# Problem Statement:

The [link](https://docs.microsoft.com/en-us/azure/machine-learning/team-data-science-process/cortana-analytics-playbook-predictive-maintenance) explains one of the many approaches that we take to create predictive maintenance application. For the problem, please use the dataset downloaded from [this Azure ML Studio Experiment](https://gallery.cortanaintelligence.com/Experiment/Predictive-Maintenance-Modelling-Guide-Data-Sets-1). The R notebook from this [link](https://gallery.cortanaintelligence.com/Notebook/Predictive-Maintenance-Modelling-Guide-R-Notebook-1) has the entire approach coded in R notebook for this data if you need any reference architecture.

# Evaluation:

Initial proposal is to use precision and recall as evaluation metric if this is solved as a classification/regression problem. Important CTQ for the PoC is to identify as many failures as possible before it happens, “earlier the better” in time. Minimise false positives and false negatives. We can agree upon an evaluation criteria during our initial discussion before you start working on the POC.

## Training and test dataset:

The dataset comes with 100 assets/machines, its description, sensor data for couple of years, errors data with timestamp, maintenance data for these machines with timestamp etc. The dataset can be split in two ways.

1. Along timeline for all assets as training and test period. Identify failures for all assets during test period
2. Across assets. Use complete data available for 70% of assets and test the performance of algorithm for remaining 30%