**Developing an online interactive map for**

**Chicago Urban Flooding Research**

#### **Introduction**

Urban flooding is an urgent problem that needs to be studied and solved especially as the acceleration of climate change. More and more extreme weather leads to more frequent storms that exceed the capacity of city sewer systems and cause urban flooding. However, the reasons for urban flooding are not just climate change and obsolete sewers. And the flooding spots are also not everywhere. Where are the flooding susceptible areas? Why are some areas more inclined to have floodings than other areas? To solve this problem, basically, we need to figure out WHERE and WHY.

This project is going to create two online interactive maps and plot two related chats. One of the maps is the choropleth map of City Chicago flooding insurance claims data that is geographically collected by zip codes. From this map, we can know which area has the highest and lowest flooding risk in Chicago. And by overlaying with related demographic data, by plotting and reading the Parallel Coordinates Plot, we can know the correlations and patterns among flooding risks, percentage households of color, and income. The scatter plot is used to support the PCP plot further. The other map displays the overlay of several layer groups, which are supposed to include Chicago river systems, land use, flooding susceptibility index, and surface total imperviousness level. All the layer groups mentioned would help users geographically figure out the correlations among flooding, land use, urban surface, and hydrology.However, due to unexpected errors when parsing geojson, currently, only the Chicago river systems layers are displayed successfully. The obstacles of data processing would be described in a later methodology section. Through the two maps and two plots, users can explore the question WHERE and WHY.

In addition to the geodata display and analysis, this project would take advantage of the interactivity and accessibility of the online mapping technology. Compared to professional softwares, like ArcGIS, this online mapping program produced by leafletJS, D3JS, PlotlyJS is much lighter and easier to play with by more broad non-professional users.

#### **Background**

The background includes two perspectives: One is the author's academic background and personal hobbies. The other one is about the popularization of GIS and ambitions to use interactive online maps to facilitate the communication between citizens and the municipality.

The author’s major is landscape architecture. The simplified interpretation of this discipline is Architecture plus Nature, which is about the relationships between built environments and nature. It’s a broad and diverse area consisting of designing living space, topography, ecology, construction, buildings, urban planning, stormwater management, aesthetics, history, etc. And the author is currently a research assistant focusing on urban stormwater management and green infrastructure, which is also the area the author is interested in. Therefore, the author decides to use urban flooding as a topic for this project, and much data used in this project comes from the author's stormwater research as well.

Bringing the mapping from closed software to open-sourced online interactive format would be revolutional for the communication between professionals and average citizens. In the author’s academic class, like, design studio, ArcGIS is usually used for site analysis at the early stage of a design project. However, from the author’s observation and experience, the bar of using ArcGIS is relatively high. Users need to worry about software installation, license authorization, extremely complex interface and toolboxes, steep learning curves, etc. It’s not friendly to newbies and not accessible for non-professionals. However, the designing/building of a city environment does not just account for professionals. The citizens are important as well or even more weighted. Citizens also want and need to know the site conditions, and communicate with professionals. The traditional approach for that communication is that the professionals plot a static map from ArcGIS and insert it into a powerpoint, and display that to citizens and persuade citizens by words. This is not efficient, and easily leads to misunderstanding. What if you bring all the data and mapping to an online interactive map? Citizens could access and play with the map anywhere if they have the internet and a browser. The bar becomes much lower and citizens get more engaged into the site analysis process and more informed about the site conditions. This is the reason why the author takes this class and develops this final project.

#### 

#### **Methodology**

The methodology of this project is:

1. Collecting and parsing data. Data are downloaded from multiple Chicago data institutions, and are processed in ArcGIS, and converted to GeoJSON by QGIS and adjusted dataframe in Jupyter notebook.

2, Develop the online interactive map by several open source mapping library LeafletJS. 3, Finally, the corresponding chart would be plotted by graphing libraries D3js and PlotlyJS.

In this section, methodology would be discussed by the topics below, and errors and obstacles happened in that topic would also be discussed:

**Data Collection and Parsing**

**Choropleth Map**

**Layer Group Map**

**Parallel Coordinates Plot**

**Scatter Plot**

**Data Collection and Parsing**

Data source:

Chicago flooding insurance claims and demographic by zipcode

<https://www.cnt.org/urban-flooding/flood-equity>

Chicago Zip Code Boundaries

<https://data.cityofchicago.org/Facilities-Geographic-Boundaries/Boundaries-ZIP-Codes/gdcf-axmw>

Census tract and zip code conversion

<https://www.huduser.gov/portal/datasets/usps_crosswalk.html>

CMAP Flood Susceptibility Index

<https://datahub.cmap.illinois.gov/dataset/on-to-2050-layer-flood-susceptibility-index>

CMAP Land Use

<https://www.cmap.illinois.gov/data/land-use>

CMAP Imperviousness Area

<https://www.cmap.illinois.gov/2050/indicators/impervious-area>

Chicago Water Way

<https://data.cityofchicago.org/Parks-Recreation/Waterways/eg9f-z3t6>

All that data is downloaded and imported into ArcGIS. They need to be clipped because most dataset are super large with complex polygons and make the online maps super laggy. And due to the special projection coordinates of most Chicago geological data, they cannot be converted to correct geoJSON directly by an online converter. So the author also uses QGIS to convert the shapefile to geoJSON to get the coordinates of WSG84 projection. This is a very time-consuming process. The land use data and flooding susceptibility index still fail to be processed at the end.

