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: March 11, 2026

1 Introduction to the Linear Regression R Shiny tool 線性迴歸R Shiny 工具簡介

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.
R 是一種功能强大的統計程式語言，讓我們能夠進行數據分析，對數學建模十分有用。更理想的是，我們利用R Shiny 建立了一系列線上工具，讓使用者即使不具備程式設計背景，也能夠執行各種數學建模任務。

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.
具體而言，線性迴歸是數學建模中的一項常見任務。以下是有關如何使用我們的線性迴歸R Shiny 工具的指引。

2 Using the Linear Regression R Shiny tool 使用線性迴歸R Shiny 工具

To use the tool, you can go to our main website:
如要使用此工具，可前往我們的主網站：

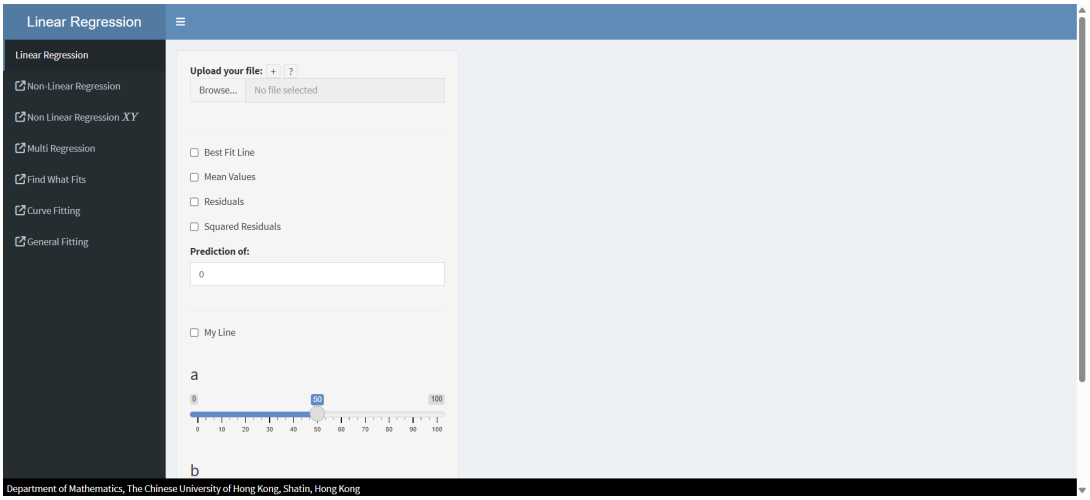
https://www.math.cuhk.edu.hk/app/mathmodel/tool.html

for the list of tools, or simply go to
瀏覽工具列表，或直接前往

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

for the Linear Regression R Shiny tool.
使用線性迴歸R Shiny 工具。

After getting into the website, you can see the following:
進入網站後，您會看到以下畫面：



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.
如果你覺得這個界面有點複雜，不用擔心！我們會逐步引導你，讓你可以使用我們的R Shiny 工具進行線性迴歸。

