2025.12.02.

**Temporal Community Dynamics in Urban Mobility: A Network Analysis of NYC Citi Bike**

Github Repository Link:

https://github.com/adamburkus64/ANS---project

Applied Network Science  
Group Project

Authors:

Ujvári Botond Rudolf

Garas Fábián Máté

Szabó Levente

Burkus Ádám

Nagy Zalán Zsolt

**Contents**

[Introduction 3](#_Toc215589723)

[Literature Review 4](#_Toc215589724)

[Data 6](#_Toc215589725)

[Methodology 6](#_Toc215589726)

[Results and findings 8](#_Toc215589727)

[Bibliography 11](#_Toc215589728)

[Appendix 11](#_Toc215589729)

# Introduction

Community detection has long been a significant area of research in modern network science in order to broaden the understanding of complex systems, via visualizing and interpreting its community structure or clustering, according to Fortunato (2010). Communities, by definition are groups of densely connected nodes, forming subgraphs, which are generally less connected to the other parts of specific network. They are not necessarily considered as components due to this lack of global connectivity. Community detection aims to partition the network into these internally cohesive and externally sparse subgraphs, revealing hidden structural or functional organization.

Urban mobility analysis is the study of how people move within cities using data-driven, spatial, and temporal methods. It is often used to examine the patterns, structures, and possible determinants of individual movement. These systems, combined with the development of technology, produces massive volumes of high-resolution data – information for tracing location – that offer unprecedented opportunities to investigate how people navigate areas and reveal the above-mentioned trends through space and time.

Among these sources, New York City’s Citi Bike service has emerged as a valuable proxy for urban mobility, providing millions of documented (time intervals, origin and destination coordinates) bicycle trips each month. New York City is the largest city by population of the United States with – based on most recent estimations – an approximately 8.5 million people, and one of most densely populated cities in the world with its 29 thousand people per square mile. It has also one of the most complex transportation systems in the world, which could all contribute to a highly interconnected and rich network of data points.

There are many possible approaches for such research, many focusing on either spatial or temporal differences of urban movements, using various data and methodologies. As for the purpose of our scrutiny, we are determined to compose our models around the latter time dimension: **investigating whether community structures and mobility patterns are different for weekdays and weekends, additionally throughout specific periods of the day, namely morning and evening.** Studying these temporal dynamics, they might reveal how communities in general emerge, dissolve and evolve throughout the week, and throughout a day. Additionally, as part of this observation, we are using two of the most popular community detection methods, Louvain and Infomap frameworks, on one hand facilitating ourselves to capture different aspects of the structure, on the other hand, to deepen our understanding of the strengths and limitations of each method throughout a systematic comparison.

# Literature Review

During the previous years, much research has been carried out and socioeconomic papers were published analysing urban mobility and spatio-temporal differences, using public transport, mobile GPS or other data sources, contributing to the optimization of urban planning and transportation engineering. As for setting the basis of our examination, we were looking for such literature, in order to better understand the general commuting behaviour of individuals in big cities and be able formulate our expectations accordingly.

Kinoshita et al. (2024) studied exactly those spatio-temporal patterns in bike sharing system usage across six major cities including New York (Kinoshita, Bando, & Sayama, 2024). Using 30 days of data, they quantify weekday-weekend variability using Jensen-Shannon divergence and show strong, universal temporal rhythms: sharp morning and evening commute peaks on weekdays, contrasted with smoother, midday-oriented patterns on weekends, with Friday acting as a transitional hybrid. They further show that station usage rankings are surprisingly stable between weekdays and weekends in most cities, indicating persistent structural demand patterns despite temporal variation. Their approach, clustering days based on similarity, demonstrates that bike-sharing data naturally form meaningful temporal communities, even without explicit network-based methods.

