

UNIVERSIDADE FEDERAL DE MINAS GERAIS  
CURSO DE MESTRADO EM GEOTECNIA E TRANSPORTES

Luísa Tavares Muzzi de Sousa

**INFLUÊNCIA DOS FATORES SOCIOECONÔMICOS E DO ACESSO AO VAREJO  
TRADICIONAL NAS ENTREGAS DO COMÉRCIO ELETRÔNICO: análise  
econométrica e espacial para Belo Horizonte, Uberlândia e Patos de Minas**

Belo Horizonte

2022

Luísa Tavares Muzzi de Sousa

**INFLUÊNCIA DOS FATORES SOCIOECONÔMICOS E DO ACESSO AO VAREJO  
TRADICIONAL NAS ENTREGAS DO COMÉRCIO ELETRÔNICO: análise  
econométrica e espacial para Belo Horizonte, Uberlândia e Patos de Minas**

Dissertação apresentada ao Curso de Mestrado em Geotecnia e Transportes da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Mestre em Geotecnia e Transportes.

Área de concentração: Transportes

Orientadora: Leise Kelli de Oliveira

Coorientador: Bruno Vieira Bertoncini

Belo Horizonte

2022

S237a

Sousa, Luísa Tavares Muzzi de.

Influência dos fatores socioeconômicos e do acesso ao varejo tradicional nas entregas do comércio eletrônico [recurso eletrônico]: análise econométrica e espacial para Belo Horizonte, Uberlândia e Patos de Minas/ Luísa Tavares Muzzi de Sousa. – 2022.

1 recurso online (113 f. : il., color.) : pdf.

Orientadora: Leise Kelli de Oliveira.

Coorientador: Bruno Vieira Bertoncini.

Dissertação (mestrado) - Universidade Federal de Minas Gerais, Escola de Engenharia.

Bibliografia: f. 105-113.

Exigências do sistema: Adobe Acrobat Reader.

1. Transportes - Teses. 2. Comércio eletrônico – Teses.  
3. Econometria – Teses. 4. Transporte de mercadorias – Teses.  
I. Oliveira, Leise Kelli. II. Bertoncini, Bruno Vieira.  
III. Universidade Federal de Minas Gerais. Escola de Engenharia.  
IV. Título.

CDU: 656(043)



UNIVERSIDADE FEDERAL DE MINAS GERAIS  
ESCOLA DE ENGENHARIA  
COLEGIADO DO CURSO MESTRADO EM GEOTECNIA E TRANSPORTES

### FOLHA DE APROVAÇÃO

#### INFLUÊNCIA DOS FATORES SOCIOECONÔMICOS E DO ACESSO AO VAREJO TRADICIONAL NAS ENTREGAS DO COMÉRCIO ELETRÔNICO: ANÁLISE ECONOMÉTRICA E ESPACIAL PARA BELO HORIZONTE, UBERLÂNDIA E PATOS DE MINAS

LUÍSA TAVARES MUZZI DE SOUSA

Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em GEOTECNIA E TRANSPORTES, como requisito para obtenção do grau de Mestre em GEOTECNIA E TRANSPORTES, área de concentração TRANSPORTES. Aprovada em 19 de janeiro de 2022, pela banca constituída pelos seguintes professores:

Prof.ª. Leise Kelli de Oliveira - Orientadora - UFMG

Prof. Bruno Vieira Bertoncini - Coorientador - UFC

Prof. Cassiano Augusto Isler - USP

Prof.ª. Ana Margarita Larrañaga Uriarte - UFRGS

Belo Horizonte, 19 de janeiro de 2022.



Documento assinado eletronicamente por **Cassiano Augusto Isler, Usuário Externo**, em 19/01/2022, às 11:32, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Leise Kelli de Oliveira, Professora do Magistério Superior**, em 19/01/2022, às 11:33, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Bruno Vieira Bertoncini, Usuário Externo**, em 19/01/2022, às 11:33, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Ana Margarita Larranaga Uriarte, Usuário Externo**, em 19/01/2022, às 11:50, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site [https://sei.ufmg.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **1155424** e o código CRC **56F35B88**.

## **AGRADECIMENTOS**

Ao final desta etapa, quero expressar meu sentimento de gratidão à minha orientadora Leise, que tanto me incentiva, me acolhe com paciência e sabedoria e me inspira na minha trajetória. Gratidão também ao meu coorientador Bruno pela disponibilidade e conselhos.

Agradeço aos meus pais, Adriana e José Geraldo, e ao meu irmão Guilherme pelo amor, apoio e presença constante. Agradeço também ao Bernardo por me apoiar e acreditar em mim.

Agradeço a todos os amigos que fiz e a todas as pessoas que colaboraram comigo durante o mestrado, em especial à Isabela.

Gratidão também à FAPEMIG pela concessão da bolsa de estudos durante o mestrado.

E, por fim, gratidão a Deus por tornar tudo isso possível.

## RESUMO

SOUSA, Luísa Tavares Muzzi de. **Influência dos Fatores Socioeconômicos e do Acesso ao Varejo Tradicional nas Entregas do Comércio Eletrônico: Análise Econométrica e Espacial para Belo Horizonte, Uberlândia e Patos de Minas**. 113f. Dissertação (Mestrado em Geotecnia e Transportes) – Universidade Federal de Minas Gerais, Belo Horizonte, 2022.

O comércio eletrônico tem crescido devido às facilidades associadas à realização de compras *online*, como maior comodidade e menores preços, gerando um aumento do número de entregas. Este fenômeno tem consequências negativas para a mobilidade urbana, visto que, muitas vezes, o planejamento urbano não engloba o transporte urbano de cargas (TUC). As externalidades negativas do TUC incluem o crescimento do número de viagens de veículos de carga, poluição ambiental e sonora, congestionamentos, aumento do risco de acidentes de trânsito e danos à infraestrutura viária. O conhecimento das características relacionadas às entregas contribui para o planejamento urbano no que tange o TUC, auxiliando na implementação de políticas públicas visando uma maior sustentabilidade da cidade. Estudos prévios indicaram uma relação entre fatores socioeconômicos e o comércio eletrônico em determinada região. No entanto, identificou-se que há uma carência de trabalhos que explorem fatores espaciais e sua relação com o *e-commerce*, como o acesso a oportunidades de compra no varejo tradicional e a distribuição geográfica dos perfis demográficos de e-consumidores. Neste contexto, torna-se relevante o estudo da influência dos fatores socioeconômicos e do acesso às oportunidades de compra nas entregas do comércio eletrônico. Assim, o objetivo desta dissertação foi identificar a relação entre fatores socioeconômicos e oportunidade de acesso ao varejo tradicional e o número de entregas do comércio eletrônico em Belo Horizonte, Uberlândia e Patos de Minas. Para tanto, primeiramente verificou-se a influência destes fatores no número de entregas por meio de modelos de regressão binomial negativa. Para analisar a correlação espacial entre as entregas e os fatores socioeconômicos e a oportunidade de compra foi utilizada Regressão Geograficamente Ponderada em Belo Horizonte, onde, dentre as três cidades analisadas, se detectou dependência espacial das entregas. Os resultados obtidos para os modelos econométricos indicam a influência da renda e da idade como sendo um ponto em comum para as três cidades; o modelo espacial demonstrou a influência da renda e idade, bem como gênero, número de estabelecimentos comerciais e cor ou raça.

**Palavras-Chaves:** transporte urbano de carga, comércio eletrônico, fatores socioeconômicos, análise econométrica, regressão geograficamente ponderada

## ABSTRACT

SOUSA, Luísa Tavares Muzzi de. **Influência dos Fatores Socioeconômicos e do Acesso ao Varejo Tradicional nas Entregas do Comércio Eletrônico: Análise Econométrica e Espacial para Belo Horizonte, Uberlândia e Patos de Minas**. 113f. Dissertação (Mestrado em Geotecnia e Transportes) – Universidade Federal de Minas Gerais, Belo Horizonte, 2022.

The growth of online purchases is associated with the benefits of online shopping, such as greater convenience and lower prices, which generates an increase in the number of deliveries. This phenomenon has negative consequences for urban mobility since urban planning often does not include urban freight transport (UFT). The negative externalities of UFT include growth in the number of freight vehicle trips, environmental and noise pollution, congestion, increased risk of traffic accidents, and damage to the road infrastructure. Knowing the characteristics related to deliveries contributes to an UFT inclusive urban planning, which helps to implement public policies for greater city sustainability. Previous studies have indicated a relationship between socioeconomic factors and e-commerce in a particular region. However, there is a lack of studies that explore spatial factors and their relationship with e-commerce, such as access to shopping opportunities in traditional retail and the geographic distribution of demographic profiles of e-consumers. In this context, it is relevant to study the influence of socioeconomic factors and access to purchase opportunities in e-commerce deliveries. Therefore, the objective of this dissertation was to identify the relationship between socioeconomic factors and the opportunity to access traditional retail and the number of e-commerce deliveries in Belo Horizonte, Uberlândia, and Patos de Minas. To do so, firstly, the influence of these factors on the number of deliveries was verified using negative binomial regression models. To analyze the spatial correlation between deliveries and socioeconomic factors and purchase opportunity, Geographically Weighted Regression was used in Belo Horizonte, where, among the three cities analyzed, spatial dependency of the number of deliveries was detected. Results for the econometric models indicate the influence of income and age as being a common point between the three cities. The spatial model disclosed the influence of income and age, as well as gender, number of retail shops, and race.

**Keywords:** urban freight transport, online shopping, sociodemographic factors, econometric analysis, geographically weighted regression



## LISTA DE FIGURAS

Figura 1: Localização dos municípios de Belo Horizonte, Uberlândia e Patos de Minas. ....	35
Figura 2: Distribuição dos dados populacionais, renda, dados de domicílios, entregas e estabelecimentos comerciais em Belo Horizonte (IBGE, 2010; Transportadora, 2019; BHMap, 2020). ....	39
Figura 3: Distribuição da população de Belo Horizonte por cor ou raça (IBGE, 2010; BHMap, 2020). ....	40
Figura 4: Distribuição da população de Belo Horizonte por faixas de idade (IBGE, 2010; BHMap, 2020). ....	41
Figura 5: Distribuição dos dados populacionais, renda, dados de domicílios, entregas e estabelecimentos comerciais em Uberlândia (IBGE, 2010; Transportadora, 2019; OpenStreetMap, 2021). ....	43
Figura 6: Distribuição da população de Uberlândia por cor ou raça (IBGE, 2010). ....	44
Figura 7: Distribuição da população de Uberlândia por idade (IBGE, 2010). ....	45
Figura 8: Distribuição dos dados populacionais, renda, dados de domicílios, entregas e estabelecimentos comerciais em Patos de Minas (IBGE, 2010; Transportadora, 2019; OpenStreetMap, 2021). ....	47
Figura 9: Distribuição da população de Patos de Minas por cor ou raça (IBGE, 2010). ....	48
Figura 10: Distribuição da população de Patos de Minas por idade (IBGE, 2010). ....	49
Figura 11: Número de entregas realizadas em 2019 em Belo Horizonte, Uberlândia e Patos de Minas (Transportadora, 2019). ....	50
Figura 12: Fluxograma das etapas do método proposto. ....	51
Figura 13: Representação cartográfica de bairros analisados e bairros retirados da análise econométrica de Belo Horizonte. ....	63

Figura 14: Coeficientes espaciais das variáveis número de estabelecimentos comerciais, renda média e habitantes por domicílio para o Modelo GWR1. ....	68
Figura 15: Coeficientes espaciais das variáveis número de homens e número de mulheres para o Modelo GWR1. ....	69
Figura 16: Coeficientes espaciais das variáveis relacionadas às faixas etárias para o Modelo GWR1.....	70
Figura 17: Coeficientes espaciais das variáveis relacionadas à cor ou raça para o Modelo GWR1.....	72
Figura 18: R <sup>2</sup> local do Modelo GWR1. ....	74
Figura 19: Coeficientes espaciais das variáveis número de estabelecimentos comerciais, renda média e número de mulheres para o Modelo GWR2. ....	75
Figura 20: Coeficientes espaciais das variáveis relacionadas às faixas etárias para o Modelo GWR2.....	76
Figura 21: Coeficientes espaciais das variáveis relacionadas à cor ou raça para o Modelo GWR2.....	77
Figura 22: Coeficiente R <sup>2</sup> local do Modelo GWR2.....	78
Figura 23: Coeficientes espaciais das variáveis número de estabelecimentos comerciais e renda média para o Modelo GWR3.....	79
Figura 24: Coeficientes espaciais das variáveis relacionadas às faixas etárias para o Modelo GWR3.....	80
Figura 25: R <sup>2</sup> local do Modelo GWR3. ....	81
Figura 26: Representação cartográfica de bairros analisados e bairros retirados da análise econométrica de Uberlândia. ....	82
Figura 27: Representação cartográfica de bairros analisados e bairros retirados da análise econométrica de Patos de Minas. ....	88

## LISTA DE TABELAS

Tabela 1: Oportunidades de pesquisa .....	7
Tabela 2: Métodos utilizados nos estudos descritos .....	19
Tabela 3: Síntese dos principais aspectos dos estudos discutidos .....	20
Tabela 4: Métodos utilizados pelos estudos descritos .....	32
Tabela 5: Síntese dos principais aspectos dos estudos discutidos .....	33
Tabela 6: Características dos municípios analisados .....	37
Tabela 7: Estatística descritiva dos dados de Belo Horizonte .....	42
Tabela 8: Estatística descritiva dos dados de Uberlândia .....	45
Tabela 9: Estatística descritiva dos dados de Patos de Minas .....	49
Tabela 10: Resumo dos dados utilizados .....	52
Tabela 11: Resumo das variáveis utilizadas .....	54
Tabela 12: Estatística descritiva dos 142 bairros analisados em Belo Horizonte .....	64
Tabela 13: Dados estimados para o Modelo BH1 .....	65
Tabela 14: Dados estimados para o Modelo BH2 .....	66
Tabela 15: Estatística descritiva dos 26 bairros analisados em Uberlândia .....	82
Tabela 16: Dados estimados para o Modelo U1 .....	83
Tabela 17: Dados estimados para o Modelo U2 .....	84
Tabela 18: Dados estimados para o Modelo PM1 .....	86
Tabela 19: Dados estimados para o Modelo PM2 .....	87
Tabela 20: Estatística descritiva dos 25 bairros analisados em Patos de Minas.....	88

Tabela 21: Dados estimados para o Modelo PM3 .....	89
Tabela 22: Dados estimados para o Modelo PM4 .....	90
Tabela 23: Comparação entre modelos econométricos .....	92
Tabela 24: Comparação entre modelos espaciais .....	97

## LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

AIC	Crítério de Informação de Akaike
AICc	Crítério de Informação de Akaike corrigido
ANOVA	Análise de variância
B2C	<i>Business to Consumer</i>
CV	<i>Cross-validation</i>
GWR	<i>Geographically Weighted Regression</i>
IBGE	Instituto Brasileiro de Geografia e Estatística
IDHM	Índice de Desenvolvimento Humano Municipal
PIB	Produto Interno Bruto
TUC	Transporte Urbano de Carga
VIF	<i>Variance Inflation Factor</i>

# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>4</b>
<b>2</b>	<b>REVISÃO DA LITERATURA.....</b>	<b>9</b>
2.1	INFLUÊNCIA DOS FATORES SOCIOECONÔMICOS NO COMÉRCIO ELETRÔNICO NAS ÁREAS DE MARKETING, ECONOMIA E GESTÃO, COMUNICAÇÃO E CIÊNCIAS SOCIAIS.....	9
2.2	INFLUÊNCIA DOS FATORES SOCIOECONÔMICOS E ACESSO AO VAREJO TRADICIONAL NO COMÉRCIO ELETRÔNICO EM PESQUISAS NAS ÁREAS DE GEOGRAFIA E TRANSPORTES .....	21
<b>3</b>	<b>ÁREA DE ESTUDO .....</b>	<b>35</b>
3.1	DESCRIÇÃO DAS ÁREAS DE ESTUDO .....	36
3.2	DADOS.....	37
<b>4</b>	<b>MÉTODO DE PESQUISA .....</b>	<b>51</b>
4.1	DADOS.....	51
4.2	ANÁLISE ECONOMÉTRICA .....	52
4.3	ANÁLISE ESPACIAL.....	55
4.3.1	<i>Verificação da existência de dependência espacial das entregas do comércio eletrônico .....</i>	<i>55</i>
4.3.2	<i>Regressão Geograficamente Ponderada .....</i>	<i>57</i>
<b>5</b>	<b>RESULTADOS .....</b>	<b>63</b>
5.1	BELO HORIZONTE.....	63
5.1.1	<i>Análise econométrica .....</i>	<i>63</i>
5.1.2	<i>Análise espacial .....</i>	<i>66</i>
5.2	UBERLÂNDIA .....	81
5.2.1	<i>Análise econométrica .....</i>	<i>81</i>
5.2.2	<i>Análise espacial .....</i>	<i>85</i>
5.3	PATOS DE MINAS.....	86
5.3.1	<i>Análise econométrica .....</i>	<i>86</i>
5.3.2	<i>Análise espacial .....</i>	<i>91</i>

<b>6</b>	<b>DISCUSSÃO .....</b>	<b>92</b>
<b>7</b>	<b>CONCLUSÃO.....</b>	<b>102</b>
<b>8</b>	<b>REFERÊNCIAS.....</b>	<b>105</b>

## 1 INTRODUÇÃO

Em países da Europa como Dinamarca, Alemanha, Holanda, Noruega, Suécia e Reino Unido, calcula-se que 80% dos usuários da *Internet* são e-consumidores (UNCTAD, 2019). As facilidades advindas da realização de compras *online*, como menores preços, maior comodidade (Zhou *et al.*, 2007; Bucko *et al.*, 2018) e economia de tempo (Harris *et al.*, 2017; Rahman *et al.*, 2018) são motivos que contribuíram para o aumento da preferência do consumidor pelo *e-commerce*. Estima-se que em 2017, aproximadamente 1,3 bilhões de pessoas maiores de 15 anos fizeram uso do comércio eletrônico B2C (*business to consumer*) no mundo; 12% a mais que em 2016 (UNCTAD, 2019). Em 2019, estima-se que este número tenha sido de 1,48 bilhões de pessoas (7% a mais que em 2018), com as vendas totalizando \$4.870 bilhões (UNCTAD, 2021).

Em 2018, as vendas *online* do varejo no Reino Unido totalizaram \$84 bilhões; nos Estados Unidos o valor foi de \$519,6 bilhões, enquanto na China, o comércio eletrônico varejista arrecadou \$1.060,4 bilhões. Em 2019, esses números aumentaram para \$89 bilhões no Reino Unido, \$598 bilhões nos Estados Unidos e \$1.233,6 bilhões na China. Em 2020, devido à COVID-19 e às restrições à circulação de pessoas impostas em muitos países, a demanda pelo comércio eletrônico B2C cresceu 22% segundo dados dos países responsáveis por 65% do uso do *e-commerce* no mundo (UNCTAD, 2021).

O Brasil fez parte do grupo de 20 países com o maior número de vendas do comércio eletrônico B2C em 2019, ocupando a 20ª posição (UNCTAD, 2021). O aumento no número de pedidos alavancou o crescimento do comércio eletrônico B2C no país de 2011 a 2019, passando de uma média de 54 milhões de pedidos em 2011 para 148 milhões em 2019. Em 2011, o valor total das vendas realizadas por meio dessa modalidade foi de 18,7 bilhões de reais, alcançando R\$61,9 bilhões em 2019. Com relação ao ano de 2018, no ano de 2019, o *e-commerce* no país apresentou um crescimento de 16% no número de vendas. Enquanto o valor médio das compras diminuiu 4% de 2018 para 2019, a média anual de pedidos cresceu 21%. Apesar de as regiões brasileiras estarem em diferentes estágios de desenvolvimento do comércio eletrônico, todas apresentaram um aumento nas vendas, sendo o número de novos e-consumidores 10,7 milhões em 2019 (E-Bit, 2020).

Em 2020, as medidas restritivas adotadas em decorrência da pandemia do Coronavírus, como o isolamento social e as restrições ao comércio tradicional, contribuíram para que as vendas *online* B2C no Brasil alcançassem o valor de R\$87,4 bilhões, um crescimento de 41% em



relação a 2019. Ainda em 2020, o número de novos e-consumidores foi de 13,2 milhões, um aumento de 23% em relação a 2019, enquanto o número de pedidos aumentou 30%, alcançando 194 milhões de pedidos, e o valor médio das compras aumentou 8% (E-Bit, 2021).

Nota-se, portanto, que, de forma geral, o comércio eletrônico tem crescido no mundo nos últimos anos, crescimento que também se fez visível no Brasil, conforme dados apresentados. Embora a pandemia tenha contribuído para um forte crescimento do *e-commerce* em 2020, trata-se de uma situação atípica. Portanto, esta dissertação considerou dados “pré-pandemia”, pois tem a intenção de analisar as entregas do comércio eletrônico em uma situação de normalidade.

O aumento no número de pedidos impulsiona o aumento no número de entregas domiciliares (Cardenas *et al.*, 2016; Allen *et al.*, 2018). Cada pedido do comércio eletrônico pode gerar uma (ou mais entregas) que, em geral, tem por destino o local de moradia ou de trabalho do comprador. Atributos relacionados às entregas têm grande importância para o consumidor, que busca entregas rápidas e espera receber seus pedidos em dois ou três dias (Brusch e Stüber, 2014). O consumidor tem preferência por valor de frete baixo, sendo o frete grátis uma grande motivação para compras *online* (E-Bit, 2021) e para pedidos de entregas domiciliares (Nguyen *et al.*, 2019). Tais exigências do consumidor propõem um desafio às transportadoras e aos varejistas, que precisam manter a eficiência do serviço ao mesmo tempo em que equilibram os custos das entregas da última milha (Cardenas *et al.*, 2016; Allen *et al.*, 2018), que são os mais onerosos dentre os custos da distribuição urbana (Gevaers *et al.*, 2014).

Além do impacto financeiro às transportadoras e aos varejistas, as entregas domiciliares têm consequências negativas para a mobilidade urbana. O aumento do número de entregas acarreta no crescimento do número de viagens de veículos de transporte de carga (Allen *et al.*, 2018). O aumento da circulação de veículos, por sua vez, leva a externalidades negativas como o aumento dos congestionamentos e da poluição, tanto ambiental quanto sonora (Moroz e Polkowski, 2016). Além disso, a alta circulação de veículos de carga aumenta os riscos de acidentes e os danos à infraestrutura urbana (Park *et al.*, 2016). O problema é ainda maior quando se considera as taxas de entregas não realizadas, que requerem novas tentativas (Cardenas *et al.*, 2016).

As características das operações das entregas domiciliares muitas vezes estão em conflito com o planejamento urbano, voltado para outros setores da dinâmica do transporte urbano, como o transporte privado, o transporte público e o transporte não motorizado (Allen *et al.*, 2018). É

preciso, portanto, que exista uma preocupação com o planejamento do transporte de cargas para um aumento da qualidade de vida nas cidades (Taniguchi *et al.*, 2014).

As características da distribuição das entregas de produtos do comércio eletrônico estão associadas a fatores relacionados a vários *stakeholders* envolvidos no comércio eletrônico e no planejamento urbano, como por exemplo os consumidores, as transportadoras e os agentes governamentais (Tadić e Veljović, 2020). Além disso, estudos prévios indicaram uma relação entre fatores socioeconômicos e o comércio eletrônico em determinada região. Alguns estudos que investigaram essa relação foram feitos para outras áreas, principalmente o *marketing* (Bellman *et al.*, 1999; Li *et al.*, 1999; Sim e Koi, 2002; Brown *et al.*, 2003; Keng Kau *et al.*, 2003; Swinyard e Smith, 2003; Black, 2007; Lennon *et al.*, 2007; Soopramanien e Robertson, 2007; Weltevreden, 2007; Brashear *et al.*, 2009, Maan e Kumar, 2020). No *marketing*, os fatores que influenciam o comércio eletrônico são analisados para investigar o comportamento do consumidor e detectar o público alvo para campanhas de publicidade (Keng Kau *et al.*, 2003). Além disso, pouca ou nenhuma consideração existe sobre os aspectos espaciais (Frag *et al.*, 2006b), principalmente relacionados à entrega.

As áreas de transportes e de geografia possuem um enfoque semelhante na análise da influência de fatores socioeconômicos no comércio eletrônico, abordando a temática do planejamento urbano e da problemática relacionada às entregas da última milha (Beckers *et al.*, 2018). Porém, mesmo nessas áreas ainda há poucos estudos sobre fatores espaciais e sua relação com o *e-commerce*, como o acesso a oportunidades de compra no varejo tradicional e a distribuição geográfica dos vários perfis demográficos de e-consumidores (Frag *et al.*, 2003; Frag *et al.*, 2005; Krizek *et al.*, 2005; Frag *et al.*, 2006a; Frag *et al.*, 2006b; Frag *et al.*, 2007; Ren e Kwan, 2009; Cao *et al.*, 2013; Clarke *et al.*, 2015; Loo e Wang, 2017; Beckers *et al.*, 2018; Kirby-Hawkins *et al.*, 2018; Hood *et al.*, 2020; Song, 2021). No entanto, sabe-se que a homogeneização da distribuição do comércio eletrônico numa área de estudo, bem como a adoção de um baixo nível de desagregação de dados pode influenciar negativamente a análise (Lee *et al.*, 2017; Beckers *et al.*, 2018).

Além disso, há escassez de estudos que avaliem dados reais sobre o comércio eletrônico, como dados de entregas ou de compras efetivamente realizadas (Ren e Kwan, 2009; Kirby-Hawkins *et al.*, 2018; Cheng *et al.*, 2021; Song, 2021). Em sua maioria, os estudos utilizaram dados autodeclarados provenientes de questionários respondidos pelo potencial consumidor. Esses dados são, portanto, sujeitos a vieses inerentes ao comportamento humano, como a memória e

expectativas sociais, bem como da interpretação do questionário (Bellman *et al.*, 1999). O uso de dados reais de entregas domiciliares permitiria uma melhor compreensão do fenômeno estudado, o que possibilitaria a adoção de melhores soluções para as externalidades associadas às entregas da última milha. Além disso, há ainda uma carência de estudos voltados à realidade de países em desenvolvimento (Loo e Wang, 2017; Song, 2021), especialmente a América Latina.

Considerando o contexto apresentado e as oportunidades de pesquisa detalhadas na Tabela 1, torna-se relevante o estudo da influência dos fatores socioeconômicos dos bairros e do acesso às oportunidades de compra no varejo tradicional no número de entregas do comércio eletrônico. Portanto, a pergunta de pesquisa que se pretende responder por meio deste estudo é: Quais fatores socioeconômicos influenciam na quantidade de entregas de produtos do comércio eletrônico em uma área urbana? Como esses fatores e o acesso a oportunidades de compra no varejo tradicional influenciam na quantidade de entregas de produtos do comércio eletrônico em uma área urbana?

Tabela 1: Oportunidades de pesquisa

<b>Categoria</b>	<b>Oportunidade de pesquisa</b>	<b>Referência</b>
Fenômeno	Uso de dados reais	Ren e Kwan (2009) Kirby-Hawkins <i>et al.</i> (2018) Cheng <i>et al.</i> (2021) Song (2021)
	Análise no contexto de países em desenvolvimento	Loo e Wang (2017) Song (2021)
Metodológico	Incorporação da análise espacial nas análises	Song (2021)

Neste trabalho, foram analisados dados de entregas do comércio eletrônico realizadas por uma transportadora. Portanto, neste estudo não foram incluídos dados de entregas instantâneas, como aquelas provenientes de restaurantes, lanchonetes ou supermercados.

De acordo com a literatura, os fatores socioeconômicos mais recorrentes para o comércio eletrônico são o gênero, a idade e renda do consumidor (por exemplo Farag *et al.*, 2006a; Clarke *et al.*, 2015; Beckers *et al.*, 2016; Hood *et al.*, 2020; Jaller e Pahwa, 2020). Alguns estudos também identificaram como fatores relevantes a cor ou raça (Ren e Kwan, 2009; Saphores e Xu, 2020) e o número de habitantes por domicílio (Sener e Reeder, 2012; Zhou e Wang, 2014). Além disso, o acesso a oportunidades de compra no varejo tradicional também se mostrou

relevante para o comércio eletrônico (por exemplo Farag *et al.*, 2006b; Clarke *et al.*, 2015; Kirby-Hawkins *et al.*, 2018; Hood *et al.*, 2020).

Diante disso, foi estabelecido como hipótese de pesquisa que os fatores gênero, idade, renda, cor ou raça, habitantes por domicílio e oportunidade de acesso ao varejo tradicional influenciam na quantidade de entregas do comércio eletrônico. Destaca-se que a oportunidade de acesso ao varejo tradicional, neste trabalho, é expressa pelo número de estabelecimentos comerciais varejistas nas regiões de estudo. Ainda, assim como para os dados de entrega, desconsiderou-se restaurantes, lanchonetes e supermercados.

Assim, o objetivo desta dissertação foi analisar a relação entre fatores socioeconômicos e oportunidade de acesso ao varejo tradicional com o número de entregas do comércio eletrônico. Para tanto, os objetivos específicos são:

- Identificar a relação entre as entregas do comércio eletrônico e os fatores socioeconômicos e oportunidade de acesso ao varejo tradicional;
- Analisar a correlação espacial entre as entregas do comércio eletrônico e os fatores socioeconômicos e oportunidade acesso ao varejo tradicional;
- Comparar a influência dos fatores considerados na análise para três cidades: Belo Horizonte, Uberlândia e Patos de Minas.

A análise foi desenvolvida a nível de bairro, pois sabe-se que dados demográficos podem variar dentro de uma área de estudo das dimensões de uma cidade, o que faz com que a relação entre as variáveis analisadas possa se alterar dependendo do local da cidade em que a análise é conduzida. Assim, uma análise com agregação dos dados ao nível de cidades poderia levar a resultados que não fossem de fato representativos das condições reais (Lee *et al.*, 2017; Beckers *et al.*, 2018). As cidades escolhidas para análise foram selecionadas considerando seus diferentes portes e também a disponibilidade de dados de entrega junto à transportadora.

Este trabalho está estruturado da seguinte forma: Após esta introdução, no capítulo 2 é apresentada a revisão da literatura, onde se deixa evidente a contribuição deste estudo. No capítulo 3 são descritas as áreas de estudo e os dados. No capítulo 4 é detalhado o procedimento metodológico. Os resultados são apresentados no capítulo 5 e discutidos no capítulo 6. A conclusão e sugestão para trabalhos futuros são apresentados no capítulo 7.

## 2 REVISÃO DA LITERATURA

A primeira compra pelo comércio eletrônico foi realizada em 1994 nos Estados Unidos (Lewis, 1994). Desde então, são vários os fatores que contribuem para o crescimento deste canal de varejo. Os primeiros estudos que avaliaram a influência de fatores socioeconômicos nas compras *online* eram da área do *marketing*, desenvolvidos por Bellman *et al.* (1999) e Li *et al.* (1999), tratando da relação do entre o *e-commerce* e o comportamento/percepção do consumidor.

Considerando fatores relacionados a transportes, Farag *et al.* (2003) analisaram a relação entre o *e-commerce* e a realização de viagens de compra pelos consumidores. Esse estudo também foi um dos precursores na análise da influência do acesso ao varejo tradicional na realização de compras pela *Internet*.

A partir destes estudos, outros foram realizados em diversas áreas do conhecimento, como economia e gestão (Vrechopoulos *et al.*, 2001; Stafford *et al.*, 2004; Burkolter e Kluge, 2011; Pérez-Hernández e Sánchez-Mangas, 2011; Seitz *et al.*, 2017; Van Droogenbroeck e Van Hove, 2017; Lubis, 2018; Morganosky e Cude, 2020; Valarezo *et al.*, 2020), comunicação (Li *et al.*, 1999; Fernández-Bonilla, 2020; Opesade, 2020; Pérez-Amaral, 2020), ciências sociais (Hashin *et al.*, 2009) e geografia (De Blasio, 2008; Ren e Kwan, 2009; Clarke *et al.*, 2015; Kirby-Hawkins; Hood *et al.*, 2020; Song, 2021), além das pesquisas na área de *marketing* (Sim e Koi, 2002; Brown *et al.*, 2003; Keng Kau *et al.*, 2003; Swinyard e Smith, 2003; Black, 2007; Lennon *et al.*, 2007; Soopramanien e Robertson, 2007; Weltevreden, 2007; Brashear *et al.*, 2009, Maan e Kumar, 2020) e transportes (Farag *et al.*, 2005; Krizek *et al.*, 2005; Farag *et al.*, 2006a; Farag *et al.*, 2006b; Farag *et al.*, 2007; Mokhtarian *et al.*, 2009; Cao, 2012; Sener e Reeder, 2012; Cao *et al.*, 2013; Zhou e Wang, 2014; Lee *et al.*, 2015; Lee *et al.*, 2017; Loo e Wang, 2017; Motte-Baumvol *et al.*, 2017; Becker *et al.*, 2018; Jaller e Pahwa, 2020, Saphores e Xu, 2020; Cheng *et al.*, 2021). As seções abaixo detalham esses estudos.

### 2.1 Influência dos fatores socioeconômicos no comércio eletrônico nas áreas de marketing, economia e gestão, comunicação e ciências sociais

Buscando entender o comportamento do e-consumidor, Bellman *et al.* (1999) desenvolveram um estudo para identificar características pessoais deste tipo de consumidor que poderiam influenciar o comércio eletrônico. Foram usados dados de um questionário por uma empresa privada que obteve 10.180 respostas válidas de participantes de 82 países, cuja pesquisa foi

realizada em 1998. As perguntas abrangeram comportamentos e atitudes do consumidor em relação à realização de compras online, como experiência com a *Internet* e percepção de riscos associados ao comércio eletrônico. Embora o foco do estudo fosse em variáveis comportamentais, também foram avaliadas variáveis socioeconômicas, como idade, grau de escolaridade e renda. Por meio de regressão logística, foi estimado um modelo para prever se uma pessoa compraria ou não *online*. O modelo foi capaz de realizar previsões corretas para 66% dos respondentes. Os autores concluíram que embora as variáveis socioeconômicas analisadas influenciassem a probabilidade de uma pessoa realizar compras online, essa influência era pequena quando comparada com a influência dos outros tipos de variáveis analisadas.

Li *et al.* (1999) estimaram um modelo para verificar a forma que certos fatores impactam na maneira que o consumidor utiliza a *Internet* como ferramenta de compra. Para tanto, foi testada a influência de fatores socioeconômicos (gênero, renda, idade e grau de escolaridade), fatores relacionados à percepção de canais de *marketing* (canais de comunicação, de transação e de distribuição) e fatores relacionados a orientações de compra (predisposições gerais do indivíduo em relação ao ato de comprar) no comportamento do consumidor em relação ao *e-commerce*. Os dados foram obtidos por meio de um questionário realizado por uma empresa privada que obteve 999 respostas válidas de usuários nos Estados Unidos. A análise de variância (ANOVA) foi usada para testar a significância de cada variável separadamente, e demonstrou que, dentre os fatores socioeconômicos, gênero, grau de escolaridade e renda têm influência no comércio eletrônico, sendo homens, pessoas de maior grau de escolaridade e pessoas de maior renda, mais propensas a comprar *online*. Além disso, o estudo também utilizou um modelo de regressão linear múltipla para verificar quais variáveis são preditoras do ato de comprar *online*, encontrando como resultado que, dentre as variáveis socioeconômicas, o grau de escolaridade se mostrou significativo.

Com o objetivo de entender a demanda dos consumidores por serviços de compra de produtos de mercearia *online*, Morganosky e Cude (2000) analisaram dados comportamentais e socioeconômicos de um grupo de 243 consumidores dos Estados Unidos. Os dados utilizados foram obtidos por meio de uma pesquisa com pessoas que costumavam comprar produtos de mercearia *online* em 1998 de uma grande loja do ramo. Os participantes responderam perguntas relacionadas ao tempo que gastavam realizando as compras na *Internet*, se usavam o serviço de entrega domiciliar ou retirada na loja e se também compravam produtos de mercearia em loja física, além de dados socioeconômicos (idade, grau de escolaridade, gênero, renda e número de

crianças no domicílio). Por meio de teste qui-quadrado, verificou-se que o grau de escolaridade era a principal variável socioeconômica associada a realização da compra de todos os produtos de mercearia *online*, e idade e renda eram as principais variáveis socioeconômicas associadas às razões primárias de se comprar *online*.

Com o objetivo de traçar o perfil de consumidores do comércio eletrônico e de consumidores que pensam em adotar o *e-commerce* e identificar as principais diferenças entre eles, Vrechopoulos *et al.* (2001) analisaram as características socioeconômicas dos dois grupos. Os dados do estudo foram obtidos por meio de questionários aplicados a clientes de uma grande rede de supermercados em Atenas, Grécia. Além de interrogados sobre características comportamentais, percepções e preferências, os consumidores detalharam suas características socioeconômicas (gênero, idade, estado civil, grau de escolaridade e renda). Em seguida, os dados foram analisados por meio de testes de significância estatística, e assim, cada grupo foi caracterizado de acordo com as características socioeconômicas que os representam. Verificou-se que o grupo dos adeptos ao *e-commerce* era constituído principalmente de homens, pessoas jovens (25 a 44 anos), pessoas de maior grau de escolaridade, solteiros e pessoas de maior renda.

Sim e Koi (2002) analisaram as características demográficas e psicológicas dos consumidores do *e-commerce* e do varejo tradicional em Singapura. Os dados foram obtidos por meio de um questionário realizado com consumidores. Quanto aos dados socioeconômicos, foi analisada a influência de gênero, idade, renda, grau de escolaridade e estado civil. As análises foram realizadas usando as seguintes técnicas estatísticas: correlação de Pearson, análise fatorial, teste t, tabulação cruzada e qui-quadrado. Os resultados indicaram que homens tem mais chances de comprar *online* que mulheres, compradores do comércio eletrônico são majoritariamente solteiros, e que usuários do *e-commerce* são pessoas mais jovens, pessoas de maior renda e pessoas de maior grau de escolaridade do que os compradores do varejo tradicional.

Brown *et al.* (2003) estudaram a predisposição do consumidor em comprar *online* por meio dos fatores que influenciam a intenção de compra. Dentre os fatores analisados incluem-se experiência anterior com *e-commerce*, tipo de produto e motivos que orientam a compra (como conveniência, preço ou diversão, por exemplo). Foi investigado ainda o fator gênero e seu efeito na propensão do indivíduo de comprar *online*. Os dados analisados foram obtidos em 1999 por meio de um questionário feito por uma empresa privada da área de *marketing* que alcançou 437 respostas de usuários da *Internet* nos Estados Unidos. Por meio de uma análise de variância,

verificou-se que o gênero tem efeito na intenção de realização de compras pela *Internet*, sendo os homens mais propensos a comprar *online* que as mulheres.

