

```

import pandas as pd
import matplotlib.pyplot as plt
import numpy as np

##### DATA PREPARATION
data_file = 'data_ex1_wt.csv' # absolute path respect to when the script is run
df = pd.read_csv(data_file, header=None, names=['time', 'metric'])

x = df['time'].values
y = df['metric'].values

##### REMOVE TREND

m = 5 #best degree to fit the trend

def least_squares_methods(x, y, degree):

    #compute Vandermonde matrix
    A = []

    for i in range(len(x)):
        A.append([])

    for j in range(len(x)):
        for i in range(degree+1):
            A[j].append(x[j]**i)

    #transpose of A
    A_transpose = np.transpose(A)

    # Compute A' A
    A_transpose_A = np.dot(A_transpose, A)
    # Compute (A' A)^-1
    A_transpose_A_inv = np.linalg.inv(A_transpose_A)
    # Compute (A' A)^-1 A'
    A_transpose_A_inv_transpose_A = np.dot(A_transpose_A_inv, A_transpose)
    # Compute (A' A)^-1 A' y (= b)
    b = np.dot(A_transpose_A_inv_transpose_A, y)

    return b

p = least_squares_methods(x, y, m)
yy = np.zeros(len(x))
for j in range(len(p)):
    yy += p[j] * (x**j)

#so let's compute the residuals by subtracting the polynomial of degree 5 from the data
residuals = y - yy

##### PLOT RESULTS

BINS=200

fig = plt.figure()
grid = plt.GridSpec(2, 2)
ax_old1 = plt.subplot(grid[0, 0])
ax_old2 = plt.subplot(grid[1, 0])
ax_new3 = plt.subplot(grid[0, 1])
ax_new4 = plt.subplot(grid[1, 1])

ax_old1.scatter(x, y, color='gray', s=20, alpha=0.3)
ax_old1.set_title('Scatter plot of data WITH trend')
ax_old1.set_ylim(min(y)-1, max(y)+1)
ax_old1.set_xlabel("Time")
ax_old1.set_ylabel("Metric")

```

```
ax_old2.hist(y, bins=BINS, color='blue', alpha=0.5, density=False)
ax_old2.set_xlabel("Metric values")
ax_old2.set_ylabel("Frequency")
ax_old2.set_title('Histogram of data WITH trend')

ax_new3.scatter(x, residuals, color='violet', s=20, alpha=0.3)
ax_new3.set_title('Scatter plot of data WITHOUT trend')
ax_new3.set_ylim(min(residuals)-1, max(residuals)+1)
ax_new3.set_xlabel("Time")
ax_new3.set_ylabel("Metric")

ax_new4.hist(residuals, bins=BINS, color='red', alpha=0.5, density=False)
ax_new4.set_title('Histogram of of data WITHOUT trend')
ax_new4.set_xlabel("Metric values")
ax_new4.set_ylabel("Frequency")

plt.tight_layout()
plt.show()
```