Bluebikes Usage in Boston: A Study

Jake Phelan, McKenna Quam

Abstract

This report is an investigation of the usage of Bluebikes in the greater Boston metro from January 2019 to March 2024. The database includes data from 15 million rides, between 558 stations and in 13 separate municipalities. We find that Bluebikes are not an effective solution to the commuting needs of most working professionals due to the changes in Boston's weather. The bikes are used heavily by those who need to make short and otherwise inaccessible trips (e.g. college students crossing the Charles River), but are less popular for the longer trips of subscription-paying users.

Introduction

Public transportation is an important part of a city's infrastructure. According to data from the 2014 American Community Survey, 34% of Boston work commutes are done using public transit. Boston's public transit system includes the T subway and tram system, the commuter rail, and a bus network. Newer to the scene—the focus of our investigation—is a bike-sharing system called Bluebikes.

Bluebikes is a public bike rental network in Boston and its surrounding towns. Established in 2011 as "Hubway", it adopted its current name after being bought out by the Blue Cross Blue Shield federation. Currently, the network operates over 400 stations with a fleet of over 4,000 bikes in partnership with the rideshare company, Lyft.

To rent a Bluebike, one has the choice between purchasing a single-use pass, an unlimited day pass, or an unlimited annual pass. It costs an initial flat-rate of $2.95 to check out a bike for 30 minutes, and an extra $4.00 for every 30 minutes beyond the original half hour. This price is comparable to a one-way ride on the T, including station transfer. However, the one-time Blue Bike fare actually exceeds the $2.40 T ticket by 55 cents. On the other hand, a monthly pass for the T is much more expensive than a monthly pass for Bluebikes. The MBTA, the public agency responsible for managing the T, prices unlimited passes for the trains at $90 per month. A Bluebikes pass costs $23 monthly, or about $130 annually (under $11 per month). With this disparity in mind, bike-sharing seems to offer a more economical form of public transit for Bostonians with regular commutes. We'd like to remark that both the T and Bluebikes offer similarly-priced daily passes, but this report primarily considers the viability of bike-sharing for local Bostonians.

Bike-sharing is also a more environmentally friendly method of transit than traditional cars. Bikes neither consume fossil fuel nor release exhaustive fumes! Bluebikes has even recently introduced battery-powered electric bikes, requiring less man-power than a manual bike for a trip of the same distance. Bluebikes has not yet announced if it is harvesting renewable

energy to charge the bikes' batteries. Nonetheless, electric biking has one of the lowest carbon footprints of all electric transit methods to date.

Methods

Bluebikes publishes new data at the end of every month, including many features from all rides taken using its service during that month. The company also maintains an up-to-date file containing details regarding its active bike-docking stations. The data goes back to 2011 but for the purposes of this report, we sampled the data from 2019 and on; we wanted to consider only the data after the aforementioned Blue Cross Blue Shield acquisition in 2018. Although the data is only a few years old, inconsistencies in the architecture of the dataset can be found between files of almost any two years. MongoDB is therefore a better tool for managing this data over a relational database like SQL for several reasons.

For starters, the data is not very relational. The only relational subfield in a ride's features is the data for the start and end station. The monthly ride data files report the name and station ID of a given Bluebike station, and the consistently-upkept stations dataset reports any remaining relevant information. For example, a station sample contains its municipality and number of docs. We noticed that this information can be condensed into just two fields of a ride document, by tracking each trip's start and end station information in separate sub-documents. This may not be the optimal solution in respect to database memory, but seemed to be more time-efficient in its enabling of non-relational aggregation queries rather than applying JOINs in a relational database.

On a similar note, the dataset is very large—even after our sampling. The data of 15,622,370 rides has been reported from January, 2019 through March of this year. While there is not a strict limit on the size of SQL databases, the time performance of these relational tables slows tremendously for queries and aggregation at large scales.

