



Universidade de Brasília

Faculdade de Economia, Administração, Contabilidade e Gestão de Políticas Públicas

Departamento de Administração

MANOEL LUCIANO DE CASTRO NETO

**DISTRIBUIÇÃO DE INSTALAÇÕES LOGÍSTICAS E INFRA-
ESTRUTURA DE TRANSPORTE: Um estudo espacial do
Distrito Federal**

Brasília – DF

2023

MANOEL LUCIANO DE CASTRO NETO

**DISTRIBUIÇÃO DE INSTALAÇÕES LOGÍSTICAS E INFRA-
ESTRUTURA DE TRANSPORTE: Um estudo espacial do
Distrito Federal**

Monografia apresentada ao Departamento de Administração como requisito parcial à obtenção do título de Bacharel em Administração.

Professora Orientadora: Dra. Vanessa Cabral Gomes

Brasília – DF

2023

MANOEL LUCIANO DE CASTRO NETO

**DISTRIBUIÇÃO DE INSTALAÇÕES LOGÍSTICAS E INFRA-
ESTRUTURA DE TRANSPORTE: Um estudo espacial do
Distrito Federal**

A Comissão Examinadora, abaixo identificada, aprova o Trabalho de
Conclusão do Curso de Administração da Universidade de Brasília do
aluno

Manoel Luciano de Castro Neto

Profa. Dra., Vanessa Cabral Gomes

Professora-Orientadora

Prof. Dr., Victor Rafael Rezende
Celestino,

Professor-Examinador

Prof. Me., Ricardo Favaretto Garbin

Professor-Examinador

Brasília, 20 de dezembro de 2023

RESUMO

O transporte é uma das principais atividades na logística, sendo essencial para o funcionamento de atividades comerciais. A infraestrutura física é um dos componentes da estrutura de transporte de carga, dando suporte à essa atividade e exercendo influência na localização de instalações logísticas, como armazéns e instalações de transporte de mercadorias. Nesse contexto, o objetivo geral deste trabalho foi analisar a distribuição espacial das instalações logísticas do Distrito Federal (DF) em relação à disponibilidade de infraestrutura de transporte na região. Trata-se de uma pesquisa descritiva e de natureza quantitativa, com dados coletados de fontes secundárias disponíveis online em bases de dados públicas. Os procedimentos para análise de dados envolveram a utilização de gráficos, mapas temáticos e o cálculo de métricas de distância e medidas de associação espacial entre variáveis. Os principais resultados apontam que os estabelecimentos logísticos se concentram no sudoeste do DF. Há uma autocorrelação espacial positiva e estatisticamente significativa para a densidade de empresas (empresas/km²) das regiões administrativas (RA) do DF. Além disso, a maioria das empresas está localizada próximo a uma rodovia, o que sugere que há influência da proximidade das rodovias na distribuição espacial das empresas. No entanto, a correlação espacial entre a densidade de empresas (empresas/km²) e densidade de rede viária (km/km²) das RA do DF não alcançou significância estatística, isto é, não há evidências suficientes para rejeitar a hipótese nula de aleatoriedade espacial.

Palavras-chave: Análise espacial. Instalações logísticas. Infraestrutura de transporte. Distrito Federal.

LISTA DE ILUSTRAÇÕES

Figura 1 – Localização das empresas no DF	34
Figura 2 – Centro médio e SDE	36
Figura 3 – Regiões de localização das empresas da SDE.....	38
Figura 4 – Densidade de empresas por RA	40
Figura 5 – <i>Spatial Lags</i> das densidades de empresas das RA	41
Figura 6 – Diagrama de Dispersão de Moran para a variável Densidade de empresas	42
Figura 7 – Mapa de identificação de <i>clusters</i> para a variável densidade de empresas	43
Figura 8 – Sistema viário do DF	45
Figura 9 Densidade da rede viária (km/km ²)	48
Figura 10 – Rodovias, <i>Buffers</i> e empresas	49
Figura 11 – Diagrama de Dispersão de Moran para as variáveis Densidade de empresas e Densidade da rede viária	51
Figura 12 – Mapa de identificação de <i>clusters</i> e regiões isoladas para as variáveis Densidade da rede viária e Densidade de empresas	52
Figura 13 – Empresas e zoneamento	53

LISTA DE TABELAS

Tabela 1 – Localização das empresas por RA	34
Tabela 2 – Localização das empresas dentro da SDE por RA.....	36
Tabela 3 – Área e Densidade de empresas das RA	39
Tabela 4 – Extensão de rodovias em km por RA	45
Tabela 5 – Rodovias, área e densidade da rede viária	46
Tabela 6 – Empresas localizadas a até 1 km de uma rodovia	49
Tabela 7 <i>Clusters</i> e regiões isoladas para variáveis densidade de rede viária e densidade de empresas	52
Tabela 8 – Empresas por zona	54

LISTA DE QUADROS

Quadro 1 – Objetos espaciais e variáveis de análise.....	25
Quadro 2 – Classificações CNAE.....	26

SUMÁRIO

1. INTRODUÇÃO	10
1.1. Contextualização.....	10
1.2. Formulação do problema	11
1.3. Objetivo Geral	11
1.4. Objetivos Específicos	11
1.5. Justificativa.....	12
2. REFERENCIAL TEÓRICO.....	13
2.1. Infraestrutura de transportes	13
2.2. Armazenamento e distribuição.....	15
2.3. Padrões espaciais das instalações logísticas	16
2.4. Políticas de planejamento, uso do solo e zoneamento	19
2.4.1. Planejamento urbano do Distrito Federal	20
3. MÉTODOS E TÉCNICAS DE PESQUISA	22
3.1. Tipo e descrição geral da pesquisa.....	22
3.2. Conceitos de análise espacial.....	23
3.3. Objetos espaciais e variáveis de análise.....	24
3.4. Procedimentos de coleta de dados	25
3.4.1. Empresas	26
3.4.2. Regiões Administrativas e Zoneamento.....	27
3.4.3. Sistema viário.....	28
3.5. Procedimentos de análise de dados	29
3.5.1. Métricas de distância.....	29
3.5.2. Medidas de associação espacial.....	30
3.6. Ferramentas utilizadas	32
4. RESULTADOS E DISCUSSÃO	33
4.1. Mapeamento das instalações logísticas no Distrito Federal.....	33
4.2. Distribuição espacial das instalações logísticas em diferentes regiões administrativas	39
4.3. Disponibilidade de infraestrutura de transporte no Distrito Federal.....	44
4.4. Relação entre a localização das instalações e infraestrutura de transporte.....	48

4.5. Influência da política de zoneamento do DF na localização das instalações logísticas	53
4.6. Discussão.....	54
5. CONCLUSÕES E RECOMENDAÇÕES	55
REFERÊNCIA	58
APÊNDICE	63

1. INTRODUÇÃO

1.1. Contextualização

A infraestrutura de transportes é um fator essencial para o crescimento econômico de uma região (Kin; Verlinde; Macharis, 2017; Vieira *et al.*, 2021a). Na logística empresarial, o transporte é uma atividade-primária, que envolve, por exemplo, a determinação dos modais de transporte a serem utilizados, dos roteiros e a programação de veículos, sendo um dos principais componentes dos custos logísticos (Ballou, 2009). Nas áreas urbanas, que concentram a maior parte do consumo, o transporte de cargas é fundamental para a manutenção da atividade econômica, mas traz também impactos negativos como a emissão de gases poluentes, o congestionamento do tráfego e acidentes de trânsito (Roumboutsos; Kapros; Vanellander, 2014).

No Brasil, os principais modais de transporte são o rodoviário e ferroviário (Brasil, 2021). A extensão da malha rodoviária é de 1.720.909,0 km, com uma frota de 111.446.870 veículos, e o transporte de cargas é realizado por 291.134 empresas, 556 cooperativas e 917.742 autônomos (CNT, 2022). O transporte ferroviário movimentou 1,21 bilhão de toneladas de cargas em 2021, enquanto no transporte aeroviário foram transportadas cerca de 1,4 bilhão de toneladas de carga paga e correios, em voos domésticos e internacionais. O manuseio e reorganização das cargas ao longo do processo de transporte é realizado em diversos terminais de carga espalhados pelo país, entre *gateways* (terminais de entrada e de saída de mercadorias), terminais domésticos e portos secos (CNT, 2021).

No Distrito Federal, a extensão da rede de estradas e vias é de 13.780,40 Km, 82,71% de vias urbanas/rurais e 17,29% de rodovias, sendo o modal rodoviário o mais utilizado para transporte de mercadorias. O modal ferroviário, apesar de ser mais eficiente ambientalmente (Dantas Gabriele *et al.*, 2013), é pouco utilizado na região, representando menos de 1% do tráfego de carga, enquanto o transporte aéreo representa 1,12% (Distrito Federal, 2018).

Devido à sua importância econômica e estratégica, além dos impactos sociais e ambientais, conhecer a infraestrutura de transportes é um passo importante para um bom planejamento e implantação de políticas públicas de mobilidade urbana

(Cuchava Rocha; Gouvea Campos; De Mello Bandeira, 2019; De Castro, 2013). Para as empresas, o conhecimento das vias e modais de transporte e terminais de carga disponíveis, além de fatores geográficos, espaciais e de uso de terra que impactam as atividades de transporte, é imprescindível no gerenciamento da cadeia de suprimentos (Allen; Browne; Cherrett, 2012).

1.2. Formulação do problema

Nos centros urbanos, o transporte de carga desempenha um papel fundamental no abastecimento e no atendimento às demandas de consumo, sendo uma das atividades com maior visibilidade na logística de suprimentos e fundamental para a manutenção da atividade econômica local. Sendo assim, a infraestrutura de transportes das cidades é um dos fatores que pode influenciar o padrão locacional das instalações logísticas, visto que a disponibilidade de vias urbanas e rodovias, modais de transporte, terminais de carga, etc. são essenciais para o exercício das atividades logísticas.

Diante disso, é necessário que se responda: Qual é a relação entre a distribuição espacial das instalações logísticas no Distrito Federal e a disponibilidade de infraestrutura de transporte na região?

1.3. Objetivo Geral

Analisar a distribuição espacial das instalações logísticas do Distrito Federal em relação à disponibilidade de infraestrutura de transporte na região.

1.4. Objetivos Específicos

- Mapear as instalações logísticas no Distrito Federal;
- Mapear a distribuição espacial das instalações logísticas em diferentes áreas do Distrito Federal;

- Mapear a disponibilidade de infraestrutura de transporte na região;
- Investigar a relação entre a localização das instalações e a infraestrutura de transporte existente;
- Investigar o impacto da política de zoneamento do Distrito Federal na localização das instalações logísticas.

