

# Boston Bus Performance

## Team F

Jialu Li, Junyi Li, Laksanawist Mutiraj, Qinfeng Li, Yifei Zhou



# Project Background and Goal



- ***Background:***

- Historical Roots:** Public transport in Boston dates back nearly 400 years, with ferry services originating in the 1600s.

- MBTA's Impact:** Serving over 1 million people daily, the MBTA contributes an estimated \$11.5 billion annually to the greater Boston area's economy.

- ***Goal:***

- Examining MBTA bus service performance trends by geography with important insights into potential disparities by neighborhood and other demographic factors.

# Work done so far

Dataset used for base questions:

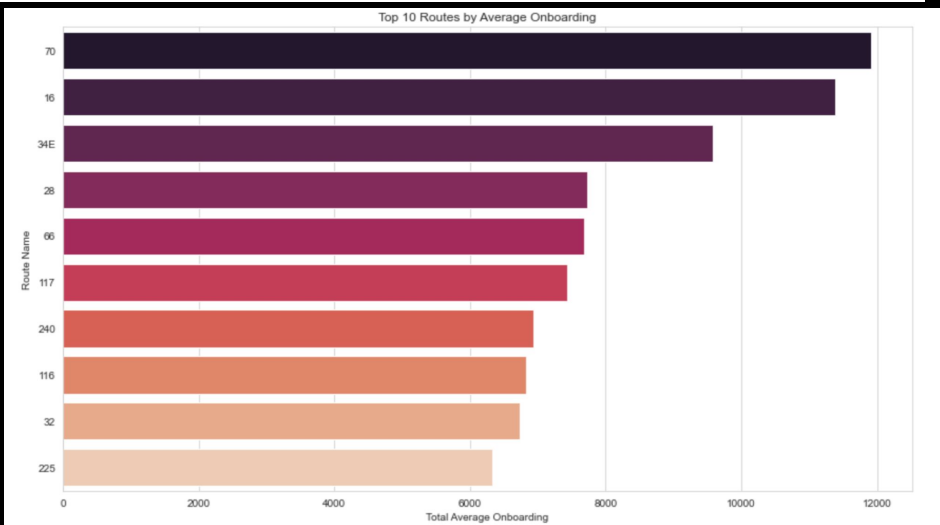
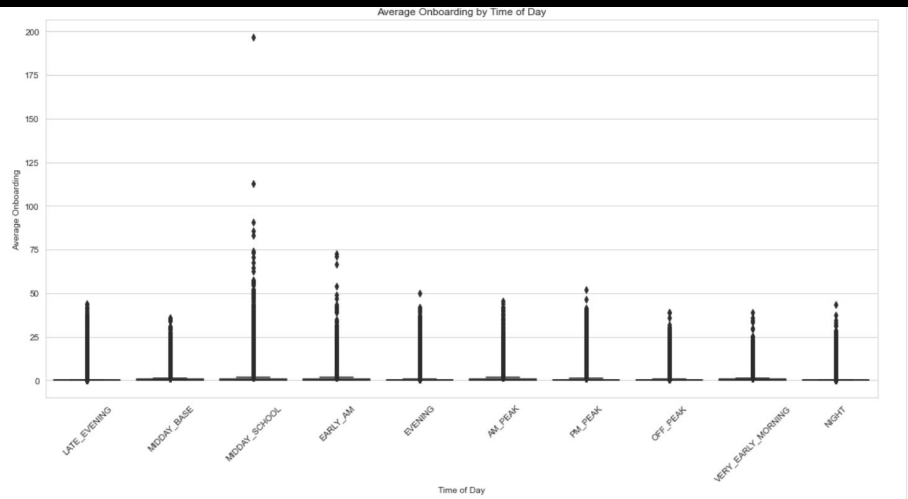
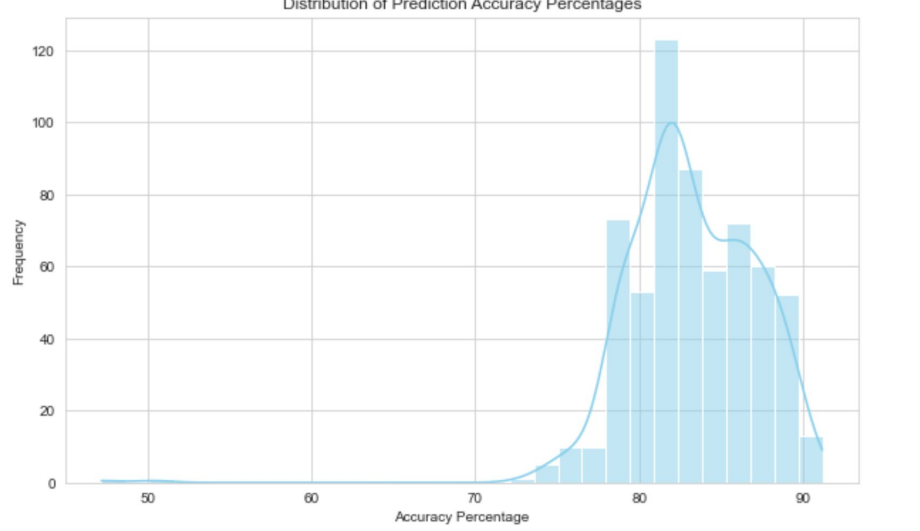
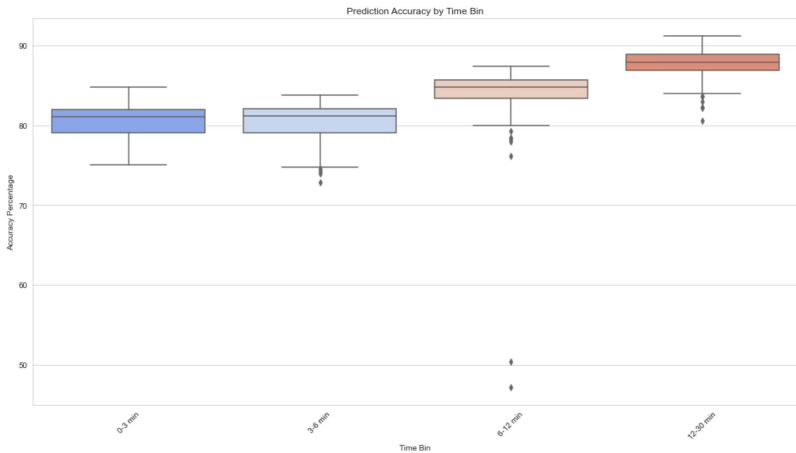
- Rapid\_Transit\_and\_Bus\_Prediction\_Accuracy\_Data
- MBTA\_Bus\_Ridership\_by\_Time\_Period%2C\_Season%2C\_Route\_Line%2C\_and\_Stop
- Boston\_Neighborhood\_Boundaries\_approximated\_by\_2020\_Census\_Block\_Groups
- Bus\_Network\_Redesign\_Draft\_Bus\_Routes
- Commuter\_Rail\_Reliability