Finally, the data is not in a consistent schema. Features reported in the files have changed several times over the data repository's existence, particularly for the stations. In April of 2023, the IDing system for stations was changed: what was previously a plain integer became an alphanumeric value (e.g. M32026). If we tried to perform a SQL JOIN on these stations, it would fail for any trip before April, 2023. For this analysis, we wanted to cover ride data before, during, and after the COVID-19 pandemic, so this lack of consistency would not work within the confines of a relational database. Within the ride's data, there exist some features which were historically recorded but are no longer tracked—like gender and bike ID—and some features which were recently added—like bike type (classic or electric).

Before putting the data into a MongoDB database, some preprocessing measures in Python had to be taken. First was the task of transforming each data file from CSV to JSON formatting. As previously mentioned, the methodology of station ID-tracking was not consistent between current and older stations, and the current stations file did not even include out-dated stations. Instead of trying to use the problematic ID values, the station name was used as the station-identifying feature of choice, as those stayed consistent when the IDs were changed. To do this, we created a dictionary with station names as keys and each value was another dictionary with the remaining desirable data (name, coordinates, municipality, total docks, and seasonal status) as entries. We iterated over the remaining recent trips datasets, adding the

historical stations to the stations dictionary; however the historical entries only contained the station name and the coordinates. Due to MongoDB's lack of schema enforcement, the missing subfields did not pose any issue to the schema. Finally, we assigned one of these nested dictionaries to the value of the "station start" and "station end" fields while creating the rides JSON file. The recent stations considered comprise 96% of the data overall; if only the most recent stations were included, 488,070 rides would have been excluded.

Ride duration was calculated from the datetime object which was provided in the trips data CSV and the start and end times were broken up into a dictionary containing each part of the datetime as a field to make for easier aggregation.

One ride document in the database looks like this:

```
{
        _id: ObjectId('66107737428bd6a4e7baf628'),
        subscribed: 0,
        start_time: { year: 2024, month: 2, day: 16, hour: 11, minute: 30, second: 45 },
        end_time: { year: 2024, month: 2, day: 16, hour: 11, minute: 38, second: 50 },
        duration: 485,
        start_station: {
                name: 'Forsyth St at Huntington Ave',
                coordinates: [ -71.090511, 42.339202 ],
                municipality: 'Boston',
                total_docks: 15,
                seasonal_status: 'Year Round'
        },
        end_station: {
                name: 'Boylston St at Massachusetts Ave',
                coordinates: [ -71.08678415, 42.34740621 ],
                municipality: 'Boston',
                total_docks: 15,
                seasonal_status: 'Year Round'
        },
        bike_type: 'classic_bike'
}
```

(This is a ride which one of our authors took to get groceries)

Analysis

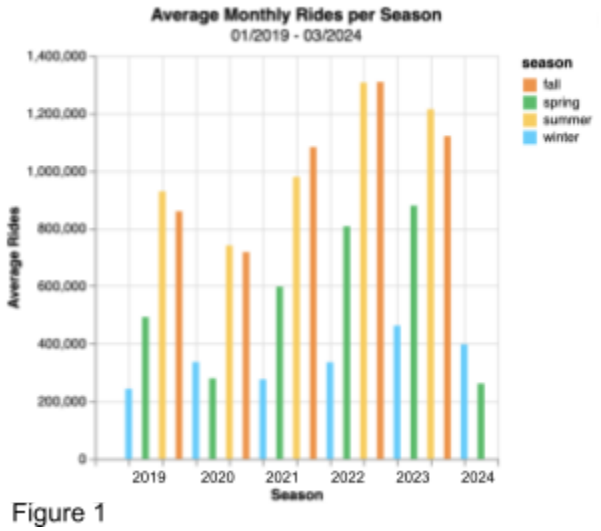

Figure 1

Seasons:

Boston's weather conditions vary vastly between seasons; this is one of the biggest hurdles to user accessibility for biking as a method of transportation in the city. Figure 1 shows the average amount of monthly rides divided across the seasons in the data (Fall is September – November, Spring is March – May, Summer is June – August, and Winter is December -– February). Ride counts are at their greatest when the weather is nicest in the summer and fall seasons, before taking a sharp dip in winter and then starting to rise again in the spring. Within the graph, one can see the effect of the pandemic on people's ability to go out of the house: the spring 2020 season sees a downturn in average monthly rides as opposed to the usual annual pattern of a spring increase. We have not yet completed the spring 2024 season, but it is likely that the average number of rides will rise above the previous winter season, as seen in nearly every year previously (excluding 2020).

