

UNIVERSIDADE DE SÃO PAULO
ESCOLA POLITÉCNICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TRANSPORTES

THAIS MALTA BARACAT

**Análises espaço-temporais do perfil de uso das estações de bicicletas
compartilhadas**

Versão Corrigida

SÃO PAULO
2022

THAIS MALTA BARACAT

**Análises espaço-temporais do perfil de uso das estações de bicicletas
compartilhadas**

Versão Corrigida

Dissertação apresentada à Escola
Politécnica da Universidade de São Paulo
para obtenção do título de Mestre em
Ciências

Área de concentração:
Engenharia de Transportes

Orientador: Prof. Dr. Orlando Strambi

Co-orientadora: Dra. Patricia Sauri Lavieri

SÃO PAULO

2022

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

Assinatura do autor: _____

Assinatura do orientador: _____

Catálogo-na-publicação

Baracat, Thais

Análises espaço-temporais do perfil de uso das estações de bicicletas compartilhadas / T. Baracat -- versão corr. -- São Paulo, 2022.
68 p.

Dissertação (Mestrado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Transportes.

1.Bicicletas compartilhadas 2.Perfil temporal 3.Espaço urbano 4.Clusters 5.Análise multivariada I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Transportes II.t.

BARACAT, T.M. **Análises espaço-temporais do perfil de uso de estações de bicicletas compartilhadas**. Dissertação (Mestrado em Engenharia de Transportes) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2022.

Aprovado em: 22/08/2022

Banca examinadora

Prof. Dr. Orlando Strambi

Instituição: Escola Politécnica da Universidade de São Paulo

Julgamento: Aprovado

Prof. Dr. Cira Souza Pitombo

Instituição: Escola de Engenharia de São Carlos da Universidade de São Paulo

Julgamento: Aprovado

Prof. Dr. Ana Margarita Larrañaga Uriarte

Instituição: Universidade Federal do Rio Grande do Sul

Julgamento: Aprovado

AGRADECIMENTOS

Esta pesquisa teve como principal motivação a construção de um espaço-urbano mais amigável e sustentável para as gerações atuais e futuras de ciclistas. Este trabalho foi fruto de muita dedicação, e só possível de ser realizado com a sinergia dos membros envolvidos.

Em primeiro lugar, gostaria de agradecer ao Prof. Dr. Orlando Strambi, pela constante discussão e troca de ideias, bem como no auxílio da estruturação e do desenvolvimento da pesquisa. Gostaria de agradecer à Prof. Dra. Patrícia Lavieri, pela contribuição em entender o propósito da pesquisa e no desenvolvimento do modelo confirmatório. Agradeço também à empresa TemBici, em especial à Renata Rabello, pela disponibilização dos dados do sistema de bicicletas compartilhadas de São Paulo. À Prof. Dra. Mariana Giannotti, agradeço pela contribuição sobre o entendimento da análise espacial e pelo contato com outros pesquisadores.

Agradeço aos meus colegas de pesquisa: à Silvia Mauad, que iniciou e esteve presente desde o início desta jornada comigo, ao Fernando Gomes, pelo entusiasmo em transformar e manipular dados espaciais e ao Flávio Soares, por me introduzir as plataformas voltadas para os ciclistas.

Em seguida agradeço à minha família e aos meus amigos pela motivação e palavras de apoio, assim como por fornecer outros pontos de vista para transformar o produto deste trabalho mais acessível e amigável para novos leitores. À Engimind Consultores, agradeço pela flexibilidade e compreensão durante o mestrado.

Por fim, sou extremamente grata às instituições de ensino da qual fui aluna, e especialmente à Universidade de São Paulo, que me auxiliou a conquistar desde a graduação diversas conquistas pessoais e profissionais.

RESUMO

Os sistemas de bicicletas compartilhadas vêm ganhando popularidade nos últimos anos, de forma que precisam ser responsivos à demanda. Logo, faz-se necessário entender o comportamento espaço-temporal dessas viagens. Esta pesquisa estudou o sistema de estações fixas Bikesampa, da cidade de São Paulo, e teve como primeiro objetivo classificar as estações segundo a demanda horária de retiradas e devoluções. O algoritmo de agrupamento k-médias foi utilizado, resultando na identificação de três grupos de estações: (i) balanceado, (ii) desbalanceado, com maior número de retiradas na manhã, e (iii) desbalanceado, com maior número de devoluções pela manhã. Através de uma análise de autocorrelação espacial, verificou-se que as estações de cada grupo não se distribuem aleatoriamente no espaço, indicando a necessidade de diferentes estratégias de reposicionamento de bicicletas entre estações dependendo da área onde estão localizadas e sugerindo a existência de associação com características do entorno das estações. Esta hipótese foi confirmada por meio de um modelo probit, que demonstra a associação do comportamento temporal da demanda com atributos da operação, do espaço urbano e características socioeconômicas. Conhecer o comportamento temporal das viagens de bicicletas compartilhadas permite o desenvolvimento de políticas de operação e de incentivo ao usuário para melhorar a eficiência desses sistemas. Também, permite prever o comportamento temporal das estações diante de novas estratégias de reposicionamento de bicicletas e implantação de estações no curto e médio prazo, ou após mudanças socioespaciais de longo prazo.

Palavras-chave: Bicicletas compartilhadas, Perfil temporal, Espaço urbano, Clusters, Análise multivariada

ABSTRACT

Bikesharing has gained popularity over the years and now faces the challenge to be responsive and meet the growing demand. Thus, understanding the temporal pattern of bikesharing trips is paramount. This study examined data from the Bikesampa bikesharing system (a fixed station system operating in the Brazilian city of São Paulo). First, a k-means clustering was applied to group bikesharing stations according to hourly demand of bike pick-ups and returns. The results revealed three clusters of stations with well-defined and distinct temporal patterns: (i) balanced, (ii) unbalanced, with higher rates of bike pick-up in the morning and (iii) unbalanced, with higher rates of bike return in the morning. A spatial autocorrelation analysis showed that stations belonging to each cluster are not randomly distributed over space, indicating that the system may require different rebalancing strategies depending on the area where stations are located and suggesting the association with built environment characteristics around the stations. This hypothesis was confirmed through a probit model, which shows the association between the temporal demand patterns and operational, spatial, and socioeconomic attributes. Such understanding can help guiding the development of operational strategies and user incentive policies to improve the efficiency of bikesharing systems. Also, it allows for the prediction of temporal patterns when analyzing new strategies regarding bicycle repositioning or implementation of stations in the short and medium term, or after socio-spatial changes in the long-term.

Keywords: Bikesharing, Temporal profile, Urban environment, Clusters, Multivariate analysis

LISTA DE FIGURAS

Figura 1 – Lacunas, hipóteses, objetivos e contribuições	17
Figura 2 – Sequência de atividades da pesquisa	18
Figura 3 – Distribuição da população, emprego e renda na cidade de São Paulo.....	20
Figura 4 – Localização das 239 estações do Bikesampa selecionadas para análise	23
Figura 5 – Passos para o cálculo dos perfis temporais de uso das estações	30
Figura 6 – Curva de silhueta para diferentes valores de número de grupos	31
Figura 7 – Perfis temporal de uso de cada grupo de estações	33
Figura 8 – Distribuição espacial dos grupos de perfis temporais	34
Figura 9 – Uso do solo e pontos de referência relevantes na região das estações	35
Figura 10 – Zonas utilizadas para o cálculo do Índice de Moran.....	39
Figura 11 – Mapa de cluster e mapa de significância LISA.....	40
Figura 12 – Sugestão de rebalanceamento Berrini-Brooklin/Vila Cordeiro	42
Figura 13 – Gráfico <i>boxplot</i> da distribuição da variável distância média para cada grupo.....	43
Figura D1 – Gráficos <i>boxplot</i> e colunas das variáveis explicativas quantitativas	66

LISTA DE TABELAS

Tabela 1 – Temas de estudo sobre sistemas de bicicletas compartilhadas da 3ª geração	14
Tabela 2 – Características positivamente correlacionadas com o uso de bicicletas compartilhadas ..	14
Tabela 3 – Quadro resumo das atividades e ferramentas utilizadas.....	18
Tabela 4 – Características do sistema de transportes da cidade de São Paulo	21
Tabela 5 – Descrição das variáveis explicativas.....	25
Tabela 6 – Fonte de informação e anos de coleta e referência das variáveis explicativas.....	26
Tabela 7 – Áreas de serviço das estações [ha]	28
Tabela 8 – Estatística descritiva das variáveis explicativas.....	29
Tabela 9 – Hipóteses H_0 e H_1 do I de Moran	37
Tabela 10 – Classificação LISA	38
Tabela 11 – Índice de Moran Global	38
Tabela 12 – Valores médios e contagem das variáveis para os diferentes grupos	44
Tabela 13 – Resultados da estimação do modelo probit.....	48
Tabela 14 – Resultados da matriz de covariância	49
Tabela 15 – Verossimilhança e qui-quadrado.....	49
Tabela 16 – Quadro resumo com interpretações do modelo.....	54
Tabela B1 – Especificações alternativas formação de grupos de perfil temporal de uso das estações	64
Tabela C1 – Número de estações nas zonas da área de cobertura do sistema Bikesampa.....	65

LISTA DE SIGLAS

CEM	Centro de Estudos da Metr�pole
CET	Companhia de Engenharia de Tr�fego
CPTM	Companhia Paulista de Trens Metropolitanos
EMTU	Empresa Metropolitana de Transportes Urbanos
FAPESP	Funda�o de Amparo � Pesquisa do Estado de S�o Paulo
G1	Grupo 1 – Balanceado
G2	Grupo 2 – Retiradas Manh�
G3	Grupo 3 – Devolu�es Manh�
GTFS	<i>General Transit Feed Specification</i>
IBGE	Instituto Brasileiro de Geografia e Estat�stica
ITDP	Instituto de Pol�ticas de Transporte e Desenvolvimento
LISA	Local Indicator of Spatial Autocorrelation
METR�	Companhia do Metropolitano de S�o Paulo
OSM	Open Street Map
POD	Pesquisa Origem e Destino
RMSP	Regi�o Metropolitana de S�o Paulo
SM	Sal�rio m�nimo
SPTRANS	S�o Paulo Transportes
UCB	Uni�o de Ciclistas do Brasil

SUMÁRIO

1.	INTRODUÇÃO	12
2.	REVISÃO DA LITERATURA	13
3.	OBJETIVOS E CONTRIBUIÇÕES	17
4.	MATERIAIS E MÉTODOS	18
4.1.	Caracterização do local de estudo.....	18
4.2.	Caracterização do objeto de estudo	21
4.3.	Coleta e tratamento das variáveis explicativas.....	24
4.3.1.	<i>As variáveis explicativas.....</i>	<i>24</i>
4.3.2.	<i>Coleta de informação</i>	<i>24</i>
4.3.3.	<i>Tratamento dos dados.....</i>	<i>27</i>
4.3.4.	<i>Estatística descritiva das variáveis explicativas.....</i>	<i>28</i>
4.4.	Grupos de estações	30
4.4.1.	<i>Cálculo dos perfis temporais médios das estações.....</i>	<i>30</i>
4.4.2.	<i>Classificação das estações segundo perfis temporais</i>	<i>30</i>
4.5.	Análise da autocorrelação espacial dos grupos de estações.....	37
4.6.	Relação entre grupos de estações e variáveis explicativas	42
4.6.1.	<i>Pré-seleção de variáveis</i>	<i>42</i>
4.6.2.	<i>Estimação de modelo confirmatório.....</i>	<i>45</i>
5.	LIMITAÇÕES E SUGESTÕES	55
6.	CONCLUSÕES	56
	REFERÊNCIAS.....	58
	ANEXO A – FONTE DE INFORMAÇÃO DAS VARIÁVEIS EXPLICATIVAS	62
	ANEXO B – ESPECIFICAÇÕES PARA CONSTRUÇÃO DO PERFIL DE USO.....	64
	ANEXO C – NÚMERO DE ESTAÇÕES NAS ZONAS DE COBERTURA DO SISTEMA	65
	ANEXO D – DISTRIBUIÇÃO DAS VARIÁVEIS EXPLICATIVAS NOS GRUPOS.....	66

1. INTRODUÇÃO

O uso de bicicleta vem crescendo gradativamente, seja pelo aumento do uso recreacional e desportivo (LEVY; GOLANI; BEN-ELIA, 2019), ou pelos benefícios para a saúde e meio ambiente (SHAHEEN; COHEN; MARTIN, 2013; ZHANG et al., 2015), ou ainda por ser economicamente acessível para diferentes classes sociais (CHARDON; CARUSO; THOMAS, 2017; PRITCHARD et al., 2019). Também, os sistemas de bicicletas compartilhadas – transporte que oferece o aluguel temporário de bicicletas em regiões estrategicamente localizadas – explicam esse aumento (FAGHIH-IMANI et al., 2017; FISHMAN, 2016; WANG et al., 2016).

Ainda, a pandemia do Covid-19 impulsionou o uso de modos de transporte ativo, como a bicicleta, como solução para o deslocamento de curtas distâncias e redução da possibilidade de difusão do vírus nas rotas de transporte, e também pela implantação de ciclovias temporárias, mas que permaneceram em operação no período pós-pandemia (ABREU; ALMEIDA; TURINI, 2021).

Em particular, as bicicletas compartilhadas conferem conveniência ao usuário e tornam este meio de transporte mais acessível e flexível, pois dispensam a posse de bicicleta e diminuem a barreira de entrada, introduzindo novos ciclistas ao sistema (BENEDINI; LAVIERI; STRAMBI, 2019; EREN; UZ, 2020; FISHMAN; WASHINGTON; HAWORTH, 2013). Os sistemas de bicicletas compartilhadas se dividem em dois tipos: (i) de estações fixas, em que o usuário realiza a retirada e a devolução da bicicleta na estação e (ii) sem estações, que permite o deslocamento porta a porta (WU; KIM; CHUNG, 2021). O foco da dissertação são os sistemas de estações fixas.

Com a crescente popularidade das bicicletas compartilhadas, é fundamental entender a dinâmica do compartilhamento de bicicletas para desenvolver sistemas eficientes e responsivos à demanda (SHAHEEN; COHEN; MARTIN, 2013). Estudos recentes apontam que o espaço urbano (tipo de uso do solo e ambiente construído) e características socioeconômicas afetam o comportamento de viagem dos usuários (CHARDON; CARUSO; THOMAS, 2017; EREN; UZ, 2020; GEHRKE; WELCH, 2019; FAGHIH-IMANI et al., 2014; FAGHIH-IMANI et al., 2017; FLORINDO et al., 2018; RADZIMSKI; DZIĘCIELSKI, 2019; SARAIVA et al., 2019; SCHIMOHR; SCHEINER, 2021; TRAN, OVTRACHT e D'ARCIER, 2015; WANG et al., 2016; WU; KIM; CHUNG, 2021). Em geral, estas características são variáveis explicativas de modelos confirmatórios, nas quais a variável dependente Y é usualmente quantitativa e

expressa pelo número de viagens, seja em números totais ou desagregadas em retiradas e devoluções, em cada estação e em algum período estabelecido.

Assim como o espaço, o tempo também se relaciona com o comportamento de sistemas de bicicletas compartilhadas. Boa parte delas é utilizada no pico da manhã e da tarde, frequentemente para viagens pendulares de curta distância de primeira e última milha (GU; KIM; CURRIE, 2019; WU; KIM; CHUNG, 2021). Há também usos mais difusos ao longo do dia, especialmente por motivos de educação e de lazer (FAGHIH-IMANI et al., 2017; GU; KIM; CURRIE, 2019). Os estudos que avaliam conjuntamente aspectos espaciais e temporais geralmente incorporam a dimensão temporal na variável dependente Y, a qual passa a variar de hora em hora (FAGHIH-IMANI et al., 2014, 2017), ou criam modelos distintos para os picos (MIX; HURTUBIA; RAVEAU, 2022; TRAN; OVTRACHT; D'ARCIER, 2015), ou acrescentam variáveis *dummy* indicando o período do dia (SCHIMOHR; SCHEINER, 2021).

Em geral, estes estudos sintetizam os resultados temporais através de um único perfil diário de demanda para o sistema de bicicletas compartilhadas e falham em testar a hipótese de existência de mais de um comportamento temporal no sistema em análise. Estudos como o de Gu, Kim e Currie (2019) e Lathia, Ahmed e Capra (2012), ou então o estudo de Verma et al. (2021), este último no âmbito de estações de metrô, capturam justamente a existência de diferentes perfis diários de uso em um mesmo sistema, mas não os correlacionam estatisticamente com características do espaço urbano. Eventualmente, se limitam a comentar, de maneira subjetiva, como os comportamentos temporais estão relacionados com o uso do solo.

Neste sentido, o objetivo geral do estudo é ampliar o entendimento do comportamento espaço-temporal de sistemas de bicicletas compartilhadas com estações fixas, visando orientar a elaboração de políticas de operação e incentivo ao usuário. O objeto de estudo foi o sistema Bikesampa da cidade de São Paulo.

O texto está estruturado da seguinte forma. A revisão da literatura está no Capítulo 2. O Capítulo 3 apresenta os objetivos e contribuições. Os materiais e métodos estão no Capítulo 4. O Capítulo 5 contém as limitações do trabalho e sugestões para pesquisas futuras. As conclusões estão no Capítulo 6.

2. REVISÃO DA LITERATURA

As pesquisas sobre os sistemas de compartilhamento de bicicletas aumentaram consideravelmente nos últimos anos. Si et al. (2019) analisaram os

tópicos mais recorrentes na literatura entre 2010 e 2018 e identificaram as categorias mais estudadas no período (ver Tabela 1). Em particular, apontam que o problema mais urgente dos sistemas da 3ª geração¹ é o desequilíbrio da oferta e demanda.

Tabela 1 – Temas de estudo sobre sistemas de bicicletas compartilhadas da 3ª geração

Categorias	Temas principais
Fatores e barreiras	Ambiente construído, infraestrutura, clima, campus universitário, número e localização de estações e processo de adesão
Otimização do sistema	Reposição, redistribuição e rebalanceamento de bicicletas
Comportamento e impacto	Benefícios diretos (ex. conveniência, redução custos/tempo) e indiretos (ex. uso conjugado com transporte público, redução de emissões)
Segurança e saúde	Risco, saúde, uso de capacete e atividade física

Fonte: Adaptado de SI et al. (2019).

A literatura apresenta estudos que estabelecem a relação entre a demanda de bicicletas compartilhadas e os fatores e barreiras que estimulam ou restringem o seu uso. Estes estudos buscam entender como características operacionais, socioeconômicas e do espaço urbano impactam o uso das estações. Frequentemente, utilizam uma métrica agregada no tempo (dia, mês ou ano) para representar a demanda pelas bicicletas compartilhadas. A Tabela 2 apresenta alguns estudos e as características com correlação positiva com a métrica de uso de bicicletas compartilhadas. O leque de variáveis indicadas nos diversos estudos da Tabela 2 orientou a busca de informações disponíveis para a análise em São Paulo.

Tabela 2 – Características positivamente correlacionadas com o uso de bicicletas compartilhadas

Autores	Local	Características
América do Norte		
Wang et al. (2016)	Minneapolis – St. Paul	Percentual da população branca Existência de via pavimentada no entorno da estação Distância até a estação de bicicleta mais próxima Número de empregos em um raio de 30 minutos Número de restaurantes no entorno da estação
Faghih-Imani et al. (2014)	Montreal	Extensão de ciclovias e ciclofaixas no entorno da estação Extensão de vias locais no entorno da estação Número de estações no entorno da estação Número de restaurantes no entorno da estação Presença de metrô no entorno da estação Presença de universidade no entorno da estação Densidade populacional na zona da estação Densidade de empregos na zona da estação
América Latina		
Saraiva et al. (2019)	Passo Fundo	Quantidade de instituições de ensino no entorno da estação Áreas verdes no entorno da estação Distância até a estação mais próxima Distância até a ciclovia/ciclofaixa mais próxima População potencial no entorno da estação
Mix; Hurtubia; Raveau (2022)	Santiago	Número de residências no entorno da estação Número de escritórios no entorno da estação

¹ Os sistemas da 3ª geração possuem dispositivos que previnem o furto de bicicletas e tecnologias para identificar o usuário e a localização de bicicletas (Demaio e Gifford, 2004).

		Extensão de infraestrutura cicloviária no entorno da estação Tamanho médio da rua
Ásia		
Wu; Kim; Chung (2021)	Suzhou	Tamanho da estação Densidade populacional no entorno da estação Número de shopping center no entorno da estação Número de restaurantes no entorno da estação Número de serviços financeiros no entorno da estação Número de centros acadêmicos no entorno da estação Número de empregos no entorno da estação Acessibilidade da estação relativa ao entorno Acessibilidade da estação de bicicleta até as estações de metrô mais próximas Número de pontos de ônibus no entorno da estação
Lin et al. (2019)	Shangai	Presença de terminal de ônibus no entorno da estação Entradas de metrô no entorno da estação de bicicleta Distância média até as 4 estações mais próximas Número de estações de ônibus no entorno da estação Número de viagens no pico da manhã Razão do número de usuários e número de viagens Localização da estação
Europa		
Schimohr; Scheiner (2021)	Colônia	Percentual de tipos de uso do solo no entorno da estação Número de shoppings center, restaurantes e bares no entorno da estação Número de serviços de saúde no entorno da estação Número de escolas infantis no entorno da estação Número de instituições públicas no entorno da estação Número de centros esportivos no entorno da estação Número de atrações turísticas no entorno da estação Número de igrejas no entorno da estação Número de parques infantis no entorno da estação Distância até estações de metrô com mais de uma linha População no entorno da estação Percentual da população em certas faixas de idade
Tran; Ovtracht; d'Arcier (2015)	Lyon	Tamanho da estação Densidade da rede viária Número de habitantes no entorno da estação Número de empregos no entorno da estação Número de estudantes no entorno da estação Número de residência estudantil no entorno da estação Número de estações de trem no entorno da estação Número de restaurantes no entorno da estação Número de cinemas no entorno da estação Presença de via pavimentada ao longo do rio
Radzimski; Dzięcielski (2019)	Poznań	Número de partidas diárias de transporte público Extensão de ciclovias/ciclofaixas no entorno da estação Disponibilidade de vagas na estação Presença de áreas industriais no entorno da estação Presença de shopping center no entorno da estação Presença de parque no entorno da estação Presença de cinema no entorno da estação Presença de corpo d'água no entorno da estação Densidade populacional
Outros		
Chardon; Caruso; Thomas (2017)	Mundialmente	Tipo de operador, privado Densidade de estações População

Fonte: Elaboração própria (2022).

O uso da métrica agregada de demanda dificulta a compreensão de sua variação horária (FAGHIG-IMANI et al., 2014; TRAN; OVTRACHT; D'ARCIER, 2015). Conhecer essa variação é relevante para identificar fatores que propiciam o uso em horários específicos, auxiliando a minimizar o risco de saturação ou esvaziamento de estações e reduzindo o custo da redistribuição de bicicletas, melhorando a qualidade e disponibilidade do serviço (FRICKER; GAST, 2016; TRAN; OVTRACHT; D'ARCIER, 2015). Logo, auxilia no desenho de políticas, como mecanismos de incentivo para que os ciclistas participem do rebalanceamento de bicicletas (FAGHIG-IMANI et al., 2017).

Em localidades na Europa, Ásia e América, foram constatados picos de demanda por bicicletas compartilhadas nos horários de pico da manhã e da noite, e eventualmente, no horário do almoço (FAGHIG-IMANI et al., 2017; GU; KIM; CURRIE, 2019; HU et al., 2021; SCHIMOHR; SCHEINER, 2021; ZHU et al., 2020).