Com o objetivo de examinar o comportamento do e-consumidor, Keng Kau *et al.* (2003) conduziram um estudo com dados obtidos por meio de um questionário aplicado por uma empresa privada realizado com usuários da *Internet* em 2000. Foram obtidas 3.712 respostas válidas sobre os hábitos de compra (preferências e frequência de compra) de consumidores majoritariamente de Singapura, além de dados socioeconômicos (gênero, estado civil, idade e grau de escolaridade). Por meio de uma análise de *cluster*, separou-se os consumidores em seis grupos baseados em suas atitudes em relação ao *e-commerce* (*on-off shopper, comparison shopper, traditional shopper, dual shopper, e-Laggard, information surfer*) e, com base em testes qui-quadrado, pôde-se verificar as características socioeconômicas de cada grupo.

Swinyard e Smith (2003) analisaram as características e os motivos de grupos de pessoas que compravam *online* e grupos de pessoas que não o faziam. Para tanto, foram obtidos dados de 1.738 domicílios dos Estados Unidos por meio de um questionário a respeito da adesão ou não ao *e-commerce* e das atitudes e concepções acerca dessa modalidade de compras, bem como características socioeconômicas (idade, renda, grau de escolaridade e gênero). Assim, com relação às características socioeconômicas, foi usada análise de *cluster* para a definição de quais influenciavam o grupo de compradores do comércio eletrônico. Concluiu-se que esse grupo é mais jovem (idade média na faixa dos 40 anos), tem maior renda e maior grau de escolaridade que o grupo dos que não realizam compras *online*. Portanto, as variáveis significativas encontradas são a idade, a renda e o grau de escolaridade.

Stafford *et al.* (2004) analisaram a influência do fator gênero no comércio eletrônico em três países: Estados Unidos, Finlândia e Turquia a fim de verificar o impacto das diferenças culturais nos resultados. Dados sobre uso do comércio eletrônico e frequência de realização de compras *online* foram obtidos por meio de questionários que obtiveram 217 respostas válidas (52 na Finlândia, 63 nos Estados Unidos e 102 na Turquia). Foi utilizada análise de variância (ANOVA) para testar a hipótese de que os homens em cada nação estudada são mais propensos a fazer uso do *e-commerce*. Os resultados confirmaram a hipótese, evidenciando a influência do fator gênero no comércio *online*.

Com o objetivo de entender o comportamento do e-consumidor, Black (2007) desenvolveu um estudo para identificar quais aspectos demográficos e geográficos podem ser utilizados para



prever o número de vendas e o valor das vendas de varejo no *site* de leilões virtuais *eBay*. Foram obtidos dados de 753 transações realizadas por consumidores americanos com dois varejistas do *site*. Os dados incluíam o gênero do e-consumidor (variável demográfica) e seu local de residência – urbano ou rural e região dos Estados Unidos (variáveis geográficas). O artigo define população rural conforme a classificação usada no censo demográfico dos Estados Unidos, que é baseada na densidade populacional da vizinhança. Por meio de testes de significância estatística com as variáveis, o estudo determinou que as variáveis analisadas exercem influência no comércio eletrônico: pessoas que vivem em áreas rurais têm maiores chances de comprar produtos pelo *eBay* que pessoas de áreas urbanas; mulheres são mais propensas a realizar compras no *site*; e pessoas de diferentes regiões americanas têm comportamentos diferentes em relação ao comércio eletrônico no *eBay*.

Lennon *et al.* (2007) analisaram o uso do *e-commerce* por habitantes de regiões rurais dos Estados Unidos. Foram obtidos dados de habitantes de áreas rurais de 11 estados dos Estados Unidos por meio de pesquisas realizadas em 2000 (2.198 respostas válidas) e 2003 (979 respostas válidas). Os questionários, focados principalmente no comércio de itens de vestuário, interrogavam os indivíduos a respeito de características socioeconômicas (grau de escolaridade, renda, gênero, idade, número de habitantes no domicílio e número de pessoas com deficiência no domicílio), experiência prévia com comércio eletrônico em geral e percepções em relação à compra de vestuário e pela *Internet*. Foi utilizada modelagem de equações estruturais para verificar quais variáveis influenciavam o comércio eletrônico, sendo verificado que, dentre os fatores socioeconômicos, grau de escolaridade e renda apresentavam influência.

Soopramanien e Robertson (2007) analisaram a influência de fatores socioeconômicos, atitudes e crenças do consumidor na decisão de comprar *online* e no uso do comércio eletrônico. Os dados usados no estudo foram obtidos por meio de questionários respondidos por 894 indivíduos de diversas regiões do Reino Unido em 2005. As variáveis socioeconômicas consideradas pela análise foram gênero, idade e renda. A análise, feita por meio de regressão logística, identificou que consumidores com maior renda têm maior chance de comprar *online*, portanto, o fator renda tem influência no comércio eletrônico.

Em seu estudo do impacto do comércio eletrônico nas compras realizadas por meio do varejo tradicional nos centros de cidades holandesas, Weltevreden (2007) analisou brevemente os fatores socioeconômicos que influenciam o *e-commerce*. Os dados do estudo foram obtidos por meio de uma empresa privada que realizou uma pesquisa em 2004 com 3.218 usuários da

*Internet*. Os dados socioeconômicos (gênero, idade e grau de escolaridade) foram analisados por meio de análise de variância (ANOVA), concluindo-se que homens, pessoas entre 25 e 44 anos e pessoas de maior grau de escolaridade têm maiores chances de comprar pela *Internet*.

Brashear *et al.* (2009) realizaram um estudo comparativo das características dos usuários do comércio eletrônico de seis países (Estados Unidos, Inglaterra, Nova Zelândia, China, Brasil e Bulgária) com o objetivo de traçar um perfil do e-consumidor em diferentes regiões. Os fatores investigados se relacionam com atitudes e motivações relacionadas à realização de compras pela *Internet*, além de fatores socioeconômicos (idade, renda, grau de escolaridade e gênero). Os dados para o estudo foram obtidos por meio de um questionário elaborado pelos autores e analisados separadamente para cada país pela análise de variância (ANOVA). Em relação às variáveis socioeconômicas, o estudo concluiu que a renda é um fator de influência no comércio eletrônico em todos os seis países; a idade é um fator de influência no Brasil e nos Estados Unidos; o grau de escolaridade é um fator de influência na Bulgária e na China e o gênero é um fator de influência na Nova Zelândia.

Hashin *et al.* (2009) avaliaram a influência de fatores socioeconômicos no comportamento do consumidor do comércio eletrônico na Malásia. Os dados foram coletados por meio de um questionário administrado pelos autores com alunos de faculdade de meio período do curso de contabilidade; esse grupo foi escolhido por fazer parte do mercado de trabalho e possuir cartão de crédito, critérios julgados pelo estudo como fundamentais para a realização de compras *online*. Foram respondidas perguntas sobre características socioeconômicas (gênero, idade, renda, estado civil e cargo profissional ocupado) e hábitos de compra *online* (como frequência de compra e de acesso a *sites* de compra). Os resultados foram obtidos por meio de análise de variância (ANOVA) e indicaram que os fatores analisados têm influência no comércio eletrônico. Homens, pessoas de cargos profissionais mais elevados, pessoas de maior renda, pessoas divorciadas e pessoas com idades entre 30 a 39 compravam mais *online* que outros grupos.

Burkolter e Kluge (2011) estudaram o comportamento do e-consumidor na Alemanha e sua relação com fatores socioeconômicos e motivacionais. Os dados foram obtidos por meio de um questionário *online* que resultou em 405 respostas válidas a respeito da frequência que o respondente compra pela *Internet* e seus motivos (como, por exemplo, conveniência, experiência, necessidade de emoção e desejo de estar na moda), bem como suas características individuais (gênero, idade, grau de escolaridade e situação empregatícia). Para verificar a

influência dos fatores investigados nas compras do *e-commerce* foi utilizada a correlação de Spearman, que evidenciou que os fatores gênero, situação empregatícia e grau de escolaridade influenciam nas compras *online*, sendo homens, pessoas empregadas e pessoas com maior grau de escolaridade compradoras *online* mais frequentes. Além disso, o estudo utilizou modelos de regressão linear múltipla para prever a decisão de aderir ao *e-shopping*, constatando que os fatores idade, grau de escolaridade e situação empregatícia são variáveis socioeconômicas com significância estatística no modelo.

Pérez-Hernández e Sánchez-Mangas (2011) analisaram a decisão de comprar *online* de consumidores espanhóis a partir de sua relação com fatores socioeconômicos (idade, gênero, grau de escolaridade, situação empregatícia, número de habitantes no domicílio e número de crianças no domicílio), uso da *Internet* (possuir ou não *Internet* em casa e frequência de uso da *Internet*) e fatores de localização (se o indivíduo vive em região urbana – acima de 100.000 habitantes, ou em região de convergência – regiões que têm baixa performance econômica e recebem auxílio da União Europeia). Os dados foram retirados de uma pesquisa sobre o uso de tecnologias de informação e comunicação realizada pelo Instituto de Estatística da Espanha entre os anos de 2004 a 2009. Por meio do uso de modelos *probit*, os autores concluíram que todos os fatores socioeconômicos considerados têm influência na decisão de compra pelo comércio eletrônico, sendo o grau de escolaridade o mais influente: homens, pessoas jovens, pessoas com maior grau de escolaridade, pessoas empregadas e pessoas que moram sozinhas têm maiores chances de comprar *online*, enquanto um maior número de crianças no domicílio diminui as chances de adesão ao comércio eletrônico.

Para entender o comércio eletrônico de produtos de mercearia na Alemanha, Seitz *et al.* (2017) analisaram os fatores que motivam os consumidores a aderirem à compra *online* desses produtos. Assim, foram analisados dados relacionados à percepção e comportamento dos clientes em relação ao *e-commerce*, obtidos por meio de questionários aplicados em regiões urbanas, suburbanas e rurais. Os respondentes foram também perguntados sobre suas características socioeconômicas (gênero, idade, renda e número de filhos no domicílio). Utilizando modelagem de equações estruturais, foram detectados cinco perfis de consumidor de produtos de mercearia *online*, cada um associado a diferentes motivações e necessidades. Em seguida, verificou-se os fatores socioeconômicos que estavam associados a cada perfil, demonstrando a influência de tais fatores no comportamento em relação ao *e-commerce*. Destaca-se que o questionário usado nesse estudo não perguntava aos consumidores se eles de fato compravam *online*, mas sim se eles se imaginavam comprando *online*.

Van Droogenbroeck e Van Hove (2017) analisaram como características socioeconômicas influenciam na adesão do consumidor ao *e-commerce* de produtos de mercearia. Os dados foram obtidos por meio de um questionário realizado com 468 consumidores de uma rede de supermercados na Bélgica. A análise utilizou modelos de regressão logística para avaliar as variáveis gênero, grau de escolaridade, idade, número de habitantes no domicílio, presença de crianças no domicílio e situação empregatícia. O estudo concluiu que o grau de escolaridade, a idade, a situação empregatícia e a presença de crianças têm influência no comércio eletrônico, sendo pessoas de maior grau de escolaridade, pessoas menores de 61 anos, pessoas empregadas e presença de crianças no domicílio aumentam as chances de se comprar *online*. Destaca-se também que o efeito da idade diminui quando analisado juntamente com a situação empregatícia e a presença de crianças no domicílio devido a correlação entre a variável idade e essas variáveis.

Com o objetivo de estudar os fatores socioeconômicos que influenciam as decisões de compras realizadas por meio do comércio eletrônico, Lubis (2018) analisou dados de uma pesquisa feita com frequentadores de grandes cafeterias na cidade de Medan, Indonésia. O questionário obteve 200 respostas válidas a respeito de fatores socioeconômicos (gênero, renda, grau de escolaridade, idade, morar sozinho ou com a família), preferência do consumidor (vantagens, desvantagens, riscos e benefícios de se comprar *online* e no varejo tradicional) e uso de aplicativos de compra *online*. A estimação de modelos de regressão logística indicaram que as decisões de compra pelo *e-commerce* são influenciadas por gênero, renda e uso de aplicativos de compra, sendo que homens, pessoas de maior renda e usuários de aplicativos apresentam maior tendência a comprar pela *Internet*.

Fernández-Bonilla (2020) verificou a propensão dos indivíduos de realizar compras *online* por meio do estudo de variáveis socioeconômicas e comportamentais (uso da *Internet* e percepção de risco associado ao *e-commerce*) na Espanha. Os dados foram obtidos por meio de pesquisas realizadas em 2014 e 2019 pelo Instituto Nacional de Estatística Espanhol, que obteve 17.195 respondentes válidos. As variáveis socioeconômicas avaliadas foram gênero, idade, grau de escolaridade e renda. Por meio de regressão logística, constatou-se que o gênero, a renda e o grau de escolaridade têm relação com a adoção do *e-commerce*, sendo que homens, pessoas de maior renda e de maior escolaridade compram mais pela *Internet*.

Maan e Kumar (2020) estudaram a relação entre fatores socioeconômicos e a frequência de realização de compras *online* por consumidores indianos. Os dados foram obtidos por meio de

um questionário realizado em cidades da Índia, alcançando 439 respostas válidas a respeito da frequência de compras *online* realizadas pelos respondentes e suas características socioeconômicas (idade, grau de escolaridade e profissão). A análise fez uso de tabulação cruzada e do teste qui-quadrado, verificando que o fator idade estava associado ao uso do *e-commerce*, sendo os mais jovens (entre 18 e 35 anos) compradores mais frequentes.

Opesade (2020) investigou a relação entre fatores socioeconômicos e o interesse, adoção e difusão do comércio eletrônico na Nigéria. Para tanto, foram obtidos dados do número de buscas *online* para acesso aos dois maiores *sites* de *e-commerce* do país por meio do *Google Trends*, um serviço que oferece informações sobre o que usuários estão pesquisando nas ferramentas de busca da *Google*. Esses dados foram coletados no período entre 2014 e 2017, juntamente com informações socioeconômicas do censo nigeriano sobre a população de todos os estados. As variáveis socioeconômicas usadas no estudo se traduzem em informações relacionadas a renda, pobreza absoluta, grau de escolaridade, situação empregatícia, acesso à *Internet*, população e conhecimento sobre uso de computadores. Os dados foram analisados por meio de distribuição de frequências e medidas de relação entre variáveis, resultando na conclusão de que os fatores renda, pobreza absoluta, população, acesso à *Internet* e conhecimento sobre o uso de computadores têm correlação com o volume de buscas pelos dois *sites* de *e-commerce* analisados.

Pérez-Amaral *et al.* (2020) estudaram a desigualdade de uso do *e-commerce* na Espanha analisando dados socioeconômicos, determinando assim quais fatores exerciam influência na opção pelo comércio eletrônico pelos consumidores. Os dados foram adquiridos de uma pesquisa sobre o uso de tecnologias de informação e comunicação realizada pelo Instituto de Estatística da Espanha com 133.420 consumidores entre 2008 e 2016. Por meio de regressão logística com dados em painel, foi possível determinar o impacto das variáveis socioeconômicas (idade, grau de escolaridade, gênero, renda, situação empregatícia e número de habitantes no domicílio) na adoção do *e-commerce* em regiões da Espanha. Assim, concluiu-se que todas as variáveis socioeconômicas analisadas tinham influência no *e-commerce*: homens, pessoas jovens (16 a 24 anos), pessoas com maior grau de escolaridade, pessoas com menos habitantes no domicílio, pessoas empregadas e pessoas com maior renda têm mais chance de fazer uso do *e-commerce*. Além disso, o estudo investigou o impacto do tamanho da população de uma região no *e-commerce*, concluindo que habitantes de regiões de tamanho médio (entre 100.000 e 500.000 habitantes) são menos propensos a comprar *online* que o resto da população, enquanto outras regiões não apresentaram resultados significativos. Os autores

destacam ainda que teria sido importante para o estudo ter dados desagregados para regiões com menos de 20.000 habitantes, o que teria tornado possível analisar as diferenças entre as populações rurais e urbanas.

Valarezo *et al.* (2020) analisaram os fatores socioeconômicos que influenciam a adoção do comércio eletrônico pelos consumidores. Para tanto, foram usados dados de 174.776 respondentes obtidos de uma pesquisa do Instituto Nacional de Estatística da Espanha e realizada entre 2008 e 2017. Os dados socioeconômicos considerados foram gênero, idade, grau de escolaridade, ocupação, situação empregatícia, renda e estado civil. A análise foi feita por modelos de probabilidade linear, modelos de regressão logística e pelo modelo de seleção de Heckman. Concluiu-se que os fatores gênero, idade, situação empregatícia e renda estão associados à probabilidade de se comprar *online*, sendo homens, pessoas jovens (menores de 25), pessoas empregadas e pessoas de maior renda mais propensos ao uso do *e-commerce*.

Os trabalhos apresentados nesta subseção mostram uma forte tendência de analisar as características que levam o consumidor a comprar *online* tendo como motivação o entendimento do mercado e a identificação de grupos aos quais podem ser direcionadas estratégias de *marketing* (Valarezo *et al.*, 2020). Os questionários e modelos aplicados nestes estudos foram adaptados à esta realidade, e, em sua maioria, enfatizam características comportamentais que determinam as decisões de uso do *e-commerce* (Valarezo *et al.*, 2020). Ainda assim, estes estudos tratam da influência de fatores socioeconômicos no comércio eletrônico, mesmo que alguns o façam apenas brevemente (Bellman *et al.*, 1999; Weltevreden, 2007) ou foquem em apenas um fator específico (Brown *et al.*, 2003; Stafford *et al.*, 2004; Black, 2007).

Além disso, apenas dois estudos analisam a influência de variáveis geográficas ou de localização: Black (2007) inclui em sua investigação a influência da localização urbana/rural e regional no comércio eletrônico, e Pérez-Hernández e Sánchez-Mangas (2011) consideraram a influência regional em sua pesquisa.

Percebe-se que a regressão logística foi um método fortemente utilizado pelos estudos (Bellman *et al.*, 1999; Soopramanien e Robertson, 2007; Van Droogenbroeck e Van Hove, 2017; Lubis, 2018; Fernández-Bonilla, 2020; Pérez-Amaral *et al.*, 2020; Valarezo *et al.*, 2020). Além deste, destaca-se a recorrência do uso da análise de variância (Li *et al.*, 1999; Brown *et al.*, 2003; Stafford *et al.*, 2004; Weltevreden, 2007; Brashear *et al.*, 2009; Hashin *et al.*, 2009) e do teste

qui-quadrado (Morganosky e Cude, 2000; Sim e Koi, 2002; Keng Kau *et al.*, 2003; Maan e Kumar, 2020). A Tabela 2 resume os métodos utilizados nos estudos descritos nesta subseção.

Tabela 2: Métodos utilizados nos estudos descritos

Autor	Regressão logística	ANOVA	Qui-quadrado	Testes estatísticos	SEM	Análise de <i>cluster</i>	Regressão linear	Modelo <i>probit</i>	Distribuição de frequências
Bellman <i>et al.</i> (1999)	•								
Li <i>et al.</i> (1999)		•							
Morganosky e Cude (2000)			•						
Vrechopoulos <i>et al.</i> (2001)				•					
Sim e Koi (2002)			•						
Brown <i>et al.</i> (2003)		•							
Keng Kau <i>et al.</i> (2003)			•						
Swinyard e Smith (2003)						•			
Stafford <i>et al.</i> (2004)		•							
Black (2007)				•					
Lennon <i>et al.</i> (2007)					•				
Soopramanien e Robertson (2007)	•								
Weltevreden (2007)		•							
Brashear <i>et al.</i> (2009)		•							
Hashin <i>et al.</i> (2009)		•							
Burkolter e Kluge (2011)							•		
Pérez-Hernández e Sánchez-Mangas (2011)								•	
Seitz <i>et al.</i> (2017)					•				
Van Droogenbroeck e Van Hove (2017)	•								
Lubis (2018)	•								
Fernández-Bonilla (2020)	•								
Maan e Kumar (2020)			•						
Opesade (2020)									•
Pérez-Amaral <i>et al.</i> (2020)	•								
Valarezo <i>et al.</i> (2020)	•								
<b>Quantidade de trabalhos</b>	<b>7</b>	<b>6</b>	<b>4</b>	<b>2</b>	<b>2</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>

Nota-se também, que os fatores gênero, idade, grau de escolaridade e renda são os mais recorrentes na literatura como influentes no comércio eletrônico, tendo sido apontados em diversos estudos. Destaca-se ainda, que a maioria dos estudos utilizou dados autodeclarados, provindos de questionários, com exceção de Black (2007), que usou de dados reais de compras realizadas em um *site* e Opesade (2020), que empregou dados de número de buscas por *sites* de compras. Ainda assim, esse último não trabalhou com dados reais de compra, uma vez que não há como garantir que a compra tenha sido efetuada apenas pela informação sobre a busca do produto. Sabe-se que dados autodeclarados estão sujeitos a uma série de vieses, como falhas de

memória do respondente, particularidades da escala utilizada no questionário e alteração proposital de dados para adequação em expectativas sociais (Bellman *et al.*, 1999).

A Tabela 3 apresenta uma síntese dos fatores analisados e dos fatores influentes, tipo de local de estudo e tipo de dado analisado pelos estudos descritos nesta subseção.

Tabela 3: Síntese dos principais aspectos dos estudos discutidos

Referência	Fatores							Local		Dado	
	Gênero	Idade	Grau de escolaridade	Renda	Habitantes/domicílio	Cor ou raça	Acesso ao varejo tradicional	Países desenvolvidos	Países em desenvolvimento	Dado real	Dado autodeclarado
Bellman <i>et al.</i> (1999)		•	•	•				•	•		•
Li <i>et al.</i> (1999)	*	•	*	*				•			•
Morganosky e Cude (2000)	•	*	*	*	•			•			•
Vrechopoulos <i>et al.</i> (2001)	*	*	*	*				•			•
Sim e Koi (2002)	*	*	*	*				•			•
Brown <i>et al.</i> (2003)	*							•			•
Keng Kau <i>et al.</i> (2003)	*	*	*					•			•
Swinyard e Smith (2003)	•	*	*	*				•			•
Stafford <i>et al.</i> (2004)	*							•	•		•
Black (2007)	*							•		•	
Lennon <i>et al.</i> (2007)	•	•	*	*	•			•			•
Soopramanien e Robertson (2007)	•	•		*				•			•
Weltevreden (2007)	*	*	*					•			•
Brashear <i>et al.</i> (2009)	*	*	*	*				•	•		•
Hashin <i>et al.</i> (2009)	*	*		*					•		•
Burkolter e Kluge (2011)	*	*	*					•			•
Pérez-Hernández e Sánchez-Mangas (2011)	*	*	*		*			•			•
Seitz <i>et al.</i> (2017)	*	*		*	*			•			•
Van Droogenbroeck e Van Hove (2017)	•	*	*		*			•			•
Lubis (2018)	*	•	•	*					•		•
Fernández-Bonilla (2020)	*	•	*	*				•			•
Maan e Kumar (2020)		*	•						•		•
Opesade (2020)			•	*					•	•	
Pérez-Amaral <i>et al.</i> (2020)	*	*	*	*	*			•			•
Valarezo <i>et al.</i> (2020)	*	*	•	*				•			•

Na coluna Fatores: variável analisada “•”; variável analisada e influente “\*”



## **2.2 Influência dos fatores socioeconômicos e acesso ao varejo tradicional no comércio eletrônico em pesquisas nas áreas de geografia e transportes**

Os estudos sobre os fatores que influenciam o comércio eletrônico nas áreas do transporte e da geografia têm motivações distintas dos estudos apresentados na seção anterior. As análises nas áreas de transportes e geografia, em geral, são motivadas pela necessidade de entender o perfil do e-consumidor para que se possa melhorar os problemas associados às entregas da última milha, tornando-as mais sustentáveis; além de prover ferramentas para auxiliar as autoridades no planejamento urbano (Beckers *et al.*, 2018).

Em seu estudo sobre o impacto do *e-commerce* no comportamento do consumidor em relação às viagens para compras no varejo tradicional, Farag *et al.* (2003) analisaram o impacto de variáveis socioeconômicas no comércio eletrônico em cidades da Holanda, além de variáveis relacionadas ao acesso à *Internet* e experiência com *e-shopping*. Os dados foram obtidos por meio de uma pesquisa com consumidores realizada por uma empresa privada em 2001 com 2.220 respondentes. A análise foi feita utilizando modelos de regressão logística, as variáveis socioeconômicas analisadas foram gênero, idade, grau de escolaridade, renda, tipo de domicílio (um único habitante, domicílio sem crianças ou domicílio com crianças) e situação empregatícia. Os resultados apontaram os fatores gênero e idade como influentes no comércio eletrônico, sendo os homens e as pessoas entre 25 e 35 anos mais propensos a comprar *online*. Além disso, o estudo também verificou a influência do acesso a oportunidades de compra em lojas físicas, concluindo que pessoas que vivem em regiões de maior acesso a oportunidades de compra no varejo tradicional compram pela *Internet* tanto quanto pessoas que vivem em regiões de menor acesso.

Farag *et al.* (2005) estudaram a relação entre buscas por produtos *online*, compras *online* e viagens para compras em lojas físicas. Na análise, também foi estudada a influência de fatores socioeconômicos (gênero, renda, idade, grau de escolaridade e composição do domicílio – habitante mora sozinho ou acompanhado) e do acesso ao varejo tradicional no *e-commerce*. Os dados foram obtidos por meio de um estudo realizado em 2003 com usuários da *Internet*, alcançando 826 respostas válidas de pessoas de quatro cidades (uma em região urbana e três em regiões suburbanas) na Holanda. O tipo de *e-commerce* considerado foi o que não incluía produtos alimentícios e outros produtos de uso diário, e o acesso ao varejo tradicional foi medido como o número de lojas de produtos não alimentícios/uso diário por metro quadrado em uma cidade, tendo a cidade em região urbana maior acesso e as três em regiões suburbanas

menor acesso. Os autores obtiveram os resultados por meio de análise de caminho e concluíram que os fatores gênero, idade, renda e grau de escolaridade influenciam no comércio eletrônico, sendo homens, pessoas entre 26 e 45 anos, pessoas de maior renda e pessoas de maior grau de escolaridade mais propensas a comprar pela *Internet*. Além disso, habitantes de regiões urbanas têm mais chances de comprar *online* do que residentes de cidades suburbanas.

Krizek *et al.* (2005) analisaram o uso do *e-commerce* e *e-banking* por residentes de regiões metropolitanas de cidades dos Estados Unidos. Por meio do uso de dados obtidos por um questionário realizado em 2003, o estudo avaliou a propensão dos indivíduos em realizar tais atividades por meio da *Internet* com o objetivo de verificar a influência do acesso ao serviço tradicional e do congestionamento no uso de tais serviços *online*. Verificou-se também a influência de fatores socioeconômicos (idade, grau de escolaridade, renda, número de carros no domicílio e número de adolescente no domicílio). A análise foi realizada por meio de modelos de regressão logística. Foi possível concluir que o acesso e congestionamentos não influenciam o uso de *e-commerce*, enquanto os fatores socioeconômicos exercem influência – pessoas mais jovens, com maior renda e maior grau de escolaridade têm maiores chances de fazerem uso do comércio eletrônico.

Em seu estudo sobre a relação entre *e-commerce* e compras no varejo tradicional, Farag *et al.* (2006a) investigaram os fatores que influenciam o comércio eletrônico, incluindo fatores socioeconômicos, espaciais (acesso a lojas físicas), fatores relacionados à experiência com a *Internet* e atitudes do consumidor. Dados foram obtidos em 2002 sobre consumidores da cidade de Minneapolis, Estados Unidos, pelo Departamento de Transportes do estado, e em 2003 sobre consumidores da cidade de Utrecht, Holanda, em uma pesquisa realizada pela Universidade de Utrecht. As variáveis socioeconômicas estudadas foram gênero, idade, grau de escolaridade, renda e posse de um ou mais carros, enquanto as variáveis espaciais se baseavam no acesso ao varejo tradicional medido em termos do tempo de viagem às lojas físicas. Inicialmente foram usados testes qui-quadrado para verificar se havia significância estatística nas diferenças entre compradores virtuais e tradicionais. Esses testes concluíram que os homens, pessoas de maior renda e pessoas de maior grau de escolaridade (em ambas as cidades) compram mais *online*, assim como os jovens (em Minneapolis), domicílios com dois ou mais carros (em Utrecht) e pessoas com maior acesso a lojas físicas (em Utrecht). Em seguida, os dados foram analisados por meio de regressão logística, para considerar o efeito conjunto das variáveis. Essa análise concluiu que em Utrecht, mulheres têm menores chances de comprar *online*, enquanto em Minneapolis, pessoas mais velhas têm menores chances de aderir ao *e-commerce*, em ambas as

idades, pessoas de maior renda têm mais chances de comprar pelo comércio eletrônico. Além disso, verificou-se que em Utrecht pessoas com menor tempo de viagem a lojas físicas compram mais pela *Internet*.

Farag *et al.* (2006b) investigaram o impacto de variáveis espaciais no comércio eletrônico na Holanda. Os autores analisaram dados obtidos entre os anos de 1996 e 2001 a partir de pesquisa com usuários da *Internet* realizada por uma empresa privada. Por meio de modelos de regressão logística, os autores testaram duas hipóteses: o comércio eletrônico é um fenômeno predominantemente urbano, pois o surgimento de novas tecnologias geralmente se dá no meio urbano – teoria da difusão da inovação (Hägerstrand, 1967; Rogers, 1983); e as pessoas tendem a comprar mais *online* quando têm baixo acesso a lojas físicas – teoria da eficiência (Anderson *et al.*, 2003). Os resultados obtidos confirmaram as duas hipóteses para o cenário holandês: no meio urbano, as pessoas são mais propensas a usar o *e-commerce*, enquanto no meio rural, onde há baixo acesso a lojas físicas, as pessoas compram pela *Internet* com mais frequência. O acesso foi medido como o número de lojas físicas que poderiam ser acessadas em um determinado tempo pelo consumidor de carro a partir de sua residência. Além disso, o estudo também verificou que homens, pessoas com maior grau de escolaridade, pessoas de maior renda e pessoas de até 33 anos têm maior probabilidade de comprar *online*.

Farag *et al.* (2007) estudaram a relação entre buscas *online*, compras *online* e compras presenciais de produtos não alimentícios e outros produtos de uso não-diário. No estudo, também foi analisada a influência de fatores socioeconômicos e do acesso a oportunidades de compra em lojas físicas no comércio eletrônico. Os dados foram obtidos por meio de questionários aplicados a residentes de quatro cidades holandesas (uma em região urbana e três em regiões suburbanas), alcançando 826 respostas válidas. O acesso às lojas físicas foi avaliado em termos do total de metros quadrados de lojas disponível nas cidades, sendo a cidade em região urbana a com maior acesso ao varejo tradicional. Por meio de modelagem de equações estruturais, os dados foram analisados e concluiu-se que moradores de regiões urbanas (com maior acesso ao varejo tradicional) compram mais pela *Internet* que os moradores de cidades em regiões suburbanas (com menor acesso a lojas físicas). Além disso, verificou-se que mulheres e pessoas mais velhas compram menos *online* do que homens e pessoas jovens, enquanto pessoas de maior grau de escolaridade e maior renda compram mais *online*.

De Blasio (2008) estudou as diferenças entre o uso do *e-commerce* em regiões rurais e urbanas da Itália sob a ótica da tecnologia da informação como substituta das aglomerações urbanas.

Com isso, o autor pretendia verificar a hipótese de que o uso da *Internet* pode reduzir a importância das distâncias físicas no sentido de que moradores de regiões rurais – onde há um menor acesso a lojas físicas – fariam mais compras pelo comércio eletrônico. Com base em dados de uso de tecnologia, idade, renda e grau de escolaridade de membros de 8000 domicílios de 344 cidades obtidos pelo Banco da Itália em 2002, o autor usou modelos de regressão logística para realizar a análise. Como resultado, pôde-se concluir que o grau de escolaridade, a renda e a idade do consumidor são variáveis relacionadas ao uso do *e-commerce*, sendo as pessoas de maior renda e grau de educação e mais jovens as maiores compradoras do comércio eletrônico. Além disso, verificou-se que os habitantes das cidades utilizam mais o comércio eletrônico do que os habitantes das áreas rurais, o que pode estar associado ao fato de os moradores rurais se sentirem desencorajados a adquirir produtos pela *Internet* por não poder vê-los antes em uma loja física.

Mokhtarian *et al.* (2009) analisaram atitudes e percepções individuais acerca do *e-commerce* e do varejo tradicional para a definição de segmentos de *marketing* caracterizados por diferentes características, incluindo características socioeconômicas. Para isso, os autores realizaram pesquisas com 966 habitantes de duas cidades universitárias do estado da Califórnia, Estados Unidos, obtendo dados acerca de suas atitudes de compra e características socioeconômicas (gênero, idade, situação empregatícia, renda e grau de escolaridade). Utilizando análise de *cluster*, obtiveram-se sete grupos com diferentes características em relação às suas atitudes em relação às compras e canais por onde escolhem realizar compras (*online* ou varejo tradicional), como por exemplo, “*store shopaholics*” e “*practical and leisure-oriented*”, sendo diferentes características socioeconômicas fortemente associadas a cada grupo.

Ren e Kwan (2009) estudaram a influência do contexto socioeconômico e do acesso ao varejo tradicional na adoção do comércio eletrônico na Região Metropolitana de Columbus, nos Estados Unidos. Dados relativos ao uso da *Internet* para o *e-commerce* foram obtidos em 2000 e 2004 por meio de um questionário *online*, dados socioeconômicos (gênero, cor ou raça, grau de escolaridade e renda) foram retirados do censo demográfico, dados de número de pacotes entregues foram extraídos de uma base pública, assim como dados sobre o número de lojas presentes na região. Os fatores que influenciam o comércio eletrônico foram estimados por modelos de regressão logística. A análise determinou que os fatores socioeconômicos que influenciam o *e-commerce* são gênero e cor ou raça, sendo mulheres e pessoas brancas usuários mais frequentes do comércio eletrônico. Os autores acreditam que renda e grau de escolaridade não se mostraram significativos, pois o estudo foi desenvolvido em uma região de grande

homogeneidade em relação a esses aspectos. Além disso, determinou-se também que pessoas que vivem em regiões de baixo acesso a lojas físicas tem maior propensão a comprar *online*, sendo o acesso definido em termos de tempos de viagem até lojas físicas e disponibilidade de entregas na região. Os autores destacam, no entanto, que a influência do acesso é pequena.

Em sua análise sobre a relação entre a realização de compras *online* e a geração de viagens pelo consumidor, Cao (2012) investigou brevemente a influência dos fatores socioeconômicos idade, gênero, renda, ocupação, situação empregatícia, grau de escolaridade, número de carros por domicílio, posse de cartão de crédito e posse de carteira de motorista na realização de compras *online*. Os dados foram obtidos por meio de um questionário realizado entre 2008 e 2009 com 540 usuários de *Internet* da Região Metropolitana de Minneapolis-St. Paul, Estados Unidos. Por meio de modelos de regressão logística, constatou-se que o grau de escolaridade tem influência nas compras do comércio eletrônico, com pessoas de maior grau de escolaridade realizando mais compras *online*.

Sener e Reeder (2012) estudaram os fatores que influenciam a adoção e a frequência com que indivíduos realizam *teleshopping* (compras pelo telefone ou *online*) e teletrabalho nos Estados Unidos. Por meio de uma pesquisa governamental, obtiveram-se dados socioeconômicos (gênero, idade, grau de escolaridade, presença de crianças e adolescentes no domicílio, renda, número de habitantes no domicílio e número de pessoas empregadas no domicílio), características comportamentais e frequências de realização de teletrabalho ou *teleshopping*. A análise fez uso de modelos de cópula, identificando que, dentre os fatores socioeconômicos, gênero, idade, renda e número de habitantes no domicílio influenciam no *teleshopping*, sendo mulheres, pessoas entre 22 e 36 anos, pessoas de maior renda e domicílios com mais habitantes mais propensos a adotar o *teleshopping*. Além disso, verificou-se também que indivíduos que vivem em áreas urbanas com população maior que 1 milhão de habitantes têm maior preferência pela adoção do *teleshopping*.

Cao *et al.* (2013) aplicaram modelagem de equações estruturais para analisar a distribuição geográfica dos e-consumidores nas cidades de Minneapolis e Saint Paul (Estados Unidos) a partir de dados de uma pesquisa com 585 usuários da *Internet* entre dezembro de 2008 e janeiro de 2009. Neste estudo também foram testadas as hipóteses da teoria da difusão da inovação (Rogers, 1983; Hägerstrand, 1967) e da teoria da eficiência (Anderson *et al.*, 2003). Os resultados também confirmaram ambas as teorias, demonstrando que a influência do acesso (medido, no estudo, em termos do número de empregos na área do varejo em um certo *buffer*)

no comércio eletrônico varia dentro da região de estudo. Os autores verificaram que em áreas urbanizadas os usuários de *Internet* compram mais *online*, pois usam mais a *Internet*; enquanto em áreas não urbanas com baixo acesso a lojas físicas, as pessoas tendem a comprar *online*. Portanto, verifica-se que a influência do acesso não é uniforme, mas depende da área analisada naquela região metropolitana. Além disso, o estudo também verificou que pessoas com maior grau de escolaridade, pessoas com empregos em tempo integral e pessoas de maior renda tendem a comprar *online* com mais frequência nas áreas urbanas.

Em seu estudo da relação entre o comércio eletrônico e a geração de viagens de compra realizadas pelo consumidor, Zhou e Wang (2014) analisaram as variáveis socioeconômicas e regionais que influenciam o *e-commerce*. Os dados de viagens, uso do comércio eletrônico e características socioeconômicas foram retirados de uma pesquisa governamental realizada em 2009 nos Estados Unidos com 308.901 indivíduos. A análise utilizou modelagem de equações estruturais e verificou que, dentre as variáveis gênero, idade, grau de escolaridade, número de habitantes no domicílio, renda e situação empregatícia, as que têm influência no *e-commerce* são a renda, a idade, o número de habitantes no domicílio e o grau de escolaridade. Portanto, pessoas de maior renda, domicílios com menor número de habitantes, pessoas mais jovens e pessoas com maior grau de escolaridade compram pela *Internet* com maior frequência. Além disso, o estudo verificou que populações de regiões de maior densidade populacional tendem a comprar mais *online*.

Clarke *et al.* (2015) analisaram as variações espaciais e a influência de fatores socioeconômicos no comércio *online* de produtos de mercearia no Reino Unido. O método usado consistiu em regressões logísticas para uma base de dados privada proveniente de uma pesquisa com o consumidor realizada entre 2004 e 2010 combinada com dados censitários. O estudo demonstrou que os fatores socioeconômicos analisados (idade, gênero e renda) influenciam no uso do comércio eletrônico, sendo homens, pessoas entre 25 e 44 anos e pessoas de maior renda mais propensas a comprar *online*. Demonstrou-se também que o comércio eletrônico não é mais um fenômeno exclusivamente urbano, tendo aumentado nas áreas rurais devido à melhora na cobertura de *Internet*. Além disso, concluiu-se que o acesso às mercearias em uma região também influencia nas compras do comércio eletrônico, isto é, pessoas que vivem longe de lojas físicas tendem a comprar *online*. No estudo, o acesso foi medido em termos de distância (em milhas) até lojas físicas.

