9^{th} International Conference on Computational Social Science IC^2S^2
July 17-20, 2023, Copenhagen, Denmark

Disaggregating bike-share mobility data to measure urban segregation over time

urban segregation; mobility; transportation networks; network dynamics; bike share

Extended Abstract

Globally, cities are aiming to increase accessibility and mobility. Initiatives to reduce commute times, better connect neighborhoods, and modernize transportation infrastructure come on the heels of growing economic and racial segregation and inequality in urban areas, which has been shown to harm sustainability, public health, and child wellbeing, among many societal outcomes [3, 4]. Within these initiatives, increasing the use of bicycles and the availability of bike-share programs has become a common means of improving urban mobility. Bicycles are a green, sustainable mode of transportation and are effective for connecting communities with other modes of transportation [1]. In several major U.S. metropolitan areas, bike-share programs enable city residents and tourists to bike between docking stations on demand. In this work, we analyze 8 years of publicly available ride data for the city of Boston's bike-share program, known as "Bluebikes". By merging the ride data with census demographics we are able to address the following two research questions: First, while aggregate ride activity has been studied, how has bike-riding activity differed across neighborhoods over time, especially during the COVID-19 pandemic? Second, has urban segregation, as measured through bike transportation networks, increased over time in Boston, and if so, in which areas?

To analyze urban segregation in Boston, we merge two publicly available datasets: the Bluebike ride data and the 2020 decennial U.S. census demographic data. Since its inception, the Bluebikes program has publicly recorded all rides taken, reporting the start and end location and times for each ride. We analyze data for all rides taken between January 2015 to November 2022, totaling over 15 million rides. In the same time period, the number of bike stations grew from 140 to over 400 stations. We merge the bike station data with demographic data from the 2020 decennial U.S. census, building on past works that merge mobility data with census data [2]. For each station, we identify the surrounding census tract and label the station based on the most prevalent racial group in the corresponding census tract. For the 515 stations that were in use between 2015 - 2022, 3.9% are located in predominantly Asian tracts; 9.9% are located in predominantly Hispanic and Latino tracts; 10.5% are located in predominantly Black tracts; and 75.8% are located in predominantly White tracts.

By merging the ride and census data we are able to disaggregate ride activity and verify whether previously reported patterns are consistent across census demographics. For instance, it is known that over the course of the COVID-19 pandemic, ridership has not only increased but the proportion of weekend rides has increased relative to weekday commute rides.[1] We find that the transition from weekday to weekend rides in most prominent in predominantly White census tracts. Figure 1a shows that in 2019, 77% of all rides originating from predominantly White census tracts took place on weekdays; however, by 2022, the same figure decreased to 71.6%. In concert with the growing prevalence of weekend rides, the fraction of riders who are Bluebike subscribers, as opposed to single-trip riders, has also decreased, most noticeably

[1] `https://www.wbur.org/news/2022/10/11/boston-city-bike-increase-baker-marijuana-convictions-traffic-newsletter`

in predominantly White and Hispanic or Latino tracts. Figure 1a also shows that the onset of COVID-19 lockdowns in March 2020 had varying impacts across demographics, with the number of rides originating from Hispanic or Latino census tracts least impacted.

Beyond patterns in rider activity based on station demographics, we also measure urban segregation in Boston over time. Moro et al. [2] introduces a measure of urban segregation based on random walks. Given a graph $G = (V, E)$ in which each node is labeled with a class label, Moro et al. [2] count the number of steps a random walk beginning at node i requires to encounter each class label at least once, known as the class-coverage time. Segregation can be quantified by aggregating the distribution of random-walk lengths across all origin nodes, where longer walk lengths indicate higher levels of segregation. We measure mobility segregation in Boston over time by constructing a station network for each month of ride data and aggregating the lengths of random walks; the nodes in the network are the stations, labeled by the most prevalent racial demographic in the corresponding census tract, and a directed edge from node u to v is weighted by the number of rides starting at station u and ending at station v in the given month. Weighting the edges based on ride volume ensures that the random walks reflect user activity.

Our analysis of random-walk lengths indicates that while mobility segregation in Boston was declining prior to the start of the COVID-19 pandemic, the amount of segregation increased during the pandemic. From 2015-2020, the average walk length was smallest for walks beginning at stations located in predominantly Black census tracts, possibly because these stations are located centrally in the city. Meanwhile, the average walk length for all other demographic groups decreased in the pre-COVID time period. However, during COVID-19, our results, visualized in Figure 1b, show that not only has the number of rides increased but the level of segregation has also increased. In fact, while walks beginning in predominantly Black census tracts were previously the shortest, from 2019 to 2022, the average random-walk length for predominantly Black census tracts increased by 90.2%; meanwhile, the average walk length for Hispanic or Latino tracts increased by 94.3%. For comparison, the average walk length for White census tracts increased by 43.6%. These results highlight the need for greater accessibility and diversity of contact in Boston as well as the importance of analyzing mobility data at more granular levels.

