LAB MANUAL

Unit III – Machine Learning

Unit III – Machine Learning

Lab 1. Implementation of linear regression in Excel using CO2 emissions dataset

Objective

The objective of this project is to implement Linear Regression in Excel using a CO2 Emissions dataset. The goal is to predict the CO2 Emissions based on a corresponding variable, such as Industrial Activity, and analyze the relationship between these variables using a linear regression model.

Problem

To implement Linear Regression in Excel using a CO2 Emissions Dataset, follow the steps below. We will assume you have a dataset with CO2 emissions data along with a corresponding variable.

Solution

To predict solar power output using linear regression and save the model, we'll go through the following steps:

1. Insert a Scatter Plot
   - Highlight the data range for the independent variable (Industrial Activity) and dependent variable (CO2 Emissions)
   - Go to the Insert tab in the Excel ribbon.
   - In the Charts group, click on Insert Scatter (X, Y) or Bubble Chart.
   - Choose Scatter with Straight Lines or Scatter with Markers to display the data points.
2. Add a Trendline (Linear Regression)
   - Right-click on any data point in the scatter plot.
   - Click Add Trendline.
   - In the Format Trendline pane, do the following:
   - Select Linear for the type of trendline.
   - Check Display Equation on chart.

o   Check Display R-squared value on chart.

o   To predict future CO2 emissions
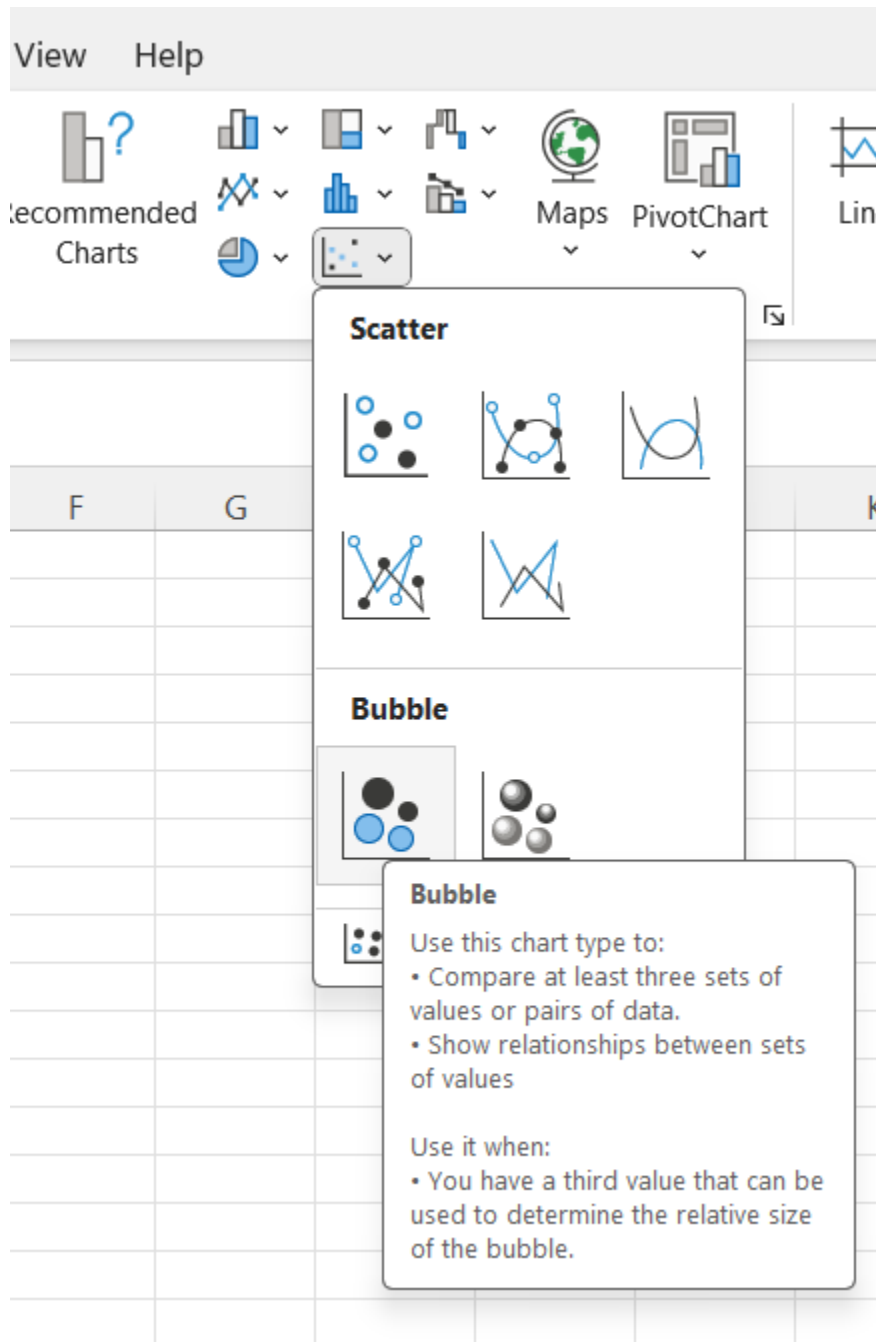
Procedures

