CO-MAP-V

COVID-19 Massachusetts Map + Visualization Tool

CSE 583 Final Presentation
Autumn 2020

Aja Sutton
Andrew Teng
Jason Thomas
Nanhsun Yuan

Background

COVID-19 has presented challenges to the public as well as researchers.

Obtaining health information can present challenges due to patient privacy.

Our data are a simulation of a COVID-19 outbreak in Massachusetts at the county level, January to March 2020.

Synthetic data may provide an avenue for researchers to better understand real-world data trends without the need to overcome the obstacles involved in obtaining real patient data.

Goals

1: To create a geovisualization tool to show how the COVID-19 outbreak has progressed over time in the state of Massachusetts.

2: To visualize aerial data (density, like case counts etc.) grouped by geographic level and to render comparison charts.

3: Utilize the synthetic aspect of the data to create a framework that can be easily implemented by others using the same common data model.

Data used

Our data is a simulation of a COVID-19 outbreak in Massachusetts at the county level, January to March 2020. Synthetic data may provide an avenue for researchers to better understand real-world data trends without the need to overcome the obstacles involved in obtaining real patient data.

Contains the following tables needed for analysis: `person`, `location`, `death`, `condition_occurence`.

- `location` contains county names which we use for spatial joins/visualization.
- COVID-19 info (cases) is in the condition table.
- Person table includes gender, race, ethnicity, and death date.

Dataset is created by Synthea and converted into OMOP common data model.

Mimics the start of pandemic by having a span of three months (Jan. to Mar. 2020).

Incorporates 2010 Census data.

GeoJSON file of Massachusetts counties.

Use cases

Use Case 1: Health Analytics Personnel who:

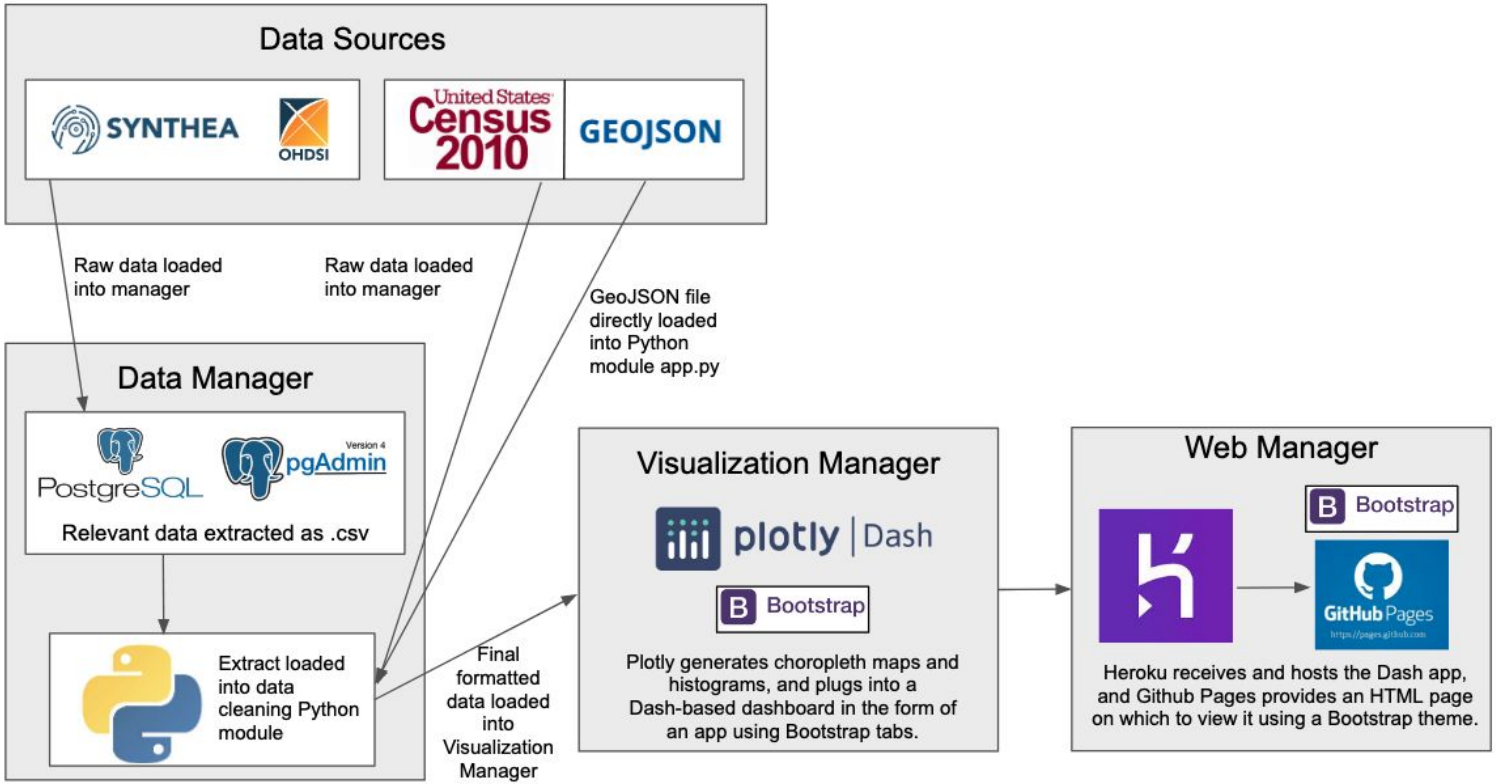
- Have limited-to-no experience with Geographic Information Science (GIS).
- Are interested in geo-visualizing disease distribution data based on raw input (e.g. counts).
- Have time-series data in discrete bins they want to interrogate.
- Want to be able to convey numerical data visually to an audience with variable expertise.

