

DEVAL MEHTA



CONNECTING QUEENS

LEVERAGING DATA-DRIVEN INSIGHTS FOR BETTER TRANSIT PLANNING

January 24, 2025

TODAY'S AGENDA

- The “Transit Desert” Problem
- Key Takeaways
- Motivation
- Our Approach
- Findings
- Recommendations
- Future Work
- Conclusion
- Q & A

Queens Bus Map



THE TRANSIT DESERT PROBLEM

QUEENS LACKS EQUITABLE PUBLIC TRANSIT ACCESS

STEP ONE

Identify Underserved
Areas

STEP TWO

Analyze transit service
gaps.

STEP THREE

Provide actionable
insights to improve transit
equity

KEY TAKEAWAYS

UNDERSERVED

Most Eastern Queens Neighborhoods are underserviced*

CONSISTENT

Bus activity throughout the day is consistent across routes, communities, and neighborhoods

SUB-OPTIMAL

Network connectivity levels suggest space for improvement



WHY SHOULD WE CARE?

CONGESTION PRICING

Queens residents have no way into Manhattan or out of the state without paying a toll

QUALITY OF LIFE

Longer commute times contribute to less productivity and can affect mental health

ENVIRONMENTAL IMPACT

Fewer cars on the road means fewer carbon emissions



METHODOLOGY

DATA COLLECTION

Static and Real-Time Data over one week from the MTA BusTime API

EXPLORATION

