In this exercise, you are given a set of individual sensor tag files with some variations in dates and format. There are 2 parts to this exercise:

Part 1

For the first part, you are tasked to perform data wrangling.

* 1. Load the files into a single suitable data structure
  2. Pivot the data to create tag names as columns, timestamp as index
  3. Filter for good quality
  4. Forward fill for missing timestamp values
  5. Use the dataset to build data normalization (normalize to range 0 to 1) model for each sensor

Part 2

In this final part, you are tasked to create a simple frontend framework. Use Docker/Kubernetes. Please write codes in object-oriented style.

Data Visualization Module:

1. Use dropdown selector to select a particular sensor name. e.g. Sensor 1
2. Load the selected sensor data and plot time-series charts with proper handling of timestamp on x-axis. Please use data caching mechanism for the visualization performance for repeated sensor selection.

We would prefer if you can implement the code with Pyspark.

Please develop a data analytics solution to the problem and present it to the team in a suitable format you desire (notebook/web application, Docker etc.) detailing the methodology, performance consideration, and justification for your solution.

Please note that the point of this exercise is to evaluate your understanding of data engineering, the capability in developing such solution, and ingenuity of your proposed solution.