2.1 Step 1: Data input
步驟一：數據輸入

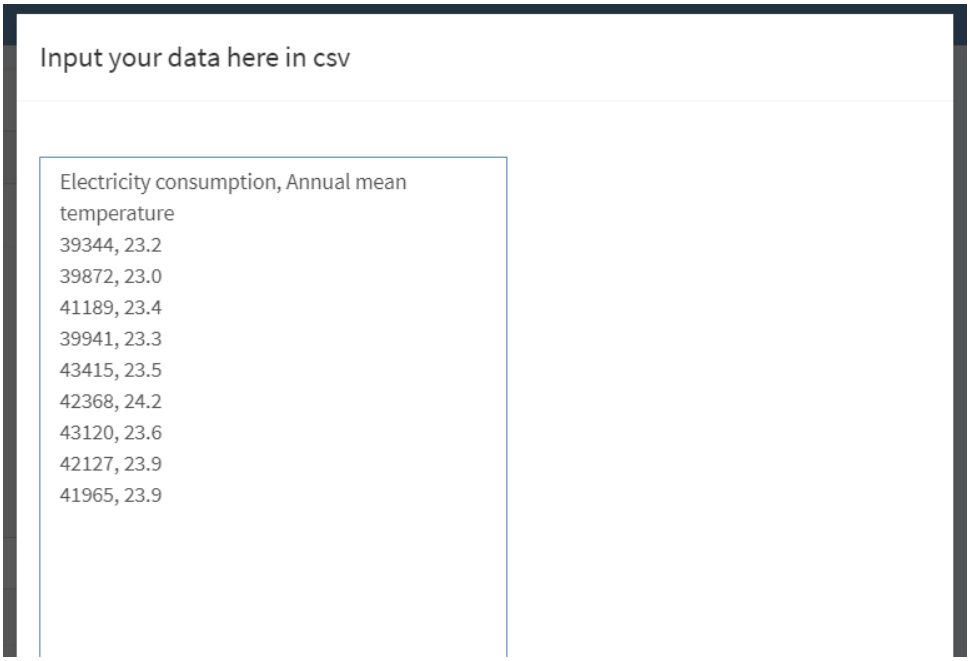
At the top left-hand corner, you can see "Upload your file", this is where we input the data. There are three different buttons.
在左上角，你會看到「Upload your file」，這就是我們輸入數據的地方。這裡有三個不同的按鈕。

Firstly, the "?" gives you the general guide on data input.
首先，「?」按鈕會提供關於數據輸入的一般指南。

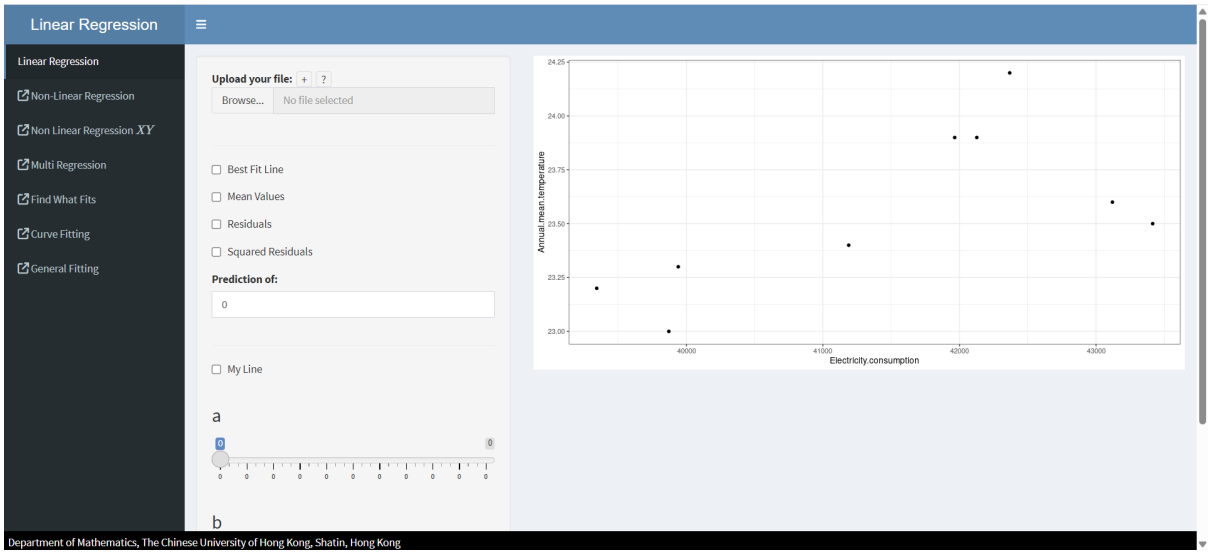
Secondly, the "+" allows you to input data directly. After clicking it, you can see the following:
其次，「+」按鈕讓你可以直接輸入數據。點擊後，你會看到以下畫面：

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".
在第一行，你可以分別輸入x 軸和y 軸的變量名稱。例如，舉例來說，我們可以將x 軸定義為「用電量」，y 軸定義為「年平均温度」，以觀察兩者之間的相關性。因此，我們在第一行輸入「Electricity consumption, Annual mean temperature」。

