

Accompanying notes for the Synengco Recruitment challenge

Saeed Azad

I decided to use Python data science libraries to tackle this challenge, including Pandas to wrangle and analyse data and Matplotlib for visualization.

By exploring the dataset for financial years 2015 to 2019, I noticed we have data from 2015-07-01 to 2020-06-30. Thus, I first decided to exclude 2020 from the analysis, so that baseline model would reflect pre-covid patterns.

As the goal is to make a predicted model based on average historic data, I grouped the data by month and day to get the average for every 48 half-hourly time intervals. Since the 2015 data start in July, averages for January-June were calculated from four values while averages for July-Dec were from five values.

To visualize actual daily electricity demand for 2020, I removed the first interval for 2020/05/01 as there is only one value for midnight, so I could easily visualize the whole day.

By plotting actual daily electricity demand for 2020 April, we can see there is a lot of variation across April, fluctuating on a weekly timescale. Aside from the weekly fluctuations, there is no clear trend to increase or decrease across the month. Total daily electricity demand for April 2020 ranges from 260186 to 303291.

We can see substantial increase between recent April 2020 actual data and the historic average. By analysing the percentage increase for each day in April, we can see the covid-related increase in energy demand was between 1% (on 20th April) and 16% (on 4th April).

To give better understanding of these changes, I also visualised an entire week to observe trends in electricity demand between business days and weekends. I also wrote a function that allows the user to specify a given date and plot the energy demand profile for that day.

If I had more time, I would build a time series model using the scikit learn python library to more accurately forecast the covid-related changes in energy demand, based on the available historic data. I was also fascinated by the energy profiles across different days of the week and in particular how they changed with covid. I would like to explore these patterns quantitatively, for example by analysing whether there were significant changes in energy use behaviour before and after covid.