Process Book:

Initial Project Proposal

Background and Motivation: Discuss your motivations and reasons for choosing this project, especially any background or research interests that may have influenced your decision.

My thesis research deals with modeling the transportation network impacts of inundation events on the Boston metro region. Inundation events (flooding events where networks are degraded or disabled) are serious threats to coastal cities around the world. Such events can have major impacts on regional transportation networks, as seen in NYC in the aftermath of Hurricane Sandy. I have sought to develop an understanding of potential inundation consequences on the Boston metro region by conducting two related analysis.

1. A GIS analysis- cataloging and quantifying likely impacts to transportation assets and driver of transportation demand (people & jobs)
2. A transportation modeling analysis using the Cube Voyager modeling platform to model the impact of inundation events on the regional multi-modal transportation system from 1ft to 6 ft.

This research has provided me with large amounts of data that I would like to display in an interactive web based format.

Thesis Defense PPT: Has more information on topic and outputs (If interested):

https://www.dropbox.com/s/o6vfy2cofbsonvr/DefensePPT_M30.pptx?dl=0

It is a very interesting data set that provides information on:

- the location, number and extent of transportation assets inundated at water levels from 1 to 6ft.
  - Roads, transit stops, land use, population, etc
- Modeled traffic flows given a damaged and degraded road network
- Locations of major congestion given an inundation event
- transit ridership by all transit routes and modes,
- number of trips not able to be completed due to inundation (Lost Trips & their origins and destinations)
  - by trip purpose (work, school, shopping, etc)
- Changes in accessibility by traffic analysis zone by walk, transit and automobile.

Project Objectives. Provide the primary questions you are trying to answer with your visualization. What would you like to learn and accomplish? List the benefits.

I have analyzed much, but not all, of this data using python, Arcgis, TransCAD and Excel already. I understand what kinds of questions the data can answer but I am trying to find ways to communicate this data more efficiently.

Furthermore I want to find ways to communicate and share the data with others who do not have the same software packages available, or possibly the technical expertise required to deal with the data.

The primary questions I would want a user to be able to answer after interacting with the visualization would be:

- The location, count, and extent of impact given inundation at different levels
- The shifts in transit ridership caused by degradation: For example of the Red Line is disabled what are some of the buses that experience increased ridership - i.e. people who are likely changing their route
- How accessibility changes for different modes given different inundation levels.
- Congestion Hot spots given inundation
- Be able to compare the results to a non inundated scenario

Benefits:

- Provide people with an understanding of possible inundation impacts on the transport network
- Provide understanding of the impact to peoples personal modes, streets, routes of interested. For example: allow a user to examine if at a certain inundation level the streets or bus/train they use would be compromised
- Provide policy makers with broad metrics of the expected impacts of such inundation events:
    - Total Lost trips by inundation level
    - Changes in:
        - Vehicle Hours Traveled
        - Vehicle Distance Traveled
        - Congestion Hot Spots
    - Other metrics of interest

Data: From where and how are you collecting your data? If appropriate, provide a link to your data sources.

- I already have the data- most of it is in tabular format (dbf & csv).
- I also have network data that I can convert in to Shapefiles and then into GeoJSON. The only concern with the network data is that it quite large (264,000 Links). I'm not sure if D3 will be able to display this data efficiently. I don't think topojson will make much of a difference since I am dealing with links, not polygons.
- Finding a way to display this data dynamically may be a challenge.

Data Processing. Do you expect to do substantial data cleanup? What quantities do you plan to derive from your data? How will data processing be implemented?

I do not expect to have substantial data cleanup but there will be some.

- I will need to convert much of my data in to JSON format, but I've already been playing around with they Python Pandas.toJSON function and believe that it can be used for most of the tabular data.
- I will need to develop a strategy for dealing with matrices – suggestions would be appreciated – I assume key value pairs (O-D) in a JSON format would likely be the fastest (?)

- Again the network data – I may have to use a smaller clipped network, or remove some of the minor links in the network.
- Some of the data is output as text report files, but I have already written a few python scripts that extract this data to PANDAS Dataframes, I may need to write a few more of these scripts, as there are different formats but I am fairly comfortable doing this.
- There will also be some data formatting and aggregation but between python & JS I do not anticipate the data cleaning to be require excessive amounts of time.

Visualization. How will you display your data? Provide some general ideas that you have for the visualization design. Include sketches of your design.

Accessibility Impacts of Inundation:

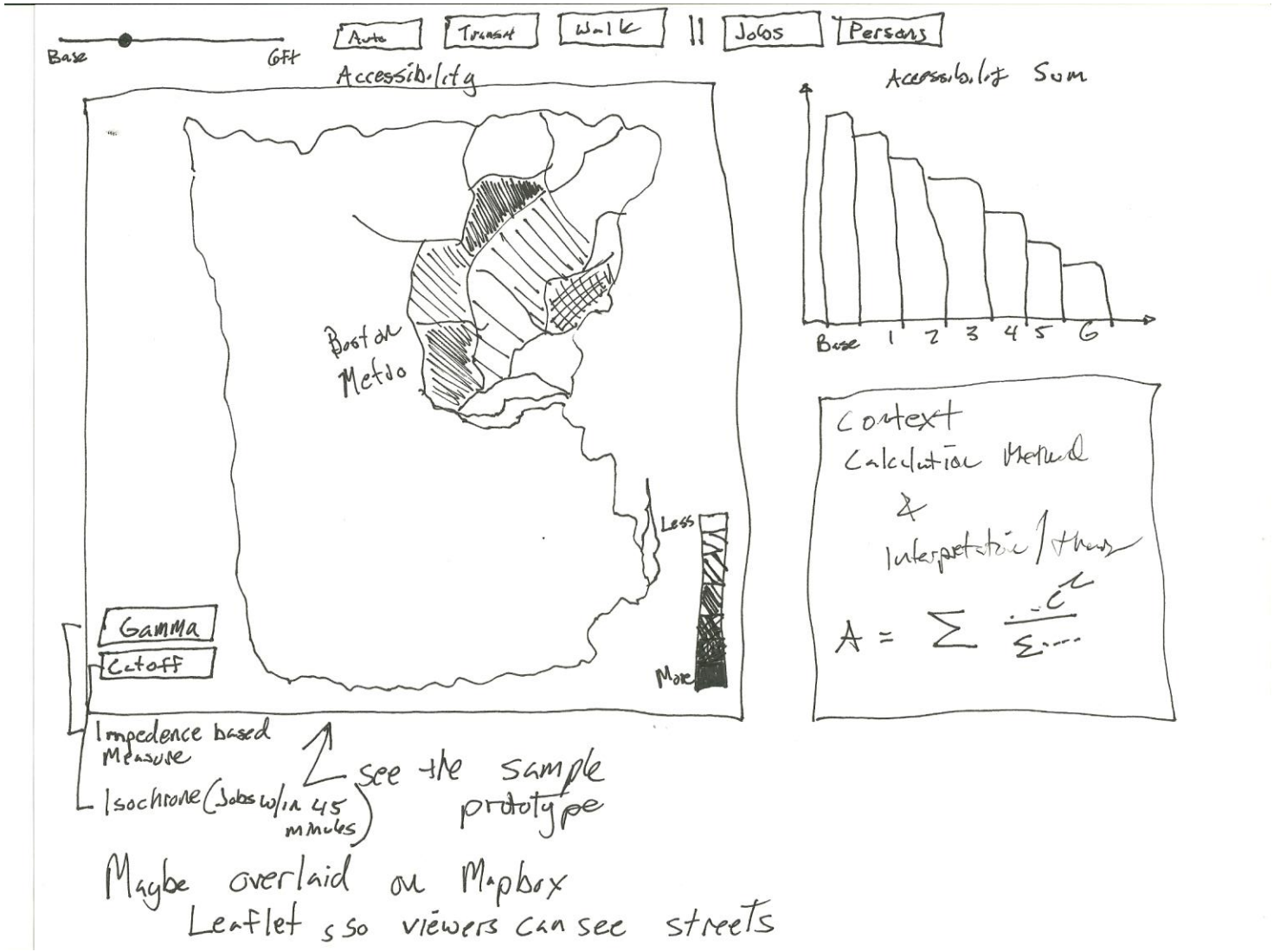


Figure 1: Accessibility Visualization Example

See Prototype Example: http://mdgis.github.io/ , still needs a lot of work but I've already started trying to create this design.

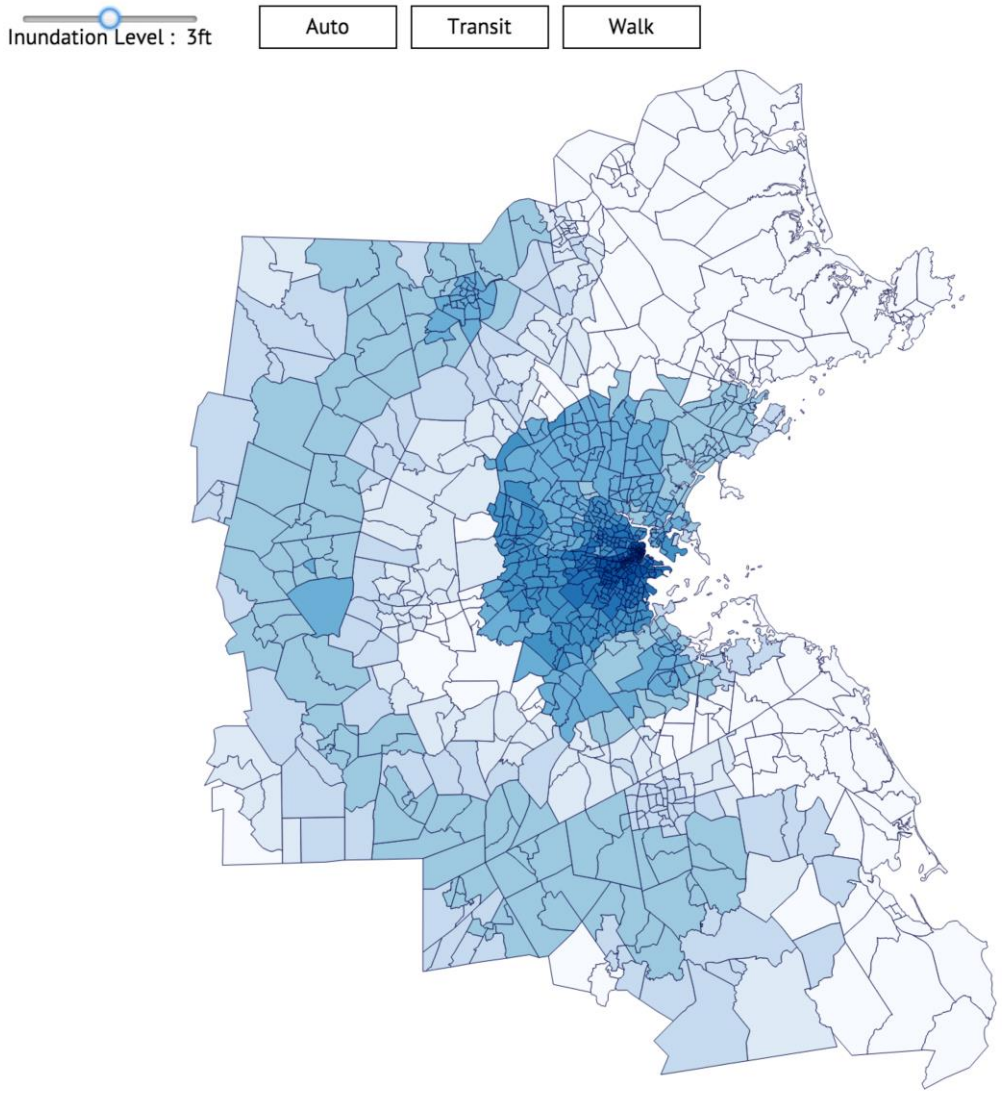


Figure 2: Screenshot of current prototype

This design will include information on the side about the how accessibility was calculated, the accessibility sum (a measure of regional accessibility) and a chart showing the change over the different inundation levels.

