electricityMap

electricityMap data analyst challenge

First off, thanks for taking the time and putting effort into completing this challenge, we really appreciate it! Remember, there's never only a single solution to a problem, so don't stress over striving for perfection. Lastly, have fun and express yourself, feel free to do stuff that is outside of the box.

Overview

One of our goals is to become the single source of truth for electricity data. To this end, we need to be able to clearly and succinctly explain what our data is, how reliable it is, and give an overview of its quality to interested customers. Sometimes, clients have further questions on the technical details of the data - schemas, distributions and/or energy modelling specific - and often ask for guidance on how to best utilise it.   Therefore, it is important for us to give clients confidence in our data quality, answer any questions they have, and give them useful insights around their use-cases.

Case background information

A Copenhagen-based Danish EV charging app is interested in using our data to help their users understand the carbon footprint of their EV usage, and reduce their emissions by charging when electricity is the greenest. They don't really understand our data, but they know that it can be used for the "smart-charging" use-case. Here are some of the questions they have:

- Why are there two electricity zones in Denmark (East Denmark and West Denmark)?
- How much and why does the carbon footprint of electricity in Denmark change throughout the day?
- How can we use your data to help users charge at the greenest times?

The last question is quite open-ended, the idea is to help the client feel confident that our data can be used to help their users reduce their charging emissions. Some ideas could include: send them example visuals that they could use in their app to engage their users, or demonstrate how much smart charging could reduce their users' emissions. Note that we do not ask you to create a complex optimisation algorithm from scratch but you should rather focus on convincing and explaining to the client what are the potential benefits of optimisation for reducing their user's emissions.

In the end, we want to convince them to use our data to have an impact!

Task

Your task is to conduct an analysis on data for Eastern Denmark from Jan 1st 2020 to December 31st 2020 to answer the client's questions. There is not a given data set to work within this task, but you can find hourly data for generation by production type for Eastern Denmark here (ENTSO-e) . For the carbon intensity of each production type, you can use the ones listed on our public Github repository. Don't focus too much on creating sophisticated models, the purpose of this case is to convince the client of our data quality and of the usefulness of the data given their use-case. You can assume that the client is savvy but has little technical knowledge.

Once you have extracted the raw data from ENTSO-e in spreadsheet format, we want you to create a simple sqlite database in which to store your data before analysing it. You can read a straightforward runthrough of how to set this up here. Once the data is in your sqlite database, you can query this raw data using SQL just as you might when working internally with electricityMap data.

Note: Typically, we deal with consumption data which is electricity generated within a zone, plus/minus imports/exports to/from that zone. Don't worry about imports and exports to/from Eastern Denmark for this assignment—just focus on the electricity generation by production type.

Your work should include:

1. The code you used to create and query the sqlite database with your data
    a. We want to understand how well you are able to navigate basic Python and SQL, and setting up a small database. We want to see how you would create a simple database schema that you can use for analysis.
2. The code you used to process/clean the data, along with clear explanations documenting your process
    a. We want to understand your process so feel free to be technical in your explanations. This would be for the benefit of the engineering team of electricityMap.
3. Materials presenting your analysis to the client (electricityMap team will play this role)
    a. You are free to decide the exact format of these materials and what to include; for example, they could be a slide deck, short 2-page report, visuals, mock dashboard (if you think this is necessary). Assume that you won't be able to walk the client through the materials, so make them as self-explanatory and clear as possible.

Logistics

- We know this is a fairly large task, but we expect you to spend max one day on this assignment, so scope it accordingly - we know that this is a task that one could spend a lot of time on, so we're not looking for something perfect!
- use the tools and packages of your choice

Deliverables

When you are done, please send us the following deliverable:

- A git repo (link or zip) containing your code, analysis and explanations (ideally git but a file/other document is fine)
- Your materials in a zipped folder (and a document explaining which materials you included and any relevant info needed).

Don't hesitate to ask us questions if anything is unclear.

Looking forward to your creation!

the electricityMap team