BSSAnalyzer: a framework for analyzing and comparing bike-sharing systems

Gabriel Campanelli Iamato[1], Higor Amario de Souza[2], Fabio Kon[3]

[1] Institute of Mathematics and Computer Sciences
University of São Paulo – São Carlos, SP, Brazil

[2]Department of Computer and Digital Systems Engineering, Polytechnic School
University of São Paulo – São Paulo, SP, Brazil

[3]Department of Computer Science, Institute of Mathematics, Statistics and
Computer Science – University of São Paulo – São Paulo, SP, Brazil

{gabriel.c.iamato, higoramario}@usp.br, kon@ime.usp.br

Abstract. *Bike-sharing systems (BSS) are a valuable public infrastructure for sustainable mobility. Understanding how BSS is used may provide insights for urban planning and for the expansion of these systems. This work proposes BSSAnalyzer, an open-source object-oriented framework that integrates and analyzes data from cycling infrastructure, public transportation, points of interest (POI), and other urban features to evaluate BSS performance. We applied the BSSAnalyzer in cities of London, New York, and São Paulo. The results indicate strong correlations between BSS usage and POIs related to food, services, and commercial establishments. BSSAnalyzer can also highlight areas of higher public interest, key urban connectivity infrastructure, connectivity between neighboring BSS stations, inactive stations, and cyclist flow patterns.*

1. Introduction

The use of bicycles as a means of transportation has grown significantly in recent decades in urban centers around the world, driven by public policies focused on sustainable mobility and mitigating environmental impacts. The benefits include improvements in the population's quality of life, and reductions of traffic congestion, air and noise pollution, and greenhouse gas emissions [Watts 2018, Duran-Rodas et al. 2019]. Thus, the bicycle is positioned as a strategic element in smart city policies and data-driven urban planning.

In addition to the private bicycle fleet, bike-sharing systems (BSS) stand out, present in hundreds of cities worldwide [Duran-Rodas et al. 2019]. These systems operate through public-private partnerships and allow temporary use of bicycles distributed across geographically organized stations. Regarding mobility, BSS play a strategic role in modal integration, enabling first- and last-mile trips, connecting users to public transport [Adnan et al. 2019]. Thus, BSS contribute to sustainable urban mobility ecosystems.

Whether private bicycles or BSS, the individual impact of cycling leads to a more qualified experience of the city and to the appropriation of public space. For instance, in São Paulo, [CEBRAP 2018] reports that there would be a significant reduction in expenses if the use of this mode of transportation were expanded, with a greater impact on lower-income classes, about 14% of the monthly income (US$ 40 on average), along with an increase in municipal GDP of approximately US$ 164 million. This survey also reports that cyclists are responsible for a 3% reduction in CO^2 emissions, which could increase to 18% if the cycling potential were reached, representing only part of the city's trips.

Regarding BSS, [Duran-Rodas et al. 2019] indicate factors that influence the popularity of docking stations, such as transportation infrastructure, distance to downtown, presence of leisure establishments, and places of public interest. These characteristics reveal macrogeographic features that serve as the basis for analysis. However, other factors may influence the BSS usage, such as user profile, strategies of the bike-sharing company, advertising, and local culture [Duarte 2016].

The strategic expansion of BSS can leverage data science and statistical techniques, using quantitative, reproducible methods to integrate and analyze data sources such as geospatial features, urban infrastructure, and historical usage patterns of BSS. These analyses may provide insights for the decision-making process. For example, to choose locations for new BSS stations, or to build new bike lanes in high-demand areas.

Considering the diverse data sources available and the variability of factors that can influence the BSS usage, it is difficult to compare these systems across cities. Data sources can vary in format, availability, or granularity, making comparison difficult. Although several studies investigate factors that influence demand in BSS, there is a lack of generalizable and reproducible tools that allow for the analysis and comparison of different systems in a standardized way, integrating multiple urban and geospatial layers.

This paper presents BSSAnalyzer, an oriented-object, open-source tool to automate data analysis of BSS. The tool enables the analysis of data related to the population, cities, and the internal characteristics of the BSS themselves. It contains predefined metrics to evaluate the performance of the BSS stations and the whole BSS. BSSAnalyzer implements geospatial statistical analysis to identify factors that drive or hinder BSS usage, facilitating comparisons across different systems. We validated the tool by analyzing BSS data from London (LDN), New York (NYC), and São Paulo (SP).

BSSAnalyzer allows researchers to uncover current usage patterns and the drivers behind system growth. The tool facilitates a multi-faceted analysis, enabling users to contrast high- and low-demand areas, investigate station inactivity, and measure the impact of existing transport and cycling infrastructure on system performance. Furthermore, by identifying strategic points of interest (POI), BSSAnalyzer provides actionable insights to guide future system expansion.

Section 2 shows the related studies, while Section 3 describes the BSSAnalyzer and our methodology. Sections 4 and 5 shows the case studies results and our conclusions.

2. Related Work

Given the relevance of urban mobility for cities, there are several studies related to bike-sharing systems. Some studies aim to understand user behavior [Adnan et al. 2019, Willberg et al. 2021], propose strategies for balancing bicycles across stations [Faghih-Imani et al. 2017, Caggiani et al. 2019], predict usage demand [Ding et al. 2022], case studies [Benedini et al. 2020, Ma et al. 2020], among others.