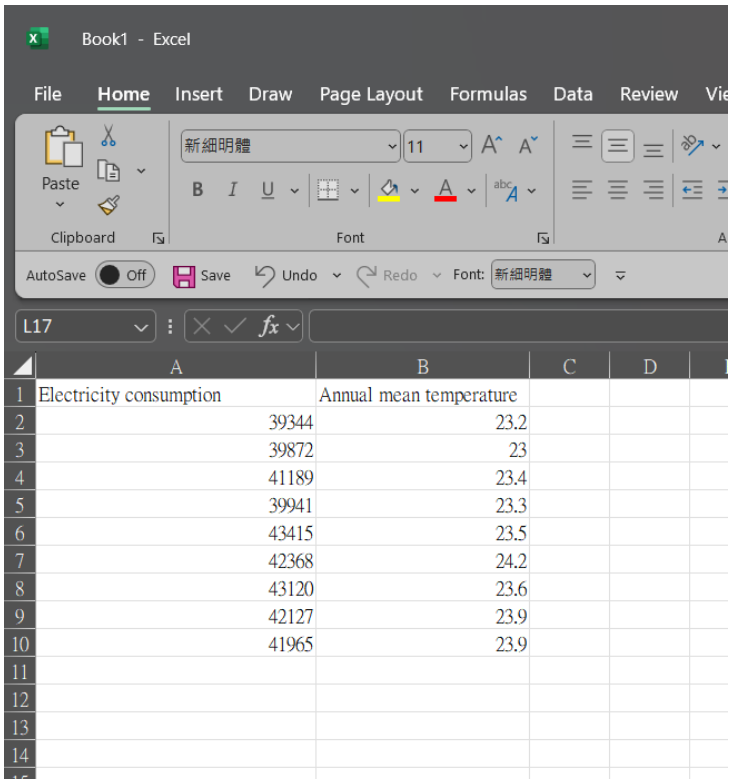
For the following lines, you can simply type in the data points respectively. You should be able to have a result that looks like this:
在接下來的行數，你可以分別輸入各個數據點。你應該會得到類似這樣的結果：



Scroll down and press "submit". Your data points should be shown on the screen like this:
向下滾動並按下「submit」，你的數據點應該會在屏幕上顯示，如下所示：



You can also upload an Excel file to the website directly. Firstly, you will need to prepare your Excel file like this:
你也可以直接將Excel 檔案上傳到網站。首先，你需要準備好Excel 檔案，格式如下：

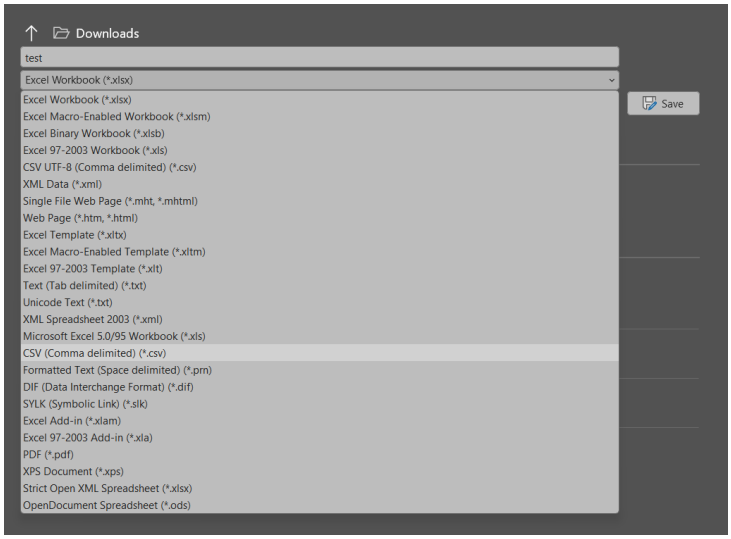


	A	B	C	D
1	Electricity consumption	Annual mean temperature		
2	39344	23.2		
3	39872	23		
4	41189	23.4		
5	39941	23.3		
6	43415	23.5		
7	42368	24.2		
8	43120	23.6		
9	42127	23.9		
10	41965	23.9		
11				
12				
13				
14				

