>
```

In [8]:

```python
import seaborn as sns
sns.distplot(data['Sorting Time'])
```
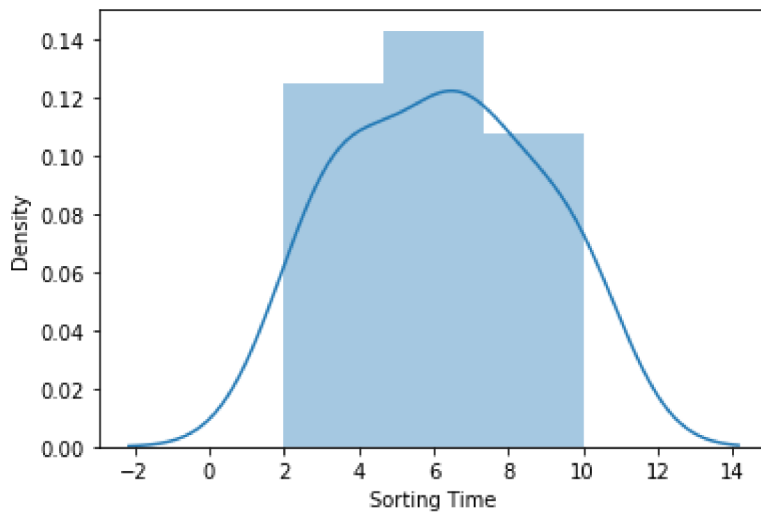
executed in 376ms, finished 08:41:50 2021-11-26

```
C:\Users\win\anaconda3\lib\site-packages\seaborn\distributions.py:2557: Futu
reWarning: `distplot` is a deprecated function and will be removed in a futu
re version. Please adapt your code to use either `displot` (a figure-level f
unction with similar flexibility) or `histplot` (an axes-level function for
histograms).
  warnings.warn(msg, FutureWarning)
```

Out[8]:

```
<AxesSubplot:xlabel='Sorting Time', ylabel='Density'>
```


In [20]:

```python
x = data['Sorting Time']
y = data['Delivery Time']
```

executed in 21ms, finished 08:51:47 2021-11-26

In [26]:

```python
b, m = polyfit(x, y, 1)
plt.scatter(x, y)
plt.plot(x, y, '.')
plt.plot(x, b + m * x, '-')
plt.title('Scatter plot Delivery Time')
plt.xlabel('Sorting Time')
plt.ylabel('Delivery Time')
plt.show()
```

executed in 367ms, finished 08:53:15 2021-11-26

In [35]:

```python
model = sm.OLS(y, x).fit()
predictions = model.predict(x)
```

executed in 26ms, finished 08:57:56 2021-11-26

In [36]:

```
model.summary()
```

executed in 47ms, finished 08:58:09 2021-11-26

Out[36]:

OLS Regression Results

Dep. Variable:	Delivery Time	**R-squared (uncentered):**	0.955
Model:	OLS	**Adj. R-squared (uncentered):**	0.953
Method:	Least Squares	**F-statistic:**	424.5
Date:	Fri, 26 Nov 2021	**Prob (F-statistic):**	6.12e-15
Time:	08:58:09	**Log-Likelihood:**	-57.349
No. Observations:	21	**AIC:**	116.7
Df Residuals:	20	**BIC:**	117.7
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Sorting Time	2.5652	0.125	20.603	0.000	2.306	2.825

Omnibus:	1.504	**Durbin-Watson:**	1.305
Prob(Omnibus):	0.471	**Jarque-Bera (JB):**	0.508
Skew:	-0.348	**Prob(JB):**	0.776
Kurtosis:	3.310	**Cond. No.**	1.00

Notes:
[1] R² is computed without centering (uncentered) since the model does not contain a constant.
[2] Standard Errors assume that the covariance matrix of the errors is correctly specified.

In [38]:

```
x_log = np.log(data['Sorting Time'])
```

executed in 13ms, finished 08:58:37 2021-11-26

In [39]:

```
model = sm.OLS(y, x_log).fit()
predictions = model.predict(x_log)
```

executed in 21ms, finished 08:59:18 2021-11-26

In [40]:

```
model.summary()
```

executed in 49ms, finished 08:59:34 2021-11-26

Out[40]:

OLS Regression Results

Dep. Variable:	Delivery Time	**R-squared (uncentered):**	0.975
Model:	OLS	**Adj. R-squared (uncentered):**	0.974
Method:	Least Squares	**F-statistic:**	791.0
Date:	Fri, 26 Nov 2021	**Prob (F-statistic):**	1.48e-17
Time:	08:59:34	**Log-Likelihood:**	-51.035
No. Observations:	21	**AIC:**	104.1
Df Residuals:	20	**BIC:**	105.1
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Sorting Time	9.6706	0.344	28.124	0.000	8.953	10.388

Omnibus:	3.656	**Durbin-Watson:**	1.453
Prob(Omnibus):	0.161	**Jarque-Bera (JB):**	2.164
Skew:	0.772	**Prob(JB):**	0.339
Kurtosis:	3.298	**Cond. No.**	1.00

Notes:
[1] R² is computed without centering (uncentered) since the model does not contain a constant.
[2] Standard Errors assume that the covariance matrix of the errors is correctly specified.

In [42]:

```
y_log = np.log(data['Delivery Time'])
```

executed in 8ms, finished 09:00:23 2021-11-26

In [43]:

```
model = sm.OLS(y_log, x).fit()
predictions = model.predict(x)
```

executed in 13ms, finished 09:00:40 2021-11-26

In [44]:

```
model.summary()
```

executed in 45ms, finished 09:00:51 2021-11-26

Out[44]:

OLS Regression Results

Dep. Variable:	Delivery Time	**R-squared (uncentered):**	0.917
Model:	OLS	**Adj. R-squared (uncentered):**	0.912
Method:	Least Squares	**F-statistic:**	219.7
Date:	Fri, 26 Nov 2021	**Prob (F-statistic):**	3.00e-12
Time:	09:00:51	**Log-Likelihood:**	-25.284
No. Observations:	21	**AIC:**	52.57
Df Residuals:	20	**BIC:**	53.61
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Sorting Time	0.4008	0.027	14.821	0.000	0.344	0.457

Omnibus:	2.572	**Durbin-Watson:**	1.446
Prob(Omnibus):	0.276	**Jarque-Bera (JB):**	1.346
Skew:	-0.275	**Prob(JB):**	0.510
Kurtosis:	1.889	**Cond. No.**	1.00

Notes:
[1] R^2 is computed without centering (uncentered) since the model does not contain a constant.
[2] Standard Errors assume that the covariance matrix of the errors is correctly specified.

In [45]:

```
model = sm.OLS(y_log, x_log).fit()
predictions = model.predict(x_log)
```

executed in 19ms, finished 09:01:08 2021-11-26

In [46]:

```
model.summary()
```

executed in 37ms, finished 09:01:28 2021-11-26

Out[46]:

OLS Regression Results

Dep. Variable:	Delivery Time	**R-squared (uncentered):**	0.972
Model:	OLS	**Adj. R-squared (uncentered):**	0.970
Method:	Least Squares	**F-statistic:**	688.7
Date:	Fri, 26 Nov 2021	**Prob (F-statistic):**	5.72e-17
Time:	09:01:28	**Log-Likelihood:**	-13.899
No. Observations:	21	**AIC:**	29.80
Df Residuals:	20	**BIC:**	30.84
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Sorting Time	1.5396	0.059	26.244	0.000	1.417	1.662

Omnibus:	1.636	**Durbin-Watson:**	1.727
Prob(Omnibus):	0.441	**Jarque-Bera (JB):**	1.137
Skew:	0.304	**Prob(JB):**	0.566
Kurtosis:	2.035	**Cond. No.**	1.00

Notes:
[1] R² is computed without centering (uncentered) since the model does not contain a constant.
[2] Standard Errors assume that the covariance matrix of the errors is correctly specified.

In [47]:

```
x_sqrt = np.sqrt(data['Sorting Time'])
```

executed in 14ms, finished 09:01:52 2021-11-26

In [48]:

```
model = sm.OLS(y, x_sqrt).fit()
predictions = model.predict(x_sqrt)
```

executed in 10ms, finished 09:02:09 2021-11-26

In [49]:

```
model.summary()
```

executed in 53ms, finished 09:02:23 2021-11-26

Out[49]:

OLS Regression Results

Dep. Variable:	Delivery Time	**R-squared (uncentered):**	0.975
Model:	OLS	**Adj. R-squared (uncentered):**	0.973
Method:	Least Squares	**F-statistic:**	772.0
Date:	Fri, 26 Nov 2021	**Prob (F-statistic):**	1.88e-17
Time:	09:02:23	**Log-Likelihood:**	-51.284
No. Observations:	21	**AIC:**	104.6
Df Residuals:	20	**BIC:**	105.6
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Sorting Time	6.9466	0.250	27.785	0.000	6.425	7.468

Omnibus:	6.818	**Durbin-Watson:**	1.334
Prob(Omnibus):	0.033	**Jarque-Bera (JB):**	4.599
Skew:	1.090	**Prob(JB):**	0.100
Kurtosis:	3.708	**Cond. No.**	1.00

Notes:
[1] R² is computed without centering (uncentered) since the model does not contain a constant.
[2] Standard Errors assume that the covariance matrix of the errors is correctly specified.

In [50]:

```
y_sqrt = np.sqrt(data['Delivery Time'])
```

executed in 12ms, finished 09:03:00 2021-11-26

In [51]:

```
model = sm.OLS(y_sqrt, x).fit()
predictions = model.predict(x)
```

executed in 7ms, finished 09:03:14 2021-11-26

In [52]:

```
model.summary()
```

executed in 41ms, finished 09:03:29 2021-11-26

Out[52]:

OLS Regression Results

Dep. Variable:	Delivery Time	R-squared (uncentered):	0.930
Model:	OLS	Adj. R-squared (uncentered):	0.927
Method:	Least Squares	F-statistic:	266.0
Date:	Fri, 26 Nov 2021	Prob (F-statistic):	5.09e-13
Time:	09:03:29	Log-Likelihood:	-31.484
No. Observations:	21	AIC:	64.97
Df Residuals:	20	BIC:	66.01
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Sorting Time	0.5926	0.036	16.309	0.000	0.517	0.668

Omnibus:	1.452	Durbin-Watson:	1.434
Prob(Omnibus):	0.484	Jarque-Bera (JB):	1.105
Skew:	-0.328	Prob(JB):	0.575
Kurtosis:	2.087	Cond. No.	1.00

Notes:
[1] R² is computed without centering (uncentered) since the model does not contain a constant.
[2] Standard Errors assume that the covariance matrix of the errors is correctly specified.

In [53]:

```
model = sm.OLS(y_sqrt, x_sqrt).fit()
predictions = model.predict(x_sqrt)
```

executed in 10ms, finished 09:03:51 2021-11-26

In [54]:

```
model.summary()
```

executed in 43ms, finished 09:04:03 2021-11-26

Out[54]:

OLS Regression Results

Dep. Variable:	Delivery Time	**R-squared (uncentered):**	0.987
Model:	OLS	**Adj. R-squared (uncentered):**	0.987
Method:	Least Squares	**F-statistic:**	1542.
Date:	Fri, 26 Nov 2021	**Prob (F-statistic):**	2.10e-20
Time:	09:04:03	**Log-Likelihood:**	-13.658
No. Observations:	21	**AIC:**	29.32
Df Residuals:	20	**BIC:**	30.36
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Sorting Time	1.6364	0.042	39.267	0.000	1.549	1.723

Omnibus:	0.176	**Durbin-Watson:**	1.461
Prob(Omnibus):	0.916	**Jarque-Bera (JB):**	0.231
Skew:	-0.179	**Prob(JB):**	0.891
Kurtosis:	2.632	**Cond. No.**	1.00

Notes:

[1] R² is computed without centering (uncentered) since the model does not contain a constant.

[2] Standard Errors assume that the covariance matrix of the errors is correctly specified.