Dock-based and dockless bikesharing systems: analysis of equitable access for disadvantaged communities Dataset

Dataset available at: https://doi.org/10.25338/B8X064

(This dataset supports report **Dock-Based and Dockless Bikesharing Systems: Analysis of Equitable Access for Disadvantaged Communities,** https://doi.org/10.7922/G2RN365J**)**

This U.S. Department of Transportation-funded dataset is preserved by the University of California in the digital repository Dryad (https://datadryad.org), and is available at https://doi.org/10.25338/B8X064

The related final report **Dock-Based and Dockless Bikesharing Systems: Analysis of Equitable Access for Disadvantaged Communities**, is available from the National Transportation Library's Digital Repository at https://rosap.ntl.bts.gov/view/dot/60052.

Metadata from the Dryad Repository record:

Author:

- Qian, Xiaodong, University of California, Davis, https://orcid.org/0000-0002-7245-3362; xdqian@ucdavis.edu
- Jaller, Miguel, University of California, Davis, https://orcid.org/0000-0003-4053-750X; mjaller@ucdavis.edu
- Niemeier, Debbie, University of Maryland, Baltimore; niemeier@umd.edu
- Hu, Miao, University of California, Davis, https://orcid.org/0000-0002-6667-3772; calhu@ucdavis.edu

Publication Date: December 8, 2021

Abstract:

Dockless bikeshare systems show potential for replacing traditional dock-based systems, primarily by offering greater flexibility for bike returns. However, many cities in the US currently regulate the maximum number of bikes a dockless system can deploy due to bicycle management issues. Despite inventory management challenges, dockless systems offer two main advantages over dock-based systems: a lower (sometimes zero) membership fee, and being free-range (or, at least free-range within designated service areas). Moreover, these two advantages may help to solve existing access barriers for disadvantaged populations. To date, much of the research on micro-mobility options has focused on addressing equity issues in dock-based systems. We have limited knowledge of whether, and the extent to which dockless systems might help mitigate barriers to bikeshare for disadvantaged populations. Using San Francisco and Los Angeles as case studies, because both cities have both dock-based and dockless systems running concurrently, we quantify bikeshare service levels for communities of concern (CoCs) by analyzing the spatial distribution of service areas, available bikes and bike idle times, trip data, and rebalancing among the dock-based and dockless systems. We find that dockless systems can provide greater availability of bikes for CoCs than for other communities, attracting more trip demand in these communities because of a larger service area and frequent bike rebalancing practices. More importantly, we notice that the existence of electric bikes helps mitigate the bikeshare usage gap between CoCs and other tracts. Besides the data analyses for bikeshare trips,

we also study the spatial distribution of online suggested station locations and find that the participants' desired destinations for work/school purposes have not been covered to the same extent in CoCs as in other communities. Our results provide policy insights to local municipalities on how to properly regulate and develop dockless bikeshare systems to improve mobility equity.

Methods:

There are three sets of data collected for this study: 1) communities of concern; 2) bikeshare data; 3) online suggestion data. We introduce the collection process for three datasets separately.

- **Communities of concern**: There is not a unified definition of disadvantaged populations or underserved communities across different regions in the US. Our previous study defines disadvantaged communities based on income, percentage of minority populations, and vehicle ownership (Qian and Niemeier 2019). In San Francisco, the MTC provides its own definition of disadvantaged communities for the transportation field. In a recent report evaluating dockless bikeshare in San Francisco, the SFMTA uses the MTC definition when conducting equity analyses (MTC 2018). The MTC terms and identifies disadvantaged populations as "Communities of Concern (CoCs)" based on the 2012-2016 American Community Survey (ACS) 5-year tract-level data. In Los Angeles, since different areas have their customized disadvantaged populations, we adopt a different definition of CoCs in Los Angeles, which is provided by the Southern California Association of Governments (SCAG) from its Plan Performance Environmental Justice Technical Report. In the report, SCAG investigated all Census Designated Places (CDPs) and City of Los Angeles Community Planning Areas (CPAs) and selected regions that fall in the top 33% of all communities in SCAG region for having the highest concentration of minority populations and low-income households (Southern California Association of Governments 2020). A person is classified as "minority" if the individual self-identified as one of the minority groups in the census (Table 4). SCAG performed poverty classification according to the income guidelines outlined by the U.S. Department of Health & Human Services.
- **Bikeshare data**: The analyses require data from both bikeshare systems. However, as would be true in cities across the U.S., the data from the two bikeshare systems in San Francisco/Los Angeles are not in the same format. Currently, Motivate and B-cycle operate most of the dock-based bikeshare systems in the US. Among all of these dock-based systems, Motivate operates Citi Bike (New York), Divvy (Chicago), Capital Bike Share (Washington DC), Ford GoBike (Bay Area), Biki (Honolulu), and Bluebikes (Greater Boston) which together contributed over 80% of all dock-based bikeshare trips in 2018 (NACTO 2019). All of the dock-based bikeshare systems operated by Motivate and B-cycle provide trip data, including information about trip start day and time, end day and time, start station, end station, bike id, and rider type (annual member or day pass user). For the members' trips, the database also includes the riders' gender and year of birth. However, the operators do not provide information on bike availability or rebalancing activities. We also found that these limitations exist for all dock-based bikeshare systems in the US because, currently, all operators only provide trip-based

