Rocco Vulpis

December 14, 2023

Smoothing Data using Java and Plotting in Excel

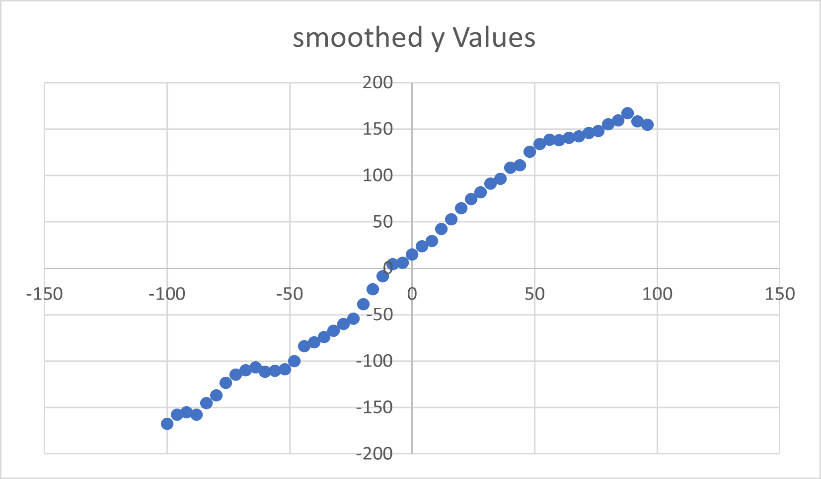
In the data smoothing portion of this project, y-values of a mathematical function were averaged using a technique known as 'smoothing'. This was achieved by modifying the data points read from a CSV file, 'saltedData.CSV', and exporting the smoothed data to a new CSV file, 'smoothedData.CSV'. The graph created in Microsoft Excel with this data illustrates the effect of smoothing, resulting in a similar representation of the original slope intercept function.

The process started with importing the function's values through the readFile method. Similar to the salting phase, the BufferedReader class was employed for sequential reading of 'saltedData.CSV'. The values from each line were parsed into doubles and stored in two ArrayLists: one for x-values and one for y-values.

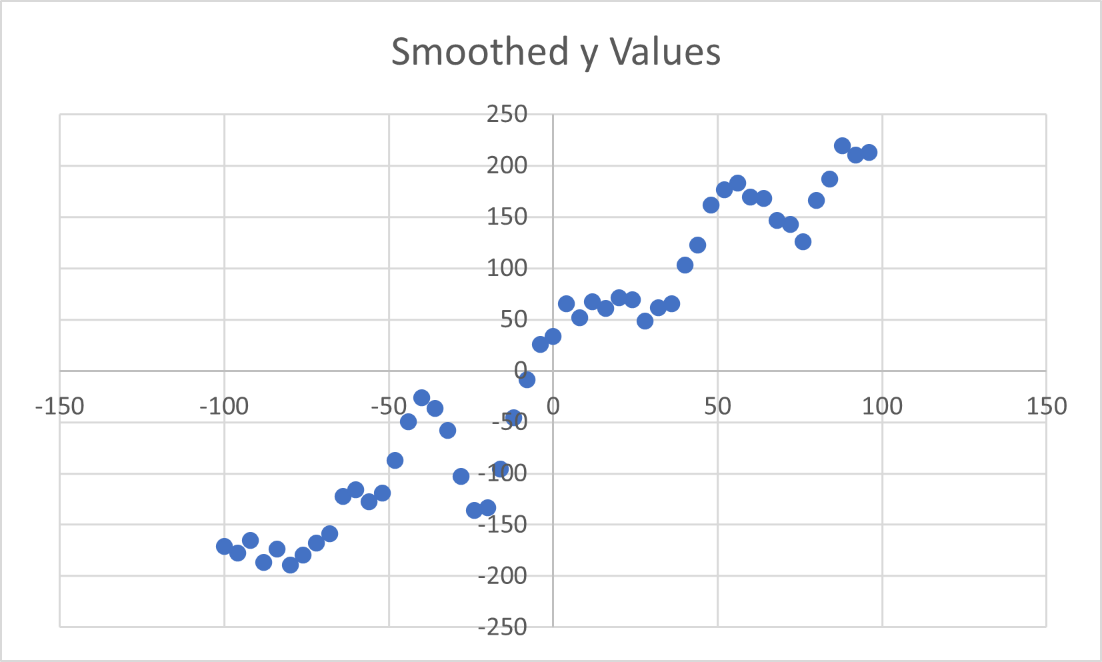
Once the data was imported and stored, the smoothValues method was called. This method implemented a moving average technique to smooth the y-values. A specified window size determined the range of values to be averaged for each point. Within this window, y-values were averaged, and the result replaced 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 involved writing the data back to a CSV file using the writeToFile method. This method wrote the smoothed x and y values to 'smoothedData.CSV' utilizing the BufferedWriter class. The generated CSV file was written and stored in the default workspace directory. In Microsoft Excel, the x and y values were stored into columns A and B of the spreadsheet.