Lee *et al.* (2015) estudaram os fatores que influenciam a probabilidade de se comprar *online*, sendo eles fatores socioeconômicos, fatores relacionados ao ambiente construído (definido no estudo como o acesso em termos de distância ao *shopping center* mais próximo), fatores relacionados aos tipos de produtos adquiridos e fatores envolvendo percepções e atitudes sobre a participação do indivíduo no *e-commerce*. Os dados foram obtidos por meio de um questionário aplicado a usuários e não-usuários do comércio eletrônico na cidade de Davis, Estados Unidos; foram obtidas 2.043 respostas válidas nos anos de 2009 e 2010. Os respondentes foram questionados sobre sua frequência de compra de produtos *online* (exceto alimentos), suas características socioeconômicas (gênero, idade, acesso a um automóvel, grau de escolaridade, situação empregatícia e renda), acesso a *shopping centers*, atitudes de compra (envolvendo fatores como uso da tecnologia, transportes, estilo de vida e meio ambiente) e percepções sobre o comércio eletrônico (como, por exemplo, em relação a preço, qualidade do produto, seleção do produto, horas de operação de lojas e *websites* de compra). Foi utilizada regressão logística para determinar quais fatores socioeconômicos e do ambiente construído têm influência no comércio eletrônico. Os resultados constataram que jovens (menores de 35 anos) e pessoas de maior renda têm mais chances de realizar compras *online*, portanto, os fatores idade e renda têm influência no *e-commerce*.

Em seu estudo das relações entre compras realizadas *online* e compras presenciais, Lee *et al.* (2017) exploraram o efeito de variáveis socioeconômicas, de atitudes e percepções do consumidor e do ambiente (uso do solo) na frequência de compras *online*. Os dados foram obtidos em 2009 e 2010 por meio de um questionário realizado com mais de 2000 habitantes de três localidades – duas na cidade de Davis (Estados Unidos) e uma nos arredores desta cidade. Os autores fizeram uso de modelos de cópula para obter possíveis fatores explicativos (renda, grau de escolaridade, idade, situação empregatícia, disponibilidade de uso de carro, percepções e atitudes em relação ao comércio eletrônico e uso do solo) para as variáveis relacionadas à frequência de compras *online* e frequência de compras presenciais nas três localidades descritas. Os resultados obtidos mostram que, dentre as variáveis socioeconômicas, a renda influencia no *e-commerce* (pessoas com maior renda utilizam mais o comércio eletrônico), enquanto a idade não se mostrou significativa. Os autores destacam, que a idade pode não ter apresentado influência no modelo desenvolvido, pois esse fator pode estar indiretamente associado aos fatores relacionados a atitudes e preferências, já que estas tendem a variar conforme a idade do indivíduo.

Com o objetivo de investigar as características de pessoas que trabalham e realizam compras em casa pela *Internet*, Loo e Wang (2017) verificaram a influência de fatores socioeconômicos, hábitos de uso da *Internet*, atitudes em relação ao *e-working* e ao *e-shopping* e acesso geográfico na realização dessas duas atividades. Os dados foram obtidos por meio de um questionário realizado com 608 trabalhadores integrais que realizavam alguma atividade de *e-working* na cidade de Nanjing, China, em 2015. As variáveis socioeconômicas consideradas foram gênero, idade, grau de escolaridade, renda e número de crianças com menos de 16 anos no domicílio, enquanto o acesso geográfico foi definido como as distâncias da casa do respondente até o seu local de trabalho, estação de metrô mais próxima e ao *shopping center* mais próximo. Por meio de modelos de regressão logística, verificou-se que os fatores gênero e idade influenciam no *e-commerce*, sendo mulheres e pessoas mais jovens (21 a 30 anos) mais engajados nessa atividade. Quanto à variável de acesso, a análise demonstrou que o acesso influencia no comércio eletrônico; pessoas com acesso ruim ao transporte público e a *shopping centers* gastam mais tempo em atividades de *e-commerce* em casa.

Motte-Baumvol *et al.* (2017) realizaram um estudo da influência do acesso a lojas físicas no comércio eletrônico a fim de validar a hipótese da eficiência proposta por Anderson *et al.* (2003). Além disso, também foi estudada a influência de fatores socioeconômicos (idade e renda) no comércio eletrônico. Os dados foram obtidos por meio de questionários realizados em 2013 e 2014 em três regiões metropolitanas francesas. Os respondentes foram perguntados a respeito de seus hábitos de compras e características socioeconômicas. As regiões estudadas foram divididas em subúrbios internos (mais próximos do centro da cidade) e externos (mais distantes do centro da cidade). Por meio de modelos de regressão logística, foi evidenciado que a maioria dos compradores de produtos não alimentícios pela *Internet* são jovens, pessoas de alta renda e que vivem em subúrbios externos ou áreas urbanas de pequeno porte. No entanto, a hipótese da eficiência foi apenas parcialmente validada, uma vez que não foi possível confirmar que o *e-commerce* substitui as compras em lojas físicas.

Beckers *et al.* (2018) analisaram a relação entre a predisposição a comprar *online* e fatores socioeconômicos na Bélgica para identificar o perfil do e-consumidor e como ele se encontra distribuído no espaço. Os dados usados foram obtidos por meio de uma pesquisa por questionários realizada pela federação do comércio belga com 1.500 respondentes em 2016, na qual foram respondidas perguntas sobre frequência de compras pela *Internet* e características socioeconômicas. Estimando modelos de regressão logística, os autores identificaram que fatores socioeconômicos – dentre gênero, idade, grau de escolaridade, renda e número de



crianças no domicílio –influenciam a probabilidade de um indivíduo realizar compras online na região de estudo. Assim, foi identificado que homens, pessoas na faixa dos 30 anos, pessoas de maior grau de escolaridade e pessoas de maior renda têm maior probabilidade de comprar *online*, portanto os fatores gênero, idade, grau de escolaridade e renda têm influência no *e-commerce*. Além disso, verificou-se que o grau de urbanização de uma região não tem influência na probabilidade de efetivar uma compra *online*. Os autores analisaram também o impacto da distribuição geográfica dos e-consumidores, concluindo que há uma subestimação ou superestimação de até 50% no cálculo do número de e-consumidores em potencial quando se considera uma distribuição homogênea do *e-commerce* numa área de estudo.

Kirby-Hawkins *et al.* (2018) investigaram o padrão de localização espacial dos consumidores de produtos de mercearia no Reino Unido. Os autores analisaram dados de compras fornecidos por uma rede de mercearias que atua na área de estudo por meio de modelos de interação espacial, concluindo que os fatores idade e grau de escolaridade influenciam no comércio eletrônico, sendo os jovens que moram em regiões urbanas e as pessoas de maior grau de escolaridade, principalmente as que vivem em regiões rurais, maiores usuárias do *e-commerce*. Além disso, semelhante a estudos anteriores, os resultados confirmam tanto a teoria da difusão da inovação quanto a teoria da eficiência na área de estudo. O acesso a lojas físicas, no estudo, foi medido com base na área disponível de lojas de produtos de mercearia por mil habitantes nas regiões analisadas.

Hood *et al.* (2020) analisaram os fatores que influenciam o comércio eletrônico de produtos de mercearia na Grã-Bretanha e a distribuição geográfica deste comércio por meio de análises estatísticas. Os dados foram obtidos de um questionário eletrônico respondido por 19.033 pessoas realizado entre 2015 e 2016 por uma empresa de pesquisa em *marketing* contendo perguntas relacionadas à frequência em que o consumidor realiza compras online e dados socioeconômicos como gênero, idade e renda. Os resultados concluíram que gênero, renda e idade influenciam na escolha pelo *e-commerce*, sendo mulheres, pessoas de maior renda e pessoas entre 25 e 44 anos e com mais de 55 anos mais propensas a comprar produtos de mercearia pelo comércio eletrônico. Além disso, semelhantemente a estudos anteriores, os autores obtiveram resultados que validam tanto a teoria da difusão inovação quanto a da eficiência para a região estudada.

Jaller e Pahwa (2020), com o objetivo de avaliar quais são os fatores de influência nas entregas do comércio eletrônico, analisaram o comportamento de diferentes grupos demográficos de e-

consumidores nos Estados Unidos a partir de dados obtidos em uma pesquisa governamental realizada em 2016. Por meio de regressões logísticas, os autores identificaram que, dentre os fatores socioeconômicos analisados (gênero, idade, grau de educação, situação empregatícia e renda), os que influenciam as entregas domiciliares são gênero, grau de escolaridade e número de filhos, sendo mulheres, pessoas de maior grau de escolaridade e domicílios com filhos mais propensos a comprar *online*. Os autores também concluíram que pessoas que vivem em cidades populosas do oeste dos Estados Unidos têm maior propensão a usar o comércio eletrônico.

Saphores e Xu (2020) analisaram o comércio eletrônico de produtos em geral e de produtos de mercearia nos Estados Unidos nos anos de 2009 e 2017 em relação às características socioeconômicas dos consumidores. Foram utilizados dados de duas pesquisas governamentais, realizadas em 2009 e em 2017, sobre entregas domiciliares e adesão ao *e-commerce*, além de dados socioeconômicos dos consumidores (estado civil, gênero, idade, situação empregatícia, grau de escolaridade, cor ou raça, renda, número de crianças no domicílio, número de adultos no domicílio). Por meio de modelos de regressão logística, determinou-se que a única variável significativa para as entregas domiciliares de produtos de mercearia comprados *online* é o gênero, sendo as mulheres as maiores consumidoras. Já para o comércio eletrônico em geral, as variáveis gênero, idade, renda, número de crianças no domicílio, grau de escolaridade e cor ou raça têm influência no número de entregas, sendo mulheres, pessoas jovens, pessoas com maior renda, pessoas com maior grau de escolaridade, pessoas brancas e domicílios com mais crianças mais prováveis de receberem entregas domiciliares providas do *e-commerce*. Além disso, os autores também usaram a densidade populacional para verificar a influência dos ambientes urbano e rural no número de entregas domiciliares, constatando que, embora a influência seja pequena, em áreas de menor densidade ( $< 300$  pessoas por milha quadrada), os domicílios têm menos chances de comprar *online*, mas os que o fazem, recebem mais entregas, e em áreas maior densidade ( $> 7000$  pessoas por milha quadrada), os domicílios têm menores chances de nunca comprar *online*, além de receberem mais entregas.

Cheng *et al.* (2021) estudaram os fatores que influenciam as entregas domiciliares provenientes do comércio eletrônico em Singapura. Foram considerados dados de entregas realizadas por uma transportadora, bem como dados de acesso ao varejo tradicional e socioeconômicos. As variáveis renda, número de habitantes por domicílio, idade, posse de carro, idade da moradia, densidade populacional, acesso a *shopping malls* e acesso ao transporte público foram analisadas por meio de regressão linear. Os resultados apontaram para a idade como sendo a única das variáveis testadas a não exercer influência no número de entregas domiciliares. Maior

densidade populacional, maior renda, maior número de habitantes no domicílio e a posse de um veículo estão associados a um maior número de entregas, enquanto uma maior idade da moradia, maior acesso a *shopping malls* e maior acesso ao transporte público indicam um menor número de entregas domiciliares.

Song (2021) analisou a influência de atributos espaciais no comércio eletrônico na China, verificando a distribuição geográfica da adoção dessa modalidade de compra no país com o objetivo de avaliar a influência de fatores socioeconômicos (gênero, idade, grau de escolaridade e renda) e geográficos (acesso ao transporte público, serviços de entrega domiciliar disponíveis e grau de urbanização) nessa distribuição, e entender a desigualdade no acesso e uso do comércio *online*. Para tanto, foram usados dados de 1.918 distritos, obtidos do maior *site* de compras da China, a respeito da realização de compras *online* em 2017. A análise foi realizada por meio de regressão geograficamente ponderada (GWR), técnica que permite atribuir pesos diferentes às variáveis dependendo de sua localização, eliminando o risco de vieses geográficos. Os resultados indicam que distritos de regiões metropolitanas fazem maior uso do *e-commerce*; esse uso no país decresce a partir das regiões da costa Leste em direção à zona rural no Oeste, indicando uma desigualdade na participação da população no comércio eletrônico, sendo as áreas com maior grau de urbanização maiores usuárias do *e-commerce*. Isso condiz com a teoria da difusão da inovação. Além disso, influenciam positivamente no comércio eletrônico na China a disponibilidade de serviços de entrega domiciliar, o acesso à *Internet* e a renda. Verificou-se também que a idade é influente, com pessoas com menos de 40 anos tendo maior propensão a comprar *online*.

Assim como nos estudos apresentados na seção anterior, nota-se que o método de regressão logística foi o mais utilizado na literatura, sendo empregado em 14 análises (Farag *et al.*, 2003; Krizek *et al.*, 2005; Farag *et al.*, 2006a; Farag *et al.*, 2006b; De Blasio, 2008; Ren e Kwan, 2009; Cao, 2012; Clarke *et al.*, 2015; Lee *et al.*, 2015; Loo e Wang, 2017; Motte-Baumvol *et al.*, 2017; Beckers *et al.*, 2018; Jaller e Pahwa, 2020; Sephores e Xu, 2020). A Tabela 4 resume os métodos utilizados pelos estudos descritos nesta subseção.

Tabela 4: Métodos utilizados pelos estudos descritos

Autor	Regressão logística	SEM	Modelos de cópula	GWR	Regressão linear	Análise de <i>cluster</i>	Análises estatísticas	Qui-quadrado	Análise de caminho	Modelo de interação espacial
Farag <i>et al.</i> (2003)	•									
Farag <i>et al.</i> (2005)									•	
Krizek <i>et al.</i> (2005)	•									
Farag <i>et al.</i> (2006a)	•							•		
Farag <i>et al.</i> (2006b)	•									
Farag <i>et al.</i> (2007)		•								
De Blasio (2008)	•									
Mokhtarian <i>et al.</i> (2009)						•				
Ren e Kwan (2009)	•									
Cao (2012)	•									
Sener e Reeder (2012)			•							
Cao <i>et al.</i> (2013)		•								
Zhou e Wang (2014)		•								
Clarke <i>et al.</i> (2015)	•									
Lee <i>et al.</i> (2015)	•									
Lee <i>et al.</i> (2017)			•							
Loo e Wang (2017)	•									
Motte-Baumvol <i>et al.</i> (2017)	•									
Beckers <i>et al.</i> (2018)	•									
Kirby-Hawkins <i>et al.</i> (2018)										•
Hood <i>et al.</i> (2020)							•			
Jaller e Pahwa (2020)	•									
Saphores e Xu (2020)	•									
Cheng <i>et al.</i> (2021)					•					
Song (2021)				•						
<b>Quantidade de trabalhos</b>	<b>14</b>	<b>3</b>	<b>2</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>

Percebe-se também que os fatores socioeconômicos idade, renda, gênero e grau de escolaridade foram os mais apontados na literatura como influentes no comércio eletrônico. Cabe destacar que, assim como na seção anterior, a maioria dos estudos realizados fez uso de dados autodeclarados provenientes de questionários; apenas quatro estudos fizeram uso de dados reais de compras (Ren e Kwan, 2009; Kirby-Hawkins *et al.*, 2018; Cheng *et al.*, 2021; Song, 2021). Ainda, apenas dois estudos das áreas de geografia e transporte abordaram a realidade de um país em desenvolvimento – no caso, a China (Loo e Wang, 2017; Song, 2021).

Por fim, observa-se que o impacto espacial no comércio eletrônico ainda merece atenção. Fatores espaciais, como a distribuição das oportunidades de compra ou a distribuição de grupos demográficos de e-consumidores no espaço, foram estudados por poucos autores (Frag *et al.*, 2003; Farag *et al.*, 2005; Krizek *et al.*, 2005; Farag *et al.*, 2006a; Farag *et al.*, 2006b; Farag *et al.*, 2007; Ren e Kwan, 2009; Cao *et al.*, 2013; Clarke *et al.*, 2015; Loo e Wang, 2017; Beckers *et al.*, 2018; Kirby-Hawkins *et al.*, 2018; Hood *et al.*, 2020; Cheng *et al.*, 2021; Song, 2021). Sabe-se que uma consideração homogênea da distribuição do *e-commerce* numa área pode levar a erros de até 50% na estimativa do número de consumidores na região (Beckers *et al.*, 2018) e um baixo nível de desagregação dos dados pode impactar em perdas de informações sobre as relações entre as variáveis que impactam o comércio eletrônico (Lee *et al.*, 2017).

A Tabela 5 apresenta uma síntese dos fatores analisados e dos fatores influentes, tipo de local e tipo de dado estudado nos estudos descritos nesta subseção.

Tabela 5: Síntese dos principais aspectos dos estudos discutidos

Referência	Fatores							Local		Dado	
	Gênero	Idade	Grau de escolaridade	Renda	Habitantes/domicílio	Cor ou raça	Acesso ao varejo tradicional	Países desenvolvidos	Países em desenvolvimento	Dado real	Dado autodeclarado
Farag <i>et al.</i> (2003)	*	*	•	•	•		•	•			•
Farag <i>et al.</i> (2005)	*	*	*	*	•		*	•			•
Krizek <i>et al.</i> (2005)		*	*	*	•		•	•			•
Farag <i>et al.</i> (2006a)	*	*	*	*			*	•			•
Farag <i>et al.</i> (2006b)	*	*	*	*			*	•			•
Farag <i>et al.</i> (2007)	*	*	*	*			*	•			•
De Blasio (2008)		*	*	*			*	•			•
Mokhtarian <i>et al.</i> (2009)	*	*	*	*				•			•
Ren e Kwan (2009)	*		•	•		*	*	•		•	•
Cao (2012)	•	•	*	•				•			•
Sener e Reeder (2012)	*	*	•	*	*			•			•
Cao <i>et al.</i> (2013)			*	*			*	•			•
Zhou e Wang (2014)	•	*	*	*	*			•			•
Clarke <i>et al.</i> (2015)	*	*		*			*	•			•
Lee <i>et al.</i> (2015)	•	*	•	*			•	•			•
Lee <i>et al.</i> (2017)		•	•	*				•			•
Loo e Wang (2017)	*	*	•	•	•		*		•		•
Motte-Baumvol <i>et al.</i> (2017)		*		*			*	•			•
Beckers <i>et al.</i> (2018)	*	*	*	*	•			•			•
Kirby-Hawkins <i>et al.</i> (2018)		*	*				*	•		•	
Hood <i>et al.</i> (2020)	*	*		*			*	•			•
Jaller e Pahwa (2020)	*	•	*		*			•			•
Saphores e Xu (2020)	*	*	*	*	*	*		•			•

Cheng <i>et al.</i> (2021)		•		*	*		*	•		•	
Song (2021)	•	*	•	*					•	•	

Na coluna Fatores: variável analisada “•”; variável analisada e influente “\*”

A partir desta revisão da literatura, identifica-se que ainda há possibilidades a serem exploradas em relação à influência dos fatores socioeconômicos e do acesso ao varejo tradicional no comércio eletrônico. Identificou-se que há uma carência de estudos que considerem a espacialidade dos dados e que analisem dados de entregas domiciliares, como realizado nesta dissertação. Destaca-se também a possibilidade do uso de técnicas espaciais na análise deste fenômeno. Além disso, este trabalho realiza também a comparação entre diferentes regiões de estudo. Por fim, um estudo como o proposto seria o primeiro, a saber, sobre o contexto brasileiro e um dos poucos sobre países em desenvolvimento.

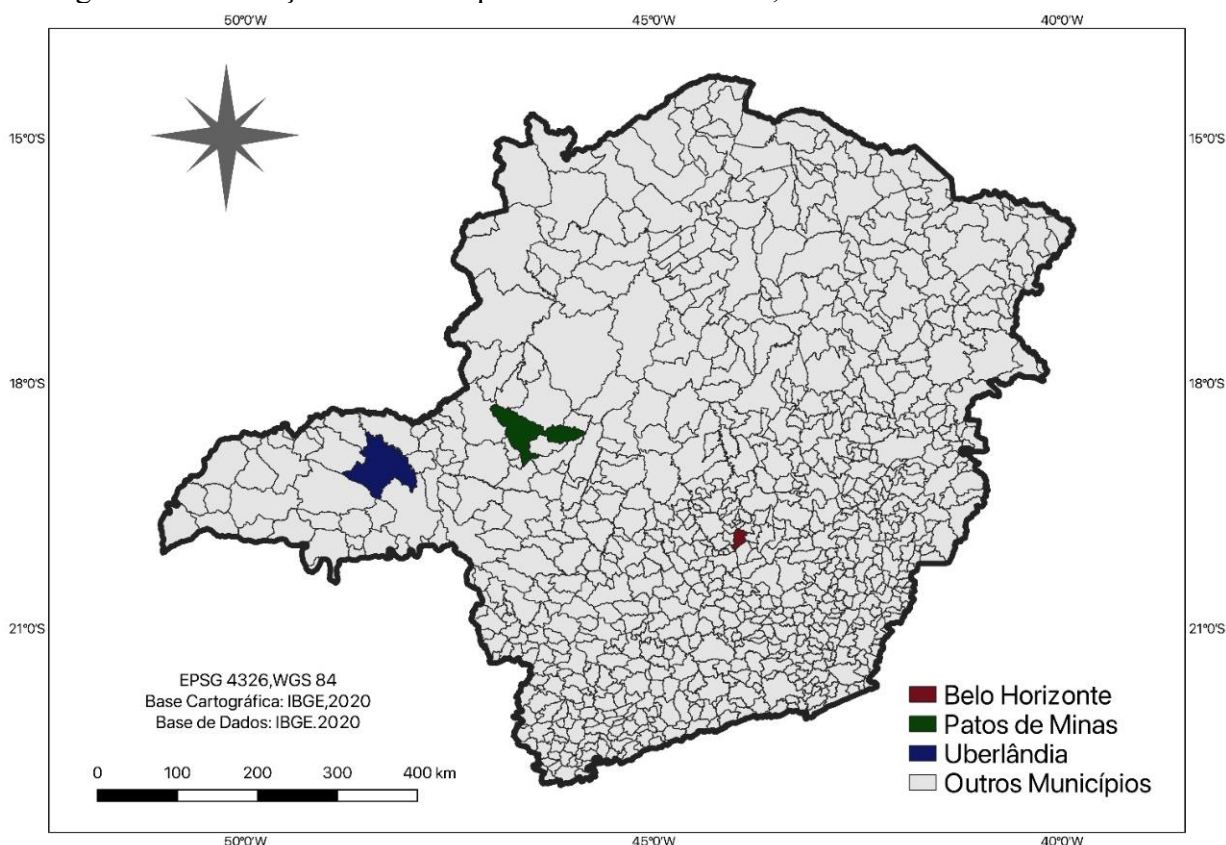
Diante do exposto, este trabalho pretende contribuir para um melhor entendimento dos fatores que influenciam as entregas do comércio eletrônico, especialmente no contexto espacial. O conhecimento aprofundado das características relacionadas às entregas pode contribuir para o planejamento urbano no que tange o transporte de carga e suas externalidades, auxiliando na implementação de políticas públicas visando uma maior sustentabilidade da cidade. Por exemplo, tal conhecimento poderia auxiliar na designação dos melhores locais para a instalação de centros de distribuição de mercadorias (Saphores e Xu, 2020) e *pick-up points* (pontos de retirada de mercadoria pelo consumidor) (Allen *et al.*, 2018) e na identificação de locais que poderiam sofrer maior impacto ambiental e no tráfego causados pelo transporte de mercadorias (Sephores e Xu, 2020).

Além disso, este estudo pode também contribuir com outras áreas do conhecimento, como o *marketing* e os negócios. O conhecimento dos fatores socioeconômicos que influenciam o comércio eletrônico poderia auxiliar na criação de estratégias de venda e/ou entregas para determinados grupos, além de estratégias de vendas direcionadas aos locais onde há um menor acesso ao varejo tradicional (Clarke *et al.*, 2015; Bucko *et al.*, 2018; Kirby-Hawkings *et al.*, 2018).

### 3 ÁREA DE ESTUDO

Neste trabalho foram analisados dados referentes a três cidades mineiras: Belo Horizonte, Uberlândia e Patos de Minas. Essas cidades foram escolhidas devido à intenção de se realizar uma comparação entre cidades com diferentes características sociodemográficas, e também devido à disponibilidade de dados sobre entregas do comércio eletrônico. A Figura 1 mostra a localização dos três municípios em Minas Gerais.

Figura 1: Localização dos municípios de Belo Horizonte, Uberlândia e Patos de Minas.



Segundo IBGE (2020a), uma rede urbana é classificada, no Brasil, em diferentes níveis hierárquicos baseados na atratividade da sua oferta de equipamentos e serviços para populações de outras cidades e na sua influência sobre outras cidades. Belo Horizonte é classificada como uma metrópole, pois exerce influência direta ou indiretamente em todas as cidades brasileiras. Uberlândia recebeu a classificação de Capital Regional B, por concentrar atividades de gestão, porém em um alcance menor que uma metrópole. Patos de Minas foi considerada um Centro Sub-regional A, pois possui atividades de gestão com menor nível de complexidade e influência do que uma Capital Regional. As subseções abaixo detalham cada uma das áreas de estudo.

### 3.1 Descrição das áreas de estudo

Belo Horizonte é a capital do estado. Sua população estimada em 2020 era de mais de 2,5 milhões de pessoas e sua área era de 331,354 km<sup>2</sup>. No censo de 2010, sua população era de mais de 2,3 milhões de habitantes, sendo a densidade demográfica no município 7.167 hab/km<sup>2</sup>. Naquele mesmo ano, a taxa de escolarização da população entre 6 e 14 anos era de 97,6% e o Índice de Desenvolvimento Humano Municipal (IDHM) era de 0,810. Em 2018, o salário mensal médio dos trabalhadores formais na cidade era estimado em 3,6 salários -mínimos, sendo o número de pessoas ocupadas mais de 1,4 milhões (58,2% da população estimada na época) e o PIB *per capita* R\$36.759,66 (IBGE, 2020b).

Uberlândia possui a segunda maior população do Estado de Minas Gerais, atrás apenas da capital. Sua população estimada em 2020 era de mais de 699 mil pessoas e sua área era de 4.115,206 km<sup>2</sup>. No censo de 2010, sua população era de mais de 604 mil habitantes, sendo a densidade demográfica no município 146,78 hab/km<sup>2</sup>. Naquele mesmo ano, a taxa de escolarização da população entre 6 e 14 anos era de 98% e o Índice de Desenvolvimento Humano Municipal (IDHM) era de 0,789. Em 2018, o salário mensal médio dos trabalhadores formais era estimado em 2,7 salários-mínimos, sendo o número de pessoas ocupadas mais de 239 mil (35,1% da população estimada na época) e o PIB *per capita* R\$54.801,25 (IBGE, 2020c).

Patos de Minas ocupa a 16<sup>a</sup> posição do *ranking* populacional do estado. Sua população estimada em 2020 era de mais de 153 mil pessoas e sua área era de 3.190,456 km<sup>2</sup>. No censo de 2010, sua população era de mais de 138 mil habitantes, sendo a densidade demográfica no município 43,49 hab/km<sup>2</sup>. Naquele mesmo ano, a taxa de escolarização da população entre 6 e 14 anos era de 98,6% e o Índice de Desenvolvimento Humano Municipal (IDHM) era de 0,765. Em 2018, o salário mensal médio dos trabalhadores formais na cidade era estimado em 2,2 salários-mínimos, sendo o número de pessoas ocupadas mais de 46 mil (30,9% da população estimada na época) e o PIB *per capita* R\$31.066,32 (IBGE, 2020d). A Tabela 6 apresenta um resumo das características de cada município.



Tabela 6: Características dos municípios analisados

<b>Dado</b>	<b>Belo Horizonte</b>	<b>Uberlândia</b>	<b>Patos de Minas</b>	<b>Ano de referência</b>
População estimada	2.521.564	699.097	153.585	2020
População censo	2.375.151	604.013	138.710	2010
Densidade demográfica (hab/km <sup>2</sup> )	7.167,00	146,78	43,49	2010
Área (km <sup>2</sup> )	331,354	4.115,206	3.190,456	2020
Taxa de escolarização (%)	97,6	98	98,6	2010
IDHM	0,810	0,789	0,765	2010
Salário médio mensal *	3,6	2,7	2,2	2018
População ocupada (%)	58,2	35,1	30,9	2018
PIB <i>per capita</i> (reais)	36.759,66	54.801,2	31.066,32	2018

\* em salários-mínimos. FONTE: IBGE (2020b), IBGE (2020c), IBGE (2020d)

### 3.2 Dados

Neste trabalho foram usados três tipos de dados: socioeconômicos, de acesso a oportunidades de compra e de entregas do comércio eletrônico. Os dados socioeconômicos foram obtidos a nível de bairro para cada município analisado por meio do último Censo do IBGE, realizado em 2010 (IBGE, 2010). Os dados utilizados foram gênero (número de homens e de mulheres), idade (dividida em faixas etárias a partir dos 18 anos), renda média, cor ou raça (branca, preta, amarela, parda e indígena), número de domicílios e habitantes por domicílio (obtido a partir do número de domicílios fornecido pelo censo). Os dados do Censo de 2010 foram extrapolados para o ano de 2019, com exceção do número de domicílios e dos habitantes por domicílio, pois não foi possível encontrar um fator que permitisse o ajuste ao nível de bairro. O ajuste dos dados referentes ao número de pessoas por bairro (população, número de homens e de mulheres, número de pessoas por faixa etária e número de pessoas por cor ou raça) foram feitos com base na razão entre a população total estimada para cada cidade em 2019 e a população total da cidade em 2010. O ajuste da renda média para Belo Horizonte foi realizado com base na razão entre a renda média da cidade no ano de 2019 e a renda média da cidade em 2010, enquanto para Uberlândia e Patos de Minas utilizou-se a razão entre a renda média no Brasil em 2019 e em 2010 devido à ausência de estimativa disponível para o ano de 2019 para estas cidades.

Destaca-se que se tentou obter dados sobre grau de escolaridade (fator que foi considerado relevante em outros estudos, conforme mencionado na Revisão da Literatura). No entanto, não foram encontrados dados no nível de detalhamento necessário para este estudo.

A Figura 2 apresenta os dados populacionais referentes aos habitantes maiores de 18 anos de Belo Horizonte, bem como a distribuição de renda, domicílios, habitantes por domicílios, entregas e estabelecimentos comerciais para os 479 bairros da cidade. A Figura 3 apresenta a

distribuição da população maior de 18 anos de Belo Horizonte segundo as cores ou raças definidas pelo IBGE. Por fim, a Figura 4 apresenta a distribuição da população da cidade por idade. Os dados socioeconômicos das figuras referem-se a estimativas para 2019.

Figura 2: Distribuição dos dados populacionais, renda, dados de domicílios, entregas e estabelecimentos comerciais em Belo Horizonte (IBGE, 2010; Transportadora, 2019; BHMap, 2020).

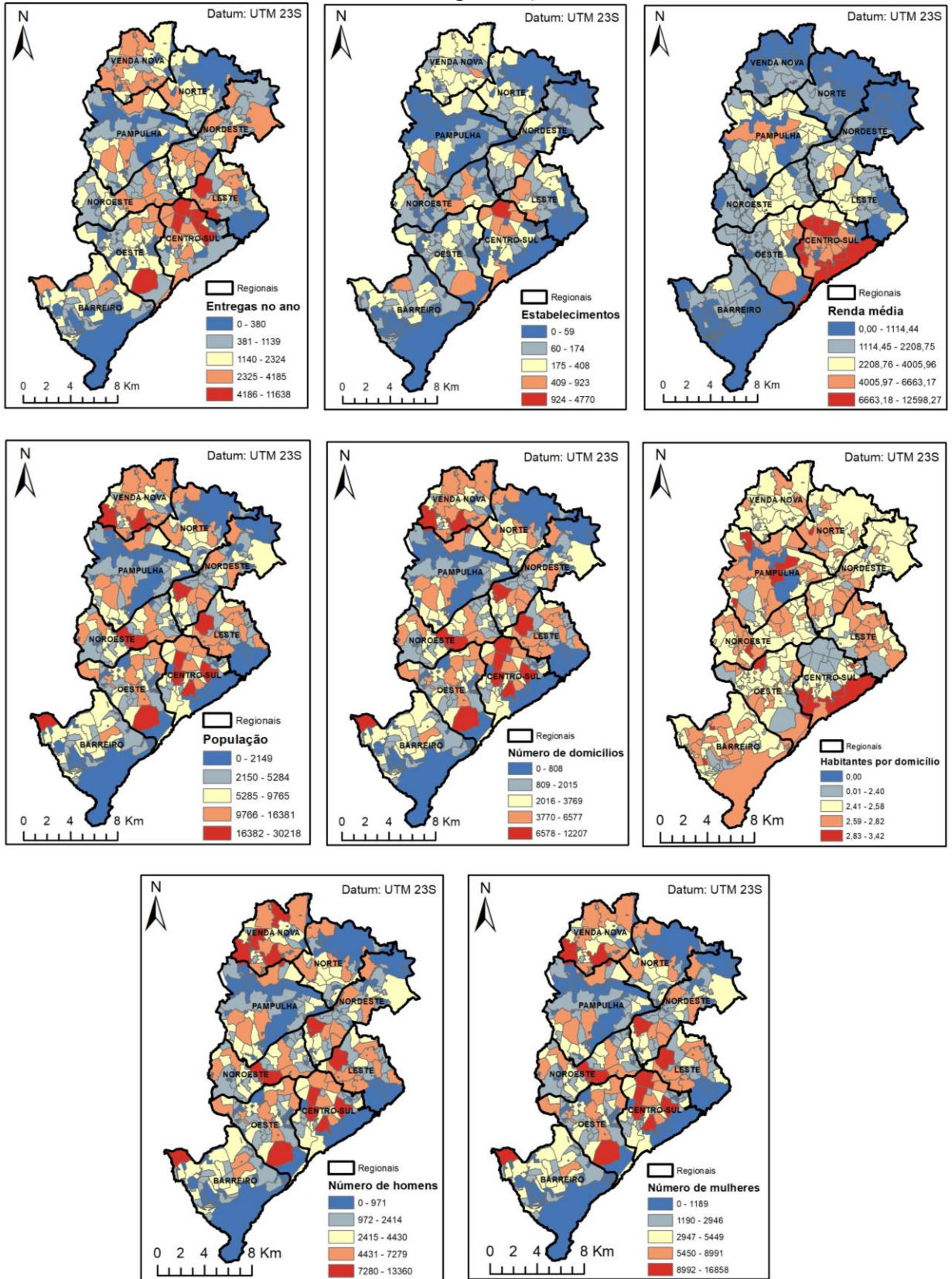


Figura 3: Distribuição da população de Belo Horizonte por cor ou raça (IBGE, 2010; BHMap, 2020).

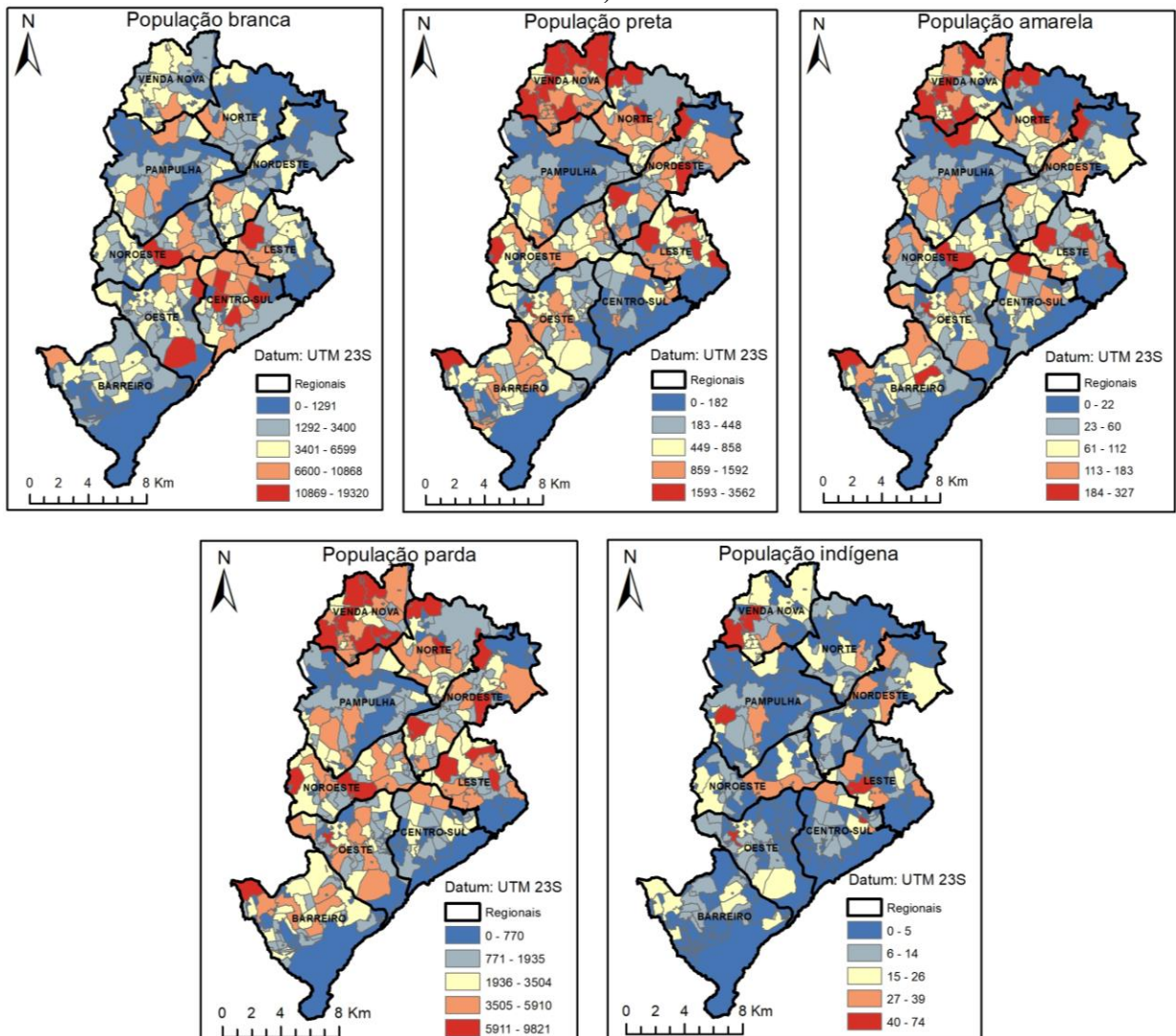
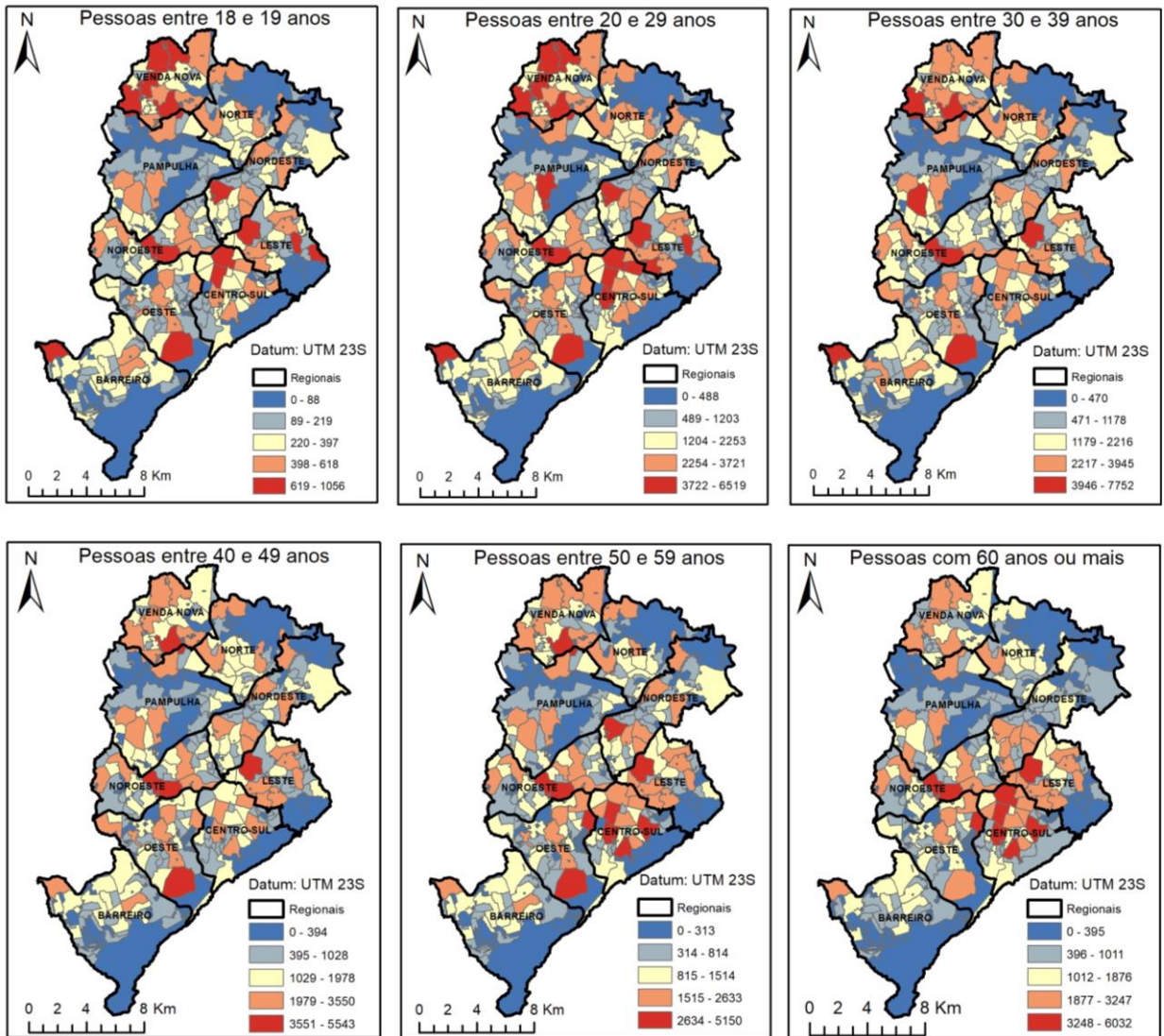




Figura 4: Distribuição da população de Belo Horizonte por faixas de idade (IBGE, 2010; BHMMap, 2020).