Trends and Seasonality in the Data

Geographic Distribution

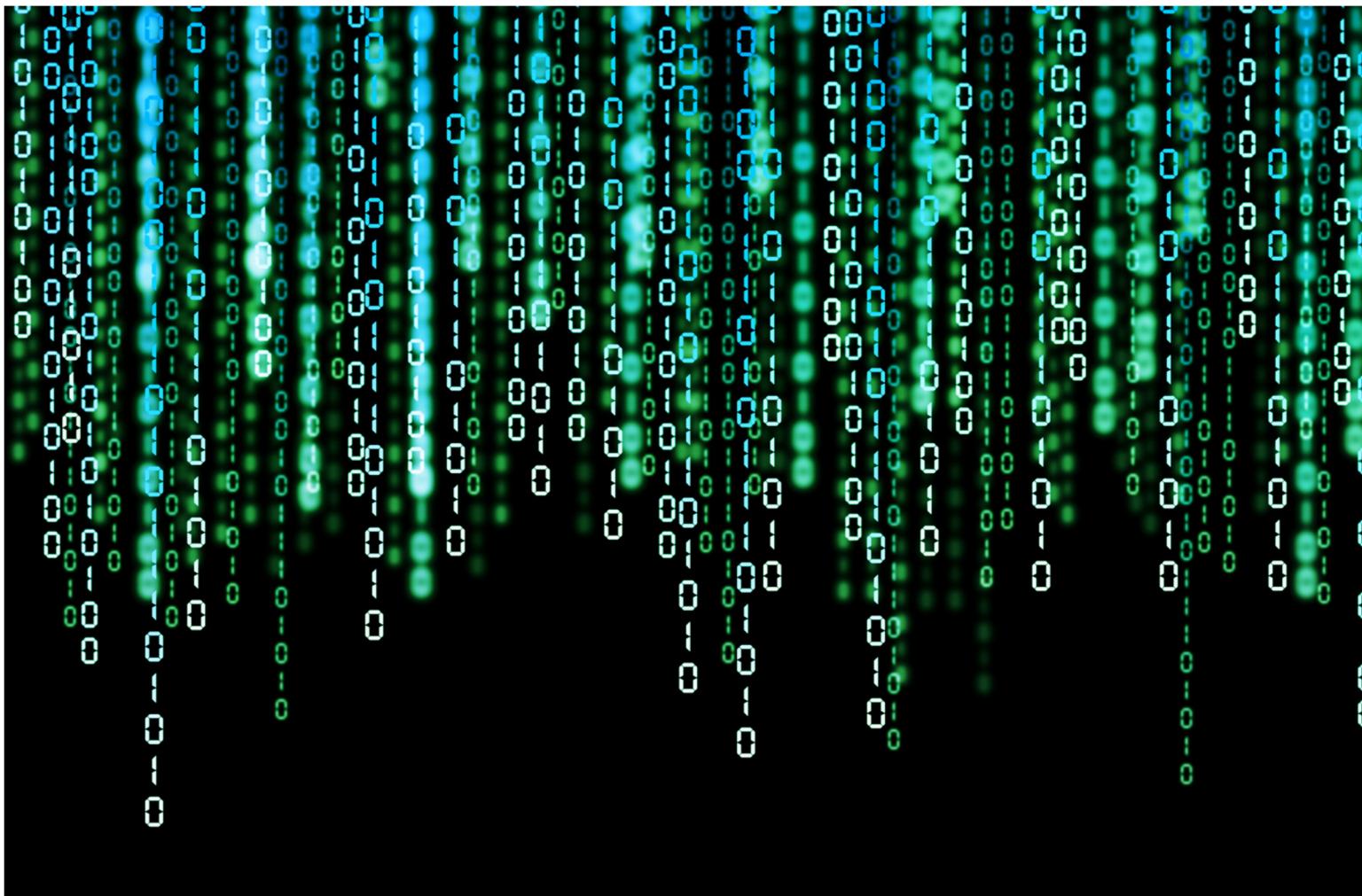
Measure Activity by Neighborhood

MACHINE LEARNING

Time-Series HDBSCAN Clustering

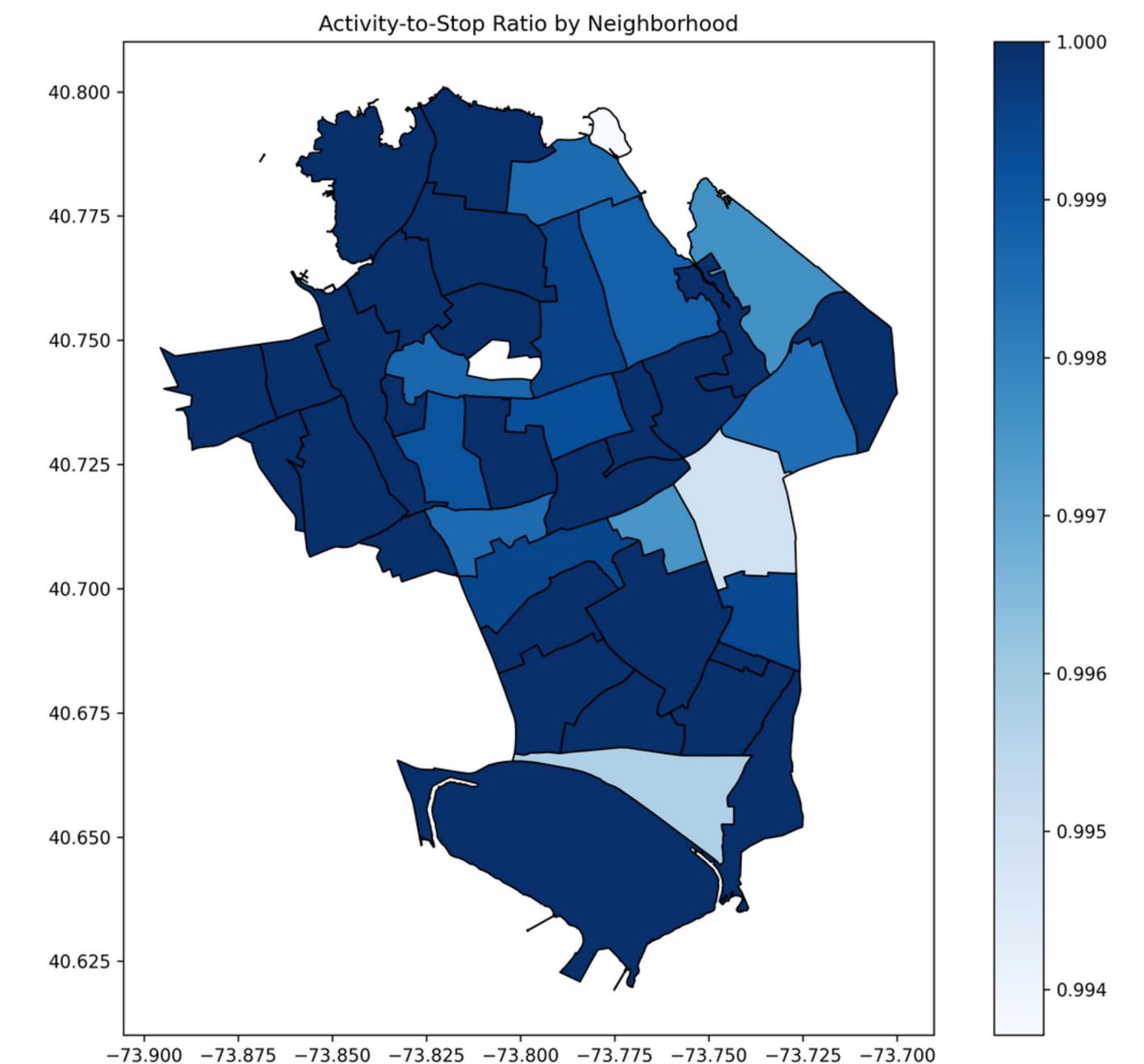
NETWORK ANALYSIS

Louvain Method

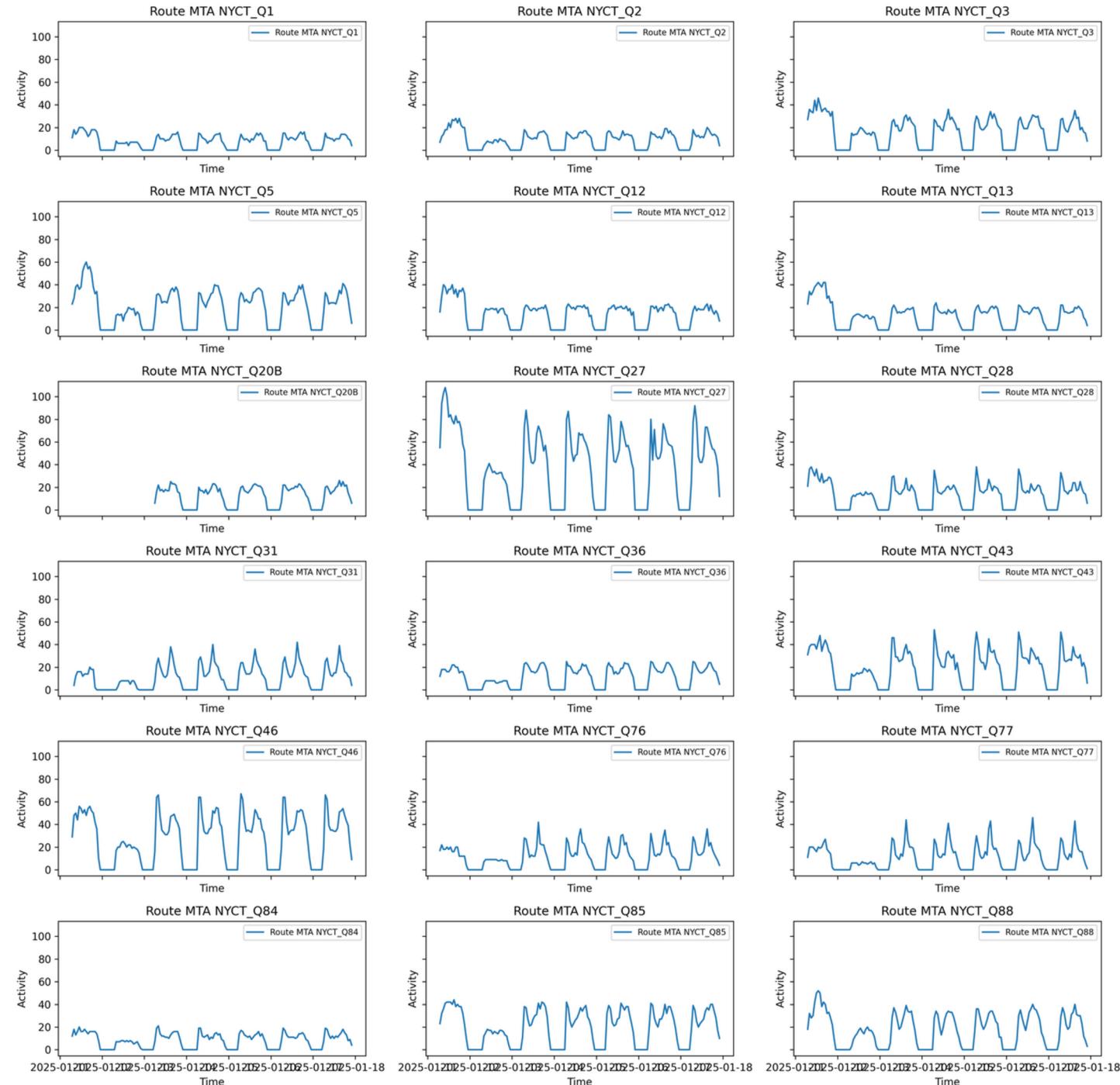


TRANSIT EQUITY OVERVIEW

Neighborhood	Stop Count	Bus Activity Count	Activity-to-Stop Ratio
Fort Totten	159	158	0.993711
Queens Village	3544	3526	0.994921
Springfield Gardens (South)-Brookville	469	467	0.995736
Hollis	1215	1212	0.997531
Douglaston-Little Neck	847	845	0.997639