Inundated Assets

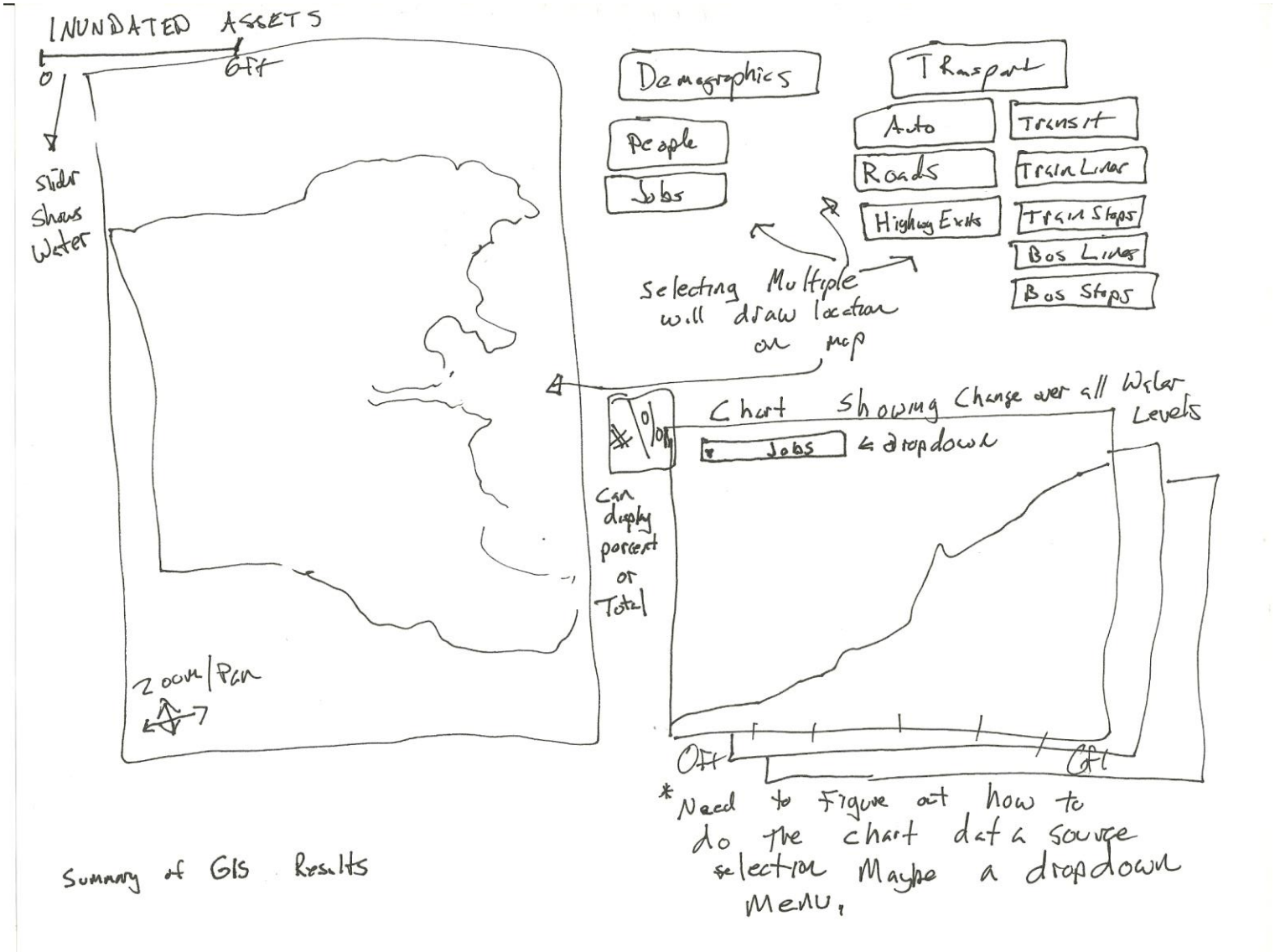


Figure 3: Inundated Assets

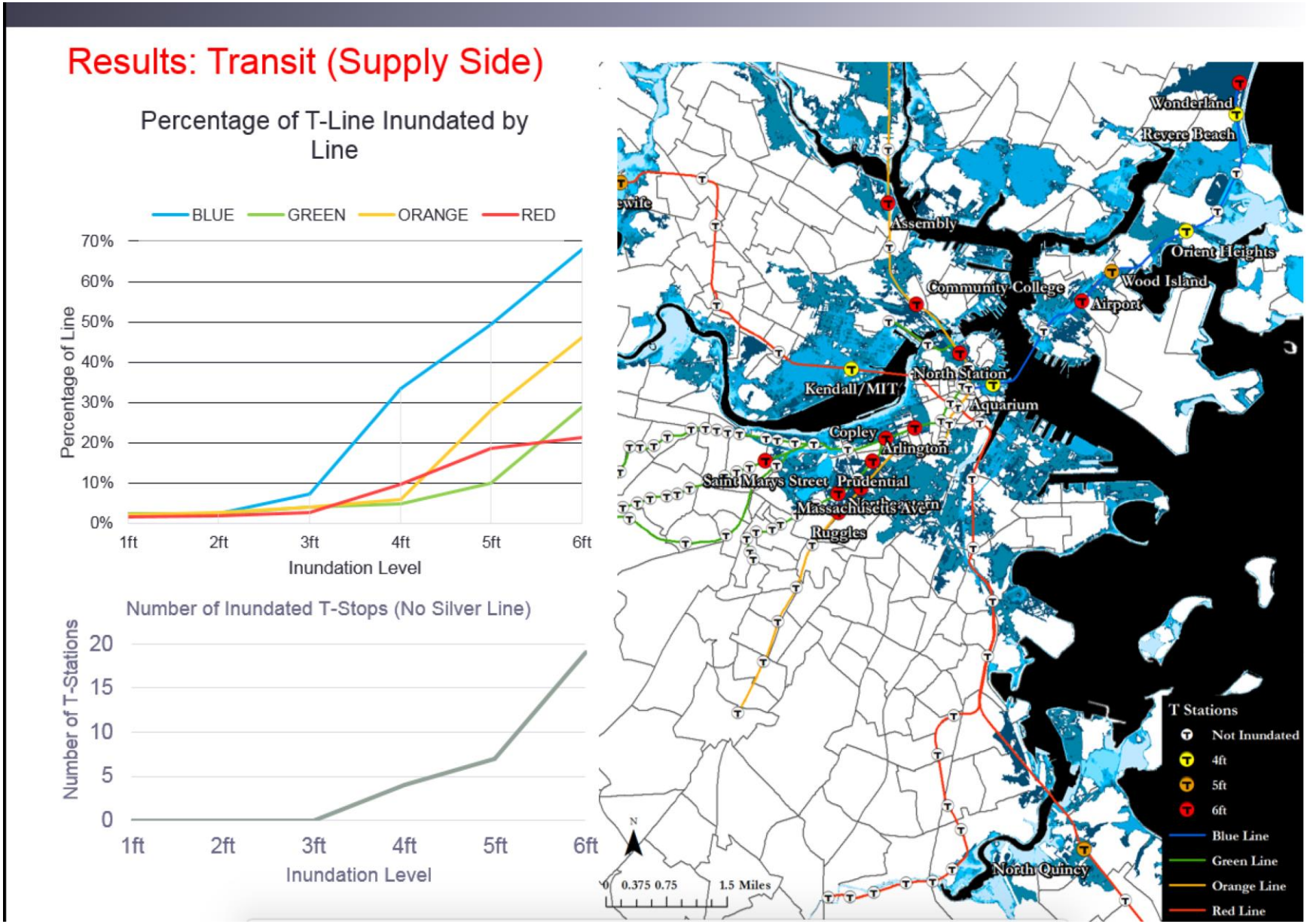


Figure 4: Example of Inundation Layers / Inundated Transit Assets / Change over Inundation Level