1.5. Justificativa

Nos grandes centros urbanos, para atender à demanda dos consumidores, o transporte de carga é intenso. Em regiões como a do Distrito Federal, altamente urbanizada e com intensa circulação de pessoas, a infraestrutura de transportes é fundamental para o crescimento econômico e impacta significativamente o ambiente e a qualidade de vida da população urbana (Vieira et al., 2021). A análise dessa infraestrutura é uma das principais atividades envolvidas no planejamento urbano e de transportes das entidades governamentais e no gerenciamento de suprimentos das organizações.

A infraestrutura de transportes exerce importante influência na localização de instalações logísticas, como armazéns, centros de distribuição e instalações de transporte de mercadorias. Os modais de transporte e terminais de carga disponíveis, acesso a rodovias interestaduais e a proximidade aos locais de origem e destino de remessas são alguns dos fatores que influenciam os padrões locacionais das instalações logísticas nos centros urbanos (Allen; Browne; Cherrett, 2012; Giuliano; Kang, 2018; Guerin et al., 2021). Além disso, as políticas de planejamento urbano, como o zoneamento, também exercem um papel relevante no padrão locacional dessas instalações (Sakai; Beziat; Heitz, 2020).

Uma análise dos padrões locacionais das instalações logísticas no Distrito Federal pode fornecer *insights* importantes para as empresas e o poder público. Além disso, permite expandir esse campo de estudo, olhando para a capital federal de um país emergente.

2. REFERENCIAL TEÓRICO

Este capítulo está dividido em quatro partes: Infraestrutura de transportes, Armazenamento e distribuição, Padrões espaciais das instalações logísticas e Planejamento urbano do Distrito Federal. Na primeira, discute-se o transporte como atividade logística e a infraestrutura que dá suporte a essa atividade. Na segunda, discute-se o armazenamento e distribuição na logística. A terceira parte aborda os padrões espaciais de instalações logísticas, definindo-as e apresentando estudos que analisam fatores de influência na sua localização. Por fim, na quarta parte é feita uma introdução ao planejamento urbano do Distrito Federal.

2.1. Infraestrutura de transportes

Segundo Bowersox et al. (2013), os transportes são responsáveis por cerca de dois terços dos custos logísticos, representando a maior despesa logística de uma empresa típica. Na cadeia de suprimentos, o transporte é o responsável pela movimentação das mercadorias e matérias-primas para destinos específicos e é essencial para o funcionamento das atividades comerciais, pois, sem uma gestão de transporte eficaz e eficiente, as compras, a fabricação e o atendimento ao cliente são comprometidos. Além disso, Bowersox et al. (2013) apresenta o armazenamento de produtos como um dos serviços principais prestado pelas empresas de transporte, visto que um veículo pode ser utilizado temporariamente para armazenamento nos pontos de origem ou destino do carregamento, o que pode reduzir custos.

De acordo com Bowersox et al. (2013) a estrutura de transporte de carga é composta por infraestrutura física, veículos e transportadoras, que operam dentro de 5 modais básicos: ferroviário, rodoviário, hidroviário, dutoviário e aéreo. Para atender melhor às necessidades do cliente e reduzir custos, as transportadoras podem utilizar uma combinação desses modais, realizando o transporte intermodal. Alguns sistemas intermodais utilizados são: *piggyback* (integra rodovias e ferrovias), carreta sobre vagão plataforma (TOFC – *Trailer on a flatcar*) e o contêiner sobre vagão plataforma (COFC – *Container on a flatcar*).

Para Rodrigue (2020), as infraestruturas são os bens de capital que dão suporte às atividades dos indivíduos, instituições e corporações na sociedade, satisfazendo, por exemplo, necessidades de transporte, mobilidade e geração de energia. De acordo com Prud'homme (2004), a infraestrutura associada aos serviços de transporte inclui estradas, pontes, túneis, trilhos ferroviários, portos, aeroportos e centros de distribuição, etc.

Rodrigue (2020) caracteriza o impacto econômico do setor de transporte em três tipos: *core* (fundamental), operacional e geográfico. O *core* se refere à infraestrutura física e aos custos de transporte, que estão ligados ao nível de produção e a economias de escala. Já o operacional diz respeito aos ganhos de tempo e à confiabilidade, que permitem melhorias na gestão de estoques e no uso dos ativos de transporte. Por fim, o geográfico está ligado à acessibilidade e localização, isto é, à capacidade de acesso de uma economia ao mercado consumidor e aos *inputs* de produção, e à proximidade das redes de transporte em relação aos terminais modais e intermodais.

Percebe-se que, no nível macroeconômico, o transporte está ligado à produção, à empregabilidade e à renda, enquanto no nível microeconômico, o transporte está ligado às atividades de produção, ao consumidor e aos custos de distribuição. O transporte é o responsável por ligar os fatores de produção em uma rede de relacionamentos entre produtores e consumidores (Rodrigue, 2020).

Zhang e Cheng (2023) investigaram a relação entre a infraestrutura de transporte e crescimento econômico no Reino Unido e observaram que o estoque de infraestrutura tem impacto positivo significativo no crescimento econômico da região no longo prazo. Segundo François e Manchin (2013) uma infraestrutura de transportes bem desenvolvida é um dos fatores determinantes no volume de transações comerciais realizadas entre países. Outros estudos também abordam a importância do setor de transporte para o comércio e crescimento econômico (Meersman; Nazemzadeh, 2017; Rahman *et al.*, 2021; Şahan; Tuna, 2021; Vlahinić Lenz; Pavlić Skender; Mirko-*vić*, 2018; Wang *et al.*, 2020).

Segundo Erkan (2014), países que melhoram sua performance logística experimentam aumento de produtividade e nível de emprego e redução de pobreza. As atividades logísticas podem conferir vantagem competitiva para as nações, visto que a eficiência logística está ligada à redução de custos de transporte e de produção. Para Erkan (2014), a infraestrutura de transporte é um dos principais determinantes

de eficiência logística, e países com infraestrutura ferroviária e portuária de qualidade conseguem ter melhor performance logística.

De acordo com Park (2020), a qualidade da infraestrutura de transporte e logística é uma fonte de vantagem comparativa para os países, isto é, é um dos fatores determinantes para o que será exportado. Park (2020) ressalta a importância desse resultado para o poder público, que têm papel crucial no fornecimento de infraestrutura de transporte. Os formuladores de políticas públicas devem direcionar esforços para aperfeiçoar os modais de transporte mais intensamente utilizados pelas indústrias de uma região.

2.2. Armazenamento e distribuição

Para Croucher, Baker e Rushton (2014), os depósitos se justificam quando eles fazem parte de cadeia de suprimentos de menor custo que pode ser desenhada para atender melhor às necessidades do cliente. Os depósitos podem estar envolvidos em vários estágios do abastecimento, produção e distribuição dos bens, sendo um dos elementos de maior custo na cadeia de suprimentos e, portanto, sua gestão é crítica tanto em termos de custo quanto de serviço.

Croucher, Baker e Rushton (2014) listam vários tipos de classificações que podem ser adotadas para os depósitos. São elas:

- por estágio na cadeia de suprimentos: materiais, *work-in-progress* ou produtos acabados;
- por área geográfica atendida: mundo, regional (vários países), nacional ou local;
- por tipo de produto: peças pequenas, grandes conjuntos (por exemplo, carrocerias de carros), alimentos congelados, perecíveis, produtos perigosos, etc.;
- por função: por exemplo, manutenção de estoque ou classificação;
- por proprietário: propriedade do usuário (por exemplo, fabricante ou varejista) ou de uma empresa logística terceira;
- por uso da empresa: uso por uma empresa ou várias empresas;
- por área; desde 100 metros quadrados ou menos até mais de 100.000 metros quadrados;

- por altura; desde armazéns com cerca de 3 metros de altura até armazéns *high-bay* que podem ter mais de 45 metros de altura;
- por equipamento: de operações altamente manuais a depósitos altamente automatizados.

Segundo Croucher, Baker e Rushton (2014), o principal objetivo de armazéns ou depósitos é facilitar a movimentação dos bens através da cadeia de suprimentos até o consumidor final. Embora existam várias técnicas e iniciativas que visam reduzir a necessidade de armazenar estoque, isto ainda é necessário especialmente nas seguintes situações: 1) a demanda pelo produto é contínua ou 2) o tempo de espera da oferta é maior que o tempo de espera da demanda, o que pode acontecer, por exemplo, em entregas com tempo de espera de apenas um dia.

Além do armazenamento de estoque, Croucher, Baker e Rushton (2014) listam diversas outras funções que um depósito pode exercer. São elas:

- centro de consolidação: garante que linhas de produtos solicitadas pelos clientes sejam entregues juntas;
- *cross-dock centre*: local onde os bens vindos de outros pontos na cadeia de suprimentos são transferidos diretamente de veículos de entrada para veículos de saída, sem ficarem armazenados;
- centro de triagem: funciona como o *cross-dock centre*, mas as mercadorias são trazidas especificamente para triagem para uma determinada região;
- instalações de montagem: utilizadas como ponto final de montagem dos produtos;
- *trans-shipment point*: utilizado para triagem de mercadorias coletadas de um centro de distribuição nacional e classificadas para veículos menores de entrega imediata aos clientes;
- centro de devolução de mercadorias: utilizado para lidar com as devoluções.

2.3. Padrões espaciais das instalações logísticas

Bowersox et al. (2013) define as instalações de um sistema logístico como todos os locais onde há manuseio ou armazenamento de estoque de materiais, de produtos em processo ou de produtos acabados, o que inclui lojas, depósitos e

fábricas. No projeto logístico da cadeia de suprimentos, as principais decisões que devem ser tomadas pelos gerentes são escolha da quantidade, tamanho e localização das instalações logísticas.

Segundo Croucher, Baker e Rushton (2014), uma vez que os centros de distribuição e armazenamento podem assumir diversas funções, as decisões de localização são extremamente importantes e de natureza complexa. A rede logística precisa minimizar custos ao mesmo tempo em que fornece serviço ao cliente da forma mais eficiente e, portanto, o serviço e o custo são cruciais na determinação da quantidade, tamanho e localização das instalações. Busca-se determinar qual é a melhor combinação de armazenamento e transporte para o alcance do nível de serviço adequado ao cliente.

Para Hesse e Rodrigue (2004), as atividades que compõem a logística podem ser incluídas em duas funções principais: a distribuição física e a gestão de materiais. A distribuição física reúne todas as funções de movimento e manuseio dos bens, o que inclui os serviços de transporte, armazenamento, comércio, atacado e varejo, enquanto a gestão de materiais reúne as atividades relacionadas à produção. Essas funções, no entanto, são interdependentes de modo que se torna difícil analisar produção, distribuição e consumo de maneira isolada. A distribuição física e gestão de materiais são, desse modo, integradas na gestão de cadeia de suprimentos.