Este comportamento é típico de viagens pendulares, hipótese confirmada por modelos de regressão realizados com dados dos sistemas de Montreal, Santiago e Lyon (FAGHIG-IMANI et al., 2014; MIX; HURTUBIA; RAVEAU, 2022; TRAN; OVTRACHT; D'ARCIER, 2015). Nesses estudos, as variáveis explicativas estatisticamente significativas para este comportamento são socioeconômicas, como número de residentes, empregos e estudantes. Outros estudos com sistemas de compartilhamento na China não utilizam modelos confirmatórios, mas atribuem o comportamento à localização das estações em regiões de uso comercial e residencial (GU; KIM; CURRIE, 2019; ZHU et al., 2020).

O uso no período da tarde e da noite é maior comparativamente aos demais períodos do dia, visto que as viagens nesse horário são estimuladas também por motivos de lazer e gastronomia. Esta hipótese foi confirmada por estudos em Montreal e Lyon, de modo que variáveis explicativas como a presença de restaurantes, cinemas, áreas de recreação são significativas (FAGHIG-IMANI et al., 2014; TRAN; OVTRACHT; D'ARCIER, 2015).

Por outro lado, a ocorrência de demanda mais difusa no tempo foi explicada pela presença de áreas com uso do solo misto ou de diferentes tipos de pontos de interesse, fazendo que o uso das bicicletas compartilhadas seja diverso, tanto para retiradas como devoluções, como verificado em Barcelona e Sevilha (FAGHIG-IMANI et al., 2017). Em Montreal e Suzhou, esse comportamento foi atribuído à presença de centros acadêmicos, já que os horários de aula, em geral, não são fixos (FAGHIG-IMANI et al., 2014; GU; KIM; CURRIE, 2019).

As características do sistema também podem explicar demandas horárias específicas. Em Barcelona e Lyon, o tamanho da estação influencia positivamente nos fluxos de retiradas e devoluções na manhã e à tarde (FAGHIH-IMANI et al., 2017; TRAN; OVTRACHT; D'ARCIER, 2015). Em Sevilha, esta influência é positiva apenas no pico da manhã (FAGHIH-IMANI et al., 2017). Em Montreal, o tamanho influencia negativamente os fluxos, mas a densidade de estações no entorno contribui positivamente com as retiradas e devoluções (FAGHIH-IMANI et al., 2014).

Em termos de geografia, maiores altitudes e elevações em Barcelona, Sevilha e Lyon reduzem os fluxos em ambos os sentidos, e em diferentes horários (FAGHIH-IMANI et al., 2017; TRAN; OVTRACHT; D'ARCIER, 2015). Por outro lado, do ponto de vista da infraestrutura cicloviária, em Montreal, a presença de ciclovias e ciclofaixas no entorno da estação tem impacto significativo nos fluxos de chegada no período da manhã e de retirada no período noturno (FAGHIH-IMANI et al., 2014).

3. OBJETIVOS E CONTRIBUIÇÕES

A Figura 1 apresenta as lacunas de conhecimento e hipóteses, bem como os objetivos e contribuições dessa pesquisa.

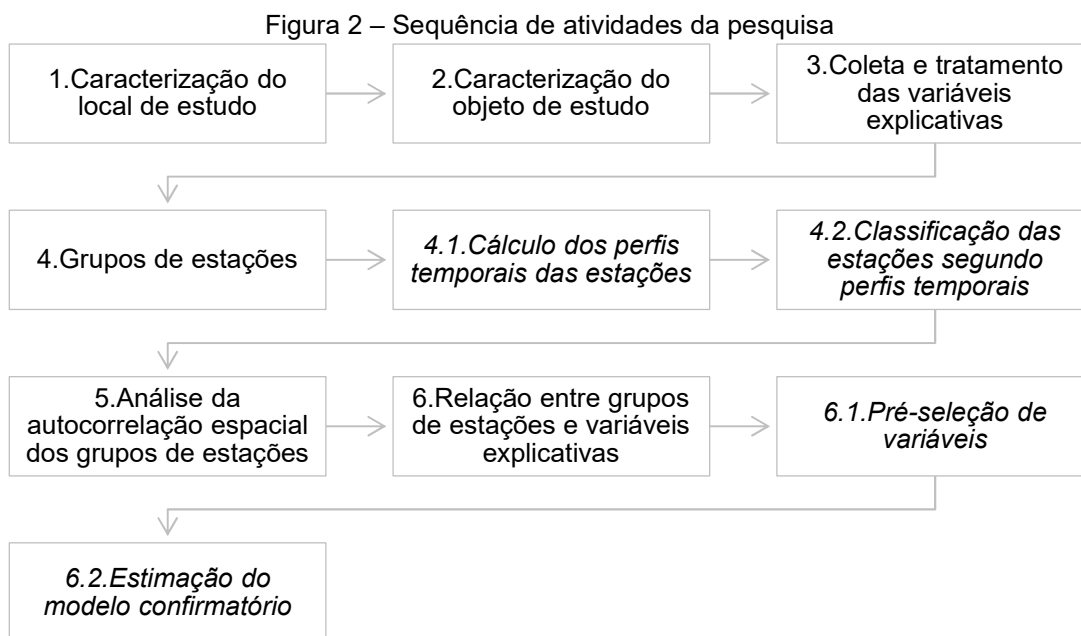
Figura 1 – Lacunas, hipóteses, objetivos e contribuições

Lacunas	Estudos recentes utilizam geralmente uma métrica agregada para análise da demanda e sua relação com atributos operacionais, socioeconômicos e do espaço urbano, e não avaliam a variação horária da demanda
Hipóteses	As estações possuem distintos perfis temporais médios de demanda, que estão associados com atributos operacionais, socioeconômicos e do espaço urbano
Objetivos	<ul style="list-style-type: none"> - Classificar as estações em grupos, a partir dos perfis temporais de demanda - Analisar se os grupos de estações estão distribuídos aleatoriamente no espaço - Estabelecer relações dos perfis temporais de cada grupo com características operacionais, socioeconômicas e do espaço urbano no entorno das estações
Contribuições	<ul style="list-style-type: none"> - Complementar os estudos da variação horária da demanda - Entender o comportamento espaço-temporal da demanda permite aumentar a eficiência operacional de redistribuição de bicicletas - Melhoria do processo de tomada de decisão dos operadores de sistemas de bicicletas compartilhadas, especialmente do ponto de vista do crítico processo de reposicionamento de bicicletas entre estações e do reequilíbrio dos fluxos de retiradas e devoluções. Isto evita que estações fiquem completamente cheias ou vazias, aumentando a confiabilidade do sistema e tornando seu uso mais atrativo - Há poucos estudos sobre o comportamento das bicicletas compartilhadas na América Latina e, especialmente, em uma megalópole como São Paulo, que sofreu rápido processo de urbanização e tem dinâmica de viagens excêntrica

Fonte: Elaboração própria, 2022.

4. MATERIAIS E MÉTODOS

Uma síntese da sequência de atividades está apresentada na Figura 2. De forma complementar, a ferramenta utilizada em cada etapa está apresentada no quadro resumo da Tabela 3.



Fonte: Elaboração própria, 2022.

Tabela 3 – Quadro resumo das atividades e ferramentas utilizadas

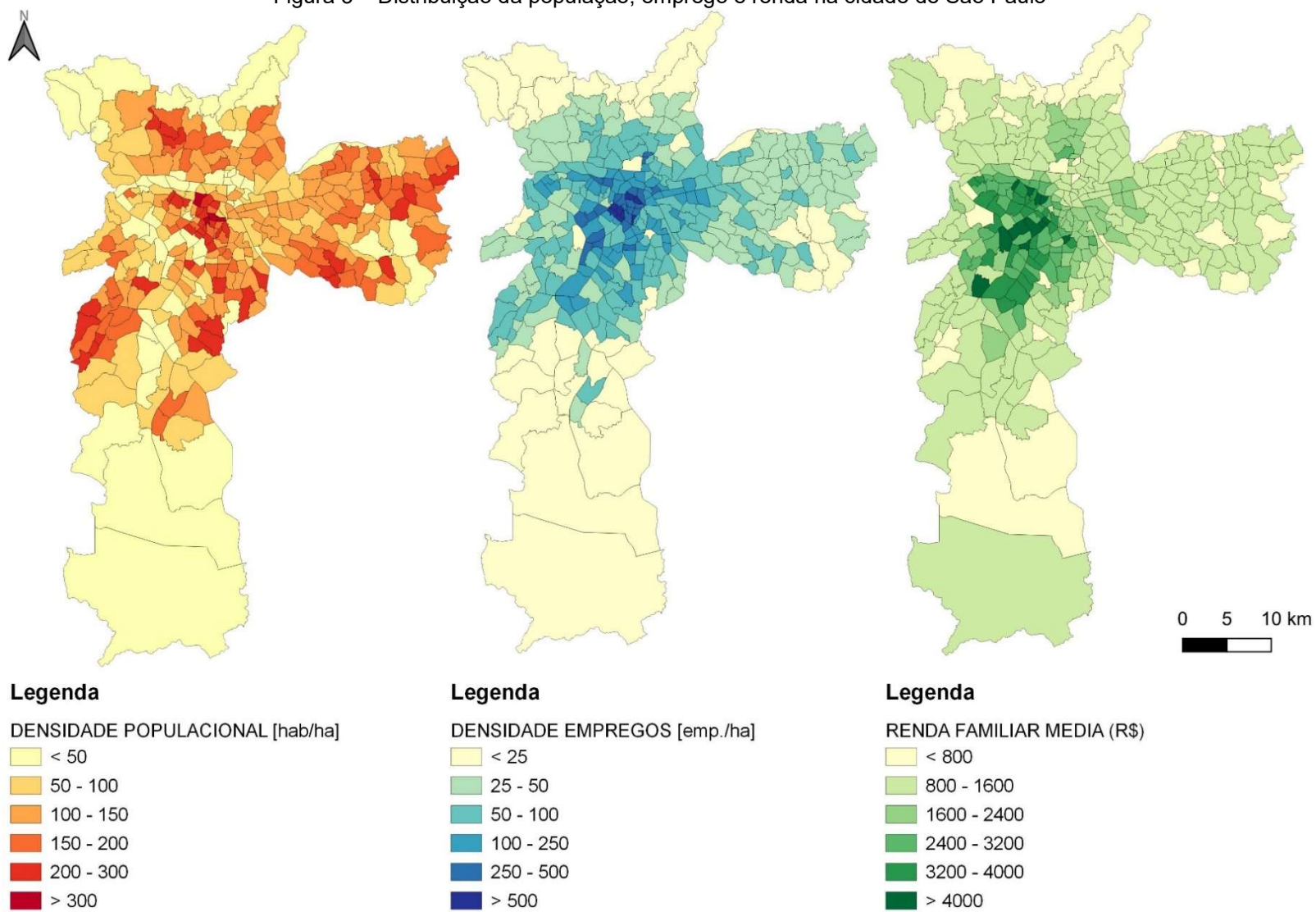
Atividade	Ferramenta
1. Caracterização do local de estudo	-
2. Caracterização do objeto de estudo	-
3. Coleta e tratamento das variáveis explicativas	QGIS
4. Grupos de estações	-
4.1. Cálculo dos perfis temporais das estações	Excel
4.2. Classificação das estações segundo perfis temporais	RStudio (pacote <i>cluster</i>)
5. Análise da autocorrelação espacial dos grupos de estações	GeoDa
6. Relação entre grupos de estações e variáveis explicativas	-
6.1. Pré-seleção de variáveis	Stata e RStudio (pacote <i>car</i>)
6.2. Estimção do modelo confirmatório	GAUSS

Fonte: Elaboração própria, 2022.

4.1. Caracterização do local de estudo

A cidade de São Paulo possui 1.521 km² e cerca de 12 milhões de habitantes (IBGE, 2020), configurando-se como a mais populosa do Brasil e da América Latina. É uma megalópole com uma dicotomia excêntrica entre centro e periferia, e teve um desenvolvimento urbano que resultou em distribuição desigual de oportunidades (BENEDINI; LAVIERI; STRAMBI, 2019; PRITCHARD et al., 2019). A Figura 3 permite observar tanto a dispersão da população, como a concentração espacial dos empregos e das rendas mais altas na cidade de São Paulo.

Figura 3 – Distribuição da população, emprego e renda na cidade de São Paulo

**Legenda**

DENSIDADE POPULACIONAL [hab/ha]

- < 50
- 50 - 100
- 100 - 150
- 150 - 200
- 200 - 300
- > 300

Legenda

DENSIDADE EMPREGOS [emp./ha]

- < 25
- 25 - 50
- 50 - 100
- 100 - 250
- 250 - 500
- > 500

Legenda

RENDA FAMILIAR MEDIA (R\$)

- < 800
- 800 - 1600
- 1600 - 2400
- 2400 - 3200
- 3200 - 4000
- > 4000

Fonte: Elaboração própria, 2022.

O sistema de transporte coletivo da cidade de São Paulo conta com ônibus municipais (SPTRANS) e intermunicipais (EMTU), assim como metrô e trem. O número de linhas, pontos de parada e estações estão na Tabela 4. Quanto à sua configuração espacial, o sistema é predominantemente radial, ligando a área central ao restante do município.

Tabela 4 – Características do sistema de transportes da cidade de São Paulo

Sistema	Linhas	Nº pontos/estações
Ônibus municipal	1.314	19.500
Ônibus intermunicipal	500	19.000
Metrô	5	64
Trem	6	90

Fonte: Adaptado de CIDADE DE SÃO PAULO (2022) e SPTRANS (2022).

As viagens da Região Metropolitana de São Paulo (RMSP) se distribuem nos modos: 36% em transporte coletivo, 31% em motorizados individuais, 32% a pé e 1% em bicicletas (METRÔ-SP, 2020b). Em comparação com o uso de bicicletas próprias, uma maior proporção das viagens com bicicletas compartilhadas é feita de forma combinada com outros modos, particularmente de transporte coletivo. Este é um resultado esperado, uma vez que não é possível, com poucas exceções, transportar a bicicleta dentro dos ônibus, trem ou metrô (BENEDINI, LAVIERI e STRAMBI, 2019).

Sobre infraestrutura cicloviária, São Paulo possui cerca de 700 quilômetros de ciclovias, ciclofaixas e ciclorrotas (CET, 2020). Contudo, 72% das viagens realizadas em bicicletas não utilizam vias segregadas. Os motivos principais para o uso da bicicleta são: trabalho (69%), escola (14%) e lazer (6%). As viagens de bicicletas são realizadas em maioria por homens (90%), de 30 a 49 anos (41%), com 2º grau completo (40%) e renda familiar entre R\$1.908 a R\$3.816² (43%) (METRÔ-SP, 2020a).

4.2. Caracterização do objeto de estudo

Atualmente, há dois sistemas de compartilhamento de bicicletas de estações fixas na cidade de São Paulo, ambos considerados de 3ª geração. Cerca de 93% das estações pertencem ao sistema Bikesampa, associado ao Banco Itaú e operado pela empresa Tembici. Dada a grande representatividade, este sistema foi escolhido como o objeto de estudo da pesquisa.

² Em R\$ de 2018.

No período de referência do estudo, que considerou os dias úteis entre 01/02/2020 e 15/03/2020³, o Bikesampa operava com 279 estações, concentradas nas zonas Oeste, Sul e Centro, logo, não havendo operação nas zonas Norte e Leste de São Paulo. É importante destacar que não foram desconsiderados dias de chuva ou de baixa temperatura, pois devido a concentração das estações, acredita-se que este efeito é transversal para todas as estações.

Para o período de análise, a Tembici disponibilizou um banco de dados de viagens, contendo o dia, a hora e a estação, da retirada, e as mesmas informações para a estação de devolução, bem como a data de nascimento informada pelo usuário. Algumas datas de nascimento não resultam em idades possíveis para ciclistas, de forma que considerou-se o idades iguais ou superiores a 18 anos.

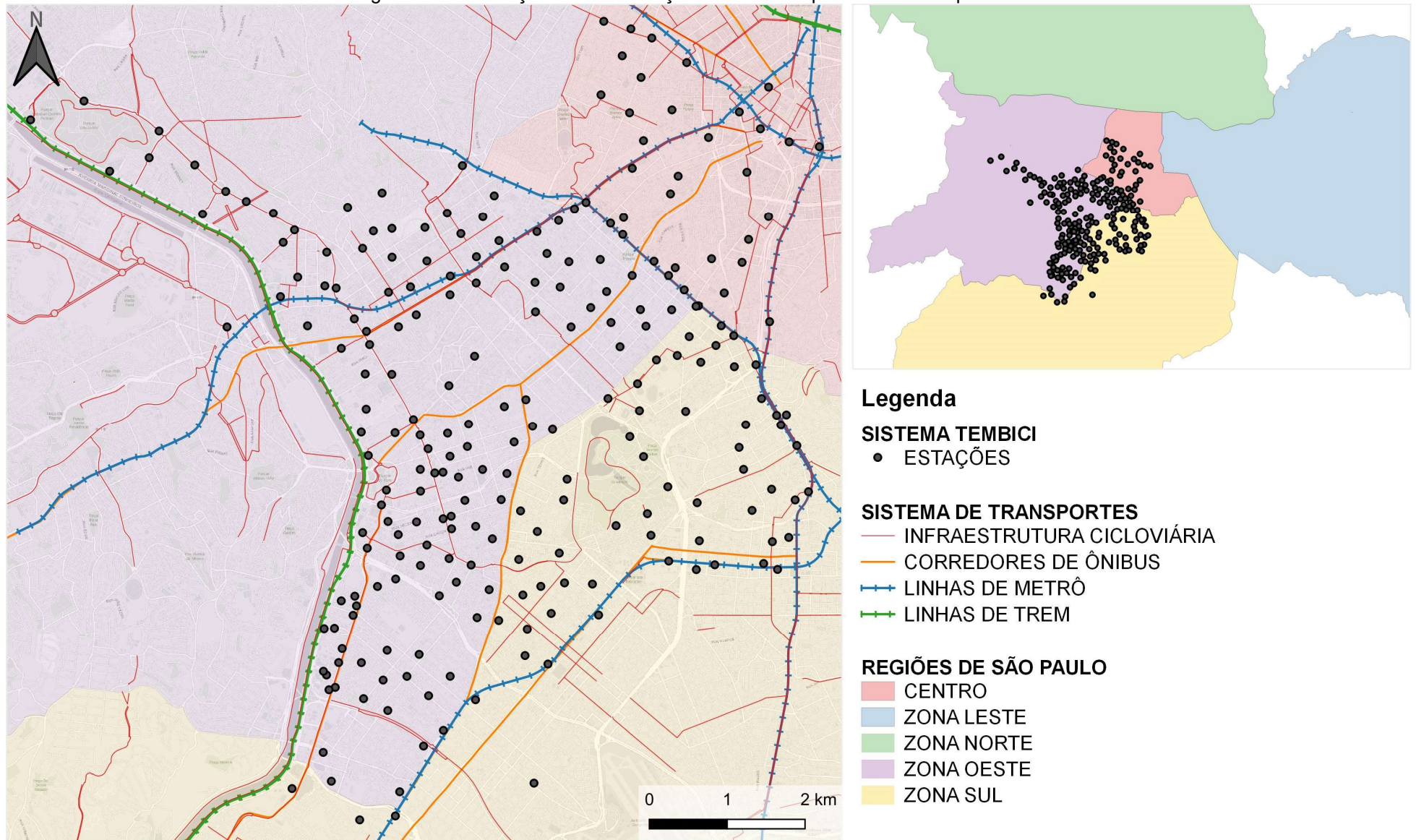
Vale destacar que não há um identificador do usuário, de modo que os indivíduos podem aparecer com frequências distintas no banco de dados. Também foram disponibilizadas, para cada estação, as coordenadas, o número de docas e presença de bolsão⁴. Das 279 estações, foram selecionadas apenas 239 (ver Figura 4), tendo sido excluídas:

- 35 estações que operaram em apenas um dos meses (fevereiro ou março). Esta remoção é essencial, pois o método de construção dos perfis temporais considera o fluxo médio do período de análise. Logo, para essas estações, os fluxos estariam subrepresentados por elas terem operado em um período inferior ao selecionado para análise;
- 1 estação que não possuía informação sobre suas coordenadas e número de docas, impossibilitando o cálculo das variáveis explicativas apresentadas no item 4.3, a seguir;
- 2 estações localizadas no campus da Universidade de São Paulo, onde as características espaciais e socioeconômicas são diferentes do restante da cidade;
- 2 estações localizadas no Largo da Batata, que possuem fluxos e tamanhos muito distintos das demais, e podem ser consideradas como observações discrepantes.

³ A Tembici informou uma queda significativa no número de viagens a partir do dia 16/03/2020 em virtude dos efeitos da pandemia do Covid-19

⁴ Reserva de capacidade de uma estação

Figura 4 – Localização das 239 estações do Bikesampa selecionadas para análise



Fonte: Elaboração própria, 2022.

4.3. Coleta e tratamento das variáveis explicativas

Para identificar as características socioeconômicas, operacionais e do espaço urbano associadas ao comportamento temporal de uso de bicicletas compartilhadas, foram levantadas informações alinhadas com os fatores encontrados na revisão da literatura, as quais foram tratadas para a definição das variáveis explicativas.

4.3.1. As variáveis explicativas

As variáveis explicativas utilizadas nesta pesquisa estão descritas na Tabela 5 abaixo. Elas foram definidas segundo a informação disponível em órgãos públicos e fontes de dados de livre acesso, assim como apresentado no item 4.3.2.

4.3.2. Coleta de informação

A fonte de informação (ver Anexo A) e o ano de referência estão apresentados na Tabela 6. Também, é indicado se a variável é do tipo espacial e, em caso afirmativo, qual a tipologia da informação. É importante destacar:

- Para a variável declividade, a informação utilizada é a inclinação da superfície do terreno em relação à horizontal, desagregada nas classes: (i) até 5%, (ii) 5 a 25%, (iii) 25 a 60% e (iv) acima de 60%. A escolha para o limiar de 5% deve-se ao fato de que esta é a categoria que mais se aproxima do terreno plano, que exige menos esforço físico do ciclista;
- As variáveis de uso do solo foram calculadas a partir da informação do uso do solo predominante da quadra, e não com base no uso do solo específico de cada lote. Segundo adotado pelo Geosampa, o uso do solo é considerado predominante se sua participação é maior ou igual a 60% na quadra. Logo, é possível que existam quadras em que os usos estejam distribuídos de forma que nenhum seja predominante. Esse procedimento pode não representar adequadamente categorias de uso do solo menos frequentes;
- As variáveis relativas à idade foram calculadas individualmente para cada usuário, visto que esta informação estava disponível no banco de dados fornecido pela Tembici.
- As variáveis relativas à renda foram obtidas da POD de 2017 e correspondem à zona OD em que está localizada a estação.