Several studies compare BSS across different cities. [Fishman et al. 2014] compares how bike-sharing systems have impacted car usage in cities in the United States, Great Britain, and Australia. [De Chardon et al. 2017] studied 75 systems with open data, investigating factors that influence trip counts. The results indicate that temperature and wind influence the number of trips and the quality of cycling infrastructure. However, the expansion of the system (more stations and bikes) did not increase trips. [Tyndall 2022] compared cyclists' commuting behavior with respect to slope and weather conditions in several U.S. cities. The results show that, contrary to [De Chardon et al. 2017] regarding the climate, these factors do not influence the choice to commute by bike.

The factors impacting the trip volume of BSS stations vary considerably across different studies, with significant variations related to regional, cultural, socioeconomic, behavioral, and infrastructure factors of each society. In this need to understand each system in its individuality, tools for urban planning and bike system analysis emerge, such as the one proposed by [Lovelace et al. 2017], which focuses on station planning based on cycling potential and travel flows. Additionally, [Kon et al. 2022] presents a tool for analyzing bike flows in Boston, using the open data from the city's BSS.

Open data initiatives have played a central role in expanding comparative analyses. The CityBikes platform[1] has an API with public information of BSS from more than 400 cities worldwide, providing standardized access to basic data such as station locations, bike availability, and parking spaces. This data allows comparative analyses between cities and encourages urban mobility research based on open data.

Despite these advancements, many analysis models still lack systematic integration between the internal characteristics of different BSS (e.g., number of trips, spatial distribution of stations, capacity) and structural variables of the urban environment (e.g., connectivity with public transport, cycling infrastructure, POI, terrain slope). The BSSAnalyzer framework proposed in this work aims to integrate multiple factors into a unified analytical architecture, offering a reproducible and scalable tool that supports comparative analyses and provides technical insights for urban planning decisions. Moreover, the tool can be adapted to the specificities of each local context.

3. Methodology

In this section, we describe the BSSAnalyzer tool, explaining how it was devised and implemented. Then, we present the features for data analysis and the metrics to evaluate BSS. After, we detail the data sources used for the analyses of our experiments.

3.1. BSSAnalyzer

The BSSAnalyzer[2] was designed to handle data from several cities, each with its own data sources. Thus, the system was structured using the object-oriented programming (OOP) paradigm, allowing for a modular and reusable organization of the code. Using the OOP paradigm eliminates redundancy in the implementation of analysis algorithms for different cities, promoting code reuse and modularity. The tool was built using Python notebooks and data science libraries, such as Pandas, Geopandas, Folium, matplotlib, and scikit-learn, among others. Additionally, algorithms previously developed in the Bike-Science project were applied for georeferenced data analysis [Kon et al. 2022].

[1] `https://citybik.es`
[2] `https://github.com/GabrielIamato/BSSAnalyser`

Figure 1 shows the BSSAnalyzer class diagram. The parent class *Preprocessing* serves as the foundation for processing the distinct datasets. It contains methods responsible for data processing tasks, such as cleaning, normalization, and transformation, preparing the data for subsequent analyses. The idea is to provide common operations that must be inherited by child classes with specific datasets from each city, sharing the same processing and analysis classes. The *Exploratory Data Analysis (EDA)* class is responsible for executing functions related to the initial data exploration. Within this class, functions are implemented to generate descriptive statistics, identify null data, create distribution and correlation plots, perform outlier analysis, and generate tables with key indicators, providing a detailed view of the data's quality and characteristics. The *Bike Science* class performs the main analyses of the system. It uses maps, indexes, and geographic data matching processes to conduct a deeper analysis of bike-sharing system behavior. The class is also responsible for handling spatial data and providing insights into areas of higher and lower usage, as well as assisting in the comparison between the different bike-sharing systems analyzed.

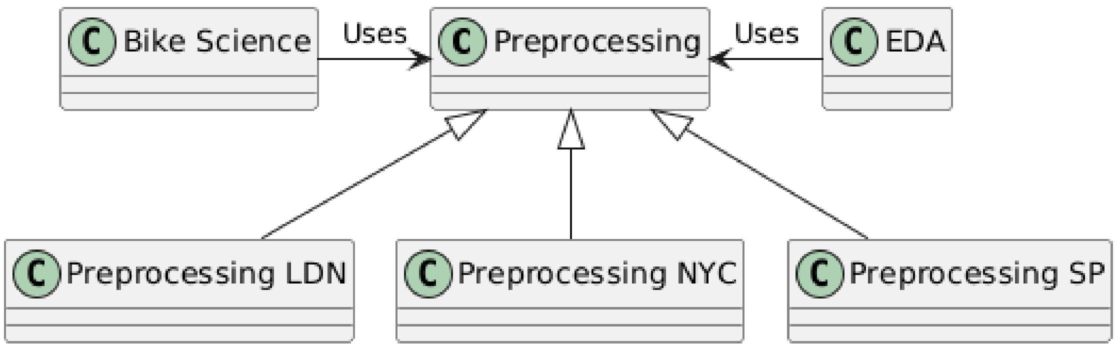


Figure 1. BSSAnalyzer class diagram. The data is processed in the *Preprocessing* class and analyzed in the *Bike Science* and *EDA* classes.