Meeting Frequency: once a week and daily communication through Discord

Task division:

- Data Collection: Yifei Zhou, Qinfeng Li, Laksanawisit Mutiraj
- Exploratory Data Analysis for base questions: Jialu Li, Junyi Li
- Extension Project: Junyi Li

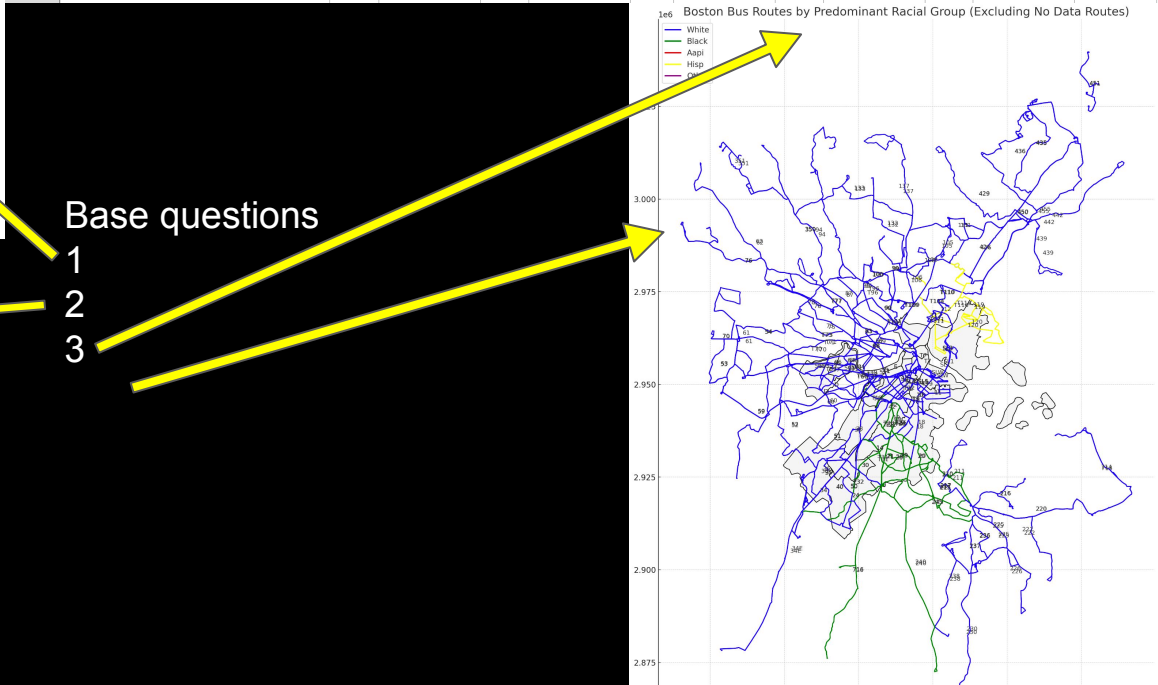




		route_id	direction_id	travel_time_seconds	average_travel_time
2	0	01	Inbound	2186.178082191781	0 days 00:36:26.178082192
3	1	01	Outbound	2036.2826523777628	0 days 00:33:56.282652378
4	2	04	Inbound	1362.3664122137404	0 days 00:22:42.366412214
5	3	04	Outbound	1295.1330798479087	0 days 00:21:35.133079848
6	4	07	Inbound	983.1116687578419	0 days 00:16:23.111668758
7	5	07	Outbound	876.9397590361446	0 days 00:14:36.939759036
8	6	08	Inbound	2973.036211699164	0 days 00:49:33.036211699
9	7	08	Outbound	3279.767441860465	0 days 00:54:39.767441860
10	8	09	Inbound	1823.0196749358427	0 days 00:30:23.019674936
11	9	09	Outbound	1403.5589421783952	0 days 00:23:23.558942178
12	10	10	Inbound	2215.960502692998	0 days 00:36:55.960502693
13	11	10	Outbound	2113.189189189189	0 days 00:35:13.189189189

gtfs_route_id	reliability_score
9703	32.00941915227626
449	40.25524468576182
448	40.630198757585696
459	42.997043635764754

bus_routes_race_data												
route_shor	route_long	white_all	white_percentage	black_all	black_percentage	aspi_all	aapi_percentage	hisp_all	hisp_percentage	other_all	other_percentage	total_population
10	modified - MF (South Boston-Andrew-Ruggles)	210242.0	38.945820181093400	145108.0	26.880214585278400	66182.0	12.259740067280200	107284.0	19.873590302168100	11016.0	2.040634864179970	539832.0
10	modified - MF (South Boston-Andrew-Ruggles)	210242.0	38.945820181093400	145108.0	26.880214585278400	66182.0	12.259740067280200	107284.0	19.873590302168100	11016.0	2.040634864179970	539832.0
100	modified - WF+ (Medford-Fellsay-Wellington)	0.0		0.0		0.0		0.0		0.0		0.0
100	modified - WF+ (Medford-Fellsay-Wellington)	0.0		0.0		0.0		0.0		0.0		0.0
105	modified - MF (Saugus-Malden)	0.0		0.0		0.0		0.0		0.0		0.0
105	modified - MF (Saugus-Malden)	0.0		0.0		0.0		0.0		0.0		0.0
106	modified - MF (Lebanon Loop-Wellington)	0.0		0.0		0.0		0.0		0.0		0.0
106	modified - MF (Lebanon Loop-Wellington)	0.0		0.0		0.0		0.0		0.0		0.0
108	modified - MF (Malden Center-Kennedy Drive)	0.0		0.0		0.0		0.0		0.0		0.0
108	modified - MF (Malden Center-Kennedy Drive)	0.0		0.0		0.0		0.0		0.0		0.0
11	modified - WF+ (South Boston-Broadway)	66908.0	79.3444489244124	3378.0	4.005881934397460	5118.0	6.069302468989400	8406.0	9.968455755045890	516.0	0.6119109171548520	84326.0
11	modified - WF+ (South Boston-Broadway)	66908.0	79.3444489244124	3378.0	4.005881934397460	5118.0	6.069302468989400	8406.0	9.968455755045890	516.0	0.6119109171548520	84326.0