Tabela 5 – Descrição das variáveis explicativas

Categoria	Variável explicativa	Descrição
Características do sistema	dist_estacao_proxima	Distância da estação até a estação mais próxima, em quilômetros
	dist_média	Distância média entre a estação e todas as outras estações, em quilômetros
	tamanho	Número de docas da estação
Características de demanda	Bolsão	Presença de bolsão na estação (sim: presente, não: ausente)
	dur_origem	Média das durações de viagens com origem na estação, em minutos
	dur_destino	Média das durações de viagens com destino na estação, em minutos
Infraestrutura cicloviária	bicicletário	Presença de bicicletários na área de serviço da estação (sim: presente, não: ausente)
	dist_vias_cicláveis	Distância até a ciclovia mais próxima, em quilômetros
Sistema de transportes	Metrô	Presença de estação de metrô na área de serviço da estação (sim: presente, não: ausente)
	CPTM	Presença de estação da CPTM na área de serviço da estação (sim: presente, não: ausente)
	pontos_ônibus	Número de pontos de ônibus na área de serviço da estação
	vias_ônibus	Extensão de corredores e faixas de ônibus na área de serviço da estação, em quilômetros
	estacionamento	Número de estabelecimentos de estacionamento na área de serviço da estação
	zona_azul	Presença de zona azul na área de serviço da estação (sim: presente, não: ausente)
Pontos de interesse	parque	Presença de parque na área de serviço da estação (sim: presente, não: ausente)
	museu	Presença de museus na área de serviço da estação (sim: presente, não: ausente)
	vida_noturna	Número de teatros, casas de show e cinemas na área de serviço da estação
	esporte	Presença de clubes e estádios na área de serviço da estação (sim: presente, não: ausente)
	shopping	Presença de shopping center na área de serviço da estação (sim: presente, não: ausente)
	mercado	Número de mercados na área de serviço da estação
	estudo	Presença de universidade na área de serviço da estação (sim: presente, não: ausente)
Geografia	restaurantes	Número de restaurantes na área de serviço da estação
	declividade	Percentual da área de serviço da estação com baixa declividade (<5%)
Características socioeconômicas	perc_idade_18_25	Percentual de usuários com idade entre 18 e 25 anos da estação
	perc_idade_26_36	Percentual de usuários com idade entre 26 e 36 anos da estação
	perc_idade_37_55	Percentual de usuários com idade entre 37 e 55 anos da estação
	perc_2_5_SM	Percentual da população residente na zona OD da estação com renda familiar entre 2 e 5 salários mínimos
	perc_6_SM	Percentual da população residente na zona OD da estação com renda familiar acima de 6 salários mínimos
Uso do solo predominante	residencial	Percentual do uso do solo residencial predominante na quadra, na área de serviço da estação
	comercial	Percentual do uso do solo comercial predominante na quadra, na área de serviço da estação
	industrial	Percentual do uso do solo industrial predominante na quadra, na área de serviço da estação
	uso_misto	Percentual do uso do solo predominante de uso misto na quadra, na área de serviço da estação
	equipamentos_públicos*	Percentual do uso do solo predominante de equipamentos públicos na quadra, na área de serviço da estação
	outros**	Percentual do uso do solo de outros tipos predominante na quadra, na área de serviço da estação

*Instalações destinadas à educação, saúde, cultura, assistência social, esportes, lazer, segurança pública, abastecimento, serviços funerários.

**Outros: usos do solo sem predominância, sem informação, escolas, garagens, terrenos vagos e outros (não especificados pelo GeoSampa).

Fonte: Elaboração própria, 2022.

Tabela 6 – Fonte de informação e anos de coleta e referência das variáveis explicativas

Categoria	Variável explicativa	Espacial*	Fonte	Ano referência
Características do sistema	dist_estação_próxima	-	Tembici	2021
	dist_média	-		
	tamanho	-		
	bolsão	-		
Características de demanda	dur_origem	-	Tembici	2020
	dur_destino	-		
Infraestrutura cicloviária	bicicletário	Ponto	Ciclomapa	2021
	dist_vias_cicláveis	Linha		
Sistema de transportes	metrô	Ponto	Geosampa	2020
	CPTM	Ponto	Geosampa	2020
	pontos_ônibus	Ponto	SPTRANS, EMTU	2020
	vias_ônibus	Linha	Geosampa	2016/2017
	estacionamento	Ponto	CEM	1999
	zona_azul	Linha	Geosampa	2017
	Pontos de interesse	parque	Polígono	Geosampa
museu		Ponto	Geosampa	2018
vida_noturna		Ponto	Geosampa	2018
esporte		Ponto	Geosampa	2018
shopping		Polígono	Geosampa	2017
mercado		Ponto	CEM	1999
estudo		Ponto	CEM	1999
restaurantes		Ponto	CEM	1999
Geografia		declividade	Polígono	Geosampa
Características socioeconômicas	perc_idade_18_25	-	Tembici	2020
	perc_idade_26_36	-		
	perc_idade_37_55	-		
	perc_2_5_SM	-	POD	2017
	perc_6_SM	-		
Uso do solo predominante	residencial	Polígono	Geosampa	2016
	comercial	Polígono		
	industrial	Polígono		
	uso_misto	Polígono		
	equipamentos_públicos	Polígono		
	outros	Polígono		

*Ponto é a entidade geográfica representada por um par de coordenadas, linha é a entidade geográfica representada por no mínimo dois pontos e polígono é uma série de coordenadas que formam segmentos de linhas que fecham uma área.

Fonte: Elaboração própria, 2022.

4.3.3. Tratamento dos dados

As variáveis que não se enquadram como dados espaciais (variáveis de operação do sistema e socioeconômicas) foram calculadas no Excel. As demais informações coletadas são georreferenciadas, de modo que o tratamento e cálculo de variáveis foi realizado no *software* QGIS. Estas variáveis podem ser divididas em dois grupos, apresentados a seguir.

4.3.3.1. Variáveis de distância

O primeiro grupo diz respeito a variáveis de distâncias. Primeiramente, para o cálculo dessas variáveis, as informações georreferenciadas foram reprojetadas para o sistema de referência de coordenadas SIRGAS2000 UTM 23S, de forma que as coordenadas fossem transferidas para o plano cartesiano, permitindo o cálculo das distâncias em quilômetros. Para as variáveis de distância do mais próximo (*dist_estação_próxima* e *dist_vias_cicláveis*), foi utilizada a ferramenta “*Distância para o ponto central mais próximo*” do QGIS. Por outro lado, para a variável de distância média, foi utilizada a ferramenta “*Matriz de distância*”. É importante destacar que todas as distâncias são euclidianas.

4.3.3.2. Variáveis de presença, contagem e percentual

O segundo grupo de variáveis considera a área de serviço da estação.

Área de serviço

Para a qualificação dos serviços e das características na vizinhança da estação, utilizou-se o conceito de área de serviço, em que foi considerada a área compreendida em uma distância na rede caminhável de 500 metros a partir da estação. Embora a literatura indique distância de acesso às estações entre 200 e 500 metros, este último valor é mais frequente (EREN; UZ, 2020; GU; KIM; CURRIE, 2019; ITDP, 2014; RADZIMSKI; DZIĘCIELSKI, 2021) e foi aqui adotado para representar a disposição de caminhada do pedestre até a estação.

É importante destacar que os estudos geralmente utilizam áreas circulares (obtidas com um raio fixo, usualmente chamadas de “*buffers*”) para a representação da vizinhança. A área de serviço, por usar a distância em rede ao invés de uma distância euclidiana, tem como vantagem a captura do comportamento real do pedestre em seu percurso de acesso à estação. Por outro lado, tem como desvantagem a necessidade de contar com a descrição de uma rede caminhável (LANDEX; HANSEN; ANDERSEN, 2006).

Até o desenvolvimento desta pesquisa, não estava disponível publicamente uma rede caminhável da cidade de São Paulo. Para sua elaboração, partiu-se da rede de

logradouros de São Paulo de 2020 disponibilizada pelo CEM, enriquecida por procedimentos manuais, utilizando como referência o mapa de fundo do *Google Maps*.

O primeiro procedimento foi a inclusão de passarelas sob a forma de segmentos lineares (*links*) em nós (*nodes*) específicos, a partir da informação de passarelas disponibilizada pelo Geosampa. Em seguida, foram incluídas as rotas caminháveis de grandes áreas, como parques municipais. Por fim, como as vias de maiores velocidades (ex. Hélio Pelegrino, República do Líbano, Faria Lima, etc.) estavam representadas por dois segmentos (um para cada sentido do tráfego), foram incluídas as faixas de travessia para pedestres.

Finalmente, as áreas de serviço foram calculadas com a ferramenta “área de serviço”, a qual tem como resultado os pontos de intersecção da rede no limite de 500 metros da estação. Para a obtenção do polígono da área de serviço, foi utilizada a ferramenta “*polígonos convexos*”, que faz a ligação dos pontos resultantes da etapa anterior. A Tabela 7 apresenta a estatística descritiva das áreas de serviço calculadas, em hectares.

Tabela 7 – Áreas de serviço das estações [ha]

Variável	Área de serviço				
	Mínimo	Média	Mediana	Máximo	Desvio Padrão
área_serviço	19,321	46,381	47,797	56,985	5,730

Fonte: Elaboração própria, 2022.

Cálculo das variáveis

Para as variáveis que representam presença, número de observações ou extensão (categorias de sistemas de transporte e pontos de interesse), utilizou-se a ferramenta “*contar pontos no polígono*”, se a entidade geográfica é ponto, ou a ferramenta “*extensão no polígono*” depois da reprojeção para o sistema cartesiano, se linha.

As variáveis que representam percentual da área de serviço (declividade e uso do solo) foram calculadas realizando primeiramente uma intersecção entre os dados de base e as áreas de serviço, usando a ferramenta “*intersecção*”, sendo que a proporção da área foi calculada subsequentemente no Excel.

4.3.4. Estatística descritiva das variáveis explicativas

A Tabela 8 apresenta o resumo estatístico das variáveis explicativas.

Tabela 8 – Estatística descritiva das variáveis explicativas

Grupo	Variável explicativa	Unidade	Variáveis quantitativas					Variáveis qualitativas	
			Mínimo	Média	Mediana	Máximo	Desvio Padrão	não	sim
Características do sistema	dist_estação_próxima	quilômetro	0,032	0,306	0,294	1,310	0,131	-	-
	dist_média	quilômetro	2,897	3,962	3,725	7,570	0,833	-	-
	tamanho	unidade	7,000	16,649	15,000	47,000	5,499	-	-
	bolsão	binária	-	-	-	-	-	220	19
Características de demanda	dur_origem	minutos	10,548	22,048	21,841	41,866	4,381	-	-
	dur_destino	minutos	9,909	22,394	22,375	40,001	4,780	-	-
Infraestrutura cicloviária	bicicletário	unidade	-	-	-	-	-	211	28
	dist_vias_cicláveis	quilômetro	0,000	0,269	0,163	1,523	0,301	-	-
Sistema de transportes	metrô	binária	-	-	-	-	-	179	60
	CPTM	binária	-	-	-	-	-	229	10
	pontos_ônibus	unidade	0,000	18,762	15,000	120,000	16,177	-	-
	vias_ônibus	unidade	0,000	1,284	0,896	4,334	1,221	-	-
	estacionamento	unidade	0,000	16,326	12,000	82,000	13,750	-	-
	zona_azul	binária	-	-	-	-	-	30	209
Pontos de interesse	parque	binária	-	-	-	-	-	201	38
	museu	binária	-	-	-	-	-	188	51
	vida_noturna	unidade	0,000	3,983	2,000	30,000	4,941	-	-
	esporte	binária	-	-	-	-	-	200	39
	mercado	binária	0,000	27,226	26,000	127,000	18,576	-	-
	shopping	unidade	-	-	-	-	-	176	63
	estudo	binária	-	-	-	-	-	119	120
	restaurantes	unidade	0,000	66,301	58,000	379,000	48,874	-	-
Geografia	declividade	percentual	0,106	0,750	0,871	1,000	0,269	-	-
Características socioeconômicas	perc_idade_18_25	percentual	0,084	0,308	0,295	0,637	0,092	-	-
	perc_idade_26_36	percentual	0,179	0,465	0,474	0,648	0,080	-	-
	perc_idade_37_55	percentual	0,049	0,202	0,192	0,509	0,066	-	-
	perc_2_5_SM	percentual	0,064	0,299	0,257	0,699	0,132	-	-
	perc_6_SM	percentual	0,185	0,666	0,693	0,936	0,159	-	-
Uso do solo fiscal	residencial	percentual	0,000	0,430	0,405	0,965	0,242	-	-
	comercial	percentual	0,000	0,125	0,085	0,782	0,140	-	-
	industrial	percentual	0,000	0,002	0,000	0,214	0,016	-	-
	uso_misto	percentual	0,000	0,219	0,196	0,642	0,145	-	-
	equipamentos_públicos	percentual	0,000	0,057	0,000	1,000	0,143	-	-
	outros	percentual	0,000	0,168	0,101	0,888	0,175	-	-

Fonte: Elaboração própria, 2022.

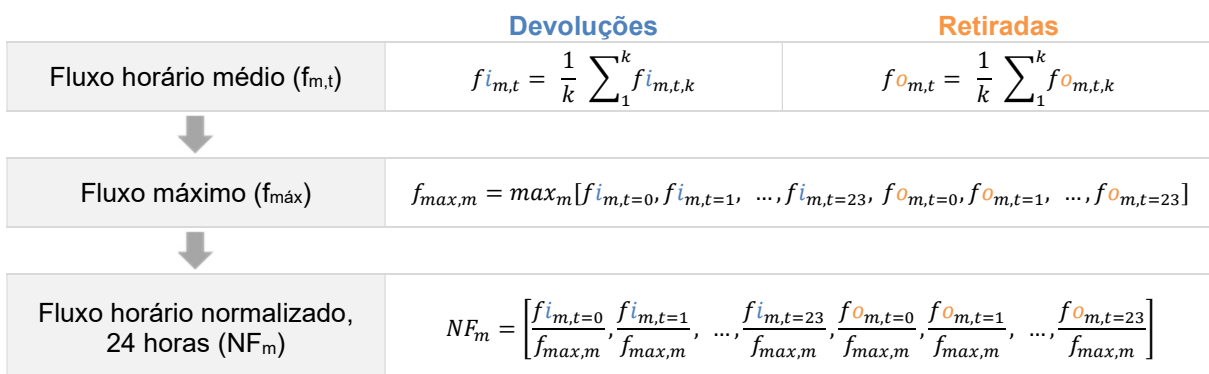
4.4. Grupos de estações

Após coletadas e tratadas as variáveis apresentadas no capítulo anterior, primeiramente foram construídos os perfis temporais médios de uso das estações, que serviram de base para a segunda etapa, o agrupamento das estações.

4.4.1. Cálculo dos perfis temporais médios das estações

Para entender o comportamento temporal das estações, foi utilizado o conceito de perfis horários médios de uso das estações segundo a metodologia proposta por Gu, Kim e Currie (2019), apresentada na Figura 5 e descrita a seguir.

Figura 5 – Passos para o cálculo dos perfis temporais de uso das estações



k: número de dias do período em análise, t: hora do dia, m: estação, i/o: devolução ou retirada

Fonte: Adaptado de Gu, Kim e Currie (2019).

Primeiro, calcula-se o fluxo horário médio (f), o qual corresponde à média ao longo do período de análise do número de bicicletas retornando ($f_{i_{m,t}}$) ou saindo ($f_{o_{m,t}}$) de uma dada estação para cada hora do dia. Como as estações possuem diferentes ordem de grandeza de fluxos, para que os perfis da variação dos fluxos horários médios ao longo do dia possam ser comparados, é realizada uma padronização, dividindo-se o fluxo horário médio pelo fluxo máximo diário ($f_{máx}$), seja de devolução ou retirada. Por fim, cria-se o vetor NF_m de 24 pares de fluxos médios horários normalizados (total de 48 variáveis), com valor entre 0 e 1.

4.4.2. Classificação das estações segundo perfis temporais

Uma vez definidos os vetores NF_m , Gu, Kim e Currie (2019) aplicaram uma técnica de análise de agrupamentos (clusterização), com o objetivo de identificar grupos de perfis temporais de uso das estações. A clusterização é uma técnica multivariada não-supervisionada que utiliza variáveis métricas ou binárias para agrupar observações, tendo como objetivo aumentar a homogeneidade interna de

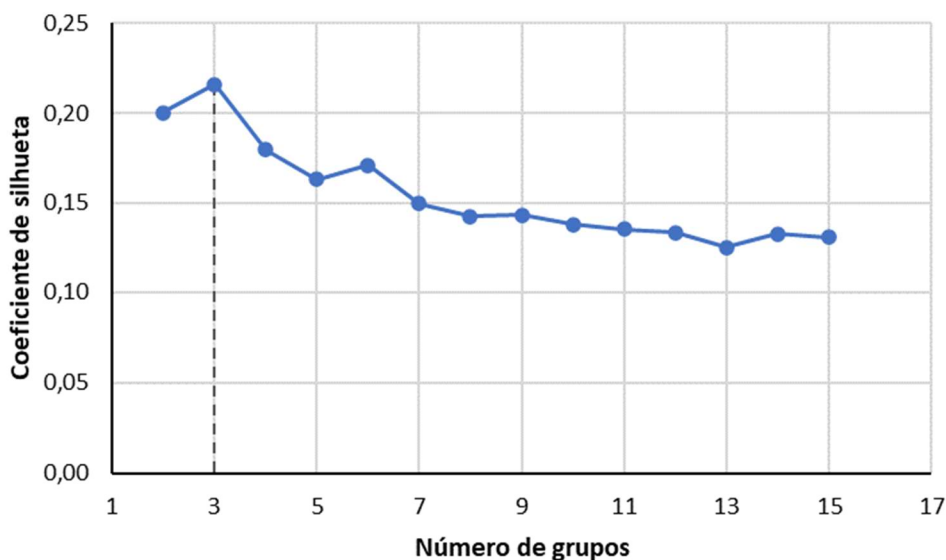
cada grupo e a heterogeneidade entre grupos (MACQUEEN, 1967; FÁVERO; BELFIORE, 2017). Em particular, a clusterização k-médias tem como objetivo a minimização da variância intragrupo (MACQUEEN, 1967).

Nas técnicas de agrupamento não-hierárquicas como a k-médias, o número de grupos é definido a priori. Para esta definição, Gu, Kim e Currie (2019) utilizaram o coeficiente de silhueta, medida de coesão que compara a distância média entre todos os elementos em um grupo e a distância média para todos os elementos em cada um dos outros grupos (ROUSSEEUW, 1987).

Este procedimento foi aplicado aos perfis temporais médios das estações do Bikesampa. Foram consideradas apenas as retiradas e devoluções realizadas entre 06:00 e 22:59, reduzindo a dimensão do vetor NF_m a 34 (ver Anexo B). A clusterização k-médias e curva de silhueta foram executadas na linguagem R no *software* RStudio com as funções *kmeans* e *silhouette*, respectivamente, do pacote *cluster*.

Para o estudo de caso, a curva de silhueta calculada está ilustrada na Figura 6. A medida de coesão é máxima na utilização de três grupos, sendo este número escolhido para classificar os perfis pela técnica k-médias.

Figura 6 – Curva de silhueta para diferentes valores de número de grupos



Fonte: Elaboração própria, 2022.

Portanto, as 239 estações do Bikesampa foram divididas em três grupos de perfis temporais (ver Figura 7). A localização das estações dos diferentes grupos no espaço urbano da região de interesse está apresentada na Figura 8 e na Figura 9. As características dos grupos são:

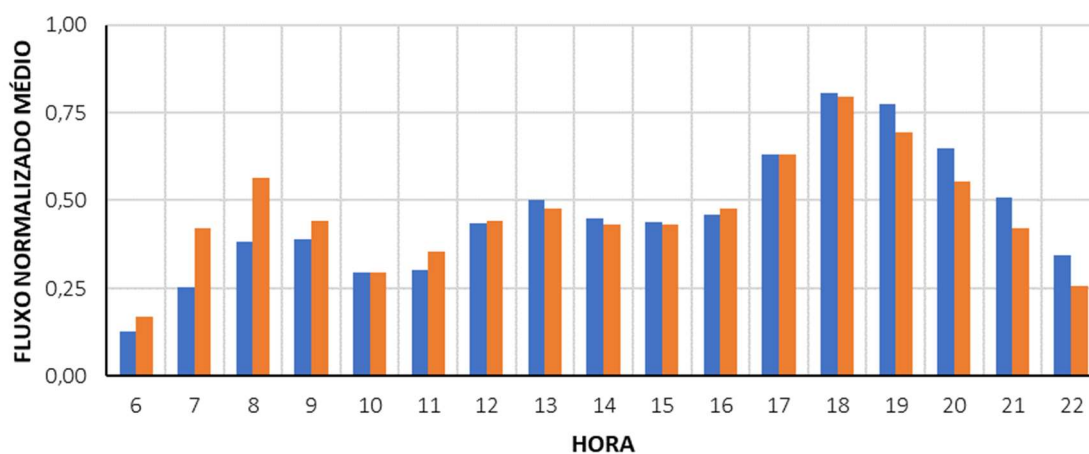
- **Grupo 1 – Balanceado (G1):** fluxo normalizado médio de retiradas e devoluções mais equilibrado ao longo do dia, exceto no pico da manhã, quando o fluxo de retiradas é ligeiramente maior. Há um aumento no fluxo entre 12 e 13 horas, e principalmente a partir do fim da tarde. O maior fluxo normalizado médio é igual a 0,80, e acontece às 18 horas. Pertencem a este grupo 95 estações (39,7% do sistema). As estações estão concentradas no eixo da Av. Paulista e Av. Rebouças, e proximidade do Parque Ibirapuera, regiões com oferta de serviços e outras atividades, mas também moradia;

- **Grupo 2 – Retiradas manhã (G2):** fluxo normalizado médio de retiradas e devoluções desbalanceados entre os períodos da manhã e da noite. Na manhã, há um fluxo maior de retiradas, atingindo um pico de fluxo normalizado igual a 0,62 às 08 horas. À noite, o fluxo de devoluções é mais expressivo, atingindo um pico de 0,73 às 18 horas. Pertencem a este grupo 87 estações (36,4% do sistema). A maioria das estações estão na periferia do sistema, em regiões residenciais;

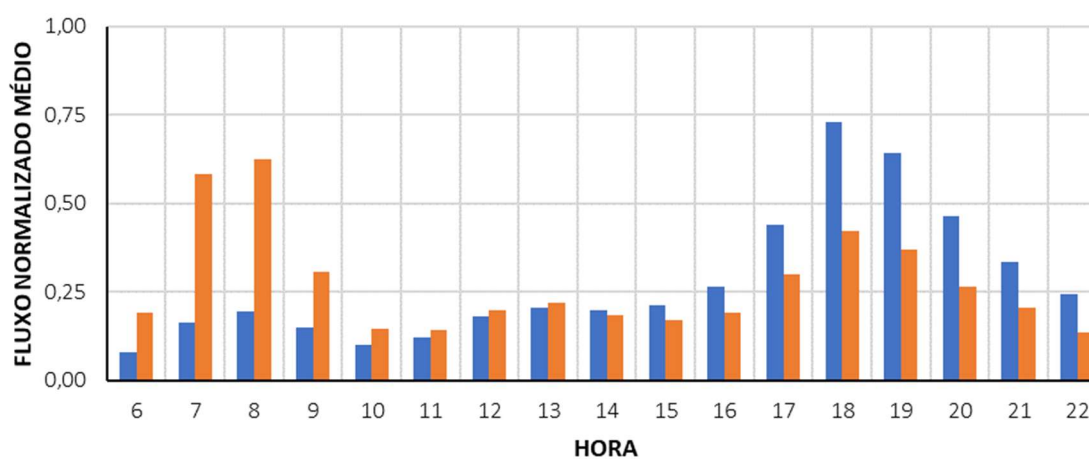
- **Grupo 3 – Devoluções manhã (G3):** fluxo normalizado médio de retiradas e devoluções desbalanceados entre os períodos da manhã e da noite. Na manhã, o fluxo maior é de devoluções, atingindo um pico de 0,88 às 08 horas. À noite, o fluxo de retiradas é maior, com um máximo de 0,62 às 18 horas. Pertencem a este grupo 57 estações (23,9% do sistema). Essas estações estão concentradas no eixo da Av. Faria Lima e Av. Berrini, regiões comerciais e de empregos em serviço.

