**Explanation:**

Set up:

* This task is focused on feature engineering, Estelle has provided a CSV dataset for you to use as a base for this task.
* You should download the notebook and CSV and start by running the cells within the notebook.
* Estelle has also provided some insight into a feature that would be interesting to add to the data. By running the cells in the notebook, this will create the feature that she described for you and provide a foundation for you to begin your own feature engineering.

Here is some context around the additional features that have been engineered in the notebook, to help you in the future:

* Firstly we have the average price changes across periods. This is a measure of the average price change by company between peak, mid-peak and off peak periods.
* We then take this idea one step further by creating another similar feature but instead of looking at the average price difference, we look at the maximum price difference across periods and months. This gives another way to look at the price changes across months.
* The reason why these 2 features could be useful is because they are another way of representing the variance of prices throughout the year. Imagine, if your utilities bill massively increased over winter, as a consumer you’d be annoyed and want to find a better deal!
* After this we continue feature engineering with some more concepts, including transformation of columns.
* To make predictions with a statistical or machine learning algorithm, all of the data must be converted to numeric data types.
* Therefore, we convert date into months and remove the raw date column, as we cannot use it in its original form.
* We also convert boolean columns into binary values.
* And we convert categorical columns into dummy variables. A dummy variable is a binary flag that indicates when a row matches the value from the categorical column that it was created from.
* As we saw during exploratory data analysis, the distribution of some columns was skewed. This is important to identify because when modeling data for prediction, based on the technique or algorithm that we use, there are sometimes assumptions within the data that we should follow.
  + One common assumption is that the columns within the data are normally distributed. Hence, if we find that columns are not normally distributed, we should treat these columns to try and transform them into a distribution that is more normal.
* Therefore, the next thing we do is transform some columns to have a closer to normal distribution. We do this using the logarithm function. As you can see from the visualisations, the newly transformed columns are much closer to a normal distribution than what they were earlier.
* Finally, we plot correlations of all the columns to see if we can identify any columns to remove. Columns that have very high correlations indicate an area to look out for. In this case, you may want to remove one of the columns, since they are likely both holding very similar information.