We also identified a paper that aligns closely with the core questions of our research. The most recent case study of Wen et al. (2025) provides a detailed examination of mobility patterns in Manhattan, New York City’s Citi Bike system, focusing specifically on how travel flows and community structures vary across multiple time periods within August. Using complex network theory, the authors construct origin-destination flow networks and apply the Louvain algorithm to identify clusters of strongly interconnected neighbourhoods. Their analysis reveals clear temporal differences: early-, mid-, and late-August show shifting hotspots concentrated around Lower- and Midtown Manhattan, with variations driven by seasonal factors, tourism intensity, land-use patterns, and the transition from summer to regular commuting routines. The study finds that modularity, node strength, and origin-destination flow distributions evolve across these periods, indicating changes in both commuting and tourism-driven travel demand.

In order to fully understand the tools of both the Louvain and Infomap method, we looked at vast amount of literature to be able to use them efficiently. The Louvain method is a community detection algorithm designed to uncover densely connected groups within large-scale networks. Its purpose is to efficiently partition nodes by iteratively performing two steps: a local modularity optimization phase, where each node is reassigned to the neighbouring community that yields the greatest positive modularity gain, and a coarsening phase, where these discovered communities are aggregated into one node to form a reduced network, after which the process repeats until no further improvement is possible. This enables the algorithm to scale nearly linearly while still identifying meaningful structural modules in massive datasets (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008). Because Citi Bike networks consist of station-to-station flows whose structure shifts across hours, days and seasons, the Louvain method is particularly valuable for detecting cohesive mobility regions at each time slice and for comparing how these clusters evolve. In our research, using Louvain on NYC Citi Bike networks allows us to identify mobility communities and quantify how commuting reshapes the city’s functional geography over time.

Infomap takes a different approach by detecting communities based on flows rather than modularity: it models a random walk on the network and finds the partition that most efficiently compresses this movement using the map equation. Instead of looking only at dense subgraphs, it identifies groups where flows circulate for long periods, which is crucial for directed, weighted mobility systems like bike trips. Algorithmically it uses a multi-phase search to minimize the flow description length and uncover multi-level structures (Jelena Smiljanić, 2023). In our NYC Citi Bike research, this makes Infomap valuable for capturing directional travel patterns (e.g., morning vs. evening flows, as seen later on [Figure 1](#Figure_1)) and for revealing how mobility evolves across time slices, giving richer insight into temporal community dynamics than structure-only methods like Louvain.

While the Louvain method detects communities by maximizing modularity and therefore emphasizes structural density, Infomap identifies communities by minimizing the map equation, capturing how flows realistically move through the network. Louvain is efficient and highlights broad structural clusters but cannot incorporate directionality, while Infomap is more sensitive to directed and weighted flows, making it better suited for uncovering functional mobility regions shaped by commuter patterns and directional bike traffic. In the context of NYC Citi Bike, Louvain helps us reveal stable structural station groupings, whereas Infomap uncovers the temporal and directional dynamics behind movement patterns. Using both methods allows us to obtain more thorough, complementary findings.

# Data

The data we have used throughout the analysis, as above mentioned was acquired from the official website of CitiBike, where all the documented rides are publicly available in monthly frequency. Considering the purpose of the study, we have decided to use the latest available dataset of 2025 October, including around 4.7 million rides across New York City.

Every ride is distinguished with a unique ID, furthermore the dataset provides information on the membership status of the user (member of casual rider), bike type (electric or classic), exact date and time, and exact coordinates of origins and destinations of renting (including station names). The geographic coordinates are provided in the EPSG:4326 reference system, which corresponds to standard GPS coordinates and allows for direct spatial matching with official administrative boundary data.

# Methodology

Given the extremely large number of stations and the resulting complexity of a raw station-level network, the analysis was conducted at an aggregated spatial scale using New York City Community Districts. These districts are administrative sub-units of the five boroughs of New York City: Manhattan, Brooklyn, Queens, the Bronx, and Staten Island. Using the geographic coordinates of each station, a spatial join was performed to assign every start and end station to its corresponding community district. This step converted every individual trip into a district-to-district movement, thereby enabling the construction of manageable and interpretable spatial mobility networks. Stations with missing geographic coordinates were removed prior to the spatial matching to avoid misclassification.