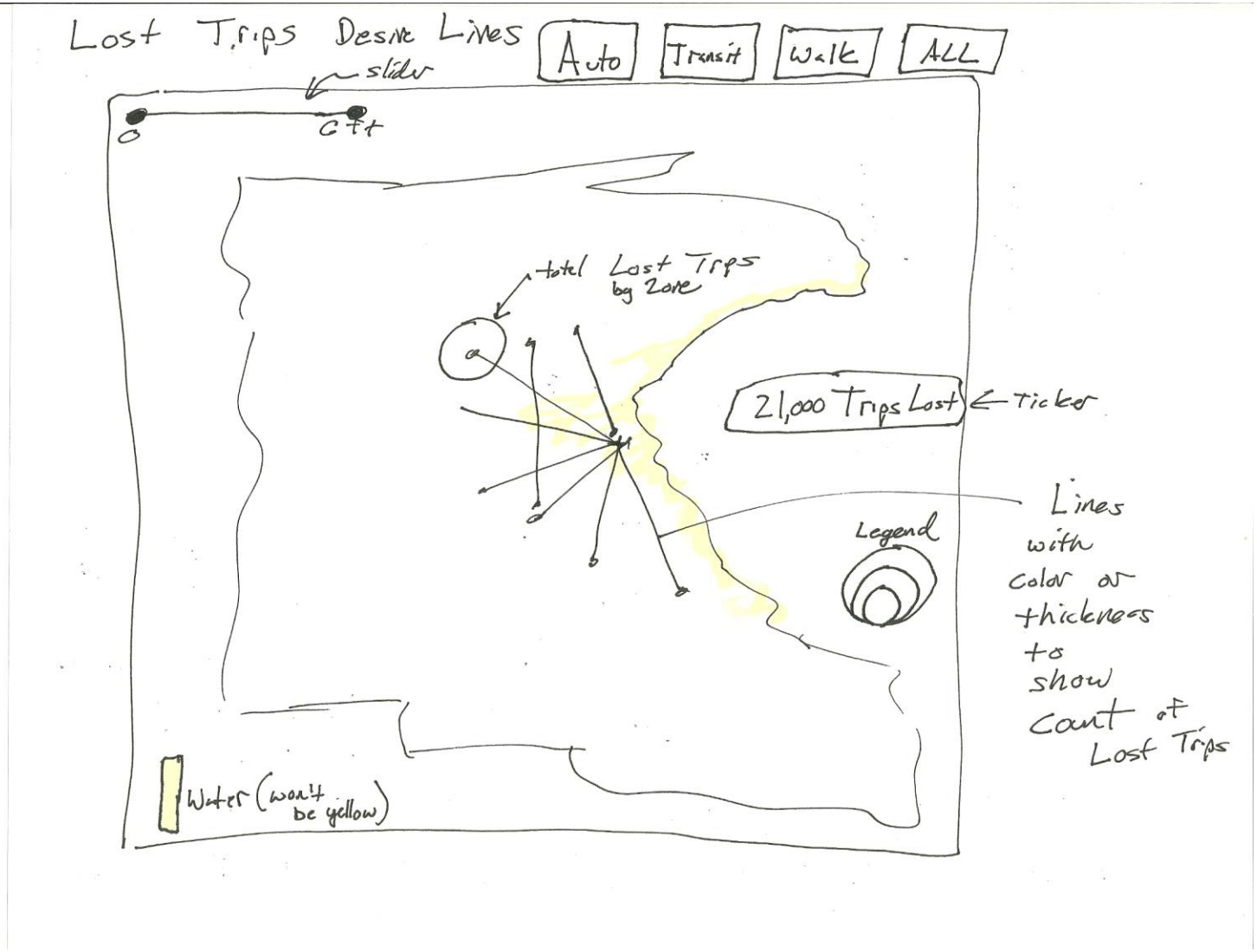


Figure 5: Lost Trips / Desire Line: Data taken from matrices with information on the trips that cannot occur because of network inundation.

Example Generated in TransCAD Transport Modeling Software:

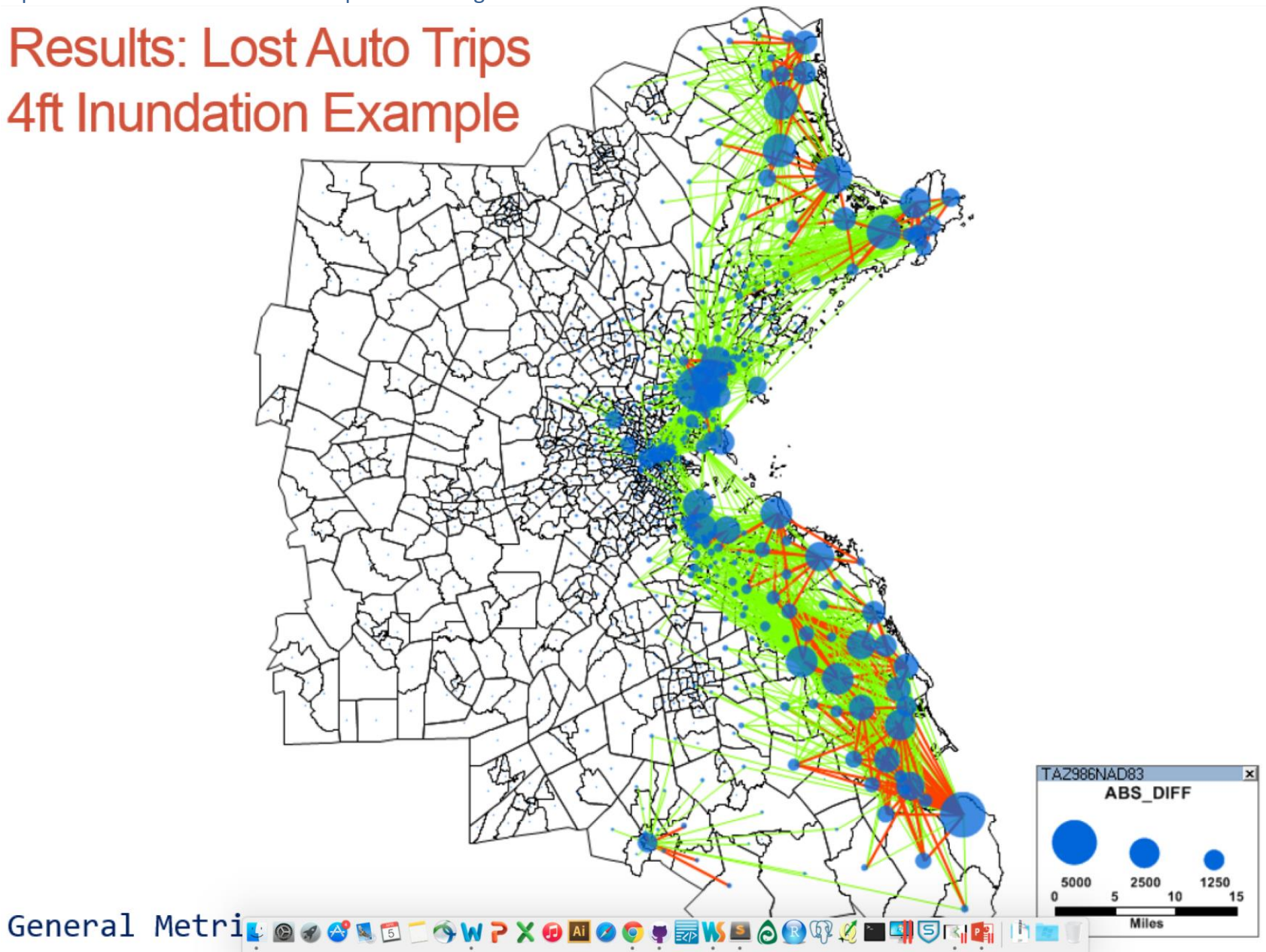


Figure 6: Example of lost trip by Auto Desire Lines – Line color indicates number of trips and circles indicate the total number of lost trips by Traffic Analysis Zone.

Lost Trips For Public Transit:

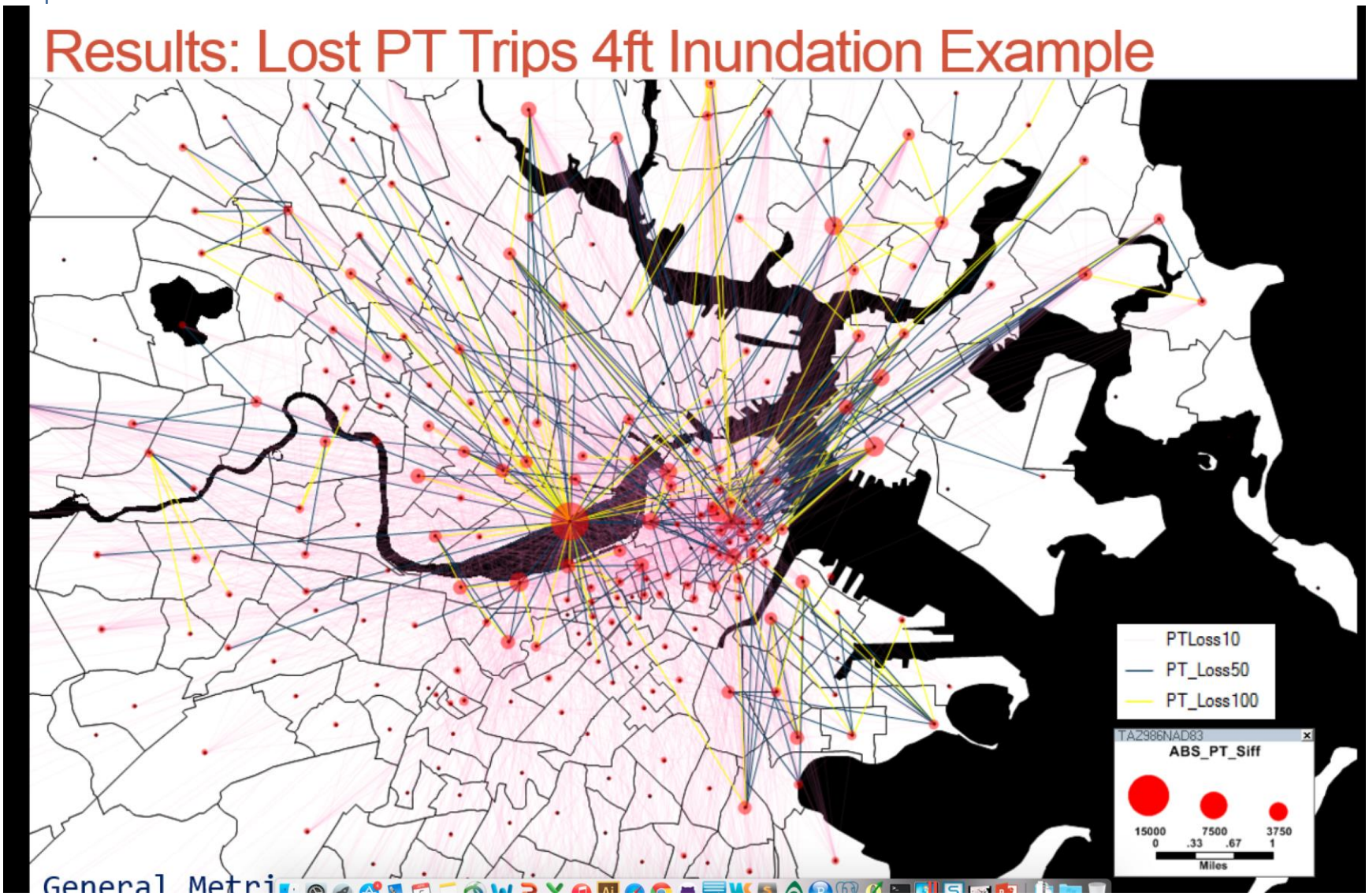


Figure 7: Same as previous but transit trips

Auto Network Summary:

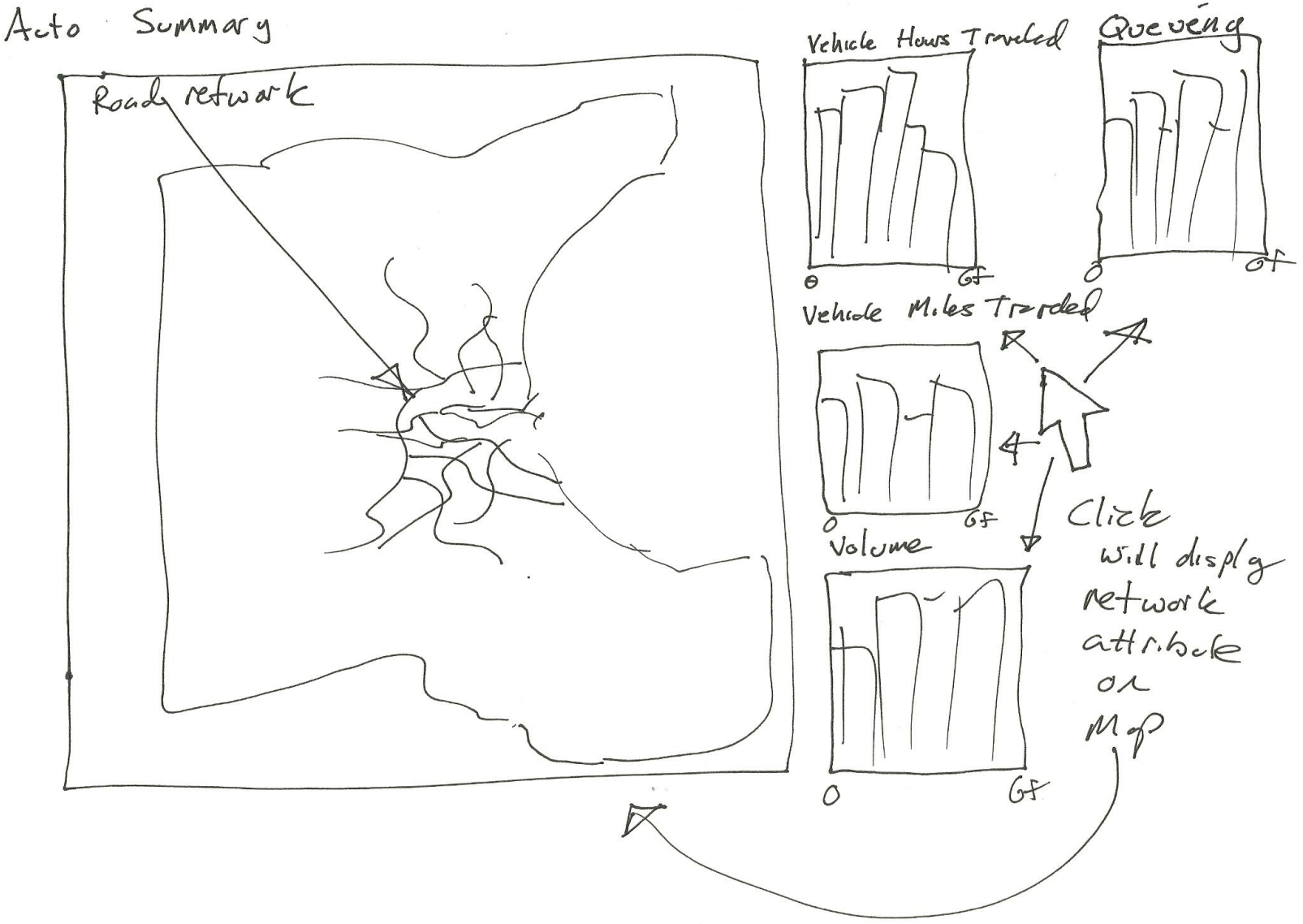
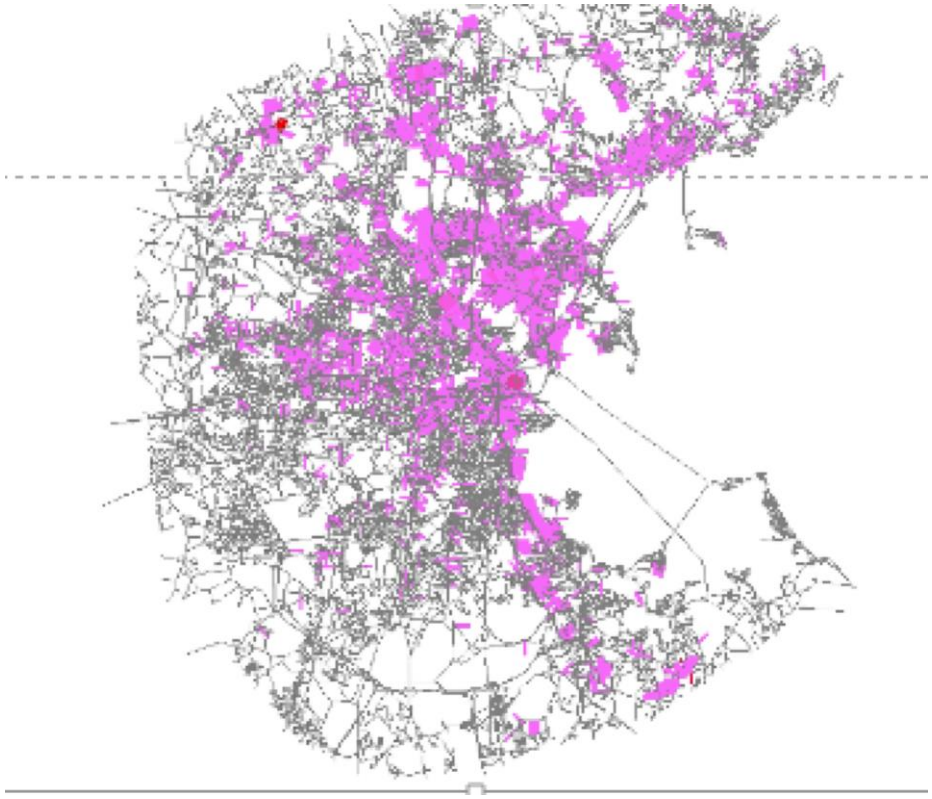


Figure 8: Auto Network Summary



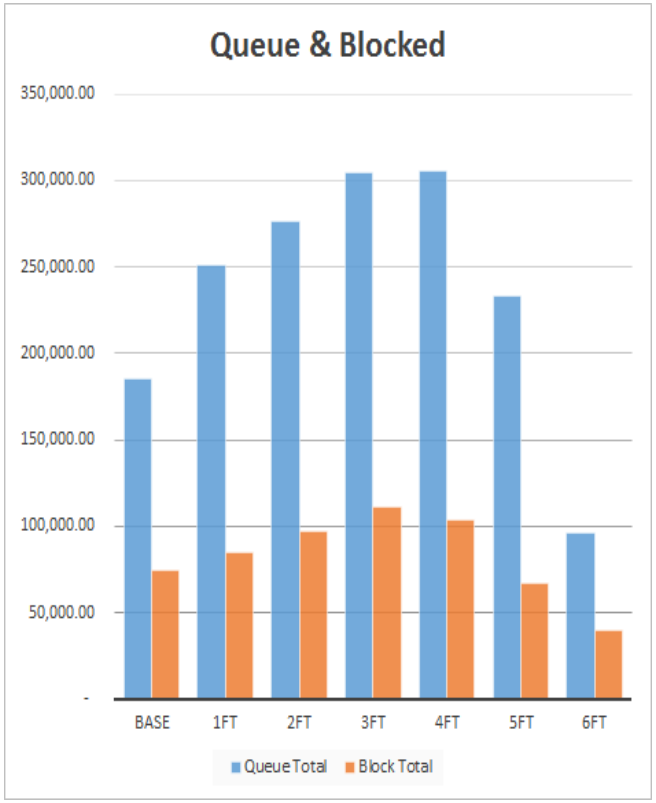


Figure 9: Network (Smaller Extracted Network) Congestion Hot Spots & Queued Vehicle Chart

Must-Have Features. These are features without which you would consider your project to be a failure.

- Transportation Assets inundated at 6 different inundation levels. This would be a map with charts on the side showing at each water level (1-6ft) where certain assets were located and in the chart showing the extent of the impact, likely showing both totals & percentages.
- Accessibility impact maps for walk, transit and auto.
- Network impacts charts – description of the congestion impact – if animated network visualization is problematic then a simplified version with more linked charts could be used.
- Some visualization of the number of trips lost by zone – if desire line are problematic then I may do it by simple circle shapes overlaid on the geography.

Optional Features. Those features which you consider would be nice to have, but not critical.

- Desire line chart
- Another visualization showing the changing ridership of different transit lines over the different inundation levels.

- My data has information on queues on network links for every 5 minutes totaling 90 minutes. Some sort of animation of these queues would be really interesting but may not be possible
- I also have information in "packets" of vehicles and the path the use on the network, the data comes in as a list of sequential nodes that packet used to complete its trip. Some sort of animation of these would also be wonderful, but I think it might be too challenging.

Project Schedule. Make sure that you plan your work so that you can avoid a big rush right before the final project deadline, and delegate different modules and responsibilities among your team members. Write this in terms of weekly deadlines.

April 5th – April 11th	• Data wrangling • Working prototype of accessibility maps • Early prototype of inundated asset maps w/chart
April 12th – April 18th	• Strategy and prototype of network impacts: ○ Figure out how to display the network while maintaining speed of visualization, highlight areas of congestion at different water levels. • Early prototype of Desire Line Lost Trips visualization
April 19th – April 25th	• Decide on Final List of Designs / if some of the optional features are not working at his point – they may have to be abandoned. • Finished Accessibility / Inundated Assets Visualization
April 26th – May 2nd	• Finalizing Designs / Polishing
May 3rd – May 5th	• Polish things up / Screen cast

Project Milestone Update

Project Website: http://mdgis.github.io/

Overview and Motivation:

My thesis research deals with modeling the transportation network impacts of inundation events on the Boston metro region. Inundation events (flooding events where networks are degraded or disabled) are serious threats to coastal cities around the world. Such events can have major impacts on regional transportation networks, as seen in NYC in the aftermath of Hurricane Sandy. I have sought to develop an understanding of potential inundation consequences on the Boston metro region by conducting two related analysis.