Em termos de localização no espaço urbano, verifica-se que os grupos de estações parecem não se distribuir aleatoriamente. As estações do **Grupo 2 – Retiradas manhã** aparentam estar localizadas na periferia do sistema. Por outro lado, as estações do **Grupo 3 – Devoluções manhã** parecem ser mais concentradas. A hipótese de aleatoriedade espacial é verificada no item 4.5 a seguir, através de uma análise de autocorrelação espacial.

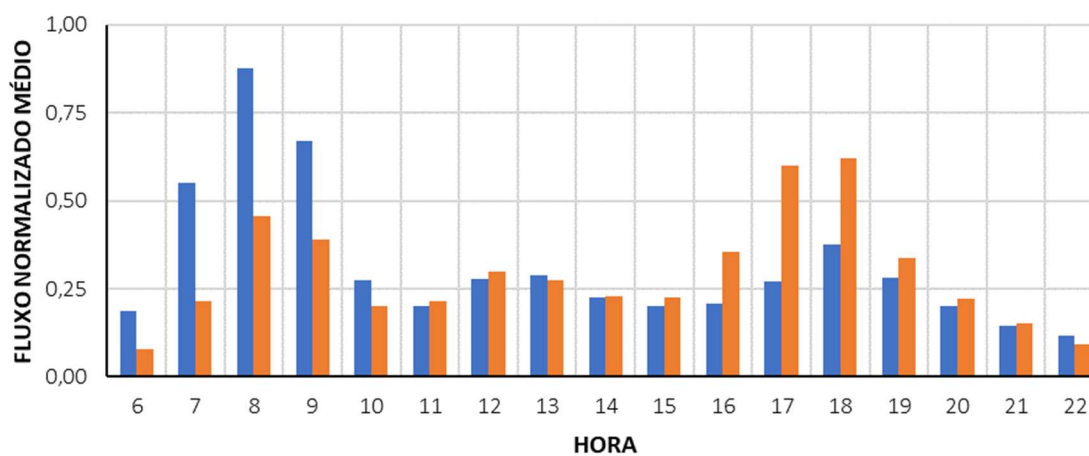
Figura 7 – Perfis temporal de uso de cada grupo de estações



(a) Grupo 1 – Balanceado



(b) Grupo 2 – Retiradas manhã

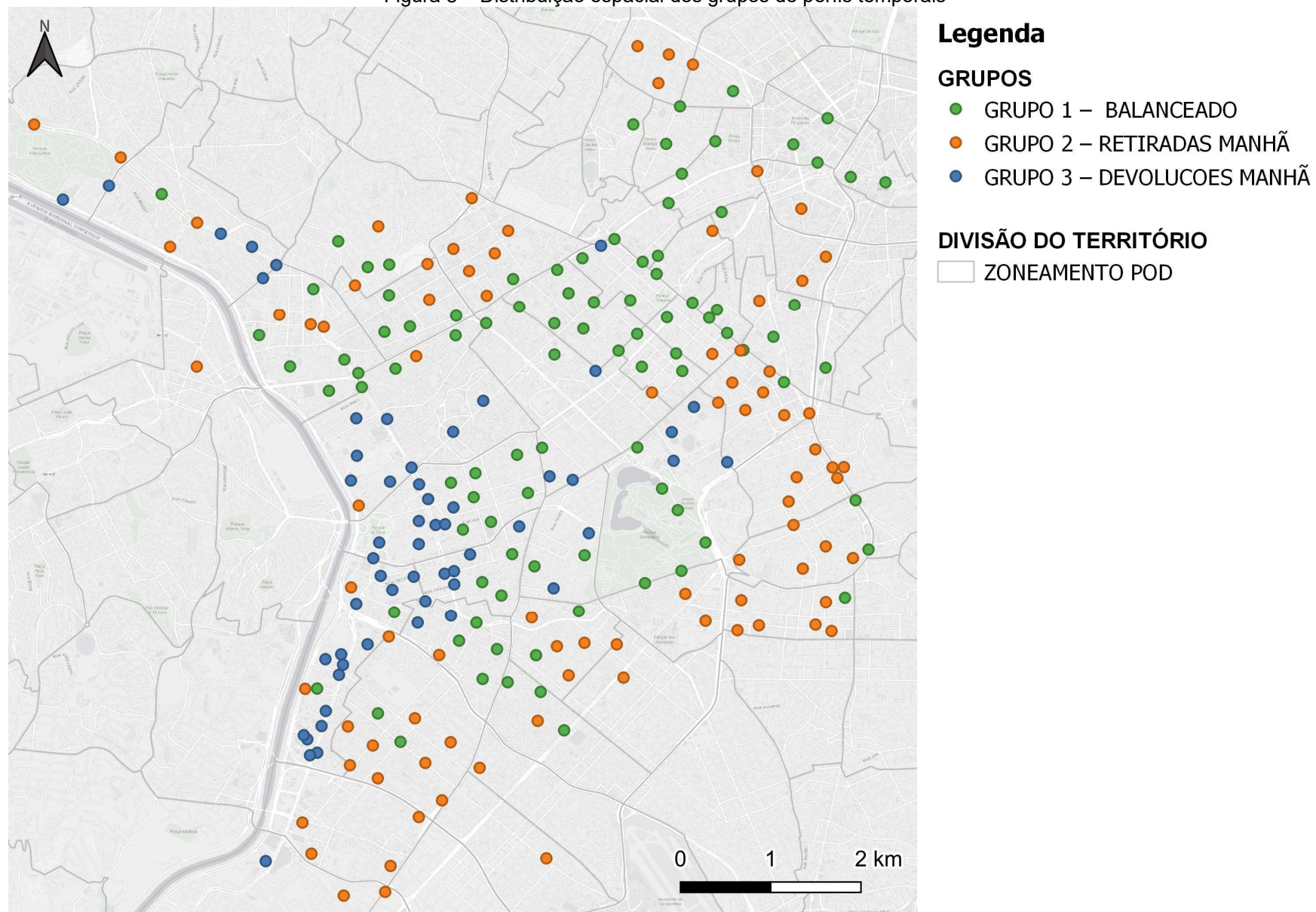


(c) Grupo 3 – Devoluções manhã

Legenda
■ Devolução
■ Retirada

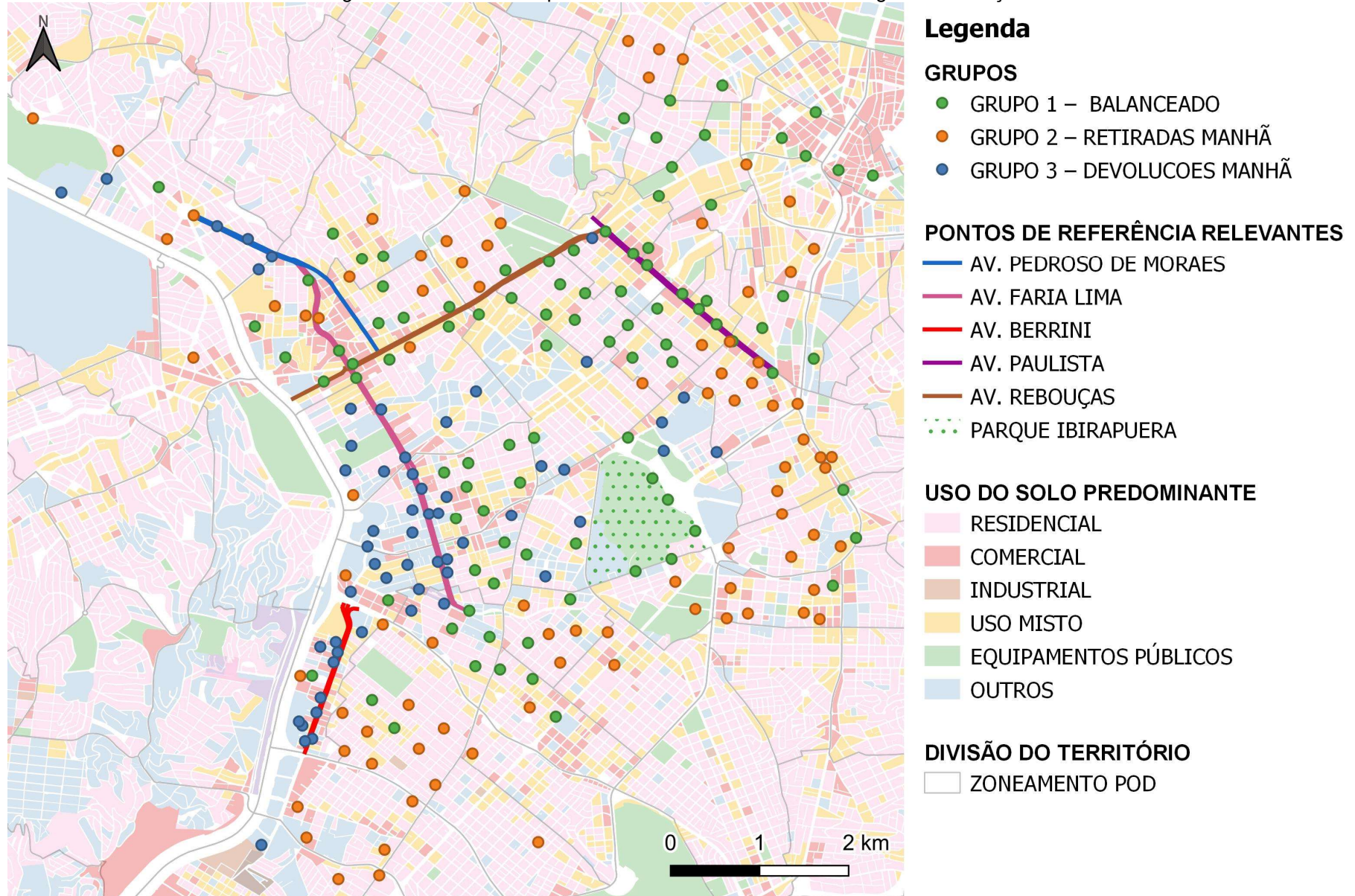
Fonte: Elaboração própria, 2022.

Figura 8 – Distribuição espacial dos grupos de perfis temporais



Fonte: Elaboração própria, 2022.

Figura 9 – Uso do solo e pontos de referência relevantes na região das estações



Fonte: Elaboração própria, 2022.

Comparativamente ao estudo de Gu, Kim e Currie (2019), o agrupamento obtido demonstra similaridade com os resultados de Suzhou, na China, visto que naquele estudo a medida de coesão também é máxima com três grupos. No entanto, os padrões temporais não são iguais. Embora ambos os sistemas tenham um grupo que apresentam aumento gradativo dos fluxos ao longo do dia, em Suzhou, este grupo é desbalanceado e conta com apenas uma estação. Por outro lado, em São Paulo, o **Grupo 1 – Balanceado** tem perfil equilibrado e detém cerca de 40% das estações.

Os dois outros grupos apresentam padrão temporal de dois picos, assim como observado em Suzhou. Contudo, em São Paulo, ambos os padrões são desbalanceados. Este comportamento pode ser explicado por diferentes configurações da estrutura urbana de São Paulo e Suzhou. A cidade de São Paulo possui polos de trabalho e residência concentrados em áreas específicas da cidade.

Os resultados apontam picos da manhã e da tarde entre 07 e 09 horas e 17 e 19 horas, respectivamente. À título de comparação, os picos de demanda de veículos motorizados na RMSP ocorrem entre 06 e 07 horas na manhã, e 17 e 18 horas à tarde (METRÔ-SP, 2020b). Logo, os picos de demanda de bicicletas compartilhadas ocorrem ligeiramente após os picos do modo motorizado.

Comparativamente aos resultados da literatura internacional, os picos de demanda encontrados para o sistema Bikesampa são coincidentes com os estudos que foram realizados em Santiago e Lyon (MIX; HURTUBIA; RAVEAU, 2022; TRAN; OVTRACHT; D'ARCIER, 2015). Por outro lado, em Suzhou o pico da tarde é mais cedo, entre 16 e 19 horas (GU; KIM; CURRIE, 2019), enquanto as cidades de Barcelona e Sevilha possuem picos de demanda mais prolongados, se estendendo até as 10 horas na manhã e até as 22 horas à noite (FAGHIH-IMANI et al., 2017).

Embora a literatura aponte fluxos maiores no período noturno, apenas dois grupos apresentam esse comportamento em São Paulo. As estações do **Grupo 3 – Devoluções manhã** apresentam um forte fluxo de devoluções na manhã, superando inclusive os valores máximos dos outros grupos em qualquer outra hora do dia. Uma hipótese para este comportamento deve-se à proximidade espacial destas estações, permitindo a realização de mais viagens em menores intervalos de tempo.

Nesse sentido, entender o padrão horário dos fluxos de retiradas e devoluções, sobretudo nos picos, é importante para prevenir que estações fiquem cheias ou vazias, desmotivando ou inibindo o uso das bicicletas compartilhadas.

Em termos práticos, os grupos com desequilíbrio podem ser alvo de reposicionamento de bicicletas. As estações do **Grupo 2 – Retiradas manhã** poderiam receber bicicletas de remanejamento na manhã, quando a demanda é maior e há maior probabilidade de as estações deste grupo estarem vazias. Inversamente, as estações do **Grupo 3 – Devoluções manhã** podem receber bicicletas de remanejamento à tarde. É esperado que as estações do **Grupo 1 – Balanceado** sejam menos dependentes de operações de reposicionamento de bicicletas.

Também, poderiam ser adotadas políticas de remanejamento entre estações executadas por parte do usuário, de modo que o ciclista seja estimulado a reposicionar a bicicleta. Um exemplo de medida seria a implantação de preço reduzido para aqueles que realizam viagens no sentido do contrafluxo no horário de pico e que, portanto, auxiliam a mitigar a falta de bicicletas ou vagas nas estações dos grupos com desequilíbrio.

4.5. Análise da autocorrelação espacial dos grupos de estações

Conforme ilustrado na Figura 8, as estações de cada grupo não parecem estar distribuídas aleatoriamente no espaço. Para confirmar essa hipótese, foi utilizado o Índice de Moran, que é uma medida de autocovariância espacial, dada pela Equação 1. Este indicador tem um valor esperado $I_{exp} = -(n-1)^{-1}$ (MORAN, 1948; ALMEIDA, 2012). As hipóteses nula (H_0) e alternativa (H_1) estão apresentadas na Tabela 9.

$$I = \frac{n \sum_i \sum_j w_{ij} z_i z_j}{S_0 \sum_{i=1}^n z_i^2} \quad (1)$$

Onde

n: número de regiões

z: valores da variável de interesse padronizada

w_{ij} : valores da matriz de peso W referente à região i e à região j

S_0 : soma dos elementos da matriz de peso W

Tabela 9 – Hipóteses H_0 e H_1 do I de Moran

Hipótese	Condição	Descrição
H_0	$I \leq I_{exp}$	O padrão espacial observado é igualmente provável comparativamente a qualquer outro – aleatoriedade espacial
H_1	$I > I_{exp}$	Autocorrelação espacial positiva (similaridade entre o valor do atributo e a localização espacial) ou negativa (caso contrário)

Fonte: Adaptado de Moran (1948) e Almeida (2012).

Para visualizar a autocorrelação espacial em cada região, utiliza-se comumente os mapas de *cluster* e significância LISA (ANSELIN, 1995). Estes mapas são coloridos nas regiões com estatísticas locais significativas (ver categorização da Tabela 10).

Tabela 10 – Classificação LISA

Classe	Descrição
Alto-Alto	A região e suas vizinhas apresentam alta concentração do fenômeno
Baixo-Baixo	A região e suas vizinhas apresentam baixa concentração do fenômeno
Baixo-Alto	A região apresenta baixa concentração do fenômeno, enquanto a vizinhança apresenta comportamento inverso
Alto-Baixo	A região apresenta alta concentração do fenômeno, enquanto a vizinhança apresenta comportamento inverso

Fonte: Adaptado de Anselin (1995) e Almeida (2012).

O I de Moran foi calculado independentemente para cada grupo de estações no *software* GeoDa. A variável de interesse utilizada foi o número de estações em cada uma das 54 zonas da POD que cobrem a área de operação das 239 estações consideradas do sistema Bikesampa (ver Figura 10 e Anexo C). Esta escolha deve-se à limitação do Índice de Moran, que requer variáveis numéricas.

A proximidade entre regiões pode ser representada por matrizes de distância ou contiguidade. A matriz de contiguidade considera apenas as zonas com fronteira direta, assumindo que zonas contíguas possuem maior interação espacial (MORAN, 1948; ALMEIDA, 2012). Foi usada a matriz de contiguidade do tipo rainha, por analogia aos movimentos no xadrez, na qual considera fronteiras com extensão maior que zero, assim como vértices contíguos.

Os resultados indicam que é possível rejeitar a hipótese nula de distribuição espacial aleatória para os três grupos de estações, uma vez que o índice de Moran calculado supera o seu valor esperado (Tabela 11).

Tabela 11 – Índice de Moran Global

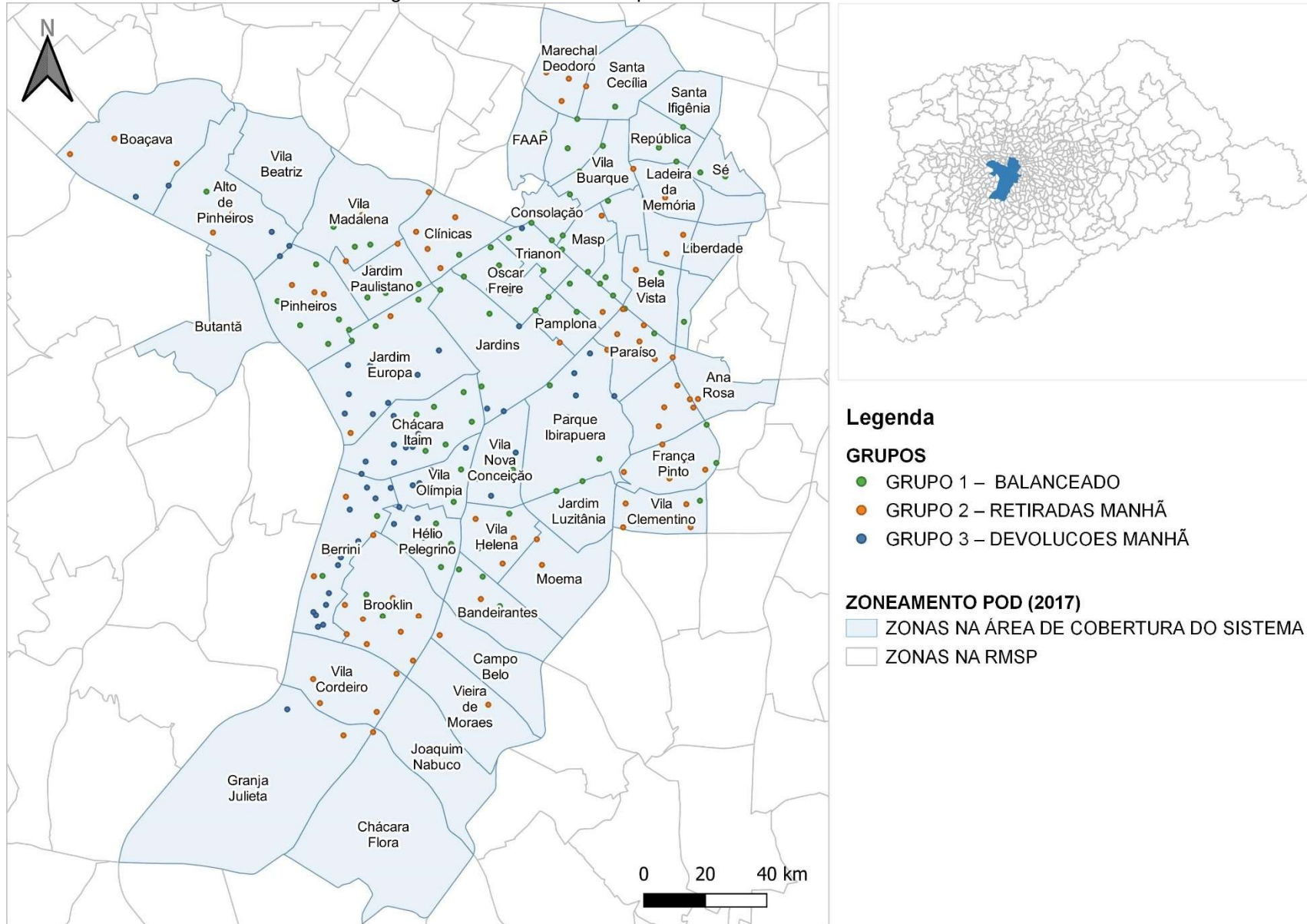
Grupo	Valor Calculado	Valor Esperado
Grupo 1 – Balanceado	0,172	
Grupo 2 – Retiradas Manhã	0,100	-0,019
Grupo 3 – Devoluções Manhã	0,373	

Fonte: Elaboração própria, 2022.

A Figura 11 ilustra os mapas LISA. Os resultados apontam que:

- **Grupo 1 – Balanceado:** a Figura 11a mostra que as zonas Jardim Europa, Jardins, Pamplona, Vila Nova Conceição e Vila Olímpia, regiões onde coexistem residências, comércio e serviços, apresentam um comportamento “Alto-Alto”. Inversamente, as zonas Granja Julieta, Chácara Flora e Joaquim Nabuco, no extremo Sul, com menores quantidades de empregos em serviços e maior uso residencial, apresentam um comportamento “Baixo-Baixo”. Já as zonas Clínicas e Campinas estão classificadas como “Baixo-Alto”, regiões onde predominam estações de outros grupos. A significância desta classificação está entre 1 e 5% (Figura 11b);

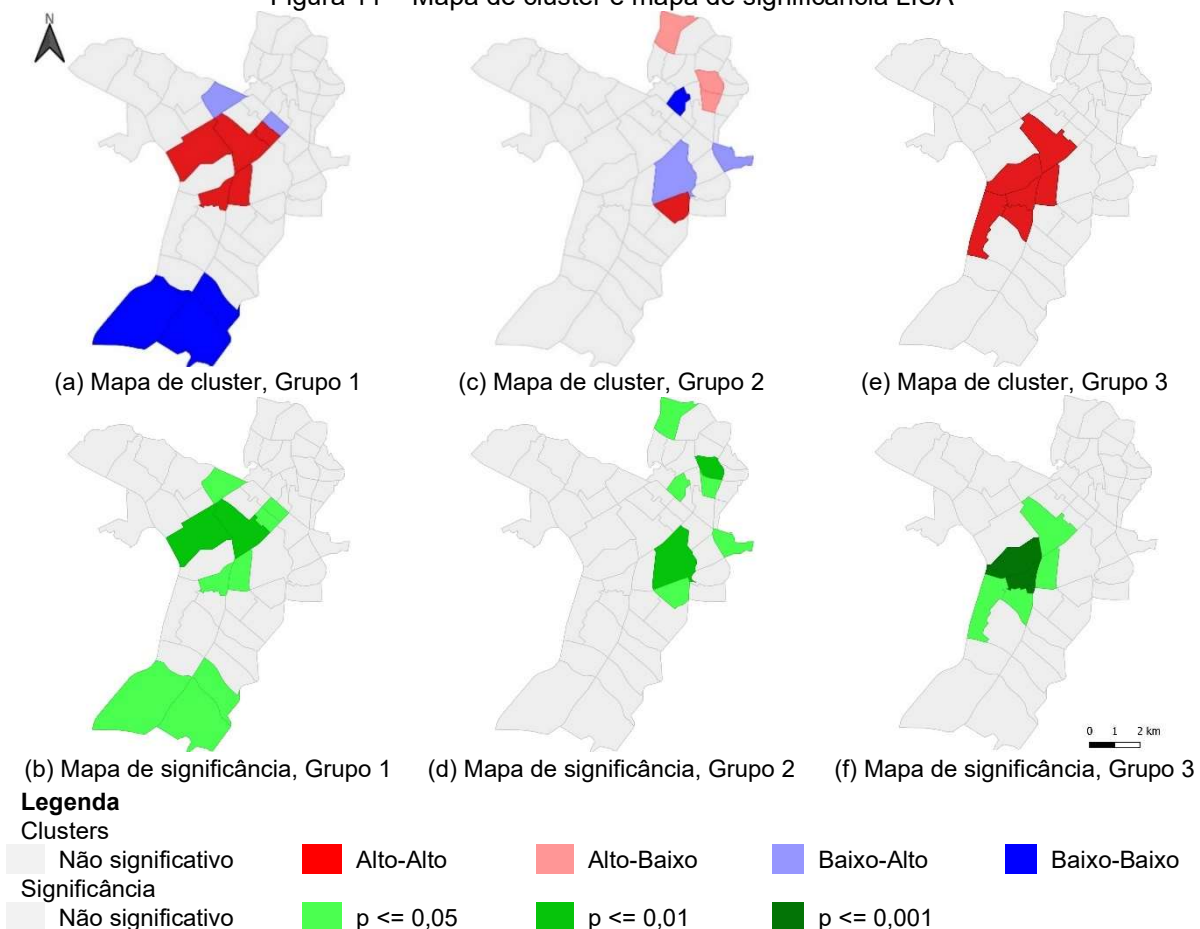
Figura 10 – Zonas utilizadas para o cálculo do Índice de Moran



Fonte: Elaboração própria, 2022. Base: METRO (2020b).