3.1.1. Preprocessing

The data sources used by the BSSAnalyzer are based on related studies, such as BSS stations, trips, and parking spots, which are directly related to the BSS's operational performance and attractiveness [Fishman 2016]. Other environmental factors, such as slope, POI, land use, public transportation, and cycling infrastructure, also influence the demand for stations, often being associated with trip volume [Faghih-Imani et al. 2014, Wang et al. 2016, El-Assi et al. 2017]. Additionally, the *Preprocessing* class collects and processes bike routes and the behavior of the k nearest stations (k = 5), including the average number of trips, average distance, and average travel time.

The *Preprocessing* class also contains functions for data cleaning, standardization, and normalization procedures. These functions must be implemented in the child classes, which must handle the specific datasets from each city. After processing in the child classes, the data is stored on disk in an organized way, making it available for use in the EDA and BikeScience classes, allowing a standard integration for the analyses.

3.1.2. Exploratory Data Analysis

The *EDA* class provides a set of essential functionalities for understanding and evaluating the data, such as central tendency, dispersion, shape, and position of the variables. In addition, the module generates a table of missing values, allowing the identification of gaps in the dataset and facilitating the treatment of these issues. The module also offers a variety of distribution plots, such as histograms, boxplots, KDE plots (kernel density estimation plots), violin plots, and bar plots.

Other EDA functionalities included are correlation matrices, scatter plots, contingency tables, and tests of independence (e.g., Chi-Square). The EDA class also provides Z-Score plots, tests of distribution and normality (e.g., Kolmogorov–Smirnov), and methods for outlier detection (e.g., interquartile range). These exploratory analysis processes are fundamental for ensuring the quality and consistency of the data, directly supporting the subsequent analyses carried out in the Bike Science class.

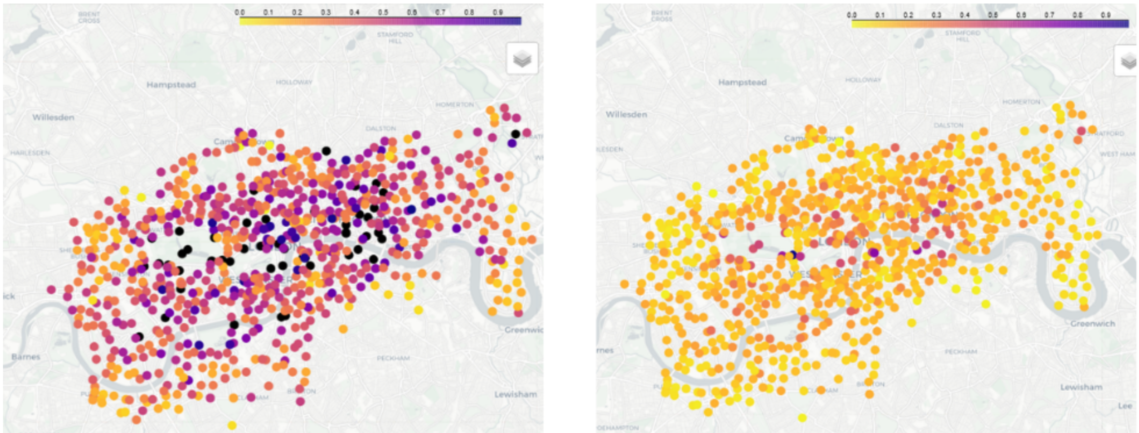
3.1.3. Bike Science

The *Bike Science* class is responsible for generating visual analyses and creating indices for geospatial analyses. This class is based on the BikeScience tool [Kon et al. 2022], which was improved to include the following features. The Bike Science class provides interactive maps in which the points represent the BSS stations, and the color of each point varies according to the intensity of a given index, which may refer to trips, public transport intersections, cycling infrastructure, among others. The visualizations follow best practices for color scales, considering accessibility for color-blind users. Figure 3 shows an example. Section 3.2 describes the main metrics and analyses of Bike Science.

Beyond providing geographic markers, BSSAnalyzer also provides heat maps and choropleth maps for the analyses. Thus, it is possible to verify the characteristics of trips aggregated by regions in different granularities. Using the trip index (Section 3.2), it is possible to identify spatial usage patterns that would be perceived through tables or aggregated statistics. These visualizations make it possible to detect areas of high and low demand, regions with potentially underutilized capacity, and possible imbalances in the spatial distribution of stations.

Most indices range from 0 to 1 to standardize the analyses and make the results more numerically interpretable, but may suffer from the effect of outliers. To address this issue, we use the Interquartile Range (IQR) method. Each gradient map includes an *outlier* identification function that separates them from the scale of 0 to 1 and plots them in distinct colors (black for upper *outliers* and green for lower ones). In this way, the maps maintain a clear visualization of the information, as shown in Figure 2.

The visualizations provided by the *Bike Science* class allow identifying usage patterns across different regions and time periods. It is also possible to identify relationships with the city's features (e.g., public transport, cycling infrastructure). Urban planners and other stakeholders can extract insights, such as the need for bicycle redistribution, system expansion in areas with suppressed demand, or greater integration with public transport, supporting informed decisions about urban mobility and the optimization of BSS.



