CS306 Data Analysis and Visualization

Lab. 2 Temperature Data Analysis

1. In previous lab, we found out several statistical measures and plotted pdf (Normalized histogram). In this lab, we will use the same dataset of temperature and do the following:-

1. Compute and plot cumulative distribution function (cdf) from previously plotted pdfs.
2. Calculate Pearson correlation coefficient (r) between T_{max} and T_{min} .
3. Interpret the r^2 value you got.
4. Plot the scatter plot from of $T_{max} \rightarrow T_{min.}$
5. Now plot the smoothed scatter plot without sorting and observe the plot.
6. Plot smoothed scatter plot with sorting and observe the difference with the plot of above. State how smoothed scatter plot is in some ways better than usual scatter plot.

Final smoothed scatter plot should look something like the above