EGB103 Assignment 2
Data Processing using Python and Pandas

Individual only (strictly no group work or collaboration allowed)
Due: Friday 20th May (Week 10)
Worth: 30%

Engineers at QUT's Gardens Point campus use a *Building Management System* (BMS) to monitor and control equipment across the campus, including air conditioning, ventilation, pool, blinds, lighting, fire, medical gases and many other aspects.

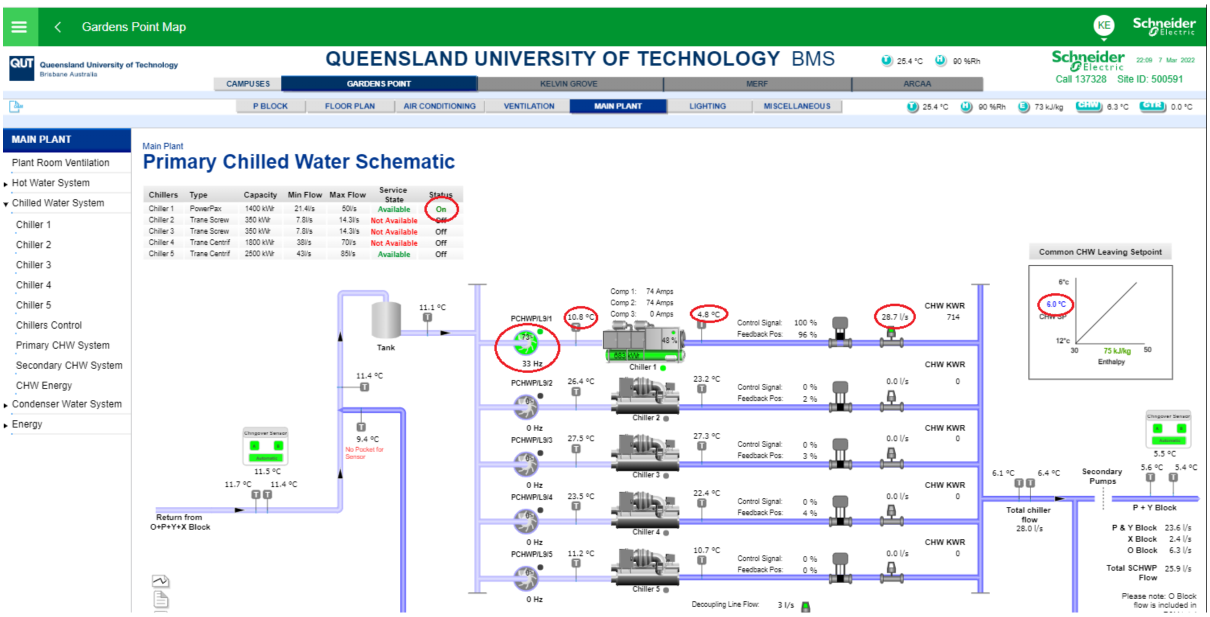


Figure 1: Building Management System (BMS)

For this assignment you will create a Juptyer notebook to explore Engineering data from the P block *Chilled Water System* which is part of the Air Conditioning system. The chilled water system in P block consists of 5 separate *chillers* that operate in parallel (see Figure 1). Individual *chillers* are automatically turned on and off as the cooling load changes throughout the day and year. The number of chillers operating at any given time will therefore vary, sometimes just one chiller is sufficient to meet the cooling demand. We see in Figure 1 that only Chiller 1 was operating at that time. The data we will be exploring is from *Chiller 1*, which is the most efficient of the 5 chillers and therefore used more often than the other chillers.

Figure 2: P Block, Chiller 1

There are two completely *separate water systems*, each is a closed circuit with its own pipes and pumps. The role of the chillers is to cool water in the *chilled water system* as it passes through the chillers. This chilled water is then circulated throughout the building to cool the air that flows through the air conditioning ducts. Each chiller is designed to *transfer heat* from the *chilled water (CHW) system* to a separate *condenser water (CW) system*. Water from the condenser water system is pumped to cooling towers on the roof where it is cooled by the air, before returning to the chillers.

