THE CHINESE UNIVERSITY OF HONG KONG
Department of Mathematics

Mathematical Modeling Project Team
mathmodel@math.cuhk.edu.hk

Modeling with R Shiny (Linear Regression)

Last updated: 29/06/2025

1 Introduction to the Linear Regression R Shiny tool

R is a powerful statistical programming language that allows us to perform data analysis, which is useful for mathematical modeling. Even better, we have set up a set of online tools using R Shiny so that you can perform various mathematical modelling tasks without any coding background.

In particular, a common task in mathematical modelling is to perform linear regression. Here is the guideline on how to use our Linear Regression R Shiny tool.

2 Using the Linear Regression R Shiny tool

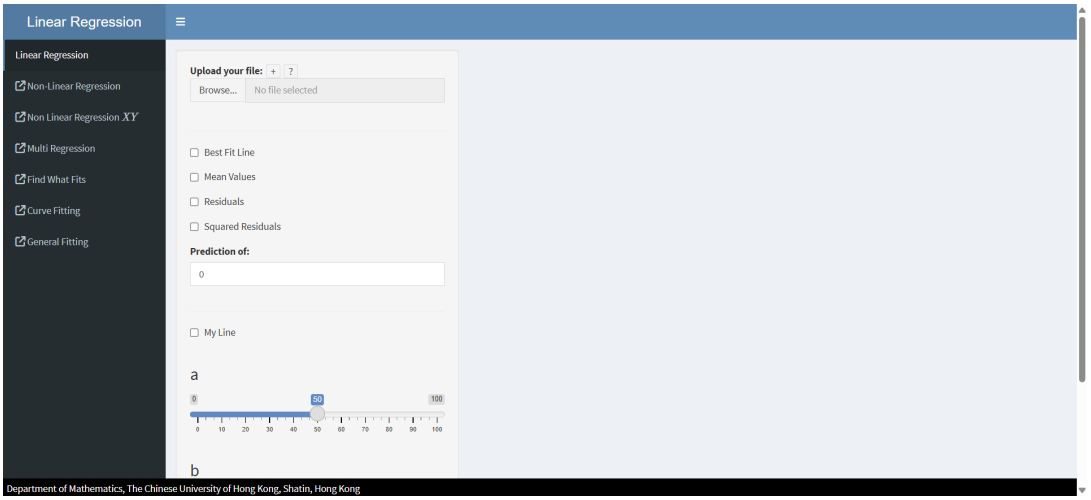
To use the tool, you can go to our main website:

https://www.math.cuhk.edu.hk/app/mathmodel/tool.html

for the list of tools, or simply go to

https://mathmodelcuhk.shinyapps.io/linear-regression/

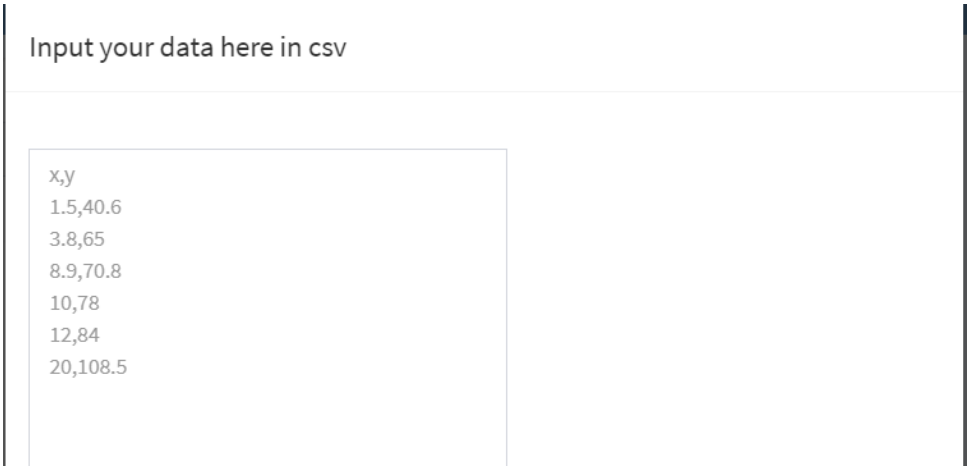
for the Linear Regression R Shiny tool.
After getting into the website, you can see the following:



