- First task: Get more familiar with python:

  Examples (List operations, string operations, functions, for loops, numpy arrays, pandas data frames…etc.)

- Create the following functions:
  1- A function that removes nan values from time series.
  2- A function that returns the shape of time series.
  3- Create a new column in our dataset that has the shape of each time series after removing NaN values from it.
  4- Create a function that takes as input a desired length (example: 4096) and returns the rows from the previous data frame created in task 3 that has that same length.
  5- A function that takes as input a fault name and returns its counts in our dataset, example output: (Do this only using list and string operations and for loops and not using pandas built in functions)

Fault name	Total count	Total count in motor	Total count in machine
“LXC”	200	80	120

- Second task: Data size required research.
  1- Do research on the internet for known deep learning architecture.
  2- How can we get an idea of how much data is required “theoretically” to train a specific model architecture?
  3- Do research on known commercially available deep learning models.
  4- How much data “was used” in training those models?
  5- What are methods of increasing data size in cases where datasets are small (hint: data augmentation)?