- Grupo 2 – Retiradas manhã:** a Figura 11c mostra que a zona Jardim Luzitânia apresenta um comportamento “Alto-Alto”, muito provavelmente pela menor oferta de empregos e maior quantidade de residências. A zona do Masp (na Av. Paulista) tem um valor “Baixo-Baixo”, em virtude da presença majoritária de estações do **Grupo 1 – Balanceado** na região. São classificadas como zonas “Baixo-Alto” o Parque Ibirapuera e a Ana Rosa, enquanto as zonas Marechal Deodoro, Ladeira da Memória e Bexiga são classificadas como “Alto-Baixo”. A significância desta classificação está entre 1 e 5% (Figura 11d). Como as estações desse grupo são mais dispersas, há mais zonas classificadas com comportamento de transição (Alto-Baixo ou Baixo-Alto);
- Grupo 3 – Devoluções manhã:** a Figura 11e demonstra que este grupo apresenta exclusivamente um comportamento “Alto-Alto” nas zonas Berrini, Chácara Itaim, Hélio Pelegrino, Jardins, Vila Nova Conceição e Vila Olímpia. Estas regiões são conhecidas por serem polos de trabalho em serviços, abrigando imóveis de lajes corporativas. A significância desta classificação está entre 0,1 e 5% (Figura 11f), sendo que as zonas Chácara Itaim e Vila Olímpia são aquelas com maior significância.

Figura 11 – Mapa de cluster e mapa de significância LISA



Fonte: Elaboração própria, 2022.

Vale destacar que a localização das estações não resulta de um processo natural, não-controlável, mas sim da decisão da operadora, a qual se baseia em expectativas de demanda e nas características de operação, de geometria viária e das atividades locais para posicionamento das estações. Logo, resulta de um evento antrópico, não-aleatório, fato confirmado pelos resultados da autocorrelação espacial.

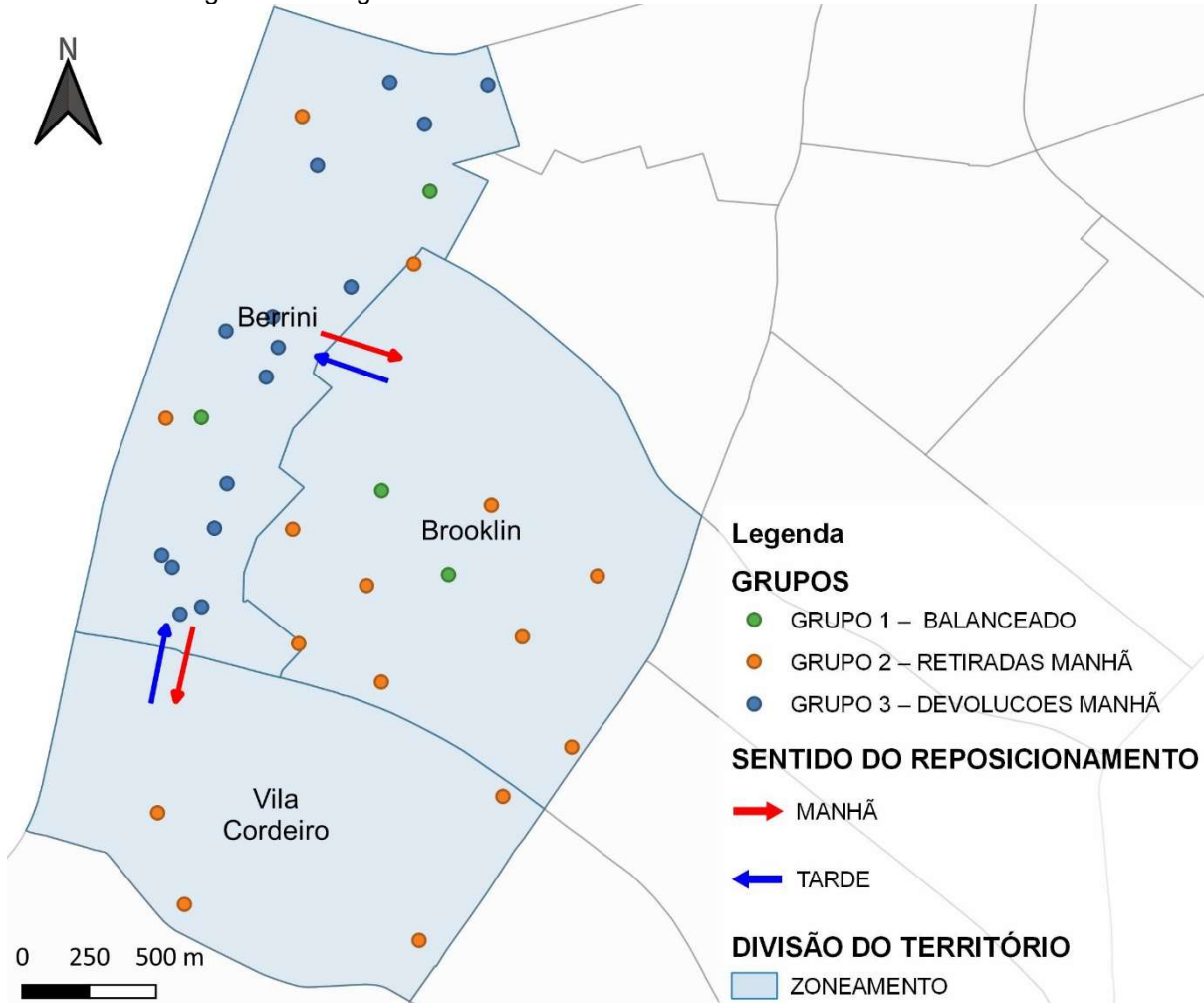
Os resultados sugerem, em primeiro lugar, a existência da associação entre as características do entorno de uma estação e a pertinência a um certo grupo de perfil temporal da demanda, hipótese avaliada no item 4.6. Conhecido ou inferido um dado grupo de perfil temporal com base nas características do entorno de uma estação, esta informação permite orientar a tomada de decisão da operadora quanto à possível necessidade de reposicionamento de bicicletas para evitar saturação das estações, em especial para os grupos com desequilíbrio de fluxos.

Em segundo lugar, os resultados contribuem para a discussão de novas possibilidades para tratar o reposicionamento de bicicletas. Em zonas com classificação “Alto-Alto”, por exemplo, há maior concentração de estações de um dado grupo na zona e na vizinhança, havendo conseqüentemente maior chance de haver déficit de bicicletas ou vagas para os grupos de estações com desequilíbrio de fluxos.

Em casos como esses, o operador pode avaliar pelo menos quatro políticas de operação para reduzir o déficit de bicicletas ou vagas. São elas: (i) aumento do número de docas da estação, (ii) implantação de bolsão na estação, com o aumento do número de bicicletas ou docas livres disponíveis, (iii) implantação de novas estações na região, e (iv) reposicionamento de bicicletas oriundas de outras estações.

A Figura 12 mostra um exemplo do caráter complementar das estações de grupos distintos em zonas vizinhas. A zona Berrini tem um alto número de estações com prevalência de devoluções pela manhã (**Grupo 3 – Devoluções manhã**). O excedente de bicicletas devolvidas nesse período pode ser remanejado para as estações com elevado fluxo de retiradas pela manhã (**Grupo 2 – Retiradas manhã**) das zonas Brooklin e Vila Cordeiro. À tarde, o processo se dá no sentido oposto.

Figura 12 – Sugestão de rebalanceamento Berrini-Brooklin/Vila Cordeiro



Fonte: Elaboração própria, 2022.

4.6. Relação entre grupos de estações e variáveis explicativas

Uma vez comprovada a existência de autocorrelação espacial, foi realizada uma investigação sobre quais características socioeconômicas, operacionais e do espaço urbano, estão associadas ao perfil temporal da demanda das estações. Inicialmente, foi realizada uma análise exploratória das variáveis para a pré-seleção de variáveis explicativas mais importantes para cada grupo. Em seguida, selecionadas estas variáveis explicativas, foi estimado um modelo confirmatório tendo como variável resposta o grupo de estação.

4.6.1. Pré-seleção de variáveis

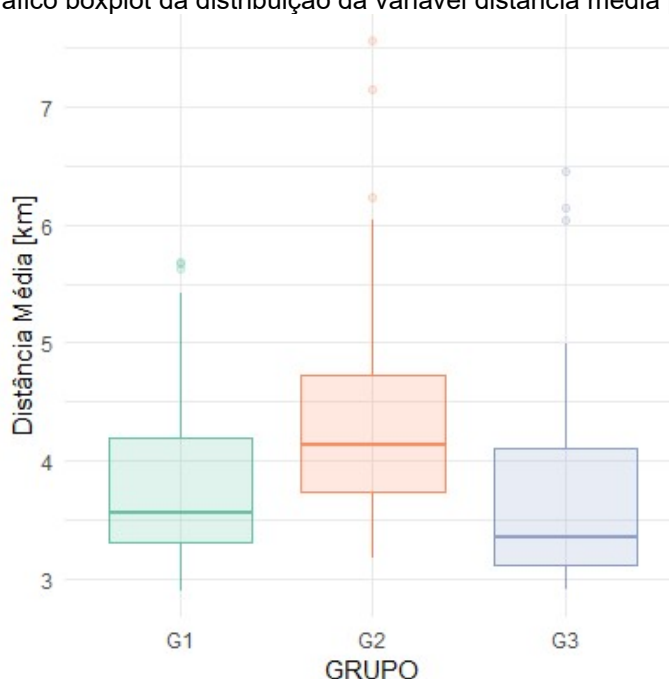
Dentre o conjunto de variáveis levantadas, foram selecionadas aquelas que individualmente apresentassem associação significativa com os grupos de perfil temporal. Para tanto, foram estimadas regressões logísticas multinomiais no *software* Stata, tendo como variável resposta os grupos de perfil temporal e como variável

explicativa cada uma das 34 variáveis listadas no item 4.3. As variáveis explicativas que não foram significativas ao nível de 5% foram desconsideradas, a saber: *CPTM*, *pontos_ônibus*, *vias_ônibus*, *parque*, *esporte*, *shopping*, *industrial* e *equipamentos_públicos*., reduzindo o número de variáveis utilizadas para 26.

Ao final, restou pelo menos uma variável de cada categoria (Tabela 1). As variáveis selecionadas foram analisadas comparativamente entre os grupos de estações, como apresentado na Tabela 12 e no Anexo D.

A título de ilustração da análise apresentada no Anexo D, a Figura 13 mostra o *boxplot* para a variável distância média (de uma estação para todas as outras). Verifica-se que a maior mediana acontece no **Grupo 2 – Retiradas manhã**, que também apresenta o maior número de valores discrepantes na porção superior do gráfico. Por outro lado, o **Grupo 1 – Balanceado** e o **Grupo 3 – Devoluções manhã** apresentam menores medianas comparativamente ao grupo anterior. Além disso, estes dois últimos grupos apresentam maior assimetria estatística (*skewness*).

Figura 13 – Gráfico boxplot da distribuição da variável distância média para cada grupo



Fonte: Elaboração própria, 2022.

Tabela 12 – Valores médios e contagem das variáveis para os diferentes grupos

Categoria	Variável explicativa	Média			Contagem			VIF
		Grupo 1 Balanceado	Grupo 2 Retiradas Manhã	Grupo 3 Devoluções Manhã	Grupo 1 Balanceado	Grupo 2 Retiradas Manhã	Grupo 3 Devoluções Manhã	
-	<i>Número de estações</i>	-	-	-	95	87	57	-
Características do sistema	dist_estacao_proxima	0,31	0,34	0,25	-	-	-	1,78
	dist_media	3,77	4,34	3,70	-	-	-	2,81
	tamanho	15,67	16,54	18,44	-	-	-	1,36
	bolsao	-	-	-	3	6	10	1,33
Características de demanda	dur_origem	23,98	21,48	19,69	-	-	-	4,17
	dur_destino	24,63	22,26	18,88	-	-	-	4,18
Infraestrutura ciclovitária	bicicletário	-	-	-	8	19	1	1,63
	dist_vias_ciclaveis	0,28	0,32	0,17	-	-	-	1,58
Sistema de transportes	metro	-	-	-	31	28	1	2,19
	estacionamento	21,00	13,46	12,91	-	-	-	3,14
	zona_azul	-	-	-	92	68	49	2,39
Pontos de interesse	museu	-	-	-	26	18	7	1,38
	vida_noturna	5,37	3,12	3,00	-	-	-	2,39
	mercado	32,39	29,01	15,90	-	-	-	1,38
	estudo	-	-	-	52	52	16	1,61
	restaurantes	85,77	55,85	49,81	-	-	-	6,66
Geografia	declividade	0,72	0,65	0,96	-	-	-	1,97
Características socioeconômicas	perc_idade_18_25	0,35	0,28	0,29	-	-	-	1,44
	perc_idade_26_36	0,44	0,50	0,46	-	-	-	3,71
	perc_idade_37_55	0,19	0,21	0,23	-	-	-	3,34
	perc_2_5_SM	0,33	0,31	0,22	-	-	-	20,67
	perc_6_SM	0,63	0,65	0,75	-	-	-	25,11
Uso do solo fiscal	residencial	0,42	0,53	0,29	-	-	-	6,03
	comercial	0,16	0,09	0,13	-	-	-	4,39
	uso_misto	0,25	0,22	0,18	-	-	-	2,89
	outros*	0,11	0,11	0,35	-	-	-	3,92

*Outros: usos do solo sem predominância, sem informação, escolas, garagens, terrenos vagos e outros (não especificados pelo GeoSampa).

Fonte: Elaboração própria, 2022.

Para a análise de multicolinearidade, foi utilizada a estatística *Variance Inflation Factor* (VIF), do pacote *car* da linguagem R. Em geral, muitos autores consideram presença de multicolinearidade quando $VIF > 10$ (FÁVERO; BELFIORE, 2017). Para as variáveis selecionadas, apenas as variáveis relativas à renda possuem VIF acima de 10. Logo, somente a variável *perc_6_SM* foi utilizada no modelo apresentado no item 4.6.2, pois estas variáveis estão correlacionadas entre si.

As características em destaque para cada grupo de estação são:

- **Grupo 1 – Balanceada:** as viagens feitas a partir dessas estações são mais longas e seus usuários são mais jovens. Este grupo apresenta a maior quantidade e presença de diversos pontos de interesse, o que explica a maior proporção de uso do solo misto de todos os grupos. Com base na literatura, era esperado que seu comportamento temporal fosse balanceado e com fluxo crescente ao longo do dia;
- **Grupo 2 – Retiradas manhã:** estas estações estão mais distantes das demais, estando localizadas em áreas residenciais e na porção mais periférica do sistema BukeSampa. Logo, seu perfil temporal pode estar associado com a demanda de primeira milha na manhã e última milha à tarde, para acesso ao transporte público;
- **Grupo 3 – Devoluções manhã:** estas estações estão em regiões de uso de solo comercial, de forma que seu uso é consistente com a demanda de última milha na manhã e primeira milha à tarde. Localizam-se em uma região plana servida de infraestrutura cicloviária, porém distante de estações de metrô, o que sugere o uso conjugado com o transporte público. Esta configuração pode intensificar o uso das bicicletas compartilhadas (BENEDINI; LAVIERI; STRAMBI, 2019; EREN; UZ, 2020; TRAN; OVTRACHT; D'ARCIER, 2015), o que explica o maior fluxo horário médio dentre os grupos, e sugere a maior capacidade (tamanho e bolsão) destas estações.

4.6.2. *Estimação de modelo confirmatório*

A análise exploratória realizada no item anterior indicou a associação entre cada variável explicativa e os grupos de perfil temporal de uso das estações. Para identificar a significância estatística do conjunto de variáveis foi estimado um modelo confirmatório. Uma vez que os grupos, considerados como variável dependente, são categóricos, é necessário um modelo apropriado para a análise de variáveis discretas.

4.6.2.1. *O modelo probit*

Os dois modelos mais usados para variáveis dependentes discretas são o logit e o probit (binomiais ou multinomiais). O probit permite relaxar algumas hipóteses restritivas do logit multinomial, principalmente quanto a flexibilização dos padrões de substituição entre alternativas (já que não utiliza a premissa de independência de alternativas) (TRAIN, 2009). Além disso, o probit tem uma abordagem mais direta do que um modelo logit hierárquico.

No contexto desta análise, o modelo probit permite estimar para cada estação, a partir de suas características, a probabilidade de pertencer a cada um dos grupos de perfil temporal. A atratividade ou utilidade de cada alternativa/grupo é dada pela Equação 2. Admite-se que o termo de erro tem distribuição normal, com matriz de covariância Ω (Equação 3). Logo, a densidade de ε_n é dada pela Equação 4. Por fim, a probabilidade de uma dada alternativa é dada pela Equação 5 (TRAIN, 2009).

$$U_{nj} = V_{nj} + \varepsilon_{nj}, \quad \forall j \quad (2)$$

$$\Omega = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \quad (3)$$

$$\Phi(\varepsilon_n) = \frac{1}{(2\pi)^{J/2} |\Omega|^{1/2}} e^{-\frac{1}{2} \varepsilon_n' \Omega^{-1} \varepsilon_n} \quad (4)$$

$$P_{ni} = \int I(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj} \quad \forall j \neq i) \Phi(\varepsilon_n) d\varepsilon_n \quad (5)$$

Onde

n: estação

j: alternativa, $j = 1, \dots, J$ (grupo de perfil temporal)

U_{nj} : atratividade/utilidade da alternativa j para a estação n

V_{nj} : atratividade/utilidade representativa da alternativa j para a estação n (função linear aditiva das variáveis explicativas)

ε_n : vetor dos termos de erro; $\varepsilon_n = [\varepsilon_{n1}, \varepsilon_{n2}, \dots, \varepsilon_{nJ}]$

Φ : função de densidade

I: 1, se condição é satisfeita, 0 caso contrário

Uma diferença importante entre os modelos probit e o logit é que, enquanto o primeiro permite uma matriz de variância-covariância genérica, com parâmetros estimáveis, no caso do logit multinomial a matriz de variância-covariância é uma matriz identidade (multiplicada pela variância do termo de erro), que não admite correlação entre as alternativas.

Segundo Train (2009), no lugar da matriz de covariância de erros do probit, é possível considerar a matriz de covariância da diferença entre erros, dada por $\tilde{\Omega}_i = M_i \Omega M_i'$, em que M_i é a matriz de identidade de dimensão (J-1) com uma coluna extra de -1 adicionada na coluna i. Então, para M_1 :

$$M_1 = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \quad (6)$$

Logo, a matriz de covariância de diferença de erro em relação à alternativa 1 (considerou-se o **Grupo 1 – Balanceado** como a categoria de referência) está apresentada na Equação 7. Os valores A, B e C são simplificações das relações entre as covariâncias. O valor de A é fixo em 1 para normalizar (ajustar a escala) as utilidades e garantir a identificação do modelo. Logo, considerar a matriz de covariância da diferença entre erros é mais vantajoso, visto que há menos parâmetros para se estimar (apenas B e C).

$$\tilde{\Omega}_i = \begin{pmatrix} \sigma_{11} + \sigma_{22} & \sigma_{11} - \sigma_{13} - \sigma_{12} + \sigma_{23} \\ \sigma_{11} - \sigma_{13} - \sigma_{12} + \sigma_{23} & \sigma_{11} + \sigma_{33} \end{pmatrix} = \begin{pmatrix} A & B \\ B & C \end{pmatrix} \quad (7)$$

Para medir a qualidade de ajuste do modelo, foi utilizado o *rho* quadrado ajustado, expresso na Equação 8 (KOPPELMAN; BHAT, 2006). Esta estatística é comumente usada para aferir a qualidade do ajuste do modelo, e tem valor entre 0 e 1, em que 1 indica que o modelo prevê corretamente todos os grupos de estações.

$$\bar{\rho}_c^2 = 1 - \frac{LL(\hat{\beta}) - K}{LL(C) - K_{MS}} \quad (8)$$

Onde

$LL(\hat{\beta})$: verossimilhança do modelo estimado

$LL(C)$: verossimilhança do modelo apenas com constantes

K: número de graus de liberdade no modelo (K = 19, neste caso)

K_{MS} : número de graus de liberdade no modelo apenas com constantes (K = 2, neste caso)

Para verificar se a escolha por um modelo probit é adequada face a um modelo mais simples, como o logit multinomial, foi realizado um teste qui-quadrado (χ^2) comparando dois modelos probit. São eles: (i) modelo restrito, em que a matriz de covariância é a matriz identidade e (ii) modelo irrestrito, em que os elementos da matriz de covariância assumem qualquer valor entre 0 e 1.

O cálculo do qui-quadrado é feito segundo a Equação 9 (KOPPELMAN; BHAT, 2006). Quando $\chi^2 > \chi^2_{crit}$, rejeita-se a hipótese H_0 de correlação entre as alternativas ocorrer por pura chance. O χ^2_{crit} é o valor do chi-quadrado crítico, que é definido a

partir dos graus de liberdade, que são a diferença entre o número de parâmetros estimados no modelo restrito e irrestrito.

$$\chi^2 = -2 \times [LL_R - LL_U] \quad (9)$$

Onde

LL_R: verossimilhança do modelo restrito (matriz de covariância é a matriz identidade)

LL_U: verossimilhança do modelo irrestrito, equivalente a $LL(\hat{\beta})$

A estimação dos modelos foi realizada no GAUSS utilizando o código desenvolvido por BHAT (2018). Foram incluídas na especificação dos dois modelos as 25 variáveis pré-selecionadas. No modelo, as variáveis do tamanho, da duração média da viagem e outras de contagem (ex.: número de restaurantes) foram padronizadas, sendo divididas pelo seu valor médio para evitar que grandes diferenças de escala entre as variáveis impactassem negativamente o modelo. Ainda, visto que ambas as variáveis de características de demanda têm comportamento similar, foi considerada apenas a variável da duração a partir da origem.

