*Data Maintenance*

This phase is implemented continuously during each iteration and then on a standard schedule of 3 months upon product deployment. However, it can also be implemented as needed. In maintaining the data, the responses to the review questions will be utilised to determine if the information for a specific charging station will be updated. Moreover, the Data Dictionary will be used for correct data definitions. The actual data updates will utilise the Content Management System or the phpMyAdmin.

It is assumed that the only the database administrator with knowledge on how to use phpMyAdmin, has access to the Content Management System, and with knowledge how to use .notebook can make changes to the data. The process of the Data Maintenance is dependent on the data, as listed below:

**For existing data**

1. Using phpMyAdmin, the User Review table will be downloaded as a csv file
   1. The file will then be loaded in the .ipnyb file for data analysis
      1. Code is available here:
   2. The review table for the charging stations will be visualised using the code, where in it will highlight charging stations with the lowest average rating and highest count of responses were charging station is not working
      1. The user should further investigate results of the visualisation and determine if these need to be updated
2. If the updates are less than 30, the Content Management System will be used to update the charging stations individually
3. If the updates are more than 30
   1. The database’s csv file will be downloaded from phpMyAdmin and updated using MS excel or python, depending on the homogeneity of the changes
      1. For location related data, code for implementing geocoding and reverse geocoding can be found here:
   2. The changes for the data will be applied and the updated file is saved in .csv format
   3. The user then can update the charging stations data, except for the charging station id
4. The updated version of the charging station database will be saved in the backup repositories

**For Victoria Libraries Data**

The Victoria Libraries Data represents around 50% of the charging stations in the current database, thus update for data from this data source will use the data wrangling code in python.

1. The libraries data (libraries.csv) is downloaded from open.gov.au
2. The latest Charging Station database form the data back-up repository
3. The .ipnyb file will be opened, and the file path of libraries.csv data will be added in the code
4. The code will be run, creating a new version of the charging station database with updated libraries data
5. The new version of the charging station database will be reuploaded in the website
6. The new version of charging station database will be saved in the backup repositories

Note: This feature was not included as a built in feature of the website. This is because the updates for this data is not as frequent due to the static nature of the data. The ROI for the work to implement for current iteration timeframe is low.

## Additional Data Notes:

* In managing the data, MS Excel as well as Python (via Google Collabs/Jupyter Notebook) were used given easily manageable size of the data (less than 300 charging stations).
* Once the charging station collected goes above 5,000 records, the system administration team should consider integrating data maintenance steps into the Content Management System