1. Insert a Scatter Plot

1. Highlight the data range for the independent variable (Industrial Activity) and dependent variable (CO2 Emissions)

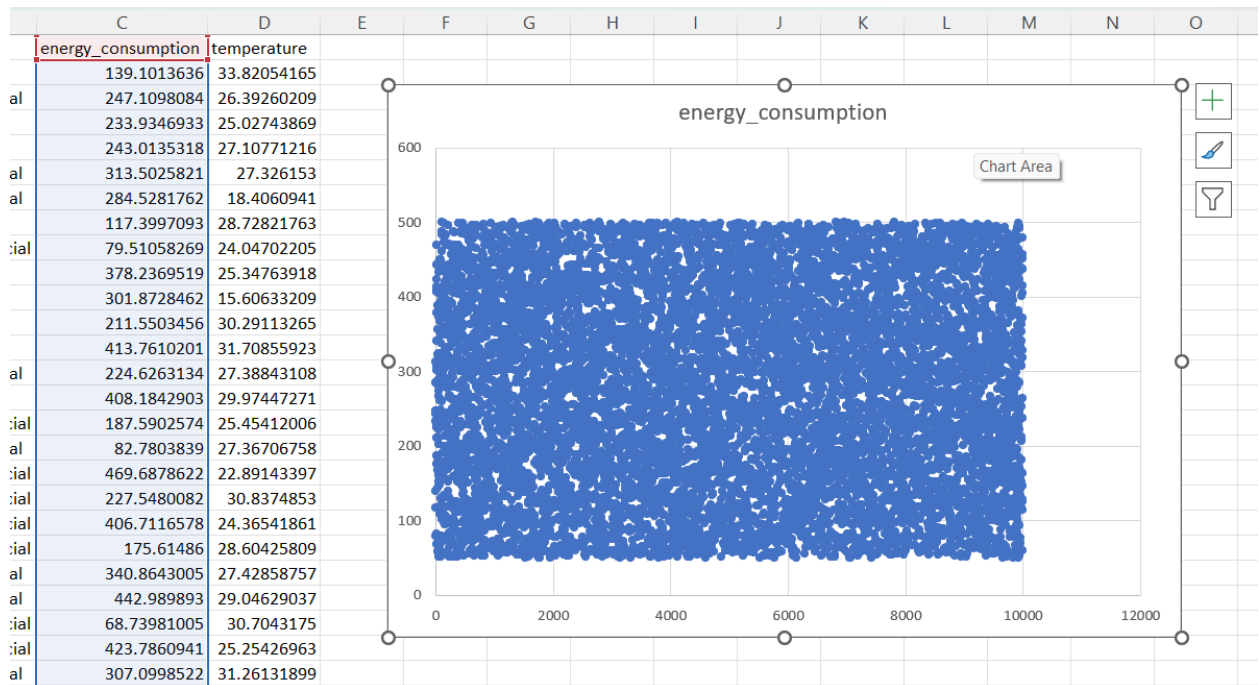| C |
| --- |
| energy_consumption |
| 139.1013636 |
| 247.1098084 |
| 233.9346933 |
| 243.0135318 |
| 313.5025821 |
| 284.5281762 |
| 117.3997093 |
| 79.51058269 |
| 378.2369519 |
| 301.8728462 |
| 211.5503456 |
| 413.7610201 |
| 224.6263134 |
| 408.1842903 |
| 187.5902574 |
| 82.7803839 |
| 469.6878622 |
| 227.5480082 |
| 406.7116578 |
| 175.61486 |
| 340.8643005 |
| 442.080802 |