#### 4.If there are service level disparities, are there differences in the characteristics of the people most impacted?

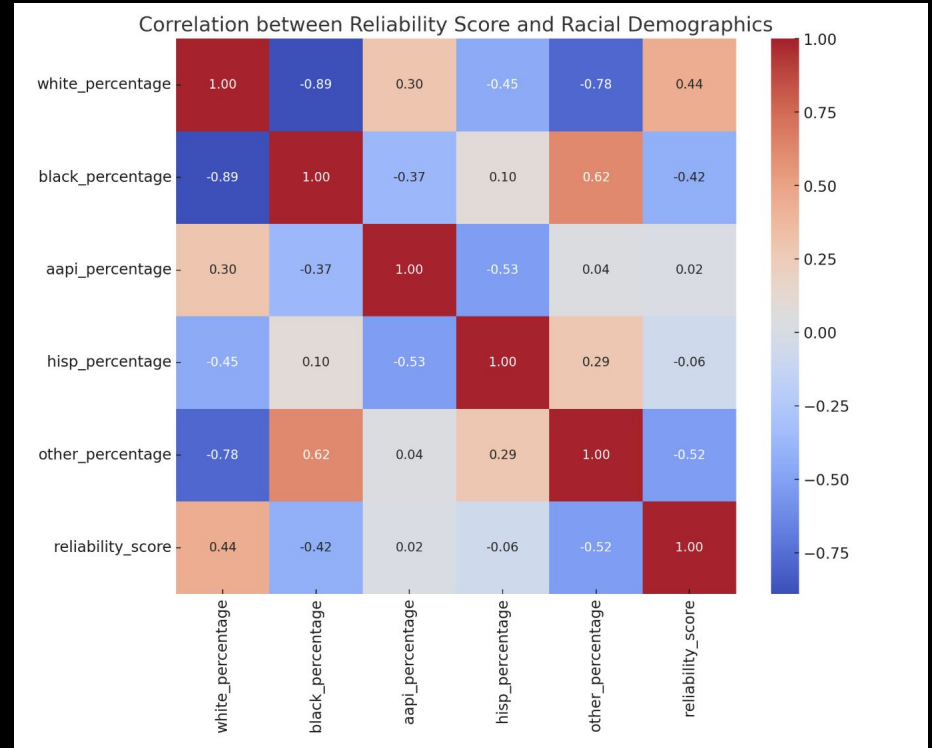
White Percentage: positive correlation (0.44)

Black Percentage: negative correlation (-0.42)

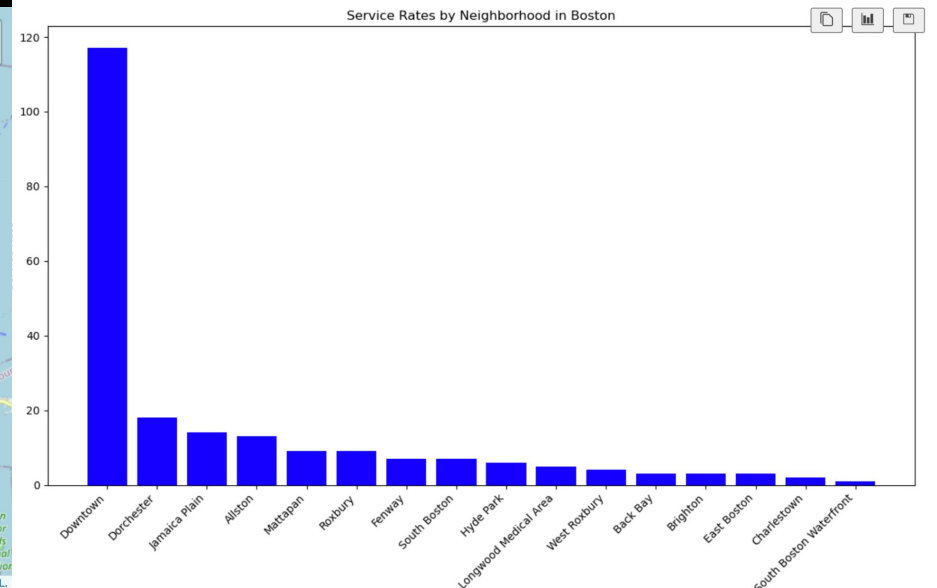
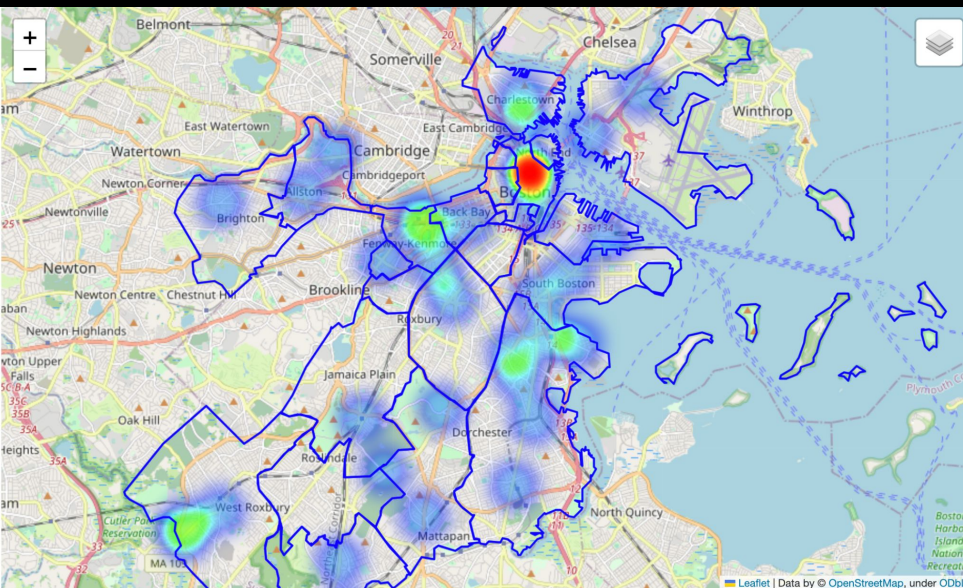
AAPI Percentage: The correlation here is negligible (0.02)