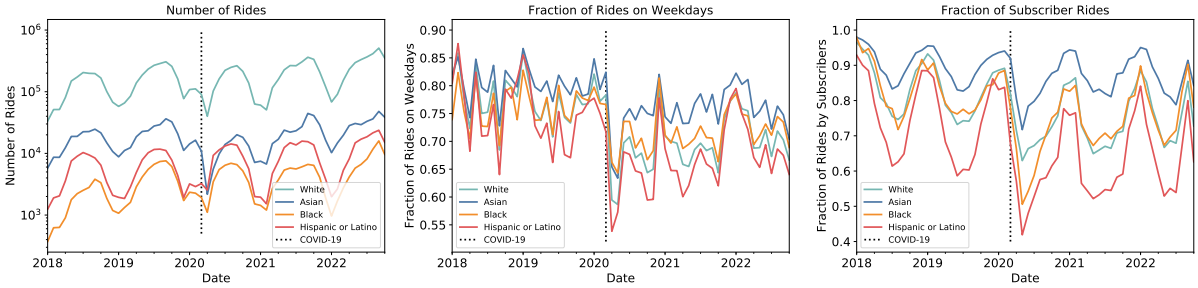
References

[1] Go Boston 2030 Mayoral Advisory Committee. *Go Boston 2030: Vision and Action Plan*. City of Boston, 2017.

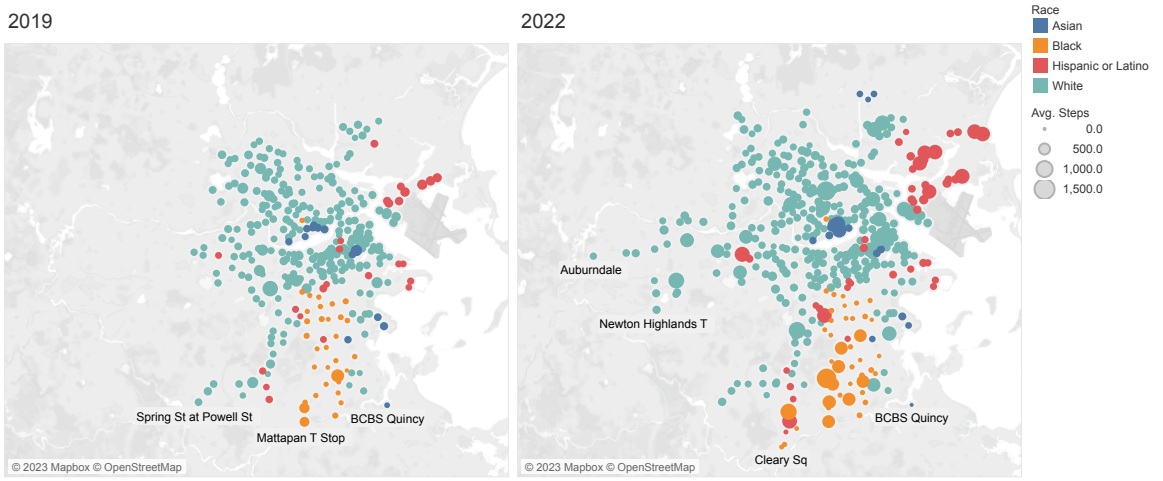
[2] Esteban Moro, Dan Calacci, Xiaowen Dong, and Alex Pentland. Mobility patterns are associated with experienced income segregation in large US cities. *Nature Communications*, 12(1):4633, July 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-24899-8.

[3] Sandro Sousa and Vincenzo Nicosia. Quantifying ethnic segregation in cities through random walks. *Nature Communications*, 13(1):5809, October 2022. ISSN 2041-1723. doi: 10.1038/s41467-022-33344-3.

[4] David Tedjopurnomo, Zhifeng Bao, Farhana Choudhury, Hui Luo, and A. K. Qin. Equitable public bus network optimization for social good: A case study of singapore. In *FAccT '22*, page 278–288, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450393522. doi: 10.1145/3531146.3533092.



(a) Disaggregation of rider activity by census tract demographic



(b) Racial segregation measured with random walks

Figure 1: Merging public bike-share data for the city of Boston with census data allows us to disaggregate ride activity patterns by racial demographics and determine regions of the city where mobility segregation has increased the most. We label each bike docking station based on the most prevalent race in the surrounding census tract. Figure 1a shows how trends in ride activity have varied over time by census-tract racial demographic, where the dashed line indicates the start of COVID-19 lockdowns in March 2020. For instance, predominantly Hispanic or Latino census tracts have the lowest fraction of weekday rides, which are generally associated with workplace commutes. Figure 1b visualizes the map of Bluebike stations in 2019 and 2022. We measure mobility segregation by counting the average length of random walks starting from each station, where walks terminate once all racial labels have been encountered. Longer walks suggest higher levels of segregation. The map for 2022 shows that the segregation of Black and Hispanic or Latino tracts noticeably increased in the three-year period.