2. Go to the Insert tab in the Excel ribbon.

File   Home   **Insert**   Draw   Page Layout   Formulas   Data   Review   View   Help

3. In the Charts group, click on Insert Scatter (X, Y) or Bubble Chart.


4. Choose Scatter with Straight Lines or Scatter with Markers to display the data points.

|	C	D	E	F	G	H	I	J	K	L	M	N	O
	energy_consumption	temperature											
	139.1013636	33.82054165											
al	247.1098084	26.39260209											
	233.9346933	25.02743869											
	243.0135318	27.10771216											
al	313.5025821	27.326153											
al	284.5281762	18.4060941											
	117.3997093	28.72821763											
:ial	79.51058269	24.04702205											
	378.2369519	25.34763918											
	301.8728462	15.60633209											
	211.5503456	30.29113265											
	413.7610201	31.70855923											
al	224.6263134	27.38843108											
	408.1842903	29.97447271											
:ial	187.5902574	25.45412006											
al	82.7803839	27.36706758											
:ial	469.6878622	22.89143397											
:ial	227.5480082	30.8374853											
:ial	406.7116578	24.36541861											
:ial	175.61486	28.60425809											
al	340.8643005	27.42858757											
al	442.989893	29.04629037											
:ial	68.73981005	30.7043175											
:ial	423.7860941	25.25426963											
al	307.0998522	31.26131899											