BUS ACTIVITY TRENDS



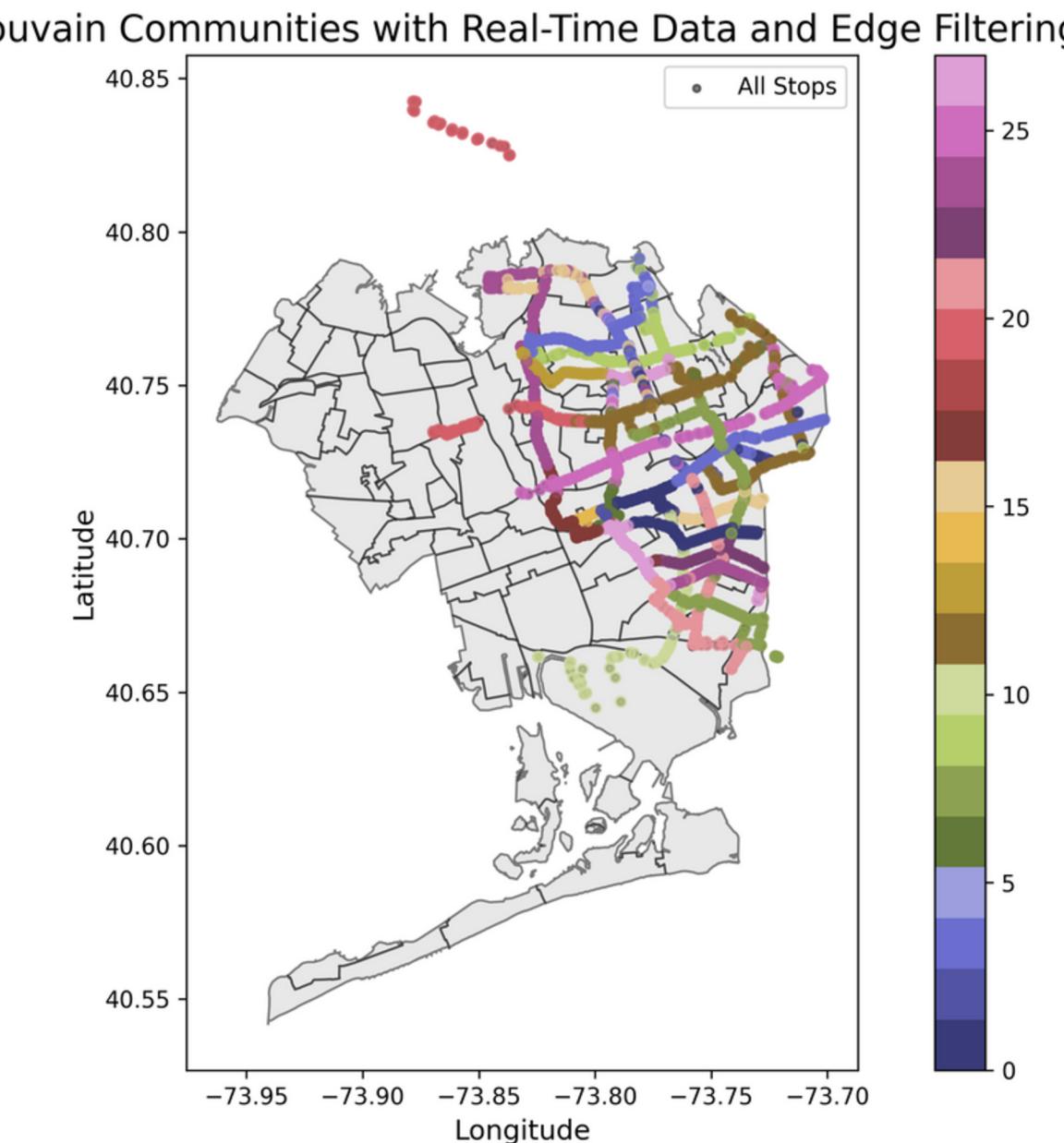