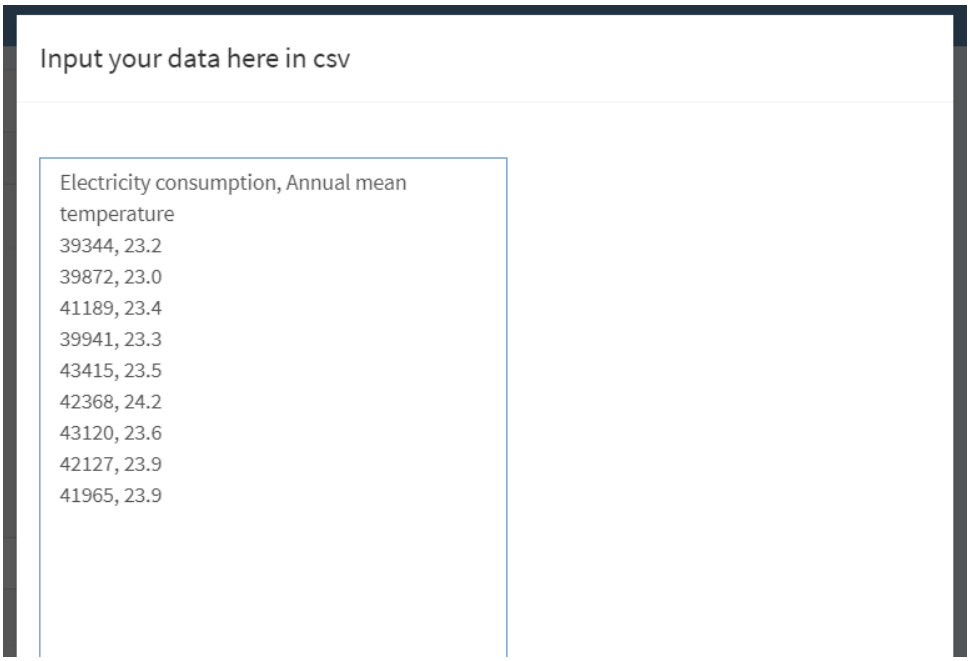
Don't worry if you think this is complicated! We will guide you step by step and you can do linear regression using our R Shiny tool.

2.1 Step 1: Data input

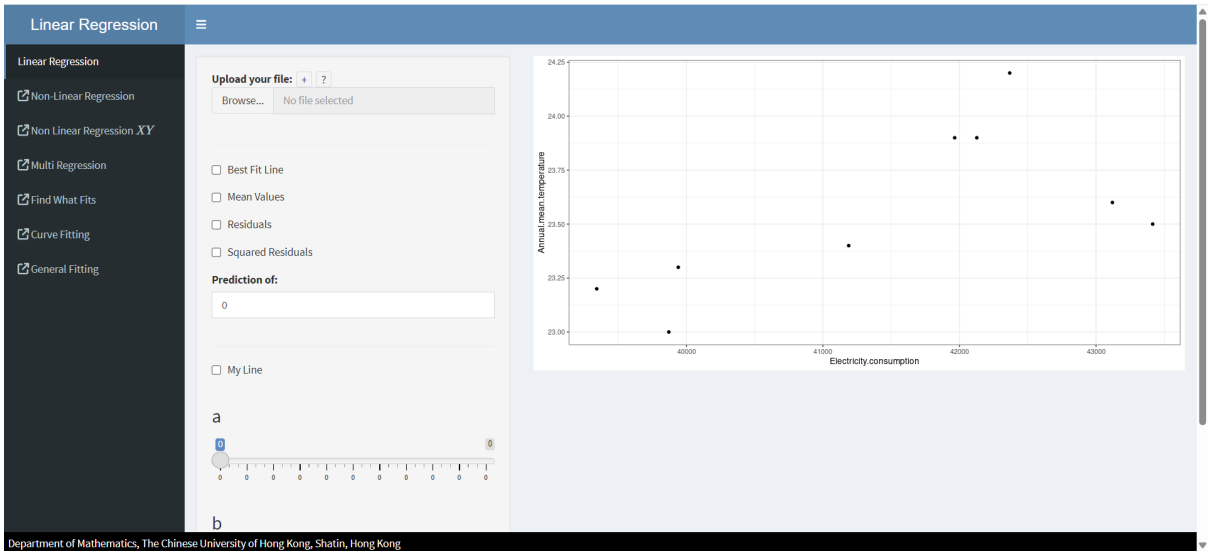
At the top left-hand corner, you can see "Upload your file", this is where we input the data. There are three different buttons.
Firstly, the "?" gives you the general guide on data input.
Secondly, the "+" allows you to input data directly. After clicking it, you can see the following:



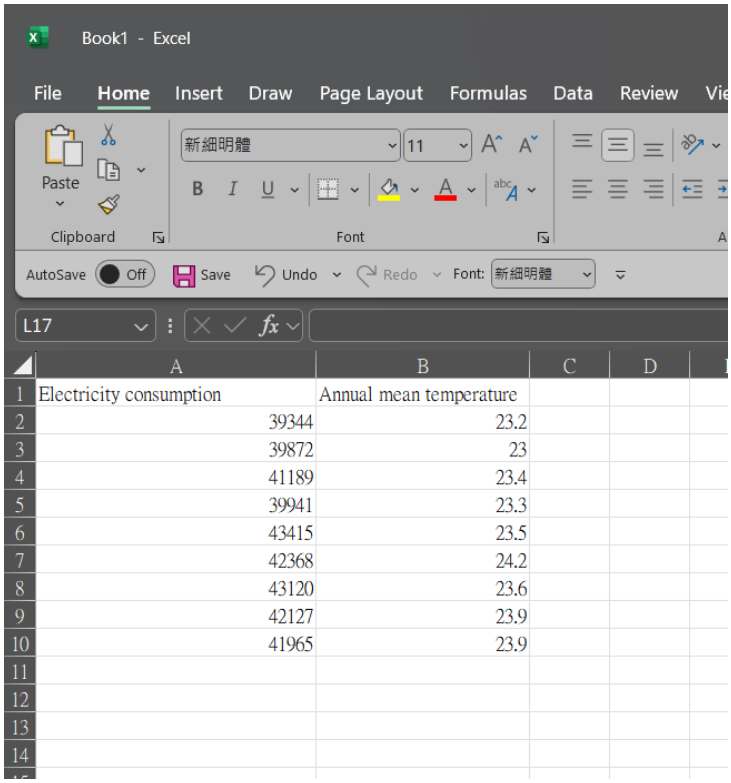
On the first line, you can input the names of the variables for the x-axis and y-axis respectively. For example, we can define the x-axis as "Electricity consumption" and the y-axis as "Annual mean temperature" to see the correlation between them. Therefore, on the first line, we type "Electricity consumption, Annual mean temperature".
For the following lines, you can simply type in the data points respectively. You should be able to have a result that looks like this:



Scroll down and press “submit”. Your data points should be shown on the screen like this:



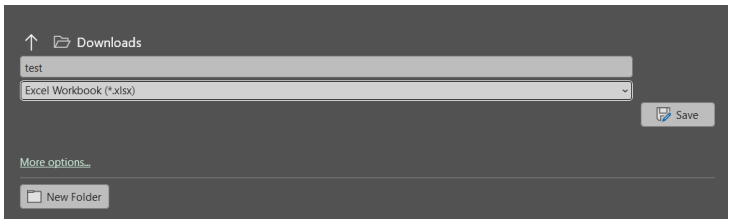
You can also upload an Excel file to the website directly. Firstly, you will need to prepare your Excel file like this:



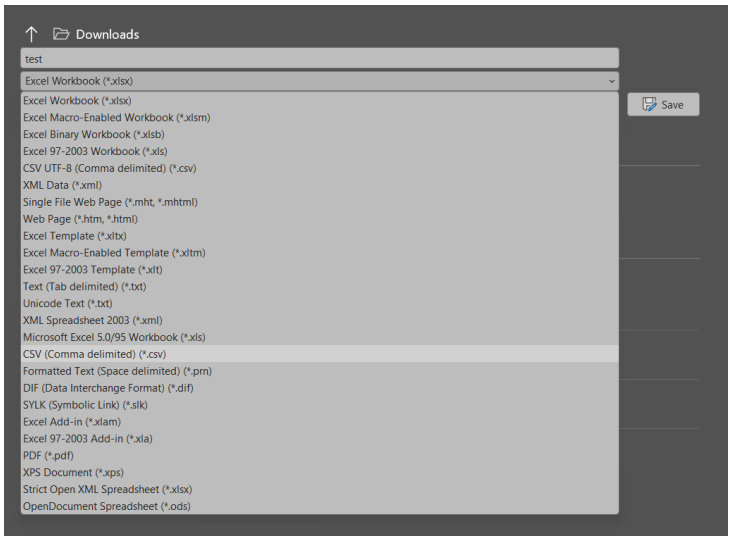
	A	B	C	D
1	Electricity consumption	Annual mean temperature		
2	39344	23.2		
3	39872	23		
4	41189	23.4		
5	39941	23.3		
6	43415	23.5		
7	42368	24.2		
8	43120	23.6		
9	42127	23.9		
10	41965	23.9		
11				
12				
13				
14				

Then go back to our website, press the “Browse...” button and look for the file stored on the computer. After uploading, you will see your data points shown on the graph.

If you face any problem regarding the upload, you may need to save your Excel file in .csv format. To do so, you can go to file → save as, and you can look for the following:



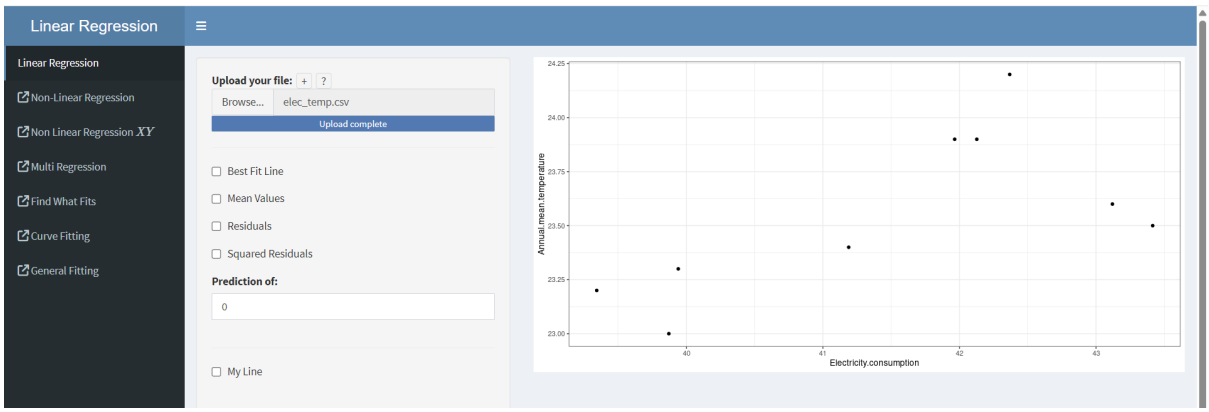
Expand the dropdown list, and look for the .csv format.



After that, you can upload the file onto our website, and it will work normally.
Remark: For the convenience of calculation, in the following demonstration, we divided the energy consumption by 1000.

