

Customizable Analysis and Visualization Tool for COVID Cases

Milestone 5

Team Members

- ▶ Calvin Burns, cburns2017@my.fit.edu (Team Lead)
- ▶ Sam Hartle, shartle2017@my.fit.edu
- ▶ Nicole Wright, nwright2017@my.fit.edu
- ▶ Stian Olsen, shagboeolsen2017@my.fit.edu

Faculty Advisor/Client

- ▶ Dr. Philip Chan, pkc@cs.fit.edu

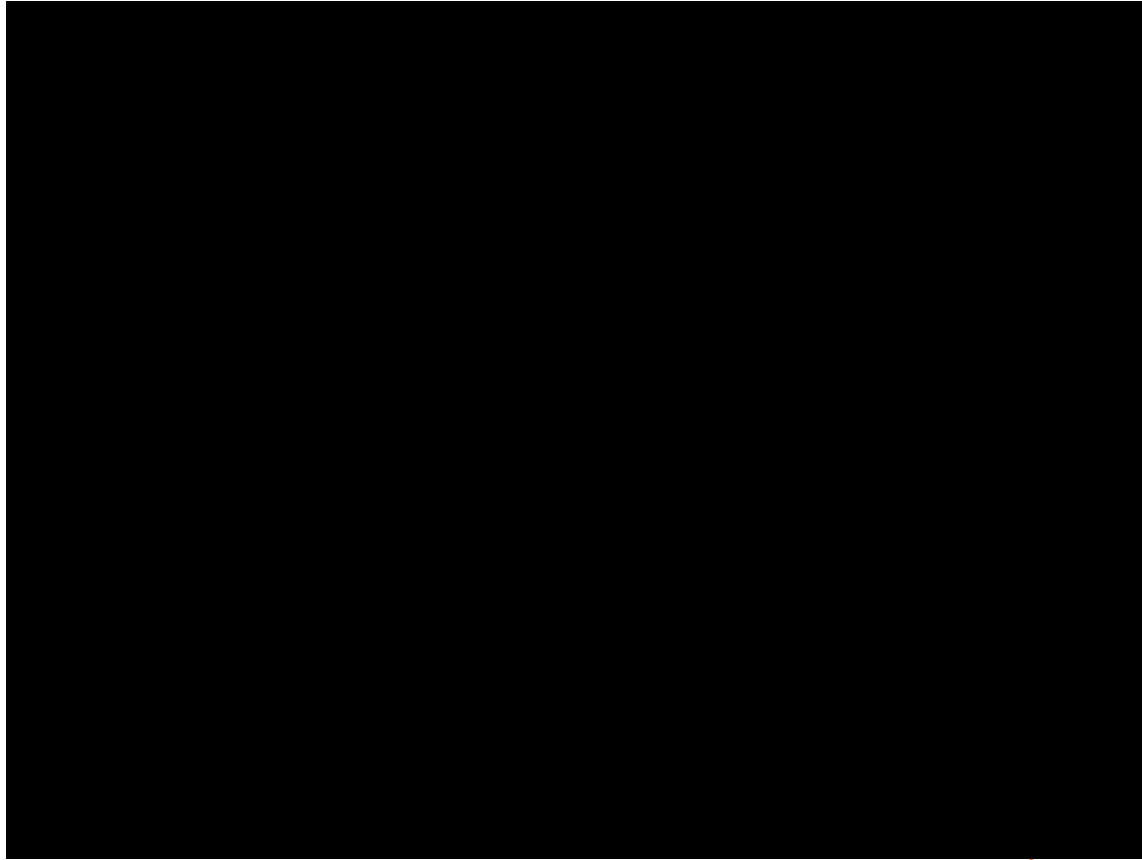
Progress Matrix

Task	Completion %	Stian	Sam	Nicol e	CJ	To do
1) Continue work on scatter plot	100%	100%	-	-	-	None
2) Operations card update to be more intuitive	80%	-	-	-	80%	
3) Save unique workspaces	0%	-	-	-	-	All
4) Layering plots	80%	80%	-	-	-	Change UI so user can create multiple plots
5) API auto updates for datasets	50%	-	50%	-	-	Convert/append JSON object received from API endpoint to a current CSV
6) Finish Application Feature	100%	-	-	80%	20%	None

Task 1 - Continue work on scatter plot

- Finding a good use for the scatter plot and finishing up the plot utility.
- Scatter plot to see if there is a correlation between positivity rate and mobility during the pandemic in Florida and in the USA.
- We also encountered a problem with partial weeks when we were resampling the data.

