

February 19, 2023.

Dear [Client point-of-contact],

The dataset was downloaded from Kaggle.com. The below table highlights the summary statistics from the downloaded datasets.

Table name	No. of records	Periods of data	Date Data Received
Brazil Energy Demand Hourly	201,312	2000 - 2023	Friday 17, 2023

Notable data quality issues that were encountered and the methods used to mitigate the identified data inconsistencies are as follows. Furthermore, recommendations have been provided to avoid the reoccurrence of data quality issues and improve the accuracy of the underlying data used to drive business decisions.

- ***Additional Index was found for January 2023.***

Mitigation: Please ensure that all data in the table are from the same period. This indicates that the data received may not be in sync with each other which may skew the analysis results if there are missing data records.

The analysis is done at the end of each year and 2023 has not come to an end.

- ***Wrong data type for Index and Hourly Demand***

Mitigation: Convert the index column to the appropriate data type and the hourly demand column to numeric.

Recommendation: Ensure the right data type for every field not to make it difficult to interpret result. And data transformations are made to ensure consistency in the data types for a given field

- ***Inconsistent values for the same attribute***

Mitigation: Please, ensure that all hourly demand field have a consistent decimal or are rounded up.

Recommendation: Ensure all field either have the same decimal place or are roundup for easy interpretation of data.

Moving forward, the I will continue with the data cleaning, standardisation and transformation process for the purpose of model analysis.

Kind regards,
[AO Sanusi]