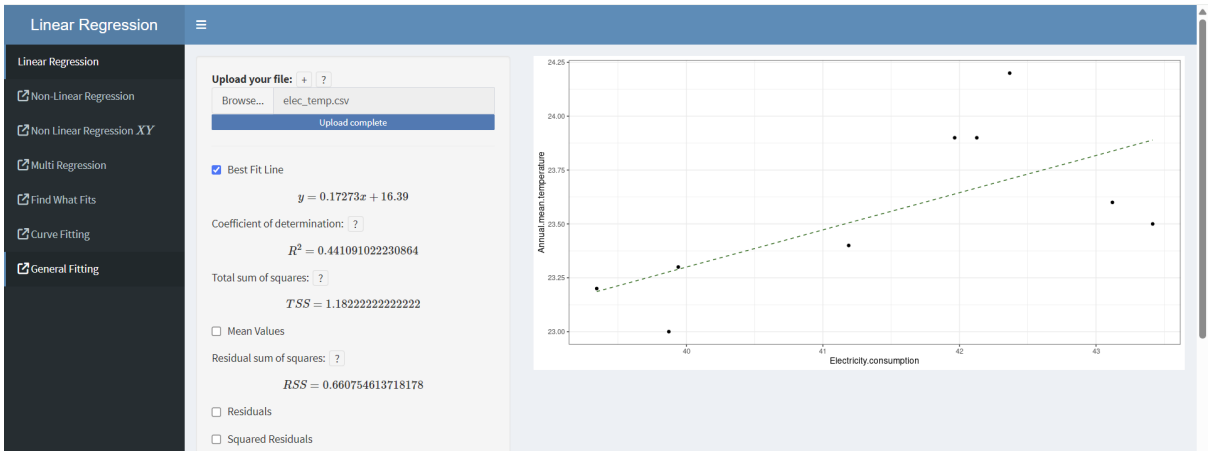
2.2 Step 2: Data analysis

After you have inputted the dataset, now we can do linear regression on it easily. Try to click different buttons on the page, and now we will briefly explain their usage.

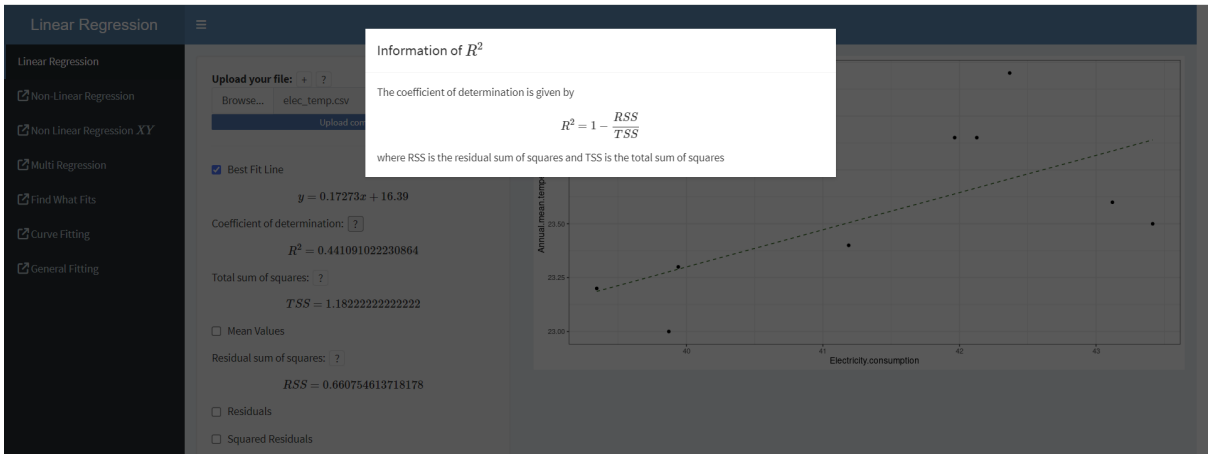


2.2.1 Best Fit Line

This directly gives you the 'best fit line', which is the linear regression of the data points. It also tells you different related information, such as the coefficient of determination, total sum of squares, and residual sum of squares. They measure how well the data points can be represented by the regression line.