4.6.2.2. Estimação e resultados

Os resultados da estimação estão apresentados na Tabela 13. Observa-se que 11 das 25 variáveis explicativas são significativas (5%). A estimação do modelo probit resultou em um ρ_c^2 de 0,404, o que pode ser considerado um bom ajuste do modelo.

Tabela 13 – Resultados da estimação do modelo probit

Grupo	Coefficiente	Erro Padrão	Estatística t
Grupo 1 – Balanceado	(Referência)		
Grupo 2 – Retiradas Manhã			
Constante	3,725	1,113	3,347
dist_média	0,294	0,125	2,354
bolsão	0,903	0,300	3,006
dur_origem	-2,545	0,725	-3,511
bicicletário	1,322	0,343	3,857
metrô	-0,631	0,258	-2,448
mercado	0,918	0,269	3,413
restaurantes	-1,087	0,284	-3,828
perc_idade_18_25	-3,045	1,441	-2,114
Grupo 3 – Devoluções Manhã			
constante	1,715	0,725	2,366
dur_origem	-1,564	0,561	-2,789
dist_vias_cicláveis	-0,785	0,282	-2,783
perc_idade_18_25	-3,711	1,301	-2,853
perc_6_SM	1,205	0,551	2,185
outros	1,560	0,367	4,247
Verossimilhança do modelo final, $LL(\hat{\beta})$ ou LL_U			-136,514
Verossimilhança do modelo apenas com constantes, $LL(C)$			-257,269
ρ_c^2			0,404

Fonte: Elaboração própria, 2022.

Para avaliar a pertinência da estimação por meio do probit, foi estimada a correlação entre as alternativas da matriz de covariância da diferença de erros. A Tabela 14 apresenta os valores estimados dos coeficientes da matriz, conforme indicados na Equação 7 acima, bem como os respectivos erros-padrão e as estatísticas t. A partir dela, verifica-se que a correlação entre as alternativas é significativa, resultado que sustenta o uso do modelo probit.

Tabela 14 – Resultados da matriz de covariância

Termo da matriz	Coefficiente	Erro Padrão	Estatística t
A	1		
B	0,498	0,111	4,482
C	0,248	0,115	2,156

Fonte: Elaboração própria, 2022.

Com relação aos resultados da comparação entre o modelo acima, irrestrito, e outro restrito, a Tabela 15 mostra os valores necessários de verossimilhança para o cálculo do chi-quadrado. Como $\chi^2 > \chi^2_{crit}$, rejeita-se a hipótese H_0 de ausência de correlação entre as alternativas, confirmando novamente a pertinência do probit para modelar o fenômeno de interesse.

Tabela 15 – Verossimilhança e qui-quadrado

Variável	Valor
Verossimilhança do modelo final, $LL(\hat{\beta})$ ou LL_U	-136,514
Verossimilhança do modelo restrito (LL_R)	-146,006
χ^2	18,983
χ^2_{crit}	5,990

Fonte: Elaboração própria, 2022.

4.6.2.3. Interpretação dos coeficientes

A interpretação dos valores estimados dos coeficientes deve ser feita por comparação com o grupo de referência (**Grupo 1 – Balanceado**). É importante destacar que o modelo captura associações entre os grupos de perfil temporal e as variáveis explicativas, sem necessariamente estabelecer relações causais. Ainda, deve-se ressaltar que as comparações são qualitativas, sendo que a magnitude dos coeficientes não apresenta significado específico. Os principais resultados são analisados a seguir, segundo as categorias de variável explicativa apresentadas anteriormente (Tabela 15). Para compreender melhor as relações espaciais descritas a seguir, recomenda-se rever a Figura 8 e a Figura 9.

Características do sistema

Exceto pelas variáveis de distância da estação mais próxima e tamanho, as demais variáveis das características do sistema (*dist_média* e *bolsão*) entraram no

modelo final, porém, apenas na função de utilidade do **Grupo 2 – Retiradas manhã**. Os coeficientes destas variáveis são positivos, indicando que há maior probabilidade de uma estação pertencer ao segundo grupo com o aumento do valor dessas variáveis, comparativamente ao grupo de referência.

Primeiramente, do ponto de vista da variável distância média, maiores valores estão associados às estações localizadas na periferia do sistema, onde prevalecem áreas residenciais, em que o padrão de retiradas pela manhã e devoluções a tarde são mais frequentes, em linha com o esperado.

Em seguida, do ponto de vista da presença de bolsão, uma primeira hipótese para o sinal positivo do coeficiente para o segundo grupo seria o provável uso destas estações para viagens de primeira milha para o trabalho, situação em que o risco de encontrar estações vazias é mais relevante para o usuário.

Também, acredita-se que, pelo maior equilíbrio entre os fluxos, as estações com fluxos balanceados do grupo de referência têm menor probabilidade de ficarem totalmente cheias ou vazias, reduzindo a necessidade de maior capacidade para acomodar picos de demanda em um dos sentidos (retirada ou devolução). De outro ponto de vista, as estações dos grupos desbalanceados possuem maior déficit de bicicletas nos horários de pico, de modo que uma reserva de capacidade é essencial para não limitar a oferta e o número de usuários que utilizam as bicicletas compartilhadas nesses horários.

Características da demanda

A variável de duração média das viagens a partir da estação de origem foi significativa, com coeficiente negativo para os grupos G2 e G3. Ou seja, à medida que aumenta a duração média das viagens com origem na estação, é menos provável que ela pertença a esses grupos. Como as estações do **Grupo 1 – Balanceado** estão próximas a áreas de lazer e recreação, como o Parque Ibirapuera, é possível que os usuários utilizem suas bicicletas também para viagens recreativas, que provavelmente são mais longas do que viagens pendulares. É esperado que as viagens à trabalho ou escola sejam mais curtas, pois o usuário busca minimizar o tempo em deslocamento nestas atividades. Isto reforça a hipótese de que as estações do **Grupo 2 – Retiradas manhã** e do **Grupo 3 – Devoluções manhã** são utilizadas para viagens pendulares.

Infraestrutura cicloviária

Ambas as variáveis desta categoria foram significativas no modelo. A variável bicicletários é significativa apenas no **Grupo 2 – Retiradas manhã**. O sinal positivo indica que estações deste grupo possuem maior probabilidade de estarem presentes em áreas com mais bicicletários do que as estações do grupo de referência. Uma hipótese para este comportamento é que, nas regiões residenciais onde estão essas estações, é menor a presença de estações de metrô e trem, mas estas possuem bicicletário, um mecanismo de incentivo para usuário usar a bicicleta própria para as viagens de primeira e última milha.

Por sua vez, a variável de proximidade de vias com infraestrutura cicloviária é significativa apenas no **Grupo 3 – Devoluções manhã**. O sinal negativo indica que as estações deste grupo têm menor (maior) probabilidade de estarem mais distantes (próximas) de ciclovias e ciclofaixas comparativamente ao grupo de referência. É importante destacar que as estações do Grupo 3 estão localizadas em áreas comerciais, que têm maior concentração de demanda e corredores mais bem definidos, como as avenidas Faria Lima, Pedroso de Moraes e Berrini. Ainda, estas regiões receberam investimento para a construção e melhoria de ciclovias no âmbito do Plano Cicloviário do Município de São Paulo (CET, 2020).

Sistemas de transportes

Apenas a variável indicativa da presença de estação de metrô na área de serviço de uma estação foi significativa no modelo, e exclusivamente no **Grupo 2 – Retiradas manhã**. O sinal negativo indica que a probabilidade de as estações pertencerem a este grupo é menor quando há a presença de metrô em sua área de serviço, comparativamente ao grupo de referência. Uma possível justificativa para este comportamento é que as estações do grupo de referência estão mais próximas ao sistema de metrô, principalmente na região da Av. Paulista.

Pontos de interesse

Da categoria de pontos de interesse, são significativas apenas as variáveis de mercados (coeficiente com sinal positivo) e restaurantes (sinal negativo), exclusivamente para o **Grupo 2 – Retiradas manhã**. A presença de mercados nas proximidades da estação aumenta a probabilidade de que pertença ao segundo grupo. De fato, é esperado que esse tipo de estabelecimento busque localização próxima aos consumidores, em áreas residenciais. Por outro lado, um maior número de

restaurantes é característica de áreas comerciais e de serviços, reduzindo a probabilidade de que uma estação tenha perfil temporal pertencente ao **Grupo 2 – Retiradas manhã**, em relação ao grupo de referência.

Geografia

O modelo não inclui a variável declividade, única representante desta categoria.

Características socioeconômicas

Entre as categorias de idade, apenas o percentual de usuários entre 18 e 25 anos foi significativo no modelo. Em ambos os grupos de estação (**Grupo 2 – Retiradas manhã** e **Grupo 3 – Devoluções manhã**), esta variável apresentou um coeficiente com sinal negativo, indicando que a probabilidade de uma estação pertencer a esses grupos é menor quando há uma maior proporção de usuários mais jovens que usam a estação.

Uma hipótese para este comportamento pode estar associada às viagens possivelmente realizadas a lazer e recreação no **Grupo 1 – Balanceado**, visto que algumas estações deste grupo estão próximas a áreas de lazer (Parque Ibirapuera) e de cultura (há diversos museus e instituições da área de cultura na Av. Paulista), sugerindo uma possível associação entre a população jovem, o uso da bicicleta e esse tipo de atividades. Adicionalmente, a maior presença de jovens pode estar associada a viagens com fins de educação. Este comportamento também reforça a hipótese de que as estações do **Grupo 2 – Retiradas manhã** e **Grupo 3 – Devoluções manhã** sejam utilizadas para viagens motivo trabalho, visto nesses grupos prevalecem faixas etárias superiores (ver Tabela 12), compatível com a maior parcela da população economicamente ativa.

Quanto às faixas de renda, apenas a categoria de renda familiar acima de 6 salários mínimos foi significativa e com coeficiente positivo, exclusivamente no **Grupo 3 – Devoluções manhã**. Boa parte destas estações encontra-se próxima a locais com emprego em serviços de alto nível, compatíveis, portanto, com a demanda de usuários que recebem rendimentos do trabalho na faixa superior de renda.

Uso do solo

Dentre as variáveis de uso do solo, a única que é significativa no modelo é a categoria “outros”, que inclui áreas sem predominância de um tipo de uso específico, áreas sem informação sobre o uso do solo, escolas, terrenos e outros tipos não especificados pelo Geosampa. Esta variável é significativa para o **Grupo 3 –**

Devoluções manhã, e possui um sinal positivo. Ou seja, a probabilidade de uma estação pertencer ao terceiro grupo é maior com o aumento da parcela de uso do solo “outros” em sua área de serviço, comparativamente ao grupo de referência. Uma hipótese é que as regiões desse grupo não possuem um uso do solo predominante nas quadras, havendo maior diversidade de usos entre lotes, que atraem viagens pela manhã (e no sentido contrário no pico da tarde/noite).

Comentários finais

Assim como já era esperado da análise realizada no item 4.6, estes resultados demonstram que os diferentes grupos de estações apresentam comportamentos distintos a depender de características socioeconômicas, da operação e do espaço urbano no seu entorno. A Tabela 16 apresenta um quadro resumo com as interpretações apresentadas anteriormente.

Compreender os fatores associados a cada perfil temporal permite que os operadores entendam como estas características estão relacionadas com os fluxos horários de retiradas e devoluções. E, dessa forma, podem prever o perfil temporal da demanda em uma estação em diferentes horizontes de tempo, a saber:

- **Curto prazo**, onde a operadora pode prever o perfil temporal nos cenários em que há mudança de localização de uma estação pré-existente, ou então das características operacionais (aumento do tamanho, ou bolsão), algo que é facilitado pelo caráter modular do sistema da TemBici, ou desencadeado por alterações de políticas sobre o uso do espaço público em áreas específicas;

- **Médio prazo**, no qual a operadora pode realizar uma previsão preliminar do perfil temporal de uso de novas estações, a partir das características do entorno do local proposto para instalação. É importante destacar que esta previsão deve ser realizada com cautela, visto que os grupos são resultado de uma análise exploratória que considerou um conjunto específico de estações. Além disso, para novas estações, é inexistente a informação da duração média da viagem, a qual foi identificada por meio do histórico de viagens;

- **Longo prazo**, em que é possível realizar uma previsão do perfil temporal face a mudanças no ambiente construído e nas características socioeconômicas da população. Ou seja, alterações como implantações de novas estações de metrô, ou novos empreendimentos imobiliários.

Tabela 16 – Quadro resumo com interpretações do modelo

Categoria	Variável	Coeficiente		Interpretação*
		G2	G3	
Características do sistema	dist_média	+		O Grupo 2 – Retiradas manhã está concentrado na periferia do sistema, em áreas residenciais. Estas estações estão mais distantes das demais
	bolsão	+		O Grupo 2 – Retiradas manhã provavelmente são utilizadas para viagens de primeira milha para o trabalho, situação em que o risco de encontrar estações vazias é mais relevante para o usuário
Características da demanda	dur_origem	-	-	O Grupo 2 – Retiradas manhã e o Grupo 3 – Devoluções manhã possuem durações de viagens mais curtas, provavelmente porque as viagens dessas estações são por motivo trabalho e educação, nas quais o usuário busca minimizar o tempo de deslocamento. Ainda, a proximidade do Grupo 1 – Balanceado a áreas de lazer e recreação pode indicar a presença de viagens mais longas
Infraestrutura cicloviária	bicicletário	+		O Grupo 2 – Retiradas manhã pode receber maior incentivo para o usuário utilizar a bicicleta própria para viagem de 1ª milha até esse sistema
	dist_vias_cicláveis		-	O Grupo 3 – Devoluções manhã está mais próximo de ciclovias e ciclofaixas. Uma hipótese para este comportamento é a localização das estações desse grupo nas proximidades de avenidas que foram abrangidas no Plano Cicloviário da Cidade de São Paulo e contam com melhor infraestrutura cicloviária
Sistema de transportes	metrô	-		O Grupo 1 – Balanceado está localizado em regiões centrais, com maior presença de estações de metrô
Pontos de interesse	mercado	+		A região do Grupo 2 – Retiradas manhã possui mais mercados na área de serviço. Uma hipótese é que as este tipo de estabelecimento busque localização próxima ao cliente final, em áreas residenciais
	restaurantes	-		A região do Grupo 2 – Retiradas manhã possui menos restaurantes na área de serviço. Acredita-se na hipótese de que os restaurantes estão localizados em regiões de comércio e serviços, como aquelas encontradas nas proximidades do Grupo 1 – Balanceado
Características socioeconômicas	perc_idade_18_25	-	-	A região do Grupo 1 – Balanceado possui maior incidência de população jovem do que os outros grupos, o que pode ser associada à presença de áreas de lazer e recreação
	perc_6_SM		+	A região do Grupo 3 – Devolução manhã concentra mais serviços e empregos de alto qualificação, o que explica maior parcela de população com rendas maiores
Uso do solo predominante	outros		+	A região do Grupo 3 – Devolução manhã apresenta maior diversidade de uso do solo

*Comparativamente ao grupo de referência **Grupo 1 – Balanceado**

Fonte: Elaboração própria, 2022.

Em especial, quando falamos de nova alocação de estações, é importante fazer duas ressalvas. Em primeiro lugar, com a realocação ou introdução de novas estações, a variável distância média teria de ser recalculada, bem como as variáveis padronizadas, o que poderia levar a diferenças nos efeitos dessas variáveis. Ainda, novas estações podem alterar as características da distribuição espacial das estações, capturadas pelo agrupamento. Um exemplo seria as estações atualmente mais periféricas do sistema deixarem de apresentar essa característica ao serem introduzidas novas estações.

5. LIMITAÇÕES E SUGESTÕES

Algumas limitações desta pesquisa devem ser consideradas, em sua maioria relativas aos dados disponíveis. A informação fornecida pela Tembici - dados de viagens, sem identificação de usuário – não permite distinguir se a viagem foi realizada por um usuário frequente ou eventual. Um usuário frequente pode contribuir com muitas viagens para o banco de dados, enquanto o eventual, poucas, fato que provoca viés na amostra. Tampouco é possível conhecer o tempo que o usuário vem utilizando o sistema, fator que afeta o comportamento em relação ao uso de bicicletas compartilhadas (BENEDINI; LAVIERI; STRAMBI, 2019; TRAN; OVTRACHT; D'ARCIER, 2015). Ainda, não foi considerado nenhum tratamento para dias com climas adversos, como fazem alguns estudos (FAGHIH-IMANI et al., 2014, 2017; SCHIMOHR; SCHEINER, 2021).

Do ponto de vista das áreas de serviço, como não existe uma rede caminhável disponível para a cidade de São Paulo, o tratamento manual da rede eventualmente pode não ter considerado alguns elementos efetivamente usados por pedestres. Com relação às variáveis relativas ao número de estabelecimentos como estacionamentos, universidades, restaurantes e mercados, o banco de dados usado é datado de 1999. Informações mais recentes não estão disponíveis devido a questões ligadas à confidencialidade dos dados. Por outro lado, comparativamente à fonte do OSM, embora a informação usada seja mais antiga, ela é mais completa.

Como sugestão de trabalhos futuros, recomenda-se avaliar os perfis temporais em sistemas com menores restrições sobre sua área de atuação. Em São Paulo, a ausência de operação da Tembici nas Zonas Norte e Leste não permite captar os aspectos que permitam avaliar questões de acessibilidade e equidade, e como estas podem afetar o comportamento de retiradas e devoluções. Sugere-se também avaliar

possível mudança do comportamento pré e pós-pandemia do Covid-19, visto que a Tembici indicou que o uso recreativo das bicicletas compartilhadas aumentou, embora o total de viagens no sistema tenha se reduzido no período.

Ainda, visto que os dados fornecidos possuem os locais de início e fim das viagens, é possível realizar uma análise dos pares de origem e destino das viagens, não apenas fluxos de retirada e devolução. Por fim, já que os grupos de estações não se distribuem aleatoriamente no espaço, sugere-se a abordagem de modelagem espacial, permitindo captar os efeitos de autocorrelação espacial aqui verificados.

6. CONCLUSÕES

Esta pesquisa estudou o sistema de bicicletas compartilhadas Bikesampa, que opera na cidade de São Paulo. Primeiramente, as estações foram classificadas em três grupos a partir dos fluxos médios horários de retiradas e devoluções: (i) balanceadas, (ii) retiradas manhã e (iii) devoluções manhã. Estes grupos não estão distribuídos aleatoriamente no espaço, havendo, para cada grupo, um padrão que sugere associação com características do entorno das estações. Em seguida, os grupos de estações foram associados a variáveis explicativas relativas a características socioeconômicas, operacionais e do espaço urbano. Os três grupos estão descritos de forma sucinta abaixo.

O **Grupo 1 – Balanceado** tem um padrão balanceado entre retiradas e devoluções, com aumento de demanda gradual ao longo do dia. Este comportamento deve-se provavelmente à localização em áreas de maior proporção de uso do solo misto. Por sua vez, o **Grupo 2 – Retiradas manhã** possui um padrão desbalanceado de dois picos, com maior número de retiradas pela manhã e devoluções à tarde, o que é possivelmente explicado pelo uso do solo majoritariamente residencial no entorno dessas estações. Por fim, o **Grupo 3 – Devoluções manhã** possui padrão desbalanceado de dois picos, porém com maior número de devoluções pela manhã e retiradas à tarde, possivelmente explicado pelas estações com área de serviço com maior presença de uso do solo comercial.

Os resultados contribuem para o desenho de políticas de incentivo e operação para o melhor gerenciamento do sistema pelo operador, de modo a torná-lo mais eficiente, confiável e amigável ao usuário. Para contornar o problema crítico de déficit de bicicletas ou vagas nas estações nos horários de pico, são feitas algumas recomendações principalmente para as estações em que há desequilíbrio de fluxos

de retiradas e devoluções (**Grupo 2 – Retiradas manhã** e **Grupo 3 – Devoluções manhã**). Por possuir fluxos mais equilibrados, espera-se que as estações do **Grupo 1 – Balanceado** sejam menos dependentes de remanejamento de bicicletas.

Como proposição decorrente dos resultados, sugere-se a criação de uma política de incentivo para bonificar os ciclistas que realizam viagens no contrafluxo no horário de pico, uma vez que estes usuários participam indiretamente do processo de reposicionamento de bicicletas. Para tanto, a operadora poderia oferecer preços reduzidos ou outros tipos de vantagem no plano de assinatura.

Adicionalmente, conhecendo o período do dia em que há maior probabilidade de déficit de bicicleta ou vagas, pode ser realizado o reposicionamento de bicicletas nos picos da demanda para cada tipo de grupo. Neste sentido, é esperado que as estações do **Grupo 2 – Retiradas manhã** recebam bicicletas remanejadas de outras estações ou contem com uma reserva de capacidade no período da manhã, e as estações do **Grupo 3 – Devoluções manhã**, na parte da tarde.

Finalmente, dada a verificação que os grupos de estações não se distribuem aleatoriamente no espaço, é razoável admitir que exista uma associação entre o perfil temporal de demanda de uma estação e as características de seu entorno. Esta hipótese foi confirmada por um modelo probit, que identificou a relação entre os grupos de perfil temporal e as características do(a): oferta e demanda do sistema, infraestrutura cicloviária, sistema de transportes, pontos de interesse, socioeconômicas e uso do solo.

Este resultado permite que o operador faça previsões do perfil temporal de uso de uma dada estação, face a mudanças de curto, médio e longo prazo. Assim, é possível prever o equilíbrio e a magnitude dos fluxos de retiradas e devoluções, de modo a identificar necessidades específicas de reposicionamento de bicicletas e orientar as políticas de operação e de incentivo.

REFERÊNCIAS

ABREU, V. H. S. D.; ALMEIDA, I. M. D.; TURINI, L. R. Utilização do transporte ativo para aumento da qualidade de vida da população urbana durante a pandemia do Covid-19 à luz de abordagem bibliométrica. **Brazilian Journal of Production Engineering**, v. 7, n. 3, p. 139-152, 2021.

ALMEIDA, E. **Econometria espacial aplicada**. Campinas: [s.n.], 2012.

ANSELIN, L. Local Indicators of Spatial Association—LISA. **Geographical Analysis**, v. 27, p. 93-115, 1995. Disponível em: <<https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>>.

BENEDINI, J. D.; LAVIERI, P. S.; STRAMBI, O. Understanding the use of private and shared bicycles in large emerging cities: The case of Sao Paulo, Brazil. **Case Studies on Transport Policy**, v. 8, p. 564-575, 2019. Disponível em: <<https://doi.org/10.1016/j.cstp.2019.11.009>>.

BHAT, C. R. New matrix-based methods for the analytic evaluation of the multivariate cumulative normal distribution function. **Transport Research, Part B: Methodological**, v. 109, p. 238-256, 2018.

CET. Mapa de Infraestrutura Cicloviária. **Companhia de Engenharia de Tráfego**, 2020. Disponível em: <<http://www.cetsp.com.br/consultas/bicicleta/mapa-de-infraestrutura-ciclovitaria.aspx>>. Acesso em: 05 mai. 2022.

CET. **Plano Cicloviário do Município de São Paulo**. Companhia de Engenharia de Tráfego. São Paulo, p. 57. 2020.

CHARDON, C. M. D.; CARUSO, G.; THOMAS, I. Bicycle sharing system 'success' determinants. **Transportation Research Part A**, v. 100, p. 202-214, 2017. Disponível em: <<https://doi.org/10.1016/j.tra.2017.04.020>>.

