****

2020 Recruitment Challenge

Electricity Demand Challenge



**Overview**

This challenge is part of a larger project to assess the impact of Covid-19 on Queensland electricity demand. The purpose of this challenge is to assess: i) knowledge & skills relating to data analysis and ii) ability to interact effectively with colleagues and mentors.

The challenge involves:

* creating a baseline of seasonally adjusted 2020 Queensland electricity by averaging the actual electricity demand over the past 5 years.
* downloading electricity demand for past 4 weeks from AEMO
* using visualisation to compare recent actual demand with the same time interval of the baseline demand.

To complete this task, please fork this GitHub repository https://github.com/Synergetic-Engineering/recruitment-challenge-2020 and share your forked repo with us when you’re done.

**Goals**

The overall goal is to provide a visualisation of how electricity demand in 2020 compares to the historic average. At a minimum, provide this visualisation for a single day of your choosing.

For bonus points:

* Identify patterns and/or trends in the data
* Allow user to select which day to visualise
* Provide visualisation for an entire week
* Create your own functions
* Write associated test(s)

**Any Questions?**

* If you run into any issues when completing this task, please don’t hesitate to contact steve.buckman@synengco.com or steve.buckman@synengco.com for help.

**Data Source**

1. **Yearly Demand**

The dataset includes 5 CSV files of actual Queensland electricity demand for the financial years 2015 to 2019. These files were downloaded from the AEMO website. Each file contains a row for each day in the year and 48 columns for each half-hourly time interval from midnight.

1. **Daily Demand**

Recent daily actuals may be downloaded from the AEMO Data Portal: <http://nemweb.com.au/Reports/CURRENT/Operational_Demand/ACTUAL_DAILY/>

Notes on these files:

* Filter on REGIONID column to obtain Queensland actuals
* Demand is recorded over 48 half-hourly time intervals, with the ‘day’ starting and ending at 4:00am
* The structure of these Daily Demand files is different to the Yearly Demand Files (e.g. single column of demand measurements vs multiple columns)

**Tools**

You may complete this challenge using your preferred development environment (ideally R or Python).

**Deliverables**

The expected deliverables for this challenge are:

* CSV file of predicted half-hourly Queensland electricity demand for 2020 (based on historic average)
* Daily electricity demand for past 4 weeks
* Visualisation of recent electricity demand compared to historic average over same time period
* Your accompanying notes on this challenge
* GitHub repository of your code

**Visualisation**

How you visualise the data is entirely up to you! It needs to show how recent actual demand compares to the average historic demand over the same period of time.

**Accompanying Notes**

Provide a short set of notes (1-3 pages) on this challenge. Topics to consider:

* Your thoughts on how to tackle this challenge
* Your rationale as to why you did things the way you did
* Observations about the data
* Your questions (with or without answers)
* What you would do differently if you had more time
* Suggestions for progressing this work further