1. A GIS analysis- cataloging and quantifying likely impacts to transportation assets and driver of transportation demand (people & jobs)
2. A transportation modeling analysis using the Cube Voyager modeling platform to model the impact of inundation events on the regional multi-modal transportation system from 1ft to 6 ft.
3. This research has provided me with large amounts of data that I would like to display in an interactive web based format.

Related Work: Anything that inspired you, such as a paper, a web site, visualizations we discussed in class, etc.:

Below are some examples of various mapping visualizations that I used. These examples provide both inspiration, as well as technical hints for my implementation. The examples below are currently all some version of a map. I expect that I will have more "inspiration" sources by the time the final project is done. For now though, I have been focusing on mapping because it is not something we covered in the course, and it is something that is very important to my project. Furthermore, I have a great deal of experience with GIS and mapping, and I really wanted to use this project to develop web mapping skills.

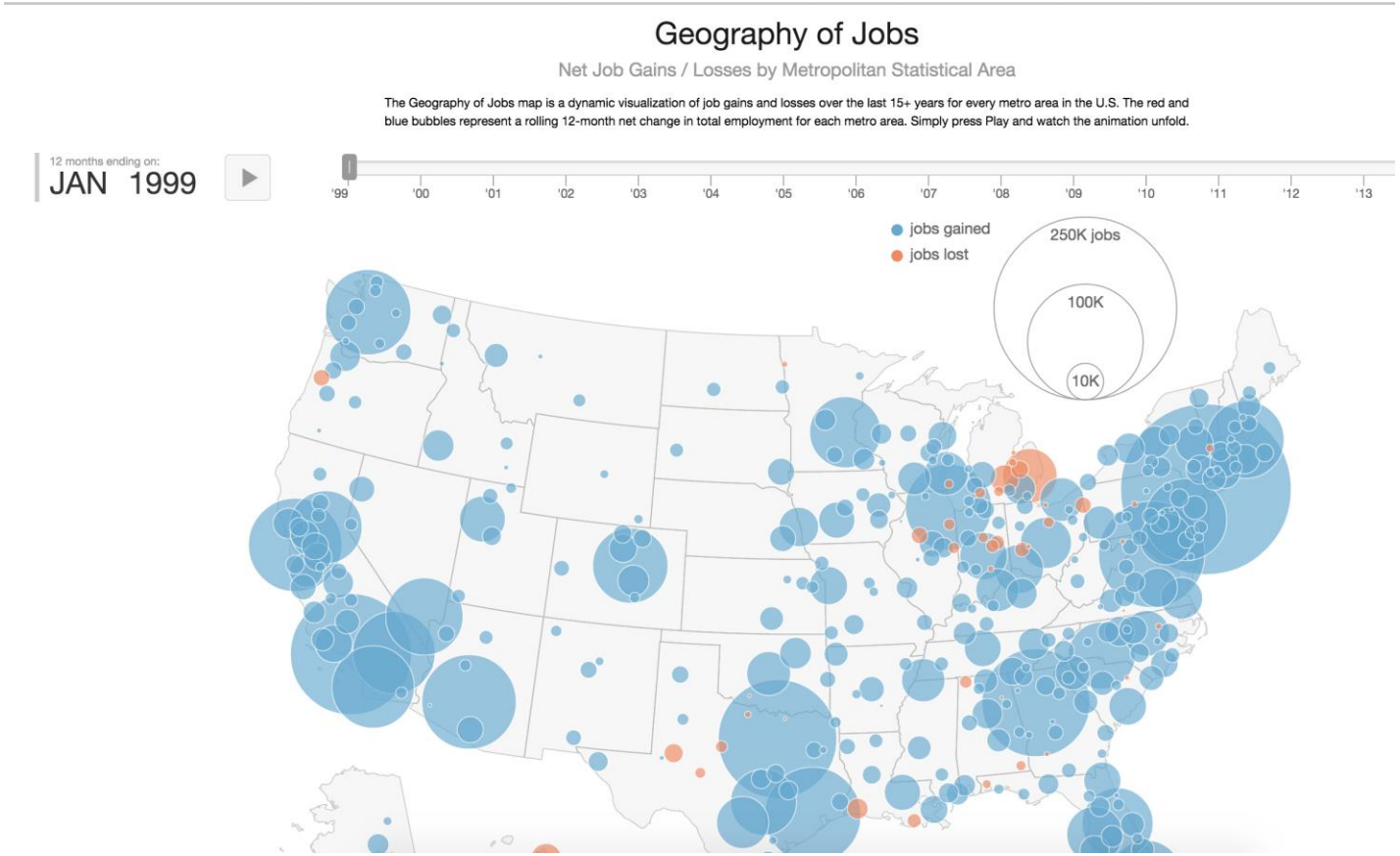


Figure 10: **http://tipstrategies.com/interactive/2014_map/**

This visualization is pure D3 (as in no leaflet, etc). I found this helpful both in the application of the slider, symbol levels, and legend. I do not have time series data but I am presenting 6 different sets of data at times (1-6ft of inundation). This has a very clean and clear presentation and message, something I need to strive for in my project.

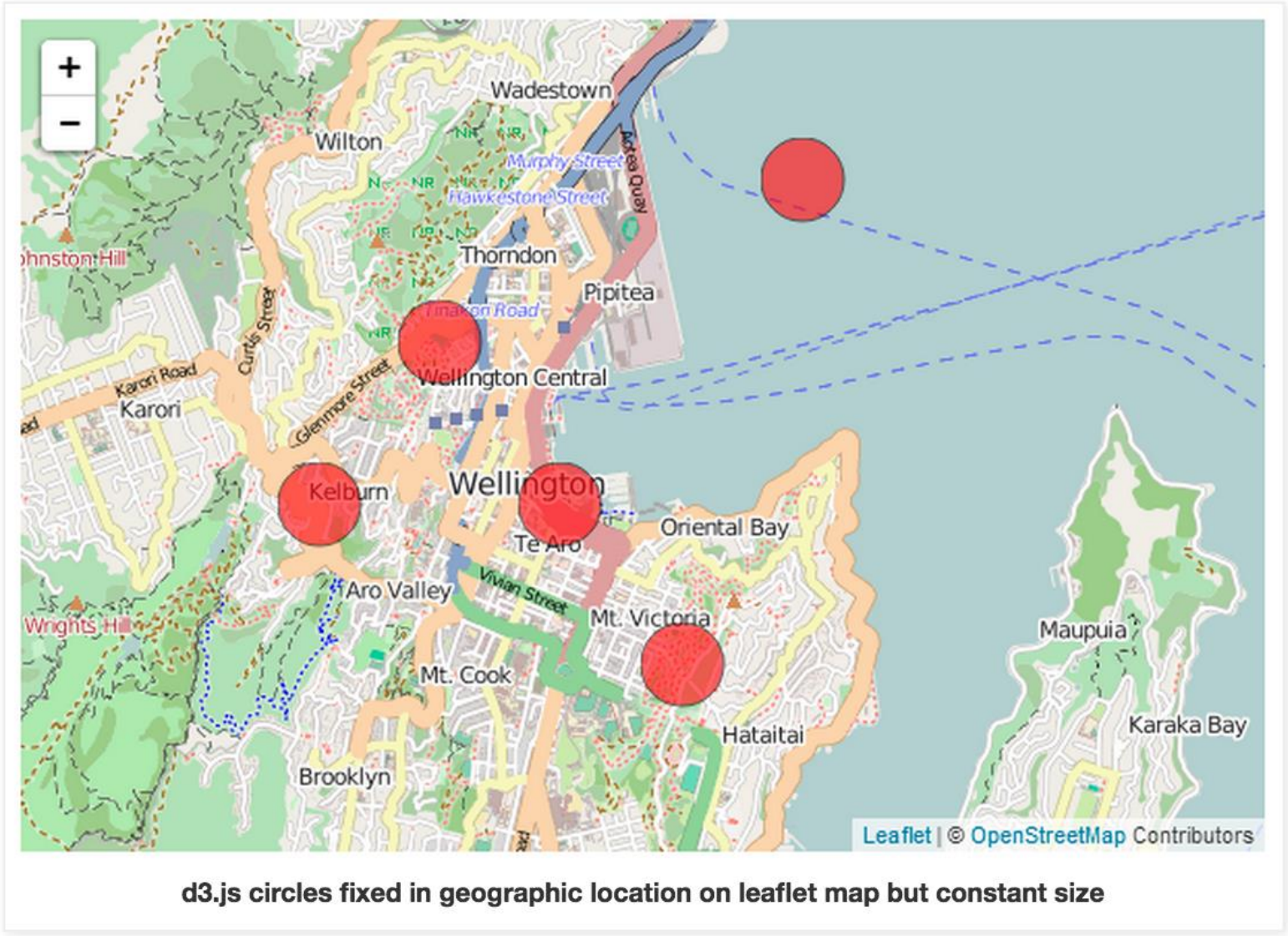


Figure 11: http://www.d3noob.org/2014/03/leaflet-map-with-d3js-elements-that-are.html

Some Leaflet / D3 examples that were helpful.

Here is what the finished product looks like:

Figure 12: http://zevross.com/blog/2014/09/30/use-the-amazing-d3-library-to-animate-a-path-on-a-leaflet-map/

Interesting piece on animating D3 on top of Leaflet.

Figure 13: http://www.tnoda.com/flightanimation

Another interesting piece on animating D3, though not on leaflet.

