

Polymath Regression tutorial on Polynomial fitting of data (Example 16-1)

The following table shows the raw data for experimental tracer concentration from a reactor which you need to fit using Polymath (refer Example 16-1, Table E16-1.1, Elements of chemical reaction engineering, 6th edition)

Time (t)	0	0.5	1	2	3	4	5	6	7	8	9	10	12	14
C_tracer (C)	0	0.6	1.4	5	8	10	8	6	4	3	2.2	1.6	0.6	0

The polynomial fit is of the form:

$$C(t) = a_0 + a_1t + a_2t^2 + a_3t^3 + a_4t^4 + \dots + a_nt^n$$

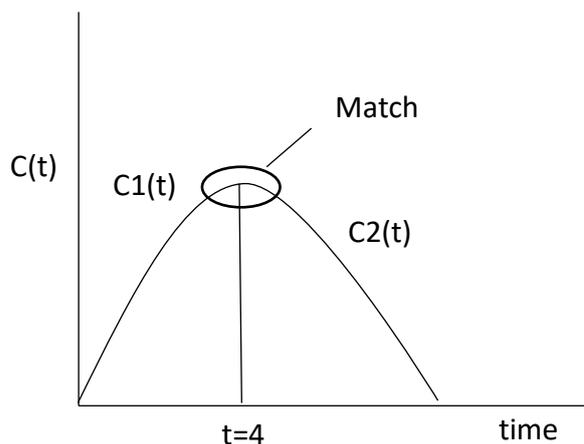
The given C(t) data is first increasing and then decreasing with time. So, we will use two polynomials to fit the C-curve, one for the ascending portion, $C_1(t)$, and another for the descending portion, $C_2(t)$. We will split our data into two parts, an increasing and a decreasing part, and regress both the data sets separately. Split the data at its maximum, and that happens at $t = 4$, when $C(t)$ is 10. Our data now looks like this:

T1	0	0.5	1	2	3	4
C1	0	0.6	1.4	5	8	10

and

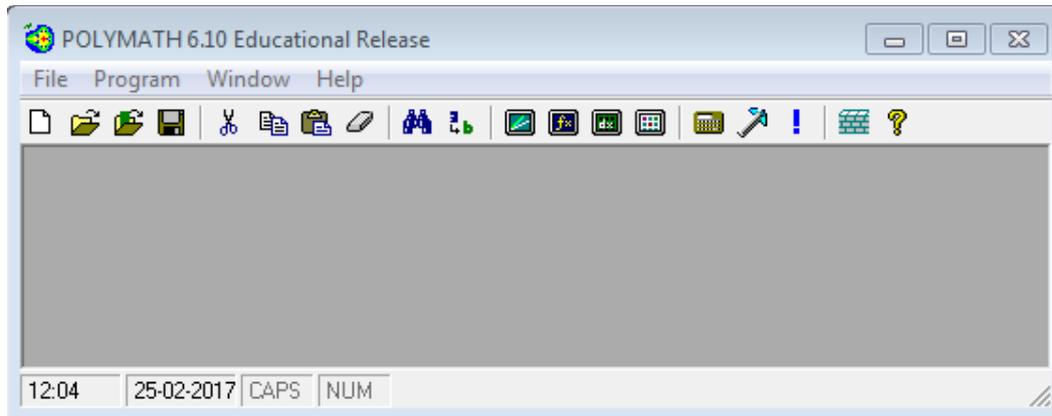
T2	4	5	6	7	8	9	10	12	14
C2	10	8	6	4	3	2.2	1.6	0.6	0

Both the polynomial must meet and give same value at $t=4$ and should look like this

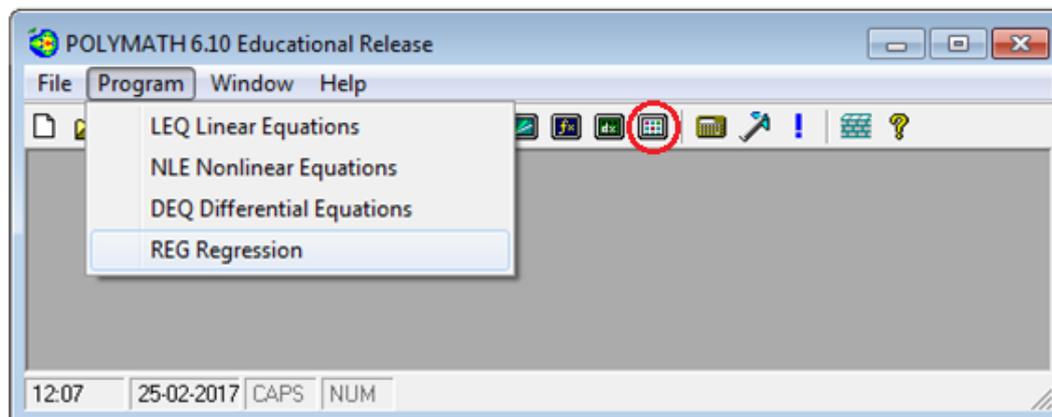


We will first regress the first portion of $C(t)$ curve and then regress the second portion

