

# Customizable Analysis and Visualization Tool for COVID Cases

## Milestone 6

# Team Members

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

# Faculty Advisor/Client

- ▶ Dr. Philip Chan, [pkc@cs.fit.edu](mailto:pkc@cs.fit.edu)

# Progress Matrix

Task	Stian	Sam	Nicole	CJ
1) Selecting a variable (column in a dataset) used in the different plots below	-	-	-	100%
2) Plot template for trends over time (line)	50%	-	-	50%
3) Plot template for proportion among categories (pie)	-	-	-	100%
4) Plot template for relationship between possible factors and situations (scatter)	45%	10%	-	45%
5) Plot template for distribution over FL counties (map)	-	-	-	-
Showcase Deliverables and User Manual	10%	30%	30%	30%

# Showcase Poster



## Custom COVID-19 Dashboard

Calvin Burns, Nicole Wright, Sam Hartle, Stian Olsen

Faculty Advisor: Dr. Philip K. Chan, Dept. of Computer Science, Florida Institute of Technology



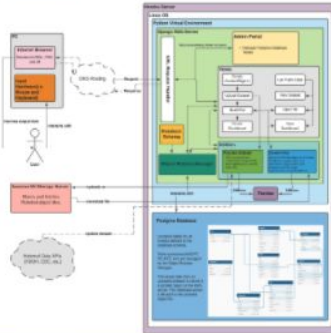
### MOTIVATION

- During the COVID-19 pandemic, our advisor noticed the lack of interactive and customizable COVID-19 statistics websites/dashboards
- The majority of dashboards have predefined datasets, graphs, and maps
- No dashboards allowed the user to perform their own statistical analysis or create their own plots
- No dashboards allowed users to upload additional data to use in conjunction with the sites existing data
- Few dashboards have layered plots, none allow the user to create their own layered plots

### GOAL

- Develop an interactive and customizable COVID-19 statistics dashboard that allows users to perform custom analysis, create custom plots, upload their own datasets/share them, layer multiple plots, and save a configuration of plots to a dashboard

### DESIGN



### FEATURES

- Upload/View Dataset
  - Users can upload a CSV dataset of COVID-related data
  - Dataset is compressed into a pickle object and can be viewed as a table on a separate page
- Dataset/Variable Selection
  - Users can filter their dataset selection to include the 3 types of datasets
    - Private
    - Curated (Reviewed by admin)
    - Shared (Not reviewed by admin)
  - Users can search for a variable in their selected dataset(s)
- Build Plot
  - Users can choose a visualization method and build different types of visualizations from variables in their chosen dataset
    - Trends over time
      - Users can select frequency of time range
        - Daily, weekly, or monthly
    - Proportion among categories
      - Visualization of a selected variable among a set of categories
    - Relationships between factors and situations
      - "Cause and effect" visualization
    - Distribution over FL counties
      - Heat map of Florida counties showing the prevalence of a variable in that county
  - Each visualization method has a customizable time range based off of the chosen dataset
- View Plot
  - Users can view a list of their created plots and select individual ones for more detailed viewing

### EVALUATION

- Majority of time was spent on design decisions
  - Crucial to success of project due to large amounts of COVID-19 related data
- Utilized the Python pickle module to serialize and deserialize datasets for more efficient storage



Interface for plot creation (Trends over time)



Resulting plot

### FUTURE WORK

- Allow users more customizations with operations they can carry out on variables in a dataset
  - Ex. Calculating positivity rate per 1,000 people in a specific Florida county
- Expand the scope of the dashboard to the entire United States instead of just the state of Florida

### ACKNOWLEDGEMENTS

- Dr. Chan for his guidance throughout the project
- Open source tools pandas, Django, and Highcharts.JS for aiding in development

# Task 1 - Selecting a variable used in the different plots

- Scope of our project was reduced
- Focus on producing 4 main plot types and reduce the customizability of our product.
- Adjust the UI to have the user select dataset/variable pairs

The screenshot shows a web application interface with two main panels. The left panel, titled "Select Plot", contains a "Name:" text input field, a "Plot Type:" dropdown menu currently set to "Table", and a "Load Plot Settings" button. The right panel, titled "Select Dataset and Variables", contains a "Filter Datasets:" section with checkboxes for "Private" (checked), "Curated", and "Shared". Below this is a "Select Variables:" section with a "+" button and two dataset selection dropdowns: "Dataset US states population - Private" and "Dataset Testing data - Private". To the right of these is a "Variable" dropdown menu currently set to "state\_name", with a list of variables shown below it: "timestamp", "reporting\_lab", "negative" (highlighted in blue), "positive", and "total". A "Create Plot" button is located at the bottom right of the interface. A copyright notice "Copyright © Senior Design Covid 2020" is visible at the bottom center.