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.
然後返回我們的網站，按下「Browse...」按鈕，尋找儲存在電腦上的檔案。上傳完成後，你會看到數據點顯示在圖表上。

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:
如果在上傳時遇到任何問題，你可能需要將Excel 檔案儲存為.csv 格式。方法是前往「檔案」→「另存新檔」，然後尋找以下選項：

Expand the dropdown list, and look for the .csv format.
展開下拉列表，尋找.csv 格式。

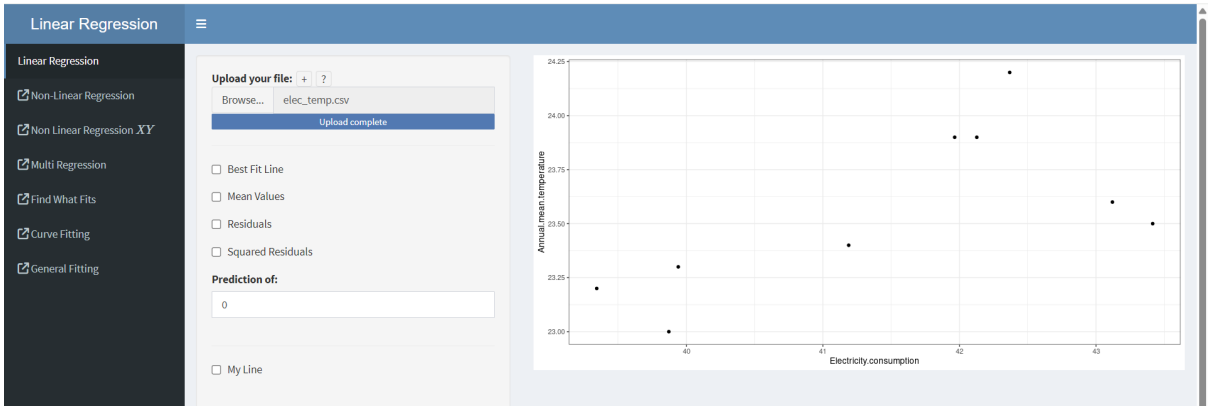


After that, you can upload the file onto our website, and it will work normally.
完成後，你可以將檔案上傳到我們的網站，它就能正常運作了。

Remark: For the convenience of calculation, in the following demonstration, we divided the energy consumption by 1000.
備註：為方便計算，在接下來的示範中，我們將用電量除以1000。

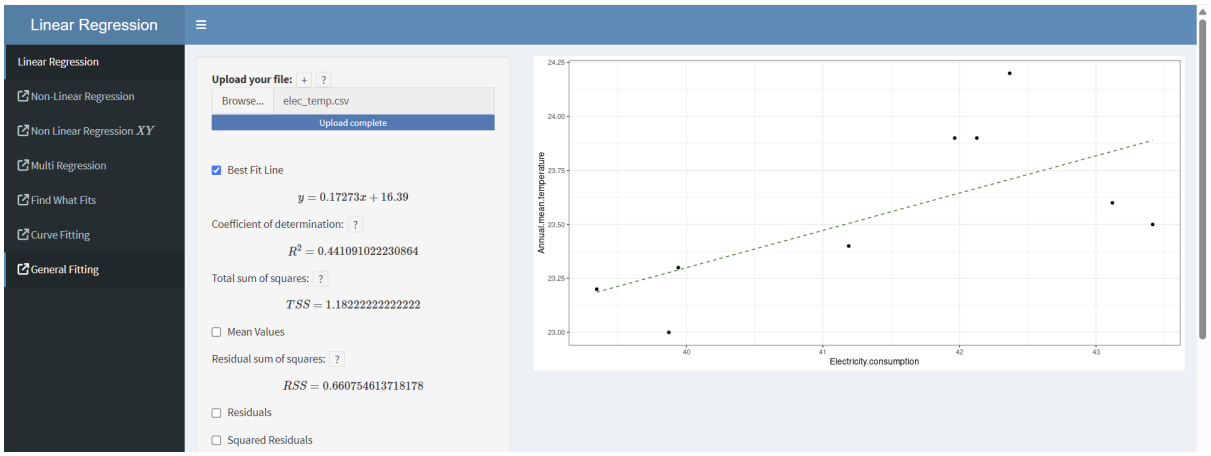
2.2 Step 2: Data analysis
步驟二：數據分析

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.
輸入數據集後，我們現在可以輕鬆地進行線性或非線性迴歸。嘗試點擊頁面上的不同按鈕，現在我們將簡要說明它們的用途。