data, not operational data. On the other hand, there is no trip data for dockless bikeshare systems; companies have not shared the data in this form because of privacy concerns or commercial advantages. However, they provide information through the General Bikeshare Feed Specification (GBFS), which is an open data standard for bikeshare. The GBFS provides real-time bike information (including bike id, location, battery level, and service status), and the number of available bikes in available hubs in a city. Unfortunately, the standard does not provide the bike id when the bike is at a hub. Additionally, the real-time bike data does not include any user data. Currently, many cities have required dockless bikeshare companies (e.g., JUMP, Bird, Lime, Lyft, Skip, Spin) to share real-time data in GBFS format. If a dockless bikeshare company provides data as required, it will be information in this format and available through an application programming interface (API). We developed a web-scraping (web data extraction) tool for the systematic and continuous collection of the real-time information from GBFS (e.g., JUMP Bike). Despite its limitations, the GBFS is very useful, and we developed a robust framework based on reasonable assumptions to infer other bikeshare data (e.g., bike availability and rebalancing operations) to support our analyses. For the study, we use historical bikeshare trip data provided by Ford GoBike; their database includes all bikeshare trips between 2013 and 2019. JUMP Bike (dockless), as mentioned, does not directly provide historical bikeshare trip data, thus, we use the web-scraping tool to gather minute-by-minute data from January to March 2019. We use this three-month sample data because, by March 2019, JUMP Bike had already been operating for over one year; thus, users were familiar with the service. Moreover, although there are some declines in bike ridership in San Francisco at the end of the year, ridership (based on data from bike counters) does not significantly fluctuate throughout the year (T. Winters 2017). For example, during 2018 the average number of bike counts (at the available bike counters) between January and March was 15,385 per month, which is 93% of the overall monthly average of 16,533. We found similar trends when analyzing the monthly trip numbers for Ford GoBike in San Francisco. The average number of monthly dock-based bikeshare trips between January and March was 210,598 in 2019, which is 3% above the average monthly usage (204,063 trips). As these findings were an indication that the three-month sample collected was representative, the analyses compare both systems within this period (January to March of 2019).

- **Online suggestion data**: The bikeshare operation companies (Baywheels in San Francisco or Metro Bike Share in Los Angeles) develop public online portals where users can suggest potential bikeshare station locations and comment on existing ones (Metro Bike Share 2021; Baywheels 2021). In this portal, users placed a dot (suggested location) on the maps of San Francisco and Los Angeles, providing a detailed description of their reasons (e.g., home, work/school, shopping, and fun) for choosing the locations, and their home zip codes. We applied the web scraping technology to download those suggestion data from these online portals. In San Francisco, we collected the historical information in the portal until the end of 2020. However, the data contained a number of duplicate locations by the same users and test data, which were removed. In the end, there is a total of 721 records of new bikeshare station suggestions available in San Francisco. In Los

Angeles, after removing duplicated data, there are 2,354 suggestions of new bikeshare stations.

Usage Notes:

Please refer to the data method section

Funding:

California Department of Transportation, Award: 65A0686 Task Order 023

Recommended citation:

Qian, Xiaodong; Jaller, Miguel; Niemeier, Debbie; Hu, Miao (2021), Dock-based and dockless bikesharing systems: analysis of equitable access for disadvantaged communities, Dryad, Dataset, https://doi.org/10.25338/B8X064

Dataset description:

This dataset contains 1 .zip file collection described below.

doi_10.25338_B8X064_v3.zip:

This collection contains 4 .zip file collections, 1 .csv file, 1 .xlsx file, and 1 .json file listed below.