Demo Task 1



Task 2 - Update Operations Card

- Suggestion by Dr. Chan to use a similar format to SQL for querying and manipulating a dataset
- To move closer to the SQL method, we removed our filters and operations card and replaced them with a single “SQL Query” card
 - Has fields for SELECT, FROM, WHERE, GROUPBY, and ORDERBY
- We found a pandas plugin called “dataframe_sql”
 - Allows us to pass a SQL query statement and it will return a filtered and manipulated dataframe

Visual for Task 2

SQL Statment

Select

From

Where

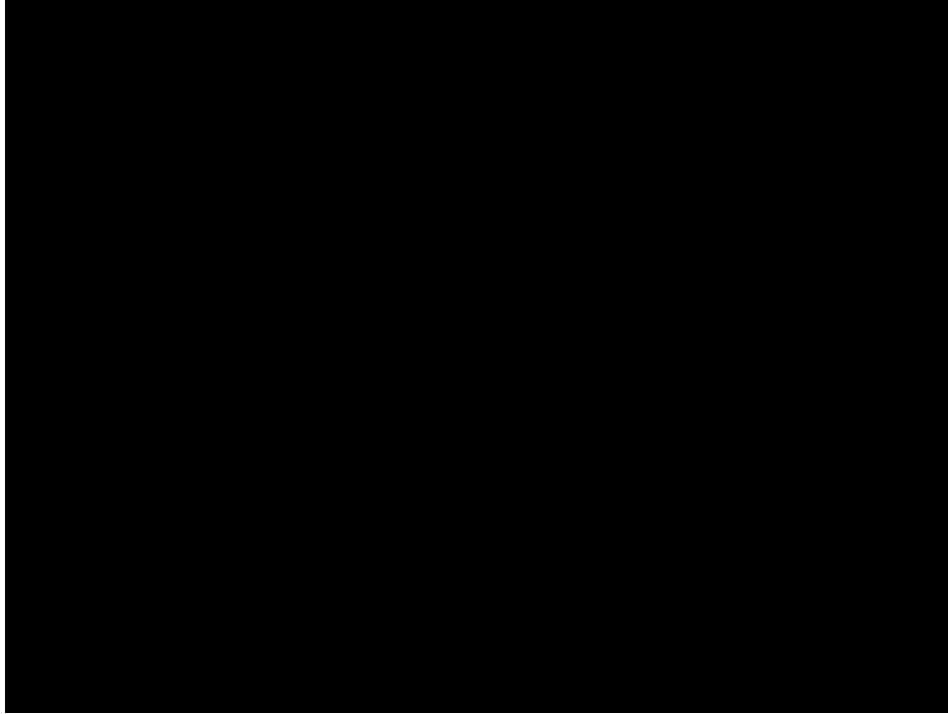
Groupby

Orderby

Task 4 - Layer plots

- Figure out which plots can be layered.
- We can layer pie charts, bar charts, line charts, timelines. We still need to figure out if we can layer scatter plots.

Demo Task 4



Task 5 - Auto Update Datasets

- User uploads a dataset and sets a URL field that points to an API endpoint
- Then a file is fetched from that endpoint and processed to update the dataset
- The research phase of this task is nearly complete
- Backend implementation will include querying the API endpoint given by the user and converting the returned JSON object to a CSV object
 - Append/Reprocess the dataset
- Frontend changes include changing the “Upload Dataset” page to add a text field for setting the URL that points to an API endpoint