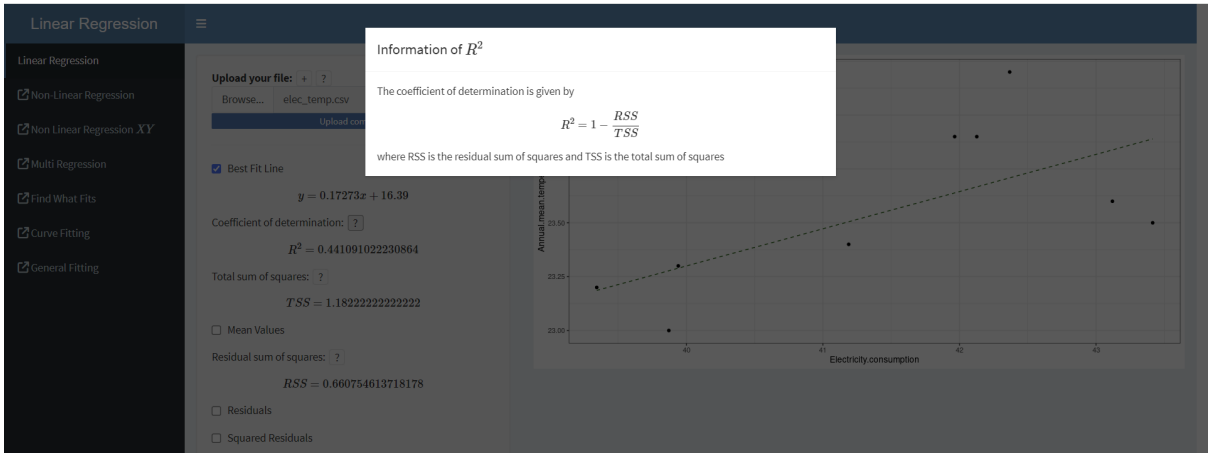


2.2.1 Best Fit Line
最佳擬合直線

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.
此功能會直接顯示「最佳擬合直線」，即數據點的線性迴歸線。它還會提供各種相關資訊，例如決定係數、總平方和及殘差平方和。這些指標用於衡量迴歸線對數據點的擬合程度。