To capture temporal heterogeneity in urban mobility, the dataset was further divided into multiple time-based subsets. Two main temporal dimensions were considered. First, trips were classified by time of day into morning (06:00-10:59), and evening periods (18:00-22:59) based on the ending time of each trip. Second, trips were categorized by day type into weekday (Monday–Friday) and weekend (Saturday–Sunday) observations. This classification resulted in separate datasets for morning, evening, weekday, and weekend traffic, allowing the network structure to be examined under distinct commuting and leisure-related mobility regimes.

For each temporal subset, a district-level flow network was constructed by aggregating the number of trips between all ordered pairs of community districts. The resulting data takes the form of weighted edge lists, where each vertex represents a community district and each edge represents the total number of bike trips between two districts within a given temporal subset. The edge weights therefore capture the intensity of mobility flows across the city. Two types of graph representations were created from each edge list. We constructed weighted and directed networks based on temporal differences, scrutinized the key centrality measures (in- and out-degree, PageRank, betweenness, closeness). Additionally, examined those on network level: checked reciprocity, density, assortativity, ratio of self-loops and other descriptive metrics ([Table 1](#Table_1)). We found that the detected communities are overlapping the borough structure in New York City (see Appendix [Figure 2-5](#Figure_2)).

*A screenshot of a graph

AI-generated content may be incorrect.  
Table 1 Comparison of networks: correlation between centrality measures and other descriptive statistics*

As for the comparison on basis of the above-mentioned aspects, we found that as far as weekday-weekend bike routes are concerned, those are pretty much consistent based on the correlation of the centrality measures. Based on network level findings, there are more edges during the week, than on the weekends (more and different routes are taken), however there is no directional bias. We can observe very similar assortativity (close to 0) and loop-ratios (0.38). Additionally, while on average cross-district rides are 3-times more frequent during weekdays than during the weekend (meaning more per-day average rides during weekdays), within-district rides do not show difference.

Comparing morning and evening rides, those show some variance of the mobility structure during the day, while based on betweenness and in/out-degree correlations, paths are rather consistent. On network level, we find very similar statistics. However, both inter- and intra-district rides are more frequent during the evening hours.

Infomap was applied separately to the morning, evening, weekday, and weekend networks. (see Appendix [Figure 2-5](#Figure_2)) After community detection, the membership vector for each community district was extracted. To ensure comparability across different temporal subsets, a canonical ordering of districts was maintained so that each district was consistently matched with itself in all comparisons. The similarity between the Infomap partitions obtained from different temporal subsets was evaluated using three complementary measures: the Adjusted Rand Index (ARI), which measures agreement between clusterings, Normalized Mutual Information (NMI), which captures the amount of shared information between two partitions and Variation of Information (VI), which measures dissimilarity between partitions. ARI and NMI take values between 0 and 1, where higher values indicate stronger similarity between community structures, while VI also ranges between 0 and 1, with lower values indicating greater similarity. These metrics were used to quantify how functional mobility communities differ between morning and evening periods as well as between weekdays and weekends.

To assess the robustness of the Infomap-based results, the Louvain community detection method was applied as a secondary, validation-oriented approach. For this purpose, the same district-level flow networks were converted into weighted, undirected graphs, as required by the Louvain algorithm. The resulting Louvain membership vectors were extracted using the same canonical node ordering as for Infomap and were compared across temporal subsets using the same ARI, NMI, and VI similarity measures. In addition, modularity scores were computed for all Louvain partitions to verify that the detected communities exhibit a meaningful modular structure.