With outlier treatment **Without outlier treatment**

Figure 2. A map generated without and with outlier treatment.

3.2. BSS metrics and analyses

The BSSAnalyzer provides several metrics and analyses to evaluate the performance and characteristics of BSS, from individual stations to the whole system and its environment. We describe these analyses in this section.

3.2.1. Trip Index

The *Trip Index* (TI) is a ratio between the number of trips (α), the days of operation for each station (β), and the number of bike slots (γ), as shown in Equation 1. It can be applied to a BSS station or to the whole system, and allows us to understand the proportion of trips regarding the available slots over time.

$$TI = \frac{\frac{\alpha}{\beta}}{\gamma} \tag{1}$$

Comparing usage patterns of stations introduced in different periods may introduce bias into the analyses, as older stations tend to have more trips. It occurs because older stations were already counting trips, which can distort direct comparison between stations implemented in distinct periods. If the number of slots is not deemed necessary for the analysis, it is simply disregarded, and the number of trips is divided only by β.

3.2.2. POI proximity

The POI (Points of Interest) proximity analysis aims to verify correlations between these POI (e.g., subway stations, restaurants, schools, parks) and BSS stations' usage. The tool counts the points contained within a certain radius around a BSS station. We adopted a default radius of 500 meters for our analyses, as it is considered a reasonable distance for someone walking from or to her/his destiny [Wei et al. 2023].

For example, when analyzing the POI related to *food*, the algorithm counts food establishments within a 500-meter radius and returns a colored map in which each station is colored according to the number of food-related POIs surrounding it, as shown in Figure 3. It is also possible to generate a heat map following the same logic.

Example station in New York | **A 500-meter radius is drawn around the station** | **The number of subway stations located within this radius is counted (two).**

Figure 3. POI proximity search algorithm for points of interest.

3.2.3. Station inactivity

We proposed an analysis to identify stations with little or no activity. The analysis considers periods of inactivity in which a station is in its regular operation. Inactivity means days without any registered trips as origin, destination, or both, depending on a configurable parameter. Additionally, there is a parameter called "grace period days", which is a tolerance window after the first recorded trip of a station. This period is excluded from the inactivity count, accounting for scenarios in which newly deployed stations undergo an initial phase of adaptation and demand consolidation. [Fishman 2016] indicates that new systems or stations may experience a maturation period before reaching stable or growing usage patterns. For example, the BSS in New York recorded approximately four trips per bike per day in its first month of operation (May 2013), almost doubling by September.

This approach allows users to visualize, through color-coded maps, both the total number of inactive days and the longest consecutive inactivity period for each station. Unlike most of the literature, which predominantly focuses on trip volume, demand patterns, and operational rebalancing [Faghih-Imani et al. 2017, Caggiani et al. 2019, Ding et al. 2022], the proposed metric provides a complementary perspective on operational stability and consistency of use. This information may assist in identifying areas with low system consolidation, potential failures in urban integration, or structural issues, contributing to more informed planning and management decisions.

3.2.4. Cycling infrastructure proximity

There are different approaches to analyze the coverage and quality of cycling infrastructure, such as analyzing bike lane density per area and spatial buffer-based metrics to measure accessibility [Buehler and Dill 2015]. The BSSAnalyzer provides both approaches.

The first consists of counting the number of bike lanes located within a predefined radius (e.g., 500 meters), providing a simple proximity indicator. The second approach advances by incorporating the total length of bike lanes within this surrounding area, producing a weighted index that reflects not only the presence but also the intensity of the available infrastructure. For each station, BSSAnalyzer provides an index that represents the extent of nearby cycle paths throughout the study area.

Thus, one can identify stations that are well located in terms of cycling connectivity, or, conversely, stations located in regions with limited cycling infrastructure. These results may provide insights for public policies, such as expanding the cycling network, reallocating BSS stations, and so on.

3.2.5. Slope

Slope is recognized as one of the main factors influencing cyclists' route choice, since steeper gradients increase physical effort, reduce comfort, and affect route attractiveness. Studies based on route choice modeling with observed data indicate that slope-related variables present significant negative coefficients, evidencing a preference for flatter routes [Broach et al. 2012]. Spatial analyses of BSS indicate that stations located in areas with lower slope tend to present higher trip volumes, reinforcing the importance of topographic variables in evaluating accessibility and system performance [El-Assi et al. 2017]. Thus, beyond the geographic location of stations, it is important to consider the surrounding slope to understand how accessible and attractive each station is for users.

The BSSAnalyzer implements the analysis of slope along the routes between a target station and its k nearest stations (by distance), considering both the destination and origin roles. This design reduces directional bias (uphill/downhill), since the same connection may involve different perceived effort depending on the direction of travel, and brings the analysis closer to what actually impacts cyclists: the topography along the route, rather than only at the station location. From the routes between a station $st1$ and its k neighbors, an aggregated statistic (e.g., average slope) is computed, summarizing the station's "local topographic barrier". Figure 4 shows an example of the slope calculation.