Na Tabela 7 se encontra a estatística descritiva dos dados socioeconômicos, do número de entregas e do número de estabelecimentos comerciais em Belo Horizonte.

Tabela 7: Estatística descritiva dos dados de Belo Horizonte.

Variável	Estatística descritiva					
	Mínimo	1º quartil	Mediana	Média	3º quartil	Máximo
Número de entregas	0,0	0,0	378,0	822,3	1200,0	11638,0
População	0	608	2256	4016	5726	30218
Renda média	0,0	720,5	1008,6	1549,4	1608,1	12598,3
Número de homens	0	290	1064	1838	2626	13360
Número de mulheres	0	323	1216	2178	3123	16858
Número de pessoas brancas	0,0	152,5	740,0	1936,9	2566,5	19320,0
Número de pessoas pretas	0,0	86,5	243,0	423,5	575,0	3562,0
Número de pessoas amarelas	0,00	4,00	21,00	42,31	61,50	327,00
Número de pessoas pardas	0	296	976	1607	2280	9821
Número de pessoas indígenas	0,000	0,000	2,000	6,255	8,000	74,000
Número de pessoas entre 18 e 19 anos	0,0	32,0	103,0	167,8	241,5	1056,0
Número de pessoas entre 20 e 29 anos	0,0	166,5	609,0	992,6	1405,5	6519,0
Número de pessoas entre 30 e 39 anos	0	138	498	867	1230	7752
Número de pessoas entre 40 e 49 anos	0,0	114,5	426,0	738,0	1055,5	5543,0
Número de pessoas entre 50 e 59 anos	0,0	79,5	304,0	589,2	845,5	5150,0
Número de pessoas com 60 anos ou mais	0,0	64,0	297,0	661,4	907,0	6032,0
Número de domicílios	0	238	882	1591	2270	12207
Habitantes por domicílio	0,000	2,440	2,540	2,534	2,630	3,420
Estabelecimentos comerciais	0,00	4,00	33,00	90,48	105,00	4770,00

A Figura 5 apresenta os dados populacionais dos habitantes maiores de 18 anos de Uberlândia, assim como a distribuição de renda, domicílios, habitantes por domicílios, entregas e estabelecimentos comerciais. As distribuições de cor ou raça para a população maior de 18 anos de Uberlândia se encontram na Figura 6. As distribuições de idades se encontram na Figura 7. Os dados socioeconômicos das figuras referem-se a estimativas para 2019.

Figura 5: Distribuição dos dados populacionais, renda, dados de domicílios, entregas e estabelecimentos comerciais em Uberlândia (IBGE, 2010; Transportadora, 2019; OpenStreetMap, 2021).

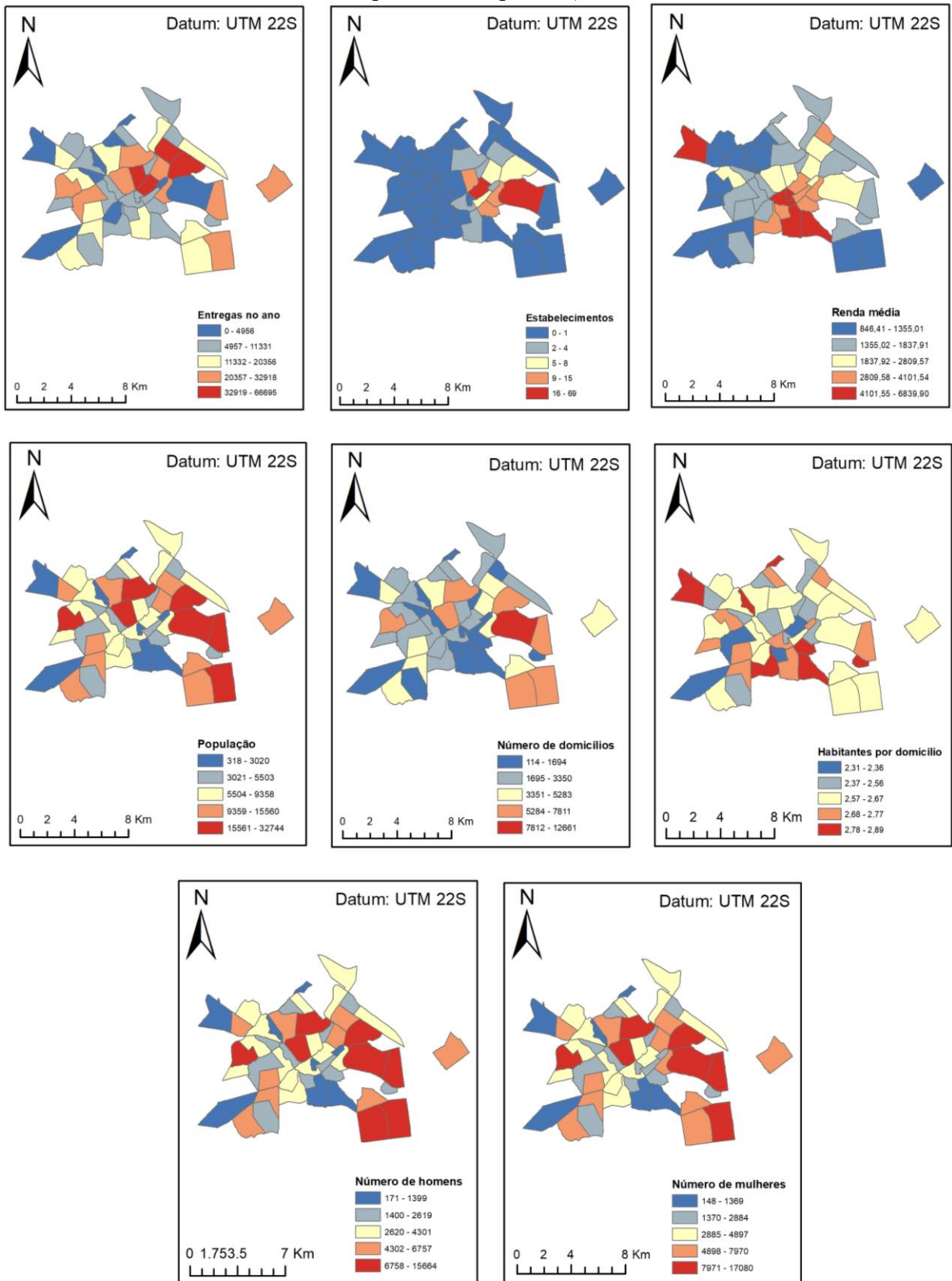


Figura 6: Distribuição da população de Uberlândia por cor ou raça (IBGE, 2010).

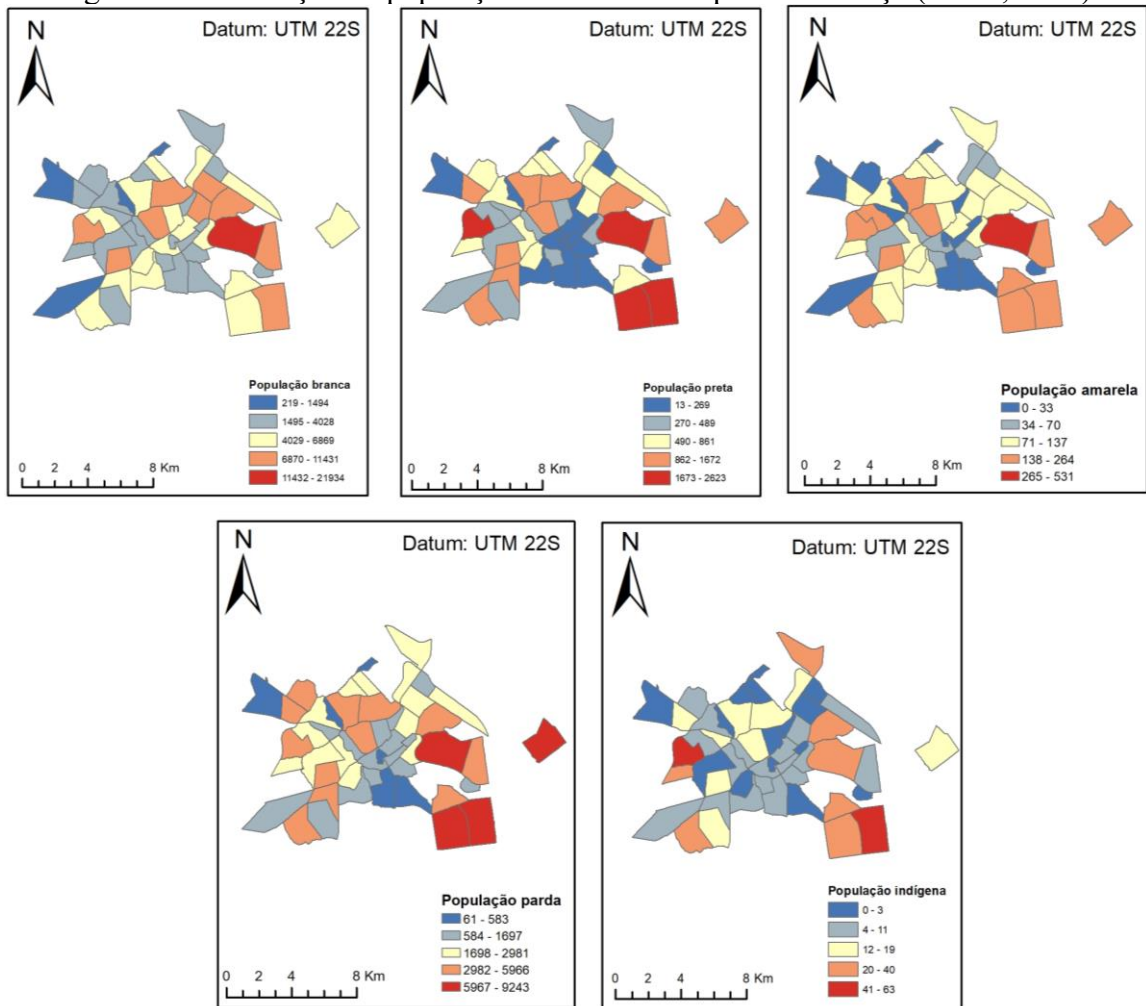
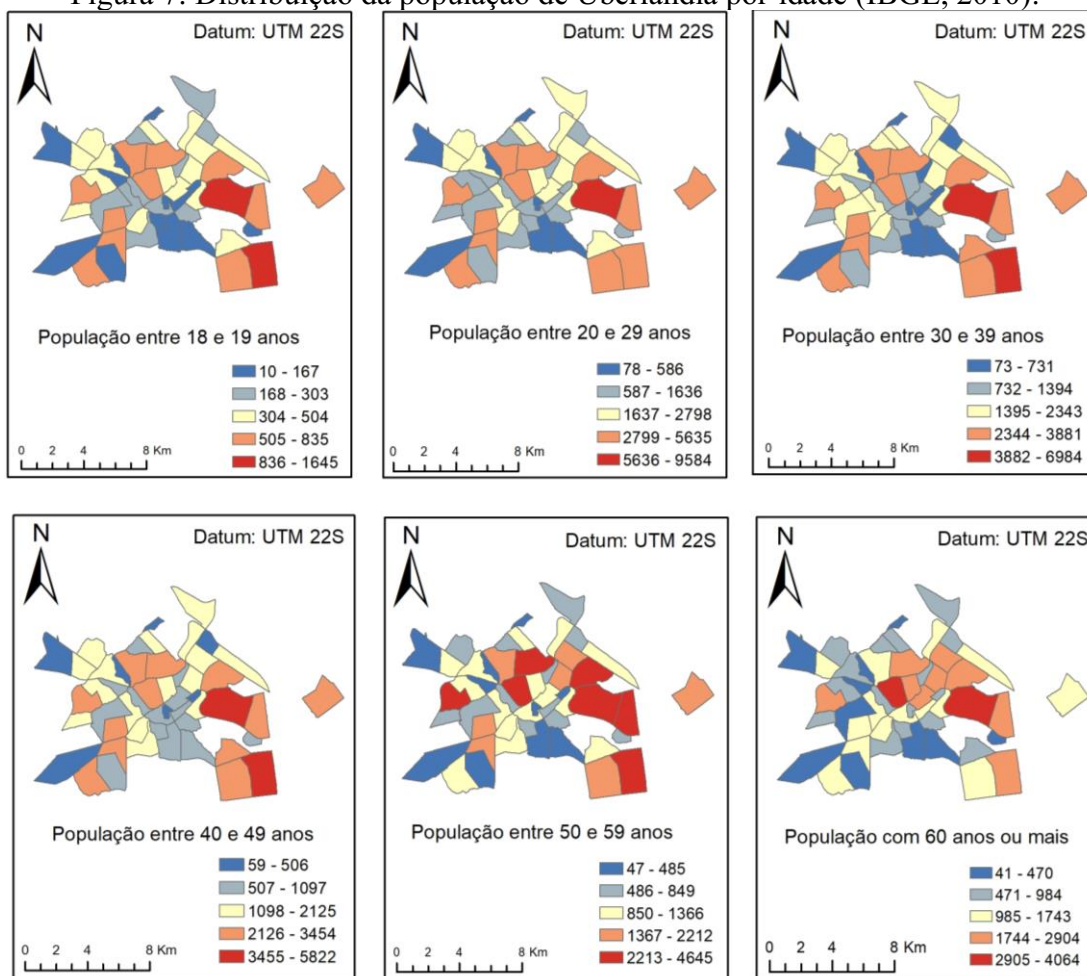




Figura 7: Distribuição da população de Uberlândia por idade (IBGE, 2010).



Na Tabela 8 se encontra a estatística descritiva dos dados socioeconômicos, do número de entregas e do número de estabelecimentos comerciais em Uberlândia.

Tabela 8: Estatística descritiva dos dados de Uberlândia

Variável	Estatística descritiva					
	Mínimo	1º quartil	Mediana	Média	3º quartil	Máximo
Número de entregas	0	7514	10664	15322	21122	66695
População	318	3996	7020	8473	11500	32744
Renda média	846,4	1347,1	1751,3	2363,7	2753,6	6839,9
Número de homens	171	1860	3394	4028	5485	15664
Número de mulheres	148	2136	3633	4444	5914	17080
Número de pessoas brancas	219	2588	4148	4915	6219	21934
Número de pessoas pretas	13,0	240,5	560,0	731,1	1172,0	2623,0
Número de pessoas amarelas	0,0	37,0	84,0	100,8	136,0	531,0
Número de pessoas pardas	61,0	885,5	2337,0	2713,1	4397,5	9243,0
Número de pessoas indígenas	0,00	3,00	8,00	12,35	15,50	63,00
Número de pessoas entre 18 e 19 anos	10,0	183,5	348,0	405,5	543,5	1645,0
Número de pessoas entre 20 e 29 anos	78	1026	1844	2165	2972	9584
Número de pessoas entre 30 e 39 anos	73	827	1618	1858	2742	6984

Número de pessoas entre 40 e 49 anos	59	730	1472	1636	2263	5822
Número de pessoas entre 50 e 59 anos	47,0	582,5	1114,0	1211,6	1612,5	4645,0
Número de pessoas com 60 anos ou mais	41,0	543,5	984,0	1196,7	1498,5	4064,0
Número de domicílios	114	1590	2679	3234	4352	12661
Habitantes por domicílio	2,310	2,560	2,630	2,630	2,725	2,890
Estabelecimentos comerciais	0,000	0,000	0,000	3,824	2,000	69,000

A Figura 8 apresenta os dados populacionais dos habitantes maiores de 18 anos de Patos de Minas, assim como a distribuição de renda, domicílios, habitantes por domicílios, entregas e estabelecimentos comerciais. As distribuições de cor ou raça para a população maior de 18 anos de Patos de Minas se encontram na Figura 9. As distribuições de idades para a cidade se encontram na Figura 10. Os dados socioeconômicos das figuras referem-se a estimativas para 2019.

Figura 8: Distribuição dos dados populacionais, renda, dados de domicílios, entregas e estabelecimentos comerciais em Patos de Minas (IBGE, 2010; Transportadora, 2019; OpenStreetMap, 2021).

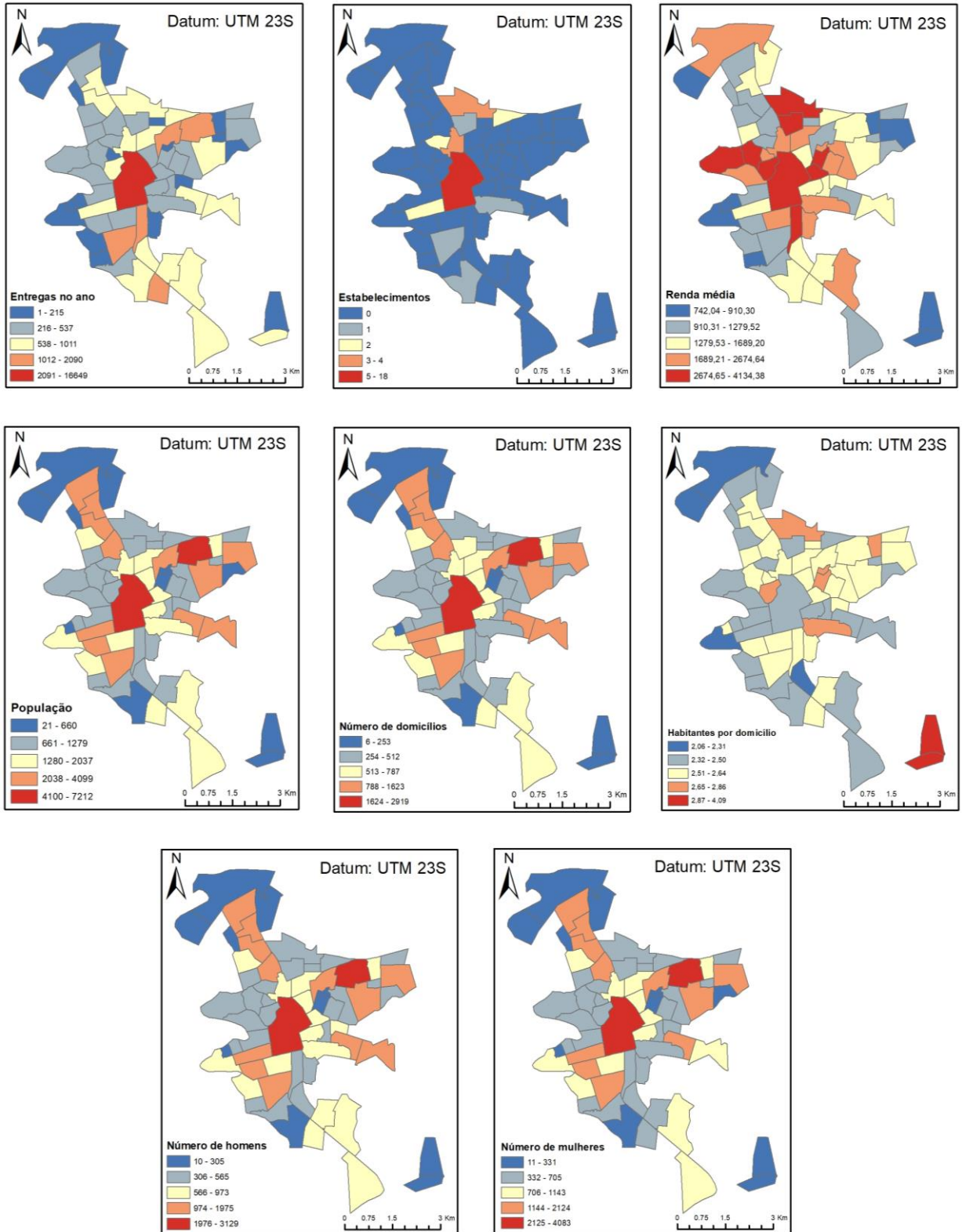


Figura 9: Distribuição da população de Patos de Minas por cor ou raça (IBGE, 2010).

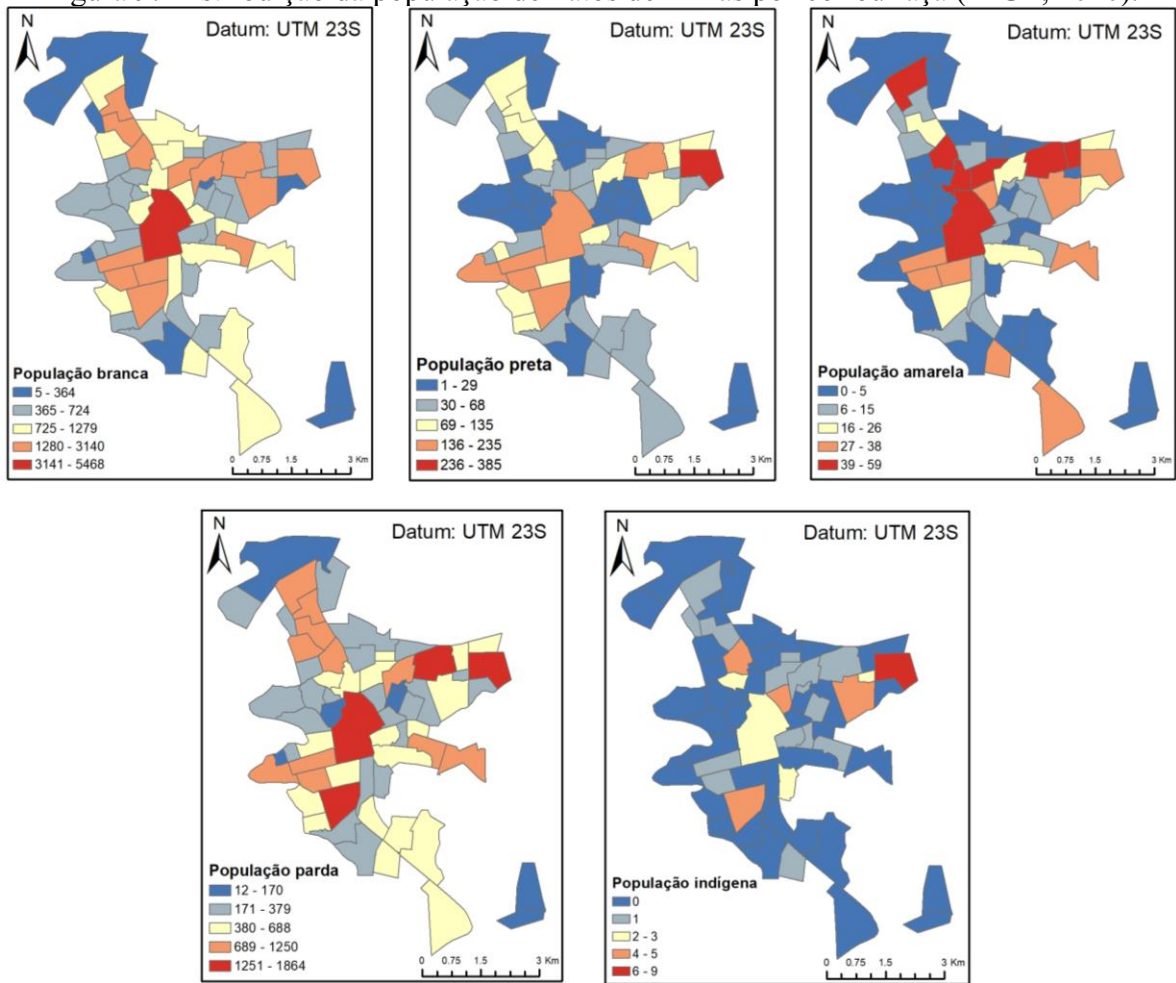
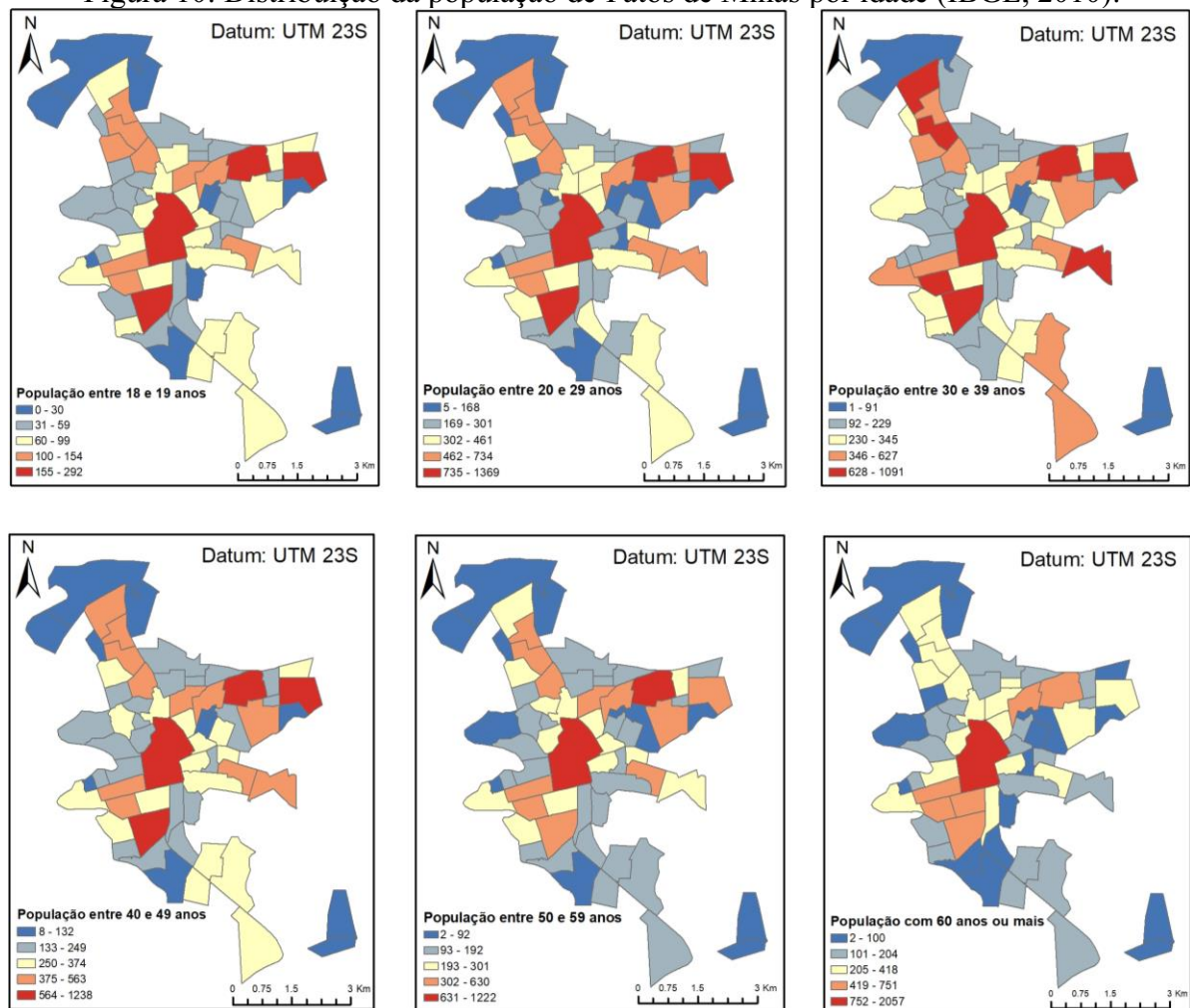


Figura 10: Distribuição da população de Patos de Minas por idade (IBGE, 2010).



Na Tabela 9 se encontra a estatística descritiva dos dados socioeconômicos, do número de entregas e do número de estabelecimentos comerciais em Patos de Minas.

Tabela 9: Estatística descritiva dos dados de Patos de Minas

Variável	Estatística descritiva					
	Mínimo	1º quartil	Mediana	Média	3º quartil	Máximo
Número de entregas	1,0	199,0	401,5	726,4	664,2	16649,0
População	21,0	830,8	1193,5	1527,6	1956,0	7212,0
Renda média	742	1082	1365	1694	1996	4134
Número de homens	10,0	401,5	563,5	725,8	905,5	3129,0
Número de mulheres	11,0	427,5	633,0	801,8	1001,0	4083,0
Número de pessoas brancas	5,0	481,5	705,5	908,9	1173,2	5468,0
Número de pessoas pretas	1,00	21,75	50,50	70,70	89,00	385,00
Número de pessoas amarelas	0,00	2,00	7,50	14,47	26,75	59,00
Número de pessoas pardas	12,0	237,0	414,0	532,4	684,2	1864,0
Número de pessoas indígenas	0,0000	0,0000	0,0000	0,8125	10,000	90,000
Número de pessoas entre 18 e 19 anos	0,00	37,50	62,50	75,72	90,00	292,00
Número de pessoas entre 20 e 29 anos	5,0	180,0	280,5	357,1	458,8	1369,0
Número de pessoas entre 30 e 39 anos	1,0	176,5	260,5	320,6	404,8	1091,0

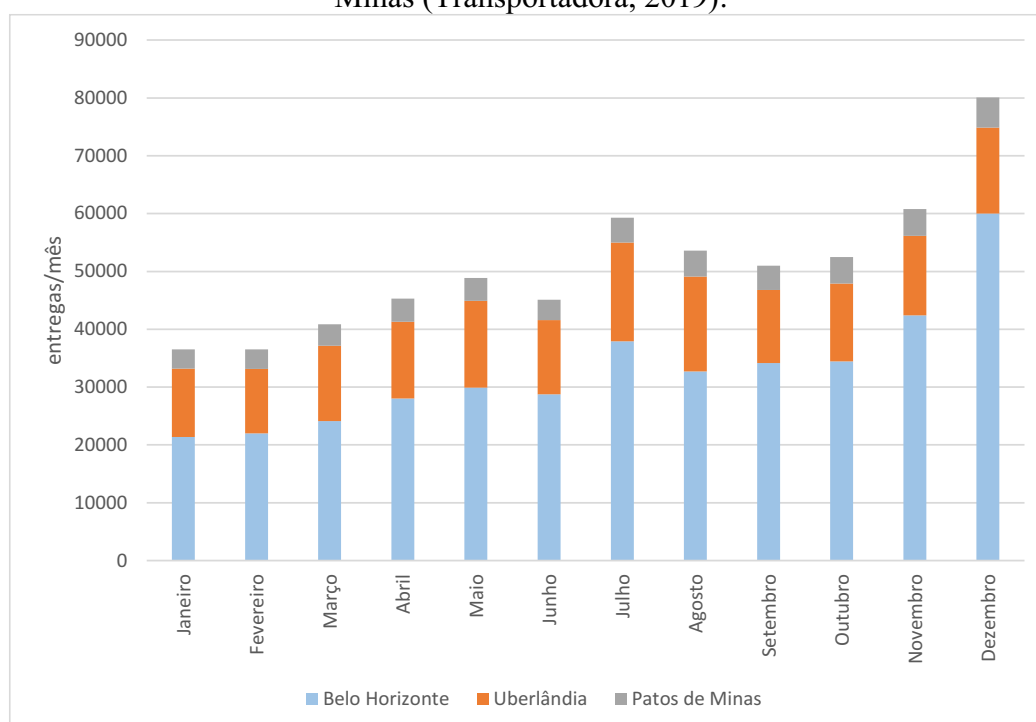


Número de pessoas entre 40 e 49 anos	8,0	183,0	255,0	305,6	373,2	1238,0
Número de pessoas entre 50 e 59 anos	2,0	121,0	170,5	228,9	264,8	1222,0
Número de pessoas com 60 anos ou mais	2,0	95,0	159,0	239,6	294,8	2057,0
Número de domicílios	6,0	327,2	477,5	607,8	775,8	2919,0
Habitantes por domicílio	2,06	2,45	2,53	2,55	2,58	4,09
Estabelecimentos comerciais	0,00	0,00	0,00	0,55	0,00	180,00

Os dados de oportunidade de acesso ao varejo tradicional, isto é, número de estabelecimentos comerciais por bairro foram obtidos de duas formas. Para Belo Horizonte, foram obtidos por meio do BHMap (2020) da Prefeitura de Belo Horizonte. Para Uberlândia e Patos de Minas utilizou-se o *plugin* QuickOSM do QGIS para extrair os dados do OpenStreetMap.

Por fim, o número de entregas do comércio eletrônico por bairro para cada cidade foi obtido de uma transportadora com forte atuação nas três cidades analisadas. Esses dados são referentes a cada mês de 2019. A Figura 11 mostra o número de entregas realizadas no decorrer do ano de 2019 para cada cidade.

Figura 11: Número de entregas realizadas em 2019 em Belo Horizonte, Uberlândia e Patos de Minas (Transportadora, 2019).



O método de análise escolhido para a verificação das hipóteses de pesquisa é detalhado na seção seguinte.

## 4 MÉTODO DE PESQUISA

O método de pesquisa deste estudo consiste de três etapas: (i) obtenção dos dados, (ii) análise econométrica e (iii) análise espacial, conforme apresentado na Figura 12. As etapas são detalhadas nas subseções a seguir.

Figura 12: Fluxograma das etapas do método proposto.



### 4.1 Dados

A Tabela 10 apresenta um resumo dos dados utilizados neste estudo, cujo detalhamento foi apresentado na seção anterior. Ressalta-se que a escolha das cidades de Belo Horizonte, Uberlândia e Patos de Minas se deu devido aos objetivos de pesquisa deste estudo, que incluem a comparação da influência dos fatores socioeconômicos e de acesso ao varejo tradicional em áreas urbanas de diferentes portes; além da disponibilidade dos dados de entregas do comércio eletrônico para essas cidades cedidos pela transportadora.

Tabela 10: Resumo dos dados utilizados

Variável		Fonte do dado
Entrega	Entrega/bairro	Transportadora (2019)
Socioeconômico	População	IBGE (2010)
	Renda	
	Gênero	
	Cor ou raça	
	Idade	
	Domicílios	
Oportunidade de compra	Estabelecimentos comerciais/bairro	BHMap (2020) OpenStreetMap (2021)

## 4.2 Análise econométrica

Para verificar a influência dos fatores socioeconômicos e do acesso a oportunidades de compra no número de entregas do comércio eletrônico, realizou-se uma análise econométrica para cada uma das três cidades selecionadas para o estudo. O método escolhido para tanto foi a regressão de Poisson e a regressão binomial negativa. A estimação dos modelos foi realizada no ambiente R.

A regressão de Poisson e a regressão binomial negativa são modelos lineares generalizados, sendo os métodos mais tradicionais de análise de variáveis dependentes que são dados de contagem, ou seja, valores inteiros não-negativos (Washington *et al.*, 2010), como é o caso do número de entregas do comércio eletrônico. Na distribuição de Poisson, a média e a variância são iguais. Quando a variância é significativamente maior que a média, os dados apresentam superdispersão. Nestes casos, é recomendada, então, a modelagem por meio da regressão binomial negativa (Washington *et al.*, 2010).

Em um modelo de regressão de Poisson, a probabilidade de em uma observação  $i$  ocorrerem  $y_i$  eventos, sendo  $y_i$  um inteiro não-negativo, é dada pela equação (1) (Washington *et al.*, 2010), em que  $P(y_i)$  representa a probabilidade de ocorrência de  $y_i$  eventos e  $\lambda_i$  é o parâmetro de Poisson para o período  $i$ . Este parâmetro é igual ao número esperado de eventos por período  $i$ ,  $E[y_i]$ .

$$P(y_i) = \frac{EXP(-\lambda_i) - \lambda_i^{y_i}}{y_i!} \quad (1)$$

Um modelo de regressão de Poisson é estimado por meio da especificação do parâmetro de Poisson como função das variáveis independentes, como por exemplo, no modelo log-linear apresentado na equação (2) (Washington *et al.*, 2010), em que  $X_i$  é o vetor de variáveis independentes e  $\beta$  é o vetor dos parâmetros estimados. O modelo  $E[y_i] = \lambda_i$  é, então, estimado por modelos de máxima verossimilhança (Washington *et al.*, 2010).



$$\lambda_i = EXP(\beta X_i) \text{ ou } LN(\lambda_i) = \beta X_i \quad (2)$$

No ambiente R, modelos de regressão de Poisson podem ser estimados utilizando a função `glm()`. Para tanto, os argumentos da função são definidos como `glm(y ~ x, family = "poisson")`, sendo  $y$  a variável dependente e  $x$  as variáveis independentes; `family = "poisson"` especifica a distribuição de Poisson e a função de ligação como a função logaritmo natural.

