Stage II: Project Proposal Pitch and Specifications (Group 4)

Group Members: Jason Kantner, Gordon Petry, Michael Williams, Casey Lishko, Yanaja Joyner, Len La Rocca

Link to GitHub Wiki: https://github.com/gpetry37/CivicStats

For this project, the main problem that we want to address and find a solution for is the lack of easily accessible/understandable sustainability data on the CivicStory website. Currently, the CivicStory website is a hub for relevant news articles relating to sustainability, however direct information on sustainability efforts based on geographic location is not accessible to the common user.

The objective of the module that we will be developing is to provide a tool for the CivicStory website where normal users will be able to access detailed data on sustainability in their area. Another objective with this is to ensure that this data is tailored to what the user desires to know and easily understandable. This will be done through the use of data visualization so the user can physically see and understand the data they want to be informed of.

In general, the product we plan to design can be used by many different websites and is not only limited to CivicStory. The module that will be designed will be a page where the user can choose from many different variables relating to data in a connected database. Once the user specifies the data they want to see, they then get to choose a graph layout for the data the be visualized as (Ex. pie chart, bar graph, line graph, etc). After all of the selections are complete, queries will pull data from the connecting database in PostgreSQL and the desired data will be visualized onto the screen of the user. For this class, the data in the database will be relating to sustainability in the area. This makes the planned module a great tool to be utilized on the CivicStory website because it strongly relates to the purpose of the website and the goals it wants to accomplish.

Right now CivicStory is mostly a stream of news articles relating to sustainability efforts. Although informative, these do not offer the user a look at raw data that relates to the issue. Allowing the user to search for specific data gives them a true understanding of the issues revolving around sustainability and what can be done to improve it in the future. Also, they can compare and contrast data from different times or places to get the “whole picture” of where sustainability is, where it comes from, and where it needs to go.

To research the problem domain, we will look over the CivicStory website to see what type of data would be relevant to put in the database connected to the website. Not every piece of sustainability data we find should be put in the database. For example, we need to have a limited geographic location so it remains relevant to those that will access the CivicStory website. Once we research the problem domain in more depth, it will then be much easier to figure out what kind of data we will implement into our database. To obtain this data, we will collect it from a

local data collection hub. Kevin from CivicStory talked to the class about what he wanted to be added to CivicStory and mentioned that there is a hub near TCNJ that collects relevant sustainability data. We have reached out to Kevin to find out where this hub is and how to obtain this data so we may utilize it in this module.

Search bars and user interfaces that allow for specific search terms do exist in many forms (Ex. PAWS class search tool). However, the addition that we plan to make ours stand out from others in that our search tool creates a data visualization based on the information provided by the user on what data they want to view.

The search system has wide applications on many databases. It is a robust search system that visualizes the data, which is useful in any database. Possibly, instead of typing in the search criteria, the user can point out on a map where they want to search, or use a slider to determine what time period they want to look at. However, the data visualization tool will always be used after the search.

The system we plan to design will ensure efficiency between the many programs that are being utilized. Our goal is to reduce lag time as much as possible between the Database Management System and the web interface. This way, when the user prompts to create a data visualization, the task will be completed in a reasonable amount of time. Performance and efficiency should always be a top priority when creating a new web tool to use. This will be accomplished by taking time and planning out our code before going into the programs right away. We will also continue to test our system throughout the development process to see what areas we need to improve.

