The Python code is being developed using ‘modular’ approach. Hence, I have created a main pipeline script (**TurbinesPipelineMain.py**) which calls 3 modules for cleaning data, calculating summary stats and to Identify anomalies etc

I have stored processed data in to CSV files as I don’t have a database environment setup in my laptop.

**The Key Assumptions are: -**

1.Files will land in the input folder

2. No special characters as part of the data

3.Timestamp field will be in format of ‘DD-MM-YYYY HH:MM’

4.All the files in the input folder will be processed on a daily basis

5.Processing data in batch mode

**Others: -**

Extract the .zip file (TurbinesPipeline.zip), navigate to Turbines path and run using

'py Anchor.py'

Main Pipeline Python Script Name: **TurbinesPipelineMain.py**

Cleanse the data module Python script name: - **CleanseData.py**

Calculate Summary Statistics module Python script name: - **SummaryStats.py**

Identify Anomalies module Python script name: **IdentifyAnomalies.py**