Após a estimativa dos modelos de regressão de Poisson, deve-se verificar se há superdispersão. No ambiente R, isto pode ser feito por meio da função `dispersiontest` do pacote ERA (Kleiber e Zeileis, 2020). Este teste possui como hipótese nula que o modelo estimado é equidisperso (ou seja, segue a distribuição de Poisson, com a média igual à variância) e como hipótese alternativa que o modelo possui superdispersão. Assim, caso a hipótese nula seja rejeitada, deve-se descartar o modelo de regressão de Poisson e utilizar a regressão binomial negativa.

A regressão binomial negativa é desenvolvida a partir da modificação da equação (2) para que para cada observação  $i$ , tenha-se a equação (3) (Washington *et al.*, 2010), em que  $EXP(\varepsilon_i)$  é um termo de erro de distribuição gama com média 1 e variância  $\alpha^2$ .

$$\lambda_i = EXP(\beta X_i + \varepsilon_i) \quad (3)$$

Assim, nestes modelos, a variância difere da média conforme a equação (4) (Washington *et al.*, 2010), em que  $\alpha$  é o parâmetro de superdispersão, pois é o que define a diferença entre o modelo de regressão de Poisson (quando  $\alpha$  se aproxima de zero) e o modelo de regressão binomial negativa (Washington *et al.*, 2010).

$$VAR[y_i] = E[y_i][1 + \alpha E[y_i]] = E[y_i] + \alpha E[y_i]^2 \quad (4)$$

A distribuição binomial negativa segue a forma apresentada na equação (5), sendo  $\Gamma(\cdot)$  uma função gama (Washington *et al.*, 2010).

$$P(y_i) = \frac{\Gamma((1/\alpha)+y_i)}{\Gamma(1/\alpha)y_i!} \left(\frac{1/\alpha}{(1/\alpha)+\lambda_i}\right)^{1/\alpha} \left(\frac{\lambda_i}{(1/\alpha)+\lambda_i}\right)^{y_i} \quad (5)$$

Um modelo de regressão binomial negativa pode ser estimado no ambiente R por meio da função `glm.nb()` do pacote MASS (Ripley *et al.*, 2021). Para tanto, os argumentos da função

são definidos  $glm.nb(y \sim x, link = \log)$ , sendo  $y$  a variável dependente,  $x$  as variáveis independentes e  $link = \log$  a função de ligação (função logaritmo natural).

Neste trabalho, definiu-se a variável dependente como o número de entregas do comércio eletrônico e as variáveis independentes como os fatores associados à renda, gênero, idade, cor ou raça, número de habitantes por domicílio e número de estabelecimentos comerciais. A Tabela 11 apresenta um resumo das variáveis.

Tabela 11: Resumo das variáveis utilizadas

<b>Variável</b>		<b>Tipo de variável</b>
Número de entregas do comércio eletrônico		Dependente
Gênero	Número de homens	Independente
	Número de mulheres	Independente
Renda média		Independente
Habitantes por domicílio		Independente
Cor ou raça	Número de pessoas brancas	Independente
	Número de pessoas pretas	Independente
	Número de pessoas amarelas	Independente
	Número de pessoas pardas	Independente
	Número de pessoas indígenas	Independente
Idade	Número de pessoas entre 18 e 19 anos	Independente
	Número de pessoas entre 20 e 29 anos	Independente
	Número de pessoas entre 30 e 39 anos	Independente
	Número de pessoas entre 40 e 49 anos	Independente
	Número de pessoas entre 50 e 59 anos	Independente
	Número de pessoas com 60 anos ou mais	Independente
Número de estabelecimentos comerciais		Independente

Inicialmente, estimou-se os modelos de regressão descritos acima para o conjunto completo dos dados (isto é, todos os bairros) das cidades de Belo Horizonte, Uberlândia e Patos de Minas. No entanto, não foi possível obter os coeficientes estimados para os modelos de regressão para Belo Horizonte e Uberlândia, pois a rotina apresentou erro. Apesar de investigar a causa do erro, não foi possível identificar um motivo claro para a ocorrência do erro.

Desta forma, baseado no Princípio de Pareto, que afirma que 80% dos resultados são consequência de 20% das causas (Pareto *et al.*, 1964), identificou-se os bairros responsáveis por 80% das entregas em cada cidade (incluindo Patos de Minas, a título de comparação com Belo Horizonte e Uberlândia). Assim, realizou-se a análise econométrica para os bairros representantes de 80% das entregas para as três cidades, e a análise econométrica com todos os bairros para Patos de Minas.

Dessa forma, foram estimados modelos de Poisson contendo todas as variáveis, em seguida, realizado o teste de dispersão para verificar se os dados apresentavam superdispersão, o que

determinou se a regressão binomial negativa substituiria a regressão Poisson. Em seguida, foi necessário determinar o melhor modelo para cada cidade. Para tanto, foi utilizado o método *stepwise* para determinação das variáveis a serem incluída nos modelos com base no Critério de Informação de Akaike (AIC). Como o número de amostras é maior do que o número de variáveis, utilizou-se *backward stepwise*, cujo algoritmo começa com o modelo completo, ou seja, com todas as variáveis, e elimina sequencialmente variáveis a fim de minimizar o valor do AIC (Hastie *et al.*, 2009). No ambiente R, o *stepwise* pode ser realizado por meio da função `stepAIC()`, com a especificação `stepAIC(object, direction = "backward")`. O componente `object` representa um modelo de regressão previamente estabelecido no qual o algoritmo do *stepwise* será aplicado, sendo utilizando o modelo contendo todas as variáveis; `direction = "backward"` é a direção que o algoritmo trabalhará para selecionar as variáveis do modelo final, de menor AIC.

A significância estatística das variáveis independentes para cada modelo foi avaliada por meio do teste Wald. Em modelos de regressão, é comum que haja multicolinearidade entre variáveis independentes, isto é, que as variáveis independentes apresentem correlação entre si ou com variáveis omitidas do modelo, mas que são relacionadas à variável dependente. Para tanto, foi estimado o *Variance Inflation Factor* (VIF), que mede a multicolinearidade em um conjunto de variáveis regressoras. Um valor de VIF menor que 10 é desejável (James *et al.*, 2015). No ambiente R, o VIF pode ser calculado por meio da função `vif()` do pacote `car` (Fox e Weisberg, 2019)

Assim, por meio do método descrito, foi possível identificar os fatores que influenciam as entregas do comércio eletrônico em Belo Horizonte, Uberlândia e Patos de Minas e comparar as diferenças e semelhanças dos resultados para cada cidade.

### **4.3 Análise espacial**

#### **4.3.1 Verificação da existência de dependência espacial das entregas do comércio eletrônico**

Antes de realizar a modelagem espacial, deve-se verificar se o número de entregas por bairro (isto é, a variável resposta) apresenta autocorrelação espacial. A autocorrelação espacial é definida como o grau de dependência espacial entre um valor de uma observação de determinada variável e as observações vizinhas (Grekousis, 2020). Portanto, a determinação da autocorrelação espacial é utilizada para verificar se existem relações entre os valores de uma

variável no espaço e se esses valores formam padrões espaciais (Grekousis, 2020). A correlação espacial positiva ocorre quando valores similares, sejam eles altos ou baixos, de determinada variável dentro de uma determinada distância tendem a estarem próximos; enquanto a correlação espacial negativa se dá quando localidades vizinhas tendem a ter valores diferentes. Por sua vez, a ausência de autocorrelação espacial indica que os dados são distribuídos de forma aleatória no espaço (Grekousis, 2020).

No caso da presença de dependência espacial nas observações de determinada variável, é importante que o fenômeno seja avaliado por meio de técnicas de estatística espacial, uma vez que a presença da autocorrelação espacial pode afetar os resultados de modelos estatísticos não-espaciais, diminuindo a acurácia de seus resultados (Grekousis, 2020). Neste trabalho, foram estimados modelos de regressão geograficamente ponderada para avaliar a relação entre as entregas do comércio eletrônico e as variáveis socioeconômicas e de acesso ao varejo tradicional quando os dados de entrega apresentaram autocorrelação espacial. No caso de não ocorrência de autocorrelação espacial, o fenômeno não é explicado espacialmente, e o modelo de regressão espacial não pode ser estimado.

Para realizar esta verificação, o Índice Global de Moran univariado (I) foi utilizado para avaliar a dependência espacial das entregas do comércio eletrônico em Belo Horizonte, Uberlândia e Patos de Minas. Este índice foi escolhido por ser o mais utilizado em se tratando de dados com variação espacial (Grekousis, 2020). O Índice Global de Moran univariado (I) computa a autocorrelação espacial levando em consideração os valores locais dos fatores e seus atributos locais simultaneamente (Moran, 1950). O índice I obtém a dependência espacial a partir da relação entre a média, o desvio da média e da variância dos dados conforme a equação (6), em que  $n$  é o número de elementos espaciais;  $x_i$  é o valor do elemento  $i$ ;  $x_j$  é o valor do elemento  $j$ ;  $\bar{x}$  é o valor médio dos elementos;  $w_{ij}$  é o peso espacial entre os elementos  $i$  e  $j$ ;  $\sum_i^n \sum_j^n w_{ij}$  é a agregação de todos os pesos espaciais (Grekousis, 2020).

$$I = \frac{n}{\sum_i^n \sum_j^n w_{ij}} \frac{\sum_i^n \sum_j^n w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_i^n (x_i - \bar{x})^2} \quad (6)$$

Além disso, o Índice Global de Moran univariado (I) deve ser interpretado com base em seu valor esperado  $E(I)$ , conforme a Equação (7), e tendo como hipótese nula a não existência de autocorrelação espacial (ou seja, os dados se distribuem de forma aleatória no espaço) (Grekousis, 2020).

$$E(I) = \frac{-1}{n-1} \quad (7)$$

O valor esperado  $E(I)$  representa o valor que o índice teria caso os dados apresentassem aleatoriedade espacial total (Grekousis, 2020). Uma vez que o p-valor é inferior a 0,05, descarta-se a hipótese nula e compara-se o valor observado do índice I ao módulo do seu valor esperado. Caso I seja superior, em módulo, ao valor esperado, existe dependência espacial (Grekousis, 2020).

Para este trabalho, o índice foi calculado no ambiente R, utilizando o pacote `spdep` (Bivand *et al.*, 2013). Inicialmente, utiliza-se a função `poly2nb()` para construir uma lista de vizinhos a partir de polígonos com bordas contíguas e com base no conceito da contiguidade da rainha (vértices são considerados contíguos). Em seguida, utiliza-se a função `nb2listw()` para suplementar a lista de vizinhos com pesos espaciais considerando-se uma matriz espacial do tipo W (padronizada por linha para criar pesos proporcionais caso os elementos tenham um número diferente de vizinhos), recomendada para o caso de polígonos. Além disso, utiliza-se o comando `zero.policy = TRUE` para que regiões da lista de polígonos que não tenham vizinhos recebam um vetor nulo de peso. Por fim, utiliza-se a função `moran.test` para obter o Índice Global de Moran univariado (I), seu valor esperado e o p-valor, considerando-se o número de entregas por bairro. Assim, considerando-se uma lista de polígonos `pl`, tem-se:

```
nb <- poly2nb(pl, queen=TRUE)
lw_W <- nb2listw(neighbours = nb, style = "W", zero.policy =
TRUE)
spdep::moran.test(x = pl$Entregas, listw = lw_W, zero.policy =
TRUE)
```

### 4.3.2 Regressão Geograficamente Ponderada

Para analisar a correlação espacial entre as entregas do comércio eletrônico e os fatores socioeconômicos e a oportunidade de compra foi utilizada a Regressão Geograficamente Ponderada (do inglês, *Geographically Weighted Regression* - GWR). Este método é vantajoso, pois é baseado na estrutura de uma regressão tradicional, além de ser intuitivo em seus procedimentos (Fotheringham *et al.*, 2002).

Uma regressão tradicional é uma análise global, pois trata-se de um modelo em que o mesmo peso é atribuído aos dados analisados em toda a área de estudo, mesmo quando esta é

desagregada em sub-regiões. Uma análise global pode ocultar dos resultados diferenciações locais importantes nas relações entre as variáveis independentes e a variável dependente. Na análise local, as relações entre as variáveis são avaliadas em relação a sua localização na área de estudo, ou seja, se essa localização muda, os resultados podem mudar. Assim, métodos de análise local, como o GWR, podem ser usados para confirmar ou expandir os resultados de uma análise global, verificando se as relações identificadas globalmente se confirmam ou se diferenciam no espaço (Fotheringham *et al.*, 2002).

O modelo geral definido pelo GWR é uma expansão da equação de uma regressão linear múltipla para a estimação de parâmetros locais e não globais, conforme a equação (8) (Fotheringham *et al.*, 2002).

$$Y_i = \beta_{0(u_i, v_i)} + \sum_k \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \quad (8)$$

Nesta equação,  $\beta_k$  representa os parâmetros a serem estimados e  $(u_i, v_i)$  são as coordenadas espaciais do ponto  $i$ , permitindo, assim, que haja uma consideração da variabilidade espacial dos dados e, conseqüentemente, da variação espacial das relações entre as variáveis (Fotheringham *et al.*, 2002).

No GWR são definidos pontos de regressão e uma região ao redor desses pontos contendo dados. Os pontos de regressão são as subunidades espaciais onde os dados são contidos, por exemplo, no caso deste trabalho, os polígonos representando cada bairro. O modelo é aplicado para cada ponto de regressão, e são atribuídos pesos diferentes aos dados de acordo com sua distância do ponto de regressão, isto é, dados próximos ao ponto de regressão têm pesos maiores e dados afastados dos pontos de regressão têm pesos menores. Para tanto são aplicadas funções de ponderação, também chamadas de funções kernel espaciais, para a definição dos pesos e funções para estabelecimento da região ótima (banda da função kernel) ao redor dos pontos de regressão (Fotheringham *et al.*, 2002). A Equação (9) demonstra o conceito da variação da ponderação dos dados, em que os parâmetros em negrito são matrizes;  $\hat{\beta}$  representa uma estimativa de  $\beta$ ;  $\mathbf{W}_{(u_i, v_i)}$  é uma matriz  $n \times n$  cujos elementos da diagonal são os pesos para cada observação  $n$  para o ponto de regressão  $i$  e cujos elementos fora da diagonal são zero (Fotheringham *et al.*, 2002).

$$\hat{\beta}_{(u_i, v_i)} = (\mathbf{X}^T \mathbf{W}_{(u_i, v_i)} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}_{(u_i, v_i)} \mathbf{y} \quad (8)$$

Neste trabalho, a função de ponderação  $w_{ij}$  foi estabelecida como uma função Gaussiana conforme a Equação (9), em que  $d_{ij}$  representa a distância Euclidiana entre o ponto de regressão  $i$  e um ponto específico  $j$  no espaço para onde os parâmetros estão sendo estimados, e  $b$  é a banda (Fotheringham *et al.*, 2002).

$$w_{ij} = \exp \left[ -\frac{1}{2} \left( \frac{d_{ij}}{b} \right)^2 \right] \quad (9)$$

A função Gaussiana permite uma ponderação contínua, evitando problemas de descontinuidade nos pesos. Se  $i$  e  $j$  forem coincidentes, o peso atribuído aos dados naquele local será uma unidade, e à medida em que a distância entre  $i$  e  $j$  aumenta, o peso atribuído aos dados diminuirá de acordo com a curva Gaussiana; quando  $i$  e  $j$  forem suficientemente distantes, o peso será virtualmente nulo, retirando aquelas observações da estimativa de parâmetros para a localização  $i$  (Fotheringham *et al.*, 2002).

A banda  $b$  foi selecionada neste trabalho utilizando o método de validação cruzada (do inglês, *Cross-validation* - CV). Na técnica CV, a banda ótima é aquela que minimiza a função descrita na Equação (10), em que  $n$  é o número de observações;  $\widehat{y}_{\neq i}(b)$  é o valor previsto para  $y_i$  sendo que a previsão da  $i$ -ésima observação é obtida com o peso daquela observação como zero (Fotheringham *et al.*, 2002).

$$CV = \sum_{i=1}^n [y_i - \widehat{y}_{\neq i}(b)]^2 \quad (10)$$

O cálculo do CV só é possível quando as localizações dos pontos de regressão são as mesmas das localizações dos dados observados (Fotheringham *et al.*, 2002).

A estimativa dos parâmetros dos modelos GWR foi realizada no ambiente R por meio dos pacotes `spgwr` (Bivand *et al.*, 2013) e `GWmodel` (Gollini *et al.*, 2015). Foram estimados três tipos de modelo: modelos utilizando todas as variáveis, modelos utilizando as variáveis do melhor modelo econométrico e o melhor modelo GWR pelo critério de minimização do Critério de Informação de Akaike corrigido (AICc).

Para a seleção do melhor modelo GWR, utilizou-se a função `gwr.model.selection()`. O algoritmo de seleção funciona com as seguintes etapas. Primeiramente, calibra-se todos os possíveis modelos GWR com uma variável independente. Destes modelos, seleciona-se o modelo que possuir o menor AICc, e inclui-se a variável independente permanentemente no modelo. Em seguida, calibra-se todos os modelos possíveis contendo uma das variáveis

independentes restantes juntamente com a variável independente selecionada no passo anterior. Seleciona-se o modelo com menor AICc. Este processo é repetido até que todas as variáveis independentes sejam incluídas no modelo. O melhor modelo é aquele que possui o menor AICc após todas as variáveis independentes terem sido testadas (Lu *et al.*, 2014). Um exemplo do uso da função encontra-se a seguir:

```
gwr.model.selection(DeVar, InDeVars, data, bw, approach = "CV",  
kernel = "gaussian")
```

A variável `DeVar` representa a variável dependente do modelo, `InDeVars` representa um vetor de variáveis independentes a serem selecionadas para o modelo, `data` se refere ao *data frame* espacial onde os dados utilizados na análise estão salvos, `bw` é a banda previamente calculada por uma função específica, `approach = "CV"` se refere ao método de validação cruzada e `kernel = "gaussian"` é a função kernel gaussiana.

Em modelos GWR, a banda pode ser calculada por meio da função `gwr.sel()` conforme descrito a seguir:

```
gwr.sel(formula, data, gweight = gwr.Gauss, method = "cv")
```

A variável `formula` representa as variáveis dependente e independentes do modelo, `data` se refere ao *data frame* espacial onde os dados utilizados na análise estão salvos, `gweight = gwr.Gauss` é a função kernel gaussiana e `method = "cv"` se refere ao método de validação cruzada.

Definida a banda, ela pode ser inserida na função `gwr()` para estimativa dos parâmetros do modelo conforme exemplificado a seguir:

```
gwr(formula, data, bandwidth, gweight = gwr.Gauss, se.fit = T,  
hatmatrix = T)
```

A variável `formula` representa as variáveis dependente e independentes do modelo, `data` se refere ao *data frame* espacial onde os dados utilizados na análise estão salvos, `bandwidth` se refere à banda previamente calculada pela função `gwr.sel()`, `gweight = gwr.Gauss` é



a função kernel gaussiana, `se.fit = T` retorna o erro padrão dos coeficientes locais do modelo e `hatmatrix = T` retorna a *hat matrix* como componente nos resultados.

Para avaliar a significância estatística dos modelos, foi verificado se os resíduos de cada modelo apresentavam autocorrelação por meio do Índice Global de Moran univariado (I). A presença de autocorrelação na distribuição espacial dos resíduos é um indicativo de que o modelo pode apresentar problemas de subestimação dos erros, fazendo-o menos confiável (Fotheringham *et al.*, 2002). Portanto, é desejável que os resíduos dos modelos GWR apresentem uma distribuição aleatória. Para tal verificação, utilizou-se a função `gwr.morantest()` conforme exemplificado abaixo.

```
gwr.morantest(x, lw, zero.policy = TRUE)
```

O componente `x` representa um modelo GWR calculado pela função `gwr()`, `lw` é um objeto do tipo `listw` criado, por exemplo, pela função `nb2listw` do pacote `spdep` e `zero.police = TRUE` é usado para que regiões da lista de polígonos que não tenham vizinhos recebam um vetor nulo de peso.

Por fim, a multicolinearidade foi avaliada pelo VIF por meio da função `gwr.collin.diagno()`, seguindo a seguinte especificação:

```
gwr.collin.diagno(formula, data, bw, kernel = "gaussian")
```

A variável `formula` representa as variáveis dependente e independentes do modelo, `data` se refere ao *data frame* espacial onde os dados utilizados na análise estão salvos, `bw` se refere à banda previamente calculada e `kernel = "gaussian"` é a função kernel gaussiana.

Por meio do método descrito, foram estimados modelos GWR para as cidades cujos dados de entrega apresentaram autocorrelação espacial para verificar a influência espacial dos fatores socioeconômicos e de acesso ao varejo tradicional nas entregas do comércio eletrônico, bem como para estabelecer uma comparação entre os resultados do modelo econométrico e da análise espacial. Assim, foi possível verificar se as relações identificadas no modelo global se mantêm no espaço ou se há diferenciação dessas relações de acordo com a localidade dentro da área de estudo. Ressalta-se que, por se tratar de análise espacial, os modelos GWR foram aplicados nos dados completos, ou seja, todos os bairros, e não apenas os responsáveis por 80% das entregas.

Ainda, tentou-se aplicar modelos de regressão de Poisson geograficamente ponderada (nos quais a distribuição do erro segue uma distribuição de Poisson) aos dados, porém os modelos apresentaram erro e não puderam ser estimados. Foram feitas tentativas com dois pacotes, spgwr e GWmodel. Destaca-se também que a regressão binomial negativa geograficamente ponderada não existe nos pacotes GWR no ambiente R. Assim, optou-se pelo GWR tradicional, cuja distribuição do erro segue uma distribuição gaussiana.

## 5 RESULTADOS

### 5.1 Belo Horizonte

#### 5.1.1 Análise econométrica

Conforme descrito anteriormente, para a cidade de Belo Horizonte, foram utilizados na análise econométrica os dados correspondentes aos bairros responsáveis por 80% das entregas. Isto se deu, pois os modelos de regressão de Poisson e regressão binomial negativa apresentaram erros que não puderam ser corrigidos, inviabilizando sua estimativa para os dados completos. Ao todo, 142 bairros, o que corresponde a 30% dos bairros da cidade, foram responsáveis por 80% das entregas. A Figura 13 ilustra a representação cartográfica dos bairros analisados e dos removidos da análise e a Tabela 12 apresenta a estatística descritiva dos 142 bairros considerados na análise.

Figura 13: Representação cartográfica de bairros analisados e bairros retirados da análise econométrica de Belo Horizonte.

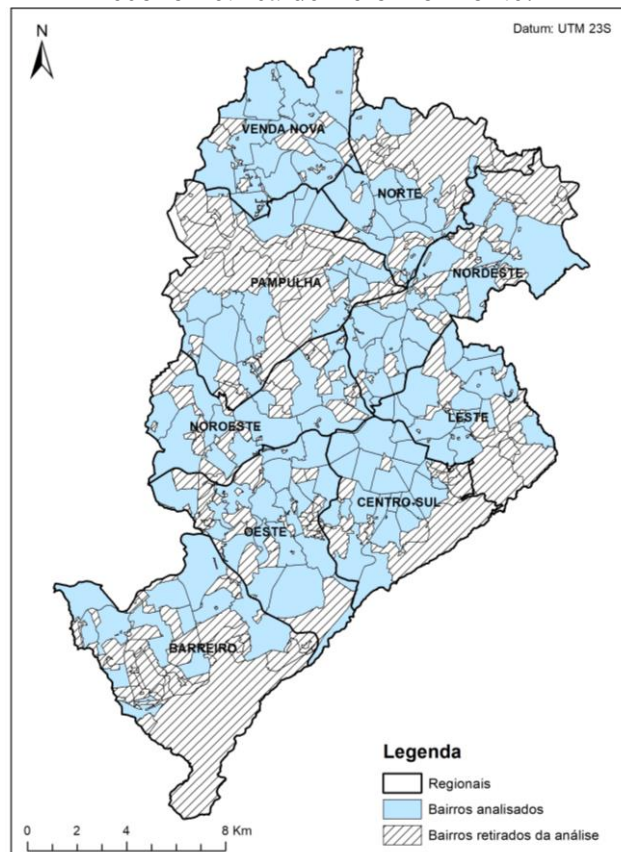


Tabela 12: Estatística descritiva dos 142 bairros analisados em Belo Horizonte

Variável	Estatística descritiva					
	Mínimo	1º quartil	Mediana	Média	3º quartil	Máximo
Número de entregas	1028	1310	1776	2219	2724	11638
População	1861	5692	8092	9326	12225	30218
Renda média	593,2	1134,7	1785,0	2470,0	3108,8	12598,3
Número de homens	842	2566	3682	4239	5602	13360
Número de mulheres	1020	3114	4394	5087	6617	16858
Número de pessoas brancas	559	2571	3972	4965	6019	19320
Número de pessoas pretas	74,0	341,2	733,0	832,7	1123,0	2415,0
Número de pessoas amarelas	13,00	51,00	77,50	94,74	123,75	327,00
Número de pessoas pardas	427	1818	3110	3419	4560	9821
Número de pessoas indígenas	0,00	4,00	10,50	13,94	19,00	74,00
Número de pessoas entre 18 e 19 anos	74,0	235,5	324,0	365,2	484,0	1056,0
Número de pessoas entre 20 e 29 anos	404	1332	1914	2238	2943	6519
Número de pessoas entre 30 e 39 anos	386	1170	1636	1983	2536	7752
Número de pessoas entre 40 e 49 anos	376	1061	1544	1711	2191	5543
Número de pessoas entre 50 e 59 anos	278	858	1236	1404	1813	5150
Número de pessoas com 60 anos ou mais	264,0	870,2	1359,5	1624,8	2072,5	6032,0
Número de domicílios	734	2306	3135	3709	4865	12207
Habitantes por domicílio	2,120	2,480	2,530	2,524	2,598	2,840
Estabelecimentos comerciais	23,0	106,0	167,5	243,1	246,0	4770,0

Inicialmente, foi estimado um modelo de regressão de Poisson contendo todas as variáveis independentes (renda média, número de homens, número de mulheres, número de pessoas brancas, número de pessoas pretas, número de pessoas amarelas, número de pessoas pardas, número de pessoas indígenas, número de pessoas entre 18 e 19 anos, número de pessoas entre 20 e 29 anos, número de pessoas entre 30 e 39 anos, número de pessoas entre 40 e 49 anos, número de pessoas entre 50 e 59 anos, número de pessoas com 60 anos ou mais, habitantes por domicílio e número de estabelecimentos comerciais). Em seguida, aplicou-se o teste de dispersão no modelo, que apresentou p-valor de 0,0007781 ( $< 0,05$ ), sendo detectado superdispersão no modelo.

Alternativamente, foi estimado um novo modelo, denominado Modelo BH1, utilizando a regressão binomial negativa. Os resultados do Modelo BH1 se encontram na Tabela 13 a seguir. Os valores do VIF menores que 10 foram marcados em verde, e os maiores em vermelho.

Tabela 13: Dados estimados para o Modelo BH1

Variável	Valor estimado	Teste Wald	Exponencial do valor estimado	VIF	AIC
Intercepto	6,979	9,909***	1073,844	-	2224,6
Número de estabelecimentos comerciais	0,000	3,457***	1,000	3,18	
Renda média	0,000	2,334*	1,000	3,55	
Habitantes por domicílio	-0,041	-0,149	0,960	1,84	
Número de homens	-0,002	-0,063	0,998	9351382,00	
Número de mulheres	-0,002	-0,068	0,998	14372030,00	
Número de pessoas entre 18 e 19 anos	0,026	1,098	1,027	39750,36	
Número de pessoas entre 20 e 29 anos	0,027	1,143	1,028	1509503,00	
Número de pessoas entre 30 e 39 anos	0,027	1,138	1,028	1379474,00	
Número de pessoas entre 40 e 49 anos	0,027	1,129	1,027	893837,20	
Número de pessoas entre 50 e 59 anos	0,027	1,136	1,028	685618,80	
Número de pessoas com 60 anos ou mais	0,027	1,139	1,028	1195810,00	
Número de pessoas brancas	-0,025	-0,921	0,975	19065200,00	
Número de pessoas pretas	-0,025	-0,916	0,975	493521,90	
Número de pessoas amarelas	-0,025	-0,906	0,976	5996,10	
Número de pessoas pardas	-0,025	-0,922	0,975	6032145,00	
Número de pessoas indígenas	-0,028	-1,019	0,972	282,93	

Nível de significância: “\*\*\*” 0,001; “\*\*” 0,01; “\*” 0,05; “.” 0,1

Verifica-se que apenas a contribuição do número de estabelecimentos comerciais e da renda é estatisticamente significativa. Ainda, essas mesmas variáveis, além de habitantes por domicílio, apresentam VIF inferior a 10. O AIC do modelo foi de 2224,6. De acordo com este modelo, o número de estabelecimentos comerciais e a renda tem influência positiva no número de entregas, isto é, um aumento em uma destas variáveis, mantendo todas as outras constantes, ocasiona um aumento do número de entregas.

No Modelo BH1, foi aplicado o *stepwise*, que gerou o Modelo BH2. Este modelo está descrito na Tabela 14.

Tabela 14: Dados estimados para o Modelo BH2

Variável	Valor estimado	Teste Wald	Exponencial do valor estimado	VIF	AIC
Intercepto	6,860	113,027***	953,367	-	2207,4
Número de estabelecimentos comerciais	0,000	4,905***	1,000	1,19	
Renda média	0,000	3,954***	1,000	1,37	
Número de pessoas entre 20 e 29 anos	0,000	5,304***	1,000	3,45	
Número de pessoas com 60 anos ou mais	0,000	1,439	1,000	3,82	

Nível de significância: “\*\*\*\*” 0,001; “\*\*\*” 0,01; “\*\*” 0,05; “.” 0,1

Assim como no Modelo BH1, no Modelo BH2, o número de estabelecimentos comerciais e a renda média têm contribuição estatisticamente significativa e positiva. Além dessas variáveis, o número de pessoas entre 20 e 29 anos também é estatisticamente significativo e apresenta influência positiva no número de entregas. Neste modelo, o VIF de todas as variáveis é menor que 10 e o AIC do modelo é 2207,04, o menor possível de acordo com o algoritmo do *stepwise*, e menor em 17,2 unidades em relação ao AIC do Modelo BH1.

### 5.1.2 Análise espacial

Inicialmente, verificou-se a autocorrelação espacial das entregas, utilizando os dados completos (todos os bairros da cidade). O cálculo do Índice Global de Moran univariado indicou dependência espacial das entregas do comércio eletrônico para o caso de Belo Horizonte, pois seu valor foi de 0,086, superior em módulo ao valor esperado (-0,002), e com p-valor de 0,003 (< 0,05). Desta maneira, torna-se fundamental o uso de técnicas espaciais para análise do fenômeno.

Assim, foram estimados três modelos de regressão geograficamente ponderada. Inicialmente, estimou-se o Modelo GWR1 contendo todas as variáveis independentes. Em seguida, estimou-se o Modelo GWR2, obtido por meio da função `gwr.model.selection()`, que conforme descrito na seção do Método, selecionou o modelo ótimo com base na minimização do AICc. Por fim, foi estimado o Modelo GWR3 contendo as variáveis independentes do modelo econométrico ótimo (Modelo BH2). As variáveis de cada modelo são sumarizadas a seguir.

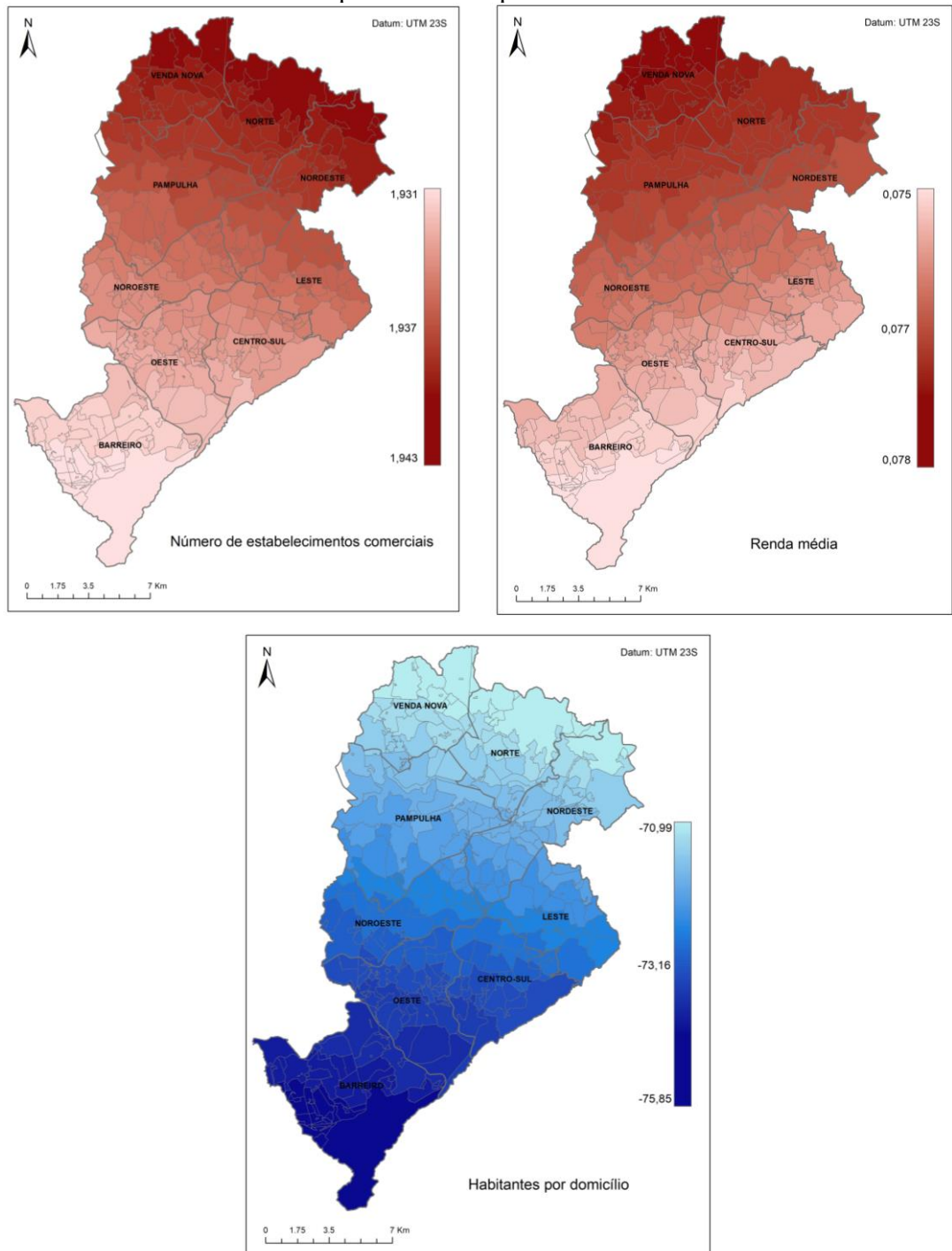
- **Modelo GWR1:** Número de entregas ~ Número de estabelecimentos comerciais + Renda média + Habitantes/domicílio + Número de homens + Número de mulheres + Número de pessoas entre 18 e 19 anos + Número de pessoas entre 20 e 29 anos + Número de pessoas entre 30 e 39 anos + Número de pessoas entre 40 e 49 anos + Número de pessoas entre 50 e 59 anos + Número de pessoas com 60 anos ou mais + Número de pessoas brancas + Número de pessoas pretas + Número de pessoas amarelas + Número de pessoas pardas + Número de pessoas indígenas
- **Modelo GWR2:** Número de entregas ~ Número de estabelecimentos comerciais + Renda média + Número de mulheres + Número de pessoas entre 18 e 19 anos + Número de pessoas entre 20 e 29 anos + Número de pessoas pretas + Número de pessoas amarelas
- **Modelo GWR3:** Número de entregas ~ Número de estabelecimentos comerciais + Renda média + Número de pessoas entre 20 e 29 anos + Número de pessoas com 60 anos ou mais

O Modelo GWR1 obteve um AIC de 7218,33 e um AICc de 7239.65. Os resíduos do modelo foram analisados por meio do cálculo do Índice Global de Moran univariado e apresentaram p-valor de 0,779 ( $> 0,05$ ), o que indica uma distribuição aleatória dos resíduos, validando o modelo.

Os resultados obtidos para este modelo se encontram nas representações cartográficas a seguir. As tonalidades escuras retratam os maiores valores estimados dos coeficientes, enquanto tonalidades mais claras retratam menores valores estimados dos coeficientes, em módulo para ambos os casos. Tons vermelhos retratam efeito espacial positivo, enquanto tons azuis retratam efeito espacial negativo. Em geral, os coeficientes estimados apresentaram pequena variabilidade espacial.

Os resultados do modelo se encontram na Figura 14, Figura 15, Figura 16 e Figura 17 a seguir.

Figura 14: Coeficientes espaciais das variáveis número de estabelecimentos comerciais, renda média e habitantes por domicílio para o Modelo GWR1.