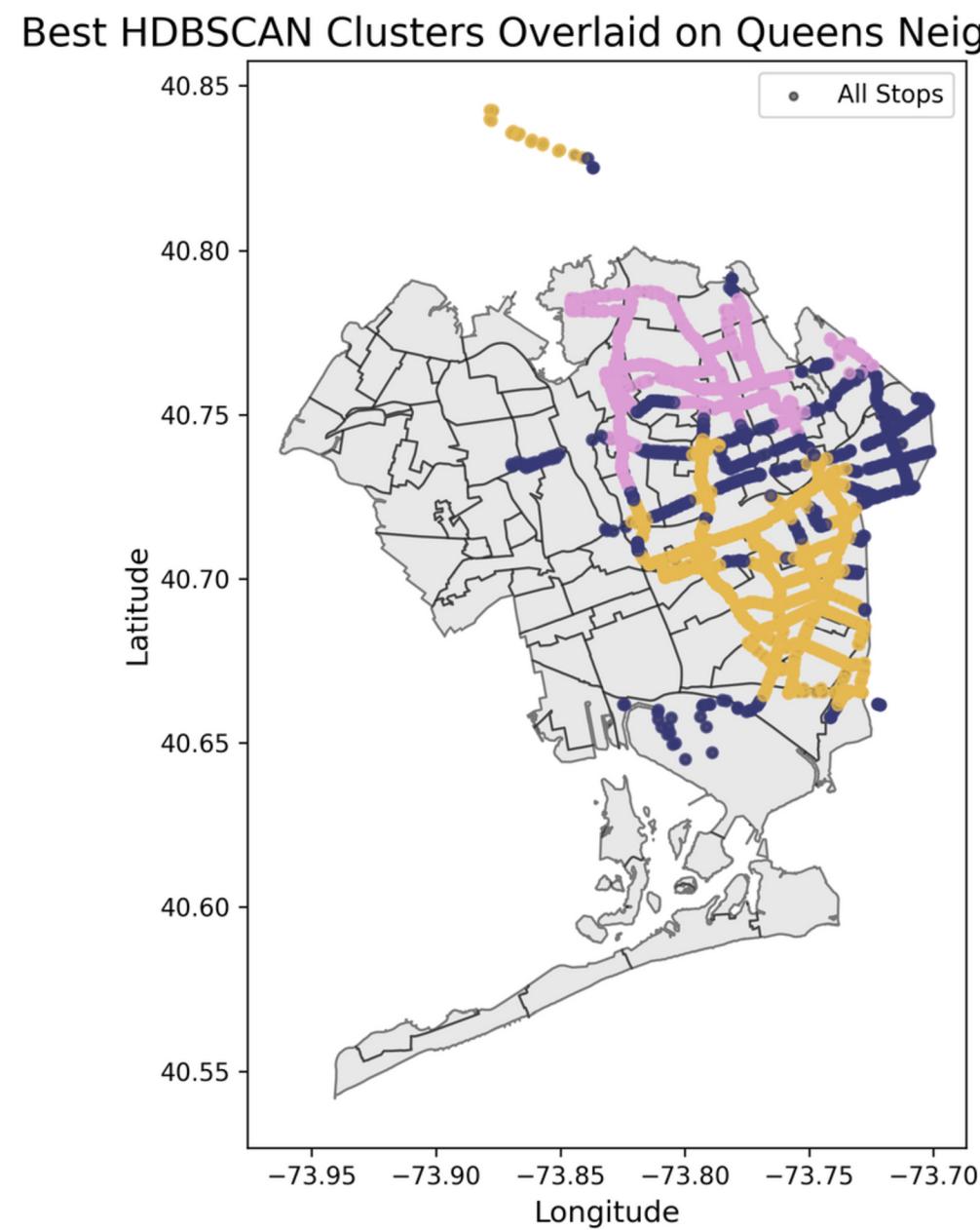
**Root Mean
Squared Error**

