EGB103 Assignment 2 Checklist

- [ ] Everything should be included in a **single** Jupyter notebook.
- [ ] The raw CSV data may require "**cleaning**" before it can be processed
- [ ] We are only interested in looking at these quantities when the chiller is **actually operating** (kWE > 0).
- [ ] Your first task is to conduct research to **determine the formula** for deriving the kWR (and COP).
- [ ] Having **calculated** those quantities (kWE, kWR and COP), your job is to investigate how those quantities tend to **vary** over time:
  - [ ] hour of the day
  - [ ] day of the week
  - [ ] month of the year
- [ ] Use appropriate **plots** to visualize all relationships.
- [ ] Add **markdown** to summarize your observations/conclusions for each plot.
- [ ] How does the **percentage** of time that chiller 1 is operating **vary** depending on the hour of the day, the day of the week etc?
- [ ] Do a detailed **analysis** showing how the system performed over the 24 hours of your actual birthday. Did you detect any **atypical** behaviour?
- [ ] **Examine** the relationship between COP and the cooling load (kWR) and identify the load that tends to produce the **best** COP.
- [ ] Also identify any other statistical **correlations** between any of the above variables.
- [ ] The Python code included should follow **best practices** as outline in the lectures, including
  - [ ] Using well chosen identifier **names**
  - [ ] Writing **clear** simple code
  - [ ] **Not repeating** yourself.
- [ ] All data processing should be done using the **Pandas** library and should make use of the following Pandas features:
  - [ ] **Reading** input data files.
  - [ ] **Parsing** dates.
  - [ ] Selecting appropriate column(s) of a Data Frame to act as the **index**.
  - [ ] **Computing** new **columns**.
  - [ ] Using Python **functions** to compute new columns.
  - [ ] **Filtering** rows by condition.
  - [ ] At least two different kinds of Matplotlib **plots**.
  - [ ] **Group By** to explore relationships between variables.
  - [ ] **Error bars** in plots to show variation (e.g., standard deviation).
  - [ ] All plots should be appropriately **titled**, and axis appropriately **labelled**.
  - [ ] User **friendly** axis labels, e.g., Mon, Tue, Wed or Jan, Feb, Mar.
  - [ ] Appropriately **sized** figures that are large enough to easy read.
  - [ ] **Binning** of continuous values so as to investigate relationships between continuous variables.
  - [ ] Using Python functions to compute **aggregate** information for groups of rows.
- [ ] **DO NOT** use any other programming language, library or system (such as R, MATLAB or Excel).