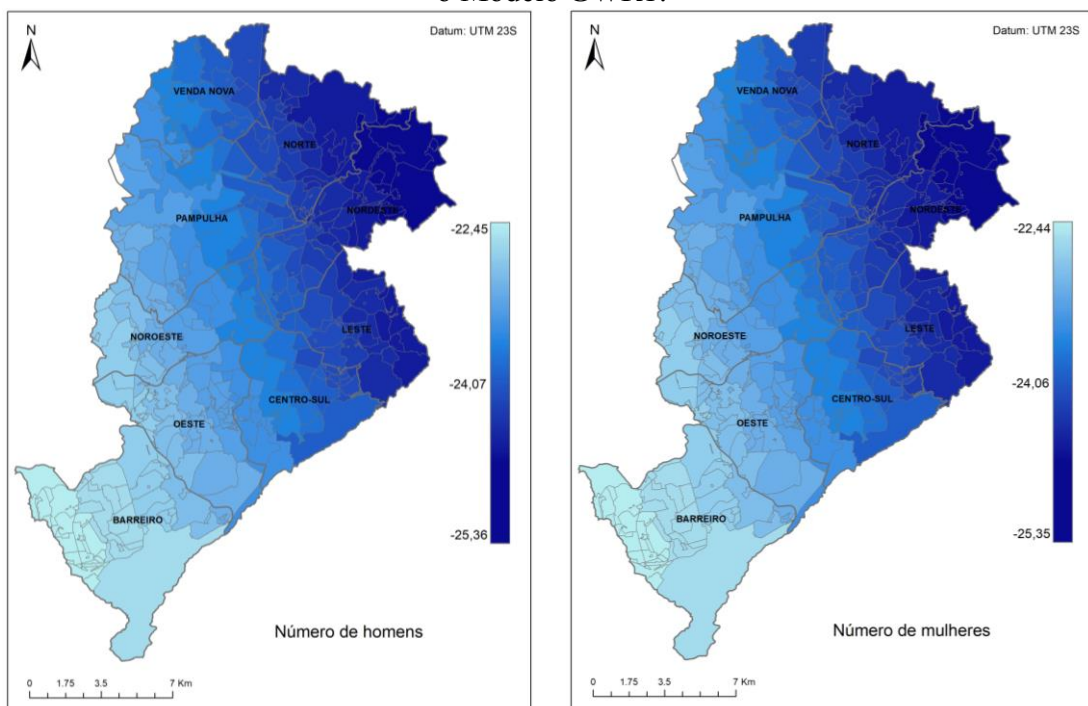


O número de estabelecimentos comerciais apresentou efeito positivo no número de entregas, indicando que em bairros onde há maior número de comércios tende-se a receber mais entregas do *e-commerce*. Esse efeito é maior nas regionais de Venda Nova, Norte e Nordeste da cidade, sendo esta primeira uma regional com um considerável número de comércios. A renda média também tem influência positiva nas entregas, o que está associado ao fato de bairros com maior renda possuírem maior poder de compra. No entanto, a influência dessa variável é maior nas



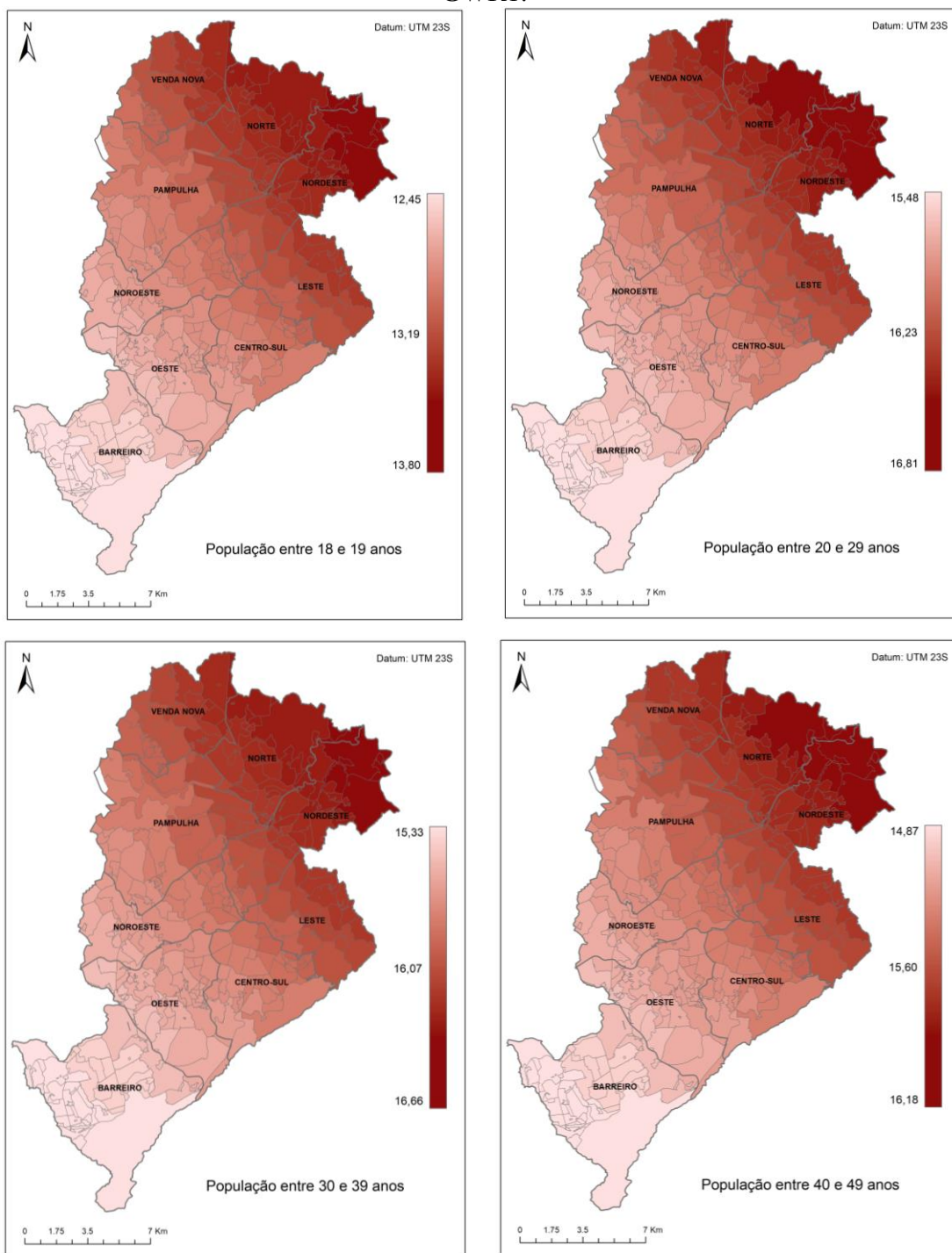
regionais de Venda Nova e Norte, regiões onde a renda média é menor, o que pode indicar que outras variáveis associadas à renda podem influenciar a distribuição espacial das entregas do comércio eletrônico na cidade. Em contrapartida, o número de habitantes por domicílio tem efeito negativo nas entregas; em domicílios com menor número de habitantes compra-se mais *online*. Este efeito é maior nas regionais Barreiro e Oeste. Apenas as variáveis número de estabelecimentos comerciais, renda média e habitantes por domicílio apresentaram VIF interior a 10.

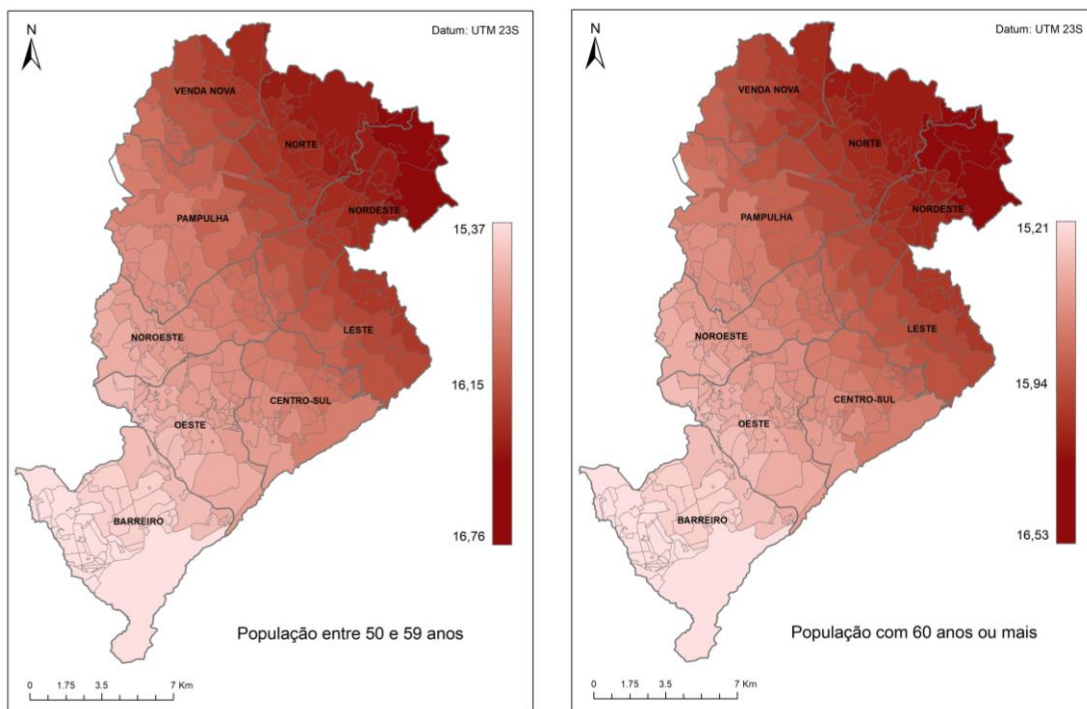
Figura 15: Coeficientes espaciais das variáveis número de homens e número de mulheres para o Modelo GWR1.



A proporção de homens e a proporção de mulheres apresentaram efeito negativo nas entregas do *e-commerce*. Os coeficientes para ambas as variáveis apresentaram valores próximos, com o coeficiente para o número de homens sendo ligeiramente superior ao das mulheres (em módulo), indicando que os homens tendem a receber menos entregas que as mulheres.

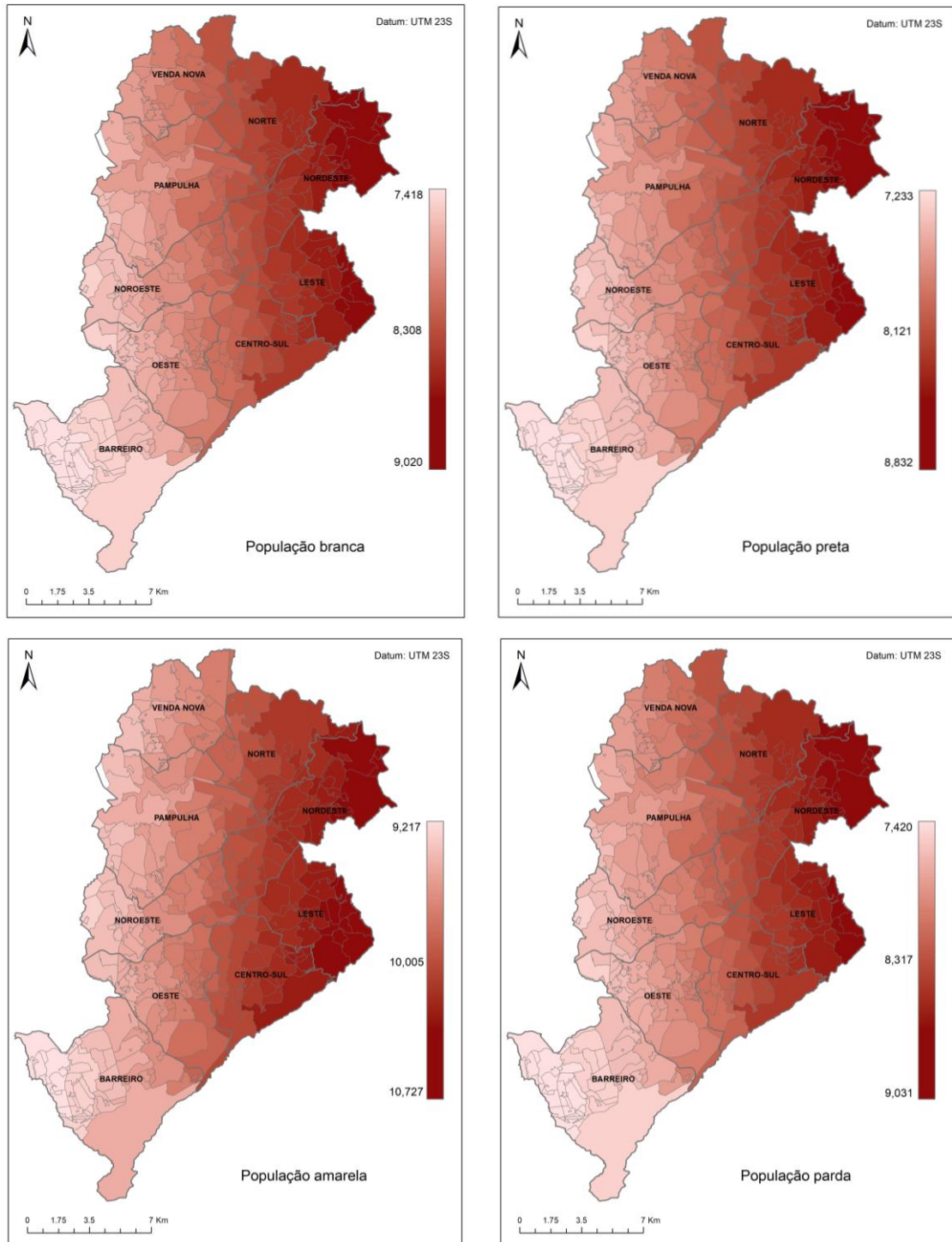
Figura 16: Coeficientes espaciais das variáveis relacionadas às faixas etárias para o Modelo GWRI.

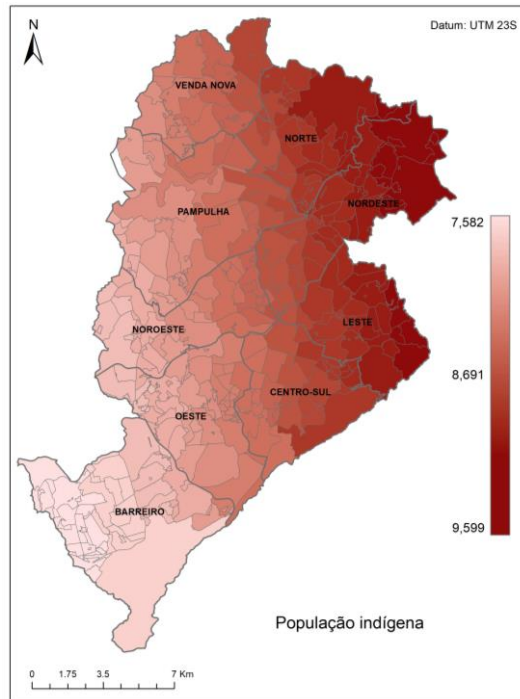




As variáveis referentes às faixas etárias apresentaram influência positiva no número de entregas. Os coeficientes estimados das variáveis tiveram maior valor nas regionais Norte e Nordeste da cidade, regiões onde não há uma grande concentração populacional em comparação com outras regionais. Os coeficientes estimados apresentam maior valor para a faixa etária dos 20 aos 29 anos, indicando que as pessoas dessa faixa etária recebem mais entregas que as outras faixas.

Figura 17: Coeficientes espaciais das variáveis relacionadas à cor ou raça para o Modelo GWR1.



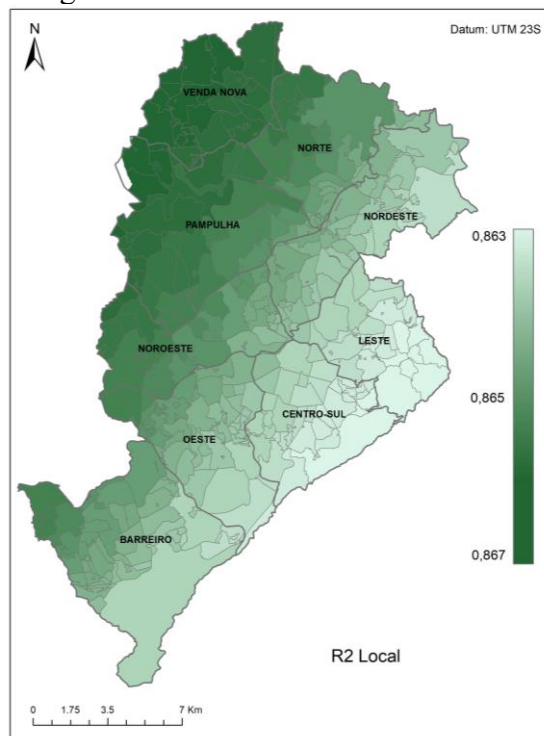


As variáveis relacionadas à cor ou raça tiveram influência positiva nas entregas, sendo essa influência maior nas regionais Nordeste e Leste. O coeficiente estimado para a variável população amarela apresentou o maior valor, indicando que esse grupo tende a receber mais entregas de compras *online*.

Conforme pode ser visto da Figura 18, o modelo apresentou valores satisfatórios do coeficiente de determinação espacial.



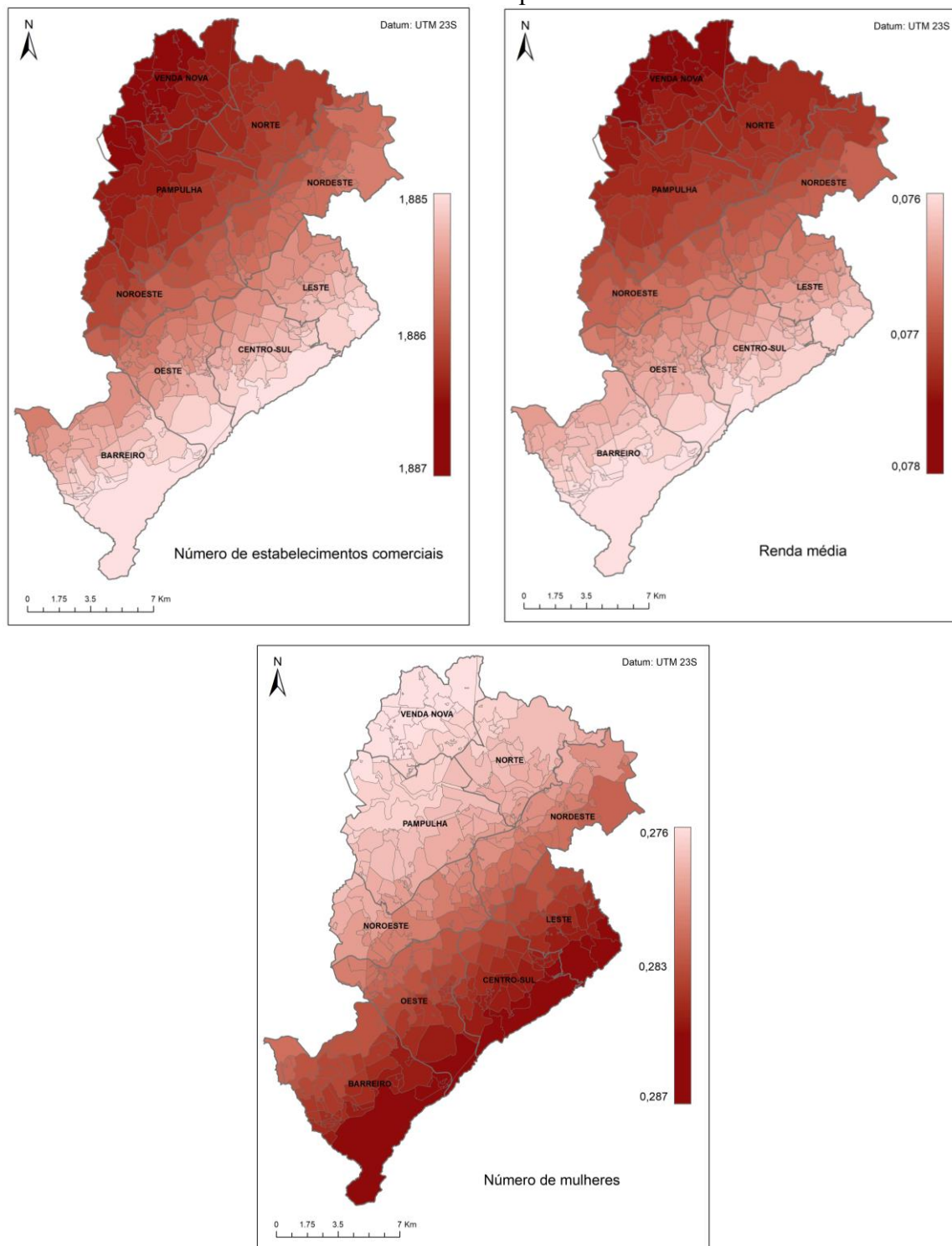
Figura 18:  $R^2$  local do Modelo GWR1.



O Modelo GWR2 obteve um AIC de 7212,775, menor que o do Modelo GWR1 em 5,556 unidades; e um AICc de 7223,523, menor que o Modelo GWR1 em 16,124 unidades. Os resíduos do modelo foram testados por meio do cálculo do Índice Global de Moran univariado, que resultou em um p-valor de 0,812 ( $> 0,05$ ), o que indica que a distribuição dos resíduos é aleatória e o modelo é válido.

Os resultados obtidos para o Modelo GWR2 se encontram na Figura 19, Figura 20 e Figura 21 a seguir. Em geral, os valores dos coeficientes apresentaram baixa variabilidade espacial, assim como no modelo anterior.

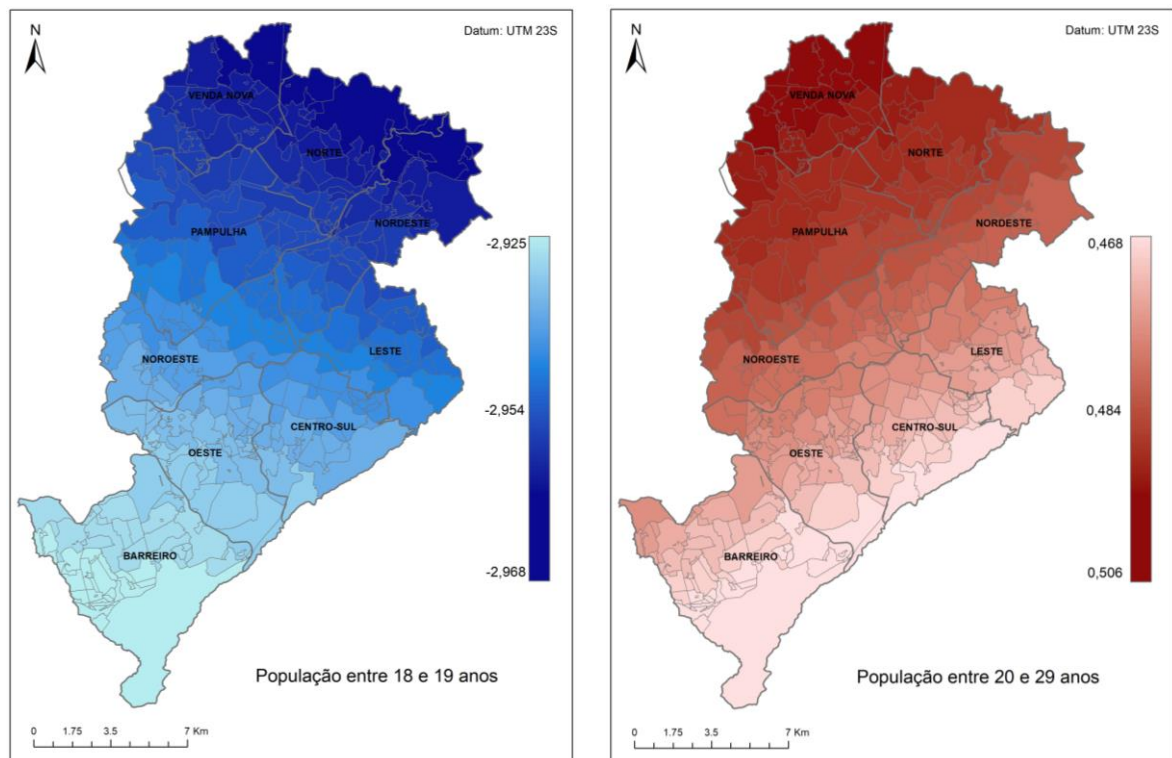
Figura 19: Coeficientes espaciais das variáveis número de estabelecimentos comerciais, renda média e número de mulheres para o Modelo GWR2.



Assim como no modelo anterior, o número de estabelecimentos comerciais teve influência positiva no número de entregas. No Modelo GWR2, no entanto, além da regional Venda Nova, a Pampulha é onde a influência dos estabelecimentos comerciais é maior. Além disso, neste modelo, o valor dos coeficientes para essa variável é menor que no anterior. A renda média, tal

qual no Modelo GWR1, apresentou influência positiva nas entregas e uma distribuição e valores dos coeficientes semelhantes aos daquele modelo. O número de estabelecimentos comerciais e a renda média apresentaram VIF menor que 10. A proporção de mulheres, ao contrário do modelo anterior, apresentou contribuição positiva no Modelo GWR2; sendo essa contribuição mais intensa nas regionais Leste, Centro-Sul, Oeste e Barreiro. O VIF para esta variável foi maior que 10; a presença de multicolinearidade pode ser uma possível explicação para a variação no sinal dos coeficientes estimados de um modelo para o outro para a mesma variável (Hair *et al.*, 2019).

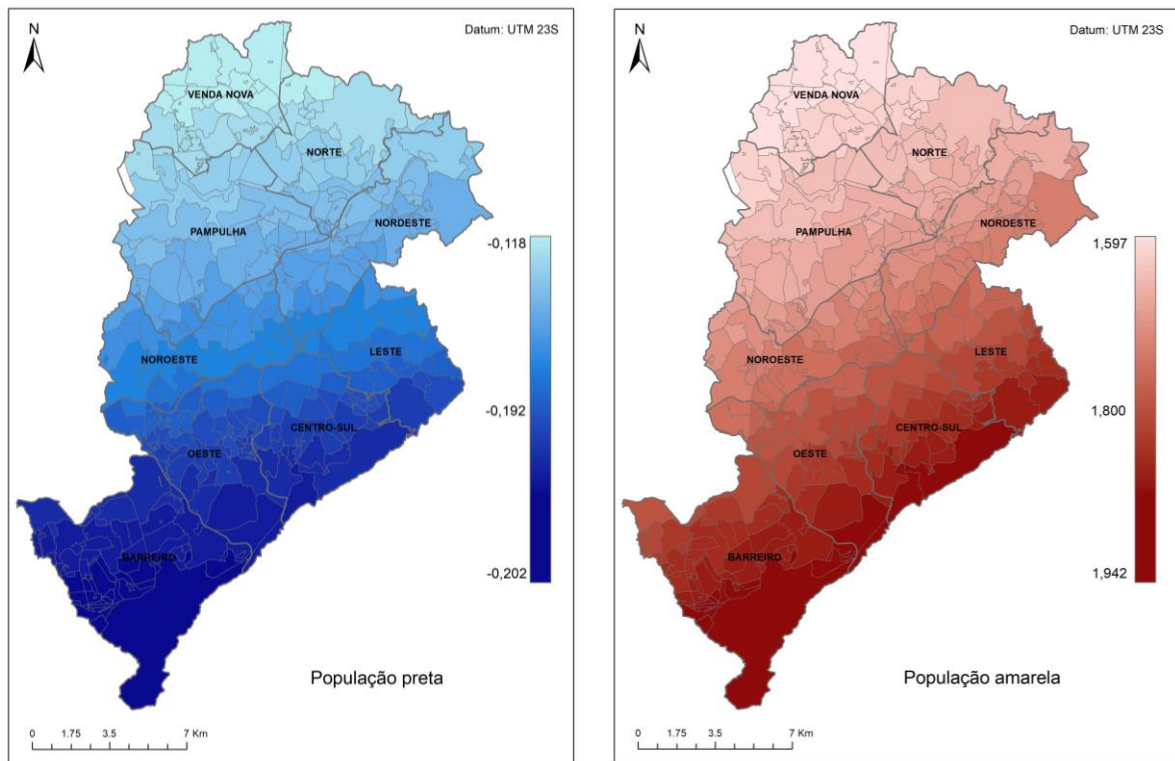
Figura 20: Coeficientes espaciais das variáveis relacionadas às faixas etárias para o Modelo GWR2.



Ao contrário do modelo anterior, a parcela da população entre 18 e 19 anos teve influência negativa no número de entregas do comércio eletrônico, sendo esta influência mais intensa nas regionais Venda Nova (onde há grande concentração de pessoas nessa faixa etária), Norte e Nordeste. A parcela da população entre 20 e 29 anos apresentou influência positiva no número de entregas, tal qual no modelo anterior. No entanto, diferentemente do Modelo GWR1, a influência dessa faixa etária foi maior nas regionais de Venda Nova, Pampulha e Norte. Além disso, o valor dos coeficientes para essa variável foi significativamente menor que no modelo anterior. Ambas as variáveis de faixa etária do Modelo GWR2 apresentaram VIF maior que 10.



Figura 21: Coeficientes espaciais das variáveis relacionadas à cor ou raça para o Modelo GWR2.

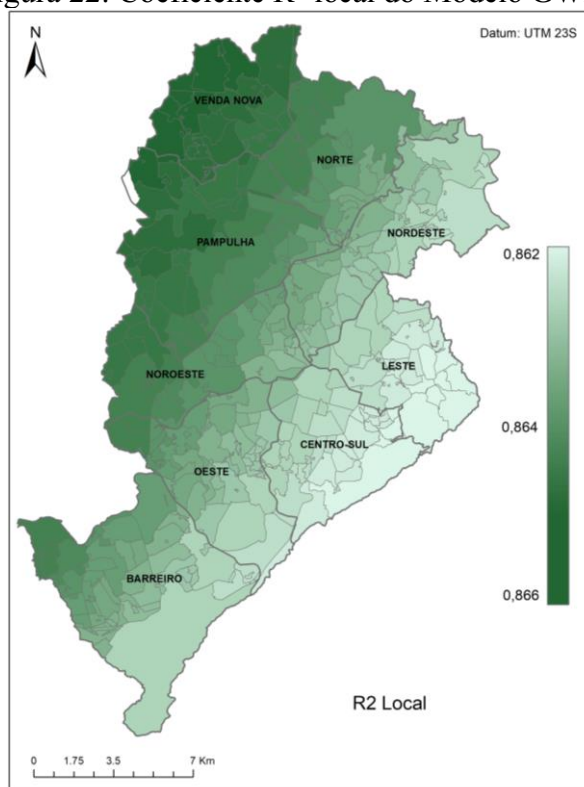


Ao contrário do modelo anterior, no modelo GWR2, a influência da população preta no número de entregas é negativa, sendo essa influência maior, em módulo, nas regionais Barreiro, Oeste e Centro-Sul. A população amarela, assim como no modelo GWR1, exerce influência positiva no número de entregas. No entanto, os coeficientes estimados para essa variável no Modelo GWR2 são maiores nas regionais Barreiro, Oeste, Centro-Sul e Leste. Além disso, os valores dos coeficientes dessa variável são menores no Modelo GWR2. No modelo GWR2, tanto a variável população preta, quanto a variável população amarela apresentaram VIF inferior a 10, ao contrário do modelo anterior.

O coeficiente de determinação espacial apresentou valores satisfatórios, como pode ser visto na Figura 22.

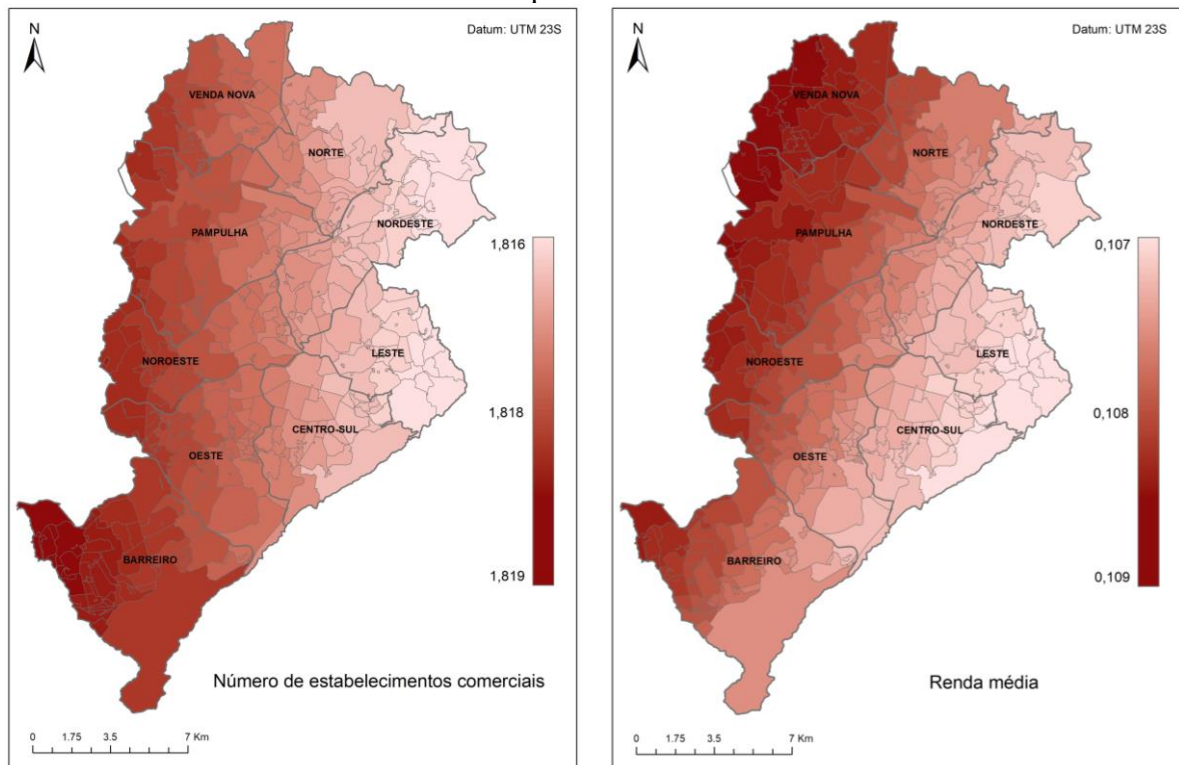
O Modelo GWR3 obteve um AIC de 7246,67, maior que o do Modelo GWR1 em 28,339 unidades e que o do Modelo GWR2 em 33,895 unidades. Apresentou também um AICc de 7254,036, maior que o do Modelo GWR1 em 14,389 e que o do Modelo GWR2 em 30,513 unidades. Os resíduos do modelo foram testados por meio do cálculo do Índice Global de Moran univariado e os resultados obtidos demonstraram uma distribuição aleatória dos resíduos, validando o modelo ( $p\text{-valor} = 0,9314 > 0,05$ ).

Figura 22: Coeficiente  $R^2$  local do Modelo GWR2.



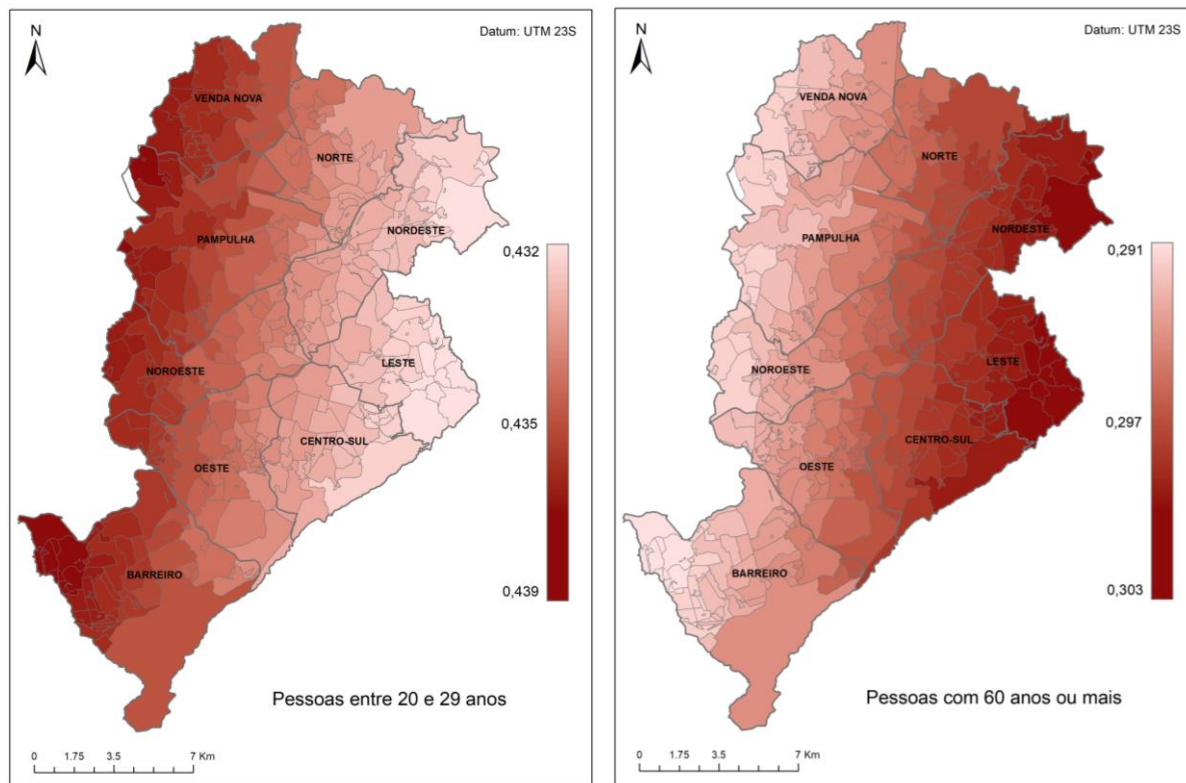
Os resultados estimados para o Modelo GWR3 se encontram na Figura 23 e na Figura 24. Todas as variáveis do modelo apresentaram VIF inferior a 10. Em geral, assim como nos modelos anteriores, os valores estimados para os coeficientes espaciais apresentaram pouca variabilidade espacial.

Figura 23: Coeficientes espaciais das variáveis número de estabelecimentos comerciais e renda média para o Modelo GWR3.



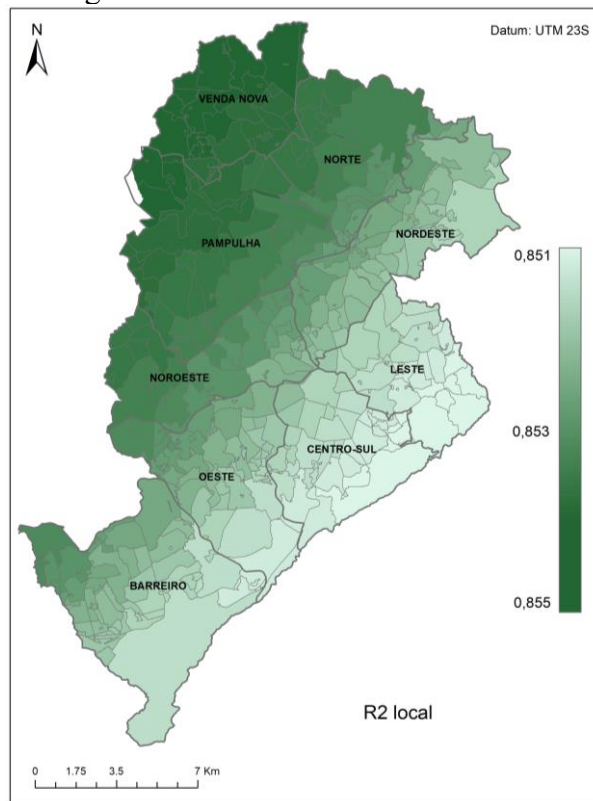
Assim como nos modelos espaciais anteriores, o número de estabelecimentos comerciais apresentou influência positiva no número de entregas do comércio eletrônico. Ao contrário destes modelos, no entanto, os coeficientes estimados têm maior valor na regional Barreiro. Ainda assim, os valores dos coeficientes são bem próximos nos três modelos. A renda média apresentou influência positiva nas entregas, assim como nos modelos GWR1 e GWR2. No entanto, além das regionais Venda Nova e Pampulha, a renda também apresentou valores maiores dos coeficientes na regional Nordeste e em parte do Barreiro, diferentemente dos modelos anteriores. Todas essas regiões, com exceção da Pampulha e parte da Nordeste, são regiões de menor renda média.

