

Interactive Urban Bike Sharing Visualization

Sam Longenbach & Minkun Liu

1 Introduction

Bike sharing systems (BSS) have been becoming a more popular alternative mode of transportation in urban areas. In general, users are able to checkout a bike from a docking station, ride the bike for a given amount of time, and return it to another docking station. While sharing platforms may differ in cost or max ride time, all docking BSS must always achieve two criteria to meet user demand. First there must always be bikes at all stations and secondly docking stations must not be completely full or a user can not return his or her bike rental.

There has been a decent amount operations research carried out in the area of tackling the problem of bicycle re-positioning problem [2]. The issue of overnight bicycle re-positioning is often viewed as static, while intraday bicycle re-positioning is dynamic as users real-time demand must be accounted for when attempting to balance the system [2]. Unsurprisingly the dynamic re-positioning problem is far more challenging than the static re-positioning. In reality in high volume BSS, the re-positioning decisions are made by humans and is not completely automated.

Existing research in the visualization domain looks to implement a dashboard to highlight bike activity patterns at various stations depending on time of day or types of attraction in close proximity [5]. Building off prior research we also look to apply clustering algorithms to highlight stations that behave similarly in regards their demand over the time of day [4]. For example, BSS stations tend to experience high demand in residential areas during morning during week day hours. We view creating a digestible visualization of station demand as a starting block for a tool that would better inform human decisions in the problem of dynamic bike re-positioning.

2 Key Takeaway

Due to the scale of the demand in large bike sharing systems, human are often involved in the problem of bicycle re-positioning. To better inform these decisions we look to create a visualization to allow for better understanding of temporal and spatial features that contribute to particular station demand.

3 Project Type

Interactive D3 Visualization

4 Audience

While users and BSS managers could both extract interesting information from our visualization, the intended audience would be managers of BSS. We would like to build off of prior works to create an interactive visualization supplemented with the implementation of clustering algorithms [1].

Previous research has shown unsurprisingly that during the weekday in the mornings bicycles are ridden from residential areas to commercial areas and vice versa in the afternoon [1]. Understanding the natural flows and patterns of the system can lead to BSS managers to make more calculated decisions when it comes to on the fly bicycle re-positioning.

5 Approach

In particular we will be working with data from [Divvy bike share](#) system located in Chicago. Each row of data contains information on a particular trip. In our case, Divvy provides the start and end time as well as the start and end station. Lastly, if the person is a Divvy member we have additional

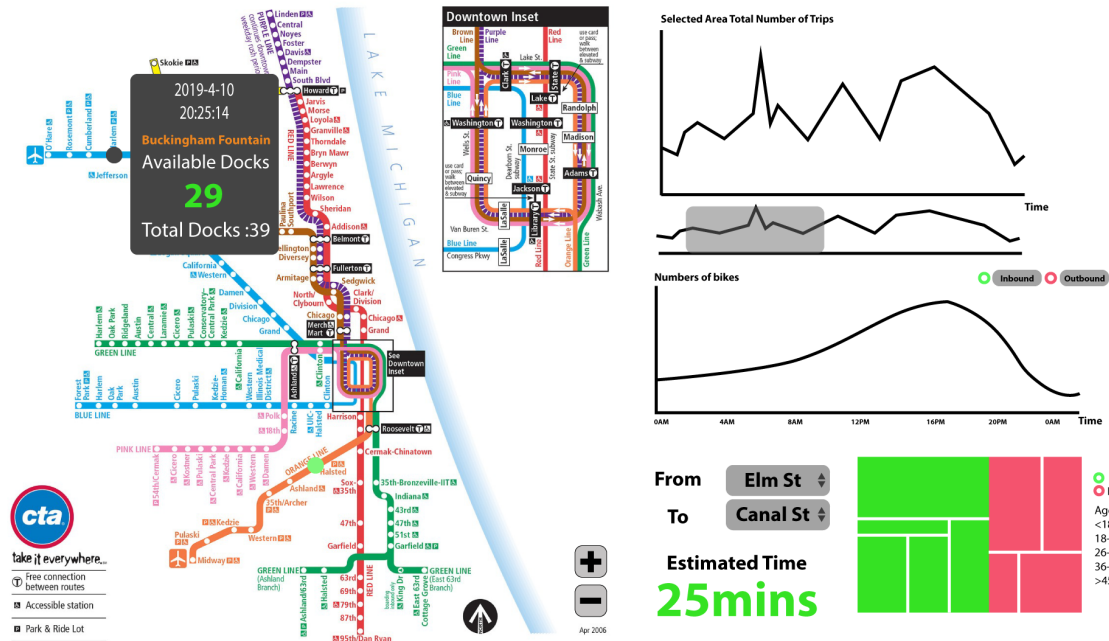


Figure 1: Rough idea of possible interface. *Note that the map is displaying Subway stations in Chicago.*

information of the gender and age of the rider.