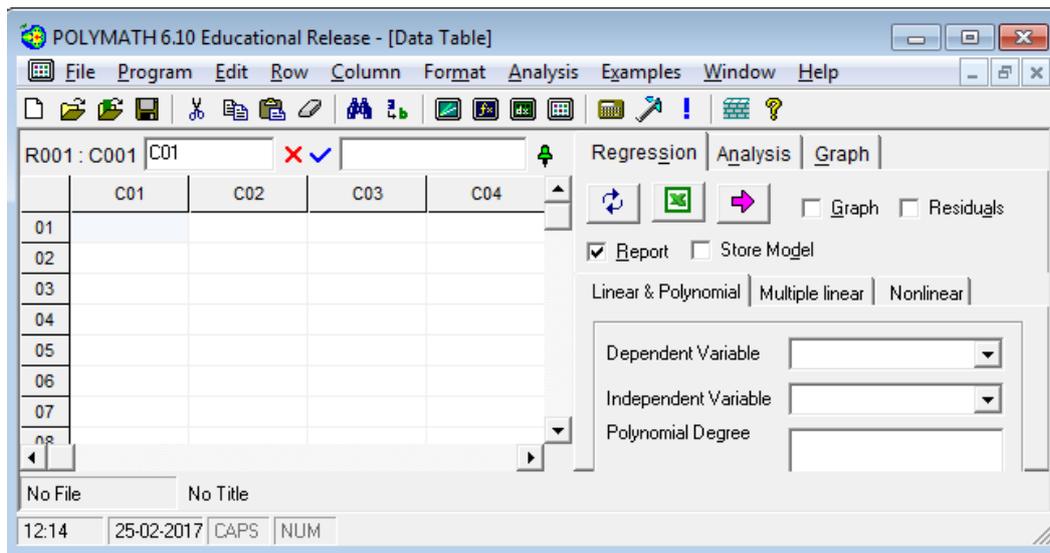
Step 1: First, launch Polymath which you can download from <http://www.polymath-software.com> You will see a window that looks like this.



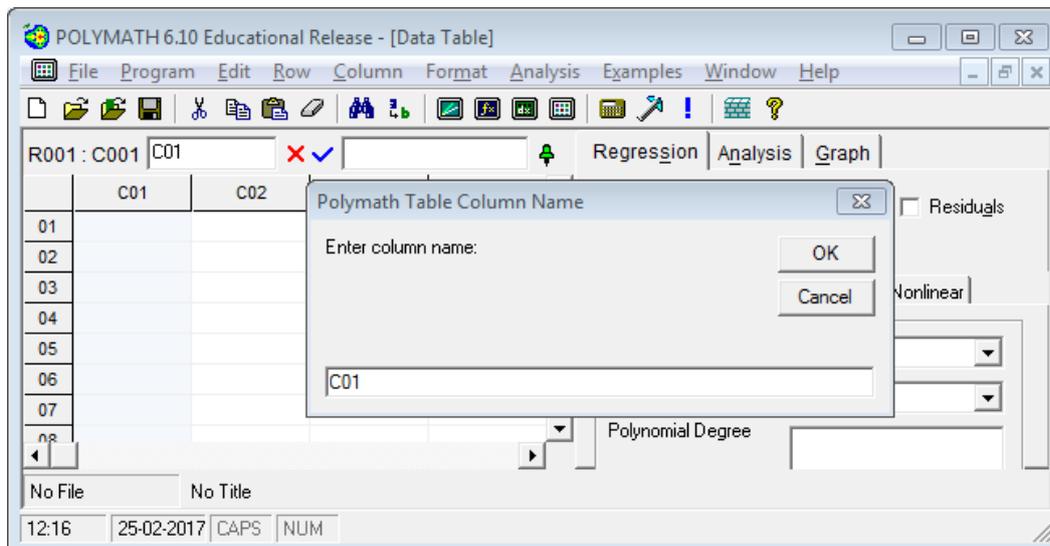
Step 2: To regress the data in Polymath, first click on the “Program” tab present on the toolbar. This will bring up a list of options from which you need to select. In this case we want to perform polynomial fitting, so select "REG Regression". The shortcut button for Regression solver is also present on the menu bar () as shown by red circle in below screenshot



