

Final Project Proposal: Ibadat Jarg & Helen Wang

Project Premise and Aims

Capital Bikeshare is a bicycle-sharing system currently operating across Washington D.C., Maryland, and Virginia. Launched in August 2008, Capital Bikeshare has risen to serve over 7 different jurisdictions, containing over 700+ stations and 5000 bicycles. The current system covers over 180+ square miles and is rapidly expanding.

Our project aims to understand the key impacts that opening a Capital Bikeshare location has on areas in the DMV with a particular focus on:

- Traffic Congestion
- Mobility
- Emissions

We plan on building inferential models to identify the key impacts that Capital Bikeshare can have on the greater DMV metropolitan area. We then aim to utilize these metrics to identify areas that may benefit the most from Capital Bikeshare expansion using predictive modeling. We hope to then compare our findings to the current expansion plans reported by Capital Bikeshare.

Data Sources

- Location data of Bikeshare stations (XY coordinates)
- Congestion data prior to and after new station from TomTom traffic API
- Emissions data from EPA
- Mobility data from Travel Time API

Data Plan

We will use a variety of data sources and multiple APIs. We will use the Capital Bikeshare's open source data that provides precise coordinates for each of the bikeshare locations. We will have to use an approximation of when these stations were opened (Or we could reach out to Bikeshare directly) to see what effect they had in their locality. We will specifically filter for stations that were built in 2016

(tentatively) to limit our sample size. The coordinates for these stations can be used to assess their impact in terms of traffic from the TomTom traffic API, Mobility data from Travel Time API and Emissions data from the EPA. We will use traffic data from these coordinates roughly 3 years pre- and 3 years post-station installation.

Methodology

We will definitely be using API's for our data collection of traffic data. Additionally, we will have to engage in data wrangling to approximate the time in which the bikeshare location opened. We will be utilizing visualizations to demonstrate what impact a bikeshare had on the locality, using heat maps (found in Plotly) to identify places that are in need of more bikeshare infrastructure. We may also use GIS if it proves to be useful for the study. Inferential models will be developed using the statsmodels package.

Moreover, we will try to find areas in the DMV that could use bikeshare stations based on whether they suffer from the negative externalities that bikeshares are best at addressing (traffic, emissions & mobility) using predictive modeling. To develop our predictive model, we will utilize cross-validation and minimize the mean squared error.

Definition of Success

A successful project would be able to accurately identify what are the positive (and potentially) negative outcomes of installing a new bikeshare station in the DMV. We suspect that stations in conjunction with biking infrastructure would decrease congestion and emissions while increasing mobility for residents in the area.

The ideal project would be able to produce a heatmap that can identify areas that could most benefit from bikeshare infrastructure and then use our qualitative assessments to give recommendation to whether or not to expand there. A result that may surprise us is that the effects of Bikeshare stations are not consequential to public transit initiatives in the city.

References & Data

[23397 Capital Bikeshare Plan Update v4 051220 WEB.pdf](#)

[Bicycle infrastructure and traffic congestion: Evidence from DC's Capital Bikeshare - ScienceDirect](#)

[API Explorer | Traffic API | TomTom Developer Portal](#)

[Capital Bikeshare Locations | Open Data DC](#)

[System Data | Capital Bikeshare](#)

[Travel Time HTTP API Documentation | Walk Score](#)