This slope analysis allows to identify stations with good spatial connectivity but low topographic accessibility (e.g., stations that are close in straight-line distance but separated by steep segments), or highlighting regions where system expansion may require mitigation strategies (e.g., alternative flatter routes, or incentives for e-bikes). It also enables fairer comparisons between areas or cities by incorporating a physical determinant that is often overlooked.

3.2.6. k nearest stations

The analysis of the k nearest stations aims to investigate whether the behavior of a station, in terms of trip volume and operational performance, exhibits spatial patterns that are reflected in neighboring stations, indicating possible effects of spatial dependence or competition/local complementarity [Faghih-Imani et al. 2014]. This analysis operates by

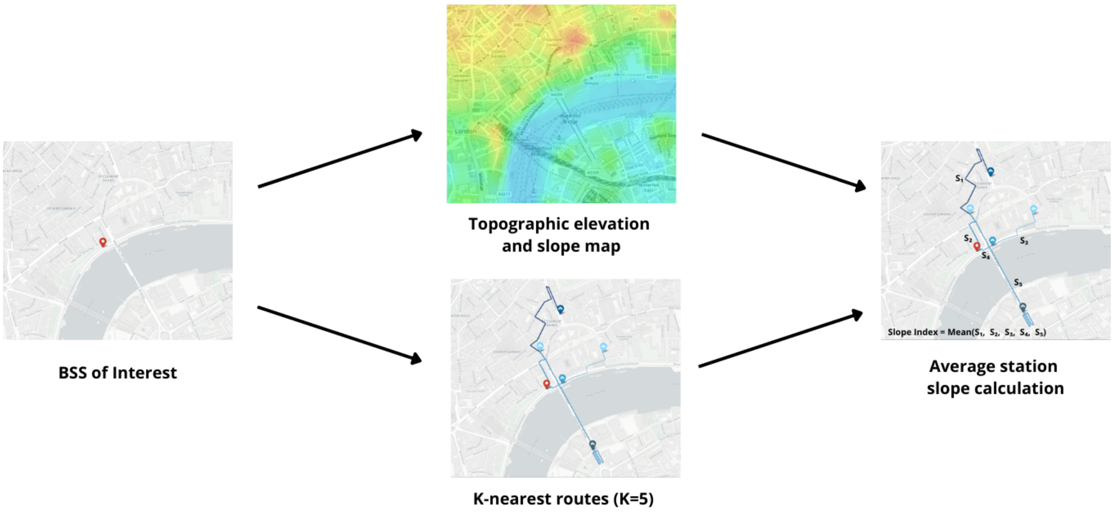


Figure 4. Example of slope calculation for a BSS station. The average slope for each of the 5 nearest routes is calculated to obtain the overall average.

fixing each station as the origin and filtering the k nearest stations according to a parameter (e.g., distance, travel time), collected via public routing APIs using OSM data. Then, metrics from the nearest stations are computed, such as the average number of trips, distance, travel time, and the trip index. Thus, the k nearest stations analysis enables the identification of clusters of high or low demand, possible saturation effects among stations located very close to each other, and regions where the aggregated behavior suggests the need for bicycle redistribution or network expansion.

3.3. Data sources

The data sources we used to build and analyze the BSSAnalyzer tool are mostly public and available on government, company platforms, and other open source organizations. The data from the São Paulo' BSS was provided by its manager company (Tembici / Bike Itaú). Table 1 shows the main data sources used in this study.

Table 1. Data sources of London (LDN), New York City (NYC), and São Paulo (SP).

Category	LDN	NYC	SP
BSS and trips	Transport for London	CitiBike	Bike Itaú
POI	OpenStreetMap	OpenStreetMap	OpenStreetMap
Public Transportation	Transport for London	NYC OpenData / GTFS	GeoSampa
Cycling Infrastructure	London DataStore	NYC OpenData	GeoSampa
Routes	Open Route Service	Open Route Service	GraphHopper
Slope	Open Route Service	NYC OpenData / DEM	GeoSampa
Land Use	London DataStore	NYC OpenData / PLUTO	GeoSampa

4. Case study results and analysis

In this section, we present a case study using data from New York, London, and São Paulo to illustrate some of the possible analyses that the BSSAnalyzer tool supports. These cities represent distinct urban contexts in terms of population density, spatial structure, and the maturity of bike-sharing systems (BSS). New York hosts one of the largest BSS in the world, strongly concentrated in Manhattan and Brooklyn, regions characterized by high population density, mixed land use, and strong integration with public transportation. London has a BSS implemented in 2010, with a strong presence in the central region and adjacent areas, embedded within a dense and consolidated urban network and supported by structured active mobility policies. The BSS of São Paulo was implemented in 2018, being spatially concentrated in business areas. Regarding BSS stations, New York has 2,218 stations, London has 799, and São Paulo has 268 stations.

4.1. Trip Index

Figure 5 illustrates the spatial distribution of stations and the intensity of trips recorded in London, New York, and São Paulo. In New York, a higher concentration of trips can be observed in the regions of Manhattan and Brooklyn, whereas in London, a more distributed pattern emerges, with greater intensity in the central zones. In São Paulo, the trips are highly concentrated in services-related business areas, such as Pinheiros, Bela Vista, and the city center. Figure 6 shows a closer look at the scale of the Trip Index for each station for both systems. The blue background in the maps indicates population density by district. The dark colored areas are districts with higher density.