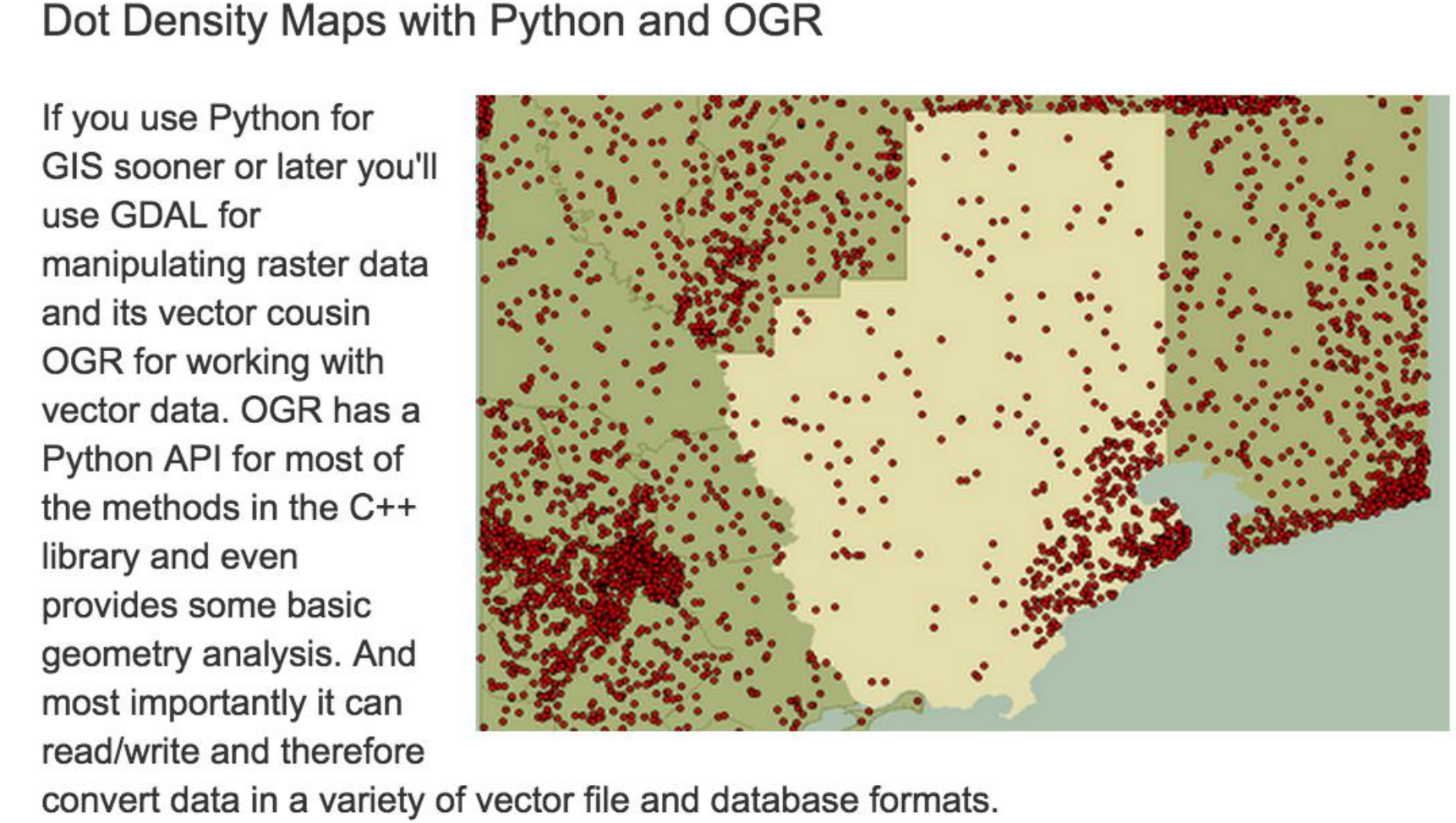


Figure 14: http://geospatialpython.com/2010/12/dot-density-maps-with-python-and-ogr.html

I wanted to do a dot density map for the inundated population, jobs and households visualization, but I think it will be too many points for D3 to handle.

Voronoi Arc Map

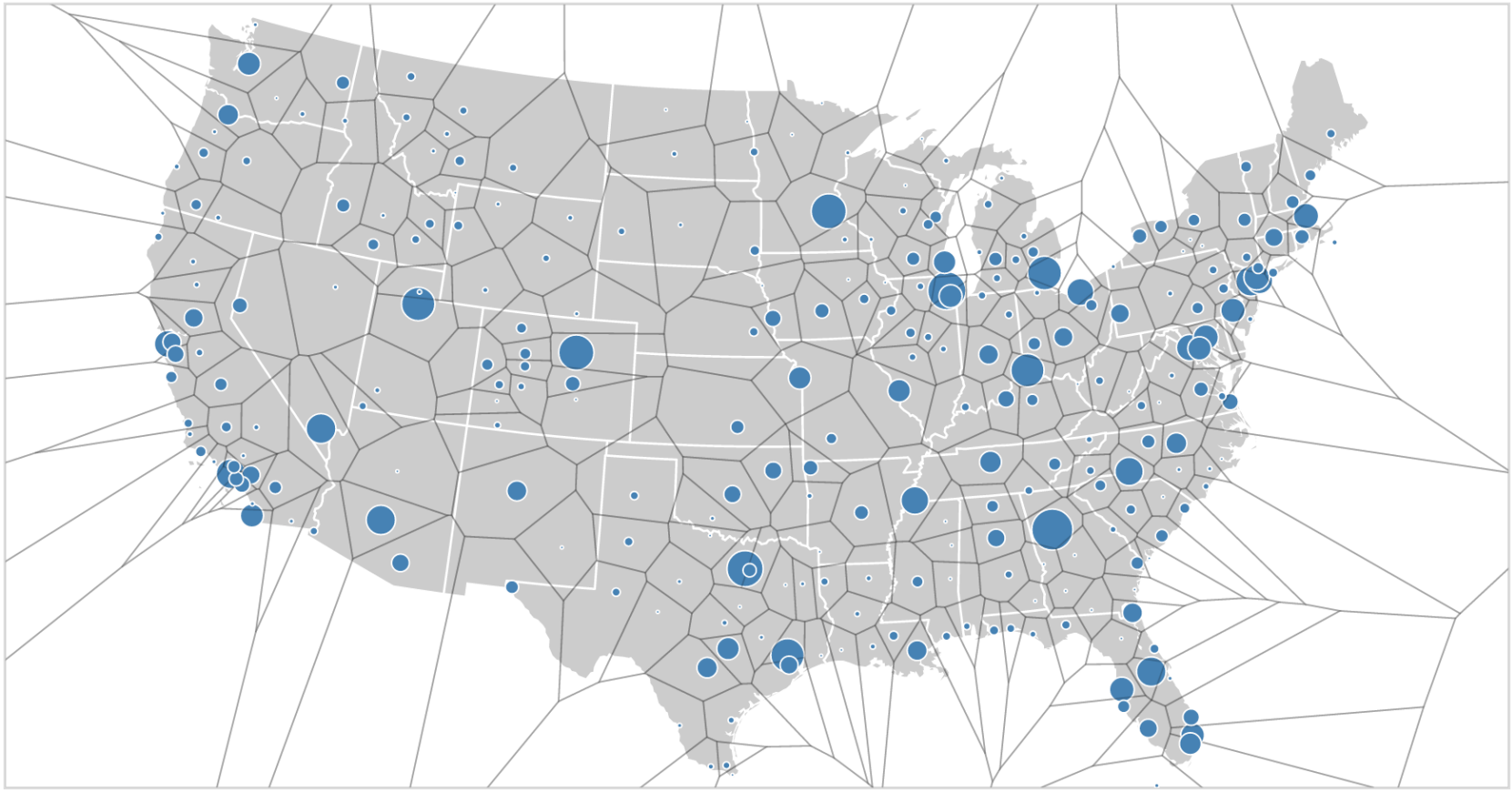


Figure 15: http://bl.ocks.org/mbostock/7608400

This was very helpful for setting up the Desire Line / Spider Line chart. I would like to implement to voroni polygons but just use them for selecting specific nodes. Initial tests did not work, it might be because of Leaflet overlay. I'm still thinking about this.

Questions: What questions are you trying to answer? How did these questions evolve over the course of the project? What new questions did you consider in the course of your analysis?

The primary questions I would want a user to be able to answer after interacting with the visualization would be:

- The location, count, and extent of impact given inundation at different levels

- The shifts in transit ridership caused by degradation: For example of the Red Line is disabled what are some of the buses that experience increased ridership - i.e. people who are likely changing their route
- How accessibility changes for different modes given different inundation levels.
- Congestion Hot spots given inundation
- Be able to compare the results to a non inundated scenario

Data: Source, scraping method, cleanup, etc.

Source:

The data is taken from my thesis work, primary source is myself and for general geographic files MassGIS.

Tables:

I used python PANDAS to export simple tables to JSON (Pandas.to_JSON(data, orient = "records")
For small tables I have also used D3's built in csv & tsv functions

Geographic Layers

Geographic data has proved to be much more complicated. I am using a combination of methods for dealing with this data.

In general all geographic data is taken from Shapefiles I already possessed. I have exported these to GeoJSON via Ogr2Ogr &/or Mapshaper [http://www.mapshaper.org/]

For larger layers I have then exported the GeoJSON into TopoJSON via the TopoJSON command line interface.

Leaflet:

I am using Leaflet to load some simple reference GeoJSON. Generally I am using Leaflet for layers that do not require interactivity on the layer itself. For example, that the map layer will not actually be clicked or interacted with by the user, the user may interact with other features that will cause changes to the layer.

Leaflet Overlay:

For interactive mapping layers I am using the Leaflet Overlay, with SVG placed in the overlay. Leaflet has faster zoom updates and is smoother than pure D3 with large geographic files. Furthermore, I have added a Mapbox base layer that provide context to users while being quite fast to update.

Exploratory Data Analysis: What visualizations did you use to initially look at your data? What insights did you gain? How did these insights inform your design?