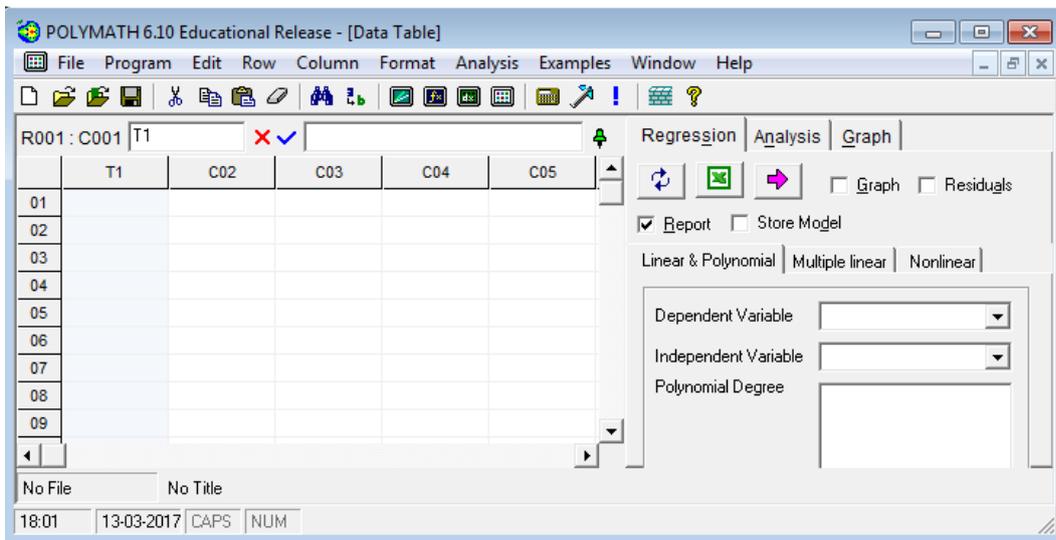
The following window will appear



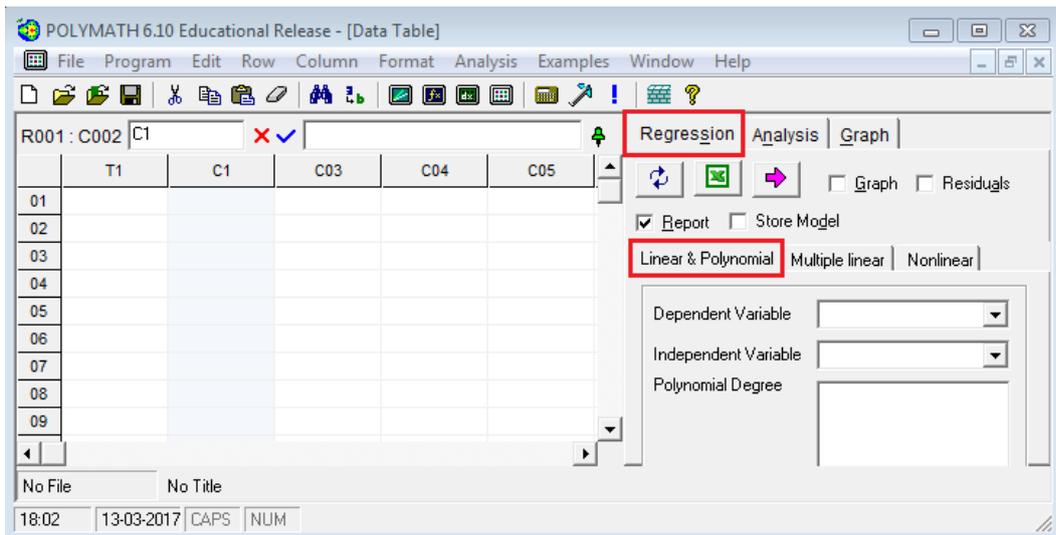
Step 3: Before inserting the data into the spreadsheet, it is recommended to change the column name with the name of the variable mentioned in the data table. This would make it easy to comprehend the polymath output. To change the column name of C01, double click on the column name “C01” or right click on C01 and select “Column Name...” A dialog box will appear where column name can be changed



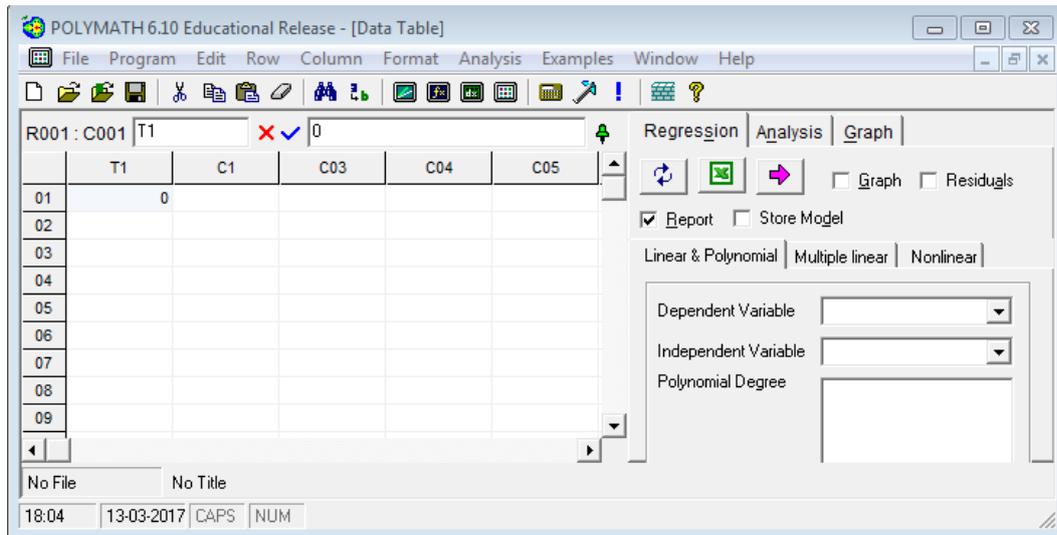
Change the column name from C01 to T1 and press OK. You will find that column name is changed to T1