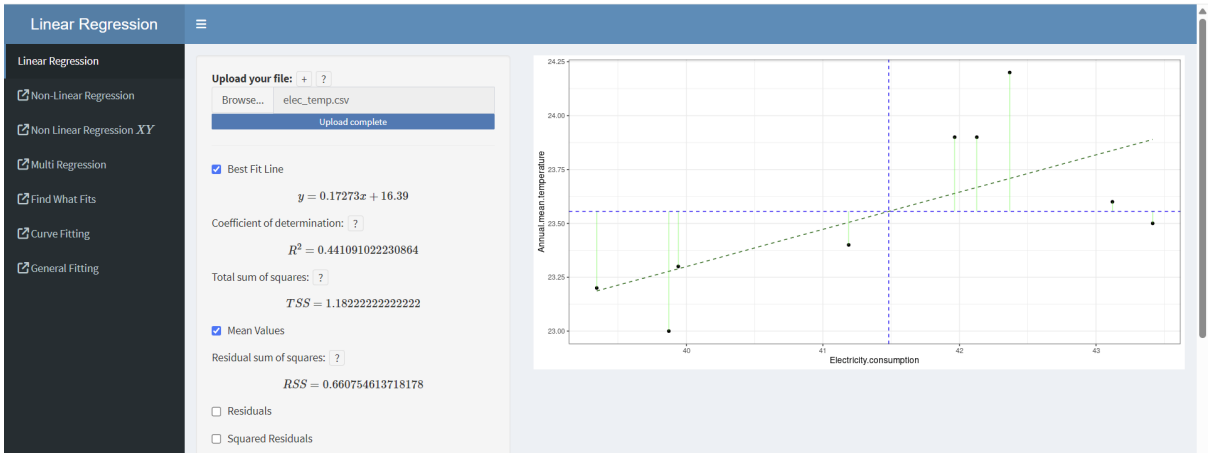


If you have any questions about the values we are showing you, of course you can always search online. But there's also a "?" next to each variable; click it, and there will be a short introduction about it.
如果你對我們顯示的數值有任何疑問，當然可以隨時在網上搜尋。不過，每個變數旁邊也有一個「?」按鈕；點擊它，便會出現相關的簡要說明。



2.2.2 Mean Values
平均值

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.
按下「平均值」按鈕後，圖表上會突然出現許多水平線和垂直線。請別慌張，以下是相關解釋。

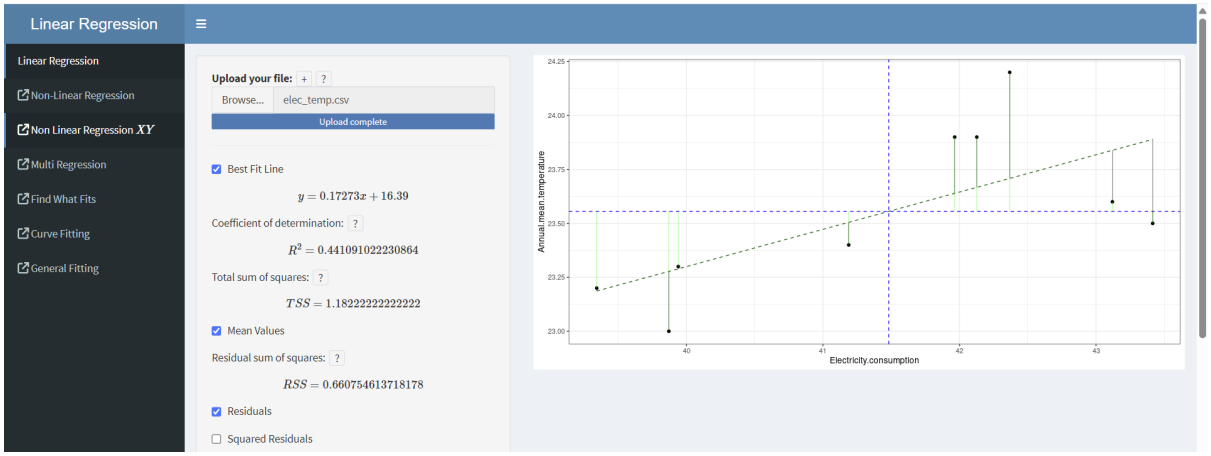


Firstly, the vertical and horizontal blue dashed lines show the mean of your data on the x-axis and y-axis, respectively. In our case, they are the electricity consumption and the annual mean temperature.
首先，藍色的垂直虛線和水平虛線分別顯示x 軸和y 軸數據的平均值。在我們的情況中，它們分別代表用電量和年平均溫度。

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.
然後，許多綠色的垂直實線將我們的原始數據點與藍色水平線連接起來。這些線段顯示每個原始數據點與y 軸數據平均值之間的差異。這個差異相當重要，因為它有助於計算總平方和。