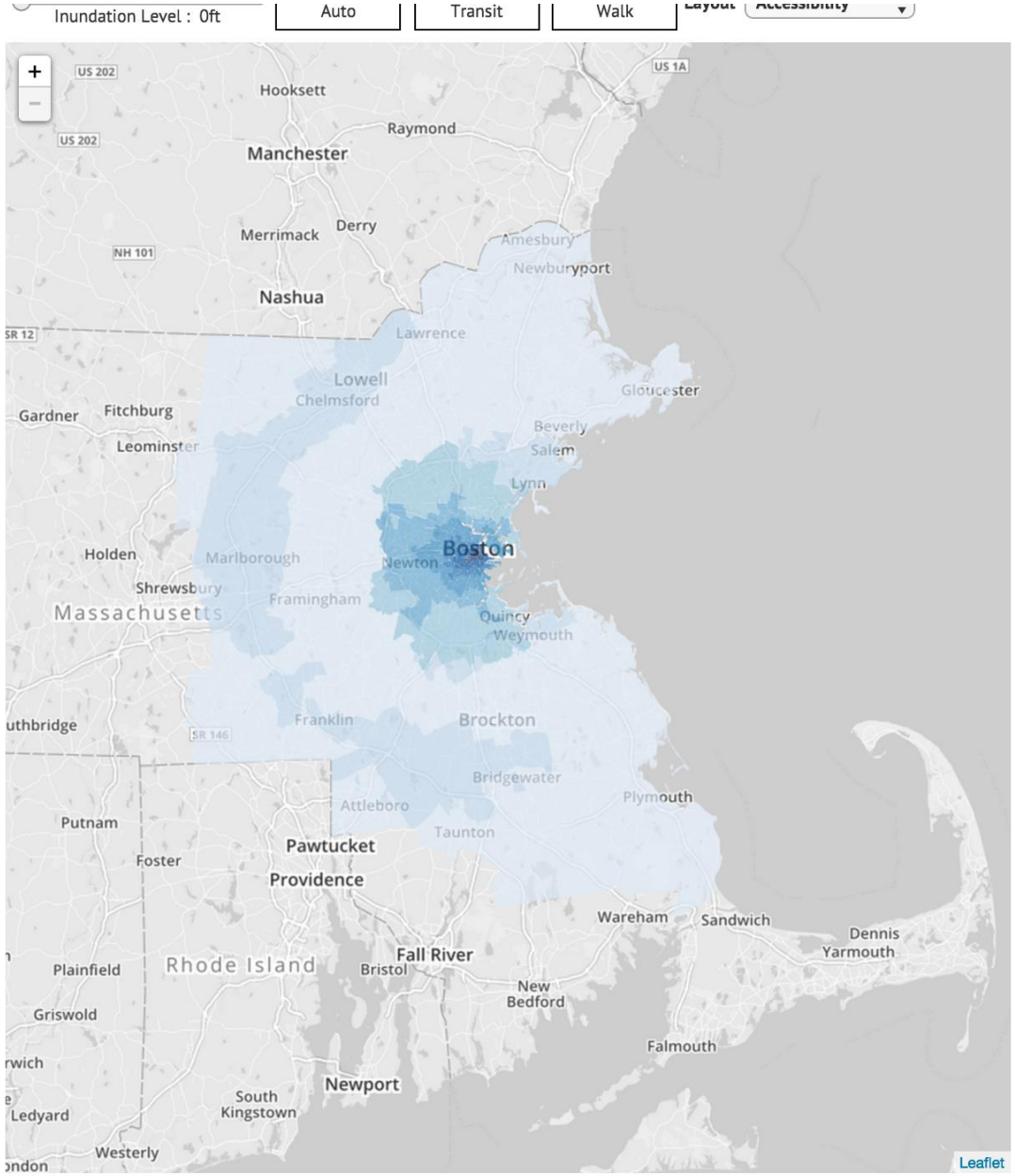
I have been working with this data intensely in excel, R, Python, ArcMAP, TransCAD and Cube Voyager. I have done a great deal of exploratory data analysis. Please reference my thesis defense presentation for more examples of how I have explored this data and presented findings.

Design Evolution: What are the different visualizations you considered? Justify the design decisions you made using the perceptual and design principles you learned in the course.

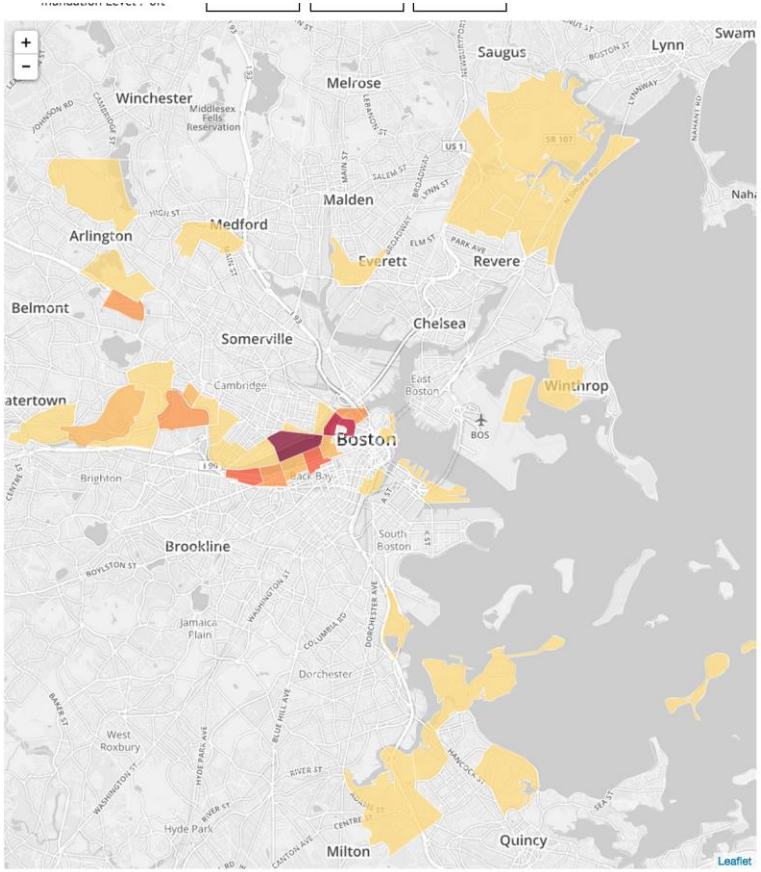
My project topic and data is very complicated so I need to develop a way of simply communicating the main concept, points and findings to a lay audience. For the past two weeks, I have focused on developing prototypes. By developing the prototypes first, I now know what visualizations I have time to develop more, and make work better. I did not want to commit to a visualization that would end up being too difficult to implement in the time available. Now that I know what basic map visualizations I can work with, I need to start thinking more about how to make the whole thing more understandable.

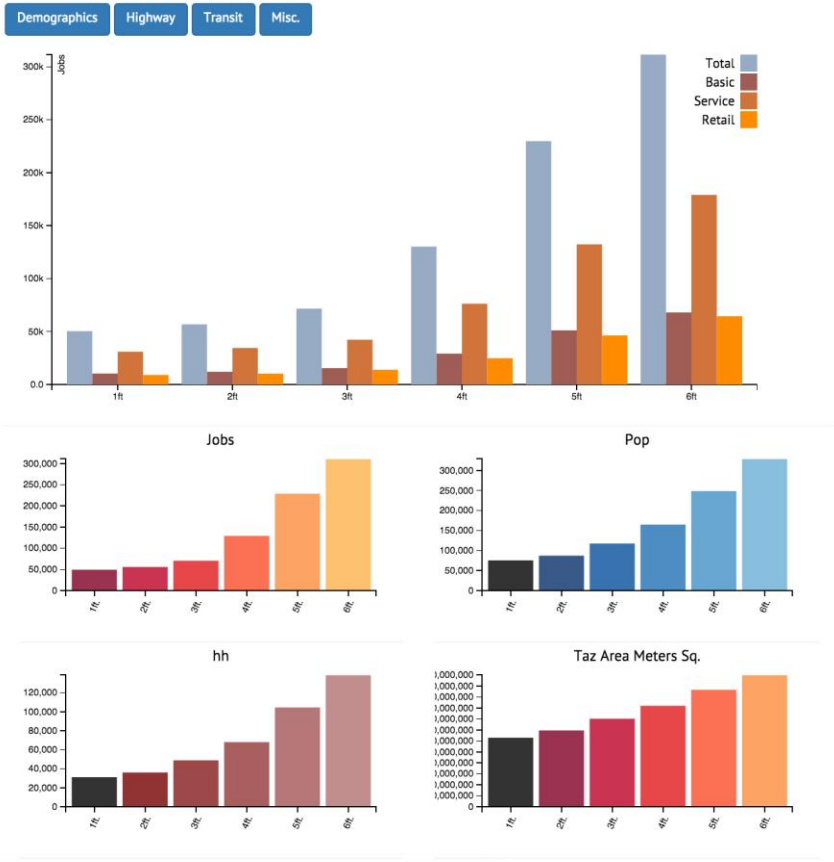
I think implementing some sort of "storytelling" feature would be helpful. Maybe, showing a specific area of the visualization and providing text commentary (sort of like a news story), then allowing the user to enter the full visualization and explore. I do not know how to implement this, as it seems to more of a web design challenge. Other alternatives would be just to have more textual description for each visualization.

Implementation: Describe the intent and functionality of the interactive visualizations you implemented. Provide clear and well-referenced images showing the key design and interaction elements.