Locations:

In trying to understand people's motivations around using Bluebikes, it is useful to get a sense of where they are traveling to and from. The top 10 most popular routes make up about 1% of the rides in the database, and they all start or end at a station on the MIT campus. The top route is from MIT at Mass Ave / Amherst St station to Beacon St at Massachusetts Ave, which has been ridden 19,985 times. The start and end points of this route are stations on either side of the Charles River bridge on Massachusetts Avenue. The inverse route, from Beason St to Amherst St, is the 4th most popular route, having been ridden 18,818 times. This route is taken by Bluebikes subscribers 90% of the time (34,969 times among subscribers vs 3,834 times by non-subscribers). For non-subscribers, paying $2.95 to cross the Charles Bridge is a superfluous expense, but for subscribers it could be a great time saver. For MIT students, being able to enter Boston over the Mass Ave bridge is more efficient than going to the Red Line T station–a 20 minute walk away from the center of campus on the Longfellow Bridge.

Interestingly, the 17th most popular route is a circular one that goes from Mugar Way at Beacon St to Mugar Way at Beacon St, which has been ridden 12,794 times. Having the same start and end station would indicate that the ride is not for transportation, but instead for leisure. The Mugar Way station is right next to the Boston Common, which is an ideal place to ride a bike leisurely. The average duration for these trips was around 55 minutes, which is almost double the average trip time of 29 minutes. The other top round trip routes are also by parks or by the waterfront: #2 is MIT at Mass Ave / Amherst St, which is by the Esplanade, and #3 is Murphy Skating Rink - 1880 Day Blvd, which is near the bike path surrounding the Pleasure Bay Bike Path. All of the round trips have an average duration of 50 minutes.

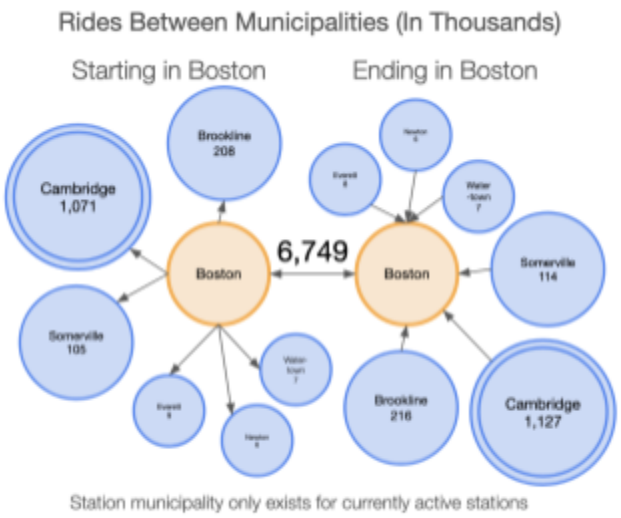

Figure 2

From a broader perspective, we can group the trips by municipality to see which area of Boston people are traveling to and from. The municipality data only is available for the currently active stations, which exclude the 4% of trips which start or end at a historical station. Overwhelmingly, people are traveling from Boston to Boston. Figure 2 focuses on all the most popular municipality connections which either start or end in Boston. The bubbles are different municipalities with the trips on the left being ones which start in Boston and the trips on the right being the ones that end in Boston. The size of the bubbles correlate to the number of trips which go between those 2 municipalities. There are over 6.7 million trips which go from a Boston station to a Boston station which is then followed by trips to/from Cambridge. The first and fourth most popular routes across the Charles are routes between Cambridge and Boston; the convenience of crossing the river using Bluebikes seems to contribute to the number of rides. Inter-Cambridge rides are the second most common, with 3.3 million rides. Biking for transportation is most practical when the distance you have to travel is short, so it follows that the most popular routes are within the same municipality or adjacent municipalities. The drop in the number of rides to other municipalities past Cambridge, Brookline, and Somerville suggests that bike-sharing is not being used heavily for commuting into the city from the suburbs. Likely, the distance is too long to justify biking over taking other kinds of transposition like the T or cars.