Then the geoJSON data needs to be converted to variables in the js file and parsed in the js code, and generate polyline or polygon by leaflet. The total imperviousness data failed to be parsed because the error invalid Latlng object. It’s super weird because I actually can console log the converted laglngs array. This part is also very time-consuming and mental exhausting.

**Choropleth Map**

is the choropleth map of City Chicago flooding insurance claims data that is geographically collected by zip codes. From this map, we can know which area has the highest and lowest flooding risk in Chicago. Tracks could be selected among Quartile of claims, total claims, total claim amount, average payout, total households, percent households of color and media income. The classification and legend would also change correspondingly.

**Layer Group Map**

The layer group map displays the overlay of several layer groups, which are supposed to include Chicago river systems, land use, flooding susceptibility index, and surface total imperviousness level. However, due to unexpected errors when parsing geojson, currently, only the Chicago river systems layers are displayed successfully.

**Parallel Coordinates Plot**

The Parallel Coordinates Plot is used to display the correlations and patterns among flooding risks, percentage households of color, and income.

**Scatter Plot**

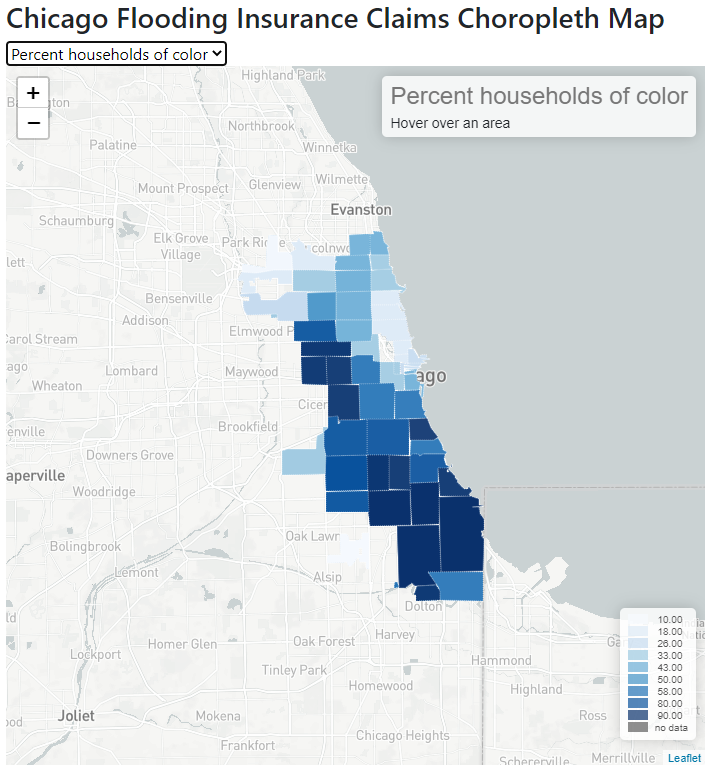
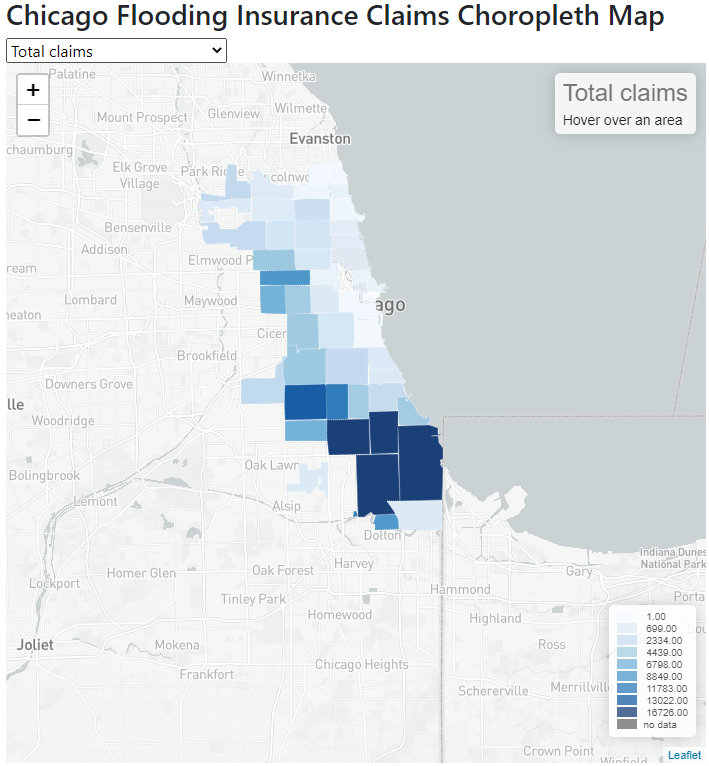
The scatter plot is used to support the PCP plot further. The x-axis is percent households of color, the y-axis is total flooding insurance claims. Users could find the correlation between these two parameters.