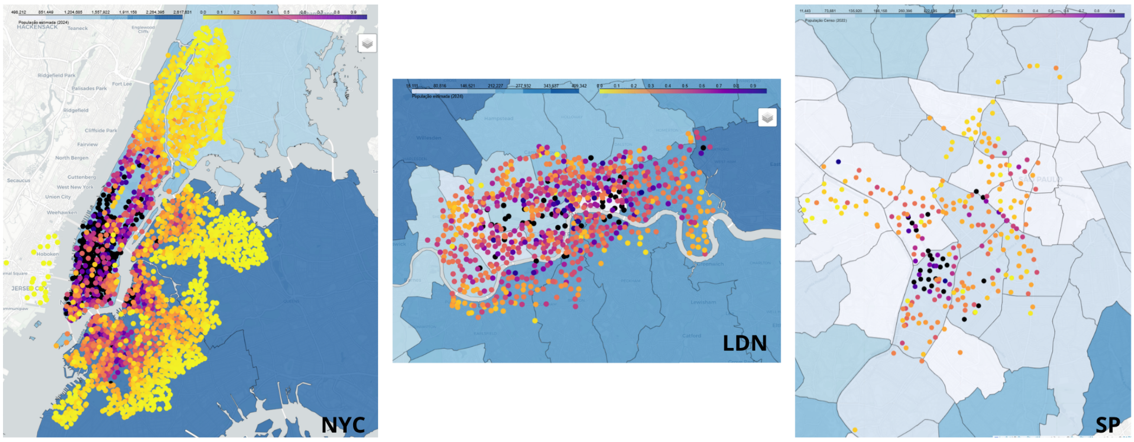


Figure 5. Trip index map overlaid on the population by district map.

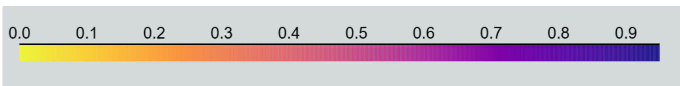


Figure 6. "Plasma" color scale used for all maps with shared bicycle stations, ranging from 0.0 (yellow) to 1.0 (dark purple)

To analyze the association between the trip index and infrastructural, geographic, and urban mobility factors, we used the Spearman correlation coefficient. This choice is justified by the non-normality of the data, the presence of outliers, and the possibility of non-linear monotonic relationships between the analyzed variables. Table 2 shows these correlations. Food-related POI show a strong positive association with the trip index for New York, with moderated positive values in São Paulo and London. This is consistent with findings in the literature, where food-related POI indicate urban centrality and attractiveness [Wang et al. 2016, El-Assi et al. 2017]. Figure 7 compares London, New York, and São Paulo for trips and different types of POI. The differences between cities should be considered in terms of specific physical and structural characteristics. Regarding slope, London has a predominantly flat topography, which may reduce its influence on travel patterns. São Paulo has a rugged topography, negatively impacting the trip index.