2. Add a Trendline (Linear Regression)

1. Right-click on any data point in the scatter plot.
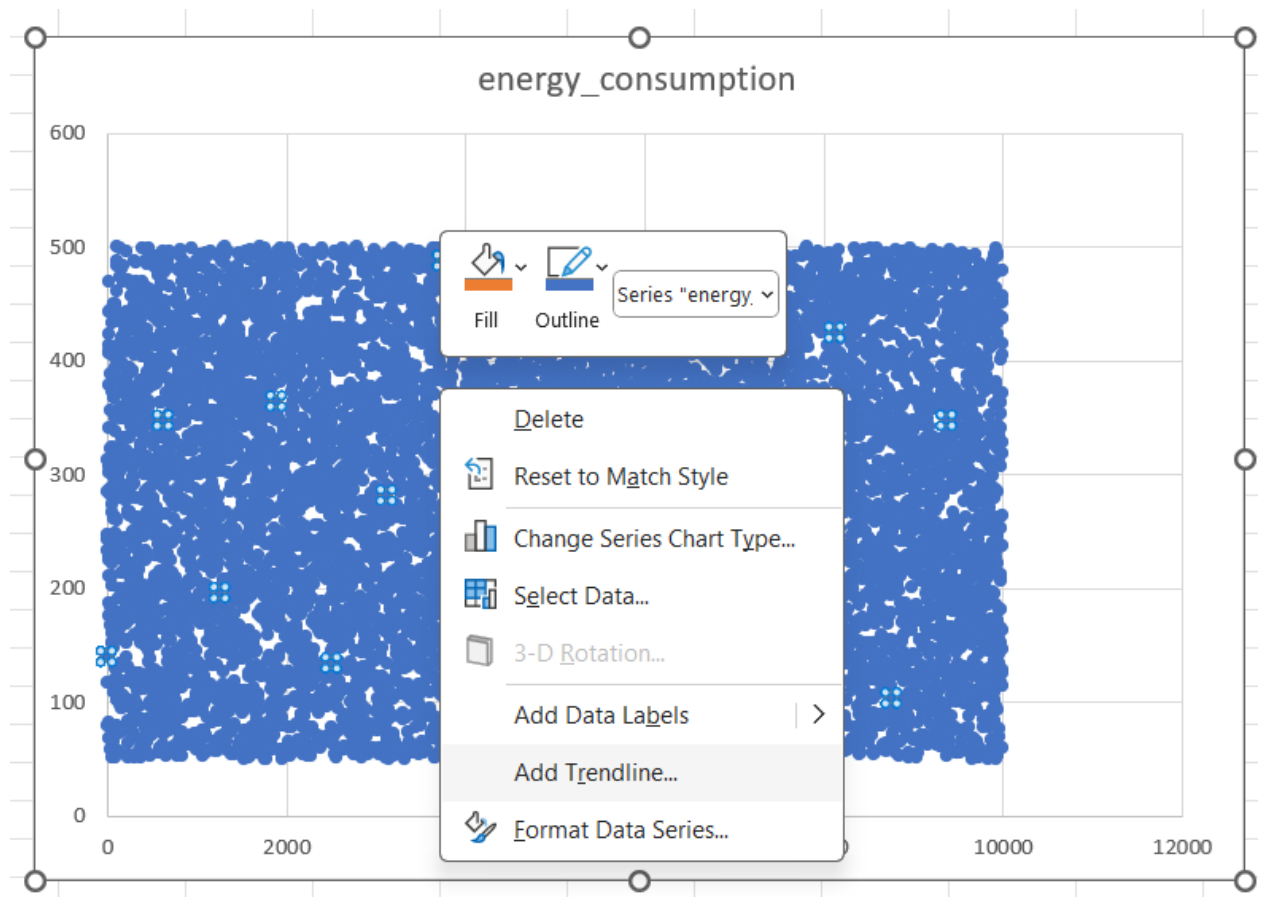
2. Click Add Trendline.
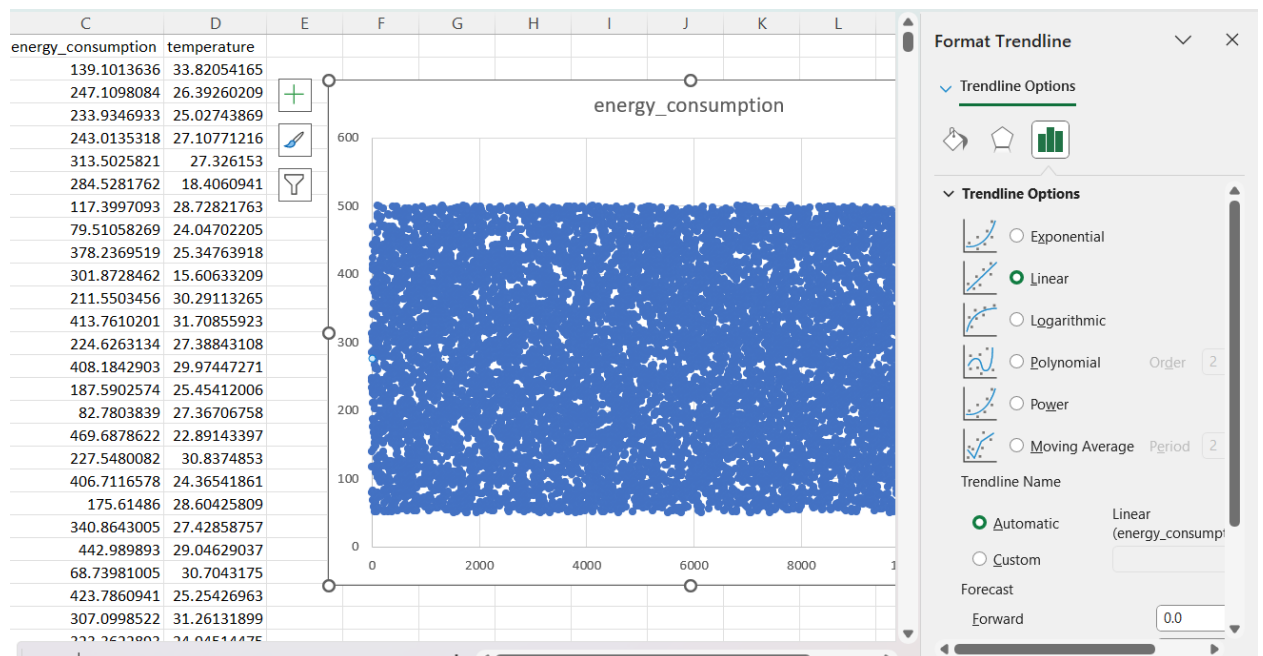
3. In the Format Trendline pane, do the following:

   ○   Select Linear for the type of trendline.

∨ Trendline Options

     ○ E*x*ponential

     ◉ Linear

     ○ L*o*garithmic

     ○ *P*olynomial   Or*d*er 2

     ○ Po*w*er

     ○ *M*oving Average P*e*riod 2

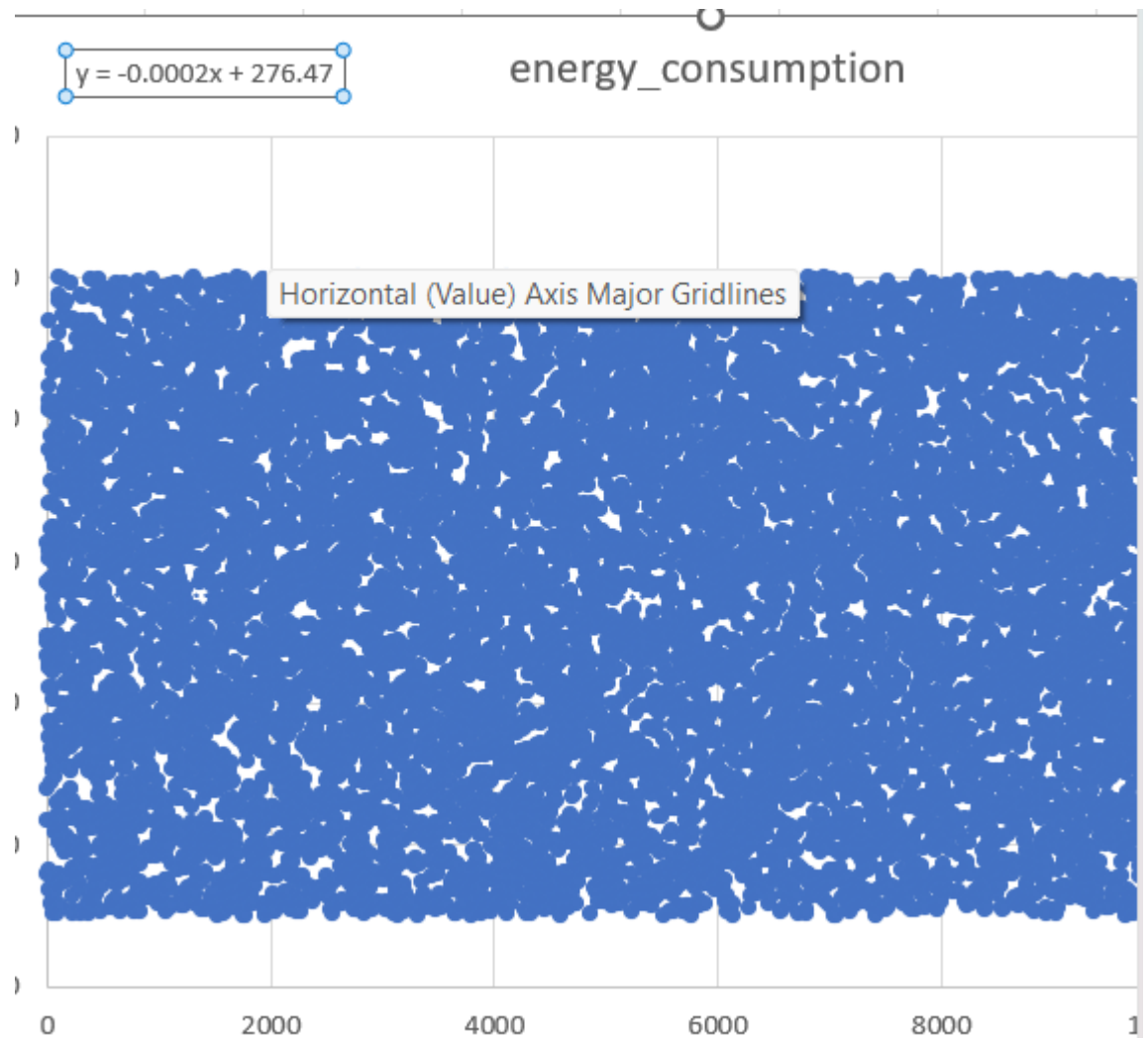   ○   Check Display Equation on chart.