The Engineers controlling the system select a target CHW Leaving Temperature. We can see in the top right side of Figure 1 that the Common CHW Leaving Setpoint is 6 °C. We can also see in Figure 1 that the chilled water is currently entering Chiller 1 at 10.8 °C and is leaving Chiller 1 having been cooled to 4.8 °C. The condenser motors inside the chillers automatically adjust to try to achieve the target CHW Leaving setpoint. The CHW pumps connected to the chillers automatically adjust to vary the water flow rate (measured in Litres per Second) through the chiller. We can see in Figure 1 that the pump attached to Chiller 1 is currently working at 73% of its capacity, spinning at 33Hz and creating a flow rate through Chiller 1 of 28.7 l/s.

The total flow achieved by all 5 chillers working in parallel is automatically adjusted by the system to match the cooling demand of the building. On a hot day or when lots of warm bodies are inside the building, the air conditioning system will need to work harder to cool the building to a constant target temperature. Since the chilled water leaves the chillers at approximately a constant target CHW Leaving Temperature, a higher building cooling demand is met by a higher net water flow rate achieved by the coolers collectively. The amount of electrical power consumed by each chiller will vary depending on how hard it needs to work to achieve the target leaving temperature at the required flow rate. We can see in Figure 3 that two of the three compressors of Chiller 1 are currently running, each consuming 44 kW for a total electrical power consumption of 88.3 kW for Chiller 1.

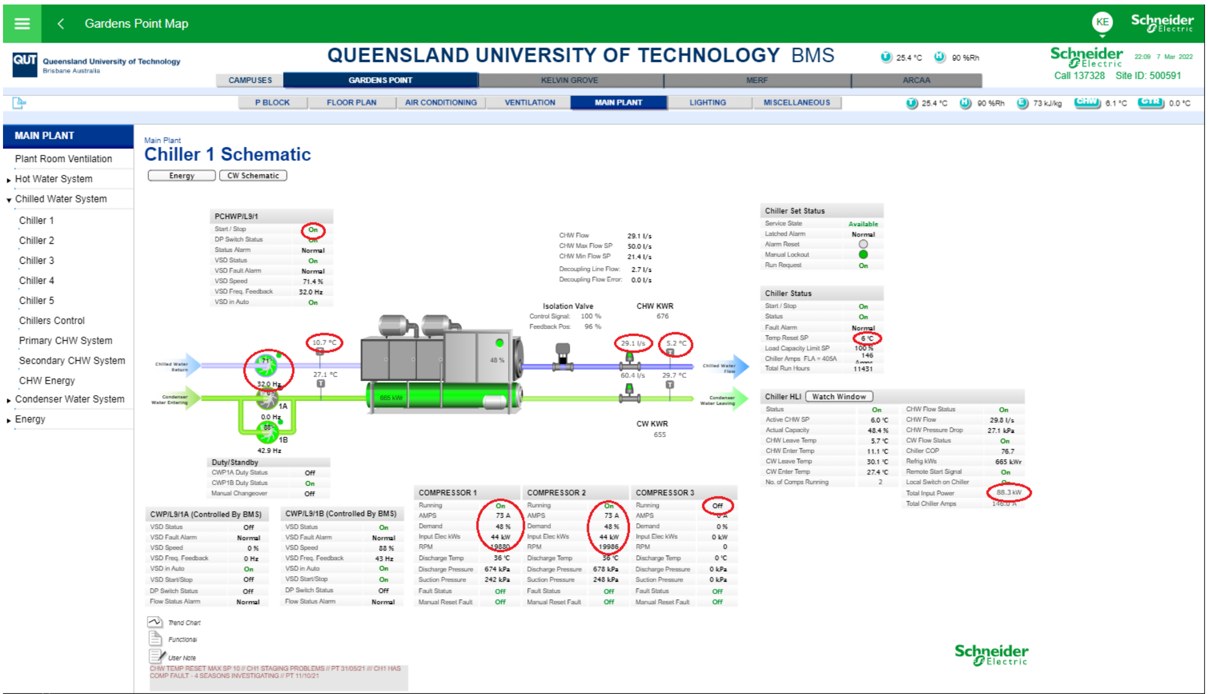


Figure 3: Detailed Schematic for Chiller 1

The data that we will study is time-series data that includes, only for Chiller 1, a historical record of the following data at approximately a 3-minute interval:

- The CHW Flow rate (litres per second)
- The CHW Entering temperature (°C)
- The CHW Leaving temperature (°C)
- The Electrical energy consumed by the chiller (Kilowatts), abbreviated kWE.