Task 6 - Dataset Application

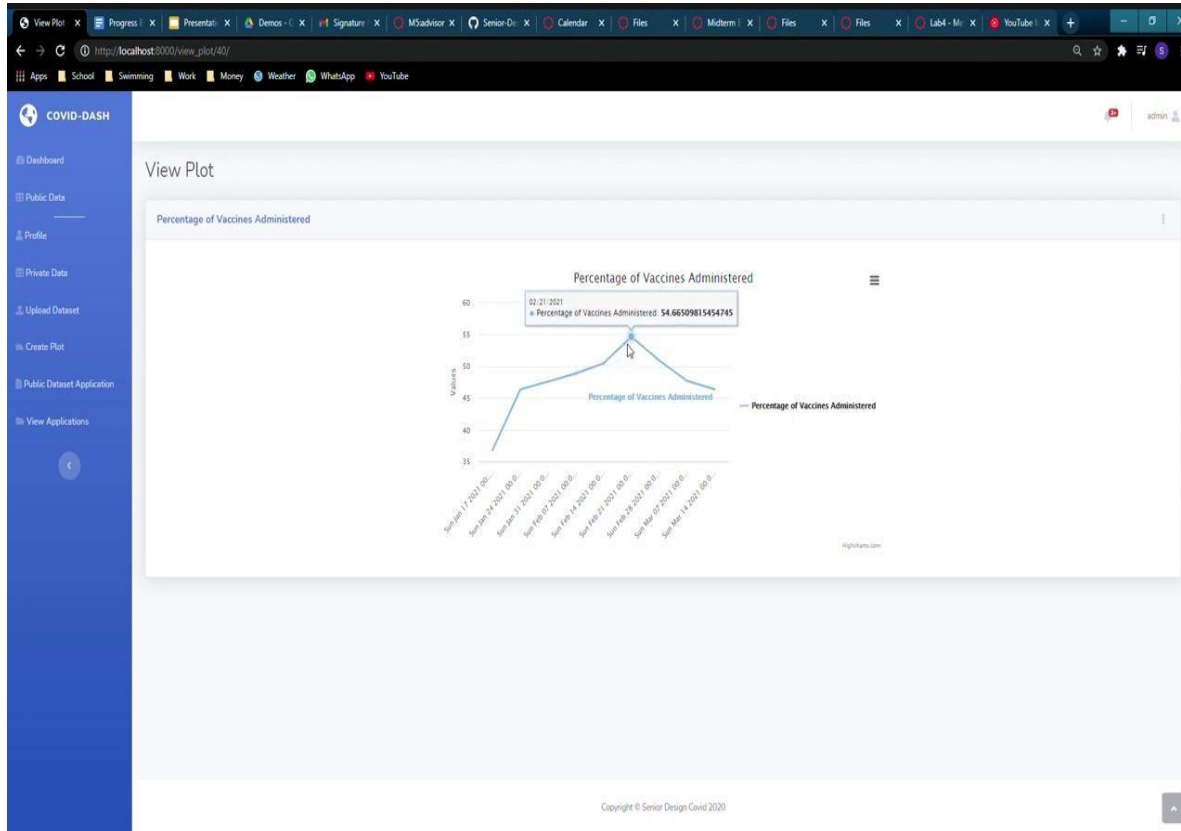
- Updated public data to include both shared and curated data
 - Shared: Data made public by a user but sent through approval process. *Use at your own risk.*
 - Curated: Data that has been submitted via an application and approved by the admin. *Verified, Safe*