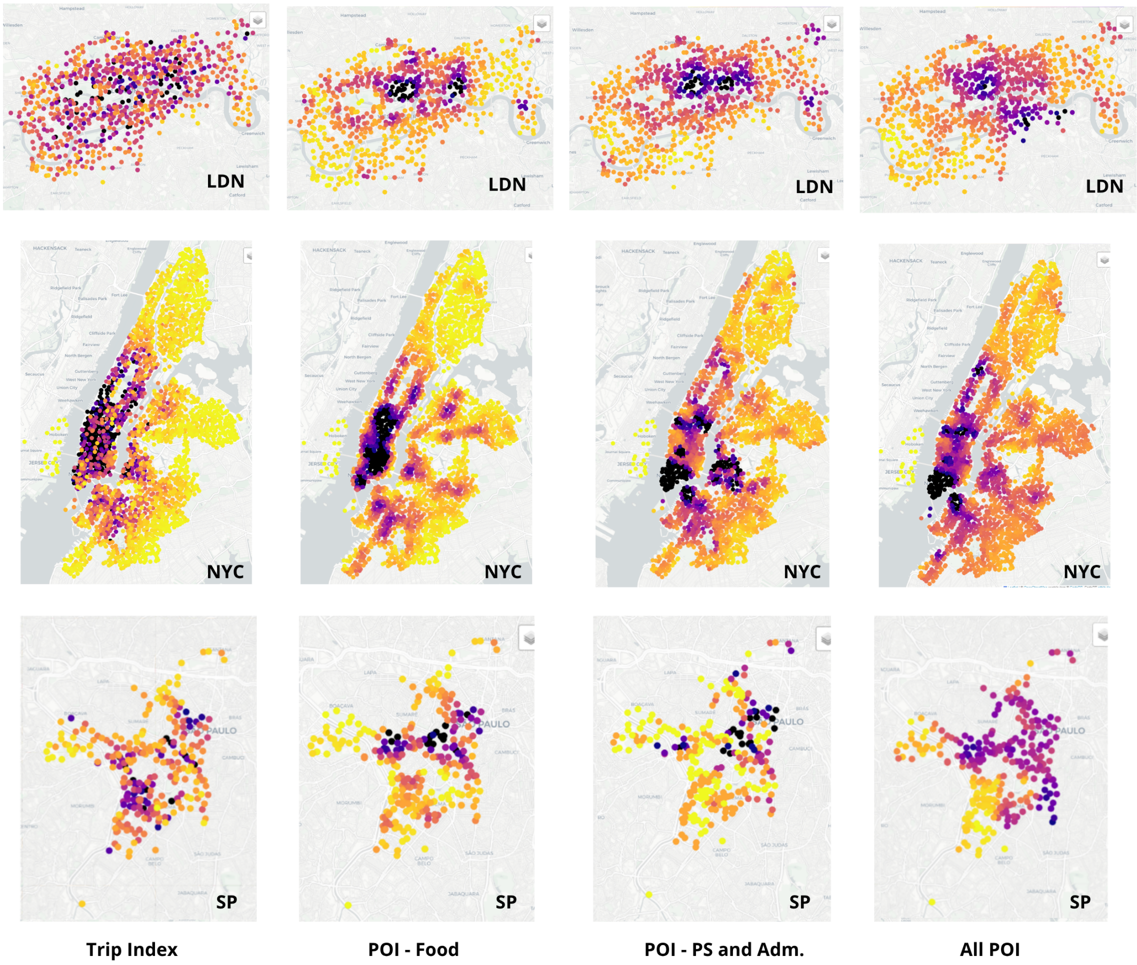


Figure 7. Maps comparing BSS stations for London and New York: trip index, food-related POI; public and administrative services POI, and all POI.

Table 2. Spearman correlation for urban features with the Trip Index in LDN, NYC, and SP. The green gradient highlights the 3 highest correlations in each column, excluding the k nearest ones (highlighted in yellow).

Features	LDN	NYC	SP
Commerce	0.18	0.64	0.49
Slope	-0.02	-0.24	-0.46
Food	0.29	0.75	0.35
Health	0.24	0.62	0.23
Public and Administrative Services	0.26	0.74	0.17
All POI	0.37	0.69	-0.04
Cycling Infrastructure	0.15	0.57	0.11
Public Transportation	0.22	0.54	0.08
k **nearest stations (k = 5)**	0.43	0.88	0.51

Another consistent result is that cycling infrastructure and proximity to public transport exhibit positive associations across the three cities. The literature highlights that

BSS stations located near public transportation corridors and stops/stations are often associated with intense usage patterns, reinforcing the role of BSS in last-mile connectivity and modal integration [Mahajan and Sánchez-Vaquerizo 2024]. Moreover, the metric k nearest stations presents high correlation in all cities, which may indicate that nearest stations tend to share similar urban contexts and correlated demand patterns, consistent with the presence of spatial dependence in BSS data [Wang et al. 2016].

The different magnitudes of the coefficients should not be interpreted as universal effects, but rather as reflections of the topographic and urban conditions of each city, reinforcing the importance of contextualized analyses in comparative studies of BSS.

4.2. Inactivity Index

The analysis of inactivity data reveals distinct patterns between the BSS of London, New York, and São Paulo, which can be seen in Figure 8. In New York, a significant concentration of inactivity periods is observed in the Manhattan region, which represents a paradoxical pattern, given that this area concentrates the highest volume of recorded trips. This phenomenon suggests the occurrence of scheduled station deactivation windows or severe operational interruptions. Regarding the London network, the analysis indicates a more homogeneous geospatial distribution of inactivity. In London, inactive stations are more evenly distributed throughout the system. Most of them exhibit a latency period of at least 23 days, suggesting a network maintenance or adjustment schedule. In São Paulo, the most inactive stations are at the edges of the BSS, concentrated on the west campus of the University of São Paulo (USP) and in the northern zone. Some stations have been inactive for up to 50 days, which may be related to vacation periods at USP.

Figure 8. Total days of inactivity per station of London, New York and São Paulo

5. Conclusions

The BSSAnalyzer is an analytical tool that aims to integrate diverse data sources and to facilitate the analysis of different bike-sharing systems. The proposed architecture, based on the oriented-object programming paradigm, was designed to operate based on widely available georeferenced data, allowing the method to be applied to different cities and bike-sharing systems without requiring complex structural adaptations. As a result, the tool presents a scalable and replicable framework, enabling its application both in academic research and in practical urban planning and mobility management contexts.

Thus, the developed model contributes to supporting strategic decisions related to the expansion of bike-sharing systems, the placement of new stations, and the evaluation of existing urban infrastructure. By enabling comparative analyses across different cities and urban contexts, the tool expands the potential for generating insights for sustainable mobility policies, offering an analytical framework capable of accompanying the evolution and expansion of these systems on a global scale.

The analysis conducted in this study showed that different urban environmental factors present relevant associations with the performance of BSS stations. Among the analyzed elements, the presence of food establishments, public and administrative services, characteristics of the surrounding urban environment, and terrain slope stood out as consistently associated with the observed usage patterns across the studied cities. As future work, we will analyze more BSS to investigate common features correlated with BSS usage that are generalizable to most cities. Future analyses should incorporate spatial autocorrelation methods to address limitations of bivariate correlation, which does not account for spatial dependence among stations. Usability tests with end users and practical comparison with other BSS analysis frameworks are also planned.