De acordo com Bowen (2008), a localização de instalações de armazenamento é fortemente influenciada pelo acesso a redes de transporte aéreo e rodoviário, o que pode ser explicado pela vantagem que esses modais podem ter quando se trata de velocidade, ao menos à nível nacional. Desse modo, regiões com uma melhor infraestrutura aeroviária e rodoviária são as que mais usufruem os benefícios da gestão de cadeia de suprimentos.

Além disso, as dimensões físicas e o tempo também irão influenciar o padrão locacional das instalações de armazenamento. Kang (2020) examinou as decisões de localização de armazéns em Los Angeles, EUA, e observou que os armazéns maiores são mais comuns em regiões com menos acesso ao mercado e menores preços de terra, mas com melhor acesso à aeroportos e terminais intermodais, e que as escolhas quanto à localização das instalações são diferentes quando se analisa diferentes intervalos de tempo.

Em estudo sobre a Região Metropolitana de São Paulo (RMSP), Guerin et al. (2021), utilizando dados sobre empresas do setor logístico em conjunto com softwares

de tratamento e visualização de dados georreferenciados, constataram que os principais fatores que influenciam a localização de armazéns são os menores preços de terra, vantagens fiscais e acesso a infraestrutura rodoviária. Com uma extensa rede de rodovias, o transporte de carga na RMSP é feito majoritariamente por via rodoviária, com menor participação dos modais aeroviário e ferroviário. Desse modo, há uma concentração maior de armazéns em regiões com intersecções de rodovias.

Giuliano e Kang (2018) analisaram os padrões locacionais das quatro maiores regiões metropolitanas da Califórnia, EUA: Los Angeles, San Francisco, San Diego, e Sacramento. No estudo, foram identificados quatro fatores que explicam a dinâmica locacional das instalações logísticas: área metropolitana, políticas de desenvolvimento econômico, estrutura econômica e geografia física. As grandes regiões metropolitanas concentram centros de distribuição nacionais e regionais, além de servirem de *hubs* do comércio global. De acordo com Giuliano e Kang (2018), a localização próxima a rodovias interestaduais podem ser uma boa alternativa a aeroportos e portos, quando estes não estão disponíveis.

A proximidade à infraestrutura de transporte é apontada por Jakubicek (2010) como um dos fatores que influenciam a localização de empresas do setor logístico. O estudo de Jakubicek (2010) analisou a indústria logística dentro do Canadá, o que engloba instalações cuja atividade primária é lida com movimentação de carga. Segundo Jakubicek (2010), a proximidade das instalações logísticas a rodovias é um fator de localização importante para diversos tipos de empresas logísticas.

O estudo de Heitz (2021) sobre instalações logísticas da área metropolitana de Paris reforça a relação entre localização e a infraestrutura de transporte. Segundo Heitz (2021), os municípios com maior taxa de construção de armazéns concentram algumas das maiores e mais importantes infraestruturas de transporte da região. Para realizar sua análise, Heitz (2021) utiliza uma classificação para os municípios de acordo com o perfil urbano e logístico, que leva em conta fatores como população, densidade populacional e características físicas dos armazéns.

Por fim, Heuvel et al. (2013), em um estudo de caso de uma província Holandesa, apontam que existência áreas de concentração logística atraem outras empresas logísticas e, assim, os estabelecimentos logísticos tendem a se localizar em determinadas áreas de uma região, se beneficiando da co-localização. Além disso, Heuvel et al. (2013) apontam que estabelecimentos de transporte são importantes para o surgimento *clusters* espaciais e que, uma vez dentro de um *cluster*, as empresas

logísticas tendem a escolher localizações similares no futuro. Diversos outros estudos abordam a localização de instalações logísticas e os fatores que a influenciam (Allen; Browne; Cherrett, 2012; De Oliveira et al., 2020; Heitz et al., 2020; Woudsma; Jakubicek, 2020).

2.4. Políticas de planejamento, uso do solo e zoneamento

Para Levy (2016), a necessidade de planejamento surge, essencialmente, por causa da interconectividade e complexidade características das cidades. Por exemplo, não é possível separar decisões sobre uso de terra daquelas sobre tráfego, pois uma irá afetar a outra, e ambas irão afetar todas as pessoas que vivem na região. Segundo Levy (2016), as decisões sobre uso territorial afetam a economia local, ajudando a determinar a demanda por serviços ou mão-de-obra, a saúde fiscal da comunidade, as qualidades visuais e sociais da região e até mesmo a saúde das pessoas. Levy (2016) acrescenta que a complexidade surge porque as relações nesse contexto urbano não são simples, o que justifica o planejamento como uma atividade separada no governo.

A regulação sobre uso de solo e zoneamento é um fator chave na localização dos empreendimentos logísticos, conforme observado por Sakai, Beziat e Heitz (2020), em estudo sobre a localização das instalações logísticas da região de Paris, na França. O poder público pode influenciar o desenvolvimento de *clusters*, evitando que atividades comerciais ou industriais intensivas sejam realizadas muito próximas de áreas residenciais, por exemplo. No entanto, Sakai, Beziat e Heitz (2020) apontam que isso pode impedir as empresas de se localizarem de maneira ótima em relação aos seus clientes. Evidentemente, conforme Buldeo Rai et al. (2022), o nível de intervenção estatal é diferente para diversas localidades, com o zoneamento sendo o mecanismo mais comum.

No Brasil, segundo Duarte (2009) todos os aspectos relacionados aos direitos sociais previstos no art. 6º da Constituição Federal de 1988, como, por exemplo, emprego, saúde, segurança, educação e cultura, devem fazer parte do planejamento urbano, visto que todos eles geram impactos ambientais. Desse modo, o planejamento

urbano irá apresentar as seguintes grandes dimensões: econômica, social, ambiental, infraestrutural, gerencial e territorial.

A Constituição Federal estabelece o plano diretor como o instrumento básico da política de desenvolvimento e de expansão urbana, obrigatório para os municípios com mais de 20 mil habitantes (Brasil, 1988). A regulamentação dos arts. 182 e 183 da Constituição Federal, que tratam da política urbana, é feita pela Lei nº 10257, de 2001, denominada Estatuto da Cidade, que estabelece as diretrizes gerais para o planejamento urbano no Brasil, com instrumentos de política urbana nacionais, estaduais e de áreas metropolitanas e municipais, dentre outros. Para os municípios com mais de quinhentos mil habitantes, a Lei nº 10257, de 2001, determina a elaboração de um plano de transporte urbano integrado, que deve ser compatível com o plano diretor ou nele inserido (Brasil, 2001).

2.4.1. Planejamento urbano do Distrito Federal

No Distrito Federal, alguns dos principais normativos relacionados ao planejamento urbano são: 1) o Plano Diretor de Ordenamento Territorial do Distrito Federal – PDOT/DF (Lei Complementar nº 803, de 2009), 2) a Lei de Uso e Ocupação do Solo do Distrito Federal – LUOS/DF (Lei Complementar nº 948, de 2019), e o 3) Código de Edificações do Distrito Federal (Lei Distrital nº 6.138, de 2018). Além disso, a Lei Distrital nº 4.566, de 2011, dispõe sobre o Plano Diretor de Transporte Urbano e Mobilidade do Distrito Federal – PDTU/DF (Distrito Federal, 2009; Distrito Federal, 2019; Distrito Federal, 2011).

Um dos principais instrumentos do PDOT/DF é o zoneamento, que, conforme explicado por Dorneles (2010), define quais as serão as zonas específicas para uso residencial, comercial e industrial. O PDOT/DF estabelece três macrozonas no Distrito Federal: 1) Macrozona Urbana, 2) Macrozona Rural e 3) Macrozona de Proteção Integral, que são definidas de acordo com as vocações intrínsecas às áreas. Desse modo, a Macrozona Rural é destinada, principalmente, às atividades do setor primário da economia, como atividades agropecuárias e extrativas, enquanto a Macrozona Urbana é destinada, principalmente, às atividades dos setores secundário e terciário,

isto é, indústrias diversas, serviços e comércio. Já a Macrozona de Proteção Integral é destinada à preservação da natureza (Distrito Federal, 2009).

As Macrozonas Rurais são divididas em duas zonas: Zona Rural de Uso Diversificado (ZRUD) e Zona Rural de Uso Controlado (ZRUC). Já a Macrozona Urbana é subdividida em seis zonas distintas: Zona Urbana do Conjunto Tombado (ZUCT), Zona Urbana de Uso Controlado I (ZUUC I), Zona Urbana de Uso Controlado II (ZUUC II), Zona Urbana Consolidada (ZUC), Zona Urbana de Expansão e Qualificação (ZUEQ) e Zona de Contenção Urbana (ZCU) (Distrito Federal, 2009).

3. MÉTODOS E TÉCNICAS DE PESQUISA

De acordo com Cervo e Bervian (1996, p. 27), o método de pesquisa é o “conjunto de processos empregados na investigação e na demonstração da verdade”. De maneira semelhante, Laville e Dionne (1999, p. 11) destacam que o método “propõe um procedimento que orienta a pesquisa e auxilia a realizá-la com eficácia”. Segundo Gil (2008), os métodos, enquanto meios técnicos de investigação, “visam fornecer a orientação necessária para a realização da pesquisa social”. Assim, em termos simplificados, a metodologia pode ser definida como o conjunto de etapas seguidas durante a condução da pesquisa para atingir seus objetivos.

Nas próximas subseções serão discutidos os procedimentos e as técnicas utilizados na coleta, tratamento e análise dos dados. As principais bases para a metodologia adotada podem ser encontradas em Guerin et al. (2021), Heitz et al. (2020) e Woudsma e Jakubicek (2020), além dos trabalhos de Longley et al. (2015) e Smith, Goodchild e Longley (2021), que discutem de forma abrangente a disciplina de análise espacial. Assim, esta seção está dividida da seguinte forma: Tipo e descrição geral da pesquisa, Conceitos de análise espacial, Objetos espaciais e variáveis de análise, Procedimentos de coleta de dados e Procedimentos de análise dos dados.

3.1. Tipo e descrição geral da pesquisa

Este trabalho tem como objetivo analisar a distribuição espacial de instalações logísticas no DF em relação à disponibilidade de infraestrutura de transporte na região, tratando-se de uma pesquisa descritiva quanto aos objetivos. Conforme Christensen et al. (2011), enquanto a pesquisa experimental visa demonstrar relações de causa e efeito, a pesquisa descritiva tem como foco examinar um fenômeno, uma situação ou um evento. De acordo com Gil (2008), esse tipo de pesquisa é utilizado para caracterizar uma população ou revelar a existência de associações entre variáveis.

Os dados foram coletados de fontes secundárias disponíveis online em bases de dados públicas (Walliman, 2021). Foram coletados quatro conjuntos de dados: empresas, regiões administrativas, zoneamento e sistema viário. Com esses dados, tornou-se possível identificar as empresas do setor de transporte e armazenamento no

DF, bem como suas localizações por regiões administrativas e zonas. Além disso, os dados do sistema viário, que englobam tanto as rodovias distritais quanto as federais, permitiram analisar a disponibilidade de infraestrutura de transporte na região.