1.67

Root Mean Squared Error:
How “far away” is the prediction on
average from the actual data?

- Closer to 0 is better
- Used here to measure day-to-day consistency

THE TRANSIT NETWORK

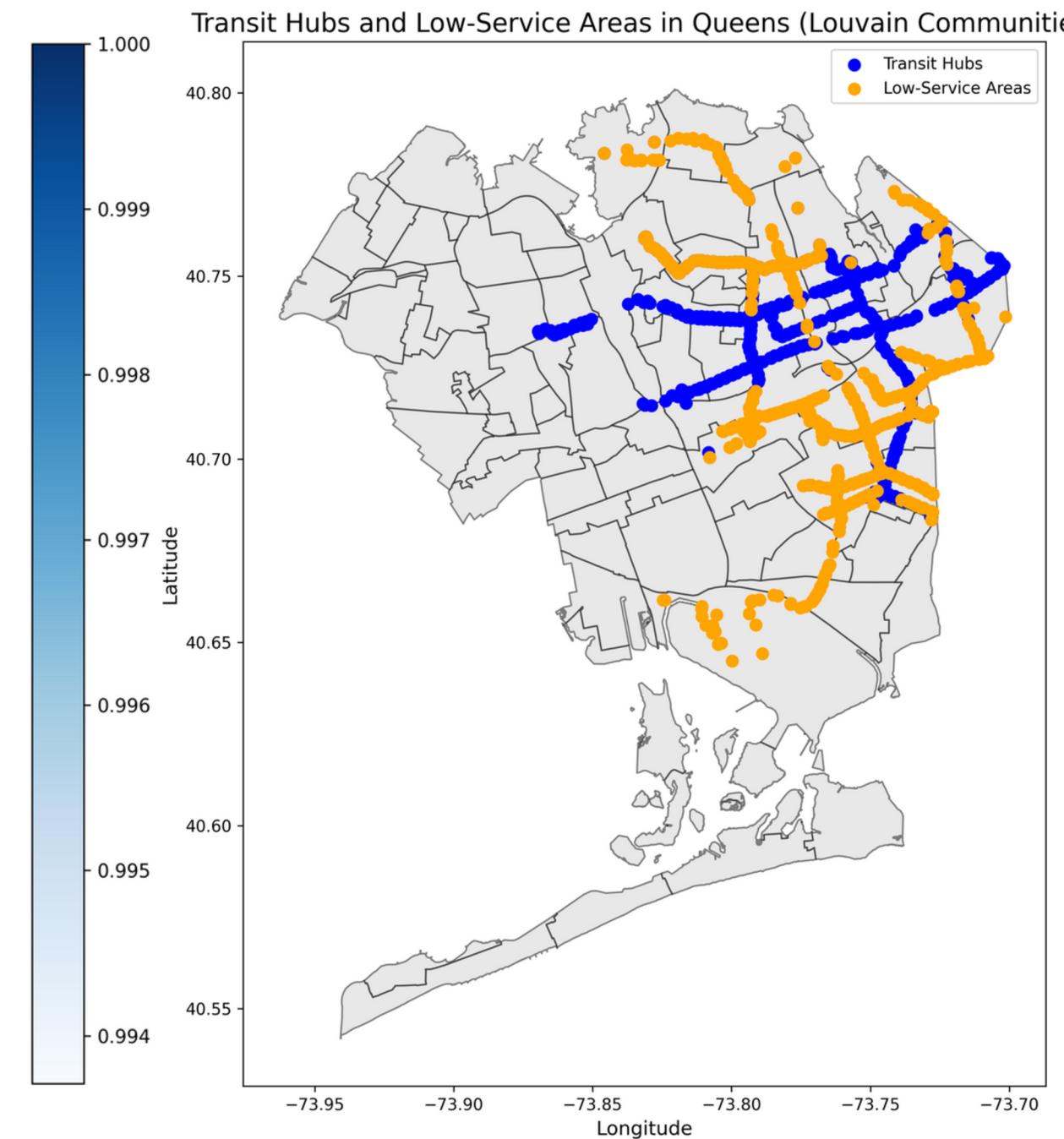
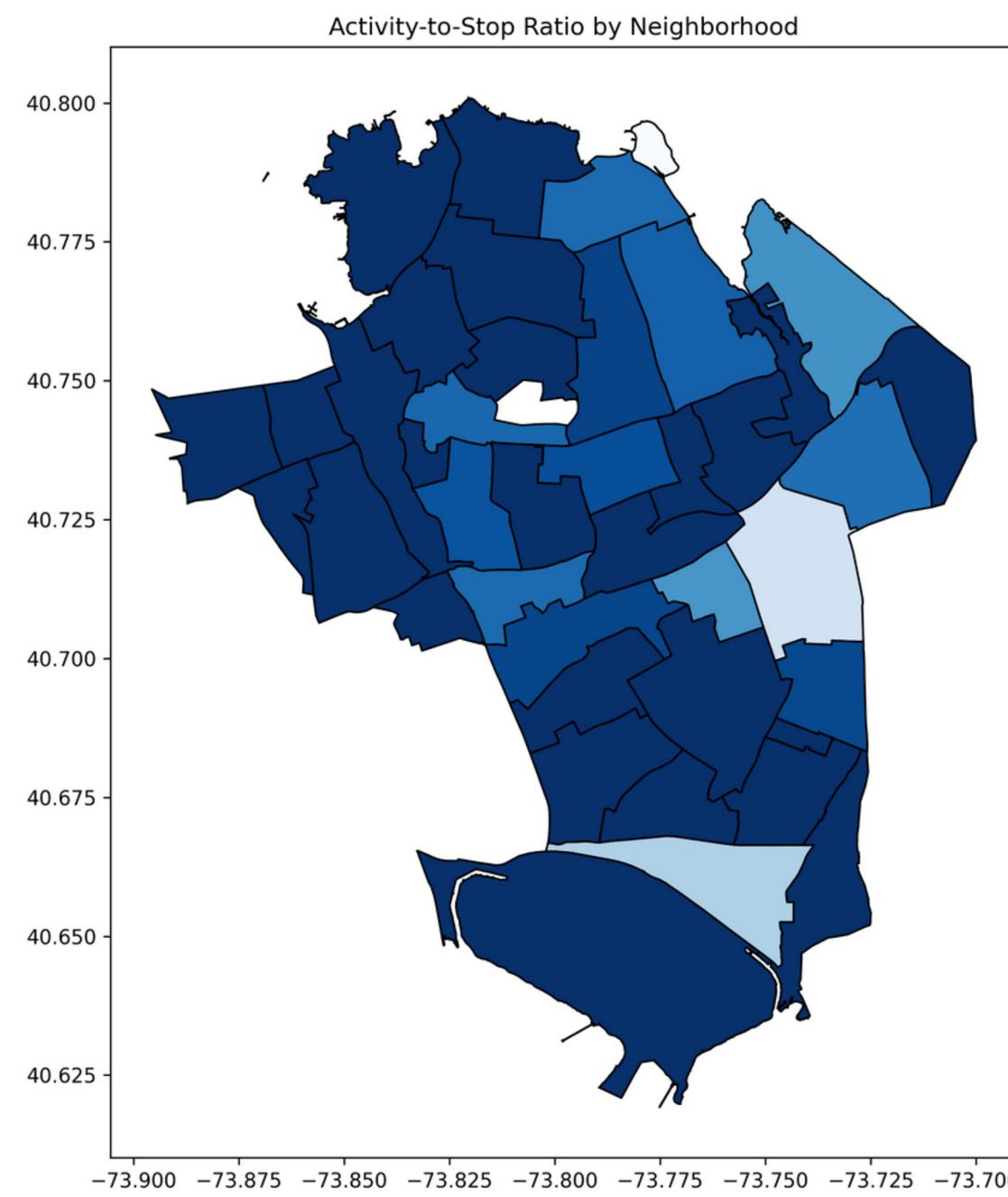


