COVID-19 Electricity Demand Analysis

Peter Dowell

p.g.dowell@googlemail.com

Introduction

In the search for developing "fast indicators" to help assess the impact of coronavirus in the UK and the global economy, electricity demand data has been identified as a candidate due to its correlation with GDP. GDP is published quarterly so an earlier approximation is of great benefit.

The ENTSO-E Transparency Platform hosts a public API that gives access to datasets that record the demand for electricity in the UK and other European countries. The data from the API provided is refreshed continually, in close to real time.

This report investigates the pandemic's effect on electricity demand, and considers how the API could be integrated into an analytical pipeline for monitoring the data daily.

Method

The following steps were performed to import prepare the data for analysis:

- Imported data into a tabular data structure;
- Converted time field into single date-time value and used it as the primary index,
- Extracted year, month, quarter & week as separate features for filtering,
- Removed data pre-2015 (3 samples from 2014).

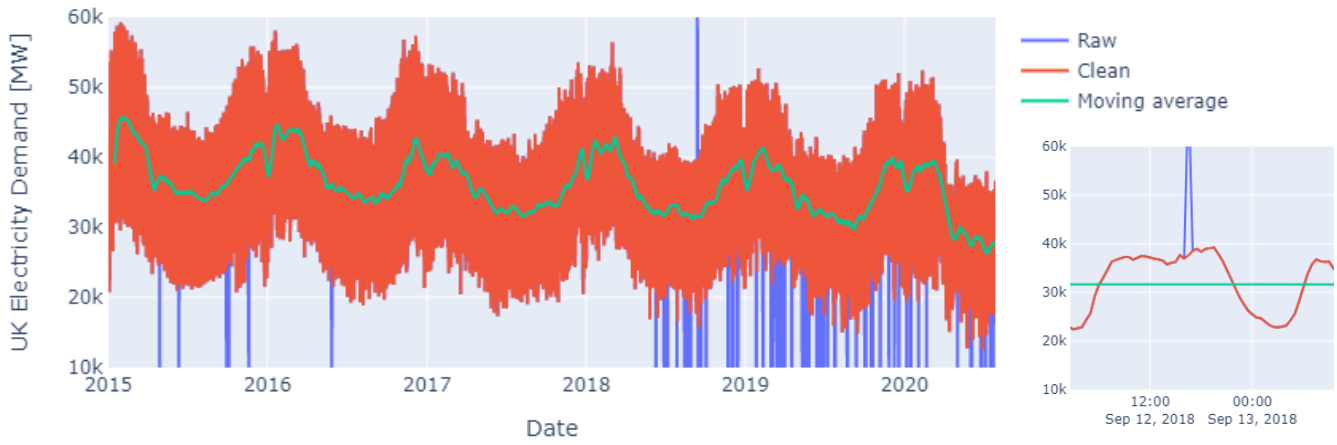


Figure 1 - UK Electricity demand 2015-2020

Figure 1 shows the raw data from the UK vs. that which has been cleaned. The raw data contained many samples with unrealistically low, or high demand rates considering the context of the surrounding data. These outliers were found using an algorithm that identified samples that were outside of the lower, or, upper fence for a given country and year. The missing values were then interpolated using a PCHIP[1] spline

[1] PCHIP Piecewise Cubic Hermite Interpolating Polynomial

to prevent overshoots around turning points in the data. The inset image shows how data is successfully identified as erroneous and filled with plausible replacements, preserving the shape of the data.

Analysis

To investigate the effect of COVID-19 on electricity demand, the median was calculated for each year.

