Divvy Bike Program

Christopher Cebra
GIScience Practicum

Research Question

- Has the Divvy bike program proved to be a successful program to connect different sections of the city of Chicago together and provide opportunities for lower-income residents, and can spatial analysis and visualization help in determining this impact?

Data

- Divvy bike usage by year (ridership)
  - https://divvy-tripdata.s3.amazonaws.com/index.html
  - Divided by year (first few years), quarter (until 2020) and month after
  - Extremely large dataset
- Network size by year
  - Contained in the usage by year data up to 4th quarter of 2017, have current data, need to manually generate when stations open if during 2018, 2019, 2020, 2021
  - Issues with low values
- Station statistics
  - https://data.cityofchicago.org/Transportation/Divvy-Bicycle-Stations-Historical/eq45-8inv
  - Contains information about station usage including percentage capacity, hourly
  - Extremely large dataset (around 50 times as large as the first one)

What we have so far

- Graphic of station access up to 2017
- News article review
  - Original purpose of Divvy program
  - Timeline for expansion
  - Struggles with expansion (for ex. E-bikes not as common in
  - Low-income communities because of risk of theft)
- R script for cleaning/extracting results for ridership

data (rides per station, when station comes up)

imgflip.com   08/09/2013
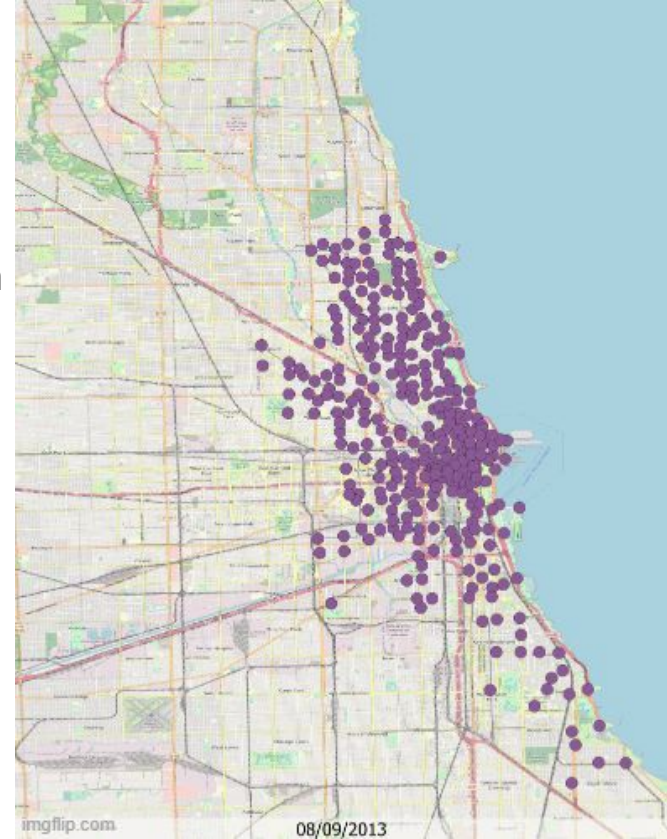
Goal

- A series of visualizations to categorize the Divvy network over time, and whether opening up new Divvy services in underrepresented areas of the city could lead to improvements in transport
- Animations including:
  - Divvy stations over time (complete up to 2017)
  - Heatmap of Divvy coverage
  - Heatmap of ridership by station and change over time
- Network size measures
  - Central point of all stations over time
  - Central point of rides over time
  - Choynowski map
- Interactive overlay
  - Demographic information
  - Preferred mode of transport information

To Do List

- Fifth week
  - Complete data generation, applying R script to all the months data CSVs
  - Finish stations and rides animation incorporating more recent data
- Sixth week
  - Complete spatial analysis (centers, Choynowski maps, etc.)
  - Complete and collate data for background layer
- Seventh week
  - Start writing up report
  - Put together visualizations into a clean, orderly manner
  - Set up interactive overlays
- Eighth week
  - Prepare for presentation

Questions?

Potential Access Metrics

- Spatial Access Package from UChicago Spatial Data Center
  - https://github.com/GeoDaCenter/spatial_access
  - Focuses on travel times and access score (focus on this one because it has bike travel times listed)
- Divvy trip data
  - https://divvy-tripdata.s3.amazonaws.com/index.html
  - Potentially use Kepler (used in GIS III) for data visualization
- PySAL access package from UChicago Spatial Data Center
  - https://access.readthedocs.io/en/latest/
  - Indicates how close locations are to important supply locations like stores, healthcare centers, etc.

Potential Research Questions

- What locations across the Chicago city limits could best benefit from improved access to services? Would constructing new forms of transit be the best benefit, or would constructing new service hubs (like grocery stores) be better?
- Has improving access to infrastructure (example: Divvy bikes) promoted an improvement in mobility and access to services across Chicago?

Bike Counts Data

https://chicagocompletestreets.org/streets/bikeways/bike-counts/

https://ride.divvybikes.com/system-data

https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0137922

https://matias-pietruszka.medium.com/an-exploratory-data-analysis-of-divvy-bike-rides-in-chicago-33a2c2f58fcb

- Check what locations have similar bike access times to their drive access times
- FFMPEG

https://data.cityofchicago.org/Transportation/Divvy-Bicycle-Stations-Historical/eq45-8inv

Challenges

- Coding component
  - New packages, more Python experience is critical
- Data temporality
  - If we are aiming to make a look at whether infrastructure modifications improve access, we need a temporal component
  - Divvy data is temporal (could be averaged over months) but the other access metrics could just promote different
- Need to clarify/cut down on amount of research. Perhaps only focus on the South Side if scope is too large
- Attempting to "streamline" multiple metrics by creating a new metric can lead to metric inflation