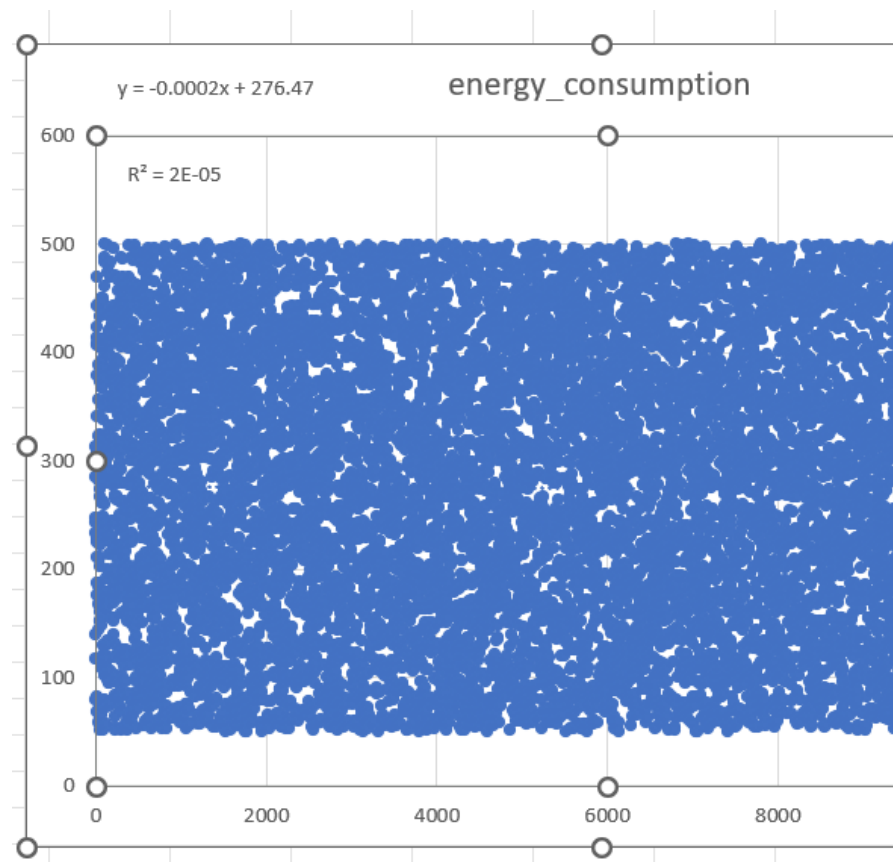   *B*ackward   0.0

     ☐ *S*et Intercept

     ☑ Display *E*quation on chart

     ☐ Display *R*-squared value on chart

y = -0.0002x + 276.47

energy_consumption

Horizontal (Value) Axis Major Gridlines

0   2000   4000   6000   8000   1

o   Check Display R-squared value on chart.