Demo Task 6

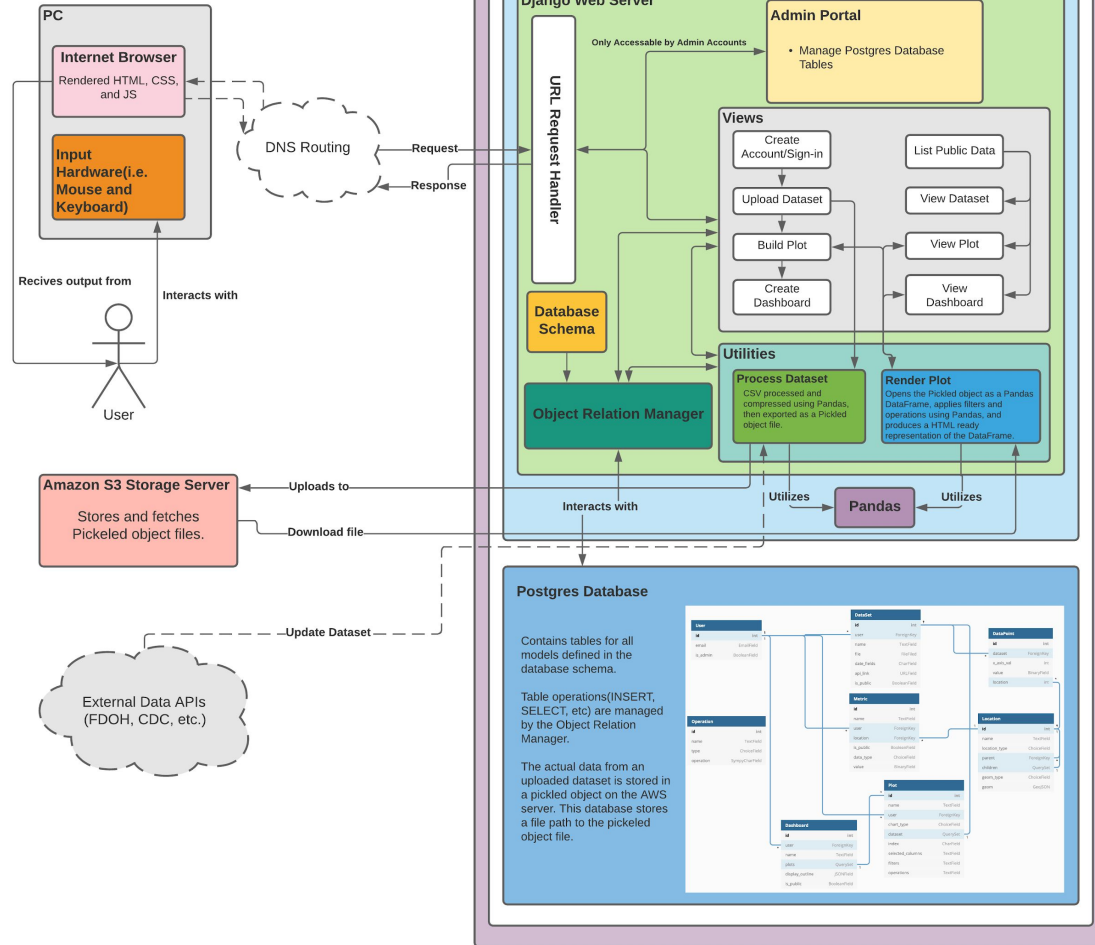
The screenshot shows a web browser window with the following elements:

- Browser Tabs:** Senior-De, Update a, Public De, Select ap, Senior De, Launch M, Milestone, Presentat, Progress.
- Address Bar:** 127.0.0.1:8000/public_data/
- Page Title:** COVID-DASH
- Sidebar Menu:**
 - Dashboard
 - Public Data** (selected)
 - Profile
 - Private Data
 - Upload Dataset
 - Create Plot
 - Public Dataset Application
 - View Applications
- Main Content Area:**
 - Curated Data:** This data has been reviewed and approved by the site admin.
 - Curated Dashboards:**
 - Show: 10
 - Search: [input field]
 - Table with columns: Name, User, Date Created, Last Modified
 - Showing 1 to of
 - Page navigation: « 1 2 3 »
 - Curated Datasets:** [Up arrow button]

Additional Demo for Vaccines



System Architecture Diagram - Current Progress



Task Matrix for Milestone 6

Task	Stian	Sam	Nicole	CJ
1) Selecting a variable (column in a dataset) used in the different plots below	User can select a dataset, then a variable (mostly implemented)	-	User chooses which of the 3 types of datasets to include (e.g. 3 check boxes, with private only checked as default)	User can search for a variable in the datasets (Possible use of the dropdown feature)
2) Plot template for trends over time (line)	User-selected trend variable from task 1 will be plotted on y-axis	User can select time range (default is max range) and frequency (default is daily)	-	-
3) Plot template for proportion among categories (pie)	-	-	User can select variable for the proportions among categories	Time range (default is max) and category selection
4) Plot template for relationship between possible factors and situations (scatter)	Time range (default is max)	User can see relationship between some x factor variable and a situation y variable	-	-
5) Plot template for distribution over FL counties (map)	Florida map with counties in heat map (implemented)	Allow for selection of time range (default is max of dataset)	-	User can see some variable when hovering over the map (cases, tests)

Questions?