- San_Francisco_Online_Suggestion_Data.csv
- San_Francisco_Communities_of_Concerns.zip
  - SF_COC.shx
  - SF_COC.shp.xml
  - SF_COC.shp
  - SF_COC.sbx
  - SF_COC.sbn
  - SF_COC.prj
  - SF_COC.dbf
  - SF_COC.CPG
  - Desktop.ini
- San_Francisco_Bikeshare_Data.zip
  - San Francisco Bikeshare Data Folder
    - Dockless Folder
      - JumpBike_Biketrip_Jan_March_1(4hour_0).csv
      - JumpBike_Bikerebalance_Jan_March.csv
      - JumpBike_Bikeidle_Jan_March.csv
    - Dockbased Folder
      - SF_Ford_GoBike_station.csv
      - GoFord_Bikerebalance_Jan_March.csv
      - GoFord_bikeidle_Jan_March.csv
      - Fordgobike-tripdata-201903.csv
      - Fordgobike-tripdata-201902.csv
      - Fordgobike-tripdata-201901.csv
- Readme.xlsx
- Los_Angeles_Online_Suggestion_Data.json
- Los_Angeles_Communities_of_Concerns.zip

- LA shapefile Folder
  - Data Souorce Document.docx
  - .DS_Store
  - Metro_Bike_Suggestions Folder
    - Metro_Bike_Suggestions.shx
    - Metro_Bike_Suggestions.shp
    - Metro_Bike_Suggestions.prj
    - Metro_Bike_Suggestions.dbf
    - Metro_Bike_Suggestions.cpg
  - Metro_Bike_Stations Folder
    - station.shx
    - station.shp
    - station.prj
    - station.dbf
    - station.cpg
    - .DS_Store
  - LA_County_Dissolved Folder
    - LA_County_Dissolved.shx
    - LA_County_Dissolved.shp
    - LA_County_Dissolved.prj
    - LA_County_Dissolved.dbf
    - LA_County_Dissolved.cpg
  - LA_County_Boundaries Folder
    - LA_County_Boundaries.shx
    - LA_County_Boundaries.shp
    - LA_County_Boundaries.prj
    - LA_County_Boundaries.dbf
    - LA_County_Boundaries.cpg
  - LA_Communities_of_Concern Folder
    - LA_Communities_of_Concern.shx
    - LA_Communities_of_Concern.shp
    - LA_Communities_of_Concern.prj
    - LA_Communities_of_Concern.dbf
    - LA_Communities_of_Concern.cpg
  - LA_Census_Designated_Places Folder
    - LA_Census_Designated_Places.shx
    - LA_Census_Designated_Places.shp
    - LA_Census_Designated_Places.prj
    - LA_Census_Designated_Places.dbf
    - LA_Census_Designated_Places.cpg
  - LA_CDP_CPA Folder
    - LA_CDP_CPA.shx
    - LA_CDP_CPA.shp
    - LA_CDP_CPA.prj

        - LA_CDP_CPA.dbf
        - LA_CDP_CPA.cpg
    - Community_Plans_Areas_(CPA)-shp Folder
        - Community_Plan_Areas_(CPA).shx
        - Community_Plan_Areas_(CPA).shp
        - Community_Plan_Areas_(CPA).prj
        - Community_Plan_Areas_(CPA).dbf
        - Community_Plan_Areas_(CPA).cpg
    - Communities_of_Concern Folder
        - Communities_of_Concern.shx
        - Communities_of_Concern.shp
        - Communities_of_Concern.prj
        - Communities_of_Concern.dbf
        - Communities_of_Concern.cpg
    - City_Boundaries Folder
        - City_Boundaries.shx
        - City_Boundaries.shp
        - City_Boundaries.prj
        - City_Boundaries.dbf
        - City_Boundaries.cpg
    - CA_Counties_Dissolved Folder
        - CA_Counties_Dissolved.shx
        - CA_Counties_Dissolved.shp
        - CA_Counties_Dissolved.prj
        - CA_Counties_Dissolved.dbf
        - CA_Counties_Dissolved.cpg
    - CA_Counties Folder
        - CA_Counties_TIGER2016.shx
        - CA_Counties_TIGER2016.shp.xml
        - CA_Counties_TIGER2016.shp
        - CA_Counties_TIGER2016.sbx
        - CA_Counties_TIGER2016.shn
        - CA_Counties_TIGER2016.prj
        - CA_Counties_TIGER2016.dbf
        - CA_Counties_TIGER2016.cpg
    - CA_Census_Designated_Places Folder
        - tl_2014_06_place.shx
        - tl_2014_06_place.shp.xml
        - tl_2014_06_place.shp.iso.xml
        - tl_2014_06_place.shp.ea.iso.xml
        - tl_2014_06_place.shp
        - tl_2014_06_place.prj
        - tl_2014_06_place.dbf
- Los_Angeles_Bikeshare_data.zip