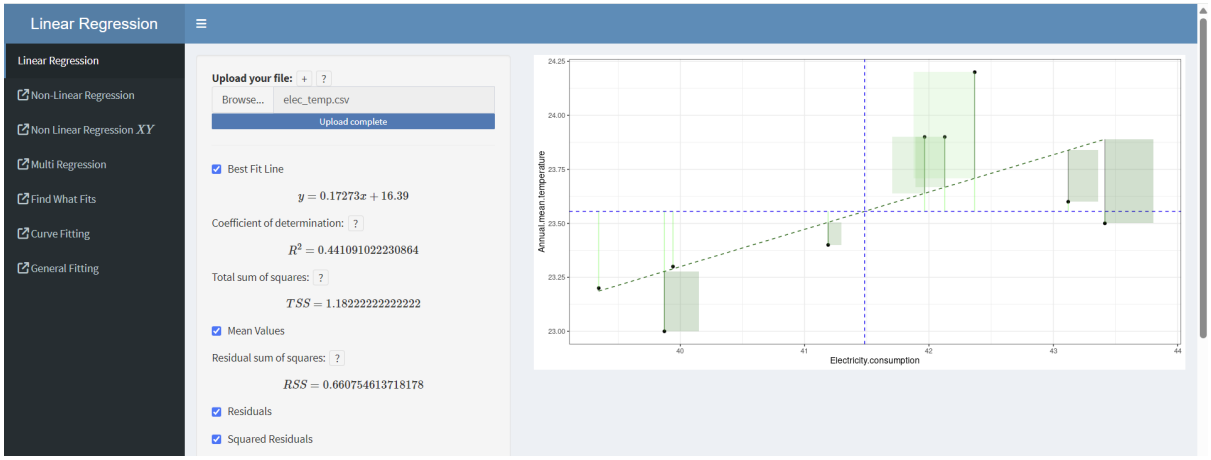
2.2.3 Residuals
殘差

This gives us another set of vertical lines, showing the difference between each of our original data points and the regression line.
此功能會顯示另一組垂直線，表示每個原始數據點與迴歸線之間的差異。



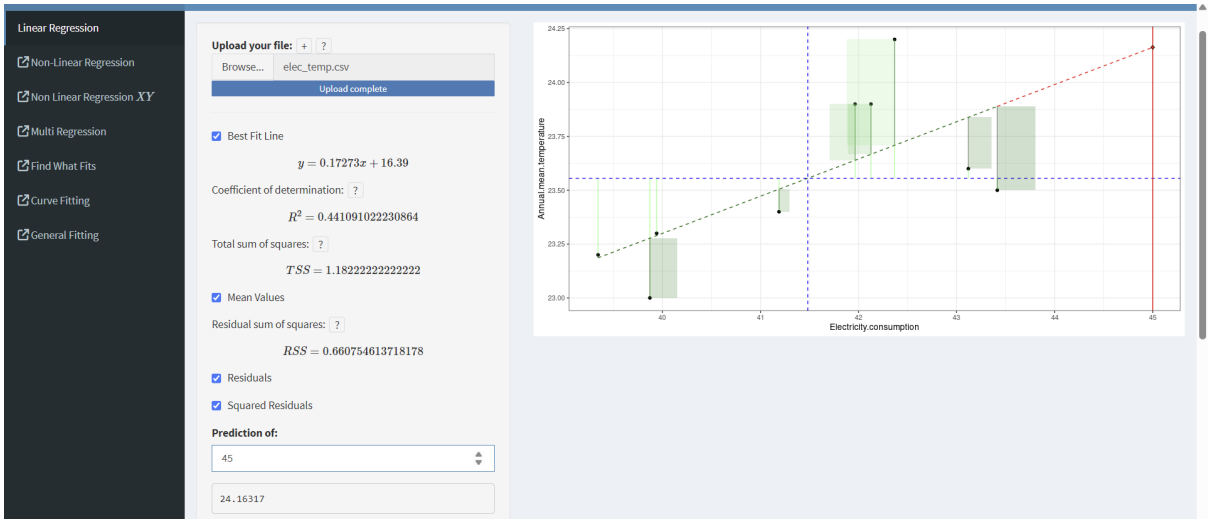
2.2.4 Squared Residuals
平方殘差

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.
此功能將殘差的平方以矩形面積的形式可視化。實際上，它原本是一個邊長等於殘差的正方形，只是由於圖表中垂直和水平的比例尺不同，使得形狀看起來像一個矩形。