Por fim, esta pesquisa é de natureza quantitativa. Nesse tipo de pesquisa, os dados coletados são quantitativos, ou seja, são dados que contêm alguma forma de magnitude e, portanto, podem ser mensurados (Walliman, 2021). Além disso, a pesquisa quantitativa aplica alguma forma de análise estatística. Neste trabalho, buscou-se realizar uma análise espacial, que combina princípios estatísticos com a análise de dados geográficos, utilizando métricas de distância, dispersão, medidas de associação espacial e ferramentas gráficas para investigar padrões e relacionamentos em dados espaciais, que representam valores numéricos em um contexto geográfico (Christensen et al., 2011; Oliveira, 2011; Longley et al., 2015).

3.2. Conceitos de análise espacial

Segundo Longley *et al.* (2015, p. 291), “a análise espacial é um conjunto de métodos cujos resultados mudam quando as localizações dos objetos analisados mudam”, ou seja, a análise espacial preocupa-se não somente com o que acontece, mas onde acontece. A seguir, são apresentados os conceitos fundamentais de lugar, atributo e objeto na análise espacial, utilizados para representar informações geográficas (Smith; Goodchild; Longley, 2021). Além disso, esses conceitos facilitam a compreensão e aplicação dos procedimentos de análise, como métricas de distância e medidas de associação espacial.

Um lugar identifica onde ocorre determinado fenômeno, sendo, portanto, definido através de um sistema de coordenadas, que permite que o lugar seja especificado. O atributo é uma característica ou propriedade de um lugar, como, por exemplo, um nome, uma medida, uma classificação ou uma estatística (Smith; Goodchild; Longley, 2021). Neste trabalho, os termos “atributo” e “variável” são utilizados de forma intercambiável, pois sempre estarão ligados a um lugar ou objeto.

O objeto é uma forma de se referir a um lugar por meio de pontos, linhas ou áreas. Um ponto é definido por um par de coordenadas, enquanto uma linha é uma sequência ordenada de pontos conectados por linhas retas. Já as áreas são anéis

ordenados de pontos conectados por linhas retas, ou seja, polígonos. Pontos, linhas e polígonos têm, respectivamente, dimensionalidade 0, 1 e 2 (Smith; Goodchild; Longley, 2021). No contexto deste trabalho, a localização das empresas é especificada por pontos, a rede viária por linhas e as regiões administrativas por polígonos, assim como as zonas. É por meio dessa representação que são construídos os mapas e são aplicados os métodos de relação e análise espacial. Desse modo, torna-se necessário explicar que no contexto deste trabalho foram analisados quatro conjuntos de objetos espaciais, os quais constituem quatro objetos de estudo, que serão descritos a seguir.

3.3. Objetos espaciais e variáveis de análise

Neste trabalho, foram considerados quatro conjuntos de objetos espaciais: 1) empresas, 2) regiões administrativas, 3) zoneamento e 4) sistema viário. Cabe ressaltar que cada um desses conjuntos é composto por diversos objetos espaciais, aos quais associam-se cada uma das variáveis de análise informadas nesta subseção.

O primeiro conjunto engloba todas as empresas do setor de transporte e armazenamento presentes no DF, que são o foco central da pesquisa. As variáveis de análise associadas às empresas são: classificação CNAE, localização por Região Administrativa (RA), Macrozona e Zona. Essas variáveis foram utilizadas para identificação das empresas e na elaboração de mapas temáticos, que fornecem *insights* sobre a distribuição espacial dos estabelecimentos no DF.

O segundo, regiões administrativas, consiste nas diferentes regiões administrativas do DF e desempenha um papel essencial na compreensão da relação entre a localização de estabelecimentos logísticos e as características específicas de cada unidade geográfica. Suas variáveis incluem Nome da RA, Área (km²), Densidade da rede viária (km/km²), Densidade de empresas (empresas por km²). A variável Nome identifica cada uma das regiões de administrativas, enquanto as variáveis Densidade da rede viária e Densidade de empresas possibilitaram a análise da disponibilidade de infraestrutura de transporte e da distribuição espacial das instalações logísticas, sendo utilizadas em mapas temáticos e no cálculo de medidas de associação espacial.

O terceiro conjunto compreende as macrozonas e zonas definidas no Plano Diretor de Ordenamento Territorial do Distrito Federal – PDOT/DF, sendo estas as suas variáveis de análise. O zoneamento foi utilizado para confecção de mapas temáticos, permitindo uma análise da localização das empresas em relação às macrozonas e zonas do PDOT, o que fornece insights sobre os impactos dessa política na distribuição espacial dos estabelecimentos.

Finalmente, o quarto conjunto engloba as rodovias distritais e federais no DF, com variáveis relacionadas à extensão, em km, e sua conexão com as Regiões Administrativas. Com as rodovias e as empresas, foi possível elaborar mapas que fornecem informações sobre a relação entre os dois, além de caracterizar as regiões administrativas.

Cabe ressaltar que se trata de um estudo populacional, uma vez que foram coletados para o estudo todos os objetos espaciais disponíveis. A Tabela 1 resume os quatro conjuntos de objetos espaciais e suas respectivas variáveis de análise.

Quadro 1 – Objetos espaciais e variáveis de análise

Objetos espaciais	Variáveis de análise
Empresas	Classificação CNAE
	RA
	Macrozona
	Zona
Regiões Administrativas (RA)	Nome da RA
	Área (km ²)
	Densidade da rede viária (km/ km ²)
	Densidade de empresas (empresas/km ²)
Zoneamento	Macrozona
	Zona
Rodovias	Extensão (km)
	RA

Fonte: elaborado pelo autor (2023)

3.4. Procedimentos de coleta de dados

A seguir, será detalhado o processo de coleta dos quatro conjuntos de objetos espaciais utilizados nesta pesquisa, bem como suas respectivas variáveis de análise:

empresas do setor de transporte e armazenamento, regiões administrativas do Distrito Federal, zoneamento e o sistema viário composto por rodovias distritais e federais.

3.4.1. Empresas

As variáveis associadas aos objetos espaciais que compõem o conjunto Empresas são as seguintes: Classificação CNAE, RA, Macrozona e Zona. Aqui, os objetos espaciais são as empresas e, para obtê-las, utilizou-se a Classificação Nacional de Atividades Econômicas (CNAE) para identificação de empresas do setor de transporte e armazenamento presentes no DF, tendo sido mantidos apenas os estabelecimentos cuja atividade principal corresponda a uma das classificações descritas no Quadro 2. Além da descrição dos códigos CNAE utilizados, são apresentadas no Quadro 2 as frequências absoluta e relativa das empresas encontradas para cada código.

Quadro 2 – Classificações CNAE (continua)

CNAE	Descrição	Frequência Absoluta	Frequência Relativa
4930201	Transporte rodoviário de carga, exceto produtos perigosos e mudanças, municipal	5192	62,35%
4930202	Transporte rodoviário de carga, exceto produtos perigosos e mudanças, intermunicipal, interestadual e internacional	2367	28,43%
4930203	Transporte rodoviário de produtos perigosos	389	4,67%
5211701	Armazéns gerais - emissão de warrant	141	1,69%
5211799	Depósitos de mercadorias para terceiros, exceto armazéns gerais e guarda-móveis	83	1,00%
5212500	Carga e descarga	63	0,76%
5250803	Agenciamento de cargas, exceto para o transporte marítimo	54	0,65%

Quadro 2 – Classificações CNAE (conclusão)

CNAE	Descrição	Frequência Absoluta	Frequência Relativa
5250804	Organização logística do transporte de carga	29	0,35%
5250805	Operador de transporte multimodal – OTM	9	0,11%

Fonte: elaborado pelo autor (2023)

Os dados sobre empresas foram coletados da base da Receita Federal, publicada no Portal Brasileiro de Dados Abertos (Brasil, 2023). Devido ao grande volume de dados armazenados, a base da RFB é disponibilizada de forma segregada, em diversos conjuntos de dados que, reunidos, formam a base completa. Foram aplicados filtros para que fossem selecionadas apenas as empresas localizadas no Distrito Federal e com cadastro ativo.

Os endereços disponíveis nos cadastros das empresas foram geocodificados, isto é, convertidos em coordenadas geográficas, com o propósito de possibilitar a identificação das regiões administrativas onde elas estão localizadas, a elaboração dos mapas temáticos e a análise de sua localização. Dessa forma, cada empresa é identificada por uma coordenada geográfica, sendo mantidos apenas os seguintes atributos: CNAE, RA, Macrozona e Zona. Estes dois últimos foram adquiridos juntamente com os dados das regiões administrativas e de zoneamento.

3.4.2. Regiões Administrativas e Zoneamento

As variáveis associadas aos objetos espaciais que compõem o conjunto Regiões Administrativas são as seguintes: Nome da RA, Área (km²), Densidade da rede viária (km/km²), Densidade de empresas (empresas/km²). Já os objetos espaciais que compõem o Zoneamento são diversas áreas localizadas no DF e classificadas dentro de uma das três classificações de macrozona e das nove subclassificações de zonas previstas no PDOT/DF (Distrito Federal, 2009), sendo estas as suas variáveis de análise (Zona e Macrozona).

Tanto as Regiões Administrativas quanto o Zoneamento desempenharam um papel fundamental na criação de mapas temáticos, permitindo uma caracterização

mais abrangente da região do Distrito Federal e possibilitando a análise da localização das empresas com relação às diferentes RA, Macrozonas e Zonas de localização. Além disso, as variáveis Densidade da rede viária (km/km^2) e Densidade de empresas ($\text{empresas}/\text{km}^2$) foram utilizadas para investigar a relação entre a localização das instalações logísticas do DF e a infraestrutura de transporte na região. Ao todo, foram consideradas 35 RA, 3 Macrozonas e 9 Zonas.

Os dados sobre Regiões Administrativas e Zoneamento foram obtidos através do site da Infraestrutura de Dados Espaciais do Distrito Federal (IDE/DF), publicados pela Secretaria de Estado de Desenvolvimento Urbano e Habitação (SEDUH) do DF (Distrito Federal, 2023).

3.4.3. Sistema viário

Os objetos espaciais que compõem o conjunto Sistema Viário são todas as rodovias federais e distritais presentes no DF. Desse modo, suas variáveis de análise são: Extensão (km^2) e RA de localização. Para obtenção desses objetos espaciais foram coletados dados sobre sistema viário nacional no catálogo de metadados da Infraestrutura Nacional de Dados Espaciais (INDE). As bases geoespaciais de rodovias federais e estaduais são disponibilizadas no site da INDE pela Coordenação-Geral de Planejamento e Programação de Investimentos (CGPLAN) e pela Coordenação de Levantamentos para Planejamento (COLEP), ambas vinculadas ao Departamento Nacional de Infraestrutura de Transportes (DNIT) (Brasil, 2023).

