

```

import numpy as np
import pandas as pd
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeRegressor
from sklearn.metrics import r2_score, mean_squared_error

df = pd.read_csv('DT-Regression-Data.csv')
df.head()

```

1 to 5 of 5 entries

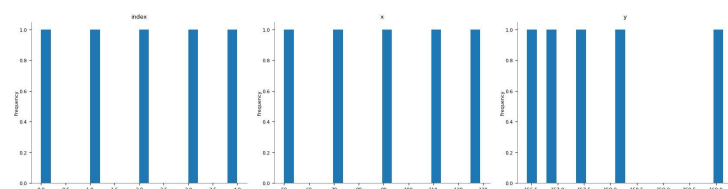
index	x	y
0	50.0	157.49342830602245
1	69.69696969696969	156.42044109462734
2	89.39393939393939	158.18931647014077
3	109.0909090909091	160.13696880372515
4	128.78787878787878	156.8195720384321

Show  per page

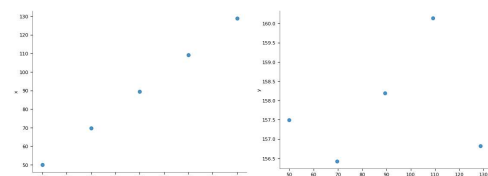


Like what you see? Visit the [data table notebook](#) to learn more about interactive tables.

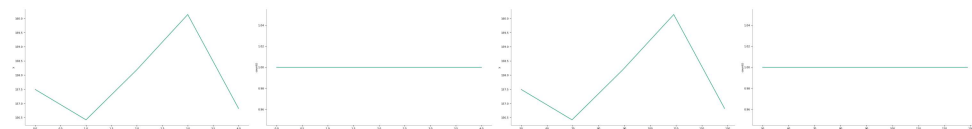
#### Distributions



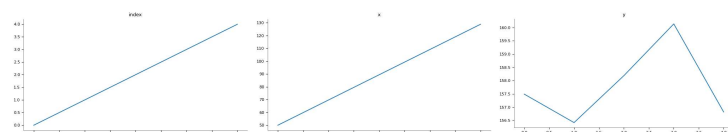
#### 2-d distributions



#### Time series



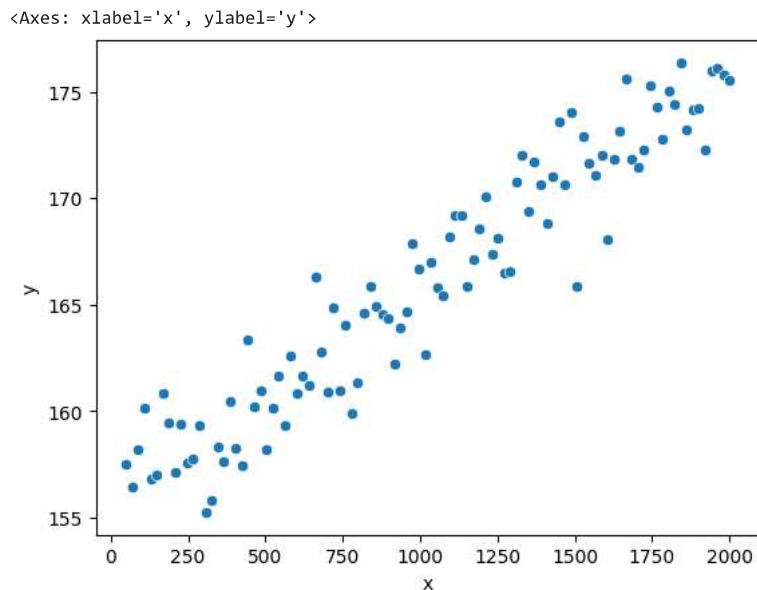
#### Values



```
df.shape
```

```
(100, 2)
```

```
sns.scatterplot(x=df.x, y=df.y, data=df)
```



```
x = df.x.values.reshape(-1, 1)
y = df.y.values.reshape(-1, 1)
x_train, x_test, y_train, y_test = train_test_split(x,
DecisionTreeRegModel = DecisionTreeRegressor()
```

```
DecisionTreeRegModel.fit(x_train,y_train)
```

```
▼ DecisionTreeRegressor
DecisionTreeRegressor()
```

```
y_pred = DecisionTreeRegModel.predict(x_test)
```

```
r2_score(y_test, y_pred)
```

```
0.7875383967595575
```

```
mse = mean_squared_error(y_test, y_pred)
```

```
mse
```

```
7.4250104983698595
```

```
rmse = np.sqrt(mse)
```

```
rmse
```

```
2.7248872450745294
```