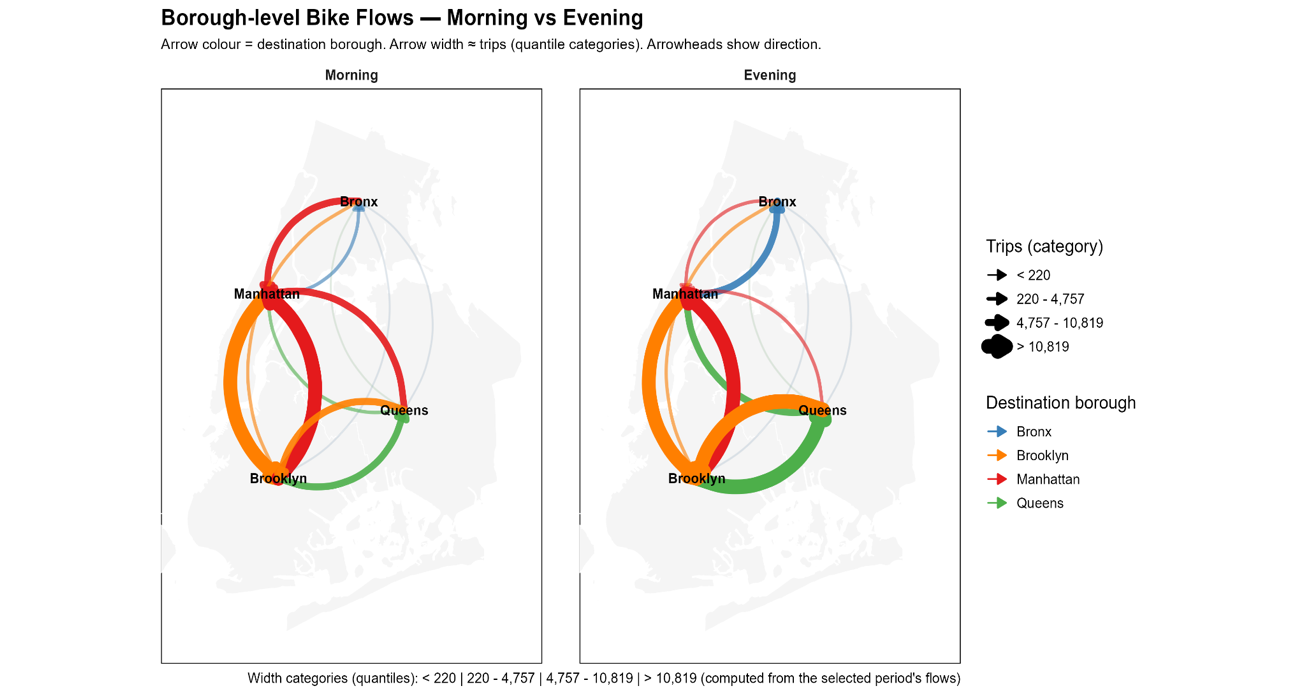
As above indicated, Louvain method is not used as the primary interpretative framework in this study, as it ignores the directionality that is intrinsic to commuting and leisure-related mobility. Instead, it serves as a robustness check to confirm that the main spatial patterns uncovered by Infomap are not artifacts of a single community detection technique. The core interpretation of community structure, temporal variation, and functional mobility regions throughout the paper is therefore based on the Infomap results, which are better suited to capturing the directional and weighted nature of bike-sharing flows in an urban environment.

# Results and findings

The initial examination of the Citi Bike network topology reveals that the system functions as a spatial interaction network heavily influenced by geography. A strong negative Spearman correlation of -0.759 was observed between spatial distance and flow weight, confirming that as the distance between districts increases, the flow of commuters decreases significantly. Further testing of the degree distributions determined that the network is not scale-free, as the power-law fit produced an exceedingly high alpha parameter indicating a poor fit. Instead, the degree distribution follows a Log-Normal distribution, which is standard for spatial transportation and flow networks. Additionally, a Small-World test comparing clustering coefficients and path lengths to a random graph indicated that none of the temporal networks exhibit small-world properties.