O primeiro passo foi filtrar os dados para que ficassem apenas as rodovias do Distrito Federal. Após isso, foi realizada a conversão do sistema de coordenadas das bases de dados, que indica localização por meio de latitude e longitude, para um sistema de coordenadas cartesiano, visto que o último é mais adequado para cálculos de área e comprimento (Longley *et al.*, 2015; Smith; Goodchild; Longley; 2021).

Finalmente, foi realizado o cálculo da extensão da rede viária do Distrito Federal por região administrativa, em quilômetros. Juntamente com os dados sobre as regiões administrativas, foi calculada a densidade da rede viária para cada RA, isto é, quantos quilômetros de rodovias cada RA possui para cada quilômetro quadrado de área.

3.5. Procedimentos de análise de dados

Os procedimentos de análise dos dados utilizados nesta pesquisa vão desde a utilização de representações visuais como gráficos mapas temáticos até cálculo de métricas de distância e medidas de associação espacial entre variáveis. Essas ferramentas analíticas são cruciais para entender e explorar a distribuição espacial das instalações logísticas do Distrito Federal e, adicionalmente, possibilitaram uma análise aprofundada das interações entre as localizações das empresas e a infraestrutura de transporte disponível na região. Ao utilizar esses métodos, que se busca é encontrar padrões, tendências e possíveis conexões, fornecendo uma visão geral completa para atender aos objetivos da pesquisa.

A seguir serão apresentados alguns dos procedimentos de análise de dados utilizados na pesquisa, que são Métricas de distância e Medidas de associação espacial. As métricas de distância são utilizadas para analisar padrões de pontos, que, no contexto deste trabalho, identificam as empresas. As medidas de associação espacial ajudam a entender como os valores de um ou mais variáveis estão geograficamente distribuídos.

3.5.1. Métricas de distância

Visto que as empresas são identificadas por coordenadas geográficas, ou seja, por pontos, é possível calcular estatísticas baseadas em distância, quais sejam, medidas de tendência central e de dispersão, que são utilizadas para analisar padrões de pontos. Neste trabalho, utilizou-se o centro médio e a Elipse de Desvio Padrão (*Standard Deviational Ellipse*, SDE). Além disso, foram utilizados *Buffers* para analisar a proximidade entre as empresas (pontos) e as rodovias (linhas).

De acordo com Smith, Goodchild e Longley (2021), o centro médio de um conjunto de pares de coordenadas (x_i, y_i) é dado pela seguinte expressão:

$$M = \left(\sum_i \frac{x_i}{n}, \sum_i \frac{y_i}{n} \right) = (\bar{x}, \bar{y})$$

O SDE é uma medida na qual uma elipse é desenhada a partir do centro médio ou outra medida de centralidade, sumarizando, assim, tanto a dispersão quanto a concentração de um conjunto de objetos (Yuill, 1971). De acordo com Yuill (1971), duas medidas podem ser utilizadas para analisar a concentração e dispersão dos dados: a área cercada pela SDE e o número de pontos dentro da SDE. Se a área da SDE é pequena em relação à área da região de estudo, isso indica concentração dos pontos. Uma proporção alta de pontos dentro da SDE também indica concentração, representando um índice de concentração dos dados (Yuill, 1971).

Por fim, a técnica de *buffering* produz áreas (*Buffers*) ao redor de objetos dentro de uma distância especificada (Smith; Goodchild; Longley; 2021). Assim, é possível determinar quais empresas estão localizadas a uma determinada distância de uma rodovia, o que ajuda a entender a influência da proximidade das rodovias em sua localização.

3.5.2. Medidas de associação espacial

As medidas de associação espacial foram calculadas considerando as variáveis Densidade da rede viária e Densidade de empresas, com o objetivo de analisar relação entre a infraestrutura de transporte disponível e a localização das empresas. Além disso, quando analisada individualmente, a Densidade de empresas também possibilita uma análise sobre a distribuição espacial dos estabelecimentos logísticos nas regiões administrativas do DF.

Foram utilizados os seguintes índices: *Moran's I* global univariado, *Moran's I* local univariado, *Moran's I* global bivariado e *Moran's I* local bivariado. Esses índices são empregados para analisar as semelhanças entre os atributos de um objeto espacial e sua localização (Smith; Goodchild; Longley, 2021).

O *Moran's I* global univariado é um índice que nos ajuda a entender como os valores de uma única variável estão geograficamente distribuídos. Ele varia entre -1 e +1, onde um índice mais próximo de +1 indica que valores semelhantes tendem a estar próximos uns dos outros e valores diferentes tendem a estar distantes uns dos outros. Em contrapartida, um índice mais próximo de -1 indica que objetos próximos tendem a ter atributos com valores diferentes. Já um valor próximo a zero indica que

não há relação entre os valores e a localização (Longley *et al.*, 2015; Smith; Goodchild; Longley, 2021).

Sejam x_i e x_j os valores de uma variável X em duas unidades geográficas i e j . Além disso sejam \bar{x} a média da variável X e w_{ij} um peso que assume valor 1 ou 0, utilizado para indicar se i e j são ou não vizinhos. Desse modo, o *Moran's I* global para n unidades geográficas é dado por:

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

Uma medida relacionada aos índices de Moran é o *lag* espacial (*Spatial Lag*) de uma unidade geográfica, que é a soma dos valores do atributo nas unidades vizinhas, ou seja, indica se uma unidade geográfica está cercada por outras com valores baixos ou altos (Anselin; Rey, 2014). É dado por:

$$L(z_i) = \sum_j w_{ij} z_j$$

onde:

$$z_i = \frac{x_i - \bar{x}}{\sigma_x}$$

O *Moran's I* local univariado é indicador local de associação espacial. Esse índice é calculado para cada unidade geográfica individualmente e os resultados, uma vez agrupados, ajudam-nos a identificar a presença de *clusters* ou aglomerações de objetos com atributos de valores altos ou baixos (Anselin, 1995; Smith; Goodchild; Longley, 2021). É dado pela seguinte expressão:

$$I_i = z_i \sum_j w_{ij} z_j$$

O *Moran's I* global bivariado calcula a autocorrelação espacial conjunta entre duas variáveis X e Y. Sua interpretação é semelhante à do *Moran's I* global univariado, indicando se os valores altos ou baixos das duas variáveis estão próximos ou distantes uns dos outros (Anselin; Syabri; Smirnov, 2002). É dado pela seguinte expressão:

$$I_B = \frac{\sum_i \sum_j w_{ij} z_1^i z_2^j}{\sum_i (z_1^i)^2}$$

onde z_1 e z_2 representam, respectivamente, as variáveis X e Y padronizadas.

O *Moran's I* local bivariado é semelhante ao *Moran's I* local univariado, porém considera duas variáveis X e Y. É calculado para cada unidade geográfica individualmente e, uma vez agrupados, os valores dos índices indicam se os valores altos ou

baixos das duas variáveis estão próximos ou distantes uns dos outros, ou seja, ajudam-nos a identificar *clusters* ou aglomerações de atributos com valores semelhantes (Anselin; Syabri; Smirnov, 2002). É dado pela seguinte expressão:

$$I_{z_1 z_2}^i = z_1^i \sum_j w_{ij} z_2^j$$

Os indicadores locais classificam uma unidade geográfica de análise em quatro categorias: HH (*High-High*), LL (*Low-Low*), HL (*High-Low*) e LH (*Low-High*). Uma unidade HH tem um atributo com valor alto e tem vizinhos com atributos altos, enquanto uma unidade LL tem valor baixo e é cercada por valores baixos. Essa classificação auxilia na identificação de *clusters* espaciais e *outliers* locais.

Para determinar se os índices calculados são estatisticamente significativos, os valores sob análise são permutados e utilizados no cálculo de outros índices, sob uma hipótese nula de independência espacial. Esses índices, então, são comparados com índice observado para obtenção de um pseudo p-valor, ou nível de significância (Smith; Goodchild; Longley, 2021). Neste trabalho, foram realizadas 999 permutações e foi utilizado um p-valor de 0,05.

3.6. Ferramentas utilizadas

A coleta e análise dos dados, a construção dos gráficos e mapas temáticos e todos os cálculos foram realizados na plataforma *Google Colaboratory*, em linguagem *Python*. Destaca-se o uso das bibliotecas *GeoPandas* e *PySAL*, utilizadas para análise espacial.

4. RESULTADOS E DISCUSSÃO

O objetivo deste trabalho é analisar a distribuição espacial das instalações logísticas do Distrito Federal em relação à disponibilidade de infraestrutura de transporte na região. Para alcançar esse objetivo, este capítulo será dividido em seis seções, nas quais serão apresentados cada um dos cinco objetivos específicos e realizada a discussão dos resultados. O capítulo está dividido da seguinte maneira: Mapeamento das instalações logísticas no Distrito Federal, Distribuição espacial das instalações logísticas em diferentes regiões administrativas, Disponibilidade de infraestrutura de transporte no Distrito Federal, Relação entre a localização das instalações e infraestrutura de transporte, Influência da política de zoneamento do DF na localização das instalações logísticas e Discussão.

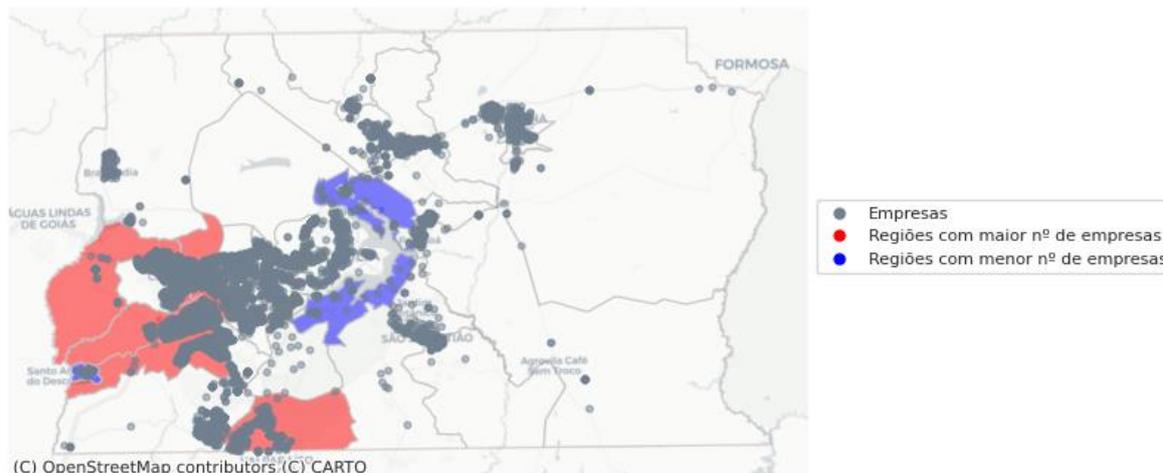
Nas cinco primeiras seções do capítulo são apresentados os resultados da pesquisa e na sexta seção é feita a discussão. Na primeira seção, são apresentadas as tabelas de frequência das instalações por região administrativa e os resultados da SDE (*Standard Deviational Ellipse*) para esses pontos. Na segunda seção é analisada a Densidade de empresas por região administrativa utilizando o *Spatial Lag* e a autocorrelação espacial (*Moran's I*) para essa variável. Na terceira seção investiga-se a disponibilidade infraestrutura de transporte (rodovias) nas diferentes regiões administrativas do DF. Na quarta seção, fez-se uma análise da relação entre a localização das instalações e infraestrutura de transporte, na qual foi verificada a proximidade física das empresas em relação às rodovias do DF e a correlação espacial entre a Densidade da rede viária e Densidade de empresas nas regiões administrativas. Na quinta seção é apresentada a distribuição das empresas por tipo de zona (zoneamento). Na sexta seção é feita a discussão dos resultados apresentados.

