311 Service Mapping

Team Members:

Matt Kaufmann (mkaufmann1)
Angela The (angelathe)
Vignesh Venkatachalam (vigneshv)
Eujene Yum (eujeneyum)

Project Overview:

Using 311 Service Request information from the Chicago Data Portal, we have analyzed each neighborhood's "responsiveness" to services requests. To assess responsiveness, we have looked at metrics such as average resolution time, median resolution time, and percent unresolved. We have pulled data from 2019 - 2021 to capture a pre-peak-decline analysis of the pandemic impacts. Additionally, we look at responsiveness and overall amount of 311 requests per capita in the context of socio-economic information such as unemployment, income, and race from the American Community Survey.

Our two-pronged approach involved:

- Visualizing 311 responsiveness vis-a-vis demographics, and
- Conducting a deep dive on the relationships between these

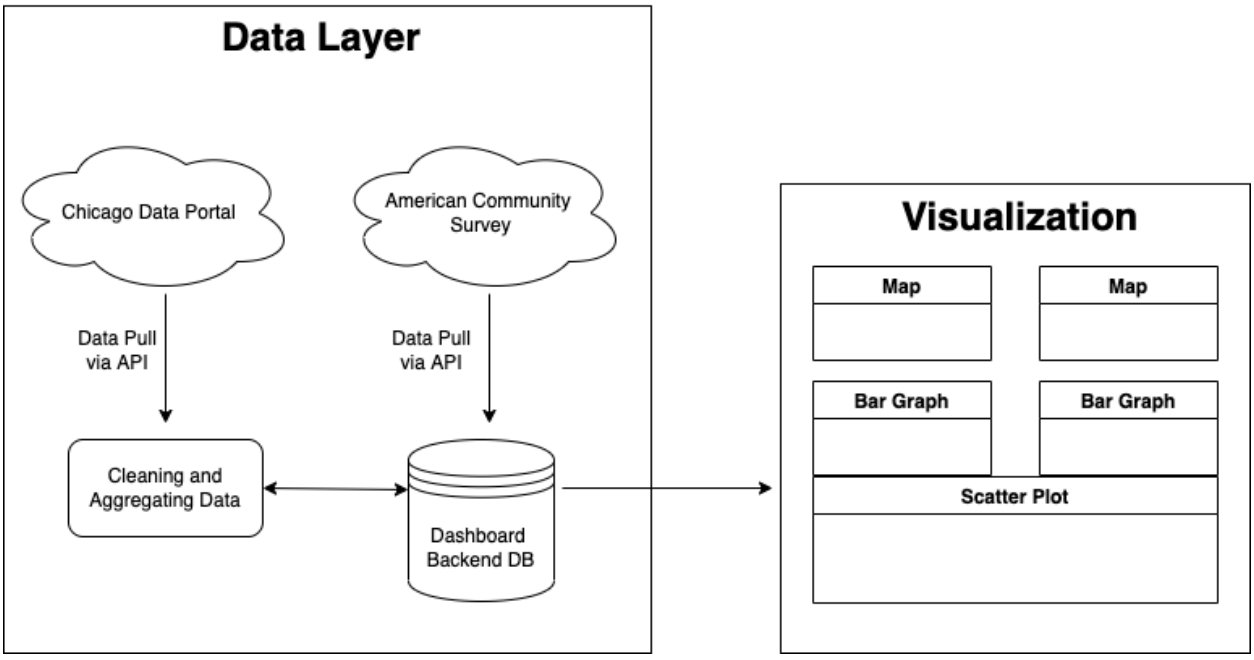
To help members of the public easily view this information, we have created a web application dashboard with Dash that allows a person to easily filter what type of demographic information or 311 responsiveness metric they want to view to create a heat map across Chicago neighborhoods.

This allows for a quick comparison across neighborhoods. To the extent a person wants to go deeper into one neighborhood, we have broken out response time by neighborhood, and compare summary responsiveness statistics to Chicago average to allow for the information to be contextualized.