Following in the steps of other research, we plan to preprocess the data in python and calculate the number of outgoing and in going trips from each station in the database. Using a clustering algorithm such as DB-Scan or K-means we can first perform some exploratory analysis on the different stations in the data set. The visual representation has potential to provide useful information on the characteristics of stations.

Again in align with research we would like to provide some multiple views to provide more information of the users and their trips. To accomplish this we will have to add interactive filter and selection features. However, the extent of these features will be limited by the timeline.

6 Best-case Impact Statement

Provide an informative visualization to inform BSS managers of any temporal or spatial patterns present in their bike station network. We strive to build a graphic so that potential patterns are easily recognizable.

7 Major Milestones

- Encode the color of stations based on clustering results.
- Root the visual encoding of our data in the principles that we have learned in class.
- Create interactive map in D3 to show real-time available bikes at station.
- As time permits add additional views. For example, estimate the expected travel between stations.

8 Obstacles

- Clustering being unsupervised by nature, it may be difficult to distinctly classify each station. One solution would be to naively classify each station on the spectrum from totally outflowing to totally inflowing of bikes at a given point of time.
- Due to limited experience in creating moveable and zoomable maps, more time than initially anticipated may be spent on this portion of the visualization.
- While the number of stations in is dataset is approximately 500, the number of trips is around 10

million. We may focus on the most recent years or sample so that are visualization can render in the browser.

- In order to implement the live status of available bikes we will need to process Python code in on an additional server.

9 Resources Needed

- Python/Pandas to preprocess and aggregate trip data.
- R or Python to run clustering algorithms to classify stations.
- Javascript code to build an interactive street map with the location of docking stations.
- Python code to process [live JSON feed](#) of available bikes in the Divvy system.
- Javascript code to click on a particular station and update views displaying station specific information.
- Continue literature review to inspire our design as we begin to advance with our current design [3].

10 Related Publications

- Bargar explored the utility of building an interactive web-based visual analytics application for comparing usage patterns between bike sharing programs. Utilized the St-DBSCAN algorithm to cluster trips as a way of categorizing flow patterns [1].
- Espergren provided an exhaustive literature survey comparing existing models as well as provided a new mathematical formulation for the static bike re-positing problem [2].
- Oliveira created pixel-oriented visualization designs that help understand the dynamics of station states and trip circulation patterns [3].
- Vogel used data mining to gain insight into the complex bike activity patterns at docking stations [4].
- Yan designed a visual analytics system to interactively explore spatial, temporal, and user dimensions patterns and compared these trends among cities [5].

11 Define Success

Due to the timeline of this project our main goal is learning. While it may sound cliché, improving our understanding of creating a smooth workable D3 visualization is our main goal. Also figuring out how to process [live JSON data](#) is something I would like to learn. Aesthetics and interpretability are very important to us while the processing efficiency is something that will addressed as time permits.

References

- [1] A. Bargar, A. Gupta, S. Gupta, and D. Ma, “Interactive visual analytics for multi-city bikeshare data analysis,” in *The 3rd International Workshop on Urban Computing (UrbComp 2014)*, New York, USA, vol. 45, 2014.
- [2] H. M. Espegren, J. Kristianslund, H. Andersson, and K. Fagerholt, “The static bicycle repositioning problem-literature survey and new formulation,” in *International Conference on Computational Logistics*. Springer, 2016, pp. 337–351.
- [3] G. N. Oliveira, J. L. Sotomayor, R. P. Torchelsen, C. T. Silva, and J. L. Comba, “Visual analysis of bike-sharing systems,” *Computers & Graphics*, vol. 60, pp. 119–129, 2016.
- [4] P. Vogel, T. Greiser, and D. C. Mattfeld, “Understanding bike-sharing systems using data mining: Exploring activity patterns,” *Procedia-Social and Behavioral Sciences*, vol. 20, pp. 514–523, 2011.
- [5] Y. Yan, Y. Tao, J. Xu, S. Ren, and H. Lin, “Visual analytics of bike-sharing data based on tensor factorization,” *Journal of Visualization*, vol. 21, no. 3, pp. 495–509, 2018.