DEMAIO, P.; GIFFORD, J. Will Smart Bikes Succeed as Public Transportation in the United States? **Journal of Public Transportation**, v. 7, n. 2, p. 1-15, 2004. Disponível em: <<http://doi.org/10.5038/2375-0901.7.2.1>>.

EREN, E.; UZ, V. E. A review on bike-sharing: The factors affecting bike-sharing demand. **Sustainable Cities and Society**, v. 54, p. 1-12, 2020. Disponível em: <<https://doi.org/10.1016/j.scs.2019.101882>>.

FAGHIH-IMANI, A. et al. How land-use and urban form impact bicycle flows: evidence from the bicycle-sharing system (BIXI) in Montreal. **Journal of Transport Research**, v. 41, p. 306-314, December 2014. Disponível em: <<https://doi.org/10.1016/j.jtrangeo.2014.01.013>>.

FAGHIH-IMANI, A. et al. An empirical analysis of bike sharing usage and rebalancing: evidence from Barcelona and Seville. **Transportation Research Part A**, v. 97, p. 177-191, March 2017. Disponível em: <<https://doi.org/10.1016/j.tra.2016.12.007>>.

FÁVERO, L. P.; BELFIORE, P. **Manual de Análise de Dados: Estatística e Modelagem Multivariada com Excel®, SPSS® e Stata®**. 1ª Edição. ed. São Paulo: GEN LTC, 2017.

FISHMAN, E. Bikeshare: A Review of Recent Literature. **Transport Reviews**, v. 36, p. 92-113, 2016. Disponível em: <<https://doi.org/10.1080/01441647.2015.1033036>>.

FISHMAN, E.; WASHINGTON, S.; HAWORTH, N. Bike share: A synthesis of the literature. **Transport Reviews**, v. 33, p. 148-165, 2013. Disponível em: <<https://doi.org/10.1080/01441647.2013.775612>>.

FLORINDO, A. A. et al. Cycling for Transportation in Sao Paulo City: Associations with Bike Paths, Trains and Subway Stations. **International Journal of Environmental Research and Pulic Health**, 2018. 562. Disponível em: <<https://doi.org/10.3390/ijerph15040562>>.

FRICKER, C.; GAST, N. Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity. **EURO Journal on Transportation and Logistics**, v. 5, n. 3, p. 261-291, 2016. Disponível em: <<https://doi.org/10.1007/s13676-014-0053-5>>.

GEHRKE, S. R.; WELCH, T. F. A bikeshare station area typology to forecast the station-level. **The Journal of Transport and Land Use**, v. 12, n. 1, p. 221-235, 2019. Disponível em: <<https://doi.org/10.5198/jtlu.2019.1395>>.

GTFS. GTFS: Making Public Transit Data Universally Accessible. **GTFS**, 2020. Disponível em: <<https://gtfs.org/>>. Acesso em: 01 jun. 2020.

GU, T.; KIM, I.; CURRIE, G. Measuring immediate impacts of a new mass transit system on an existing bike-share system in China. **Transportation Research Part A**, n. 124, p. 20-39, 2019. Disponível em: <<https://doi.org/10.1016/j.tra.2019.03.003>>.

HESS, S.; PALMA, D. Apollo: a flexible, powerful and customisable freeware package for choice model estimation and application. **Journal of Choice Modelling**, v. 32, set. 2019.

HU, S. et al. Examining spatiotemporal changing patterns of bike-sharing usage during COVID-19 pandemic. **Journal of Transport Geography**, v. 91, p. 16, fev 2021. Disponível em: <<https://doi.org/10.1016/j.jtrangeo.2021.102997>>.

IBGE. Cidades. **Instituto Brasileiro de Geografia e Estatística**, 2020. Disponível em: <<https://cidades.ibge.gov.br/brasil/sp/sao-paulo/panorama>>. Acesso em: 15 set 2020.

ITDP. Guia de planejamento de sistemas de bicicletas compartilhadas. **Instituto de Políticas de Transporte e Desenvolvimento**, p. 154, 2014. Disponível em: <https://www.itdp.org/wpcontent/uploads/2013/12/ITDP-Brasil_Guia-de-Planejamento-de-Sistemas-de-Bicicletas-Compartilhadas.pdf>. Acesso em: 10 novembro 2021.

KOPPELMAN, F. S.; BHAT, C. A Self Instructing Course in Mode Choice Modeling: Multinomial and Nested Logit Models, 2006.

LANDEX, A.; HANSEN, S.; ANDERSEN, J. Examination of catchment areas for public transport. **Traffic days**, p. 1-16, 2006.

LATHIA, N.; AHMED, S.; CAPRA, L. Measuring the impact of opening the London shared bicycle scheme to casual users. **Transportation Research Part C: Emerging Technologies**, v. 22, p. 88-102, 2012. Disponível em: <<https://doi.org/10.1016/j.trc.2011.12.004>>.

LEVY, N.; GOLANI, C.; BEN-ELIA, E. An exploratory study of spatial patterns of cycling in Tel Aviv using passively generated bike-sharing data. **Journal of Transport Geography**, v. 76, p. 325-334, 2019. Disponível em: <<https://doi.org/10.1016/j.jtrangeo.2017.10.005>>.

LIN, D. et al. The analysis of catchment areas of metro stations using trajectory data generated by dockless shared bikes. **Sustainable Cities and Society**, v. 49, 2019.

MACQUEEN, J. B. **Some methods for classification and analysis of multivariate observations**. Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics. Berkeley, Califórnia: University of California Press. 1967. p. 281-297.

METRÔ-SP. Caracterização socioeconômica dos deslocamentos. **Companhia do Metropolitano de São Paulo**, 2020a. Disponível em: <http://www.metro.sp.gov.br/pesquisa-od/arquivos/Caracteriza%C3%A7%C3%A3o_Socioecon%C3%B4mica_dos_Deslocamentos_2017.pdf>. Acesso em: 15 set. 2020.

METRÔ-SP. Pesquisa Origem e Destino de 2017. **Companhia do Metropolitano de São Paulo**, 2020b. Disponível em: <http://www.metro.sp.gov.br/pesquisa-od/arquivos/Ebook%20Pesquisa%20OD%202017_final_240719_versao_4.pdf>. Acesso em: 01 set. 2020.

MIX, R.; HURTUBIA, R.; RAVEAU, S. Optimal location of bike-sharing stations: A built environment and accessibility approach. **Transportation Research Part A**, v. 160, p. 126-142, June 2022. Disponível em: <<https://doi.org/10.1016/j.tra.2022.03.022>>.

MORAN, P. A. P. The Interpretation of Statistical Maps. **Journal of the Royal Statistical Society**, v. 10, n. 2, p. 243-251, 1948. Disponível em: <<https://doi.org/10.1111/j.2517-6161.1948.tb00012.x>>.

PAULO, C. D. S. Estrutura. **Serviços para o cidadão**, 2022. Disponível em: <<https://www.capital.sp.gov.br/cidadao/transportes/transporte-publico/estrutura>>. Acesso em: 30 ago. 2022.

PRITCHARD, J. P. et al. Potential impacts of bike-and-ride on job accessibility and spatial equity in São Paulo, Brazil. **Transportation Research Part A**, 2019. 386-400. Disponível em: <<https://doi.org/10.1016/j.tra.2019.01.022>>.

RADZIMSKI, A.; DZIĘCIELSKI, M. Exploring the relationship between bike-sharing and public transport in Poznań, Poland. **Transportation Research Part A: Policy and Practice**, v. 145, p. 189-202, 2021. Disponível em: <<https://doi.org/10.1016/j.tra.2021.01.003>>.

ROUSSEUW, P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. **Journal of Computational and Applied Mathematics**, v. 20, p. 53-65, Novembro 1987. Disponível em: <[https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)>.

SARAIVA, P. et al. Avaliação da influência do entorno no uso das estações de bicicletas compartilhadas. **Revista Brasileira de Gestão Urbana**, 2019. 15. Disponível em: <<https://doi.org/10.1590/2175-3369.011.001.AO13>>.

SCHIMOHR, K.; SCHEINER, J. Spatial and temporal analysis of bike-sharing use in Cologne taking into account a public transit disruption. **Journal of Transport Geography**, v. 92, April 2021. Disponível em: <<https://doi.org/10.1016/j.jtrangeo.2021.103017>>.

SHAHEEN, S. A.; COHEN, A.; MARTIN, E. Public Bikesharing in North America: Early Operator Understanding and Emerging Trends. **Transportation Research Record**, v. 2387, n. 1, p. 83-92, Dezembro 2013. Disponível em: <<https://doi.org/10.3141%2F2387-10>>.

SI, H. et al. Mapping the bike sharing research published from 2010 to 2018: A scientometric review. **Journal of Cleaner Production**, v. 213, p. 415-427, 2019. Disponível em: <<https://doi.org/10.1016/j.jclepro.2018.12.157>>.

SPTRANS. Pontos de Parada. **Secretaria Municipal de Transporte e Mobilidade Urbana**, 2022. Disponível em: <<https://www.sptrans.com.br/terminais-corredores-e-pontos-de-parada/pontos-de-parada/>>. Acesso em: 30 ago. 2022.

TRAIN, K. **Discrete Choice Methods with Simulation**. 2nd. ed. [S.l.]: Cambridge University Press, 2009. Disponível em: <<https://eml.berkeley.edu/books/choice2.html>>.

TRAN, D. T.; OVTRACHT, N.; D'ARCIER, F. B. Modeling Bike Sharing System using Built Environment Factors. **Procedia CIRP**, v. 30, p. 293-298, 2015. Disponível em: <<https://doi.org/10.1016/j.procir.2015.02.156>>.

VERMA, T. et al. Extracting spatiotemporal commuting patterns from public transit data. **Jornal of Urban Mobility**, v. 1, December 2021. Disponível em: <<https://doi.org/10.1016/j.urbmob.2021.100004>>.

WANG, X. et al. Modeling bike share station activity: Effects of nearby business and jobs on trips to and from stations. **Journal of Urban Planning and Development**, v. 142, n. 1, p. 04015001, 2016. Disponível em: <[https://doi.org/10.1061/\(ASCE\)UP.1943-5444.0000273](https://doi.org/10.1061/(ASCE)UP.1943-5444.0000273)>.

WANG, X. et al. Bike sharing rebalancing problem with variable demand. **Physica A: Statistical Mechanics and its Applications**, v. 591, 2021.

WU, C.; KIM, I.; CHUNG, H. The effects of built environment spatial variation on bike-sharing usage: A case study of Suzhou, China. **Cities**, v. 110, March 2021. Disponivel em: <<https://doi.org/10.1016/j.cities.2020.103063>>.

ZHANG, L. et al. Sustainable bike-sharing systems: characteristics and commonalities across cases in urban China. **Journal of Cleaner Production**, v. 97, p. 124-133, 2015. Disponivel em: <<https://doi.org/10.1016/j.jclepro.2014.04.006>>.

ZHU, R. et al. Understanding spatio-temporal heterogeneity of bike-sharing and scooter-sharing mobility. **Computers, Environment and Urban Systems**, v. 81, mai 2020. Disponivel em: <<https://doi.org/10.1016/j.compenvurbsys.2020.101483>>.

ANEXO A – FONTE DE INFORMAÇÃO DAS VARIÁVEIS EXPLICATIVAS

Este apêndice tem como objetivo apresentar as fontes de informação utilizadas para a obtenção das variáveis explicativas.

Tembici

Algumas variáveis explicativas foram calculadas com base nas informações fornecidas pela Tembici. É importante destacar que a distribuição da variável idade pode apresentar um desvio devido à falta de informação da chave identificadora do usuário no banco de dados de viagem. Os indivíduos podem então aparecer com frequências diferentes no banco de dados, viesando a distribuição da variável.

Ciclomapa

As variáveis de infraestrutura cicloviária foram coletadas na plataforma Ciclomapa, que é uma ferramenta desenvolvida pelo ITDP Brasil em parceria com a União de Ciclistas do Brasil (UCB). O acesso e importação das informações é realizado dentro do próprio site: <https://ciclomapa.org.br/>. É importante destacar que esta informação foi preferida à informação da CET por estar mais atualizada.

Geosampa

O Geosampa é o mapa digital da Prefeitura da Cidade de São Paulo, contendo informações de limites administrativos, população, dados ambientais, redes de transporte, patrimônio histórico, entre outros. Ele possui livre acesso por meio do site: <http://geosampa.prefeitura.sp.gov.br/>.

SPTRANS E EMTU

Para as informações do sistema de ônibus de São Paulo, foram consultados os sites das operadoras do sistema municipal (SPTRANS) e intermunicipal (EMTU). As informações coletadas são provenientes de arquivos GTFS⁵, que contém os pontos de parada do sistema.

Centro de Estudos da Metrópole (CEM)

O CEM é um dos centros de pesquisa da FAPESP. Este centro disponibiliza para acesso público algumas bases de dados, por meio do site:

⁵ GTFS: arquivo digital contendo dados de transporte, como tarifas, itinerário, partidas, linhas e pontos de paradas, previsões de chegada, posições dos veículos, avisos (GTFS, 2020).

<https://centrodametropole.fflch.usp.br/pt-br/download-de-dados>. É importante destacar que os dados consultados dessa fonte de informação são referentes ao Cadastro Central de Empresas, e têm como ano de referência 1999. A informação mais recente deste cadastro não é disponibilizada para acesso público por questões de confidencialidade. A alternativa para essa fonte de informação é o OSM, que embora seja mais recente, não tem o mesmo volume de dados registrados nas bases do CEM.

Pesquisa Origem e Destino (POD)

A POD da Região Metropolitana de São Paulo tem sido realizada e disponibilizada pela Companhia do Metrô a cada 10 anos, desde 1967 até 2017, sendo esta última a mais recente disponível. Esta informação pode ser consultada no site: <https://transparencia.metrosp.com.br/dataset/pesquisa-origem-e-destino/resource/4362eaa3-c0aa-410a-a32b-37355c091075>. Em particular, para os dados da renda familiar, foram utilizados os resultados das tabelas gerais da POD.

ANEXO B – ESPECIFICAÇÕES PARA CONSTRUÇÃO DO PERFIL DE USO

Para a definição do período de análise e das variáveis consideradas para a construção dos perfis temporais, foram avaliadas algumas premissas alternativas. À título de conhecimento, as especificações alternativas consideradas estão apresentadas na Tabela B1.

Tabela B1 – Especificações alternativas formação de grupos de perfil temporal de uso das estações

Especificação alternativa	Descrição
Período mensal	Consideração de fevereiro e março separadamente
Semana completa	Consideração do fim de semana
Apenas 3 ^a , 4 ^a e 5 ^a feira	Desconsideração de segundas e sextas-feiras
Período horário	Redução de 24 horas para o intervalo das 06:00 às 22:59
Fluxo máximo	Adição da 49 ^a variável $f_{\max,m}$
Fluxos máximos, retirada/devolução	Adição da 49 ^a e 50 ^a variáveis $f_{\max,m}$ e $f_{0\max,m}$
Tamanho da estação	Adição da 49 ^a variável s_m

Fonte: Elaboração própria, 2022.

Primeiramente, procurou-se avaliar como o período de análise afeta o agrupamento, considerando diferentes especificações sobre a quantidade e o tipo de dias de análise, bem como a consideração de diferentes períodos horários. Em seguida, para avaliar se estações com maiores fluxos (retirada e/ou devolução) e/ou maior tamanho (número de docas) comportam-se de modo distinto, estas grandezas foram adicionadas ao vetor de 48 variáveis que descreve o perfil temporal das estações. Com isso, buscou-se introduzir uma medida absoluta do nível de atividade da estação, já que os fluxos normalizados são valores relativos.

O resultado desta análise demonstrou que em geral as especificações alternativas convergiram para os grupos similares de perfil temporal, não sendo constatada uma migração significativa de estações entre grupos. A única especificação que trouxe uma contribuição significativa e positiva foi a redução da janela horária de 24 para 17 horas, aumentando o coeficiente de silhueta em 32%. Sendo assim, esta foi a única premissa alternativa incorporada na especificação final.

ANEXO C – NÚMERO DE ESTAÇÕES NAS ZONAS DE COBERTURA DO SISTEMA

A Tabela C1 apresenta o número de estações em cada uma das zonas selecionadas da POD de 2017, que estão na área de cobertura do sistema Bikesampa.

Tabela C1 – Número de estações nas zonas da área de cobertura do sistema Bikesampa

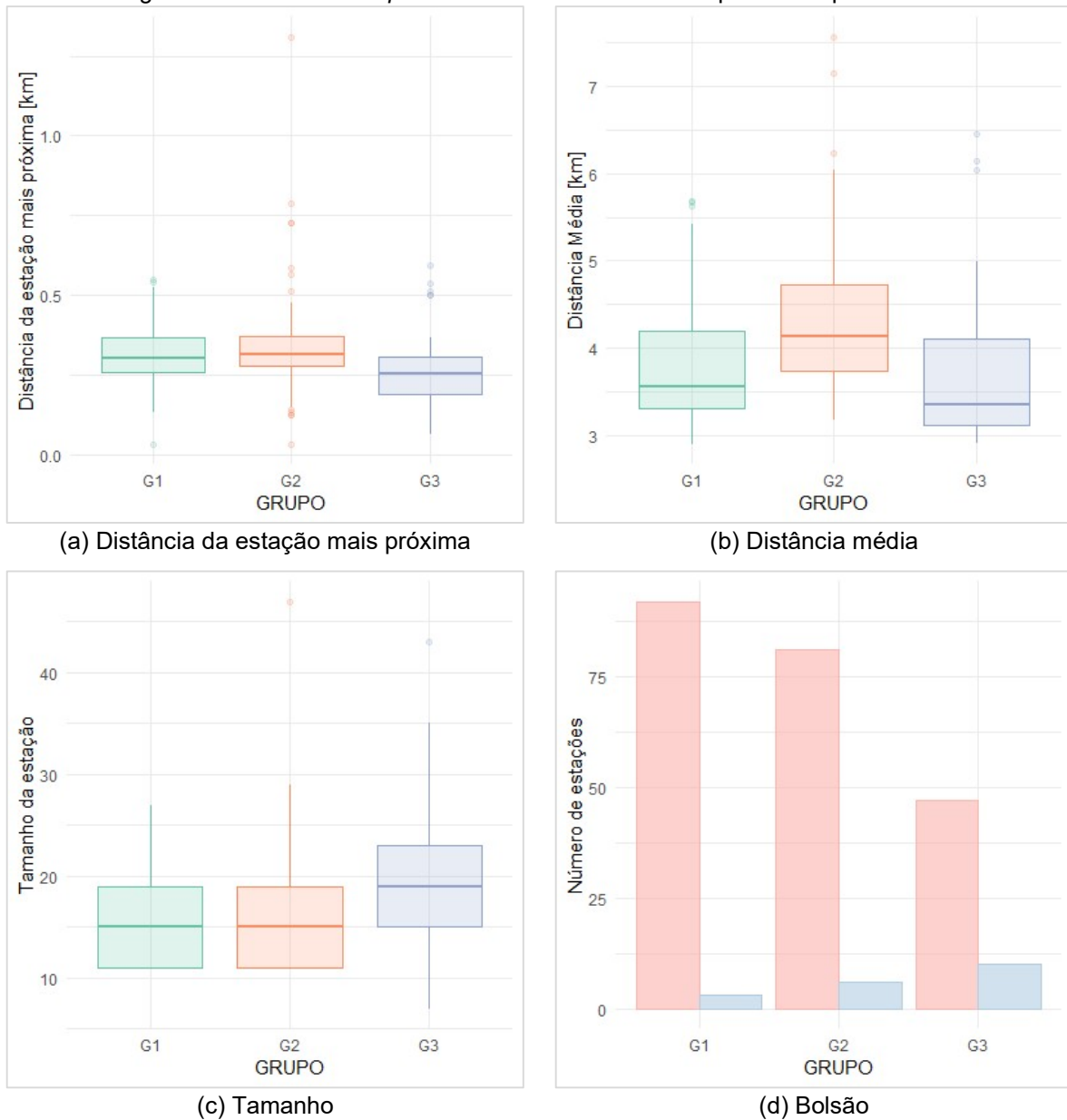
Zona	Grupo 1 – Balanceado	Grupo 2 – Retiradas manhã	Grupo 3 – Devoluções manhã
Alto de Pinheiros	1	2	3
Ana Rosa	0	1	0
Bandeirantes	3	1	0
Bela Vista	2	1	0
Berrini	2	3	15
Bexiga	0	2	0
Boaçava	0	3	2
Brooklin	2	8	0
Butantã	0	1	0
Campinas	1	1	0
Campo Belo	0	1	0
Centro Cultural	1	0	0
Chácara Flora	0	1	0
Chácara Itaim	7	0	9
Clínicas	1	5	0
Consolação	3	0	0
FAAP	1	0	0
França Pinto	2	5	0
Granja Julieta	0	1	1
Higienópolis	2	0	0
Hélio Pelegrino	4	1	2
Jardim Europa	2	2	8
Jardim Luzitânia	0	2	0
Jardim Paulistano	4	3	0
Jardins	3	1	3
Joaquim Nabuco	0	0	0
Ladeira da Memória	0	2	0
Liberdade	0	0	0
Marechal Deodoro	0	3	0
Masp	2	1	0
Moema	0	2	0
Oscar Freire	8	0	0
Pamplona	4	0	0
Paraíso	2	7	1
Parque Ibirapuera	6	1	3
Pinheiros	7	3	1
Praça João Mendes	0	0	0
República	3	0	0
Rodrigues Alves	0	5	0
Sé	2	0	0
Santa Cecília	1	1	0
Santa Ifigênia	0	0	0
São Carlos do Pinhal	3	0	0
Treze de Maio	1	0	0
Trianon	2	0	1
Vieira de Moraes	0	1	0
Vila Beatriz	0	0	0
Vila Buarque	2	0	0
Vila Clementino	1	6	0
Vila Cordeiro	0	4	0
Vila Helena	2	4	0
Vila Madalena	3	2	0
Vila Nova Conceição	2	0	2
Vila Olímpia	3	0	6

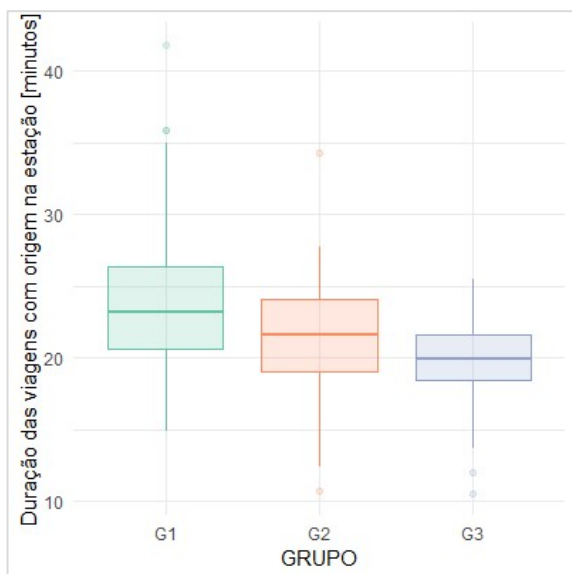
Fonte: Elaboração própria, 2022.