**Select Plot**

Name:

Plot Type: Table ▼

Load Plot Settings

**Select Dataset and Variables**

Filter Datasets: ☒ Private ☐ Curated ☐ Shared

Select Variables: +

Dataset US states population - Private ▼

Dataset Testing data - Private ▼

Variable state\_name ▼

Variable

- timestamp
- reporting\_lab
- negative**
- positive
- total

Create Plot

Copyright © Senior Design Covid 2020

# Task 2 - Trends over time

- Built a new UI/card specific to line charts.
- Select:
  - Start and end date
  - Sampling frequency
  - Display name

The screenshot displays a web application interface for creating a line chart. The interface is organized into a sidebar and a main content area. The sidebar on the left contains navigation links: 'Dashboard', 'Public Data', 'Profile', 'Private Data', 'Upload Dataset', 'Create Plot', 'Public Dataset Application', and 'View Applications'. The main content area is divided into three panels. The top-left panel, 'Select Plot', includes a 'Name' field with the text 'Total Tests and Positive Cases' and a 'Plot Type' dropdown menu set to 'Trends over time'. Below this is a 'Load Plot Settings' button. The top-right panel, 'Select Dataset and Variables', features 'Filter Datasets' with radio buttons for 'Private' (selected), 'Curated', and 'Shared'. Under 'Select Variables', there are two entries: 'Dataset' (Florida Daily Totals and Stats - Curated) and 'Variable' (totalTestResults). The bottom panel, 'Trends over time Plot Settings', contains 'Horizontal Axis' settings with 'Start Date' (01/29/2020) and 'End Date' (02/13/2021), a 'Frequency' dropdown set to 'Daily', and 'Vertical Axis' settings with 'Trend Variables' set to 'Total Tests' and 'Total Positive Cases'.

# Task 3 - Plot template for proportions among categories

- Built a new UI/card specific to pie charts.
- Select:
  - Start and end date
  - Sampling frequency
  - Variable to use as categories
  - Variable to use to aggregate

The screenshot displays a web interface for creating a plot template. It is divided into three main sections:

- Select Plot:** Contains a 'Name' input field with the value 'test', a 'Plot Type' dropdown menu set to 'Proportion among categories', and a 'Load Plot Settings' button.
- Select Dataset and Variables:** Features a 'Filter Datasets' section with 'Private' selected and 'Curated' and 'Shared' unselected. Below this is a 'Select Variables' section with a '+' button and two rows. Each row has a 'Dataset' dropdown set to 'FDOH Case Line Data - Private' and a 'Variable' dropdown. The first row's variable is 'Gender' and the second is 'Case\_'.
- Proportions among categories plot settings:** Includes 'Start Date' (03/02/2020) and 'End Date' (11/02/2020) fields with calendar icons. It also has a 'Categories' dropdown set to 'Gender - Female, Male, Unknown,' and a 'Variable for proportions' dropdown.

A 'Create Plot' button is located at the bottom right of the interface.



# Task 4 - Factors vs Situations

- Built a new UI/card specific to scatter charts.
- Select:
  - Start and end date
  - Sampling frequency
  - Possible factor and situation from the selected variables

The screenshot displays a web application interface for creating scatter plots, divided into two main panels: 'Select Plot' and 'Select Dataset and Variables'.

**Select Plot Panel:**

- Name:** A text input field containing 'test'.
- Plot Type:** A dropdown menu with the selected option 'Relationship between possible factors and situations'.
- Load Plot Settings:** A button located below the Plot Type dropdown.

