Initial Proposal

Enoch Shin, Felix Stetsenko

10/16/2019

1. **Group:** A
2. **Group Members:** Enoch Shin and Felix Stetsenko
3. **Title:** *Visualizing Transportation Network Provider and Taxi trips in Chicago*
4. **Purpose:** Our initial idea is to visualize travel patterns across the City of Chicago using Transportation Network Provider (TNP) and Taxi trip data publicly available on Chicagoâ€™s Data Portal website. We would visualize the origins and destinations (on the census tract level) of all taxi and TNP trips across the city from November 2018 to present. We could then use the data visualization for a number of possible analyses:
   - Do TNP and taxi services serve the same market, or is one service preferred over the other for certain travel markets (e.g. trips originating from Oâ€™Hare Airport)?
   - Can residential segregation be visualized through TNP and taxi service and trip patterns? For instance, is there a relationship between census tract demographics and TNP vs. taxi share of trips originating from that census tract?
   - Using a clustering algorithm on the origin / destination data, how closely would the clusters reflect neighborhood boundaries and racial and socio-economic divides?

5. Data Overview

Data: The relevant data is all publicly available, either from the City of Chicago or the U.S. Census Bureau. We would create a visualization by using an overlay on top of the Google Maps API. We would consider adding the functionality to have the data continuously update as the City of Chicago posts updates on its Data Portal.

6. Variables

Transportation Network Provider Data

City of Chicago TNP data (available from November 2018 to present).

Trip ID

A unique identifier for the trip

Trip Start Timestamp

When the trip started, rounded to the nearest 15 minutes (floating timestamp).

Trip End Timestamp

When the trip ended, rounded to the nearest 15 minutes (floating timestamp).

Trip Seconds

Time of the trip in seconds.

Trip Miles

Distance of the trip in miles.

Pickup Census Tract

The Census Tract where the trip began.

Dropoff Census Tract

The Census Tract where the trip ended.

Pickup Community Area

The Community Area where the trip began.

Dropoff Community Area

The Community Area where the trip ended.

Fare

The fare for the trip, rounded to the nearest $2.50.

Tip

The tip for the trip, rounded to the nearest $1.00.

Additional Charges

The taxes, fees, and any other charges for the trip (dollars).

Trip Total

Total cost of the trip (dollars).

Shared Trip Authorized

Whether the customer agreed to a shared trip with another passenger.

Trips Pooled

If customers were matched for a shared trip, how many trips, including this one, were pooled.

Pickup Centroid Latitude

The latitude of the center of the pickup census tract.

Pickup Centroid Longitude

The longitude of the center of the pickup census tract.

Pickup Centroid Location

The location of the center of the pickup census tract.

Dropoff Centroid Latitude

The latitude of the center of the dropoff census tract.

Dropoff Centroid Longitude

The longitude of the center of the dropoff census tract.

Dropoff Centroid Location

The location of the center of the dropoff census tract.

Taxi Data

City of Chicago Taxi data (available from 2013 to present)

Trip ID

A unique identifier for the trip.

Taxi ID

A unique identifier for the taxi.

Trip Start Timestamp

When the trip started, rounded to the nearest 15 minutes.

Trip End Timestamp

When the trip ended, rounded to the nearest 15 minutes.

Trip Seconds

Time of the trip in seconds.

Trip Miles

Distance of the trip in miles.

Pickup Census Tract

The Census Tract where the trip began.

Dropoff Census Tract

The Census Tract where the trip ended.

Pickup Community Area

The Community Area where the trip began.

Dropoff Community Area

The Community Area where the trip ended.

Fare

The fare for the trip.

Census Bureau Data

We'll grab some broader census information such as demographics, income, and information about how people commute to work (which is available on the census data). This data set will be pulled in to contextualize the two data sets above if we find that we want to add more subsets and information about the populations we're studying.

7. End Product

The final deliverable will be an interactive visualization overlaid over the Google Maps API; it will show all Taxi and TNP trips in the City of Chicago from November 2018 to present. The user will have the option of selecting specific time frames, origin / destination census tracts, and adding additional layers to the map (e.g. public transit lines, demographic and socio-economic variables).