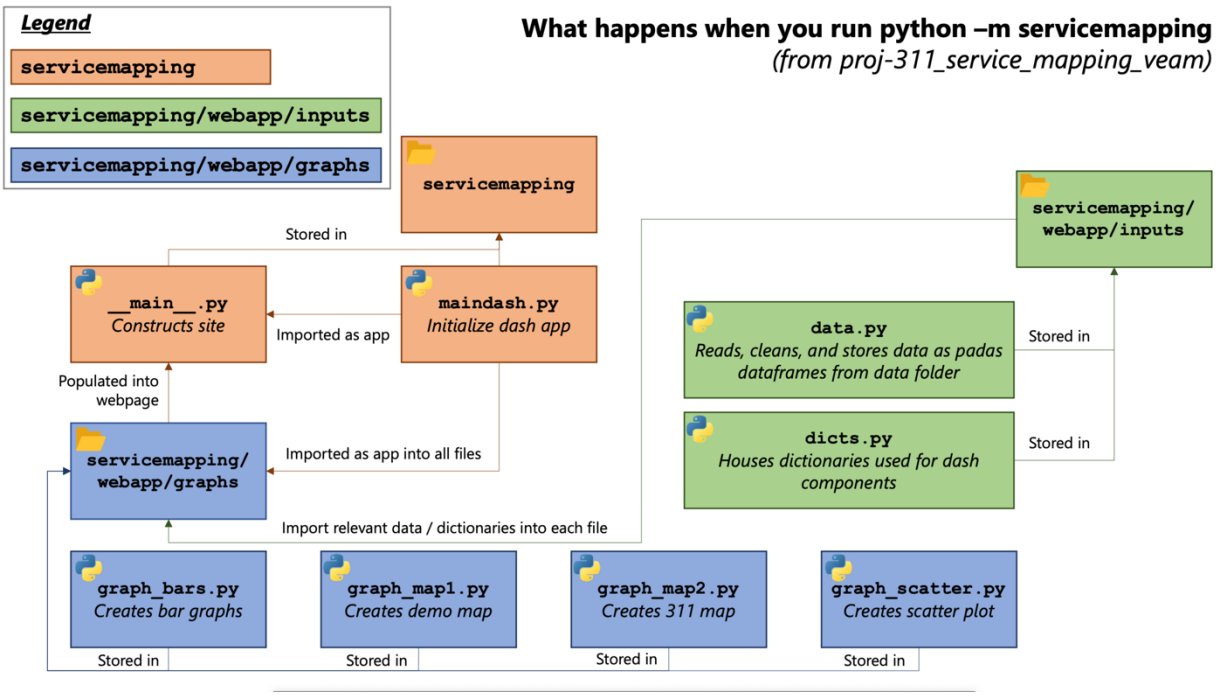
Finally, to help drive home relationships between demographics and requests, we have created a scatterplot that calculates a correlation coefficient for a user inputted demographic category and responsiveness category of Chicago neighborhoods.

Software Architecture:

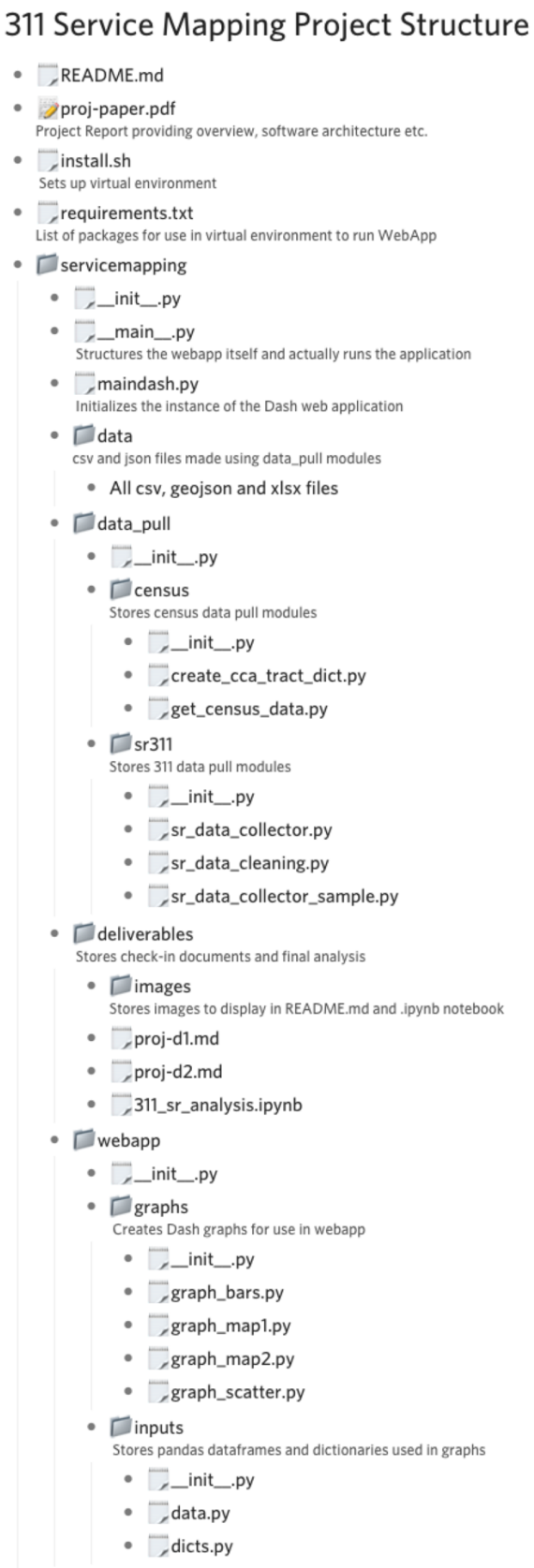
Flow Diagram



WebApp Architecture



Project Folder Structure



Code Responsibilities:

Name	Modules	Tasks	Files
Angela / Matt	Research and documentation	• Research possible dashboard platforms that would	Internal google doc with hyperlinks to resources

	Data pull	• Create dictionaries to map census tract to community area and community area number to name to be utilized in filtering and outputting data
		• Create a function to pull and filter American Community Survey utilizing the API interaction package "censusdata"
			• Data_pull folder >> census >>
			  • create_cca_tract_dict.py
			  • get_census_data.py
			• Data folder
			  • census_demos.csv
			  • census_demos_pop.csv

	Visualization	• Utilize Dash to create webpage structure and visualizations to communicate the ACS and Chicago 311 data in an intuitive, interactive way
		• Modularized the structure of the website to make code easier to digest
		• Create multi-layered dropdowns for filtering data visuals, with dynamic updating of options and data provide maximum flexibility to end user
			• __main__.py
			• Maindash.py
			• Webapp folder>> graphs
			  • graph_bars.py
			  • graph_map1.py
			  • graph_map2.py
			  • graph_scatter.py
			• Webapp folder>> inputs
			  • data.py
			  • dicts.py

Vignesh / Eujene	Data pull	• Create a function to pull filtered service request data from the Chicago data portal using an API call
		• Created multiple functions to manipulate pulled data in order to calculated responsiveness metrics
		• Tailored certain metrics to better-fit visualization requirements
			• Data_pull folder >> census >>
			  • sr_data_cleaning.py
			  • sr_data_collector.py
			  • sr_sample_data_collector.py
			• Data folder
			  • sr_census_df.csv
			  • chicago_df.csv
			  • 311_census_bar.csv

	Analysis	• Documented our thought process, assumptions, and learnings • Generated insights from the dataset as well as visualizations	• Data_pull folder >> census >> 311_sr_analysis.ipynb
	Research and documentation	• Put together cohesive guidelines on how to use the software as well as a synthesis of the project	• README.txt • requirements.txt • proj-paper.pdf • install.sh

Instructions to execute project codes

NOTE: All codes to be run from within the project root directory

Setting up Virtual Environment and installing required packages:

1. Clone the repo
2. From within project root directory `proj-311_service_mapping_veam` run `sh install.sh` (takes ~2 minutes for all packages to install)
3. Activate the virtual environment through `source env/bin/activate`

Viewing WebApp:

1. Run `python3 -m servicemapping` (takes ~2 minutes)
2. Follow the generated URL link (eg: http://127.0.0.1:8050/) by copying it and pasting in your browser (On Mac, use ⌘ and click `Follow link` on the link generated)
3. The WebApp might take a few seconds to load all the maps and charts

(Optional) Pulling Data from Data Portals

To view sample data API-pull from Chicago City Data portal:

1. Run `python3 -m servicemapping.data_pull.sr311.sr_sample_data_collector`
2. Sample dataset created is stored in `servicemapping/data/sr_sample_raw.csv`

To view census data API-pull from American Community Survey:

1. Run `python3 -m servicemapping.data_pull.census.get_census_data.py`
2. Sample dataset created is stored in `servicemapping/data/census_demos.csv`

To recreate underlying datasets - by re-pulling Service Request data from Chicago City portal: (Run Time: ~5 mins)

1. Run `python3 -m servicemapping.data_pull.sr311.sr_data_cleaning`
2. 3 datasets (`311_census_bar.csv`, `sr_census_df.csv` and `chicago_df.csv`) are created and stored in `servicemapping/data/`

Interacting with the WebApp

Responsiveness Map:

1. On the left map, user can select specific demographic filters (and sub-filters) such as Race, Income Range and % Unemployed. The heatmap (and corresponding color gradation) will dynamically update for all Chicago community areas. Hovering over any community area shows the community area name and the corresponding metric value
2. On the right map, user can choose from one of the 3 responsiveness metrics: Median Resolution Time, Average Resolution Time, Average Annual Service Requests per 1000 people, and the heat map will dynamically update for all Chicago community areas. **Hovering over any community area shows a tool-tip with community area name, the corresponding metric value and the Top-3 service requests for that particular community area in the past 3 years**

Community Area Deep dive:

User can **use the drop down to choose** the community area of interest, for which % of requests completed in different time buckets is shown for each year from 2019 to 2021. The bar chart also shows the % of issues unresolved. User can **click on any of the legends for the year** to view data for only a particular year or any pair of years. We've also provided a visual to compare the Median Request Resolution Time of the Community Area with Chicago as a whole

Scatter Plot comparing responsiveness to Demographics:

User can **use the drop down to choose** the demographic attribute (Race, Income Range and % Unemployed) and the responsiveness metrics (Median Resolution Time, Average Resolution Time, Average Annual Service Requests per 1000 people) of interest, and the **scatterplot dynamically updates to show a scatter plot of responsive metrics vs demographic attribute values, along with the trend line**. Each bubble in the scatterplot corresponds to a community area, with its size indicating the population of the community area. Specific values can be gathered by hovering over the bubble.

What the project tried to accomplish & what it accomplished:

We set out to understand the relationship between 311 service requests in Chicago and the demographic makeup of the different neighborhoods of our city.

We were able to tie census data and 311 information at a community area (neighborhood) level, allowing us to overlay service request metrics (e.g. average resolution time, # of requests/1000, etc.) with community demographics (e.g. race, income, etc.). We were able to translate that information into an interactive dashboard with:

- Two heat maps (one based on Census demographics, and one based on 311 service-request metrics)
    - The service-request metrics map includes a tool-tip that sheds light on the top three service request types of a given community area.
- Two bar graphs (one comparing request-resolution time buckets within a neighborhood, and one comparing 311 request statistics between the neighborhood to Chicago overall)
- A scatter plot with a trend line to view the correlation between 311 request statistics and Census demographics (plotting community areas with dynamic y and x axes)

We also utilized these charts to come up with preliminary analyses and gain an understanding of what potential next steps could look like in `311_sr_analysis.ipynb`. Below are some highlights.

- From amongst the unresolved service requests, below are the ones that have highest percentage of unresolved each year:
    - Tree Planting Request
    - Sidewalk Inspection Request
    - Lead Inspection Request
- From our dashboard, we can observe that...
    - In neighborhoods with a higher percentage of the population making minimal income, there are more service requests made. We see the opposite trend for high-earning neighborhoods.
    - Compared to Chicago as a whole, Hyde Park's median request resolution time is now on par, as opposed to the lag in 2019.