Hispanic Percentage: The correlation is slightly negative (-0.06))

Other Percentage: There is a negative correlation (-0.52)



## 5. Which neighborhoods are served better/worse by the MBTA bus system?





# Conclusion

1. Most of the end-to-end avg travel time is around 40 mins
2. There exists great disparity in the service level between different bus routes: e.g route 9703 has reliability score with 32.01% which is significantly lower than CR-Shuttle003 with a highest reliability of 92.59%
3. Most of the bus routes are located and serve for the neighborhoods with predominant racial group of White people
4. The correlation suggests white individuals may experience better service level than black individuals and the group of other races except while, black, hispanic and AAPI.
5. For service level, we can see the neighborhood with the highest service rate is Downtown area, followed by Dorchester, Jamaica Plain, Allston and Mattapan.



## Challenges & Limitations

1. The layer of neighborhood boundaries is only limited to Boston Area
2. Lack of detailed census data for now
3. The biggest challenge is matching data from different dataset; e.g some bus id of certain dataset are missing but it contains other useful information
4. Also, the exact locations for smaller bus stops (non-terminal) is lacking. For now, we're only able to use the data fetched from Google Maps API, which, for some of the bus terminals, is not accurate.

## Next Steps & Completion Plan

Continue working on more aspects for the extension project based on what we have achieved now

Search for more accurate and broader neighborhood boundaries and bus stop location data to refine our model. Modify Data Preprocessing Section to better handling missing data, to match the pattern across different datasets and eliminate the shortcoming of missing bus id.

Extension Proposal	
Extension Pitch	<p>We propose to extend our current geospatial analysis of Boston's transit systems by integrating additional datasets that consider factors like service disruptions and accessibility features. This effort aims to provide a more inclusive story of the city's transit dynamics, particularly focusing on how service changes and accessibility influence public transit usage and urban mobility. We also propose to delve deeper into the intersection of public transit data, specifically focusing on Blue Bike and bus data in Boston. The goal is to uncover insights into how different modes of transit interplay and affect urban mobility.</p>
Rationale	<p>Combining our existing geospatial data with information on service disruptions and accessibility will allow us to understand the resilience of the transit network. Understanding the dynamics between different transit systems is crucial for improving urban mobility. It is of particular interest to our team to explore how these modes of transit complement each other and what improvements can be made for better service integration.</p>
Questions for Analysis	<p>What are the accessibility gaps in the current transit network, and how might they affect riders with disabilities?</p> <p>What area does bus stops and blue bike stations covered, and how could that make blue bike a possible alternate for bus?</p> <p>How does the proximity of Blue Bike stations to bus stops affect ridership and transit efficiency?</p>
Data Sets & Sources	<p>Blue_Bike_Stations.geojson  Community_Centers.csv  PATI_Bus_Stops.csv  Hospitals.csv</p>
Additional Information	<p>Our preliminary analysis has shown interesting trends in transit usage across different Boston neighborhoods. This extended analysis will provide a more comprehensive view, potentially guiding city planners and policymakers in enhancing urban transit systems.</p>