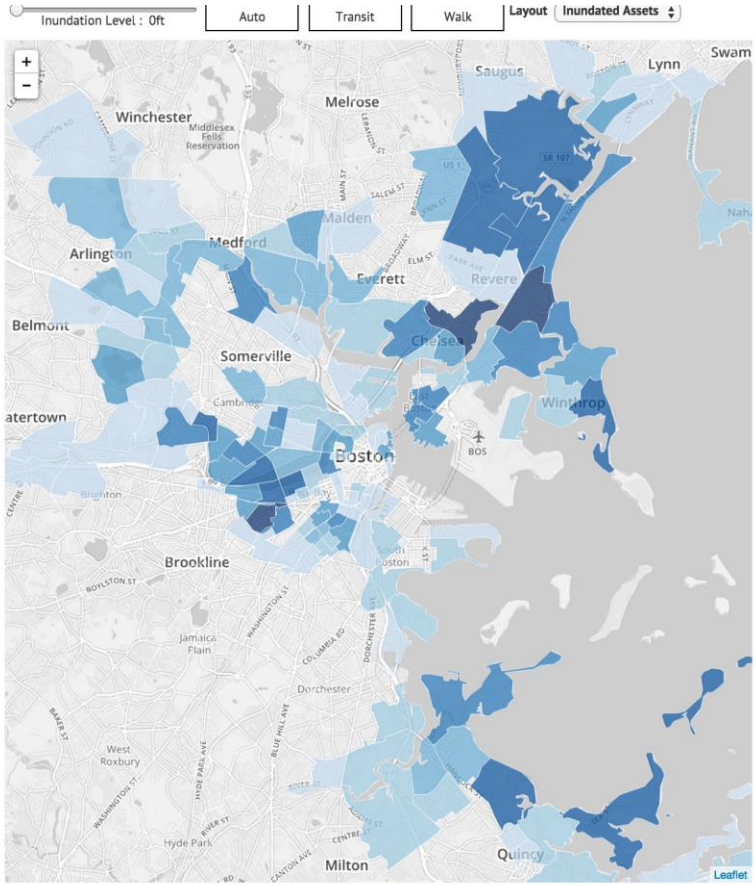


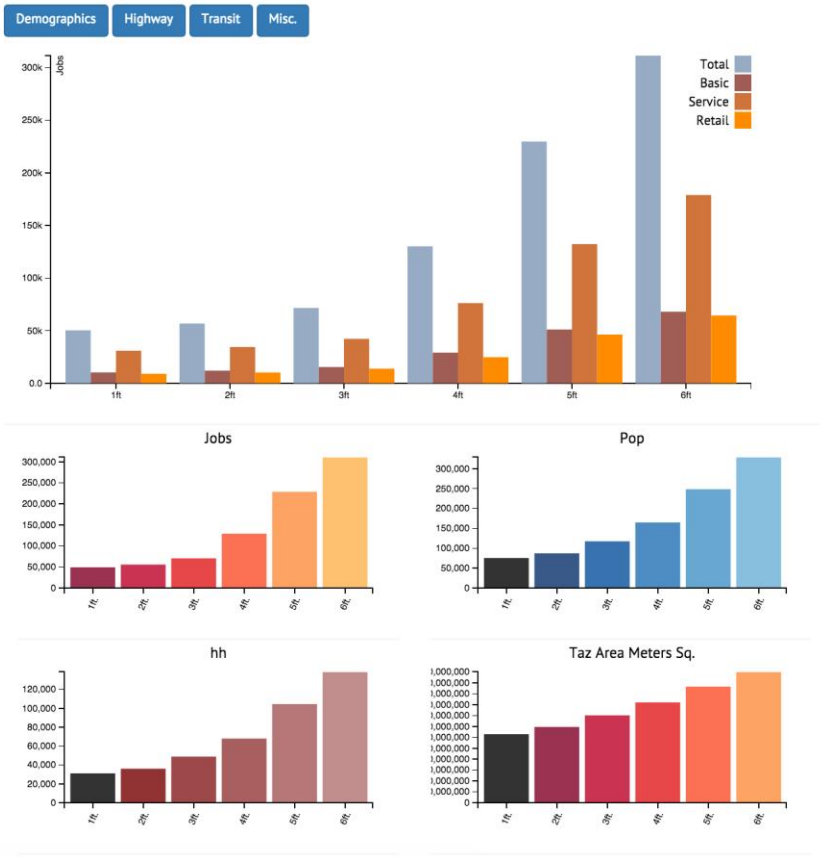
This is the accessibility chart. I have linked it to the slider and the buttons above. I need to add some description of how accessibility is calculated and how to interpret it. I also need to add legends.



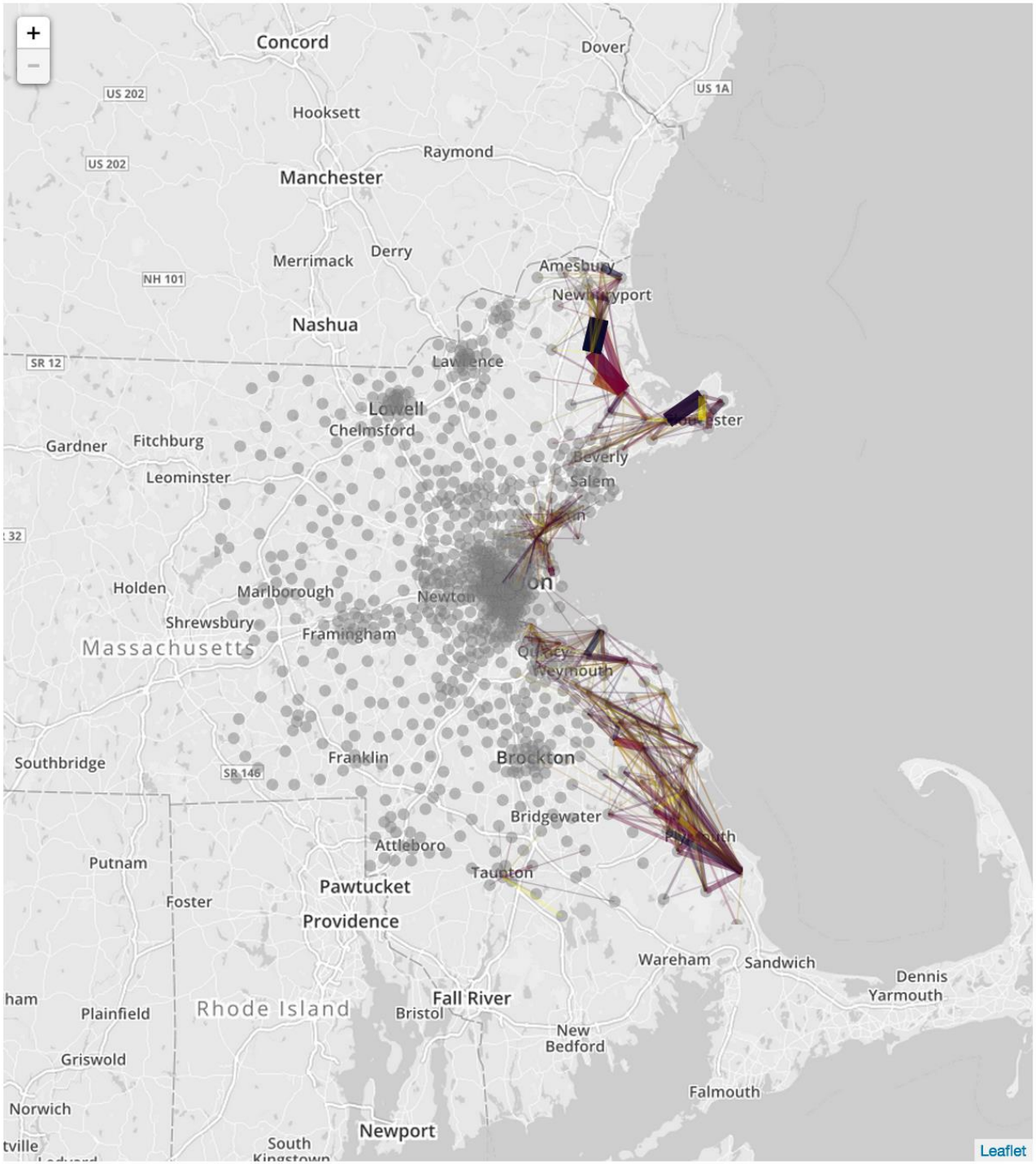


This is the current working Inundated Asset map, it shows the total number of inundated assets, if you click on the lower charts on the right the map on the left will update. I wanted to use dot density here but realized it would be too many dots for D3 to handle and I would need to pull in python. Soon I will have the buttons on the top left working and other data will be displayed in the map. For transit stations, it will be point level data, so not always chloropleth maps.



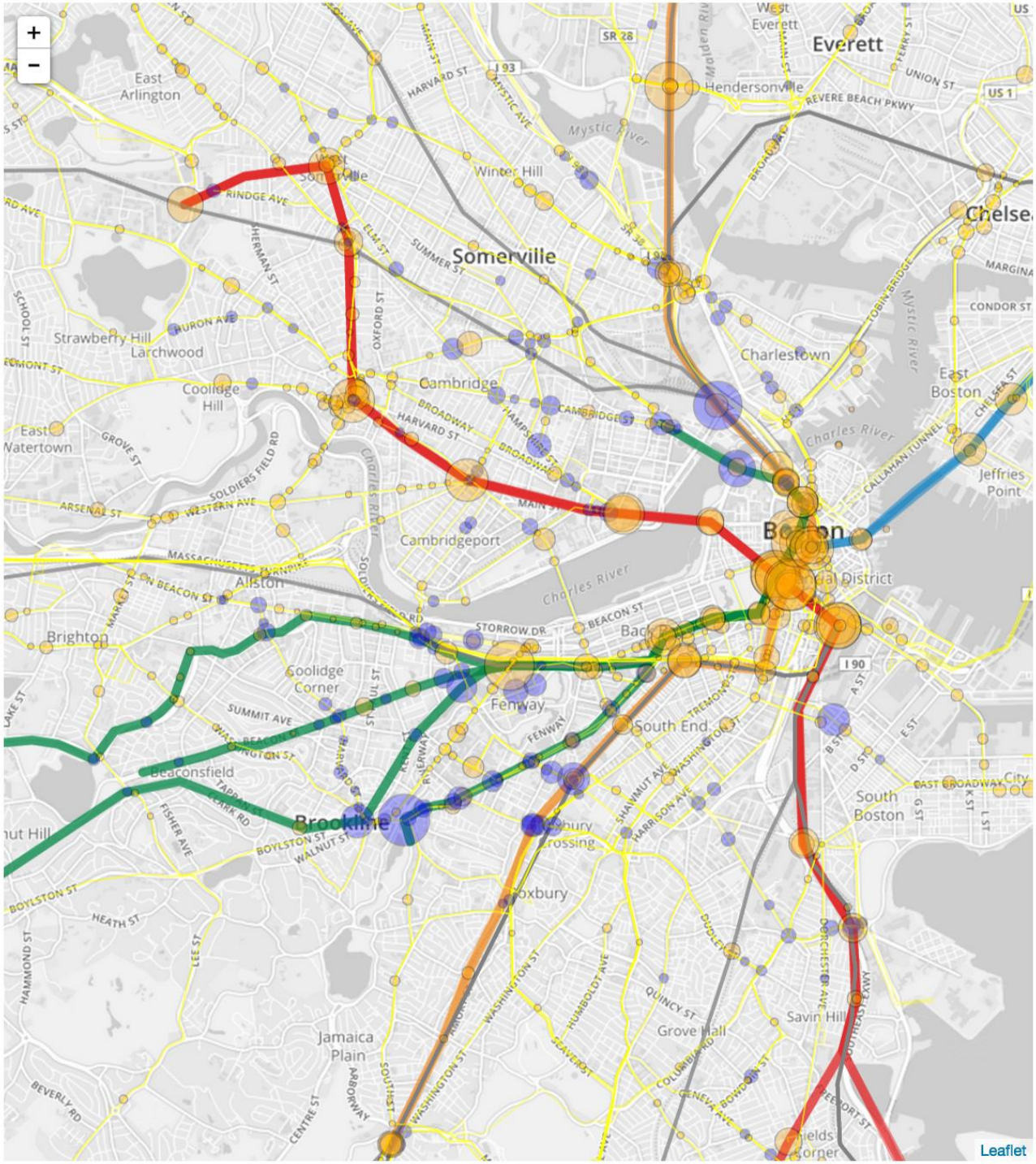


This shows an example of when you click on the population graph it changes to population.



The lines indicate areas where trips area lost. This is a traditional way to show the relationship between origin and destination flows.

I'm still figuring out how to make this more readable and interactive.



This map shows transit ridership change at stops, Orange circles are decreases in boardings at these stops and blue circles indicate increases in boardings at these stops. I will have this data linked to a display on the right of the map.

Evaluation: What did you learn about the data by using your visualizations? How did you answer your questions? How well does your visualization work, and how could you further improve it?

I'm not really ready to answer this question yet.