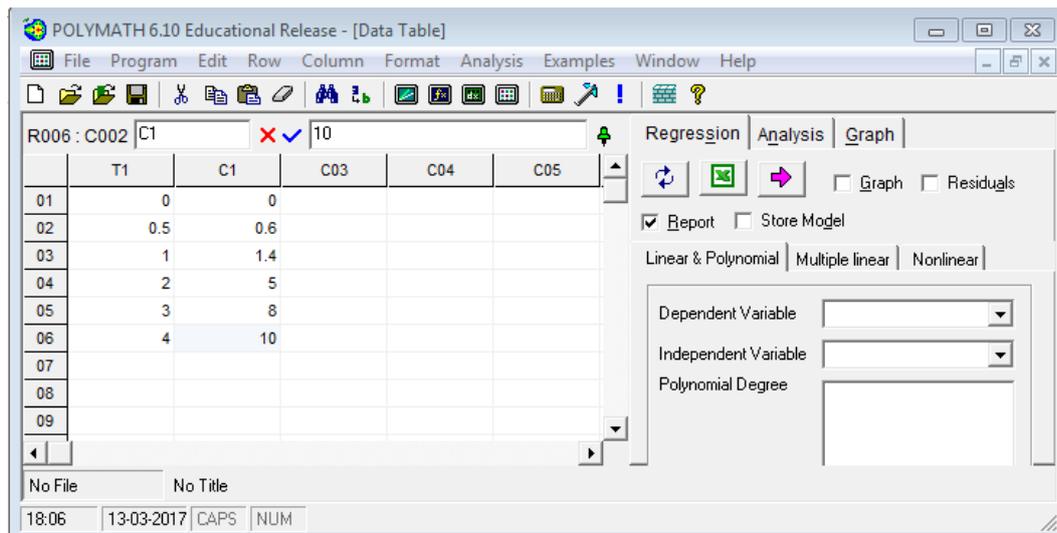
Similarly, rename C02 to C1 as shown below. By default Polymath select Regression method as “Linear and Polynomial”. If it is not so, then, click on the Regression tab on the right side of the window, and select the "Nonlinear" regression tab under the "Report" and "Store Model" check boxes (shown by red rectangle) The window should look like this:



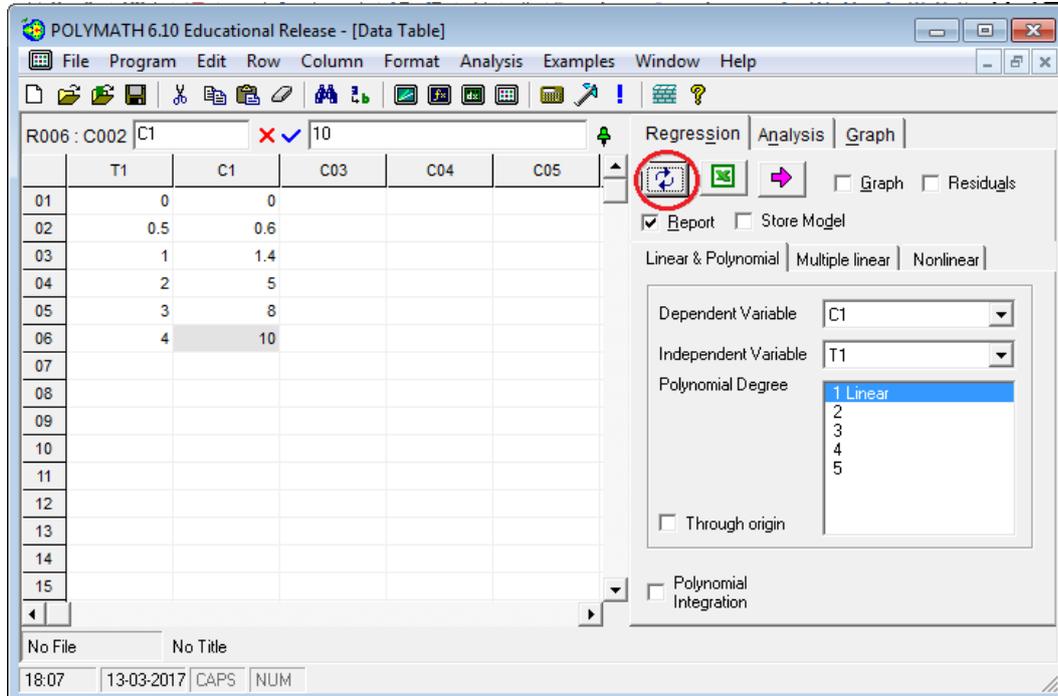
Step 4: To input the data for T1, select the first cell (row 01, column time) and enter the first data as shown below:



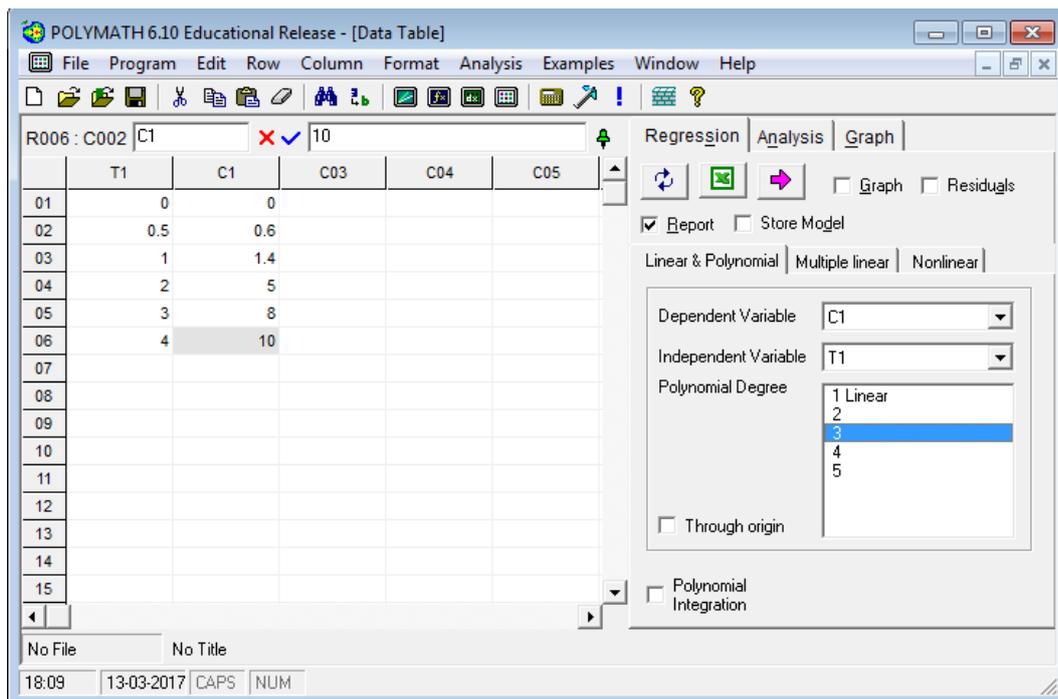
Similarly, enter the remaining data of T1 in subsequent rows. Repeat this procedure to input the data for C1. After entering the data, the spreadsheet would look like this



Step 5: Clicking the refresh button (shown below by red circle) on the left of the green excel sign updates the file. Now, you will find that Dependent variable, Independent variable and Polynomial degree is updated (shown by blue rectangle). By default, Polymath select the first column as independent variable, second column as dependent variable, and Polynomial degree as “1 Linear”. The window should look like this:

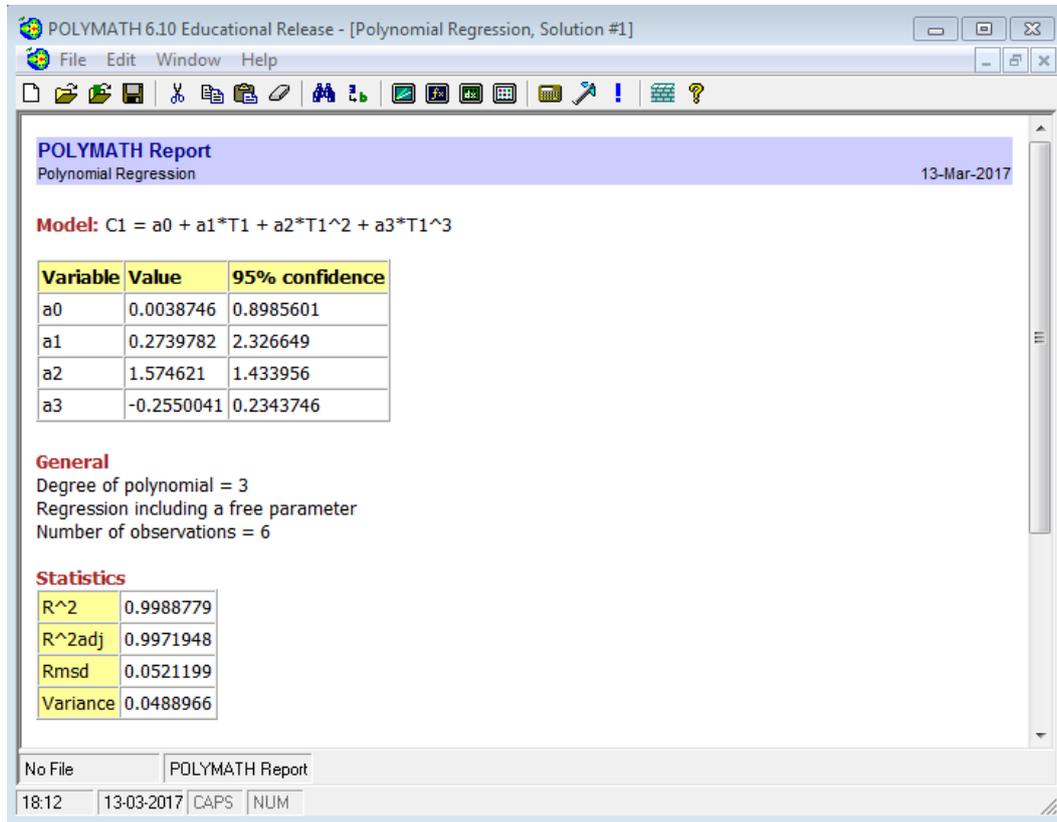


Step 6: You can also change the dependent variable, independent variable by selecting from the drop down menu next to variable name. In this case we want to change the polynomial degree only. Let’s try to fit the data by Polynomial degree 3. Select the Polynomial degree as 3 as shown below



Step 7: Now select what you want polymath to output by checking the boxes on the right side of the window. The options are Graph, Residuals, Report, and Store Model. Click on the pink arrow  to have Polymath perform the regression.

