Backend Report:

Introduction

- UCCE's mission is to provide a way for the University of California to engage with California citizens to achieve innovation in a way that offers sustainability and science literacy

Discussion

- Purpose of the website:
    - Engage the community by allowing them to enter their composting data and receive immediate feedback on how their composting is contributing to the overall composting effort
    - Collect the information as community members enter it, reducing reliance on surverys sent out three months after the workshop
- Important aspects:
    - Data must be stored in a way that is easily understandable/accessible by those without a data science background.
    - Self-contained, low maintenance, and easy to understand if someone must work on it
- Critical customers:
    - UCCE CEP and individuals entering their data
- Goal:
    - Motivate the website users and workshop participants to keep composting
- Main Specifications of the project:
    - Store Data - System to store required data
    - Shall have an exposed API
    - API should be able to handle read and write requests to the database
    - API usage needs to be well documented
- Attempts:
    - AWS
        - Faults: fee for CPU and memory utilization, complex coding (lamba functions)
    - Airtable
        - Too much focus on UX then what they are trying to accomplish
    - Google Sheets API
        - Better for those who have less coding experience to navigate and manage in the future
- Arising issues:
    - Security: can this data be stored in a publicly accessible Google sheet or if there needs to be security measures to protect the data. Google forms also have security issues.
    - Consumer Privacy Act: handling personal information like geolocation
    - Data must be adequately protected.
    - Data privacy laws to ensure data is protected and the personal data of the survey is not compromised

Results and Analysis

- Main components:
    - Google sheet worksheet
        - Data storage that takes in raw data and calculated different statistics for the data
        -
    - Google Cloud Project (GCP) API configuration
        - Access the stored data from the google sheets
            - Using built in GCP APIs
        - Two service accounts, one with read only access and one with read/modify permissions
            - Read only provides a safeguard from people trying to manipulate data
            - Read/Modify is used in secure trusted networks
    - Using python and Google API client python package

Recommendations:

- Create a proxy API that would redirect the traffic of the website from the Google Sheets API to the proxy API
    - Proxy API: sits between a client and an API, providing an access point to the API with additional functionality such as security, caching, or rate limiting, without requiring changes to the API
    - This would allow the Google Sheets to become private again and allow for more security
    - Proxy API be hosted within the GCP
- Importing the Qualtrics data using Python scripting and the development PAI key to reduce the need to have someone manually import new data into the spreadsheet