4.1. Mapeamento das instalações logísticas no Distrito Federal

Para mapear as instalações logísticas no DF, foram desenhados os pontos que representam cada uma das empresas (Figura 1) e calculadas as frequências absoluta e relativa de empresas por região administrativa (Tabela 1). A Figura 1 demonstra as localizações das empresas do setor de transporte e armazenamento do DF,

destacando as cinco regiões com maior ocorrência de empresas (cor vermelha) e as cinco regiões com menor ocorrência de empresas (cor azul). Ao todo, foram identificadas 8327 empresas, presentes em todas as 35 regiões administrativas do DF.

Figura 1 – Localização das empresas no DF



Fonte: elaborado pelo autor (2023)

Na Figura 1, as regiões em vermelho são: Ceilândia, Samambaia, Taguatinga, Recanto das Emas e Santa Maria. Juntas, elas concentram 49,48% das empresas mapeadas, conforme apresentado na Tabela 1. As três primeiras, Ceilândia, Samambaia e Taguatinga, concentram 37,83% das ocorrências de empresas. Ainda na Figura 1, são apresentadas, em azul, as 5 regiões com menor ocorrência de empresas. Conforme apresentado na Tabela 1, essas regiões são: Varjão, Água Quente, Lago Norte, Cruzeiro e Lago Sul, que juntas concentram apenas 1,91% das empresas mapeadas.

Tabela 1 – Localização das empresas por RA (continua)

Região Administrativa (RA)	Frequência Absoluta (FA)	Frequência Relativa (FR)	FR Acum.
CEILÂNDIA	1391	16,70%	16,70%
SAMAMBAIA	905	10,87%	27,57%
TAGUATINGA	854	10,26%	37,83%
RECANTO DAS EMAS	499	5,99%	43,82%
SANTA MARIA	471	5,66%	49,48%
GAMA	412	4,95%	54,43%
PLANO PILOTO	351	4,22%	58,64%
VICENTE PIRES	337	4,05%	62,69%
SIA	302	3,63%	66,31%
GUARÁ	301	3,61%	69,93%
PLANALTINA	250	3,00%	72,93%
RIACHO FUNDO II	221	2,65%	75,59%

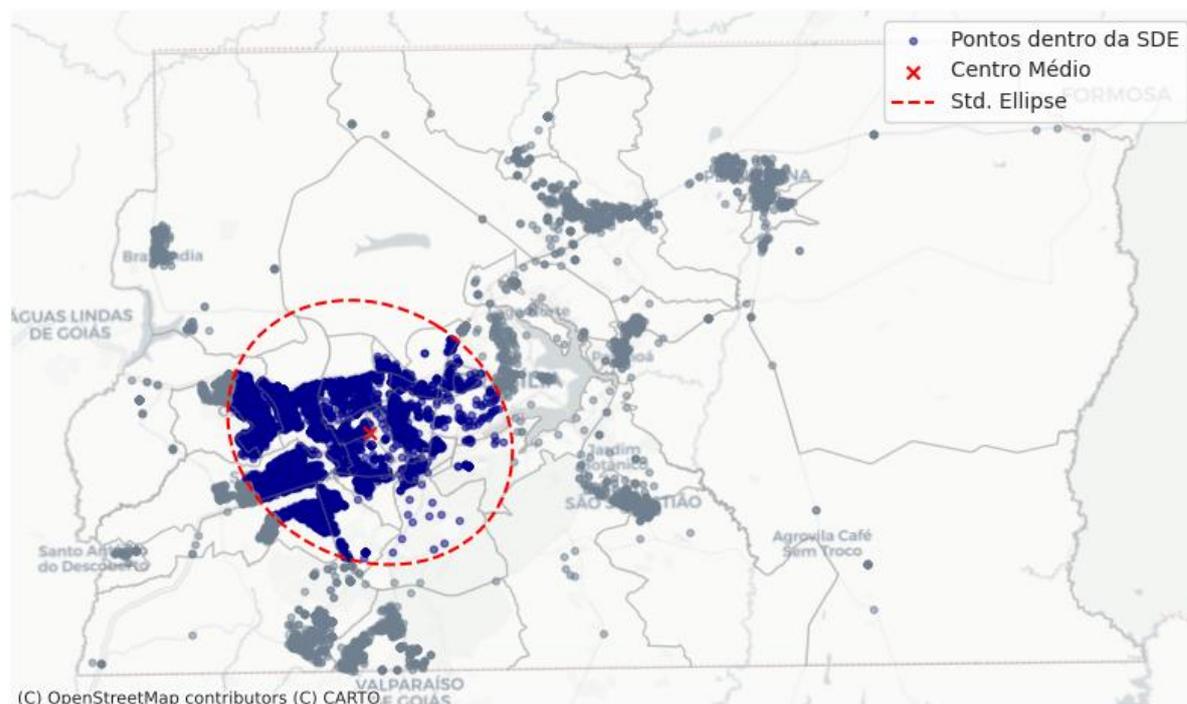
Tabela 1 – Localização das empresas por RA (conclusão)

Região Administrativa (RA)	Frequência Absoluta (FA)	Frequência Relativa (FR)	FR Acum.
SOBRADINHO	163	1,96%	82,30%
SOBRADINHO II	159	1,91%	84,21%
BRAZLÂNDIA	139	1,67%	85,88%
NÚCLEO BANDEIRANTE	118	1,42%	87,29%
RIACHO FUNDO	116	1,39%	88,69%
PARANOÁ	97	1,16%	89,85%
SCIA	97	1,16%	91,02%
ARAPOANGA	97	1,16%	92,18%
ÁGUAS CLARAS	84	1,01%	93,19%
ITAPOÃ	81	0,97%	94,16%
CANDANGOLÂNDIA	62	0,74%	94,91%
JARDIM BOTÂNICO	60	0,72%	95,63%
FERCAL	60	0,72%	96,35%
PARK WAY	55	0,66%	97,01%
SOL NASCENTE E PÔR DO SOL	47	0,56%	97,57%
SUDOESTE/OCTOGONAL	43	0,52%	98,09%
LAGO SUL	40	0,48%	98,57%
CRUZEIRO	33	0,40%	98,97%
LAGO NORTE	31	0,37%	99,34%
ÁGUA QUENTE	31	0,37%	99,71%
VARJÃO	24	0,29%	100,00%
Total	8327	100%	-

Fonte: elaborado pelo autor (2023)

Para analisar a dispersão e a concentração das empresas na região do DF, foram calculados o centro médio dos pontos que identificam essas empresas (coordenadas geográficas) e a SDE (*Standard Deviatonal Ellipse*). Em seguida, foram identificados quais os pontos estão dentro da SDE e qual é a sua área. A Figura 2 apresenta o centro médio e a SDE.

Figura 2 – Centro médio e SDE



Fonte: elaborado pelo autor (2023)

Conforme discutido na seção 3.5.1, há duas medidas geradas pela SDE que podem ser utilizadas para analisar a concentração dos dados: a área da SDE em relação à área total de estudo e a quantidade de pontos dentro da SDE. Conforme é possível observar na Figura 2, a SDE concentra a maioria dos pontos, que estão em azul. Isso também pode ser constatado na Tabela 2, que apresenta as frequências absoluta e relativa dos pontos localizados dentro da SDE por região administrativa.

Tabela 2 – Localização das empresas dentro da SDE por RA (continua)

Região Administrativa (RA)	Frequência Absoluta (FA)	Frequência Relativa (FR)	FR Acum.
CEILÂNDIA	1105	21,41%	21,41%
TAGUATINGA	854	16,54%	37,95%
SAMAMBAIA	731	14,16%	52,11%
RECANTO DAS EMAS	393	7,61%	59,72%
VICENTE PIRES	337	6,53%	66,25%
GUARÁ	301	5,83%	72,08%
SIA	242	4,69%	76,77%
ARNIQUEIRA	215	4,17%	80,94%
RIACHO FUNDO II	214	4,15%	85,08%
NÚCLEO BANDEIRANTE	118	2,29%	87,37%
RIACHO FUNDO	116	2,25%	89,62%
SCIA	97	1,88%	91,50%

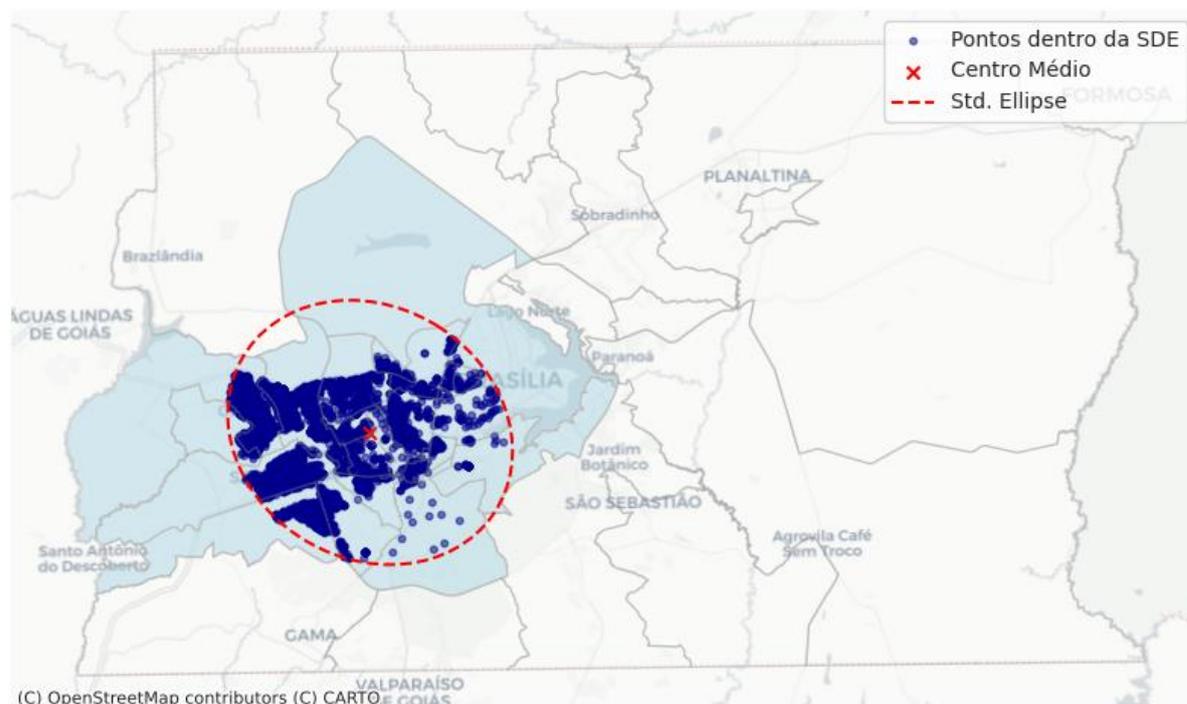
Tabela 2 – Localização das empresas dentro da SDE por RA (conclusão)