If you selected "Report" you will see a screen that details the results from the regression analysis. The report includes the model, the values of the model parameters, the statistical confidence of the parameters, and several other statistics.

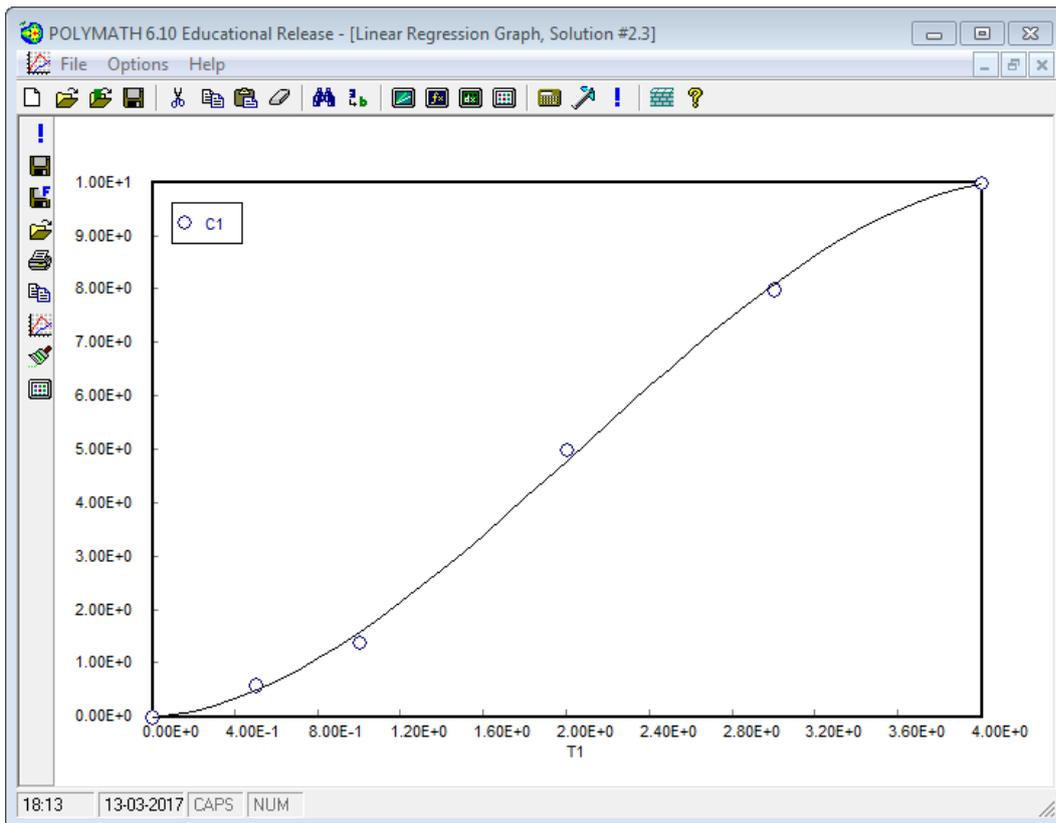


The above report shows that 3rd degree polynomial fits well to the 1st set of data as is evident from $R^2 = 0.999$

The equation for 1st data set is

$$C_1(t) = 0.0039 + 0.274t + 1.57t^2 - 0.255t^3$$

If you checked the graph option, a new window will appear like this. The graph shows the model that we fit to the data.



Step 8: To enter the second set of data in the spreadsheet, change C03 to T2 and C04 to C2, enter the data in corresponding column and then you need to select the dependent and independent variable as C2 and T2 from the drop down menu (shown by red circle in the screenshot). Select the polynomial degree as 3 as was done earlier. Your window will appear like this