When comparing the static structure of the network across different time periods, the underlying grid remains remarkably consistent. The correlation between centrality measures for weekdays and weekends is extremely high, with PageRank at 0.98, and both in-degree and out-degree correlations at 0.97. This suggests that the relative importance of specific stations and the primary routes riders utilize do not change significantly between working days and leisure days. However, a divergence appears when comparing morning and evening periods. While the out-degree correlation remains high at 0.98, implying that trips originate from the same residential hubs regardless of the time, the PageRank correlation drops to 0.58. This structural variance indicates that the hierarchy of destination nodes shifts significantly between the morning commute and evening return trips, as the "attractor" districts change based on the time of day.

Despite the consistency in the physical network structure, the intensity and functional organization of flows differ substantially across time. Weekday networks exhibit a higher density of 0.791 compared to 0.739 for weekends, and the volume of cross-district rides is approximately three times higher during the week (which shows an increased per-day average number of rides). However, the ratio of self-loops remains stable at approximately 38 percent for both periods. The network also shows high reciprocity of roughly 0.94 and near-zero assortativity, indicating no directional bias or preference for high-degree nodes to connect exclusively with each other.

The application of the Infomap algorithm, which detects communities based on directed flow, revealed distinct functional changes. On weekdays, the algorithm identified four communities (see [Figure 2](#Figure_2)), highlighted by a strong integration where the North Manhattan districts merge with the Bronx, likely driven by work-related travel flows. In contrast, the weekend network fragments differently (see [Figure 3](#Figure_3)); almost the entire borough of Manhattan forms a single community distinct from the Bronx, reflecting a leisure-oriented pattern where riders stay within the borough rather than commuting across boundaries.

*Figure 1: Morning and evening flows between boroughs*

Similar distinctions arise between morning and evening flows. As visualized in [Figure 1](#Figure_1), the morning network is dominated by a pronounced convergence, with heavy flows moving from outer boroughs onto Manhattan. In contrast, the evening panel illustrates a broad redirection outward, distributing riders back toward residential districts. This general shift accompanies a fragmentation from four communities to five (see Appendix [Figure 4-5](#Figure_4)), reflecting a transition from integrated city-wide commuting to localized "return home" circulation.

To validate these findings, the Louvain method was applied to undirected versions of the networks as a robustness check. The modularity scores were consistently positive at approximately 0.51, confirming that the community districts form a meaningful modular structure. The Louvain results largely reinforce the geography-based modularity of New York, with communities heavily overlapping with administrative boroughs. A specific, recurring anomaly was identified in Central Brooklyn, where districts 301, 304, and 405 consistently form a distinct cluster separate from the rest of the borough.

The comparison between the two methods highlights the necessity of using flow-based detection for mobility data. When comparing morning and evening partitions using Louvain, the Adjusted Rand Index was nearly perfect at 0.97, whereas Infomap showed variation with an index of 0.83. This discrepancy confirms that while the static infrastructure captured by Louvain does not change, the directional flow dynamics captured by Infomap shift significantly. Thus, the Infomap results provide a more accurate representation of the functional temporal changes in urban mobility, distinguishing between the integrated flows of morning commutes and the fragmented, localized patterns of evening leisure.

# Bibliography

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*. doi:10.1088/1742-5468/2008/10/P10008

Fortunato, S. (2010). Community detection in graphs. *Physics reports*.

Jelena Smiljanić, C. B. (2023). Community Detection with the Map Equation and Infomap: Theory and Applications.

Kinoshita, S.-i., Bando, Y., & Sayama, H. (2024). Spatio-Temporal Differences in Bike Sharing Usage: A Tale of Six Cities.

Wen, Z., Tian, D., & Wu, N. (2025). Modeling and Analyzing the Spatiotemporal Travel Patterns of Bike Sharing: A Case Study of Citi Bike in New York. *Sustainability*. doi:https://doi.org/10.3390/su17010014