**Select Dataset and Variables Panel:**

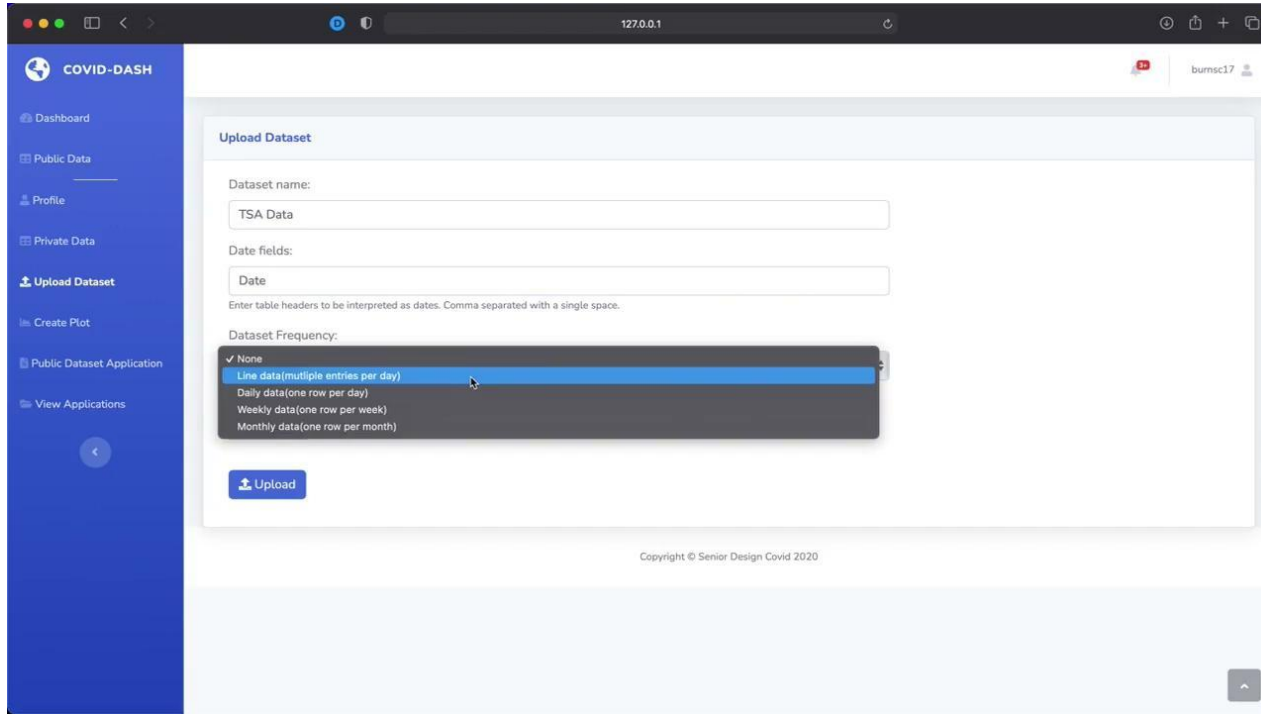
- Filter Datasets:** Three checkboxes: 'Private' (checked), 'Curated' (unchecked), and 'Shared' (unchecked).
- Select Variables:** A section with a '+' button and two dataset selection options:
  - Dataset:** 'US Google Mobility - Private' (selected).
  - Dataset:** 'TSA travelers - Private' (selected).
- Variable:** A dropdown menu with the selected option 'grocery\_and\_pharmacy\_percent\_change\_from\_baseline'.
- Variable:** A dropdown menu with the selected option 'Value'.

**Relationship between possible factors and situations plot settings Panel:**

- Start Date:** A date input field showing '02/15/2020'.
- End Date:** A date input field showing '02/23/2021'.
- Possible Factor:** A dropdown menu with the selected option 'Value'.
- Situation:** A dropdown menu with the selected option 'grocery\_and\_pharmacy\_percent\_change\_from\_baseline'.

A 'Create Plot' button is located at the bottom right of the interface.

# Project Demo



The screenshot displays the 'COVID-DASH' web application interface. On the left is a blue sidebar with navigation links: Dashboard, Public Data, Profile, Private Data, Upload Dataset (highlighted), Create Plot, Public Dataset Application, and View Applications. The main content area is titled 'Upload Dataset' and contains the following fields:

- Dataset name:** A text input field containing 'TSA Data'.
- Date fields:** A text input field containing 'Date'.
- Dataset Frequency:** A dropdown menu with the following options:
  - None (checked)
  - Line data(multiple entries per day)
  - Daily data(one row per day)
  - Weekly data(one row per week)
  - Monthly data(one row per month)

Below the frequency dropdown is an 'Upload' button with a document icon. A small text note below the date field reads: 'Enter table headers to be interpreted as dates. Comma separated with a single space.'

The browser's address bar shows '127.0.0.1' and the user's profile 'burnsc17' is visible in the top right corner. At the bottom of the page, a copyright notice states 'Copyright © Senior Design Covid 2020'.

# Lessons Learned

- *Project Scope*
  1. More complicated and involved than anticipated.
  2. Should have descoped earlier
- *Target Audience*
  1. Closer identification of what kind of user would be interested
  2. Would have allowed us to tailor development

Questions?