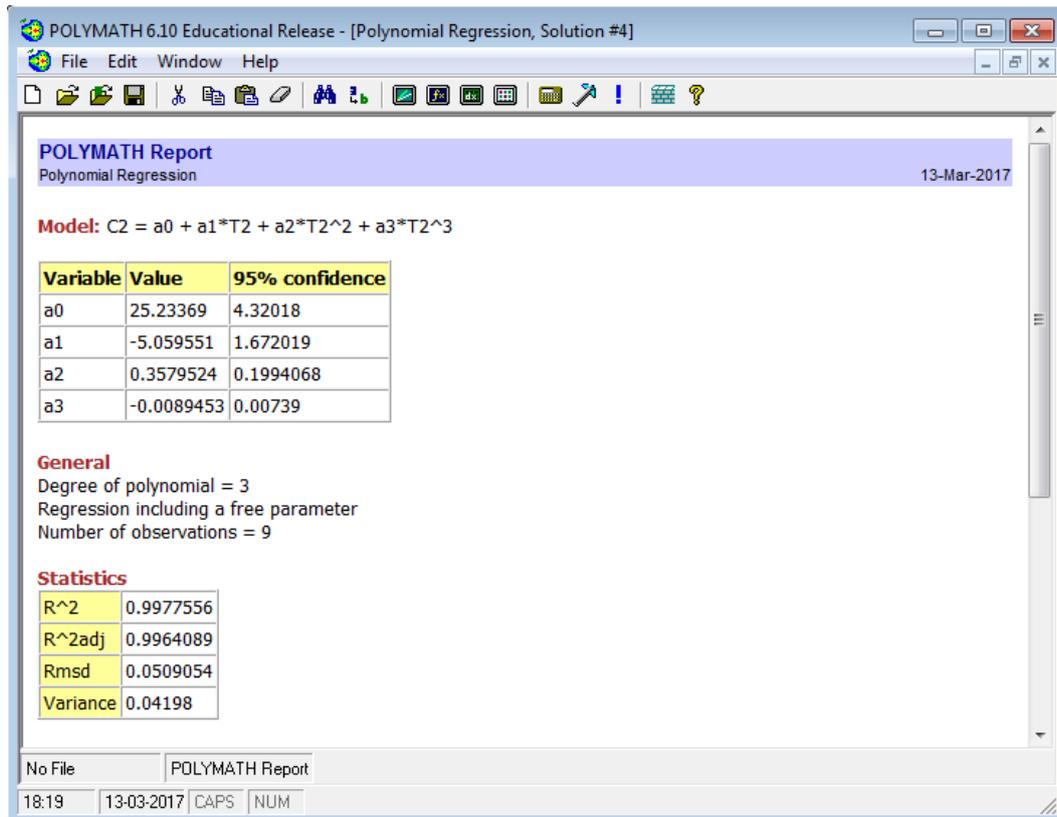
The screenshot shows the 'Data Table' and 'Regression Analysis' panels. The Data Table has the following data:

	T1	C1	T2	C2	C05
01	0	0	4		10
02	0.5	0.6	5		8
03	1	1.4	6		6
04	2	5	7		4
05	3	8	8		3
06	4	10	9		2.2
07			10		1.6
08			12		0.6
09			14		0

The Regression Analysis panel shows the following settings:

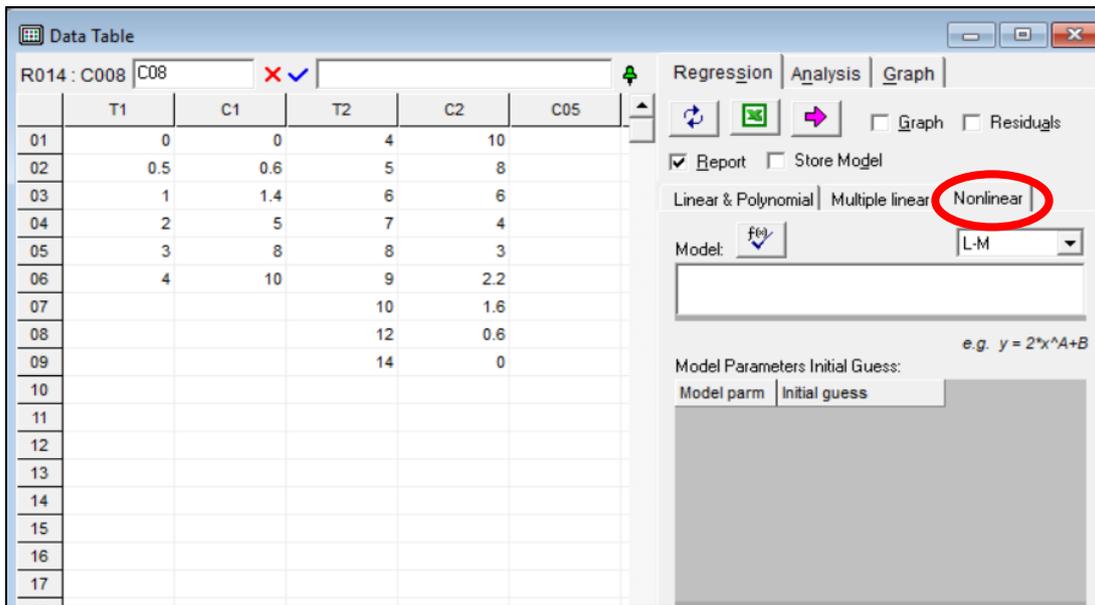
- Dependent Variable: C1
- Independent Variable: T1
- Polynomial Degree: 3
- Through origin:
- Polynomial Integration:
- Polynomial Derivative:

Run the model by clicking the pink arrow and generate the report. The report shows R² value to be 0.998 which is a good fit



To improve the accuracy of the fitting of the second data set, we can use higher order polynomial. **Let's regress using a 6th Order polynomial.** The maximum polynomial degree is limited to 5 under "Linear and Polynomial Tab". So, we will use another feature to regress polynomials with order greater than 5