Use cases

Use Case 2: Health Analytics Personnel or public health enthusiast who:

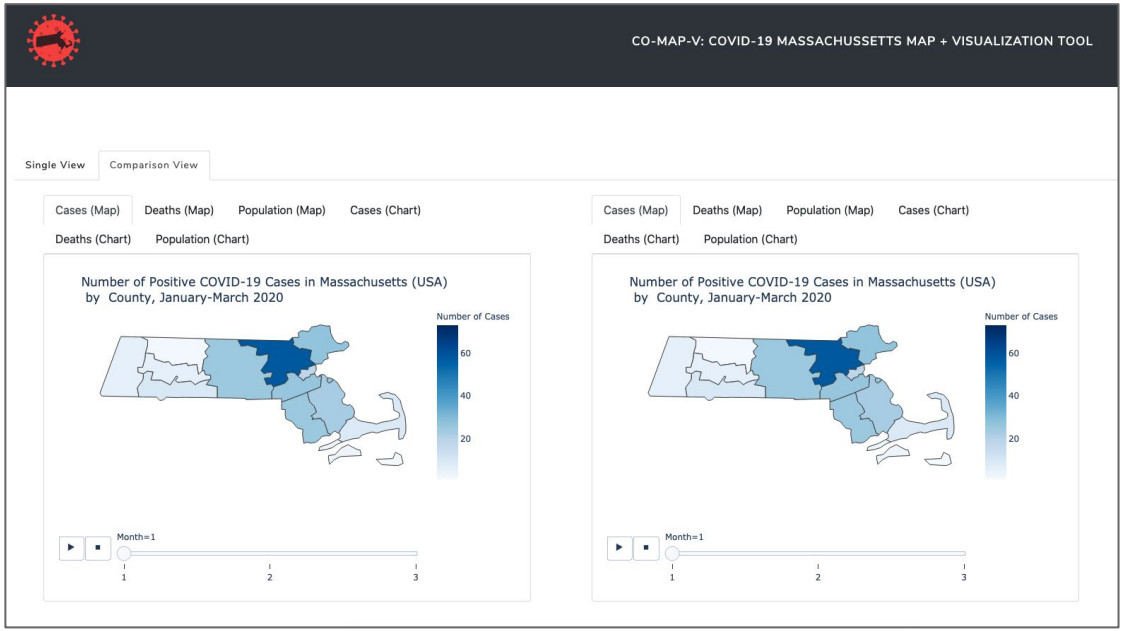
- Is interested in visualizing their own COVID-19 dataset, whether it be more rich and detailed synthetic data, real COVID-19 data, or data from a different state.
- Is interested in finding and utilizing a “plug-and-play” solution to visualize data in a manner that would require minimal effort. The user would need to download or clone or repository containing our step-by-step README file (including using our SQL query and Python scripts, and locating an appropriate GeoJSON file).

Workflow + design



Project Structure

```
co-map-v.github.io/
├── comapv/
│   ├── app.py
│   ├── __init__.py
│   ├── data/
│   │   ├── Old/
│   │   │   ├── ...
│   │   ├── __init__.py
│   │   ├── covid_ma_positive_death_counts.csv
│   │   ├── data-1605136079581.csv
│   │   ├── data_backup/
│   │   │   ├── ...
│   │   ├── data_clean.py
│   │   ├── data_cleaning.py
│   │   ├── ma_map.geojson
│   │   └── population2010.csv
│   └── tests/
│       ├── __init__.py
│       ├── smoketest_data.csv
│       ├── tests.py
│       └── tests_viz.py
├── docs/
│   ├── Component Specification.pdf
│   ├── Final presentation.pdf
│   ├── Functional Specification.pdf
│   ├── Technology review presentation.pdf
│   ├── index.html
│   └── website/
│   │   ├── ...
├── .travis.yml
├── LICENSE
├── Procfile
├── README.md
├── environment.yml
├── requirements.txt
└── setup.py
```



https://co-map-v.github.io/

Lessons learned

Lessons we went over last week:

- Wise decision to test components and research tools prior to development and implementation:
  - Plotly vs. Folium.
  - Never a bad idea to consider if your libraries are developed in association with each other (e.g. Plotly and Dash integrate well).
  - Document and test as you go.
- Have a comprehensive understanding of components before attempting implementation.
- Scoping the project is very important -- modularize goals.

Final lessons:

- Consider continuous integration options other than Travis CI (it's currently receiving so much traffic it's too slow to reliably use). One option to look into: GitHub Actions.
- Designate a project manager to make sure that everyone knows what their roles are toward the end of the project (no duplication of tasks this way).

Thanks!