Região Administrativa (RA)	Frequência Absoluta (FA)	Frequência Relativa (FR)	FR Acum.
PLANO PILOTO	92	1,78%	93,28%
ÁGUAS CLARAS	84	1,63%	94,91%
CANDANGOLÂNDIA	62	1,20%	96,11%
PARK WAY	53	1,03%	97,13%
SOL NASCENTE E PÔR DO SOL	45	0,87%	98,00%
SUDOESTE/OCTOGONAL	43	0,83%	98,84%
CRUZEIRO	33	0,64%	99,48%
LAGO SUL	27	0,52%	100,00%
Pontos na SDE	5162	100,00%	-
Porcentagem em relação ao total dos pontos analisados		61,99%	

Fonte: elaborado pelo autor (2023)

Ao todo, são 5162 pontos dentro da SDE, o que representa 61,99% do total dos pontos analisados (Tabela 2), ou seja, a maioria das empresas está dentro da SDE. Essas empresas estão localizadas em 20 regiões administrativas, sendo que 11 destas têm 100% dos seus pontos dentro da elipse (Tabela 1, Figura 2 e Tabela 2). Essas regiões são: Taguatinga, Vicente Pires, Guará, Arniqueira, Núcleo Bandeirante, Riacho Fundo, SCIA, Águas Claras, Candangolândia, Sudoeste/Octogonal e Cruzeiro. Destacam-se, ainda, as regiões Ceilândia, Taguatinga, Samambaia, que concentram 52,11% do total de pontos dentro da SDE (Tabela 2). A Figura 3 destaca todas as regiões de localização dos pontos da SDE.

Figura 3 – Regiões de localização das empresas da SDE



Fonte: elaborado pelo autor (2023)

Na Figura 3, fica evidente a concentração de empresas ao sudoeste do DF. Além de a maioria dos pontos estarem dentro da SDE, a área total da elipse é pequena quando comparada à área total do DF. A área do DF foi estimada em, aproximadamente, 5767 km², enquanto a área cercada pela SDE foi estimada em 555 km². Assim, a SDE descreve bem a localização média dos pontos, que estão concentrados em algumas regiões administrativas que representam uma parte minoritária da área total do DF.

Em resumo, a análise da frequência das instalações logísticas por região administrativa (Figura 1 e Tabela 1) e os resultados da SDE (Figura 2, Tabela 2 e Figura 3) para esses estabelecimentos revelaram um nível alto de concentração desses pontos no sudoeste do DF, com destaque para as regiões de Ceilândia, Taguatinga e Samambaia, que tem alta frequência de ocorrência de empresas e representam a maioria dos pontos localizados dentro da SDE. Junto com essas regiões, pode-se destacar as outras regiões com pontos dentro da SDE, apresentadas na Tabela 2. Assim, a análise dos pontos que representam a localização das empresas revelou a concentração desses pontos em 20 regiões administrativas (Tabela 2), que concentram a maioria dos pontos (Tabela 1).

4.2. Distribuição espacial das instalações logísticas em diferentes regiões administrativas

As regiões administrativas do DF variam muito em termos de área, isto é, o tamanho das áreas de análise não é homogêneo. Assim, foi utilizada a variável Densidade de empresas (empresas/km²) para analisar a distribuição relativa das empresas em relação às diferentes regiões administrativas. A Tabela 3 apresenta a área e a densidade de empresas de cada RA.

Tabela 3 – Área e Densidade de empresas das RA (continua)

RA	Área da RA (km ²)	Densidade de empresas
NÚCLEO BANDEIRANTE	5,05	23,38
ARNIQUEIRA	13,36	16,09
VARJÃO	1,63	14,73
TAGUATINGA	63,04	13,55
SCIA	7,39	13,13
GUARÁ	25,22	11,93
SIA	29,05	10,40
CRUZEIRO	3,19	10,34
CANDANGOLÂNDIA	6,60	9,39
ÁGUAS CLARAS	9,19	9,14
SAMAMBAIA	99,99	9,05
VICENTE PIRES	42,88	7,86
CEILÂNDIA	192,66	7,22
SUDOESTE/OCTOGONAL	6,47	6,65
RIACHO FUNDO	17,78	6,52
RIACHO FUNDO II	38,47	5,75
RECANTO DAS EMAS	93,11	5,36
ARAPOANGA	21,99	4,41
SANTA MARIA	130,89	3,60
ÁGUA QUENTE	9,51	3,26
ITAPOÃ	34,30	2,36
GAMA	275,41	1,50
SOL NASCENTE E PÔR DO SOL	40,49	1,16
SOBRADINHO II	181,35	0,88
SOBRADINHO	193,53	0,84
PLANO PILOTO	437,20	0,80
SÃO SEBASTIÃO	262,07	0,69
LAGO SUL	76,05	0,53
PARK WAY	118,41	0,46
LAGO NORTE	75,49	0,41
FERCAL	154,38	0,39
BRAZLÂNDIA	470,18	0,30

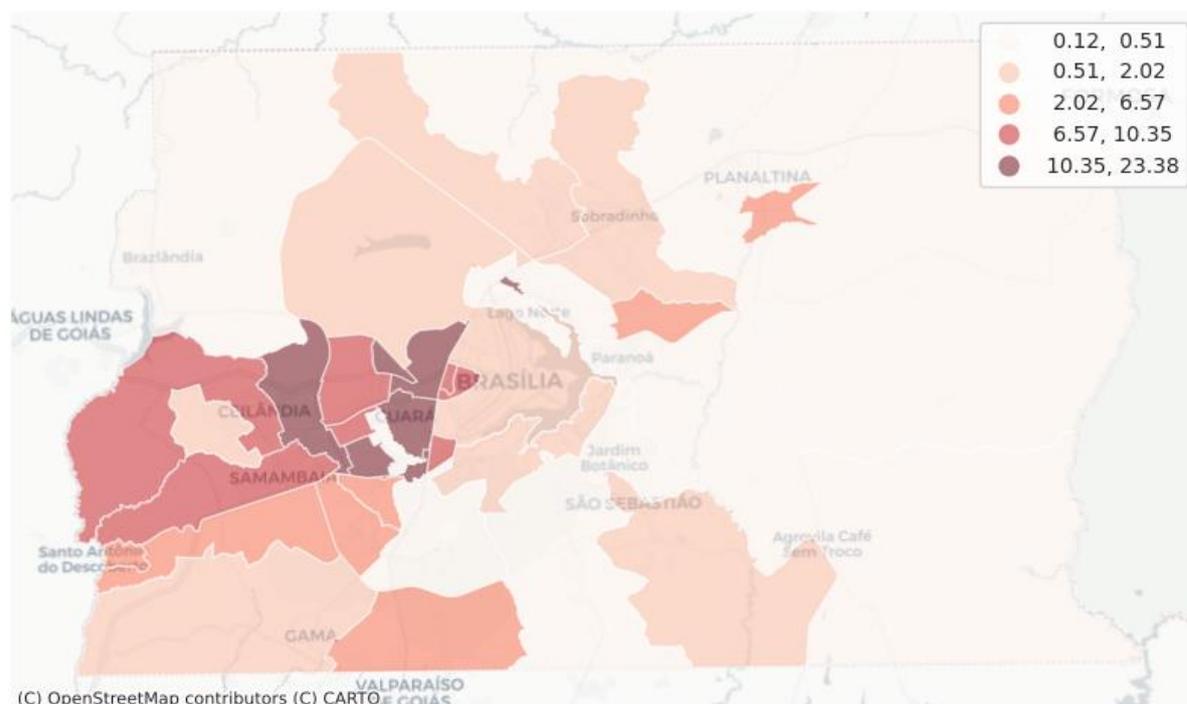
Tabela 3 – Área e Densidade de empresas das RA (conclusão)

RA	Área da RA (km ²)	Densidade de empresas
JARDIM BOTÂNICO	290,61	0,21
PLANALTINA	1509,41	0,17
PARANOÁ	830,25	0,12

Fonte: elaborado pelo autor (2023)

Conforme apresentado na Tabela 3, as 10 regiões com maior densidade de empresas são Núcleo Bandeirante, Arniqueira, Varjão, Taguatinga, SCIA, Guará, SIA, Cruzeiro, Candangolândia e Águas Claras. Com exceção do Varjão, essas regiões estão localizadas ao sudoeste do DF e são próximas umas das outras, o que pode ser observado na Figura 4, que apresenta a densidade de empresas por RA.