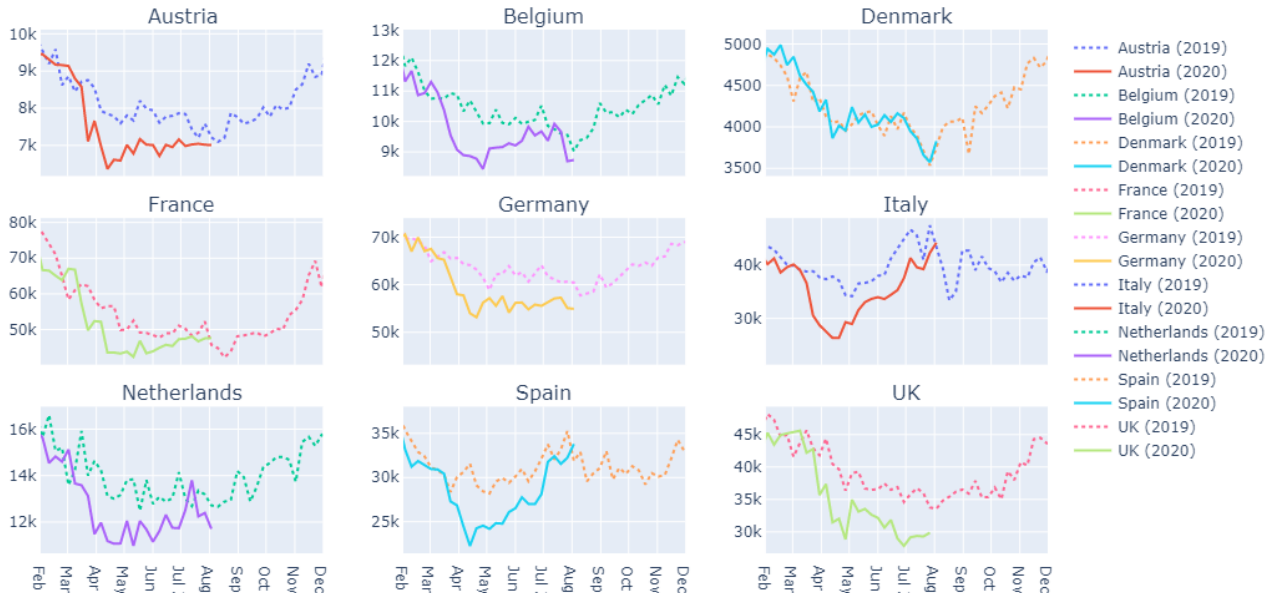
Since energy demand varies according to seasonal effects (see Figure 1), only Q1 and Q2 data was considered for comparison against 2020 data, as only data for the first two quarters is available for 2020.

To cope with large variations of demand between countries, the results for each country are expressed as a fraction of 2019 data as shown in Figure 2. This indicates that:

- Electricity demand hasn't varied more than 3% in the last 3 years for all the countries analysed;
- With the exception of Denmark, in 2020 there was a 6-16% drop in demand compared to 2019.

	Austria	Belgium	Denmark	France	Germany	Italy	Netherlands	Spain	UK
2020	0.92	0.91	1.02	0.9	0.94	0.84	0.92	0.89	0.88
2019	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2018	1.01	1.03	1.03	1.01	1.01	0.99	1.01	1.02	1.02
2017	1.01	1.03	0.97	0.99	1.0	0.97	1.01	1.01	1.01
2016	0.98	1.02	0.99	1.02	0.97	0.94	1.0	1.0	1.11
2015	0.93	1.05	1.0	1.01	0.96	0.96	0.86	1.0	1.08

Figure 2 - Median Q1 & Q2 electricity demand normalised against 2019 Q1 & Q2 data



Discussion

- Commentary on any limitations of the dataset and/or your approach, and the steps that could be taken to address these in future

Further Work

To produce a daily process to monitor this data, an analytical pipeline could be delivered to:

- Query the ENTSO-E API to acquire the raw XML provided by the web service;
- Use an XML parser to convert the data to a tabular data structure containing the time series data;
- Clean up any outliers using the method developed;
- Compare the actual data to a forecast of the demand and its confidence intervals;
- Use an anomaly detection algorithm to detect when data falls outside the forecast confidence intervals and raise a notification to the business.

This pipeline will need to be hosted on a server that is supported by IT including regular back-ups. The integrity of the code will be maintained via source control, automated unit testing and deployment scripts, on the assumption that regular changes may be required as the pandemic unfolds.

I would go back to the business and ask the following questions:

- Do you have a standard internal definition of when the pandemic started, e.g. when the WHO declared it a pandemic or when each country implemented certain policy measures? If not, would this be useful to explore with further analysis e.g. which types of policy measure have affected energy usage?
- Would you like to include in this analysis the assumption that energy demand will continue to predict GDP in these unprecedented times or will future analysis be required to confirm this still holds true given the current unprecedented circumstances?
- What kind of monitoring do you require - e.g. will you be manually checking Tableau reports or would you like an automated process that sends email alerts if a certain threshold is reached?
- Who will need access to these reports and what are the access security requirements, if any?
- Do we have a model that predicts GDP based on electricity demand and other factors? If so, should this daily process feed that existing model and what kind of data format does it require?
- What format would the electricity data have to be in to feed into that system?
- The National Grid produces their own projections of electricity demand, would it be useful to see if we can access their predictions? As the subject matter experts, their prediction of future energy demand is likely to be the gold standard; this could be valuable if it would be of use to predict GDP.