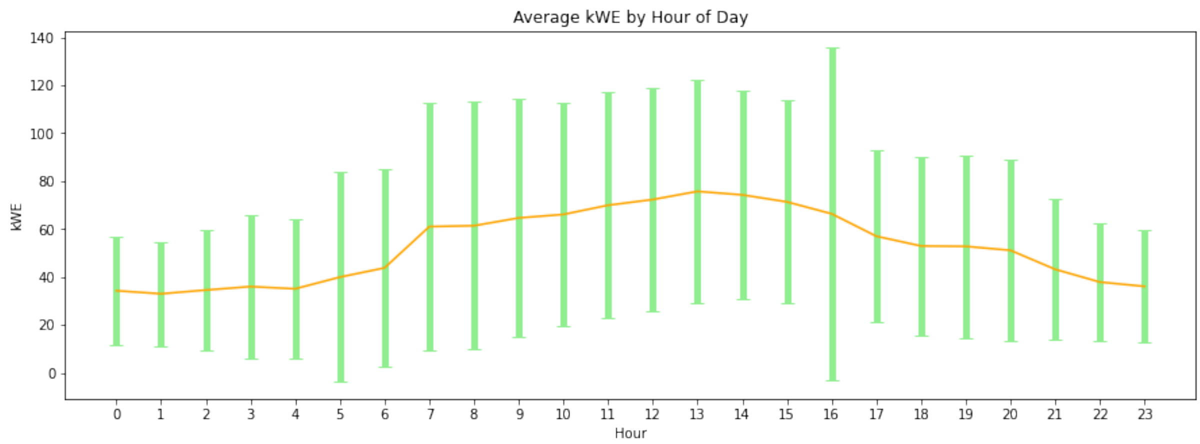
The amount of heat transferred from the chilled water system to the condenser water system can also be measured in Kilowatts, the so-called *Kilowatts of Refrigeration*, abbreviated kWR. Your first task is to conduct research to determine the formula for deriving the kWR. One of the most important metrics that HVAC (*Heating Ventilation and Air Conditioning*) Engineers monitor and try to optimize is the so-called *Coefficient of Performance* (COP) which measures how efficiently the heat energy is transferred.

Having calculated those quantities (kWE, kWR and COP), your job is to investigate how those quantities tend to vary over time. How do they tend to vary depending on the hour of the day, the day of the week or the month of the year? Note, we are only interested in looking at these quantities when the chiller is actually operating (an electrical consumption of 0 kW indicates that the chiller is not currently operating). We are also interested in how often chiller 1 is actually operating. How does the percentage of time that it is operating vary depending on the hour of the day, the day of the week etc? Remember: chiller 1 is only one of five chillers, and we don't have data for the other chillers, so we don't know when they might be operating along side of or instead of chiller 1.

We suspect that something *weird* may have happened on your birthday, so please do a detailed analysis showing how the system performed over the 24 hours of that particular day. Did you detect any atypical behaviour that day? Make sure you use your actual birthday – as documented in your QUT student record.

It is well known that the coefficient of performance tends to vary depending on the cooling load (kWR). 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. Use appropriate plots to visualize all of the above relationships. The following provides an example of what one of these plots might look like:



Add markdown to summarize your observations/conclusions for each plot.

Note that the raw CSV data may require "*cleaning*" before it can be processed.

Everything should be included in a single Jupyter notebook. The Python code included should follow best practices as outline in the lectures, including using well chosen identifier names, writing clear simple code, and 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.

All of these features are covered in the lectures and/or practical exercises, so there should be no need to use Python or Pandas features outside of what has been covered in class. Note, ***you will get Zero marks*** if you use some other programming language, library or system such as R, MATLAB or Excel.