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

Part 1

For the first part, you are tasked to create a simple backend framework with modules of load data, transform raw data, featurize data, train model and serve model. Use Docker/Kubernetes for all steps. Please write codes in object-oriented style.

* 1. Load Data Module:
     + Load the given data files with different format
     + Combine data into a single suitable data structure (outer join)
  2. Raw Data Transform Module:
     + Pivot the combined data to create tag names as columns, timestamp as index
     + Replace bad quality data as NaN
     + Forward fill for missing data
     + Split the combine data into training (70%) and testing/serving (30%) dataset
  3. Featurize Data Module:
     + Use the training dataset to build data normalization (normalize to range 0 to 1) model for each sensor
     + Use the respective data normalization model to transform raw data (both training and testing dataset) and store the normalized data
  4. Train AI/ML Model Module:
     + Use sensor 2 to sensor 5 as modeling independent features and sensor 1 as dependent target/label to build an AI/ML model using the **NORMALIZED TRAINING** dataset (range of 0 to 1). Note that the model’s accuracy is not important here and will not be used as an assessment metric in this exercise.
     + Persist the trained model
  5. Serve AI/ML Model Module:
     + Load the **NORMALIZED TESTING** dataset and perform model inference using the trained AI/ML model on docker/kubernetes using any open source tools
     + Store the inference/prediction data
     + Log the requests that are sent to the model serving pipeline on another container running a postgresdb instance or local db (Eg.postgresDB)

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.

1. Data Visualization Module:
   * + Use dropdown selector to select a particular sensor name. e.g. Sensor 1
     + Load the selected sensor’s inference/prediction 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.