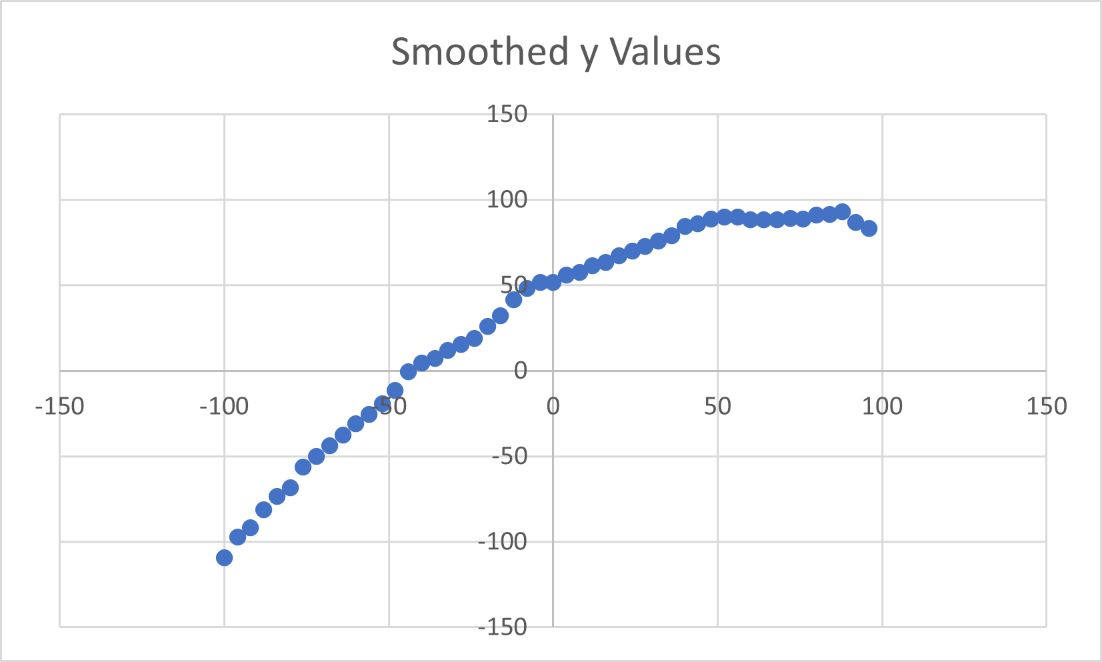
Opening the new CSV in Excel and using its charting tools, a scatter plot graph was generated. This graph displayed a pattern of points that showcased a smoother transition, compared to the salted data. The variations in the graph were significantly reduced, showing the effect of the smoothing algorithm. Below are screenshots illustrating the Excel graphs with the smoothed data points.



The result of smoothing after the y values smoothed with a window value of 10.



The result of smoothing after the y values smoothed with a window value of 2.



The result of smoothing after the y values smoothed with a window value of 25.