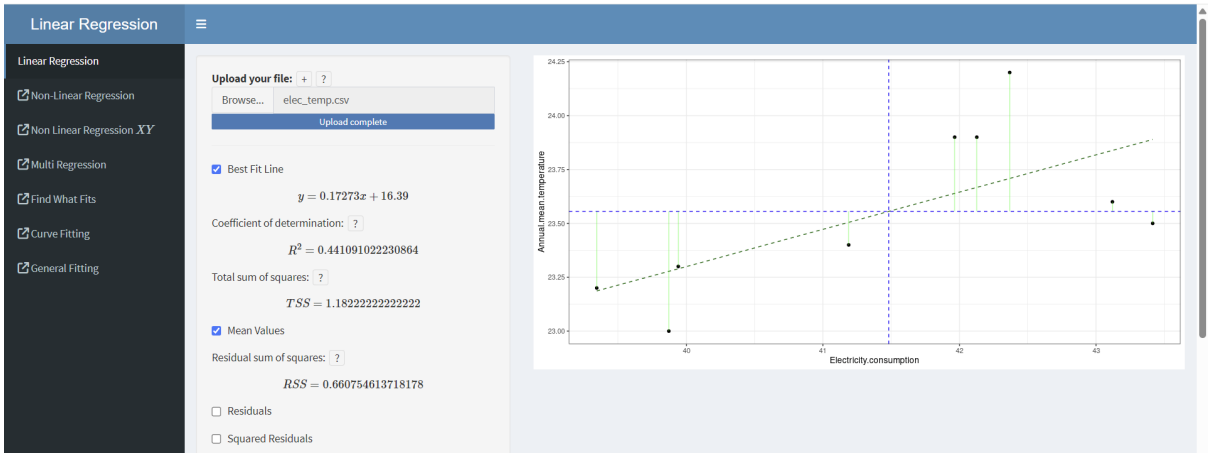


If you have any questions about the values we are showing you, of course you can always search online. But there's also a "?" next to each variable; click it, and there will be a short introduction about it.



2.2.2 Mean Values

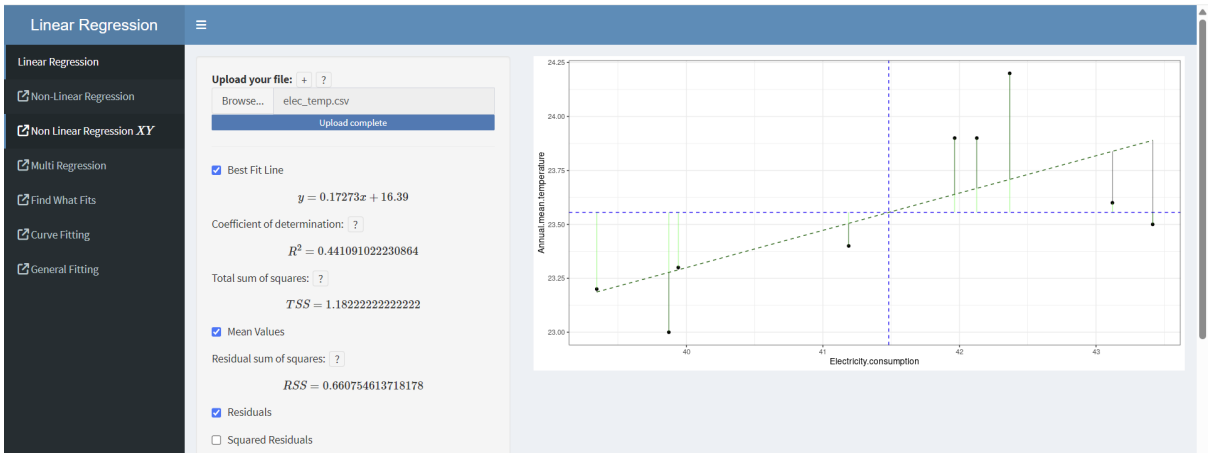
By pressing mean values, suddenly there will be a lot of horizontal and vertical lines shown on the graph. Don't panic, here is the explanation.



Firstly, the vertical and horizontal blue dashed lines show the mean of your data on the x-axis and y-axis, respectively. In our case, they are the electricity consumption and the annual mean temperature.
Then, there are many green vertical solid lines connecting our original data points with the horizontal blue line. They show the difference between each of our original data points and the mean of the data of the y-axis. This difference is important because it helps to calculate the total sum of squares.