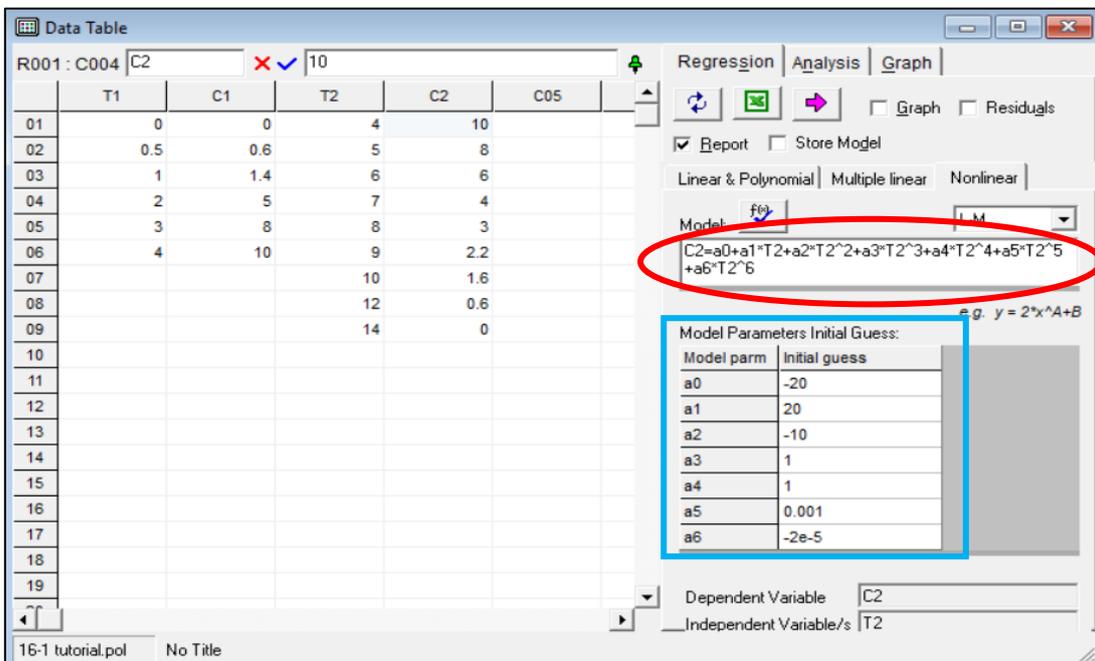
Step 9: Go back to the main window and click on “Nonlinear” Tab. In the Nonlinear Tab, we can enter equations of our choice to regress the data.



Step 10: Now enter the 6th order polynomial equation in “Model” Section (shown in red circle)

$$C2 = a_0 + a_1 \cdot T2 + a_2 \cdot T2^2 + a_3 \cdot T2^3 + a_4 \cdot T2^4 + a_5 \cdot T2^5 + a_6 \cdot T2^6$$

Next, we need to enter Initial guess values for the Model Parameters. The initial guess values are shown under blue Box. You can start with some initial guess values and check R² value. If R² is not good, then try another guess value until you obtain R² > 0.99



Step 11: Now Run the file by clicking on the pink arrow  to have Polymath perform the regression. The following report is generated. It is to be noted that R² value is >0.999 which represents a very good fit. All the model parameters are given in the table

POLYMATH Report			
Nonlinear Regression (L-M)			
Model: C2 = a0+a1*T2+a2*T2^2+a3*T2^3+a4*T2^4+a5*T2^5+a6*T2^6			
Variable	Initial guess	Value	95% confidence
a0	-20.	-33.43834	265.6643
a1	20.	37.18985	218.3198
a2	-10.	-11.58842	71.92468
a3	1.	1.69531	12.177
a4	1.	-0.1298673	1.119291
a5	0.001	0.005028	0.053058
a6	-2.0E-05	-7.743E-05	0.0010153
Nonlinear regression settings			
Max # iterations = 64			
Precision			
R ²	0.9995855		
R ² adj	0.9983418		
Rmsd	0.0218774		
Variance	0.0193842		

The equation for 2nd data set is

$$C_2(t) = -33.4 + 37.2t - 11.6t^2 + 1.7t^3 - 0.13t^4 + 0.005t^5 - 7.7 * 10^{-5}t^6$$

From Step 7 and Step 11, the polynomial equation is given by

For 1st data set , $t \leq 4 \text{ min}$, $C_1(t) = 0.0039 + 0.274t + 1.57t^2 - 0.255t^3$

For 2nd data set , $t \geq 4 \text{ min}$ and $t \leq 14 \text{ min}$,

$$C_2(t) = -33.4 + 37.2t - 11.6t^2 + 1.7t^3 - 0.13t^4 + 0.005t^5 - 7.7 * 10^{-5}t^6$$

The complete equation for C(t) curve can be written as

$$C(t) = \text{If}(t \leq 4 \text{ and } t \geq 0) \text{ then } C_1 \text{ else if } (t \geq 4 \text{ and } t \leq 14) \text{ then } C_2 \text{ else } 0$$