HDBSCAN Silhouette Score	0.71
Louvain Modularity Score	0.76

Silhouette and Modularity:
How clearly defined are our
clusters?

- Are the groups closely packed?
- How well separated are the groups?
- -1 = no identifiable clustering
- +1 = as well-defined as possible

TRANSIT NETWORK



HDBSCAN Silhouette Score	0.71
Louvain Modularity Score	0.76

Silhouette and Modularity:
How clearly defined are our clusters?

- Are the groups closely packed?
- How well separated are the groups?
- -1 = no identifiable clustering
- +1 = as well-defined as possible

RECOMMENDATIONS

NOW

Increase bus frequency to underserved neighborhoods

COMING MONTHS

Explore options to improve connectivity

BEYOND

Consider subway and express bus expansions

FUTURE EXTENSIONS

WHAT WE DO NEXT



BIG DATA

Collect data more frequently and continuously



NEW ROUTES

Implement pathfinding algorithms to improve network connectivity



POPULATION

Incorporate population-based metrics to ensure maximum utility for the most people



FULL SWING

Collect data from across the city to validate our work.

CLOSING THE GAP

IMPROVE SERVICE

Transit deserts don't have to remain dry.

CONNECT

Reducing private traffic requires improving public access

INVEST

Keep our city connected and whole



THANK YOU!

CREDITS TO

Matt Brems and Asha Mathis for
consistent feedback

Rowan Schaefer and Argishti (Argo)
Ovsepyan for support throughout
DSB

ChatGPT for aid with
troubleshooting and consolidating
new learning

My family and friends for their
empathy and encouragement