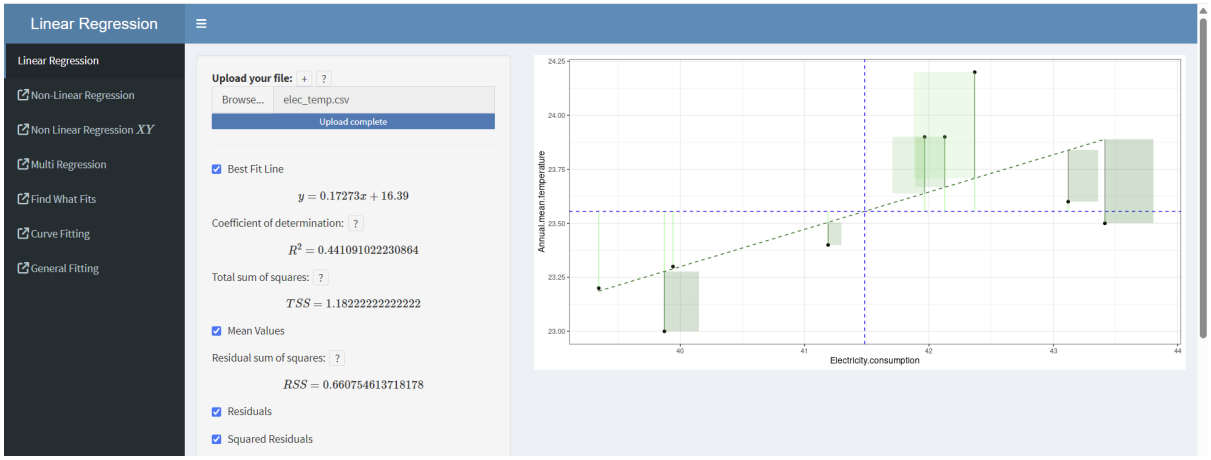
2.2.3 Residuals

This gives us another set of vertical lines, showing the difference between each of our original data points and the regression line.



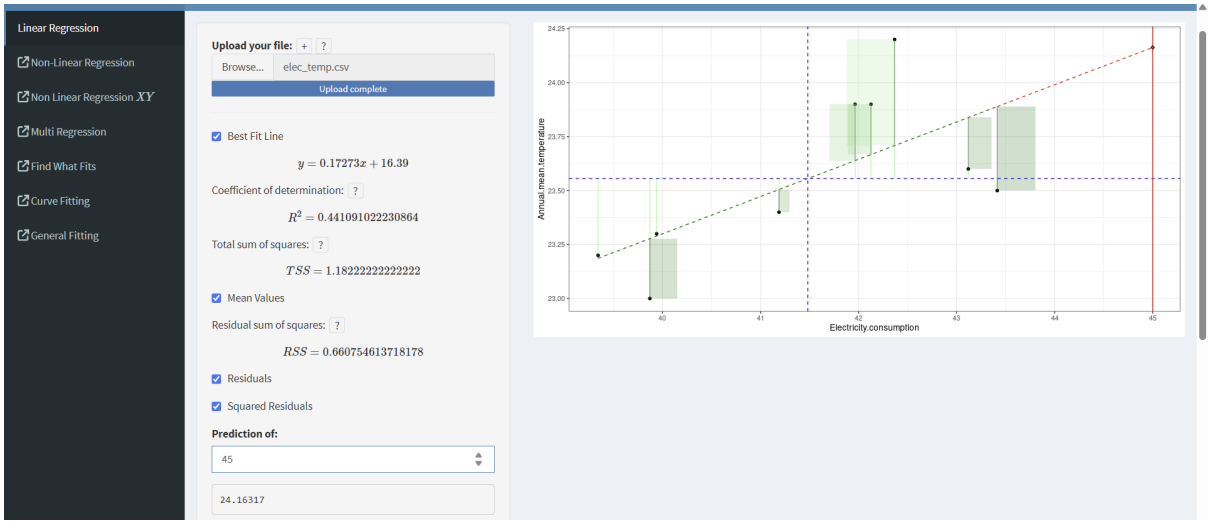
2.2.4 Squared Residuals

This function visualizes the square of residuals as the area of the rectangles. In fact, it is originally a square with side length equal to the residual, just that in the chart, the vertical and horizontal scales are different, and that makes the shape look like a rectangle.



