**Introduction**

This document includes the test we share with candidates as part of the interview process for the **Analyst** position at TerraVerde. We would like to receive your complete response **within 4 days** unless additional time is requested by you. Feel free to contact me for any clarifications.

Part of your role as an Analyst involves collecting, cleaning, and curating data. Ideally, data is sourced in a standard form but that is not always the case. Also, you need to make sure the data is cleaned and organized (i.e. formatted for usability). Finally, you need to make sure the data is complete so at times you will have to curate data to fill in missing values. We are continuously working to automate our data collecting, cleaning, and curating process by building our software tools to make work more enjoyable for our Analysts ([see article](https://www.dezyre.com/article/why-data-preparation-is-an-important-part-of-data-science/242)). However, a competent Analysts must be ready to tackle raw data and turn it into usable format on the spot.

Background – We recently received [this raw electricity data set](https://tvrp.box.com/s/tlzu629qdnm9otpqggnlc1mk20nsdybw) and the customer needs us to calculate their usage under a specific electricity tariff. The usage data is in kWh and is in a 365 x 96 format (as in approximately 365 rows of days and 96 columns of 15-minute intervals). There is a summary sheet which lists all of our customer’s electricity accounts. We are only interested in seven of the accounts which are each included in a separate sheet (ex. sheet ‘4’). In each sheet, you will notice that there are usage data (in kWh) for two channels under the CHNL\_ID column. The channels are 101 and 102 which represent electricity import for the grid and export to the grid, respectively. The reason there is export electricity from these electricity accounts is that there are behind the meter solar facilities installed at each of the sites.

Part I – We would like to convert the raw electricity data into a 2 x 35040 format as shown in [this output file](https://tvrp.box.com/s/yv3uz4x6e30jilttvddxrxvetg85ysl1). We want the data to be in **kW** (not kWh) for each 15-minute interval starting from January 1, 2017 per channel. Finally, we want the net usage in **kWh** for each electricity account. You need to find out how to convert between kWh and kW as well as calculating net electricity usage for each electricity account. You may use MS Excel or Python to perform the work. We strongly advise on performing quality check on our work at TerraVerde. We rather delay sending an analysis to a customer than sending an incorrect analysis on time. So, make sure to review your work thoroughly and check for missing or duplicate data. Good luck!

Part II – Now that the data is ready, we would like to calculate the electricity costs for each of the accounts for one year. For this analysis, we will assume the accounts are located in San Diego Gas & Electric territory and are subject to the [Schedule EECC-CPP-D](https://tvrp.box.com/s/i6xe7ri9uoumcve37kukcyzolashq02q) commodity tariff. You need to study the commodity tariff, extract relevant information, and set up a model to calculate the electricity costs for each account. Make sure you work is clear and auditable (i.e. a reviewer can see the steps you took to get to the answer as well as your assumptions). Reading the commodity tariff may be confusing which is why an average customer does not know how they are billed by their electricity provider. That is why customers come to us to help them understand if they are billed correctly and how they can save on their electricity usage and costs. The ability to understand tariffs and translate them into mathematical formulas is one of the main skillsets of an Analyst. You can do it!

Part III – As an Analyst you will often need to prepare and present a report of your findings to customers, regulators (ex. CPUC), industry groups (ex. CALSSA), etc. To show off your writing and presentation skills, please provide the following:

1. An Output.xlsx file with one sheet for each electricity account for part I
2. A document of your choosing showing the bill calculation results for part II
3. Python scripts or Excel models (either is fine; work with what you are good at)
4. A narrative explaining the steps you took to perform the test and how long it took you to finish the work
5. A list of your quality check items for part I and II

Thank you for your time and effort in responding to this test!