ANEXO D – DISTRIBUIÇÃO DAS VARIÁVEIS EXPLICATIVAS NOS GRUPOS

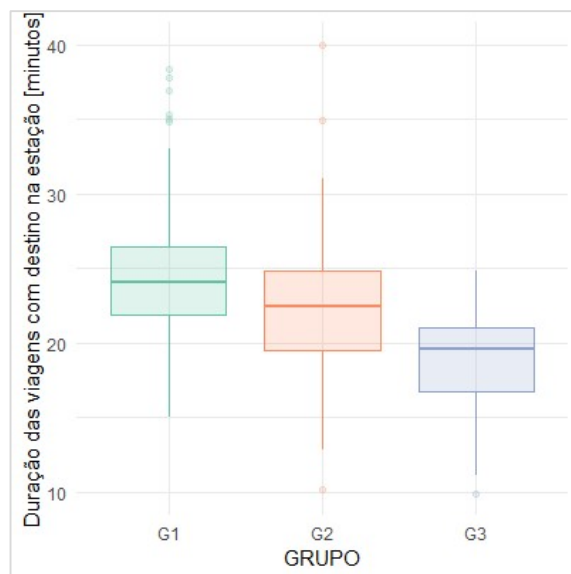
Para fins de ilustração, a Figura D1 apresenta a distribuição das variáveis explicativas segundo cada grupo de perfil temporal. Os rótulos G1, G2 e G3 representam respectivamente o **Grupo 1 – Balanceado**, o **Grupo 2 – Retiradas manhã** e o **Grupo 3 – Devoluções manhã**. Para os gráficos de colunas, a coluna vermelha representa o número de estações sem o atributo (0: não), enquanto as colunas em azul indicam o número de estações com o atributo (1: sim).

Figura D1 – Gráficos *boxplot* e colunas das variáveis explicativas quantitativas

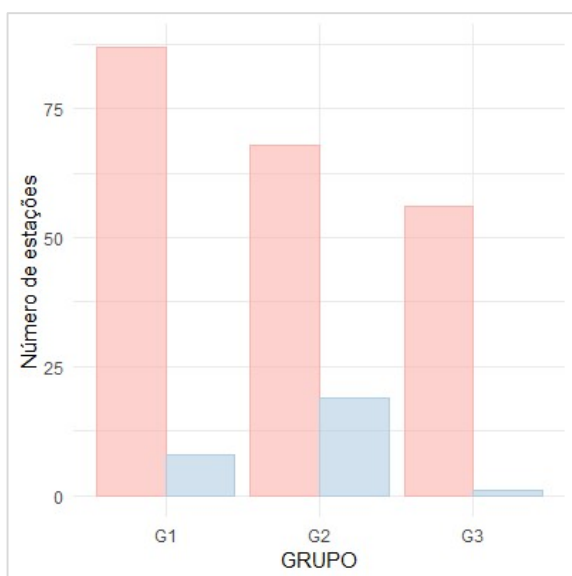




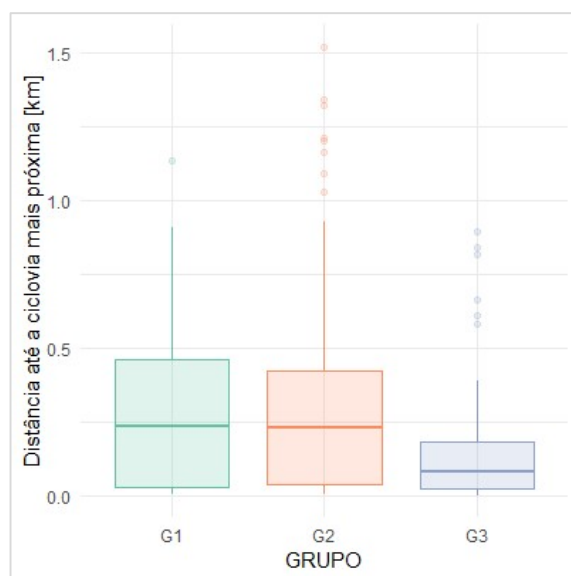
(e) Duração origem



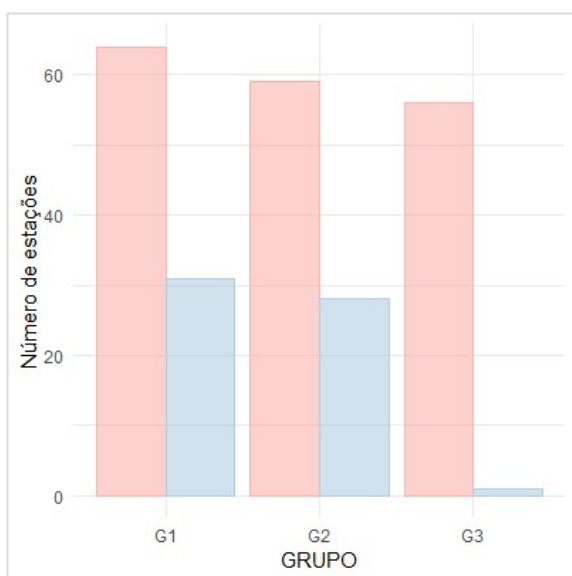
(f) Duração destino



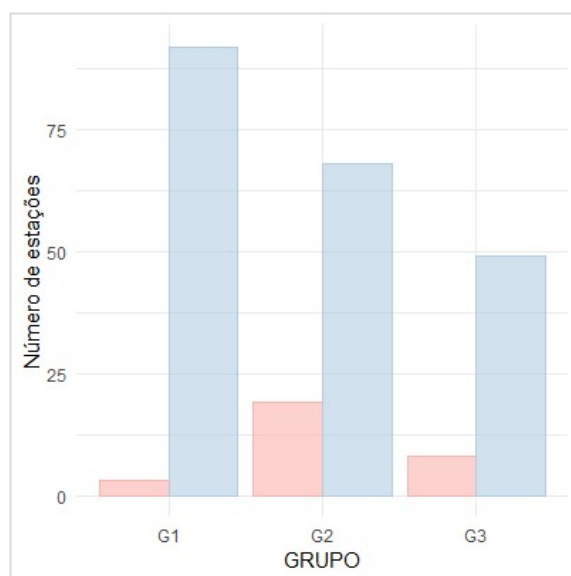
(g) Bicletário



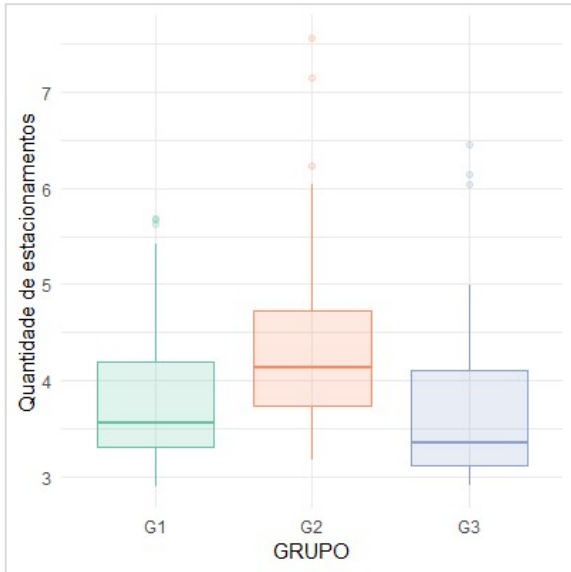
(h) Distância de vias cicláveis



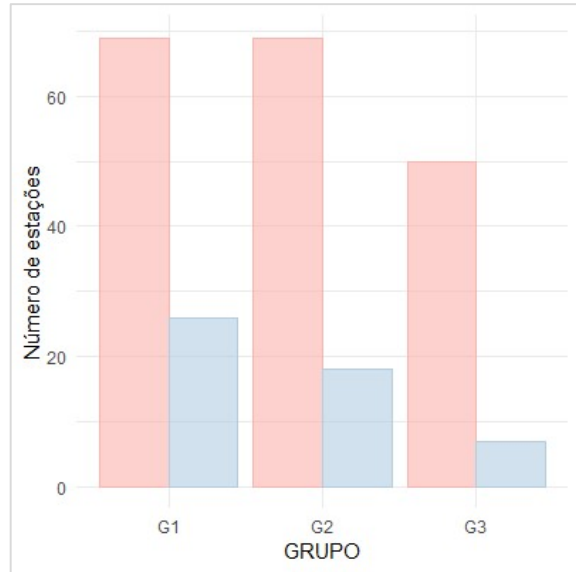
(i) Metrô



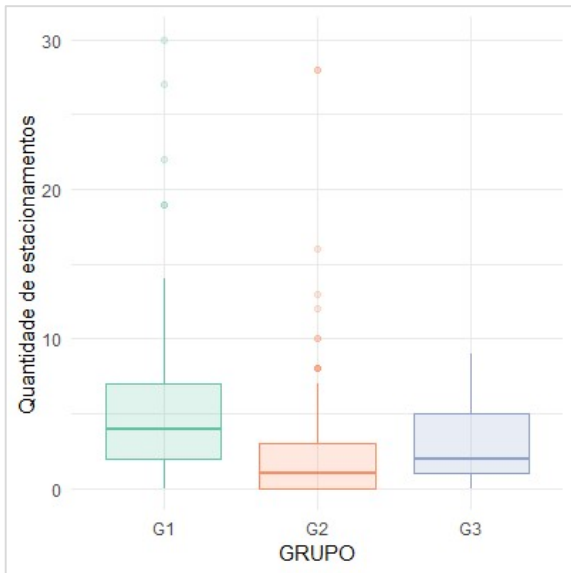
(j) Zona azul



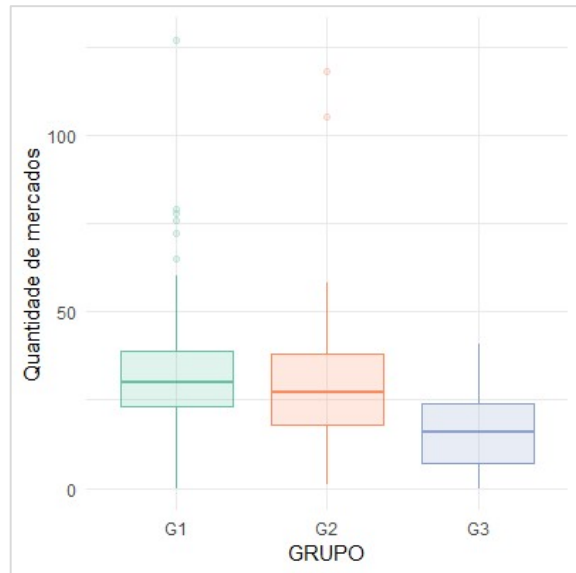
(k) Estacionamento



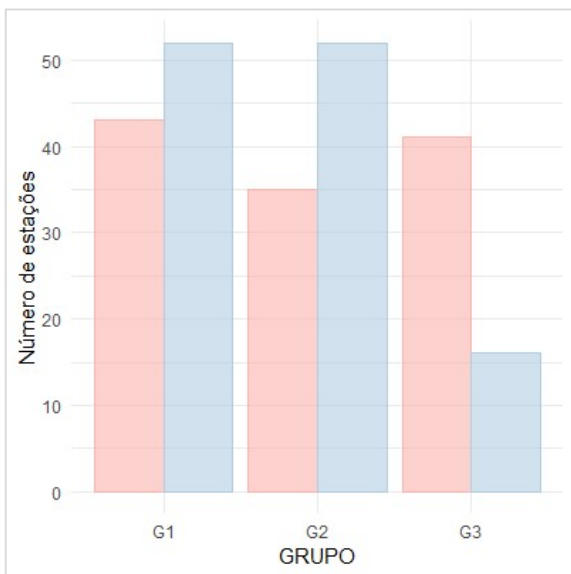
(l) Museu



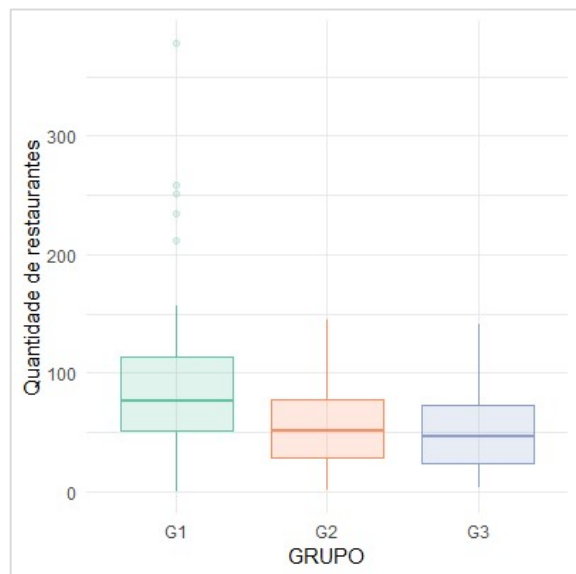
(m) Vida noturna



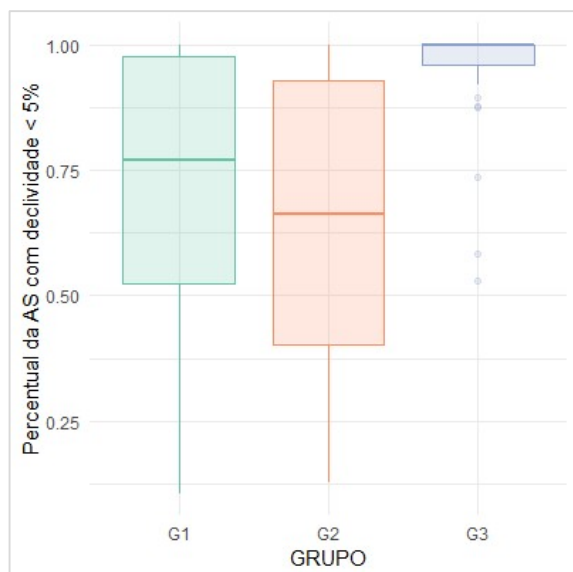
(n) Mercado



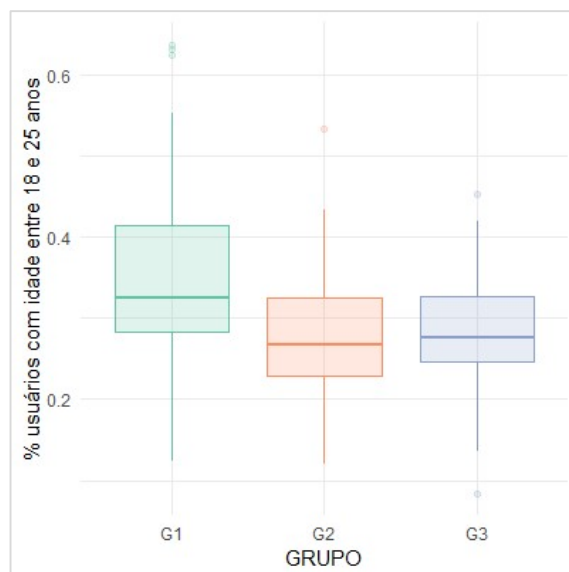
(o) Estudo



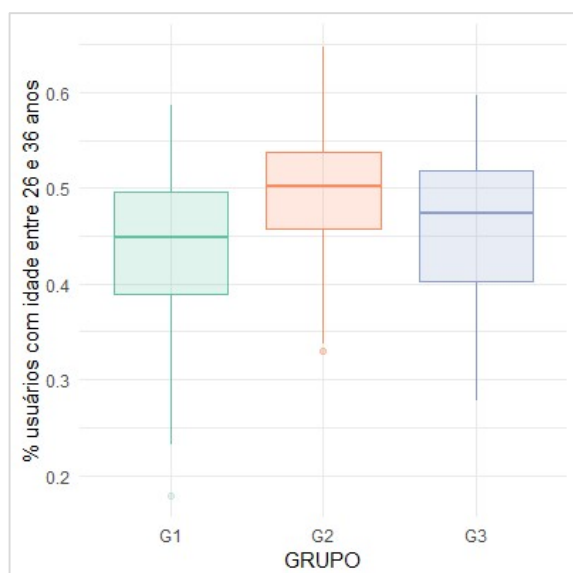
(p) Restaurantes



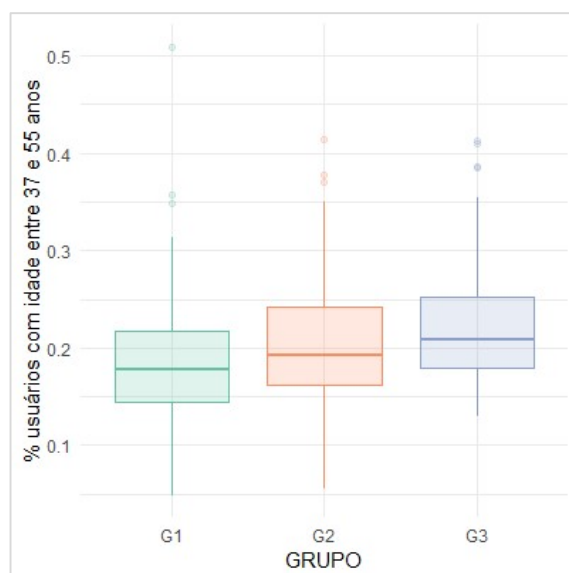
(q) Declividade



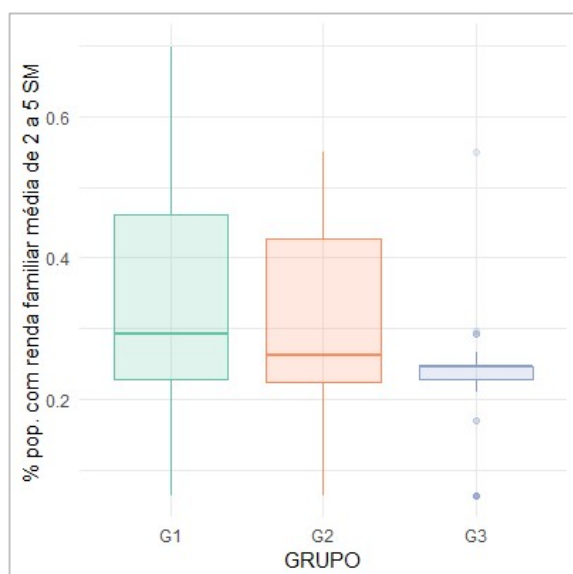
(r) Usuários entre 18 e 25 anos



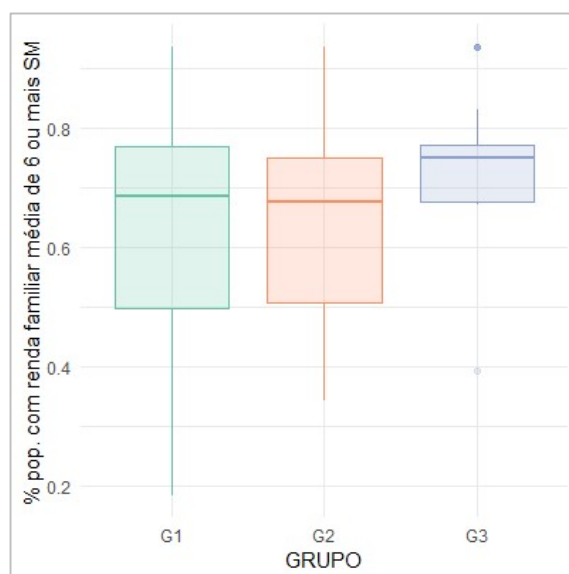
(s) Usuários entre 26 e 36 anos



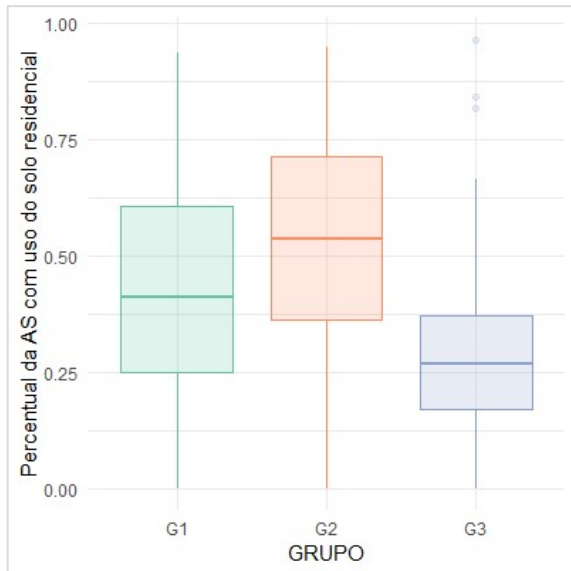
(t) Usuários entre 37 e 55 anos



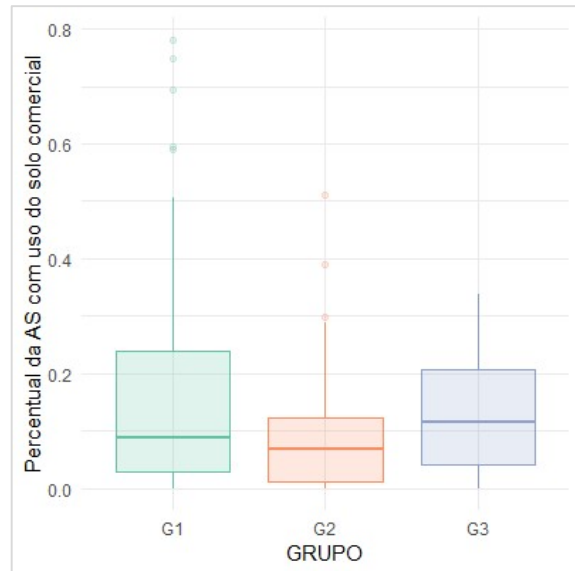
(u) Renda familiar média de 2 e 5 SM



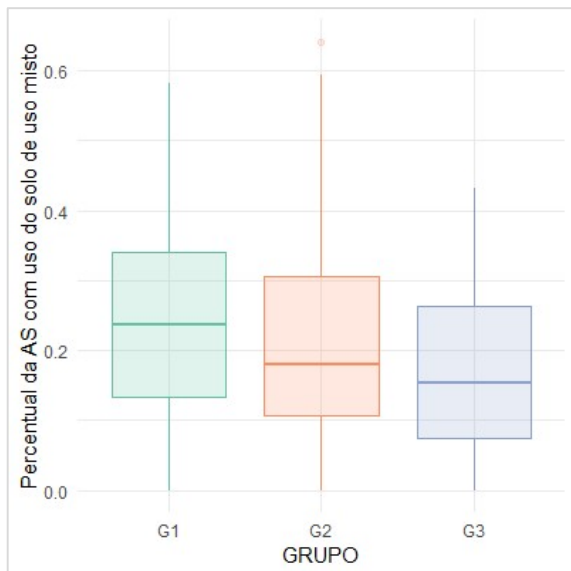
(v) Renda familiar média de 6 ou mais SM



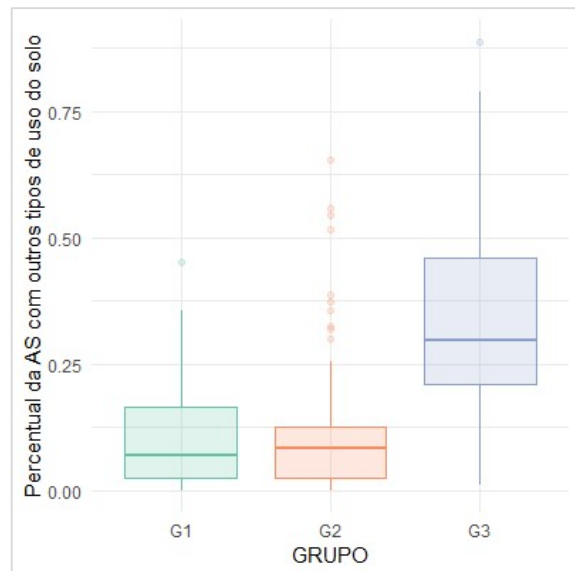
(w) Uso do solo residencial



(x) Uso do solo comercial



(y) Uso do solo misto



(z) Outros tipos de uso do solo

Fonte: Elaboração própria, 2022.