Rocco Vulpis

April 25, 2024

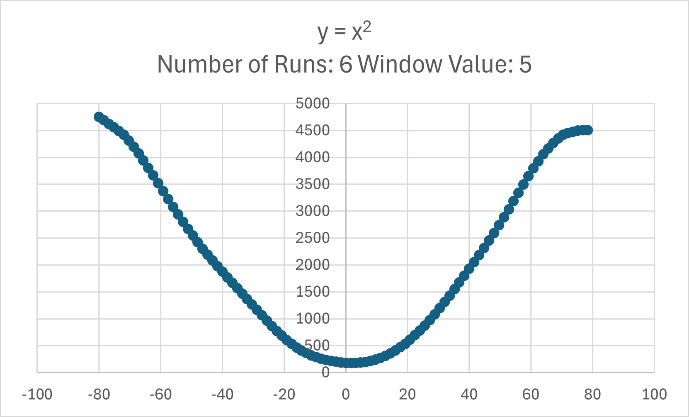
Smoothing Data using Java and Plotting in Excel

In the data smoothing portion of this project, we average the y-values of a mathematical function using a technique known as "smoothing." This process involves modifying data points read from a CSV file named "saltedData.csv" and exporting the smoothed data to a new CSV file, "smoothedData.csv." The graph created in Microsoft Excel with this data illustrates the effect of smoothing, resulting in a closer representation of the original quadratic function.

The first step is to create the function from the ExcelSmootherTester class. This process starts by creating an ExcelSmoother object and calling and passing parameters to its **run** method. The first parameter is **numberOfRuns** (The number of times the smoother will iterate over the **yValues** ArrayList) and **windowValue** (The size of the window used for smoothing, determining the range of values to average.). Inside the **run** method in the ExcelSmoother class, the **readFile** method is called. Similar to the salting phase, the BufferedReader class is utilized to read in the values from “saltedData.csv”. The values from each line are parsed into doubles and stored into two ArrayLists: one for x-values and one for y-values.

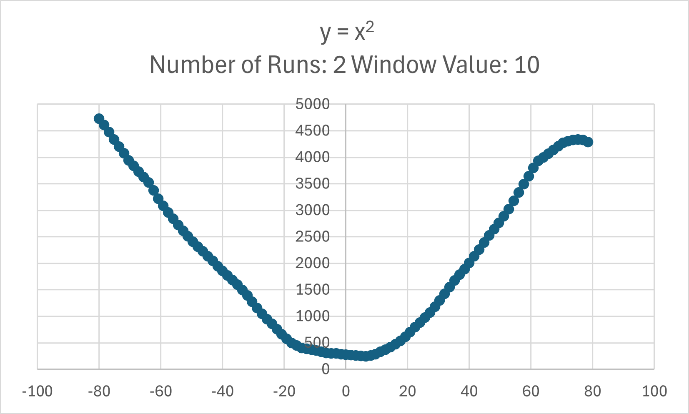
Once the data is imported and stored, the **smoothValues** method is called. This method implements a moving average technique to smooth the y-values. A specified window size determines the range of values to be averaged for each point. Within this window, the y-values are averaged, and the result replaced the original y-value in the *yValues* ArrayList. The size of the window is taken in as a parameter in the class’s **run** method.

The final stage involves writing the data back to a CSV file using the **writeToFile** method. This method writes the smoothed x and y values to “smoothedData.csv” utilizing the BufferedWriter class. The generated CSV file is written and stored in the default workspace directory. In Microsoft Excel, the x and y values are stored into columns A and B of the spreadsheet. Opening the new CSV in Excel and using its charting tools, a scatter plot graph can be generated. This graph displays a pattern of points that showcase a smoother transition, compared to the salted data. The variations in the graph are significantly reduced, showing the effect of the smoothing algorithm. Below are screenshots illustrating the Excel graphs with the smoothed data points.

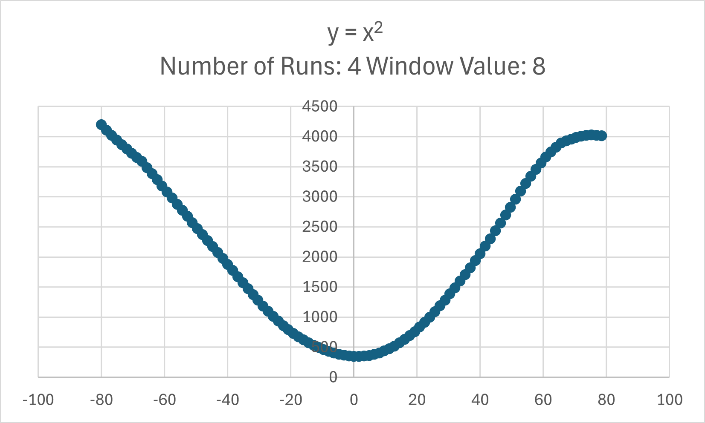


The result of y = x2 after 6 runs and smoothing the y values with a window value of 5.

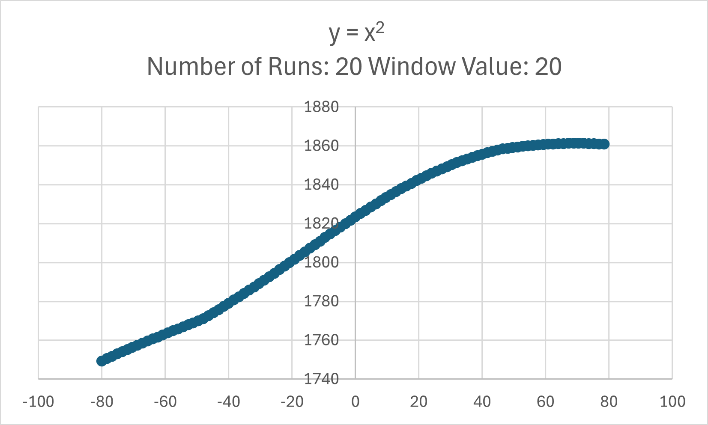
Here are additional examples of graphs created with different amounts of smoothing iterations and window values.



The result of y = x2 after 2 runs and smoothing the y values with a window value of 10.



The result of y = x2 after 4 runs and smoothing the y values with a window value of 8.



The result of y = x² after 20 runs and smoothing the y values with a window size of 20. This is an example of what happens when y values are smoothed excessively. The graph no longer resembles the original function y = x².