6. Acknowledgements

This research is funded by FAPESP procs. 25/21593-7 and 23/00811-0, CNPq 118640/2024-0, and PIBIC/USP/CNPq. We would like to thank TemBici for sharing the BSS data from São Paulo and the InterSCity project for its support during this research.

References

Adnan, M., Altaf, S., Bellemans, T., Yasar, A.-u.-H., and Shakshuki, E. M. (2019). Last-mile travel and bicycle sharing system in small/medium sized cities: user's preferences investigation using hybrid choice model. *Journal of Ambient Intelligence and Humanized Computing*, 10:4721–4731.

Benedini, D. J., Lavieri, P. S., and Strambi, O. (2020). Understanding the use of private and shared bicycles in large emerging cities: The case of São Paulo, Brazil. *Case Studies on Transport Policy*, 8.

Broach, J., Dill, J., and Gliebe, J. (2012). Where do cyclists ride? a route choice model developed with revealed preference GPS data. *Transportation Research Part A: Policy and Practice*, 46(10):1730–1740.

Buehler, R. and Dill, J. (2015). Bikeway networks: A review of effects on cycling. *Transport Reviews*, 36:1–19.

Caggiani, L., Camporeale, R., Marinelli, M., and Ottomanelli, M. (2019). User satisfaction based model for resource allocation in bike-sharing systems. *Transport Policy*, 80:117–126.

CEBRAP (2018). Impacto social do uso da bicicleta em São Paulo.

De Chardon, C. M., Caruso, G., and Thomas, I. (2017). Bicycle sharing system 'success' determinants. *Transportation Research Part A: Policy and Practice*, 100:202–214.

Ding, H., Lu, Y., Sze, N., and Li, H. (2022). Effect of dockless bike-sharing scheme on the demand for London Cycle Hire at the disaggregate level using a deep learning approach. *Transportation Research Part A: Policy and Practice*, 166:150–163.

Duarte, F. (2016). Disassembling bike-sharing systems: Surveillance, advertising, and the social inequalities of a global technological assemblage. *Journal of Urban Technology*, 23(2):103–115.

Duran-Rodas, D., Chaniotakis, E., and Antoniou, C. (2019). Built environment factors affecting bike sharing ridership: Data-driven approach for multiple cities. *Transportation Research Record*, 2673(12):55–68.

El-Assi, W., Mahmoud, M., and Nurul Habib, K. (2017). Effects of built environment and weather on bike sharing demand: A station level analysis of commercial bike sharing in Toronto. *Transportation*, 44.

Faghih-Imani, A., Eluru, N., El-Geneidy, A., Rabbat, M., and Haq, U. (2014). How land-use and urban form impact bicycle flows: Evidence from the bicycle-sharing system (BIXI) in Montreal. *Journal of Transport Geography*, 41.

Faghih-Imani, A., Hampshire, R., Marla, L., and Eluru, N. (2017). An empirical analysis of bike sharing usage and rebalancing: Evidence from Barcelona and Seville. *Transportation Research Part A: Policy and Practice*, 97:177–191.

Fishman, E. (2016). Bikeshare: A review of recent literature. *Transport Reviews*, 36(1):92–113.

Fishman, E., Washington, S., and Haworth, N. (2014). Bike share's impact on car use: Evidence from the United States, Great Britain, and Australia. *Transportation Research Part D: Transport and Environment*, 31.

Kon, F., Ferreira, É. C., de Souza, H. A., Duarte, F., Santi, P., and Ratti, C. (2022). Abstracting mobility flows from bike-sharing systems. *Public Transport*, 14:545–581.

Lovelace, R., Goodman, A., Aldred, R., Berkoff, N., Abbas, A., and Woodcock, J. (2017). The Propensity to Cycle Tool: An open source online system for sustainable transport planning. *Journal of Transport and Land Use*, 10.

Ma, X., Yuan, Y., Oort, N. V., and Hoogendoorn, S. (2020). Bike-sharing systems' impact on modal shift: A case study in Delft, the Netherlands. *Journal of Cleaner Production*, 259.

Mahajan, S. and Sánchez-Vaquerizo, J. A. (2024). Global comparison of urban bike-sharing accessibility across 40 cities. *Scientific Reports*, 14(1):20493.

Tyndall, J. (2022). Cycling mode choice amongst US commuters: The role of climate and topography. *Urban Studies*, 59(1):97–119.

Wang, X., Lindsey, G., Schoner, J. E., and Harrison, A. (2016). Modeling bike share station activity: Effects of nearby businesses and jobs on trips to and from stations. *Journal of Urban Planning and Development*, 142(1).

Watts, M. (2018). How walking & cycling is transforming cities.

Wei, J., Chen, Y., Liu, Z., and Wang, Y. (2023). Correlation between the built environment and dockless bike-sharing trips connecting to urban metro stations. *Journal of Transport and Land Use*, 16(1):131–161.

Willberg, E., Salonen, M., and Toivonen, T. (2021). What do trip data reveal about bike-sharing system users? *Journal of Transport Geography*, 91:102971.