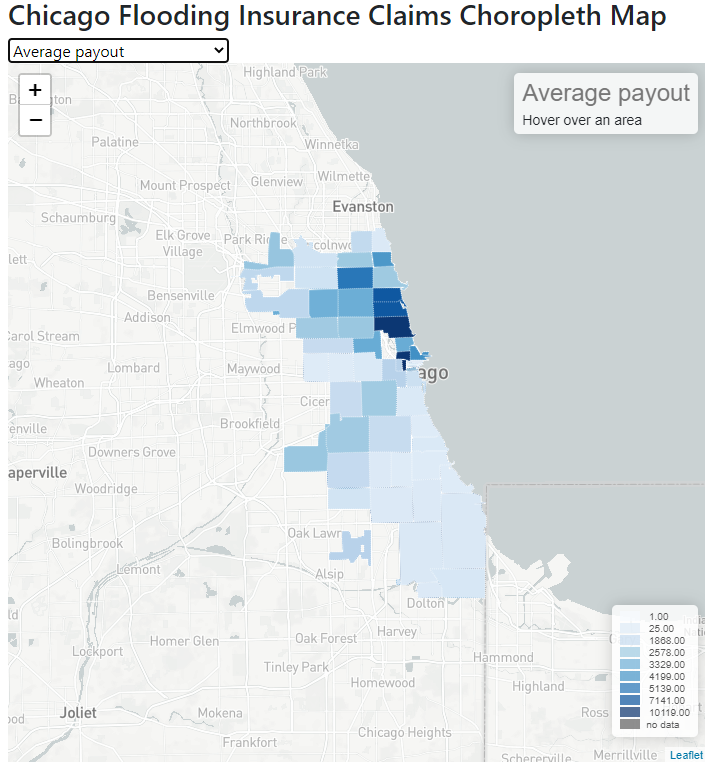
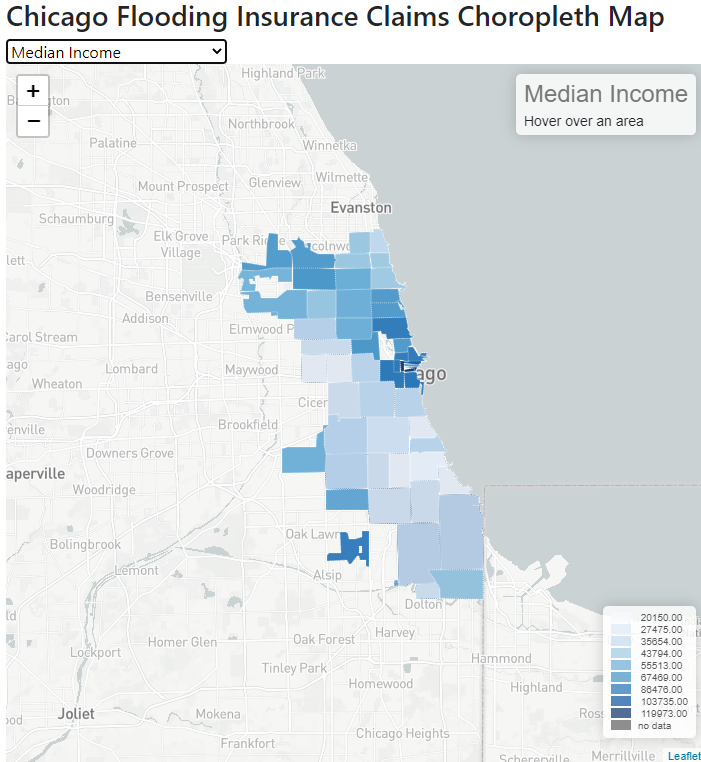
**Failed Data Parsing, Failed Connection Between Map and Failed Plot, Failed Grouped Scatter Plot**

The author tried to merge the census tract data with the flooding insurance data by the conversion standard of geoid and zip code. And the author tried to connect the map with the chart. And the author also wants to visualize the detailed census data with zip code by grouped scatterplot by D3js. However, lots of bugs happen during those processes. That’s extremely time-consuming and energy-consuming. And finally that part is not included in the product of this final project.

#### **Result and Accomplishment**

From the four choropleth maps of four factors: total claims, percentage households of color, median income, average paymout, we could get the sense that the four zip code areas: 60617, 60619, 60620, 60628 --- the south part of chicago--- suffer most from urban flooding, but get the relative high color population, low income and low flooding insurance payout, which signifies a huge inequality within the urban flooding issues in Chicago. Color and poor population are more inclined to suffer from flooding, and harder to recover from it.





For the PCP plot, every line in the chart represents the flooding claim numbers with other factors. From this plot, we could see, when the quartile of claims is high, the total claims would be high, the average payout would be low, the total households and percentage of households of color would be high, and connected median income would be low. In conclusion, the claim numbers and quartile of claims could represent the flooding risk level, and the flooding risk level and percentage of households of color have a negative relationship. And the area with a high percentage of households of color would have low median income and low insurance payment. These all signify that there is a huge inequality existing for populations with color and low-income when considering urban flooding risks.