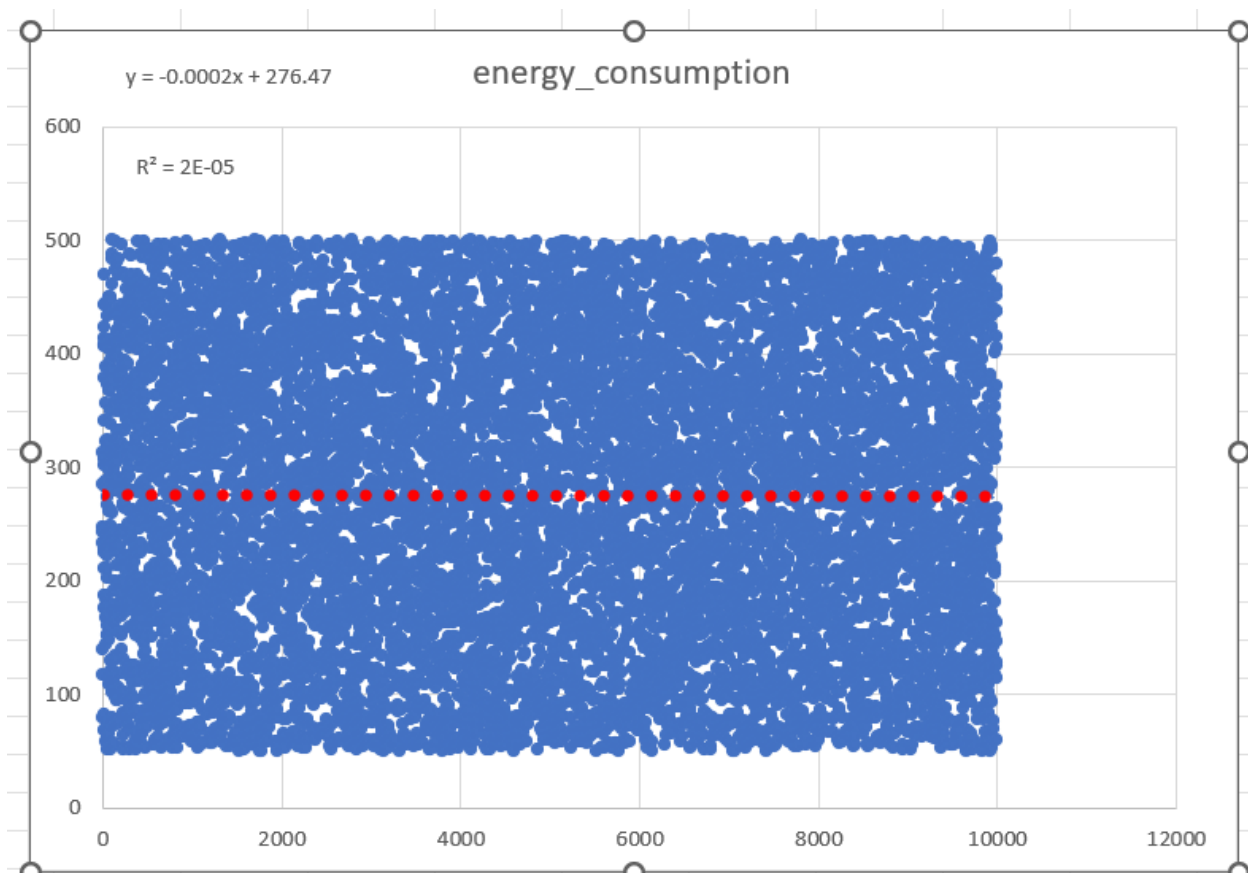☐ Display Equation on chart

✔ Display R-squared value on chart

The trendline (linear regression) will be added to your scatter plot, showing the regression equation and R-squared value.

Predict Future CO2 Emissions

To predict future CO2 emissions, use the regression equation that Excel displays on the chart.

y = -0.0002x + 276.47

energy_consumption

R² = 2E-05

y = -0.0002x + 276.47

R² = 2E-05