- Bikeshare data Folder
  - Desktop.ini
  - Dockless Folder
    - LA_JUMP_Biketrip_2020-02(4hour_0).csv
    - LA_JUMP_Biketrip_2020-01(4hour_0).csv
    - LA_JUMP_Biketrip_2019-12(4hour_0).csv
    - LA_JUMP_Bikerebalance_2020-02.csv
    - LA_JUMP_Bikerebalance_2020-01.csv
    - LA_JUMP_Bikerebalance_2019-12.csv
    - LA_JUMP_BikeIdle_2020-02.csv
    - LA_JUMP_BikeIdle_2020-01.csv
    - LA_JUMP_BikeIde_2019-12.csv
    - Desktop.ini
  - Dockbased Folder
    - metro-bike-share-trips-2020-q1.csv
    - metro-bike-share-trips-2019-q4.csv
    - Los_Angeles_Metro_Bikeshare_BikeRebalance_Data.csv
    - Los_Angeles_Metro_Bikeshare_Bikelding_Data.csv
    - Destop.ini
    - Metro Bike Stations Folder
      - station.shx
      - station.shp
      - station.prj
      - stations.dbf
      - station.cpg
      - desktop.ini
      - .DS_Store

File Type Descriptions:

- The .csv, Comma Separated Value, file is a simple format that is designed for a database table and supported by many applications. The .csv file is often used for moving tabular data between two different computer programs, due to its open format. The most common software used to open .csv files are Microsoft Excel and RecordEditor, (for more information on .csv files and software, please visit https://www.file-extensions.org/csv-file-extension).
- The .txt file type is a common text file, which can be opened with a basic text editor. The most common software used to open .txt files are Microsoft Windows Notepad, Sublime Text, Atom, and TextEdit (for more information on .txt files and software, please visit https://www.file-extensions.org/txt-file-extension).
- The .xlsx file type is a Microsoft Excel file, which can be opened with Excel, and other free available software, such as OpenRefine.
- The .json file extension is associated to JavaScript Object Notation file format, a lightweight, text-based, language-independent data interchange format. The data in a json file are stored in simple text file format and the content is viewable in any simple text

editor (for more information on .json files and software, please visit https://www.file-extensions.org/json-file-extension).

- The shp file extension is used for ESRI Shape format, a popular geospatial vector data format for geographic information systems software. A shapefile stores non-topological geometry and attribute information for the spatial features in a data set. A shapefile consists of a set of 3 mandatory files, along with several optional files. Each file in the set shares the shapefile name with a different extension. The main file .shp stores the geometry and must always have an index file shx. A dBASE file dbf stores all the attributes of the shapes in the main file. Additionally, a projection file prj stores the projection information (for more information on .shp files and software, please visit https://www.file-extensions.org/shp-file-extension).
- A .shx file contains compiled shape data (building blocks, fonts) in form of the machine language compiled version of an shp file. The file type is associated with AutoCAD , a CAD development platform, developed and sold by Autodesk, Inc (for more information on .shx files and software, please visit https://www.file-extensions.org/shx-file-extension).
- The prj file extension is traditionally used for files that contain projects. Projects contain settings, positions of saved windows, development notes and other raw data that won't be present in the final file (for more information on .prj files and software, please visit https://www.file-extensions.org/prj-file-extension).
- The dbf file extension is traditionally used for database file by many database applications. The original program, which used the DBF file extension for its database, was dBAse. A major legacy of dBase is its dbf file format, which has been adopted in a number of other applications. For example, the shapefile format developed by ESRI for spatial data in a geographic information system uses .dbf files to store feature attribute data (for more information on .dbf files and software, please visit https://www.file-extensions.org/dbf-file-extension).
- The sbx file extension is associated with the ArcView GIS application used to view and edit GIS data. The .sbx file contains spatial index for read-write shape used to fast access to shape files (for more information on .sbx files and software, please visit https://www.file-extensions.org/sbx-file-extension-arcview-spatial-index-for-read-write-shape-file).
- The sbn file extension is mainly associated with ArcView / ArcGIS geography software from ESRI. The sbn file contains binary spatial indexes, which are used only by ESRI software. The format is not documented, and is not implemented by other vendors (for more information on .sbn files and software, please visit https://www.file-extensions.org/sbn-file-extension).
- The cpg file extension is associated with the ArcGIS, a geographic information system for Microsoft Windows operating system, developed by Esri. The cpg file stores codepage for identifying a characterset (for more information on .cpg files and software, please visit https://www.file-extensions.org/cpg-file-extension-arcgis-codepage).
- The ini file extension is primarily used for configuration files in many Windows applications. A typical .ini file contains set of parameters for the program to be used when its started. Probably short for initialize. (for more information on .ini files and software, please visit https://www.file-extensions.org/ini-file-extension)

National Transportation Library (NTL) Curation Note:
As this dataset is preserved in a repository outside U.S. DOT control, as allowed by the U.S. DOT's Public Access Plan (https://ntl.bts.gov/public-access) Section 7.4.2 Data, the NTL staff has performed NO additional curation actions on this dataset. NTL staff last accessed this dataset at https://doi.org/10.25338/B8X064 on 2022-02-04. If, in the future, you have trouble accessing this dataset at the host repository, please email NTLDataCurator@dot.gov describing your problem. NTL staff will do its best to assist you at that time.