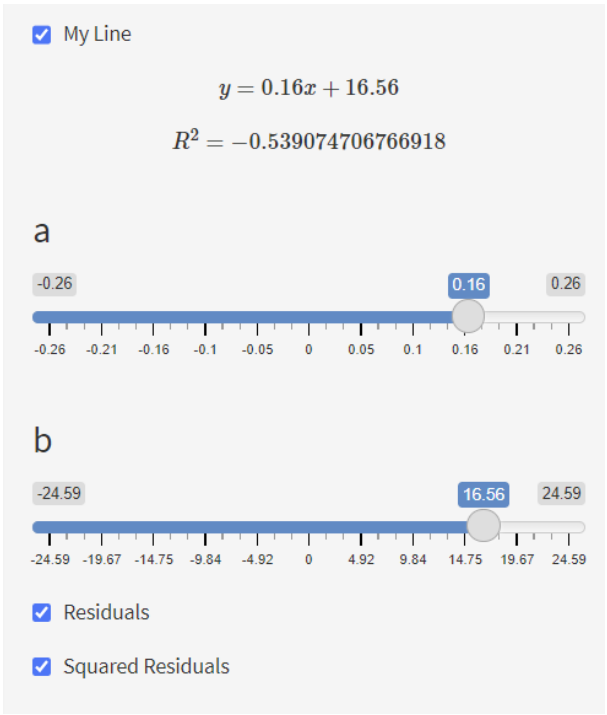
2.2.5 Prediction

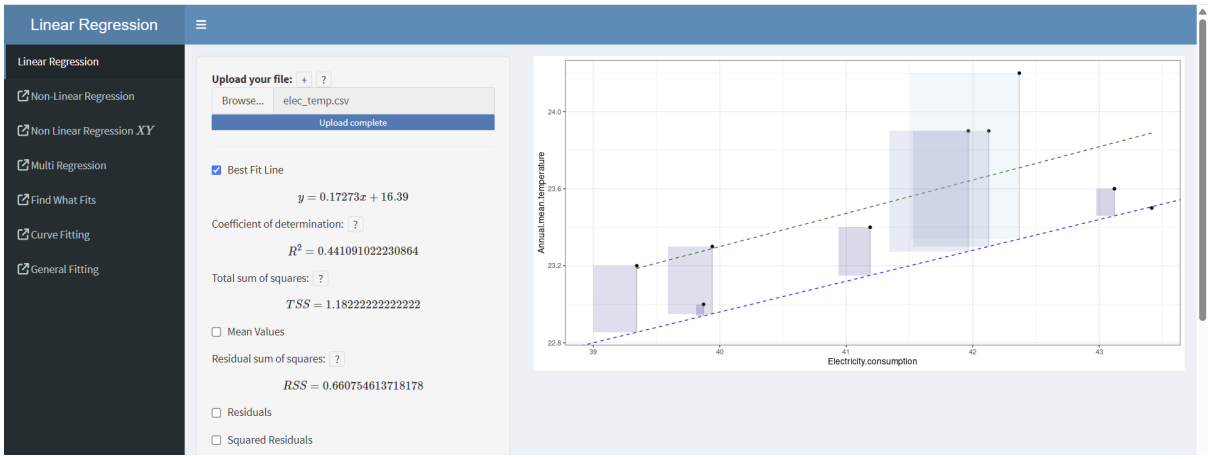
By typing any number into the box, we can know the corresponding prediction of the y-coordinate according to our regression line. You can also see how the regression line is extended, and the intersection with the red vertical line gives you the prediction of the new value.



2.2.6 My Own Line

Finally, you can even draw your own line with different slope and y-intercept so that you can compare it with the original regression line. You can also check the box of residual and squared residual, which work exactly the same as the original one above.





3 Conclusion

Linear regression is a very powerful tool in mathematical modelling. From drawing correlations to making future predictions, we usually try to go with linear regression first. We hope that this R Shiny Linear regression tool may help you with your mathematical modelling journey. Good luck!