# Appendix

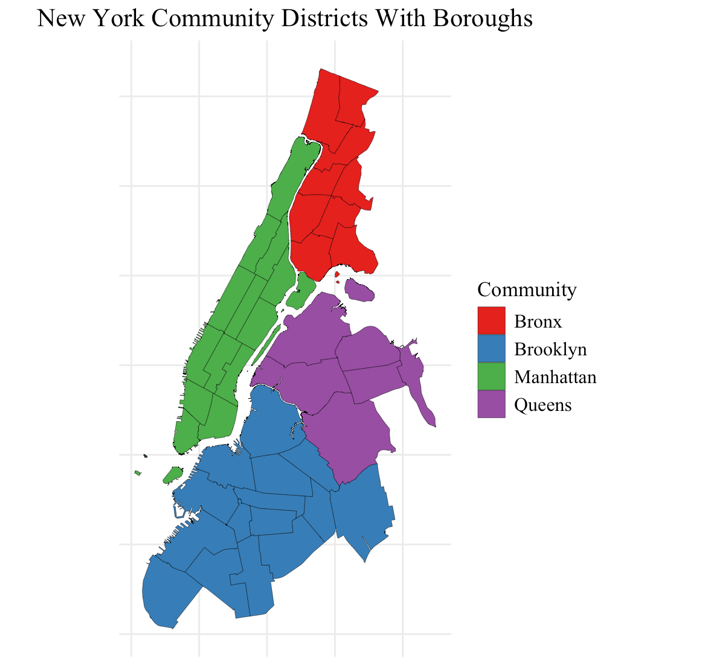
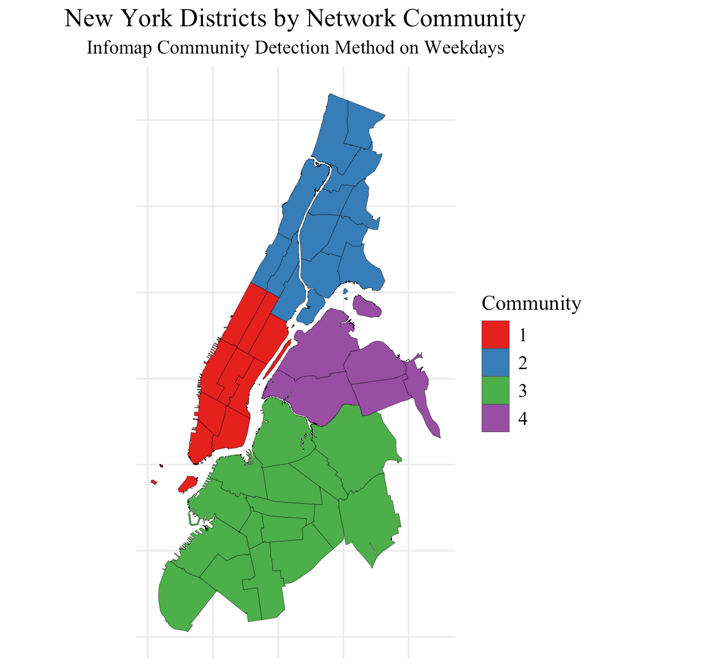


Figure 2: Community distribution of weekday commutes and the original distribution of community districts under the boroughs for comparison