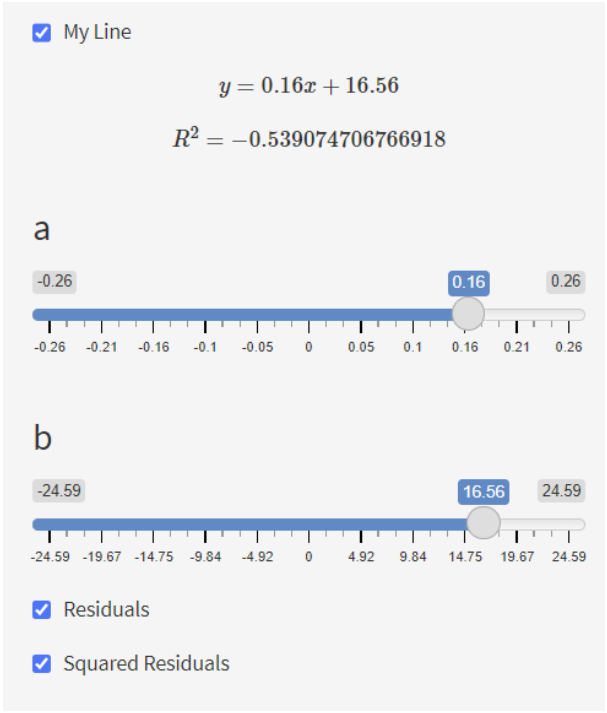
2.2.5 Prediction
預測

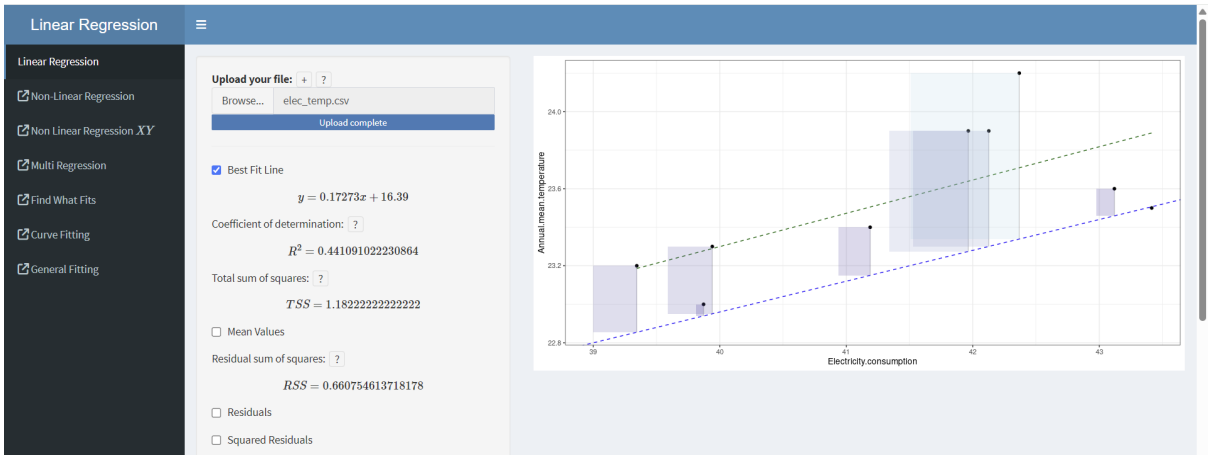
By typing any number into the box, we can know the corresponding prediction of the y-coordinate according to our regression line. You can also see how the regression line is extended, and the intersection with the red vertical line gives you the prediction of the new value.
在方框中輸入任何數值，我們便能根據迴歸線得知對應的y 坐標預測值。你還可以觀察迴歸線如何延伸，而與紅色垂直線的交點即為新值的預測結果。



2.2.6 My Own Line
自訂線條

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.
最後，你甚至可以自行繪製具有不同斜率和y 截距的線條，以便與原有的迴歸線進行比較。你也可以勾選「殘差」和「平方殘差」的選框，其功能與上方原有的選項完全相同。





3 Conclusion
總結

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!

線性迴歸是數學建模中極具價值的工具。無論是繪製相關性還是進行未來預測，我們通常會首先嘗試使用線性迴歸。我們衷心希望這個R Shiny 線性迴歸工具能夠在你的數學建模探索旅程中提供協助。祝你一切順利！