Figura 24: Coeficientes espaciais das variáveis relacionadas às faixas etárias para o Modelo GWR3.



Assim como nos modelos anteriores, as pessoas entre 20 e 29 anos tiveram influência positiva nas entregas e os valores dos coeficientes foram semelhantes entre este modelo e o Modelo GWR2. Diferentemente dos modelos anteriores, no entanto, os locais de maior valor dos coeficientes espaciais foram as porções oeste das regionais Venda Nova, Pampulha, Noroeste, Oeste e Barreiro. Assim como no Modelo GWR1, a variável pessoas com 60 anos ou mais apresentou influência positiva no número de entregas. Os valores dos coeficientes espaciais para esta variável foram maiores nas regionais Nordeste, Leste e Centro-Sul, resultado um pouco diferente do primeiro modelo espacial. Além disso, os valores dos coeficientes para essa variável foram significativamente menores no Modelo GWR3 do que no Modelo GWR1. Os coeficientes de determinação do modelo apresentaram valores significativos, como pode ser observado na Figura 25.

Figura 25:  $R^2$  local do Modelo GWR3.



## 5.2 Uberlândia

### 5.2.1 Análise econométrica

Assim como em Belo Horizonte, modelos de regressão de Poisson ou de regressão binomial negativa não puderam ser estimados para os dados completos devido à presença de erros. Assim, foram analisados os 26 bairros responsáveis por 80% das entregas. Este número corresponde a 60% dos bairros da cidade. Na Figura 26 se encontra a representação cartográfica dos bairros analisados e dos removidos da análise em Uberlândia. A Tabela 15 apresenta a estatística descritiva das variáveis referentes aos 26 bairros analisados.

Figura 26: Representação cartográfica de bairros analisados e bairros retirados da análise econométrica de Uberlândia.

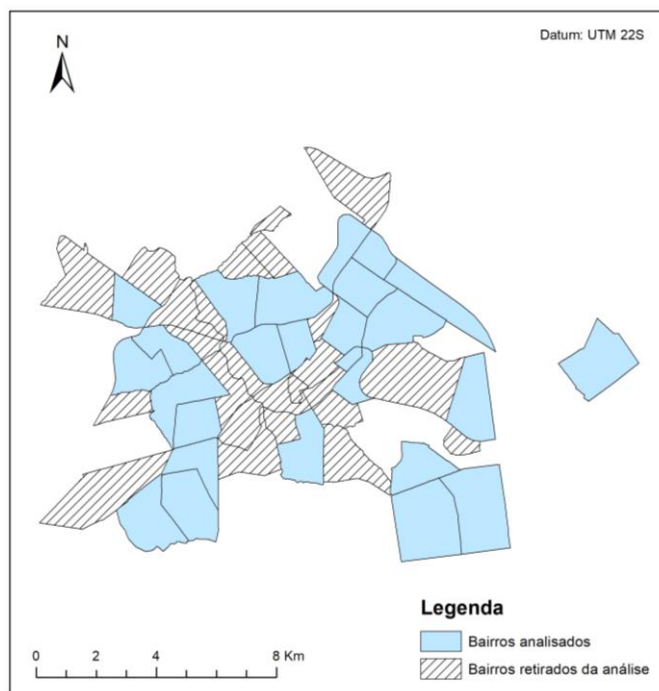


Tabela 15: Estatística descritiva dos 26 bairros analisados em Uberlândia

Variável	Estatística descritiva					
	Mínimo	1º quartil	Mediana	Média	3º quartil	Máximo
Número de entregas	10664	16282	21122	23973	26038	66695
População	2595	8324	11108	11151	15101	20698
Renda média	1021	1321	1719	2075	2394	6598
Número de homens	1278	3978	5256	5297	7381	10080
Número de mulheres	1316	4300	5832	5854	7756	10618
Número de pessoas brancas	1853	4547	5859	6231	7875	11431
Número de pessoas pretas	104,0	477,8	1015,5	1017,7	1486,2	2623,0
Número de pessoas amarelas	16,00	87,75	133,50	131,85	167,75	264,00
Número de pessoas pardas	264	1971	4138	3752	5122	9243
Número de pessoas indígenas	0,00	8,00	13,50	17,50	18,75	63,00
Número de pessoas entre 18 e 19 anos	112,0	361,0	502,0	538,0	703,2	1203,0
Número de pessoas entre 20 e 29 anos	568	2030	2780	2850	3808	5635
Número de pessoas entre 30 e 39 anos	488	1774	2539	2466	3341	4921
Número de pessoas entre 40 e 49 anos	506	1524	2182	2142	2911	4443
Número de pessoas entre 50 e 59 anos	278	1197	1477	1571	2122	2743
Número de pessoas com 60 anos ou mais	221	1044	1454	1585	2158	3575
Número de domicílios	946	3185	4252	4270	5717	7811
Habitantes por domicílio	2,310	2,560	2,620	2,606	2,665	2,760
Estabelecimentos comerciais	0,000	0,000	0,000	3,692	2,750	39,000



Primeiramente, estimou-se um modelo de regressão de Poisson contendo todas as variáveis (renda média, número de homens, número de mulheres, número de pessoas brancas, número de pessoas pretas, número de pessoas amarelas, número de pessoas pardas, número de pessoas indígenas, número de pessoas entre 18 e 19 anos, número de pessoas entre 20 e 29 anos, número de pessoas entre 30 e 39 anos, número de pessoas entre 40 e 49 anos, número de pessoas entre 50 e 59 anos, número de pessoas com 60 anos ou mais, habitantes por domicílio e número de estabelecimentos comerciais) e aplicou-se o teste de dispersão. O p-valor obtido no teste foi de 0,0000002 ( $< 0,05$ ). Logo, devido a presença de superdispersão nos dados, estimou-se o Modelo U1, contendo todas as variáveis, por meio da regressão binomial negativa. Os resultados para este modelo se encontram na Tabela 16. Os valores de VIF menores que 10 foram marcados em verde, e os maiores em vermelho e laranja (valores maiores que 10, mas relativamente próximos de 10).

Tabela 16: Dados estimados para o Modelo U1

Variável	Valor estimado	Teste Wald	Exponencial do valor estimado	VIF	AIC
Intercepto	13,460	7,300***	700815,545	-	525,27
Número de estabelecimentos comerciais	0,031	3,418***	1,031	7,85	
Renda média	-0,000	-2,235*	1,000	12,02	
Habitantes por domicílio	-1,131	-1,608	0,323	8,18	
Número de homens	-0,054	-1,077	0,947	18295970,00	
Número de mulheres	-0,054	-1,075	0,947	20930560,00	
Número de pessoas entre 18 e 19 anos	0,026	0,552	1,026	185716,00	
Número de pessoas entre 20 e 29 anos	0,023	0,473	1,023	4346331,00	
Número de pessoas entre 30 e 39 anos	0,023	0,490	1,024	3778854,00	
Número de pessoas entre 40 e 49 anos	0,025	0,513	1,025	3092713,00	
Número de pessoas entre 50 e 59 anos	0,024	0,506	1,025	1575214,00	
Número de pessoas com 60 anos ou mais	0,023	0,487	1,024	2260990,00	
Número de pessoas brancas	0,031	2,365*	1,031	1506953,00	

Número de pessoas pretas	0,032	2,386*	1,032	105046,20	
Número de pessoas amarelas	0,026	1,983*	1,026	794,93	
Número de pessoas pardas	0,031	2,345*	1,031	1071653,00	
Número de pessoas indígenas	0,038	2,502*	1,039	80,11	

Nível de significância: “\*\*\*\*” 0,001; “\*\*\*” 0,01; “\*\*” 0,05; “.” 0,1

O AIC obtido para modelo foi de 525,27 e apenas as variáveis número de estabelecimentos comerciais e habitantes por domicílio apresentaram VIF menor que 10. A variável renda média apresentou VIF relativamente próximo a 10, porém maior que este número. As variáveis significativas foram o número de estabelecimentos comerciais, a renda média, o número de pessoas brancas, pretas, amarelas, pardas e indígenas. A renda apresentou uma influência negativa no número de entregas, ou seja, um aumento na renda resulta em uma diminuição no número de entregas. As outras variáveis significativas (número de estabelecimentos comerciais, número de pessoas brancas, pretas, amarelas, pardas e indígenas) têm influência positiva no número de entregas; um aumento em uma destas variáveis, mantendo todas as outras fixadas, resulta em um aumento no número de entregas. No entanto, devido ao VIF maior que 10 nas variáveis relacionadas à cor ou raça e à renda média, a influência destes fatores no número de entregas deve ser vista com cautela devido à presença de multicolinearidade.

A partir do Modelo U1, foi obtido o Modelo U2 por meio do *stepwise*. Os resultados obtidos se encontram na Tabela 17.

Tabela 17: Dados estimados para o Modelo U2

Variável	Valor estimado	Teste Wald	Exponencial do valor estimado	VIF	AIC
Intercepto	13,420	7,247***	673336,174	-	523,5
Número de estabelecimentos comerciais	0,033	4,045***	1,033	6,20	
Renda média	-0,000	-2,240*	1,000	11,98	
Habitantes por domicílio	-1,128	-1,596	0,324	8,18	
Número de homens	-0,031	-2,362*	0,969	1232715,00	
Número de mulheres	-0,031	-2,363*	0,969	1400387,00	
Número de pessoas entre 18 e 19 anos	0,004	3,389***	1,004	101,82	
Número de pessoas entre 30 e 39 anos	0,001	1,780*	1,001	361,45	



Número de pessoas entre 40 e 49 anos	0,002	3,792***	1,002	343,50
Número de pessoas entre 50 e 59 anos	0,002	2,684**	1,002	240,64
Número de pessoas com 60 anos ou mais	0,001	2,595**	1,001	74,21
Número de pessoas brancas	0,030	2,303*	1,031	1494373,00
Número de pessoas pretas	0,031	2,324*	1,031	104204,20
Número de pessoas amarelas	0,025	1,919*	1,025	788,01
Número de pessoas pardas	0,030	2,282*	1,030	1061524,00
Número de pessoas indígenas	0,037	2,431*	1,038	78,23

Nível de significância: “\*\*\*” 0,001; “\*\*” 0,01; “\*” 0,05; “.” 0,1

No Modelo U2, apenas a variável habitantes por domicílio não foi significativa. O AIC deste modelo foi de 523,5, apenas 1,77 a menos em relação ao Modelo U1. No entanto, dentre as variáveis significativas, apenas o número de estabelecimentos comerciais apresentou VIF menor que 10, portanto, a influência das variáveis significativas cujo VIF foi maior do que 10 no número de entregas deve ser considerada com cautela. A variável renda média apresentou VIF próximo a 10, porém maior que este número. O número de estabelecimentos comerciais, assim como no Modelo U1, apresentou influência positiva no número de entregas. Também demonstraram influência positiva nas entregas o número de pessoas entre 18 e 19 anos, 30 e 39 anos, 40 e 49 anos, 50 e 59 anos, 60 anos ou mais e o número de pessoas brancas, pretas, amarelas, indígenas e pardas. A variável renda média, assim como no Modelo U1, teve influência negativa no número de entregas. Além dessa variável, o número de homens e o número de mulheres também teve influência negativa nas entregas.

### 5.2.2 Análise espacial

Inicialmente, verificou-se a autocorrelação espacial das entregas do comércio eletrônico em Uberlândia. Para tanto, foram utilizados os dados completos (todos os bairros) no cálculo do Índice Global de Moran univariado. O p-valor obtido foi de 0,180 ( $> 0,05$ ), logo a hipótese nula (os dados apresentam distribuição aleatória) não pôde ser descartada. Desta forma, um modelo de regressão espacial não é adequado para análise dos dados das entregas do comércio eletrônico em Uberlândia.

### 5.3 Patos de Minas

#### 5.3.1 Análise econométrica

Para Patos de Minas, foi possível estimar modelos utilizando os dados completos (todos os bairros da cidade). Inicialmente, estimou-se um modelo de Poisson contendo todas as variáveis independentes (renda média, número de homens, número de mulheres, número de pessoas brancas, número de pessoas pretas, número de pessoas amarelas, número de pessoas pardas, número de pessoas indígenas, número de pessoas entre 18 e 19 anos, número de pessoas entre 20 e 29 anos, número de pessoas entre 30 e 39 anos, número de pessoas entre 40 e 49 anos, número de pessoas entre 50 e 59 anos, número de pessoas com 60 anos ou mais, habitantes por domicílio e número de estabelecimentos comerciais). Neste modelo, aplicou-se o teste de dispersão. O p-valor obtido foi de 0,0001 ( $< 0,05$ ), indicando que há superdispersão. Em seguida, o modelo de regressão binomial negativa, denominado Modelo PM1 foi estimado. Os resultados obtidos para este modelo se encontram na Tabela 18.

Tabela 18: Dados estimados para o Modelo PM1

Variável	Valor estimado	Teste Wald	Exponencial do valor estimado	VIF	AIC
Intercepto	3,904	3,597***	49,601	-	929,28
Número de estabelecimentos comerciais	0,050	0,548	1,051	5,32	
Renda média	0,000	1,682*	1,000	2,08	
Habitantes por domicílio	0,367	0,957	1,443	1,22	
Número de homens	0,005	0,029	1,005	929156,10	
Número de mulheres	0,006	0,033	1,005	1286128,00	
Número de pessoas entre 18 e 19 anos	0,088	0,886	1,095	4320,62	
Número de pessoas entre 20 e 29 anos	0,087	0,846	1,091	92512,60	
Número de pessoas entre 30 e 39 anos	0,088	0,865	1,092	63541,17	
Número de pessoas entre 40 e 49 anos	0,091	0,889	1,095	62294,20	
Número de pessoas entre 50 e 59 anos	0,083	0,817	1,087	48962,91	

Número de pessoas com 60 anos ou mais	0,0888	0,871	1,093	100406,20	
Número de pessoas brancas	-0,092	-0,624	0,912	1621173,00	
Número de pessoas pretas	-0,098	-0,659	0,907	12932,50	
Número de pessoas amarelas	-0,083	-0,551	0,921	652,68	
Número de pessoas pardas	-0,091	-0,616	0,913	470293,20	
Número de pessoas indígenas	-0,127	-0,695	0,880	10,43	

Nível de significância: “\*\*\*\*” 0,001; “\*\*\*” 0,01; “\*\*” 0,05; “.” 0,1

O Modelo PM1 obteve um AIC de 929,28. A renda, única variável independente significativa, apresentou VIF menor que 10; além desta variável, apresentaram VIF menor que 10 apenas as variáveis número de estabelecimentos comerciais e habitantes por domicílio. A variável número de pessoas indígenas apresentou VIF relativamente próximo a 10. Neste modelo, a influência da renda no número de entregas é positiva. A este modelo, foi aplicado o *stepwise* que obteve o Modelo PM2, descrito na Tabela 19 a seguir.

Tabela 19: Dados estimados para o Modelo PM2

Variável	Valor estimado	Teste Wald	Exponencial do valor estimado	VIF	AIC
Intercepto	5,281	32,322***	196,477	-	908,75
Número de pessoas entre 40 e 49 anos	0,005	6,984***	1,005	2,19	
Número de pessoas pretas	-0,008	-3,749***	0,992	2,19	

Nível de significância: “\*\*\*\*” 0,001; “\*\*\*” 0,01; “\*\*” 0,05; “.” 0,1

No Modelo PM2, as duas variáveis independentes apresentaram significância estatística e VIF menor que 10. O número de pessoas entre 40 e 49 anos apresentou influência positiva no número de entregas; o número de pessoas pretas apresentou influência negativa no número de entregas. O modelo apresentou AIC de 908,75.

Além desta análise, a título de comparação com os resultados obtidos para Belo Horizonte e Uberlândia, foram analisados também os dados referentes aos bairros responsáveis por 80% das entregas do comércio eletrônico em Patos de Minas, o que corresponde a 25 bairros (39% dos bairros da cidade). A representação cartográfica dos bairros analisados e removidos da análise se encontra na Figura 27. A Tabela 20 apresenta a estatística descritiva dos dados referentes aos 25 bairros analisados.

Figura 27: Representação cartográfica de bairros analisados e bairros retirados da análise econométrica de Patos de Minas.

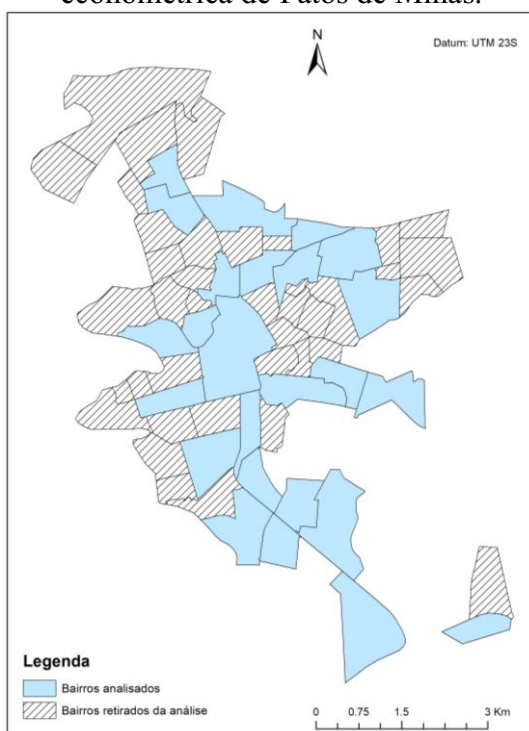


Tabela 20: Estatística descritiva dos 25 bairros analisados em Patos de Minas

Variável	Estatística descritiva					
	Mínimo	1º quartil	Mediana	Média	3º quartil	Máximo
Número de entregas	495	628	745	1478	1011	16649
População	21	1040	1548	2081	2648	7212
Renda média	883,6	1294,7	1482,8	1819,4	1901,7	4134,4
Número de homens	10,0	483,0	768,0	984,3	1240,0	3129,0
Número de mulheres	11	564	779	1096	1387	4083
Número de pessoas brancas	5	762	982	1288	1527	5468
Número de pessoas pretas	3,00	33,00	56,00	81,16	115,00	235,00
Número de pessoas amarelas	0,00	4,00	19,00	19,96	33,00	59,00
Número de pessoas pardas	12,0	348,0	542,0	690,6	967,0	1864,0
Número de pessoas indígenas	0,00	0,00	0,00	0,76	1,00	5,00
Número de pessoas entre 18 e 19 anos	1,0	53,0	81,0	100,8	118,0	292,0
Número de pessoas entre 20 e 29 anos	5,0	246,0	365,0	475,5	601,0	1369,0
Número de pessoas entre 30 e 39 anos	1,0	219,0	345,0	424,4	551,0	1091,0
Número de pessoas entre 40 e 49 anos	8,0	215,0	360,0	411,2	497,0	1238,0
Número de pessoas entre 50 e 59 anos	2	149	213	318	422	1222
Número de pessoas com 60 anos ou mais	2,0	137,0	251,0	351,1	370,0	2057,0
Número de domicílios	6,0	418,0	630,0	823,1	1016,0	2919,0
Habitantes por domicílio	2,180	2,470	2,530	2,572	2,630	3,500
Estabelecimentos comerciais	0,00	0,00	0,00	1,28	1,00	18,00

Estimou-se, então, um modelo de Poisson contendo todas as variáveis independentes, conforme foi feito anteriormente. O teste de dispersão indicou superdispersão dos dados ( $p$ -valor = 0,002 < 0,05). Assim, estimou-se o Modelo PM3 por meio de regressão binomial negativa. Os resultados deste modelo encontram-se na Tabela 21 a seguir.

Tabela 21: Dados estimados para o Modelo PM3

Variável	Valor estimado	Teste Wald	Exponencial do valor estimado	VIF	AIC
Intercepto	8,270	10,649***	3904,949	-	366,64
Número de estabelecimentos comerciais	-0,044	-1,006	0,957	15,58	
Renda média	-0,000	-1,966*	0,999	3,41	
Habitantes por domicílio	-0,640	-2,421*	0,527	2,21	
Número de homens	0,301	3,173**	1,351	2748602,00	
Número de mulheres	0,300	3,132**	1,350	4169022,00	
Número de pessoas entre 18 e 19 anos	-0,166	-3,401***	0,848	6887,40	
Número de pessoas entre 20 e 29 anos	-0,187	-3,707***	0,829	171239,00	
Número de pessoas entre 30 e 39 anos	-0,184	-3,692***	0,832	114046,50	
Número de pessoas entre 40 e 49 anos	-0,184	-3,691***	0,832	117295,10	
Número de pessoas entre 50 e 59 anos	-0,190	-3,810***	0,827	113337,20	
Número de pessoas com 60 anos ou mais	-0,182	-3,641***	0,834	255788,30	
Número de pessoas brancas	-0,115	-1,280	0,891	5785025,00	
Número de pessoas pretas	-0,117	-1,286	0,890	25658,63	
Número de pessoas amarelas	-0,111	-1,202	0,895	1422,39	
Número de pessoas pardas	-0,117	-1,297	0,889	1257317,00	
Número de pessoas indígenas	-0,016	-0,172	0,984	10,44	

Nível de significância: “\*\*\*\*” 0,001; “\*\*\*” 0,01; “\*\*” 0,05; “.” 0,1

No modelo PM3, as variáveis independentes significativas foram a renda média, habitantes por domicílio, número de homens, número de mulheres e as variáveis referentes às faixas etárias. No entanto, apenas renda média e habitantes por domicílio têm VIF menor que 10; sendo assim, a influência das variáveis estatisticamente significativas que apresentam VIF maior que 10 deve

ser considerada com cautela devido à presença de multicolinearidade. As variáveis renda, habitantes por domicílio, número pessoas entre 18 e 19 anos, 20 e 20 anos, 30 e 39 anos, 40 e 49 anos, 50 e 59 anos e 60 anos ou mais apresentaram influência negativa no número de entregas. Enquanto as variáveis número de homens e número de mulheres têm influência positiva. Ao Modelo PM3, foi aplicado o *stepwise*, que gerou o Modelo PM4, detalhado na Tabela 22.

Tabela 22: Dados estimados para o Modelo PM4

Variável	Valor estimado	Teste Wald	Exponencial do valor estimado	VIF	AIC
Intercepto	7,893	10,281***	2678,467	-	363,55
Renda média	-0,000	-1,712*	0,999	2,97	
Habitantes por domicílio	-0,501	-1,943*	0,606	1,88	
Número de homens	0,191	3,766***	1,210	699218,90	
Número de mulheres	0,189	3,710***	1,208	1044933,00	
Número de pessoas entre 18 e 19 anos	-0,155	-3,271**	0,857	5805,61	
Número de pessoas entre 20 e 29 anos	-0,183	-3,683***	0,833	148040,60	
Número de pessoas entre 30 e 39 anos	-0,177	-3,630***	0,837	97890,58	
Número de pessoas entre 40 e 49 anos	-0,183	-3,704***	0,833	102072,70	
Número de pessoas entre 50 e 59 anos	-0,184	-3,757***	0,832	97406,09	
Número de pessoas com 60 anos ou mais	-0,177	-3,595***	0,838	220560,50	
Número de pessoas brancas	-0,010	-2,393*	0,991	9963,52	
Número de pessoas pretas	-0,010	-2,024*	0,991	60,94	
Número de pessoas pardas	-0,011	-2,639**	0,989	2341,33	

O Modelo PM4 apresentou AIC de 363,55 (uma diferença de 3,09 em relação ao AIC do Modelo PM3). Todas as variáveis independentes do Modelo PM4 são estatisticamente significativas. Assim como no Modelo PM3, as variáveis cujo VIF é menor que 10 são a renda média e os habitantes por domicílio. As variáveis número de homens e número de mulheres têm influência positiva no número de entregas, assim como no Modelo PM3. Assim como no

Modelo PM3, a renda média, os habitantes por domicílio e as faixas etárias têm influência negativa no número de entregas. Além disso, a população branca, a população preta e a população parda também apresentaram influência negativa nas entregas.

### **5.3.2 Análise espacial**

Verificou-se, inicialmente, a autocorrelação espacial do número de entregas do comércio eletrônico. O p-valor obtido no cálculo do Índice Global de Moran univariado foi de 0,513 ( $> 0,05$ ). Dessa forma, não se pôde descartar a hipótese nula de que não há dependência espacial dos dados. Assim, a modelagem por regressão espacial não é adequada para analisar os dados de entregas do comércio eletrônico em Patos de Minas.

## 6 DISCUSSÃO

Os resultados apresentados na seção anterior para os modelos econométricos estão resumidos na Tabela 23.

Tabela 23: Comparação entre modelos econométricos

<b>Modelo</b>	<b>Variáveis significativas</b>	<b>VIF</b>	<b>AIC</b>
BH1	Número de estabelecimentos comerciais (+)	< 10	2224,6
	Renda média (+)	< 10	
BH2	Número de estabelecimentos comerciais (+)	< 10	2207,4
	Renda média (+)	< 10	
	Número de pessoas entre 20 e 29 anos (+)	< 10	
U1	Número de estabelecimentos comerciais (+)	< 10	525,27
	Renda média (-)	> 10	
	Número de pessoas brancas (+)	> 10	
	Número de pessoas pretas (+)	> 10	
	Número de pessoas amarelas (+)	> 10	
	Número de pessoas pardas (+)	> 10	
	Número de pessoas indígenas (+)	> 10	
U2	Número de estabelecimentos comerciais (+)	< 10	523,5
	Renda média (-)	> 10	
	Número de homens (-)	> 10	
	Número de mulheres (-)	> 10	
	Número de pessoas entre 18 e 19 anos (+)	> 10	
	Número de pessoas entre 30 e 39 anos (+)	> 10	
	Número de pessoas entre 40 e 49 anos (+)	> 10	
	Número de pessoas entre 50 e 59 anos (+)	> 10	
	Número de pessoas com 60 anos ou mais (+)	> 10	
	Número de pessoas brancas (+)	> 10	
	Número de pessoas pretas (+)	> 10	
	Número de pessoas amarelas (+)	> 10	
	Número de pessoas pardas (+)	> 10	
	Número de pessoas indígenas (+)	> 10	
	PM1	Renda média (+)	
PM2	Número de pessoas entre 40 e 49 anos (+)	< 10	908,75
	Número de pessoas pretas (-)	< 10	
PM3	Renda média (-)	< 10	366,64
	Habitantes por domicílio (-)	< 10	



	Número de homens (+)	> 10	
	Número de mulheres (+)	> 10	
	Número de pessoas entre 18 e 19 anos (-)	> 10	
	Número de pessoas entre 20 e 29 anos (-)	> 10	
	Número de pessoas entre 30 e 39 anos (-)	> 10	
	Número de pessoas entre 40 e 49 anos (-)	> 10	
	Número de pessoas entre 50 e 59 anos (-)	> 10	
	Número de pessoas com 60 anos ou mais (-)	> 10	
PM4	Renda média (-)	< 10	363,55
	Habitantes por domicílio (-)	< 10	
	Número de homens (+)	> 10	
	Número de mulheres (+)	> 10	
	Número de pessoas entre 18 e 19 anos (-)	> 10	
	Número de pessoas entre 20 e 29 anos (-)	> 10	
	Número de pessoas entre 30 e 39 anos (-)	> 10	
	Número de pessoas entre 40 e 49 anos (-)	> 10	
	Número de pessoas entre 50 e 59 anos (-)	> 10	
	Número de pessoas com 60 anos ou mais (-)	> 10	
	Número de pessoas brancas (-)	> 10	
	Número de pessoas pretas (-)	> 10	
	Número de pessoas pardas (-)	> 10	

No caso de Belo Horizonte, demonstrou-se que, nos bairros responsáveis por 80% das entregas, o número de estabelecimentos comerciais tem um efeito positivo nas entregas, ou seja, um aumento no número de estabelecimentos comerciais gera um aumento no número de entregas. Esse resultado é compatível com Farag *et al.* (2005), que verificaram que, para residentes de regiões urbanas, as compras *online* são complementares às compras presenciais, ou seja, o *e-commerce* não substitui as compras no varejo tradicional. Farag *et al.* (2007) e De Blasio (2008) também obtiveram resultados compatíveis com os deste trabalho. Os autores concluíram que residentes de regiões urbanas, com maior acesso ao varejo tradicional, compram mais *online* que moradores de regiões com menor acesso ao varejo tradicional. Ainda, segundo De Blasio (2008) e Cao (2012), isso pode estar associado ao desejo das pessoas verem os produtos fisicamente antes de adquiri-los.

O número de estabelecimentos comerciais também demonstrou uma influência positiva no número de entregas para os bairros responsáveis por 80% das entregas em Uberlândia. No

entanto, em Uberlândia, destaca-se que os bairros com o maior número de estabelecimentos comerciais – como pode ser visto na Figura 5 e na Figura 26 – fazem parte do grupo responsável por apenas 20% das entregas. Em Belo Horizonte, isso não ocorre.

Em Belo Horizonte, a renda média mostrou uma influência positiva nas entregas do comércio eletrônico nos bairros responsáveis por 80% das entregas. Esse resultado é compatível com os resultados obtidos por Li *et al.* (1999), Morganosky e Cude (2000), Vrechopoulos *et al.* (2001), Sim e Koi (2002), Swinyard e Smith (2003), Farag *et al.* (2005), Krizek *et al.* (2005), Farag *et al.* (2006a), Farag *et al.* (2006b), Farag *et al.* (2007), Lennon *et al.* (2007), Soopramanien e Robertson (2007), De Blasio (2008), Brashear *et al.* (2009), Hashin *et al.* (2009), Mokhtarian *et al.* (2009), Sener e Reeder (2012), Cao *et al.* (2013), Zhou e Wang (2014), Clarke *et al.* (2015), Lee *et al.* (2015), Lee *et al.* (2017), Motte-Baumvol *et al.* (2017), Beckers *et al.* (2018), Seitz *et al.* (2017), Lubis (2018), Fernández-Bonilla (2020), Hood *et al.* (2020), Opesade (2020), Pérez-Amaral *et al.* (2020), Saphores e Xu (2020), Valarezo *et al.* (2020), Cheng *et al.* (2021) e Song (2021). Destaca-se, no entanto, que alguns bairros de alta renda nas regionais Centro-Sul e Pampulha fazem parte do grupo responsável por apenas 20% das entregas do comércio eletrônico em Belo Horizonte. Isso pode ter relação com o fato destes bairros terem um baixo número de habitantes. No entanto, este detalhamento não foi objeto desta dissertação e sugere-se uma investigação em trabalhos futuros.

O Modelo PM1, que se refere à cidade de Patos de Minas como um todo, apresentou resultado semelhante em relação à renda média, com essa variável tendo influência positiva nas entregas do *e-commerce* na cidade. No entanto, no modelo PM2, considerado o melhor dentre PM1 e PM2 pelo critério do AIC, a renda não apresentou significância estatística e não tem influência no comércio eletrônico em Patos de Minas. Em Uberlândia, e também nos bairros responsáveis por 80% das entregas em Patos de Minas, a renda teve influência negativa no número de entregas. Isto é, um aumento na renda média gera uma diminuição no número de entregas. Este resultado não converge com o encontrado na literatura. Além disso, como o VIF é menor que 10 para a renda média em PM3 e PM4, e significativamente próximo de 10 nos modelos U1 e U2, a multicolinearidade não parece ter influenciado o resultado. Destaca-se que alguns bairros de alta renda em Uberlândia e Patos de Minas fazem parte do grupo responsável por apenas 20% das entregas, provavelmente por terem baixa ou média população.

Em Belo Horizonte, para os bairros responsáveis por 80% das entregas, a população entre 20 e 29 anos demonstrou influência positiva no número de entregas. Esse resultado é compatível

com os obtidos por Farag *et al.* (2006b) e Loo e Wang (2007). Em Patos de Minas, considerando os dados de todos os bairros, a população entre 40 e 49 anos apresentou influência positiva nas entregas do *e-commerce*, o que converge com os trabalhos de Swinyard e Smith (2003) e Beckers *et al.* (2018), e parcialmente com os trabalhos de Farag *et al.* (2005) – pessoas entre 26 e 45 anos compram mais *online* – e Vrechopoulos *et al.* (2001), Weltevreden (2007), Clarke *et al.* (2015) e Hood *et al.* (2020) – pessoas entre 25 e 44 anos compram mais *online*. A idade é um fator importante para o comércio eletrônico e está intrinsecamente ligada ao tipo de produto comercializado e ao canal de comercialização. Nos Estados Unidos, a faixa etária entre 25-34 concentra a maioria dos consumidores do comércio eletrônico (Statista, 2020). Na Europa, a faixa etária entre 16-24 anos tem igual participação no comércio eletrônico daqueles que estão na faixa etária entre 25-54 anos (Eurostat, 2021). Os nascidos a partir de 1985 são pessoas que usam a Internet de modo natural, por isto são mais propensos a usar o comércio eletrônico (Jongen, 2018). No Brasil, as faixas etárias que mais utilizam o comércio eletrônico são as pessoas entre 26 e 35 anos e entre 36 e 50 anos (Statista, 2021).

Em Uberlândia, para os bairros responsáveis por 80% das entregas, todas as faixas etárias apresentaram influência positiva no número de entregas, com exceção da população entre 20 e 29 anos, que não apresentou significância estatística. Este resultado indica que, de maneira geral, todas as faixas etárias têm propensão a receber entregas do comércio eletrônico, com exceção da população entre 20 e 29 anos. Ainda, o valor estimado do coeficiente para a população entre 18 e 19 anos é ligeiramente superior às outras faixas etárias, indicando que esse grupo tende a receber mais entregas que os outros. Isso converge parcialmente com os resultados obtido por Maan e Kumar (2020) – que concluíram que indivíduos entre 18 e 35 anos são os que mais compram *online* –, Pérez-Amaral *et al.* (2020) – que verificaram uma maior tendência dos indivíduos entre 16 e 24 anos a usar o *e-commerce* – e Valerzo *et al.* (2020) – que concluíram que pessoas menores de 25 anos compram mais *online*.

Ao contrário de Uberlândia, nos bairros responsáveis por 80% das entregas em Patos de Minas, todas as faixas etárias têm influência negativa no número de entregas. Ainda, o valor do coeficiente estimado para a população entre 50 e 59 anos é ligeiramente mais negativo que o das outras faixas etárias, indicando que esse grupo tende a receber menos entregas que os outros. Esse resultado converge com os resultados de Beckers *et al.* (2018), no entanto, diverge do obtido pelo modelo PM2, que utiliza todos os bairros, que verificou uma influência positiva da população entre 40 e 49 anos, conforme detalhado anteriormente. Por fim, destaca-se que, tanto em Uberlândia quanto em Patos de Minas (para os bairros responsáveis por 80% das entregas),

as variáveis referentes às faixas etárias apresentaram multicolinearidade (VIF maior que 10), o que prejudica a interpretação dos resultados para essas variáveis. Neste sentido, maiores análises são necessárias para entender a influência da idade nas entregas do comércio eletrônico, visto que esta variável apresenta resultados divergentes para as cidades investigadas neste estudo.

Em Uberlândia, para os bairros responsáveis por 80% das entregas, todas as variáveis referentes à cor ou raça apresentaram influência positiva no número de entregas, indicando que todas as cores ou raças têm propensão a receber entregas do comércio eletrônico. Ainda, o coeficiente estimado para a população indígena é ligeiramente maior, apontando esse grupo como tendo uma tendência um pouco maior a receber entregas, ao contrário dos resultados obtidos por Ren e Kwan (2009) e Saphores e Xu (2020) – que concluíram que a população branca tem maior propensão a comprar *online*. Em Patos de Minas, considerando-se todos os bairros, a população preta mostrou influência negativa no número de entregas, tendo propensão a receber menos entregas. Em relação aos bairros responsáveis por 80% das entregas nesta cidade, verificou-se que as populações branca, preta e parda têm influência negativa nas entregas do comércio eletrônico, tendendo a receber menos entregas, sendo a população parda a de coeficiente estimado mais negativo. Destaca-se que, tanto em Uberlândia quanto em Patos de Minas (apenas para o modelo com os bairros responsáveis por 80% das entregas), as variáveis referentes à cor ou raça apresentaram multicolinearidade (VIF maior que 10), o que prejudica a interpretação dos resultados. Os resultados em relação à cor/raça indicam que a heterogeneidade racial da população brasileira não permite uma conclusão precisa sobre sua influência nas compras do comércio eletrônico. Desta forma, não se recomenda para futuros trabalhos a inclusão desta variável em estudos relativos ao comércio eletrônico.

Em Uberlândia, para os bairros responsáveis por 80% das entregas, tanto o número de homens quanto o número de mulheres apresentou influência negativa no número de entregas. Em Patos de Minas, para os bairros responsáveis por 80% das entregas, observou-se efeito contrário, isto é, tanto o número de homens quanto o de mulheres têm influência positiva no número de entregas, ou seja, de maneira geral, ambos têm tendência a receber entregas do comércio eletrônico. Ainda, o coeficiente estimado para o número de homens foi ligeiramente superior, indicando que os homens recebem ligeiramente mais entregas do que as mulheres. Esses resultados convergem com os resultados obtidos por Vrechopoulos *et al.* (2001), Sim e Koi (2002), Brown *et al.* (2003), Stafford *et al.* (2004), Farag *et al.* (2005), Farag *et al.* (2006a), Farag *et al.* (2006b), Weltevreden (2007), Hashin *et al.* (2009), Burkolter e Kluge (2011), Pérez-

Hernández e Sánchez-Mangas (2011), Clarke *et al.* (2015), Beckers *et al.* (2018), Lubis (2018), Fernández-Bonilla (2020), Pérez-Amaral *et al.* (2020) e Valarezo *et al.* (2020). Destaca-se que, tanto em Uberlândia quanto em Patos de Minas, essas variáveis apresentaram multicolinearidade (VIF maior que 10), o que prejudica a interpretação dos resultados.

Em Patos de Minas, para os bairros responsáveis por 80% das entregas, verificou-se que o número de habitantes por domicílio tem efeito negativo no número de entregas, ou seja, domicílios com menor número de pessoas tendem a receber mais entregas. Esse resultado converge com os estudos de Pérez-Hernández e Sánchez-Mangas (2011), Zhou e Wang (2014) e Pérez-Amaral *et al.* (2020).

Os resultados obtidos com os modelos espaciais estão resumidos na Tabela 24. Apenas as entregas do comércio eletrônico de Belo Horizonte apresentaram autocorrelação espacial, indicando que a entrega do comércio eletrônico não possui um padrão espacial definido para qualquer localidade. Ainda, este resultado mostra a importância de associar técnicas econométricas com técnicas espaciais para análise do fenômeno. Para Belo Horizonte, os coeficientes estimados dos modelos apresentam pouca variabilidade espacial, o que indica que mesmo o espaço sendo heterogêneo, a combinação destes fatores apresenta um padrão homogêneo nos bairros.

