Divvy Data Document

http://j.mp/DivvyData

ALSO see the Bike Sharing Data hackpad: https://bikesharingdata.hackpad.com

Divvy is a public bicycle sharing transit system that allows annual and daily members an unlimited number of trips (free if 30 minutes or shorter). This document is maintained by Steven Vance, transportation planner and Streetsblog Chicago reporter.

Data & APIs

Data available

- Real-time station availability
- Trip data. Release of this is up to CDOT. Start, end stations. Duration, start/end time and date, bike ID, rider type (annual or 1-day). Would be released each quarter. Capital Bikeshare has data available that would likely be identical to what Divvy will release: http://www.capitalbikeshare.com/trip-history-data. 67 visualizations based on similar trip data released by Boston's Hubway: http://hubwaydatachallenge.org/
- Chris Manning: Will you record the bikeID so we can understand how often a bike breaks down? (and stations)
- Rebalancing: Staff key + bikeID

App requests (from officials, public)

- Historical trip data. Store the Divvy JSON "API" every minute, for now, just for potential apps.
- Integration with public transit. Take a trip via transit + bike sharing. For NYC. Adam Gluck is working on an iOS app for this.
- Predict when stations will be empty (for rebalancing efficiency)
- Predict when stations will have bikes (so people know when they can go grab a bike from an empty or near-empty station)

Divvy API links

1. http://divvybikes.com/stations/json - real-time feed; updated once per minute
2. https://github.com/AdamGluck/DivvyAPI - API for iOS native apps
3. https://github.com/iandees/divvyapiapi - real-time station location API; updated once per minute; returns as JSON
4. https://github.com/iandees/divvy-dump - dumps your ride history by scraping your DivvyBikes.com account page
5. http://shrouded-beach-2183.herokuapp.com/stations/87 - example API call for a station ID (which is different from the "landmark" attribute)
6. http://veloplan.net/ - bike and walk routing for Montreal using OpenTripPlanner (cibi.me

and bikeplanner.org are broken right now)

7. https://npmjs.org/package/citibike - Unofficial Citibike API for node.js. Citibike and Divvy are both operated by Alta Bicycle Share and have some similarities.
8. http://citibikedata.com - historical storage of Citibike NYC
9. http://www.reddit.com/r/Citibike/comments/1i3esv/i_work_for_citibike_ama/cb0sugw - a conversation on Reddit with a Citibike employee about the API and JSON feed
10. DivvyDataAccess API - handles request to the Divvy API and returns values in an Objective-C usable form (for iOS development in Xcode) (Adam Gluck & Andrew X
    a. Asynchronous requests
    b. Delegation - to react when asynchronous requests are fulfilled
    c. Station objects

BGL consulting's iOS/Xcode API
Adam Gluck & Andrew Beinstein
Problem: four location points to be taken into consideration to plan a trip
Asynchronous requests - multi-threading so the UI doesn't have to wait for functions to complete
iOS app available on the iTunes Store, or read Steven's review.

Other information

Other links

1. https://www.google.com/fusiontables/DataSource?docid=1nU6otJCTTGFSpTmRMtJ6AgLSs0voW0rmoZWf4cA - All planned Divvy stations (some of this information is out of date because stations have been moved and the original data set from which this came has been removed from the Divvy website). This is now highly outdated.
2. http://bikes.oobrien.com/chicago/ - Oliver O'Brien's map. Real-time visualization; historical data timeline.
3. http://api.citybik.es - Powers some of O'Brien's map
4. http://www.slideshare.net/anthonymobile/bike-hack-night-townsend - NYC
5. Predicting bike availability by Data Science for Social Good fellows in Chicago

"Press"

- John Bracken - On 10 days or sharing bikes in Chicago
- Smart Chicago Collaborative - one of the weekly posts about Hack Night

Visualizations

- Buenos Aires, Argentina - by Manuel Aristarán

Apps you can get right now

iOS apps

1. CycleFinder - official app for Divvy
This app is created by the bike sharing system's hardware vendor, Public Bike System Co. (PBSC).

http://j.mp/CycleFinder4Android
http://j.mp/CycleFinder4iPhone

2. Chicago Bike Guide
Steven Vance's iOS app as a website

http://bikechi.com/chicago

Get it for iOS in the App Store

3. Chicago Bike Route
the iTunes Store, or read Steven's review.

There are actually six more available for iOS.

Discussions

Citibike NYC information
via Frank Hebbert/Hangout
They had a #citibike hack night on June 26th... some interesting projects:

Greg Estren's city bike stats -- http://sites.google.com/site/citibikestats/

Citibike and Pandas using Wakari -- online analysis toolbox with data scraped since day 3, by Paddy Mullen. https://www.wakari.io/sharing/bundle/paddy/citibike_data_intro and http://citibikedata.com

Citibike and Foursquare - redesign NYC around bike share, by Anthony Townsend: http://www.slideshare.net/anthonymobile/bike-hack-night-townsend

Cool analysis by Alastair Coote, showing travel time impact of Citi Bike
http://experimenting.alastair.is/citibike/

Plus, two ideas that I wanted to share (but couldn't, thanks to Hangouts/network probs) --

1. data can power amazing decision support tools. Not just maps/dock occupancy/etc, but rider-centric, the info-i-need-now. E.g. if there's an empty dock, should I wait, or walk? -This is

what Scott Kubly and others were talking about with prediction models.

2. having said that, forget rider tools! what do ridership patterns tell us about how cities function? As seen by bikeshare data, is Chicago like NYC or not? What new insights about safety/urban economics/etc can we get from the data? E.g., how do different locations cluster over time (someone was looking at that in NYC last week)

And... just wait until we get rider data, GPS and detailed logs for each bike/dock/redistribution runs and more.

Scott Kubly - CDOT deputy commissioner
Build a model to predict when stations will be full or empty.

> Request that users arriving at a full station request more time via the kiosk even if sufficient time remains to arrive at another station. This will help track arrivals at full stations, if a user does not do this, the system will not record the desired use/demand, instead it will record the alternative as the primary.

All data that's not tied to a specific user should be open.
Each bike is equipped with passive GPS to show actual route people take. "Are we putting bike routes in the right spot?" How does slope affect trip routing?

Dan Gohlke - Divvy data staffer
From DC
Cabitracker - what time do I have to get to a station to ensure I get a bike?
Talked about anonymizing customer, but for low-use stations, you can identify the user.
Unless you expend funds to count bikes, you can't get any ridership/routing data. Strava.

JPV: How do you know where to rebalance? DG: NYC is blazing the way. We're starting to figure out the usage patterns. DC is very simple to understand: in/out. We don't have enough data yet in Chicago.

Data Portal

It would be useful to have the following data available on the City of Chicago Data Portal (list your ideas):

- monthly aggregate trip data (including GPS tracks)
- daily membership totals (and distinguishing daily and annual members)

Data Dictionary

http://divvybikes.com/stations/json/
Most of these are obvious, but here are the ones that aren't:

- id - unique station ID. Use this one for all intents and purposes.
- landmark - the ID the station has in the planned-stations.json file. They are different than "id". You should probably avoid using this value.

- testStation - I don't know. This probably indicates whether or not the station is in some kind of testing or temporary online mode and not for public use. I've never seen a value that's not "false"
- lastCommunicationTime - this value would probably represent a timestamp when the station last communicated with the server. I've never seen a value that's not "null"