Objective:

1. Estimate the power system risk index
2. Define the operational state

Questions:

* What kind of data are we looking at?
* What are the trends in the data?
  + Which timeframe do we perform analysis on?
  + Use Mann-Kendall test to find monotonic trends
* What is the signal to noise ratio?
* How do we deal with missing data?
  + Imputation module SKLearn
  + MIDASpy algorithm (deep learning)
* What is the best method for feature scaling and normalization?
* How is the dimensionality reduced?
  + Do we use PCA (unsupervised) or LDA (supervised)?
* How do we split the data?
  + Training, Validation, Testing data methodology?
  + Training Testing methodology?