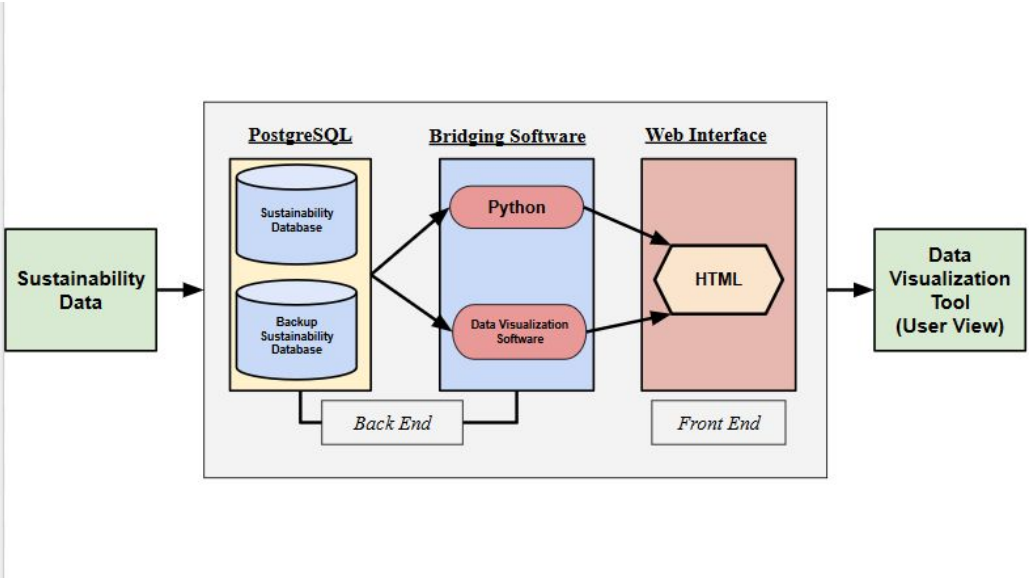
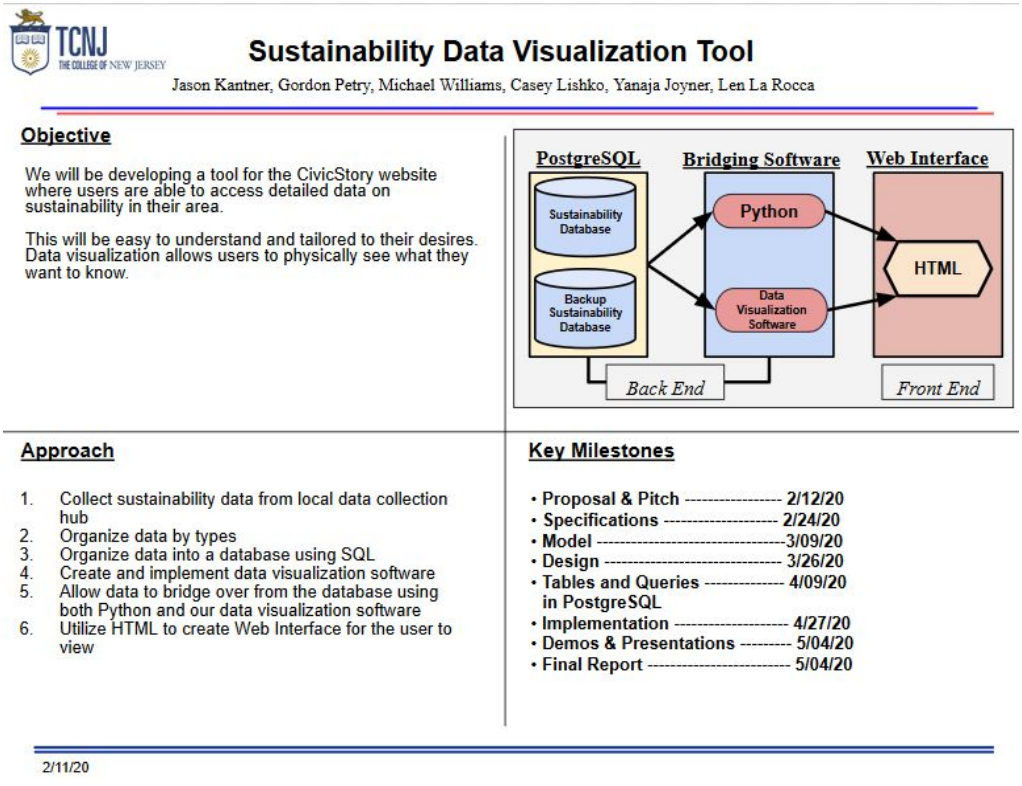
Regarding the security of our system, we don't plan to allow users to directly alter the data within the PostgreSQL database. This access will be limited to only administrators, who are allowed to do so. For the users, they will only be able to select the parameters in which they want to query the data and view the visualization that it provides. So, the users are only limited to viewing the web interface side of this module.

Backup and recovery are an important part when implementing a tool that utilizes data. For our project, we plan to have a copy of the database that will be connected to the web interface. This means that if anything happens to the database connected to the website, we will have a copy of the data just in case it needs to be utilized. This copy database will also be within PostgreSQL which makes it readily available to replace the original database if need be.

For this project, there are a few database concepts and technologies that our group will need to learn to be able to accomplish our objective. First, our team will need to learn how to utilize the Database Management System PostgreSQL. It is within PostgreSQL where our database containing the sustainability data will be located. We will also utilize HTML to create the web interface side of this project. This will create the frontend of the project where the user will view the data visualization tool. To implement this data visualization tool, we will also learn how to use data visualization software. Data visualization software is abundant and is also vital to properly creating our module. Finally, Python will be learned and utilized to bridge the gap

between the web interface and the Database Management System. Python is very important to learn because without it, we would not be able to make the connection between the front and back end portions of this module.

Diagrammatic Representation



Quad Chart