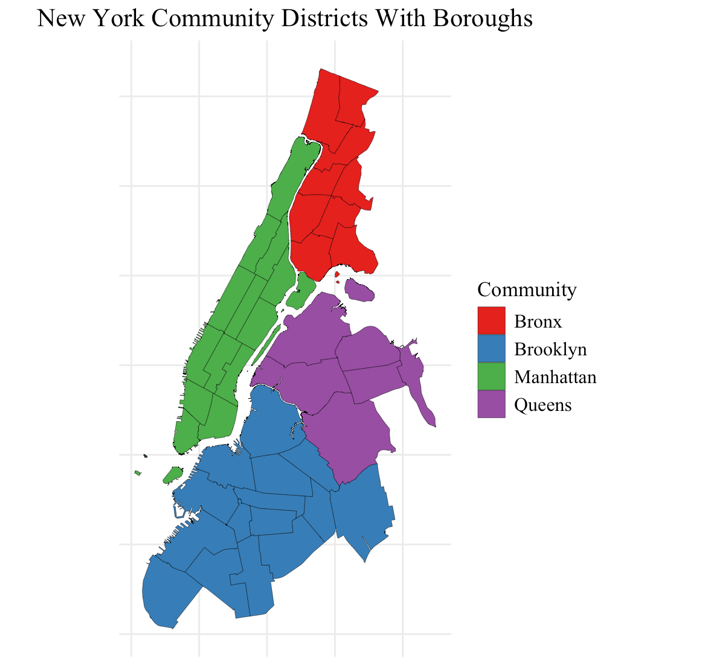
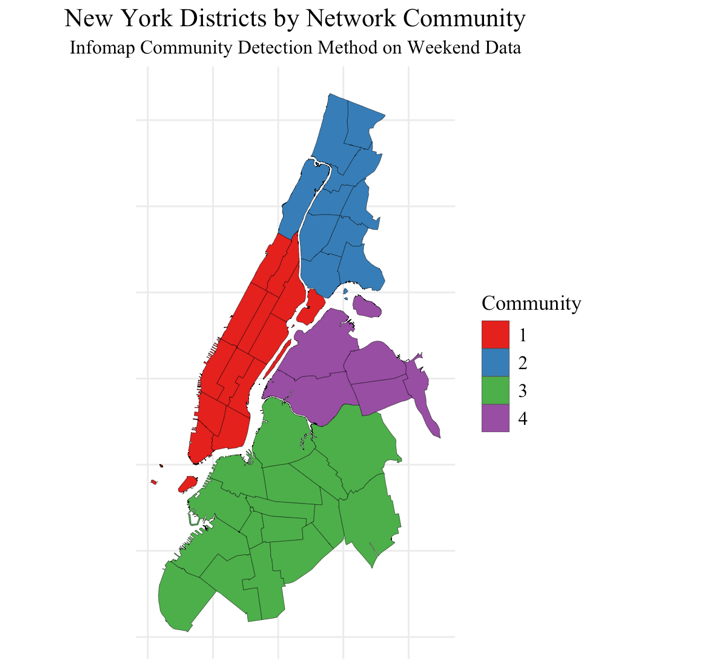


Figure 3: Community distribution of weekend commutes and the original distribution of community districts under the boroughs for comparison

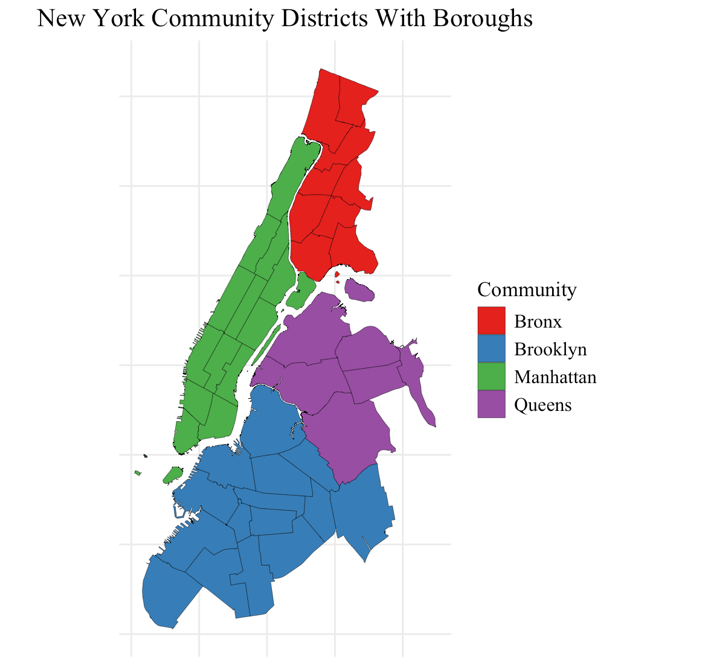


Figure 4: Community distribution of morning commutes and the original distribution of community districts under the boroughs for comparison

Figure 5: Community distribution of morning commutes and the original distribution of community districts under the boroughs for comparison