Tabela 24: Comparação entre modelos espaciais

<b>Modelo</b>	<b>Variáveis</b>	<b>VIF</b>	<b>AICc</b>
GWR1	Número de estabelecimentos comerciais (+)	< 10	7239,65
	Renda média (+)	< 10	
	Habitantes por domicílio (-)	< 10	
	Número de homens (-)	> 10	
	Número de mulheres (-)	> 10	
	Número de pessoas entre 18 e 19 anos (+)	> 10	
	Número de pessoas entre 20 e 29 anos (+)	> 10	
	Número de pessoas entre 30 e 39 anos (+)	> 10	
	Número de pessoas entre 40 e 49 anos (+)	> 10	
	Número de pessoas entre 50 e 59 anos (+)	> 10	
	Número de pessoas com 60 anos ou mais (+)	> 10	
	Número de pessoas brancas (+)	> 10	
	Número de pessoas pretas (+)	> 10	
	Número de pessoas amarelas (+)	> 10	
	Número de pessoas pardas (+)	> 10	

	Número de pessoas indígenas (+)	> 10	
GWR2	Número de estabelecimentos comerciais (+)	< 10	7223,52
	Renda média (+)	< 10	
	Número de mulheres (+)	> 10	
	Número de pessoas entre 18 e 19 anos (-)	> 10	
	Número de pessoas entre 20 e 29 anos (+)	> 10	
	Número de pessoas pretas (-)	< 10	
	Número de pessoas amarelas (+)	< 10	
GWR3	Número de estabelecimentos comerciais (+)	< 10	7254,04
	Renda média (+)	< 10	
	Número de pessoas entre 20 e 29 anos (+)	< 10	
	Número de pessoas com 60 anos ou mais (+)	< 10	

Em Belo Horizonte, a análise espacial mostrou a influência positiva do número de estabelecimentos comerciais nas entregas, indicando que em bairros onde há um maior número de comércios, tende-se a receber mais entregas do comércio eletrônico. Esse resultado converge com os estudos de Farag *et al.* (2005), Farag *et al.* (2007) e De Blasio (2008), previamente explicados nesta seção para os modelos econométricos. Considerando o Modelo GWR 2, que obteve o melhor ajuste segundo o critério do AICc, a influência do número de estabelecimentos comerciais é maior na regional de Venda Nova e parte da regional Pampulha (fronteira com Venda Nova). A regional Venda Nova concentra um número considerável de estabelecimentos comerciais. No entanto, destaca-se que a variação espacial dos valores dos coeficientes estimados é pequena, portanto, as diferenças de magnitude da influência do número de estabelecimentos comerciais nas entregas em diferentes bairros em Belo Horizonte é significativamente baixa.

A renda média também apresentou influência positiva no número de entregas do *e-commerce* em Belo Horizonte, ou seja, bairros com maior renda tendem a receber mais entregas, o que está associado ao maior poder de compra desses bairros. Esse resultado é compatível com os estudos de Li *et al.* (1999), Morganosky e Cude (2000), Vrechopoulos *et al.* (2001), Sim e Koi (2002), Swinyard e Smith (2003), Farag *et al.* (2005), Krizek *et al.* (2005), Farag *et al.* (2006a), Farag *et al.* (2006b), Farag *et al.* (2007), Lennon *et al.* (2007), Soopramanien e Robertson (2007), De Blasio (2008), Brashear *et al.* (2009), Hashin *et al.* (2009), Mokhtarian *et al.* (2009), Sener e Reeder (2012), Cao *et al.* (2013), Zhou e Wang (2014), Clarke *et al.* (2015), Lee *et al.* (2015), Lee *et al.* (2017), Motte-Baumvol *et al.* (2017), Beckers *et al.* (2018), Seitz *et al.* (2017), Lubis (2018), Fernández-Bonilla (2020), Hood *et al.* (2020), Opesade (2020), Pérez-Amaral *et*

*al.* (2020), Saphores e Xu (2020), Valarezo *et al.* (2020), Cheng *et al.* (2021) e Song (2021). Considerando Modelo GWR2, a influência dessa variável é maior na regional Venda Nova, onde os bairros possuem renda média menor, indicando que outras variáveis associadas à renda podem exercer influência na distribuição espacial das entregas em Belo Horizonte. No entanto, é importante destacar que a variação espacial do coeficiente associado à esta variável é significativamente baixa.

O número de habitantes por domicílio apresentou influência negativa nas entregas, indicando que em bairros onde há um menor número de pessoas por domicílio se recebe mais entregas. Esse resultado é compatível com os estudos de Pérez-Hernández e Sánchez-Mangas (2011), Zhou e Wang (2014) e Pérez-Amaral *et al.* (2020). A influência desta variável foi maior na regional Barreiro.

O gênero apresentou influência negativa nas entregas segundo o Modelo GWR1, sendo que os coeficientes destas variáveis têm valores próximos para ambos os gêneros, sendo ligeiramente menores para os homens, isto é, homens tendem a receber menos entregas que as mulheres. Os coeficientes apresentam influência maior nas regionais Norte, Nordeste e Leste. No entanto, é importante destacar que a variação espacial dos valores dos coeficientes para estas variáveis é baixo. No Modelo GWR2, apenas o número de mulheres foi incluído no modelo, que apresentou influência positiva nas entregas do comércio eletrônico, o que indica que bairros com maior população de mulheres tendem a receber mais entregas. Essa influência é maior nos bairros das regionais Leste, Centro-Sul, Oeste e Barreiro. No entanto, destaca-se que a variação espacial dos valores dos coeficientes estimados para esta variável é baixa. Os valores do VIF para as variáveis referentes ao gênero em ambos os modelos é maior que 10. A presença de multicolinearidade pode ser uma possível explicação para a variação no sinal dos coeficientes estimados de um modelo para o outro para a mesma variável (Hair *et al.*, 2019).

Quanto às variáveis referentes às faixas etárias, no Modelo GWR1, todas elas apresentaram influência positiva no número de entregas e os valores dos coeficientes estimados eram relativamente próximos, sendo a população entre 20 e 29 anos ligeiramente mais influente. Em contrapartida, no modelo GWR2, a população entre 18 e 19 anos tem influência negativa nas entregas, indicando que bairros com maior população nesta faixa etária tendem a receber menos entregas do *e-commerce*. Apesar da baixa variação espacial dos valores dos coeficientes estimados, a influência dessa variável é maior nas regionais Venda Nova (onde há grande concentração de pessoas nessa faixa etária), Norte e Nordeste. A influência negativa desta faixa

etária pode estar associada ao fato de serem adolescentes, grupo que, em geral, ainda não possui renda significativa. De fato, as regionais Venda Nova, Norte e Nordeste possuem menor renda média. A faixa etária entre 20 e 29 anos apresentou influência positiva nas entregas, indicando que bairros com maior presença desse grupo recebem mais entregas do comércio eletrônico. Apesar da baixa variação espacial dos coeficientes estimados para essa variável, a regional de Venda Nova (onde há grande concentração de pessoas nessa faixa etária) apresentou uma influência maior dessa variável. Os resultados do modelo GWR2 para as pessoas entre 20 e 29 anos têm convergência com os estudos de Farag *et al.* (2006b) e Loo e Wang (2007). Os valores do VIF para as variáveis referentes à idade em ambos os modelos é maior que 10. A presença de multicolinearidade pode ser uma possível explicação para a variação no sinal dos coeficientes estimados de um modelo para o outro para a mesma variável (Hair et al., 2019).

Quanto às variáveis referentes à cor ou raça, no modelo GWR1, todas tiveram influência positiva no número de entregas, sendo essa influência maior nos bairros das regionais Nordeste e Leste. Ainda, os valores dos coeficientes estimados apresentaram maiores valores para a população amarela, indicando que este grupo tende a receber mais entregas do comércio eletrônico que os outros. No entanto, a variação espacial dos coeficientes estimados para as variáveis não é alta. Em contrapartida, no modelo de melhor ajuste, GWR2, apenas a população preta e a população amarela foram incluídas. Neste modelo, a população preta apresentou efeito negativo nas entregas, ou seja, os bairros com maior população preta tendem a receber menos entregas. A população amarela apresentou influência positiva nas entregas. Desta forma, pode-se dizer que a raça tem um efeito heterogêneo nas entregas urbanas. Estes resultados convergem com os estudos de Ren e Kwan (2009) e Saphores e Xu (2020). Apesar da baixa variação espacial dos coeficientes estimados, essas variáveis exercem maior influência nas regionais Barreiro, Oeste, Centro-Sul e Leste. As regionais Barreiro e Oeste são regionais de menor renda, enquanto a regional Centro-Sul e Leste têm baixa população preta e amarela. No modelo GWR2, tanto a variável população preta, quanto a variável população amarela apresentaram VIF inferior a 10, ao contrário das variáveis associadas à cor ou raça no Modelo GWR1.

Quanto ao Modelo GWR3, estimado com as mesmas variáveis do melhor modelo econométrico para Belo Horizonte (Modelo BH2), verificou-se que as variáveis apresentaram influência positiva no número de entregas em ambos os modelos.

Por fim, comparando-se os resultados dos modelos econométricos e dos modelos espaciais em Belo Horizonte, verificou-se que os modelos espaciais permitiram uma avaliação mais completa



dos efeitos das variáveis socioeconômicas e de acesso às oportunidades de compra nas entregas do comércio eletrônico. O melhor modelo espacial incluiu mais variáveis de categorias diferentes (acesso ao varejo tradicional, renda, gênero, idade e cor ou raça), permitindo uma avaliação mais detalhada do impacto das variáveis nas entregas, enquanto o melhor modelo econométrico considerou apenas o acesso ao varejo tradicional, a renda e a idade. Além disso, o modelo espacial apresenta a variação espacial dos coeficientes estimados em todo o território, o que permitiu verificar a heterogeneidade da influência nos diferentes bairros da cidade.

Assim, os modelos espaciais permitiram um entendimento mais completo da distribuição das entregas do e-commerce na região, o que faz com que os resultados obtidos tenham potencial de auxiliar de forma mais eficaz políticas públicas para reduzir as externalidades negativas associadas ao comércio eletrônico. Por exemplo, ajudar na escolha de locais para instalação de centros de distribuição de mercadorias e *pick-up points*, além da identificação de locais com maior potencial de serem impactados pelo tráfego e poluição causados pelo transporte da última milha.

## 7 CONCLUSÃO

Nesta dissertação analisou-se a relação entre as entregas do comércio eletrônico e os fatores socioeconômicos e oportunidade de acesso ao varejo tradicional, analisou-se também a correlação espacial entre essas variáveis e as entregas, bem como comparou-se a influência dos fatores analisados para Belo Horizonte, Uberlândia e Patos de Minas. Portanto, os objetivos deste trabalho foram alcançados.

Foram utilizados dados reais de entregas do comércio eletrônico obtidos por meio de uma empresa transportadora com forte atuação nos mercados das três cidades, além de dados socioeconômicos do Censo de 2010, extrapolados para o ano de 2019, e dados relativos ao acesso ao varejo tradicional. Estes últimos foram expressos neste trabalho como o número de estabelecimentos comerciais por bairro analisado. A análise fez uso de técnicas econométricas e espaciais.

Primeiramente, conduziu-se a análise econométrica por meio de regressão binomial negativa. Os modelos estimados permitiram identificar os fatores que influenciam o número de entregas do comércio eletrônico em Belo Horizonte, Uberlândia e Patos de Minas. Em seguida, realizou-se a análise espacial. O Índice Global de Moran Univariado (I) indicou dependência espacial das entregas urbanas em Belo Horizonte, demonstrando que a distribuição das entregas não possui um padrão espacial definido em qualquer localidade. Em seguida, foram estimados modelos de regressão geograficamente ponderada para Belo Horizonte. Dessa forma, foi verificada a influência de atributos espaciais nas entregas do comércio eletrônico naquela cidade.

Quanto às hipóteses de pesquisa estabelecidas, a variável gênero apresentou influência nas entregas nas três cidades analisadas, sendo verificado, de forma geral, que homens e mulheres têm tendência a receber entregas do comércio eletrônico. As variáveis idade e cor ou raça apresentaram efeito mais uniforme nas entregas do comércio eletrônico em Uberlândia e para os bairros responsáveis por 80% das entregas em Patos de Minas. Em Belo Horizonte e Patos de Minas (todos os bairros) estas variáveis apresentaram efeito heterogêneo. A renda influenciou as entregas em todas as cidades, no entanto, esta influência se deu de maneira positiva em apenas Belo Horizonte. O número de habitantes por domicílio, considerando o modelo tido como melhor segundo o critério do AIC, foi significativo em Patos de Minas (para os bairros responsáveis por 80% das entregas), tendo influência negativa nas entregas do e-

*commerce*. Por fim, o número de estabelecimentos comerciais demonstrou influência positiva nas entregas em Belo Horizonte e Uberlândia.

Este estudo apresenta, no entanto, algumas limitações. Algumas variáveis apresentaram multicolinearidade, conforme indicado pelo VIF, apesar do uso das técnicas *stepwise* e *model selection* para seleção dos melhores modelos econométricos e espaciais, respectivamente. Isso torna a interpretação da influência de algumas variáveis menos confiável. A ausência de dados socioeconômicos recentes, devido à não realização do censo demográfico em 2020, também constitui uma limitação, embora de menor impacto. A ausência de dados oficiais disponibilizados publicamente pelos governos de Uberlândia e Patos de Minas sobre o número de estabelecimentos comerciais em cada bairro também foi uma limitação, uma vez que os dados obtidos por meio do OpenStreetMap não são tão completos.

Ainda, não foi possível estimar modelos econométricos para os dados completos em Belo Horizonte e Uberlândia, o que limita as conclusões feitas a partir destes resultados apenas aos bairros responsáveis por 80% das entregas nas cidades. Também não foi possível estimar modelos de regressão de Poisson ou binomial negativa geograficamente ponderados. Não obstante, os modelos geograficamente ponderados estimados apresentaram resultados satisfatórios.

O estudo desenvolvido neste trabalho mostra a importância das técnicas econométricas e espaciais para análise das entregas da última milha. Ainda, este trabalho contribuiu para o entendimento dos fatores influentes nas entregas do *e-commerce* não apenas em termos metodológicos. O uso de dados reais de entregas e a análise no contexto de um país em desenvolvimento latino-americano contribuem com a expansão do entendimento do fenômeno das entregas. Além disso, este estudo traz resultados que podem auxiliar no desenvolvimento de políticas públicas que visem atenuar as externalidades negativas do transporte urbano das entregas do comércio eletrônico.

Para trabalhos futuros, recomenda-se um aprofundamento na análise da influência da renda nas entregas, visto que esta variável apresentou efeito não observado na literatura para as cidades de Uberlândia e Patos de Minas, bem como da variável idade, que apresentou resultados divergentes para as cidades analisadas. Sugere-se também um aprofundamento do estudo por meio de pesquisas de preferência declarada/revelada para entender o efeito das variáveis exploradas neste estudo no comportamento do consumidor. Ainda, estudos semelhantes a este

trabalho no contexto pós pandemia podem ser interessantes para verificar se a pandemia trouxe mudanças permanentes nos fatores que influenciam as entregas do comércio eletrônico.

## 8 REFERÊNCIAS

- Allen, J.; Pieczyk, M.; Piotrowska, M.; McLeod, F.; Cherrett, T.; Ghali, K.; Nguyen, T.; Bektas, T.; Bates, O.; Friday, A.; Wise, S.; Austwick, M. (2018) Understanding the impact of e-commerce on last-mile light goods vehicle activity in urban areas: The case of London. *Transportation Research Part D*, v. 61, p. 325-338. <http://dx.doi.org/10.1016/j.trd.2017.07.020>
- Anderson, W. P.; Chatterjee, L.; Lakshmanan, T. R. (2003) E-commerce, transportation, and economic geography. *Growth and Change*, v. 34, n. 4, p. 415-432. <https://doi.org/10.1046/j.0017-4815.2003.00228.x>
- Beckers, J.; Cárdenas, I.; Verhetsel, A. (2018) Identifying the geography of online shopping adoption in Belgium. *Journal of Retailing and Consumer Services*, v. 45, p. 33-41. <https://doi.org/10.1016/j.jretconser.2018.08.006>
- Bellman, S.; Lohse, G. L.; Johnson, E. J. (1999) Predictors of online buying behavior: What personal characteristics predict whether or not people buy on the Net? Look for a “wired” lifestyle and time starvation, not demographics. *Communications of the ACM*, v. 42, n. 12, p. 32-38. <https://doi.org/10.1145/322796.322805>
- BHMap (2020) BHMap. Disponível em: <<http://bhmap.pbh.gov.br/v2/mapa/idebhgeo>> Acesso em: 20 de jul. 2020.
- Bivand, R. S.; Pebesma, E.; Gomez-Rubio, V. (2013) *Applied spatial data analysis with R*. New York: Springer.
- Black, G. S. (2007) Consumer demographics and geographics: Determinants of retail success for online auctions. *Journal of Targeting, Measurement and Analysis for Marketing*, v.15, n. 2, p. 93-102. <https://doi.org/10.1057/palgrave.jt.5750035>
- Brashear, T. G.; Kashyap, V.; Musante, M. D.; Donthu, N. (2009) A profile of the internet shopper: evidence from six countries. *Journal of Marketing Theory and Practice*. V. 17, n. 3, p. 267-281. <https://doi.org/10.2753/MTP1069-6679170305>
- Brown, M.; Pope, N.; Voges, K. (2003) Buying or browsing?: An exploration of shopping orientations and online purchase intention. *European Journal of Marketing*, v. 37, n. 11, p. 1666-1684. <http://dx.doi.org/10.1108/03090560310495401>
- Brusch, M.; Stüber, E. (2014) Developments and classifications of online shopping behavior in Germany. *International Journal of Cyber Society and Education*, v. 7, n. 2, p. 147-156. <http://dx.doi.org/10.7903/ijcse.1171>

- Bucko, J.; Kakalejčík, L.; Ferencová, M. (2018) Online shopping: Factors that affect consumer purchasing behaviour. *Cogent Business & Management*, v. 5, 1535751. <http://dx.doi.org/10.1080/23311975.2018.1535751>
- Burkolter, D.; Kluge, A. (2011) Online consumer behavior and its relationship with socio-demographics, shopping orientations, need for emotion, and fashion leadership. *Journal of Business and Media Psychology*, v. 2, n. 2, p. 20-28.
- Cao, X. (2012) The relationships between e-shopping and store shopping in the shopping process of search goods. *Transportation Research Part A*, v. 46, n. 7, p. 993-1002. <https://doi.org/10.1016/j.tra.2012.04.007>
- Cao, X.; Chen, Q.; Choo, S. (2013) Geographic distribution of e-shopping: Application of structural equation models in the Twin Cities of Minnesota. *Transportation Research Record: Journal of the Transportation Research Board*, v. 2383, n. 1, p. 18-26. <https://doi.org/10.3141/2383-03>
- Cardenas, I.; Beckers, J.; Vanelslander, T.; Verhetsel, A.; Dewulf, W. (2016). In: *6th International Conference on Information Systems, Logistics and Supply Chain*, Bordeaux.
- Cheng, C.; Sakai, T.; Alho, A.; Cheah, L.; Ben-Akiva, M. (2021) Exploring the relationship between locational and household characteristics and e-commerce home delivery demand. *Logistics*, v. 5, n. 2, p. 29. <https://doi.org/10.3390/logistics5020029>
- Clarke, G.; Thompson, C.; Birkin, M. (2015) The emerging geography of e-commerce in British retailing. *Regional Studies, Regional Science*, v. 2, n. 1, p. 371-391. <https://doi.org/10.1080/21681376.2015.1054420>
- De Blasio, G. (2008) Urban–rural differences in internet usage, e-commerce, and e-banking: Evidence from Italy. *Growth and Change*, v. 39, n. 2, p. 341-367. <https://doi.org/10.1111/j.1468-2257.2008.00422.x>
- E-Bit. (2020) Webshoppers 41<sup>a</sup> edição. Disponível em: <https://www.ebit.com.br/webshoppers> Acesso em: 07 de jun. 2021.
- E-Bit. (2021) Webshoppers 43<sup>a</sup> edição. Disponível em: <https://www.ebit.com.br/webshoppers> Acesso em: 07 de jun. 2021.
- Eurostat (2021) E-commerce statistics for individuals. Disponível em: [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=E-commerce\\_statistics\\_for\\_individuals](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=E-commerce_statistics_for_individuals) Acesso em: 30 de dez. 2021.
- Farag, S.; Dijst, M.; Lanzendorf, M. (2003) Exploring the use of e-shopping and its impact on personal travel behavior in the Netherlands. *Transportation Research Record*, v. 1858, n. 1; p. 47-54. <https://doi.org/10.3141/1858-07>

- Farag, S.; Schwanen, T.; Dijst, M. (2005) Empirical Investigation of Online Searching and Buying and Their Relationship to Shopping Trips. *Transportation Research Record: Journal of the Transportation Research Board*, v. 1926, n. 1, p. 242-251. <https://doi.org/10.1177/0361198105192600128>
- Farag, S.; Krizek, K. J.; Dijst, M. (2006a) E-shopping and its relationship with in-store shopping: empirical evidence from the Netherlands and the USA. *Transport Reviews*, v. 26, n. 1, p. 43-61. <https://doi.org/10.1080/01441640500158496>
- Farag, S.; Weltevreden, J.; Rietbergen, T.; Dijst, M.; Oort, F. (2006b) E-shopping in the Netherlands: Does geography matter? *Environment and Planning B: Planning and Design*, v. 33, n. 1, p. 59-74. <https://doi.org/10.1068/b31083>
- Farag, S.; Schwanen, T.; Dijst, M.; Faber, J. (2007) Shopping online and/or in-store? A structural equation model of the relationships between e-shopping and in-store shopping. *Transportation Research Part A*, v. 41, n. 2, p. 125-141. <https://doi.org/10.1016/j.tra.2006.02.003>
- Fernández-Bonilla, F. (2020) E-commerce: Determining factors and the importance of the e-trust. In: *23rd Biennial Conference 2020 of the International Telecommunications Society*, Calgary.
- Fotheringham, A. S.; Brunson, C.; Charlton, M. (2002) *Geographically Weighted Regression - the analysis of spatially varying relationships*. West Sussex: Wiley.
- Fox, J.; Weisberg, S. (2019) *An R Companion to Applied Regression*. Thousand Oaks: Sage.
- Gevaers, R.; Van de Voorde, E.; Vanellander, T. (2014) Cost modelling and simulation of last-mile characteristics in an innovative B2C supply chain environment with implications on urban areas and cities. *Procedia - Social and Behavioral Sciences*, v. 125, p. 398-411. <https://doi.org/10.1016/j.sbspro.2014.01.1483>
- Gollini, I.; Lu, B.; Charlton, M.; Brunson, C.; Harris, P. (2015) GWmodel: An R Package for Exploring Spatial Heterogeneity Using Geographically Weighted Models. *Journal of Statistical Software*, v. 63, n. 17, p. 1–50. <https://doi.org/10.18637/jss.v063.i17>
- Grekousis G. (2020) *Spatial Analysis Methods and Practice: Describe, Explore, Explain through GIS*. Cambridge: Cambridge University Press.
- Hägerstrand, T. (1967) *Innovation Diffusion as a Spatial Process*. Chicago: University of Chicago Press.
- Hair, J. F.; Black, W. C.; Babin, B. J.; Anderson, R. E. (2019) *Multivariate Data Analysis*. Hampshire: Cengage.

- Harris, P.; Riley, F. D.; Riley, D.; Hand, C. (2017) Online and store patronage: a typology of grocery shoppers. *International Journal of Retail & Distribution Management*, v. 45, n. 4, p. 419-445. <https://doi.org/10.1108/IJRDM-06-2016-0103>
- Hashin, A.; Ghani, E. K.; Said, J. (2009) Does consumers' demographic profile influence online shopping?: An examination using Fishbein's Theory. *Canadian Social Science*, v. 5, n. 6, p. 19-31. <http://doi.org/10.3968/j.css.1923669720090506.003>
- Hastie, T.; Tibshirani, R.; Friedman, J. (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer.
- Hood, N.; Urquhart, R.; Newing, A.; Heppenstall, A. (2020) Sociodemographic and spatial disaggregation of e-commerce channel use in the grocery market in Great Britain. *Journal of Retailing and Consumer Services*, v. 55, 102076. <https://doi.org/10.1016/j.jretconser.2020.102076>
- IBGE (2010) Sistema IBGE de Recuperação Automática – Censo demográfico 2010. Disponível em: <<https://sidra.ibge.gov.br/pesquisa/censo-demografico/demografico-2010/universo-caracteristicas-da-populacao-e-dos-domicilios>> Acesso em: 20 de jul. 2020.
- IBGE (2020a) *Região de Influência das Cidades 2018*. Rio de Janeiro: IBGE.
- IBGE (2020b) Belo Horizonte. Disponível em: <<https://cidades.ibge.gov.br/brasil/mg/belo-horizonte/panorama>> Acesso em: 12 de jun. 2021.
- IBGE (2020c) Uberlândia. Disponível em: <<https://cidades.ibge.gov.br/brasil/mg/uberlandia/panorama>> Acesso em: 12 de jun. 2021.
- IBGE (2020d) Patos de Minas. Disponível em: <<https://cidades.ibge.gov.br/brasil/mg/patos-de-minas/panorama>> Acesso em: 12 de jun. 2021.
- Jaller, M.; Pahwa, A. (2020) Evaluating the environmental impacts of online shopping: A behavioral and transportation approach. *Transportation Research Part D: Transport and Environment*, v. 80, 102223. <https://doi.org/10.1016/j.trd.2020.102223>
- James, G.; Witten, D.; Hastie, T.; Tibshirani, R. (2015) *An Introduction to Statistical Learning: With Applications in R*. New York: Springer.
- Jongen, W. (2018) *The End of Online Shopping: The future of retail in an always connected world*. World Scientific Publishing Company.
- Keng Kau, A.; Tang, Y. E.; Ghose, S. (2003) Typology of online shoppers. *Journal of Consumer Marketing*, v. 20; n. 2; p. 139-156. <https://doi.org/10.1108/07363760310464604>
- Kirby-Hawkins, E.; Birkin, M.; Clarke, G. (2018) An investigation into the geography of corporate e-commerce sales in the UK grocery market. *Environment and Planning B:*



- Urban Analytics and City Science*, v. 46, n. 6, p. 1148-1164.  
<https://doi.org/10.1177/2399808318755147>
- Kleiber, C.; Zeileis, A. (2021) AER - Applied Econometrics with R. Disponível em:  
 <<https://cran.r-project.org/web/packages/AER/AER.pdf>> Acesso em 30 nov. 2021.
- Krizek, K. J.; Li, Y.; Handy, S. L. (2005) Spatial attributes and patterns of use in household-related information and communications technology activity. *Transportation Research Record: Journal of the Transportation Research Board*, v. 1926, n. 1, p. 252-259.  
<https://doi.org/10.1177/0361198105192600129>
- Lee, R. J.; Sener, I. N.; Handy, S. L. (2015) Picture of online shoppers: specific focus on Davis, California. *Transportation Research Record: Journal of the Transportation Research Board*, v. 2496, n. 1, p. 55-63. <https://doi.org/10.3141/2496-07>
- Lee, R. J.; Sener, I. N.; Mokhtarian, P. L.; Handy, S. L. (2017) Relationships between the online and in-store shopping frequency of Davis, California residents. *Transportation Research Part A*, v. 100, p. 40-52. <http://dx.doi.org/10.1016/j.tra.2017.03.001>
- Lennon, S. J.; Kim, M.; Johnson, K. K. P.; Jolly, L. D.; Damhorst, L. M.; Jasper, C. R. (2007) A longitudinal look at rural consumer adoption of online shopping. *Psychology & Marketing*, v. 24, n. 4, p. 375-401. <https://doi.org/10.1002/mar.20165>
- Lewis, P. H. (1994). Attention shoppers: Internet is open. Disponível em:  
 <<https://www.nytimes.com/1994/08/12/business/attention-shoppers-internet-is-open.html>> Acesso em: 04 de jun. 2021.
- Li, H.; Kuo, C.; Rusell, M. G. (1999) The impact of perceived channel utilities, shopping orientations, and demographics on the consumer's online buying behavior. *Journal of Computer-Mediated Communication*, v. 5, n. 2, JCMC521. <https://doi.org/10.1111/j.1083-6101.1999.tb00336.x>
- Loo, B. P. Y.; Wang, B. (2017) Factors associated with home-based e-working and e-shopping in Nanjing, China. *Transportation*, v. 45, p. 365-384. <https://doi.org/10.1007/s11116-017-9792-0>
- Lubis, A. N. (2018). Evaluating the customer preferences of online shopping: Demographic factors and online shop application issue. *Academy of Strategic Management Journal*, v. 17, n. 2, p. 1-13.
- Lu, B.; Charlton, M.; Harris, P.; Fotheringham, A. S. (2014) Geographically weighted regression with a non-Euclidean distance metric: a case study using hedonic house price data. *International Journal of Geographical Information Science*, v. 28, n. 4, p. 660-681.  
<https://doi.org/10.1080/13658816.2013.865739>

- Maan, S.; Kumar, S. (2020) Exploring demographic differences in online purchase frequency. *Mukt Shabd Journal*, v. 9, n. 6, p. 4845-4855.
- Mokhtarian, P. L.; Ory, D. T.; Cao, X. (2009) Shopping-related attitudes: a factor and cluster analysis of Northern California shoppers. *Environment and Planning B: Planning and Design*, v. 36, n. 3, p. 204-228. <https://doi.org/10.1068/b34015t>
- Moran, P. A. P. (1950) Notes on Continuous Stochastic Phenomena. *Biometrika*, v. 37, n. 1-2, p. 17-23.
- Morganosky, M. A.; Cude, B. J. (2000) Consumer response to online grocery shopping. *International Journal of Retail & Distribution Management*, v. 28, n. 1, p. 17-26. <https://doi.org/10.1108/09590550010306737>
- Moroz, M.; Polkowski, Z. (2016) The last mile issue and urban logistics: choosing parcel machines in the context of the ecological attitudes of the Y generation consumers purchasing online. *Transportation Research Procedia*, v. 16, p. 378-393. <https://doi.org/10.1016/j.trpro.2016.11.036>
- Motte-Baumvol, B.; Belton-Chevallier, L.; Dablanc, L.; Morganti, E.; Belin-Munier, C. (2017) Spatial dimensions of e-shopping in France. *Asian Transport Studies*, v. 4, n. 3, p. 585-600. <https://doi.org/10.11175/eastsats.4.585>
- Nguyen, D. H.; de Leeuw, S.; Dullaert, W.; Foubert, B. P. J. (2019) What is the right delivery option for you? Consumer preferences for delivery attributes in online retailing. *Journal of Business Logistics*, v. 40, n. 4, p. 299-321. <https://doi.org/10.1111/jbl.12210>
- OpenStreetMap (2021) OpenStreetMap. Disponível em: <<https://www.openstreetmap.org/>> Acesso em: 18 de maio 2021.
- Opesade, A. O. (2020) Discovering patterns in electronic commerce diffusion in Nigeria using google trends web data. *Journal of Information Science, Systems and Technology*, v. 4, n. 1, p. 1-20.
- Pareto, V.; Bousquet, G. H.; Busino, G. (1964) *Cours d'Économie Politique: Nouvelle édition*. Geneva: Librairie Droz.
- Park, H.; Park, D.; Jeong, I. (2016) An effects analysis of logistics collaboration in last-mile networks for CEP delivery services. *Transport Policy*, v. 50, p. 115-125. <http://dx.doi.org/10.1016/j.tranpol.2016.05.009>
- Pérez-Amaral, T. Valarezo, A.; López, R.; Garín-Muñoz, T.; Herguera, I. (2020) E-commerce by individuals in Spain using panel data 2008–2016. *Telecommunications Policy*, v. 44, n; 4, 101888. <https://doi.org/10.1016/j.telpol.2019.101888>

- Pérez-Hernández, J.; Sánchez-Mangas, R. (2011) To have or not to have Internet at home: Implications for online shopping. *Information Economics and Policy*, v. 23, n. 3-4, p. 213-226. <https://doi.org/10.1016/j.infoecopol.2011.03.003>
- Rahman, M. A.; Islam, A. M.; Esha, B. H.; Sultana, N.; Chakravorty, S. (2018) Consumer buying behavior towards online shopping: An empirical study on Dhaka city, Bangladesh. *Cogent Business & Management*, v. 5, n. 1, 1514940. <https://doi.org/10.1080/23311975.2018.1514940>
- Ren, F.; Kwan, M. (2009) The impact of geographic context on e-shopping behavior. *Environment and Planning B: Planning and Design*, v. 36, n.2 , p. 262-278. <https://doi.org/10.1068/b34014t>
- Revelle, W. (2021) psych: Procedures for Psychological, Psychometric, and Personality Research. Disponível em: <<https://CRAN.R-project.org/package=psych>> Acesso em 06 ago. 2021.
- Ripley, B.; Venables, B.; Bates, D. M.; Hornik, K.; Gebhardt, A.; Firth, D. (2021) MASS - Support Functions and Datasets for Venables and Ripley's Modern Applied Statistics with S. Disponível em: <<https://cran.r-project.org/web/packages/MASS/MASS.pdf>> Acesso em 30 nov. 2021.
- Rogers, E. M. (1983) *Diffusion of Innovations*. New York: The Free Press.
- Saphores, J.; Xu, L. (2020) E-shopping changes and the state of e-grocery shopping in the US - Evidence from national travel and time use surveys. *Research in Transportation Economics*, v. 87, 100864. <https://doi.org/10.1016/j.retrec.2020.100864>
- Schaefer, K C.; Visser, M. L. (2003) Reverse Regression and Orthogonal Regression in Employment Discrimination Analysis. *Journal of Forensic Economics*, v. 16, n. 3, p. 283-298. <https://doi.org/10.5085/0898-5510-16.3.283>
- Seitz, C., Pokrivčák, J., Tóth, M., & Plevný, M. (2017) Online grocery retailing in Germany: an explorative analysis. *Journal of Business Economics and Management*, v. 18, n. 6, p. 1243-1263. <https://doi.org/10.3846/16111699.2017.1410218>
- Sener, I. N.; Reeder, P. R.; An examination of behavioral linkages across ICT choice dimensions: copula modeling of telecommuting and teleshopping choice behavior. *Environment and Planning A*, v. 44, n. 6, p. 1459-1478. <https://doi.org/10.1068/a44436>
- Sim, L. L.; Koi, S. M. (2002) Singapore's Internet shoppers and their impact on traditional shopping patterns. *Journal of Retailing and Consumer Services*, v. 9, n. 2, p. 115-124. [https://doi.org/10.1016/S0969-6989\(01\)00029-7](https://doi.org/10.1016/S0969-6989(01)00029-7)

- Song, Z. (2021) The geography of online shopping in China and its key drivers. *Environment and Planning B: Urban Analytics and City Science*. <https://doi.org/10.1177/23998083211002189>
- Soopramanien, D. G. R.; Robertson, A. (2007) Adoption and usage of online shopping: An empirical analysis of the characteristics of “buyers” “browsers” and “non-internet shoppers”. *Journal of Retailing and Consumer Services*, v. 14, n.1, p. 73-82. <https://doi.org/10.1016/j.jretconser.2006.04.002>
- Stafford, T. F.; Turan, A.; Raisinghani, M. S. (2004) International and cross-cultural influences on online shopping behavior. *Journal of Global Information Technology Management*, v. 7, n. 2, p. 70-87. <https://doi.org/10.1080/1097198X.2004.10856373>
- Statista (2020) U.S. digital buyer distribution 2020, by age group. Disponível em: <https://www.statista.com/statistics/469184/us-digital-buyer-share-age-group/> Acesso em: 30 de dez. 2021.
- Statista (2021) Brazil: age distribution of digital buyers 2020. Disponível em: <https://www.statista.com/statistics/519746/age-distribution-of-online-shoppers-in-brazil/#:~:text=Brazil%3A%20age%20distribution%20of%20digital%20buyers%202020&text=In%202020%2C%20more%20than%20one,from%2031.8%20percent%20in%202019.>> Acesso em: 30 de dez. 2021.
- Swinyard, W. R.; Smith, S. M. (2003) Why people (don't) shop online: A lifestyle study of the internet consumer. *Psychology & Marketing*, v. 20, n. 7, p. 567-597. <https://doi.org/10.1002/mar.10087>
- Tadić, S.; Veljović, M. (2020) Home delivery: A framework for structuring. *International Journal for Traffic and Transport Engineering*, v. 11, n. 1, p. 30-74. [http://dx.doi.org/10.7708/ijtte.2021.11\(1\).03](http://dx.doi.org/10.7708/ijtte.2021.11(1).03)
- Taniguchi, E.; Thompson, R. G.; Yamada, T. (2014) Recent trends and innovations in modelling city logistics. *Procedia - Social and Behavioral Sciences*, v. 125, p. 4-14. <https://doi.org/10.1016/j.sbspro.2014.01.1451>
- UNCTAD. (2019) UNCTAD B2C E-commerce Index 2019. Disponível em: [https://unctad.org/system/files/official-document/tn\\_unctad\\_ict4d14\\_en.pdf](https://unctad.org/system/files/official-document/tn_unctad_ict4d14_en.pdf) Acesso em: 7 de jun. 2021.
- UNCTAD. (2021) Estimates of Global E-commerce 2019 and Preliminary Assessment of Covid-19 Impact on Online Retail 2020. Disponível em: [https://unctad.org/system/files/official-document/tn\\_unctad\\_ict4d18\\_en.pdf](https://unctad.org/system/files/official-document/tn_unctad_ict4d18_en.pdf) Acesso em: 7 de jun. 2021.

- Valarezo, A.; López, R.; Amaral, T. P. (2020) Adoption of e-commerce by individuals and digital-divide. In: Alleman, J.; Rappoport, P. N.; Hamoudia, M. (eds.). *Applied Economics in the Digital Era*. Cham: Palgrave Macmillan. Cap. 4, p. 103-134.
- Van Droogenbroeck, E.; Van Hove, L. (2017) adoption of online grocery shopping: Personal or household characteristics?. *Journal of Internet Commerce*, v. 16, n. 3, p. 255-286. <https://doi.org/10.1080/15332861.2017.1317149>
- Vrechopoulos, A. P.; Siomkos, G. J.; Doukidis, G. I. (2001) Internet shopping adoption by Greek consumers. *European Journal of Innovation Management*, v. 4, n. 3, p. 142-152.
- Washington, S. P.; Karlaftis, M. G.; Mannering, F. L. (2010) *Statistical and econometric methods for transportation data analysis*. Boca Raton: Chapman & Hall/CRC.
- Weltevreden, J. W. J. (2007) Substitution or complementarity? How the Internet changes city centre shopping. *Journal of Retailing and Consumer Services*, v. 14, n. 3, p. 192-207. <https://doi.org/10.1016/j.jretconser.2006.09.001>
- Zhou, L.; Dai, L.; Zhang, D. (2007) Online shopping acceptance model - A critical survey of consumer factors in online shopping. *Journal of Electronic Commerce Research*, v. 8, n. 1, p. 41-62. <https://doi.org/10.1108/14601060110399306>
- Zhou, Y.; Wang, X (2014) Explore the relationship between online shopping and shopping trips: An analysis with the 2009 NHTS data. *Transportation Research Part A*, v. 70, p. 1-9. <https://doi.org/10.1016/j.tra.2014.09.014>