

```
# Anscombe's quartet
```

```
x=[10.0,8.0,13.0,9.0,11.0,14.0,6.0,4.0,12.0,7.0,5.0]
```

```
y=[8.04,6.95,7.58,8.81,8.33,9.96,7.24,4.26,10.84,4.82,5.68]
```

```
x=[10.0,8.0,13.0,9.0,11.0,14.0,6.0,4.0,12.0,7.0,5.0]
```

```
y=[9.14,8.14,8.74,8.77,9.26,8.10,6.13,3.10,9.13,7.26,4.74]
```

```
x=[10.0,8.0,13.0,9.0,11.0,14.0,6.0,4.0,12.0,7.0,5.0]
```

```
y=[7.46,6.77,12.74,7.11,7.81,8.84,6.08,5.39,8.15,6.42,5.73]
```

```
x=[8.0,8.0,8.0,8.0,8.0,8.0,8.0,19.0,8.0,8.0,8.0]
```

```
y=[6.58,5.76,7.71,8.84,8.47,7.04,5.25,12.50,5.56,7.91,6.89]
```

```
import matplotlib.pyplot as plt
import numpy as np
from sklearn import linear_model

X = np.array(x).reshape(-1,1)
regr = linear_model.LinearRegression()
regr.fit(X, y)
y_pred = regr.predict(X)

print('Coefficients: \n', regr.coef_, regr.intercept_)

plt.scatter(X, y, color='black')
plt.plot(X, y_pred, color='blue', linewidth=3)
plt.show()
```

Что показывают графики

1. ОК

2. Квадратичная зависимость

3. Все, кроме одного, лежат на прямой

4. Все, кроме одного, лежат на прямой

Множество 2: введем дополнительный признак

```
x=[10.0,8.0,13.0,9.0,11.0,14.0,6.0,4.0,12.0,7.0,5.0]  
y=[9.14,8.14,8.74,8.77,9.26,8.10,6.13,3.10,9.13,7.26,4.74]
```

```
X = [[z, z**2] for z in x]
```

```
regr = linear_model.LinearRegression()  
regr.fit(X, y)  
y_pred = regr.predict(X)
```

```
print('Coefficients: \n', regr.coef_, regr.intercept_)
```

```
plt.scatter(x, y, color='black')  
plt.plot(x, y_pred, color='blue', linewidth=3)  
plt.show()
```

Множество 3: используем робастные регрессоры: Тейла-Сена и RANSAC

```
x=[10.0,8.0,13.0,9.0,11.0,14.0,6.0,4.0,12.0,7.0,5.0]  
y=[7.46,6.77,12.74,7.11,7.81,8.84,6.08,5.39,8.15,6.42,5.73]
```

```
X = np.array(x).reshape(-1,1)  
#regr = linear_model.LinearRegression()  
#regr = linear_model.TheilSenRegressor(random_state=42)  
regr = linear_model.RANSACRegressor(random_state=42)  
regr.fit(X, y)  
y_pred = regr.predict(X)  
  
print('Coefficients: \n', regr.coef_, regr.intercept_)
```

```
plt.scatter(X, y, color='black')  
plt.plot(X, y_pred, color='blue', linewidth=3)  
plt.show()
```