Subscribed vs Non-Subscribed Usage:

With regular usage, purchasing a subscription plan to Bluebikes is clearly the more economical choice: taking just 1 ride per week, subscribers save $26 compared to casual users and this disparity climbs to $182 with 2 rides per week. However, not everyone uses Bluebikes

Figure 3

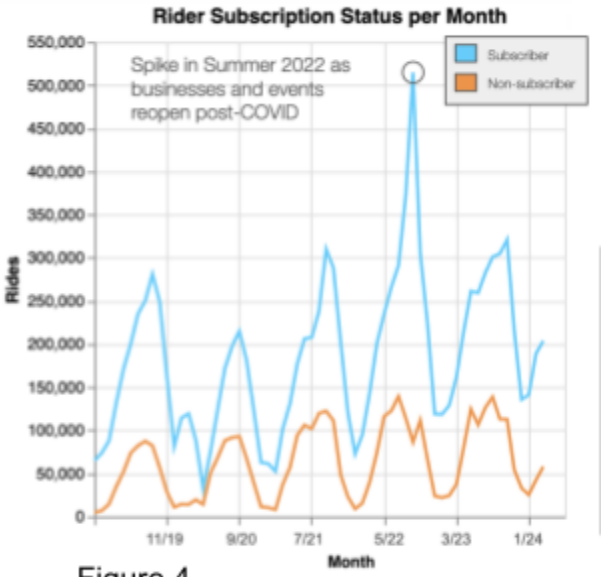

Figure 4

so often, and Boston has a high rate of tourists and visitors. It therefore follows that not all customers automatically decide to subscribe to the service.

Figure 3 tracks the monthly average duration of rides taken by subscribers and by non-subscribers over the course of the data. It is clear that non-subscribers tend to take much shorter rides than non-subscribers. At first, this may seem counterintuitive–why pay for a subscription to a service you do not use for long durations when there is à la carte pricing available? Figure 4 answers this question and clarifies the picture of subscription-based usage. Tracking the monthly number of rides of subscribers and non-subscribers shows that subscribers do take many more rides than their non-subscribed counterparts. This higher count, yet shorter duration, of rides overall by subscribers implies that the subscribers are using their subscriptions to take shorter and more frequent trips. Likely destinations for these rides include a user's local grocery store, dry cleaner, or pharmacy. Subscribers are not likely to take a long ride, say, to their place of work–it makes more sense to take a method of transit like a car or the T, due to prospective inappropriate biking weather and/or excessive muscle strain. This conclusion comes in agreement with the earlier analysis of weather and seasons in Boston.

On a grander scale, it is interesting to see how subscriber ride counts change over time as a whole. Like before, a dip in the overall popularity of Bluebikes can be observed in 2020 during the COVID-19 pandemic. However, a major spike in subscribers' ride count happened in the summer of 2022, implying that people were looking for ways to celebrate the end of the pandemic and exercise their newly found freedom to be outside of the house.

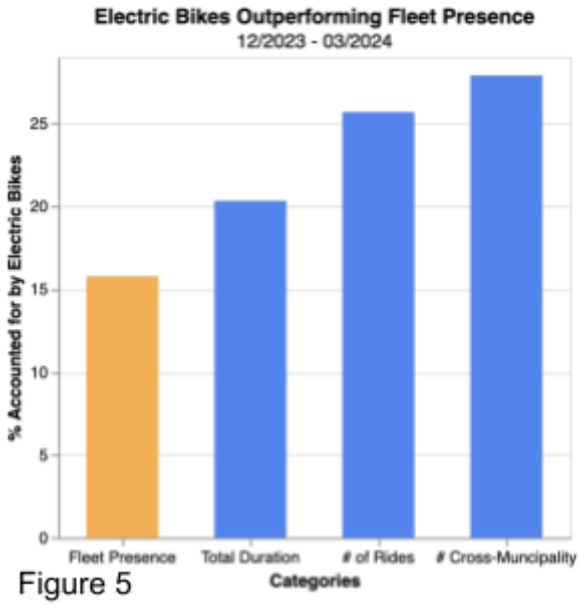

Figure 5

Electric Bikes:

In December of 2023, Bluebikes introduced electric bikes into the Boston bike-share network. These bikes can go up to 18 miles per hour, which could make them a better transport solution for people commuting longer distances. We have data for 4 months of usage so far which shows that the average trip taken on an electric bike is shorter than a trip on a regular bike. This could be due to the bikes' speed or due to the fact that Bluebikes charges by the minute to ride the electric bikes: $0.10 per minute for subscribers and an extra $2.50 per minute for non-subscribers. Despite their higher costs, people seem to be using these new electric bikes at a higher rate than expected. Within the Boston bike-share system, there are around 4,000 non-electric bikes and 750 e-bikes: e-bikes now comprise around 15% of the current fleet of bikes. Figure 5 compares the current percentage of the e-bikes in the fleet to the percentage they represent in the total ride duration (20.3%), number of rides (25.7%), and number of rides which cross municipalities (27.9%). Clearly, e-bikes are being used disproportionately to their population relative to regular bikes. The e-bikes are still a novelty item, so it remains to be seen if this trend will continue, but for

now it seems that the higher proportion of rides which cross municipalities show how e-bikes can be a tool for longer commutes where standard bikes might be inadequate.

Summary

In conclusion, our study considers the dynamics of Bluebikes usage in the greater Boston metro, spanning from January 2019 to March 2024. Through the analysis of over 15 million rides across over 5 years, several key insights emerge regarding the effectiveness of bike-sharing as a mode of transportation.

Firstly, while Bluebikes provide an economically and environmentally friendly alternative to traditional transit options, their efficacy for commuting purposes remains limited. The fluctuating weather conditions in Boston, particularly during winter months, significantly impact user accessibility and preference. Recently, the introduction of electric bikes into the Bluebikes system marks a notable shift, with unexpectedly high usage rates suggesting their potential as a solution for longer commutes.

Additionally, this study outlines the distinct usage patterns between subscribers and non-subscribers, with the former opting for more frequent but shorter trips. This highlights the importance of catering to diverse user preferences and motivations, as well as the need for targeted marketing strategies to encourage subscription adoption.

The outlook for Bluebikes is promising yet there remain improvements to be made. Primarily, expanding the electric bike fleet seems to be an essential step towards promoting bike-sharing as a viable transit option for a wider audience. Our study sheds light on the relations between user behavior, environmental considerations, and infrastructure development of bike-sharing in Boston. With a fuller understanding of the data, we can strive towards creating a more accessible, efficient, and sustainable transportation ecosystem for all residents and visitors alike.

Works Cited

Authority, Massachusetts Bay Transportation. “Main Menu.” MBTA, 2024, www.mbta.com/fares/subway-fares.

BlueShield, BlueCross. “About Bluebikes.” Bluebikes, 2024, bluebikes.com/about.
Department, Boston Transportations. “How Boston Travels to Work.” Boston Today, 2017, www.boston.gov/sites/default/files/file/document_files/2017/03/go_boston_2030_-_3_boston_today_spreads.pdf.

BlueShield, BlueCross. “Bluebikes System Data: Blue Bikes Boston.” Bluebikes, 2024, bluebikes.com/system-data.

MilNeil, Christian. “Electric Bluebikes Have Arrived in Boston.” Streetsblog Massachusetts, 22 Dec. 2023, mass.streetsblog.org/2023/12/20/electric-bluebikes-have-arrived-in-boston.

Motivate International, Inc. “Pricing.” Blue Bikes Boston, 2024, blog.bluebikes.com/pricing.

Odot, Loraine. “What Is the Carbon Footprint of Electric Bikes?” Polytechnique Insights, 19 Mar. 2024, www.polytechnique-insights.com/en/columns/energy/what-is-the-carbon-footprint-of-electric-bikes/.