Figura 4 – Densidade de empresas por RA

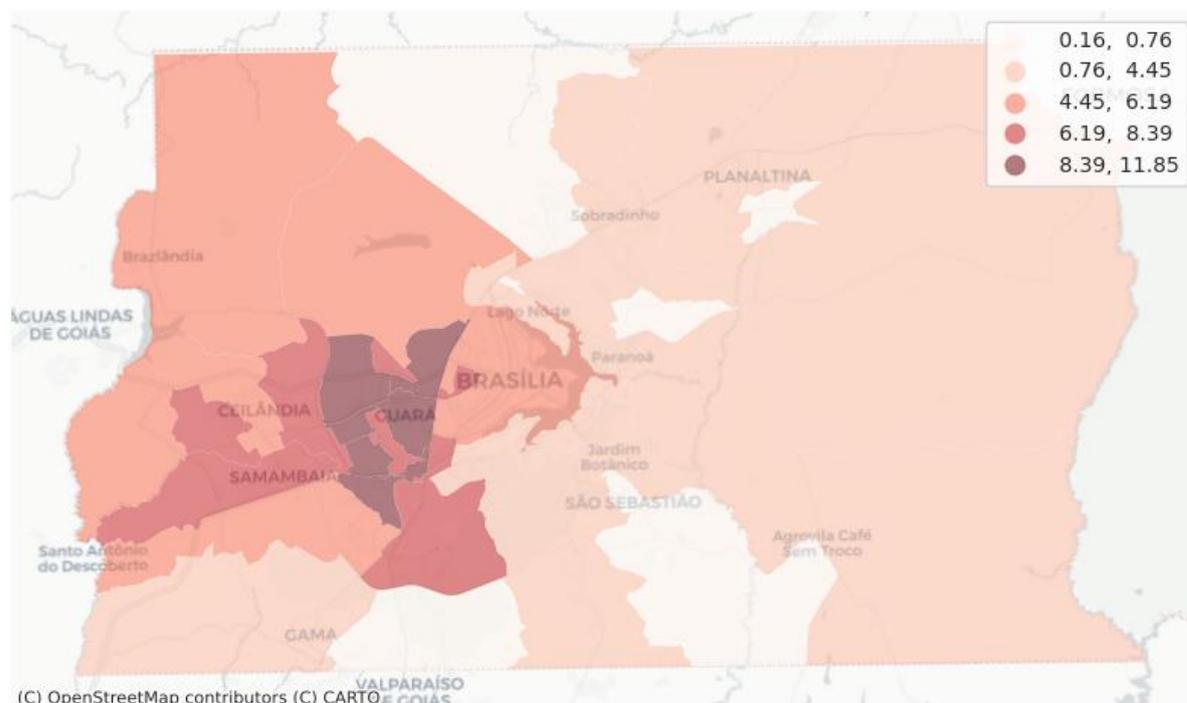


Fonte: elaborado pelo autor (2023)

Na Figura 4, além da proximidade das regiões com maior densidade de empresas, também é possível notar a proximidade das regiões com menor densidade de empresas, como o Paranoá, Planaltina e Jardim Botânico, que têm as menores densidades. Os *lags* espaciais (*Spatial Lags*) para as regiões administrativas ajudam a evidenciar a essa semelhança entre localização e densidade de empresas. Um *Spatial Lag* alto em uma RA indica que os vizinhos dessa região têm atributos com valores altos, o que pode indicar dependência espacial, que regiões com valores de densidade

de empresas semelhantes estão próximas umas das outras (Anselin; Rey, 2014). A Figura 5 apresenta os *lags* espaciais das RA.

Figura 5 – *Spatial Lags* das densidades de empresas das RA

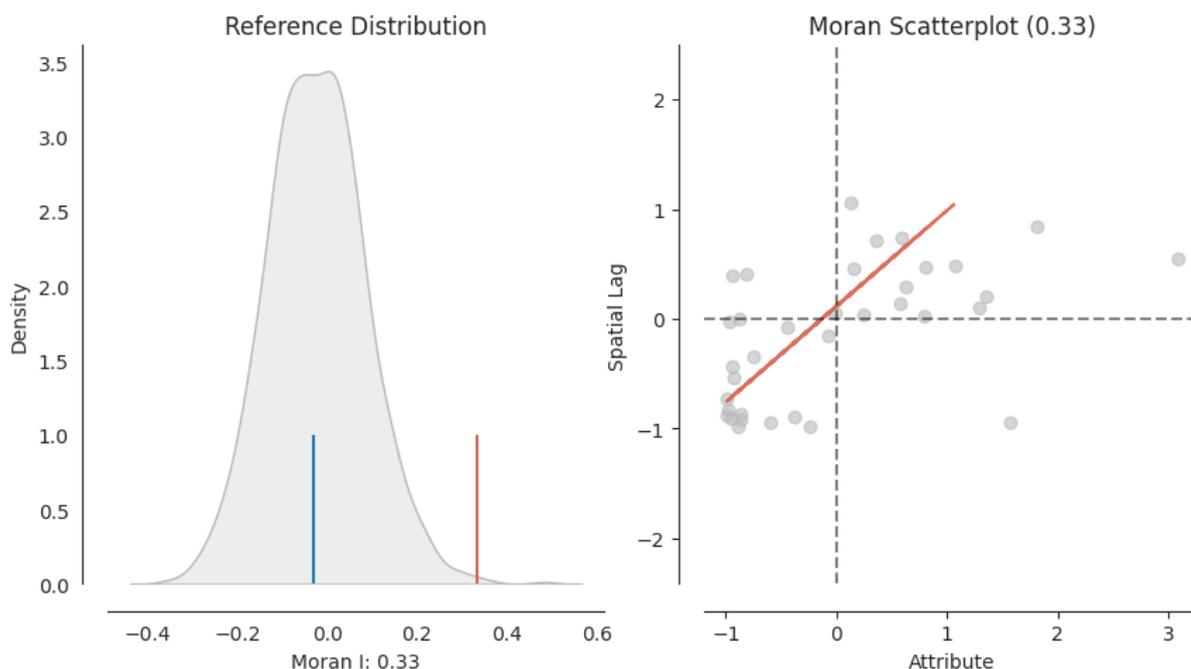


Fonte: elaborado pelo autor (2023)

Na Figura 5, as regiões em vermelho escuro são SIA, Guará, Núcleo Bandeirante, Riacho Fundo, Arniqueira, Águas Claras e Vicente Pires. Isso indica que essas regiões têm vizinhos com densidade de empresas altas. Já as regiões em cores mais claras são aquelas com regiões vizinhas com densidade de empresas baixas, que estão localizadas ao leste do DF, como as regiões de Sobradinho, Planaltina, Paranoá, São Sebastião e Jardim Botânico, entre outras.

O *lag* espacial apenas indica o valor da densidade de empresas em regiões vizinhas, sem levar em conta o valor da variável na própria região sob análise. Assim, foram calculados os índices de Moran (global e locais), que relacionam o valor da variável na região de análise aos valores da variável nas suas regiões vizinhas. A Figura 6 apresenta o Diagrama de Dispersão de Moran (*Moran Scatterplot*) e distribuição dos índices calculados nas permutações.

Figura 6 – Diagrama de Dispersão de Moran para a variável Densidade de empresas



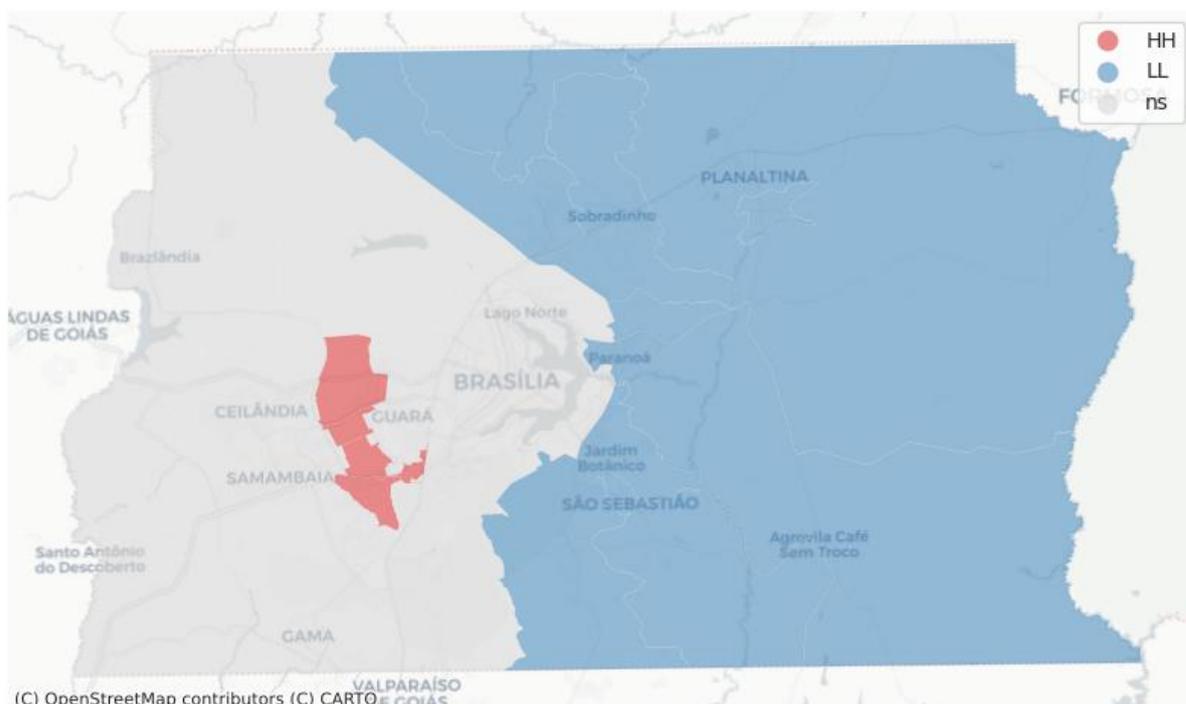
Fonte: elaborado pelo autor (2023)

No Diagrama de Dispersão de Moran (Figura 6) é apresentada a relação entre o valor da variável de análise na região e o seu *lag* espacial. Em termos práticos, quando uma região possui uma alta densidade de empresas e suas regiões vizinhas também têm valores elevados, indicados pelo *lag* espacial, o ponto representativo estará posicionado para cima e para a direita no gráfico. Analogamente, se uma região possui baixa densidade de empresas e suas regiões vizinhas também têm valores baixos, o ponto correspondente estará localizado para baixo e para a esquerda no gráfico (Figura 6).

Em relação à reta vermelha traçada no gráfico (Figura 6), ela representa a linha de referência para o padrão de dispersão dos pontos. A presença de mais pontos próximos dessa linha pode sugerir a presença ou ausência de autocorrelação espacial significativa. O índice obtido foi de 0,33 com um p-valor de 0,004, sugerindo uma autocorrelação espacial moderada e estatisticamente significativa para a variável Densidade de empresas.

Já os índices de Moran locais para a variável Densidade de empresas revelaram a presença de dois *clusters* estatisticamente significativos. Um HH, no qual as regiões com atributos com valores altos estão cercadas por outras regiões com atributos com valores altos, e um LL, no qual as regiões com atributos com valores baixos estão cercadas por outras regiões com atributos com valores baixos. A Figura 7 apresenta os *clusters* encontrados.

Figura 7 – Mapa de identificação de *clusters* para a variável densidade de empresas



Fonte: elaborado pelo autor (2023)

Na Figura 7, o *cluster* HH é formado pelas regiões Núcleo Bandeirante, Riacho Fundo, Águas Claras, Vicente Pires e Arniqueira. Essas regiões apresentam valores altos de densidade de empresas e estão localizadas próximas umas das outras. Três delas, Núcleo Bandeirante, Arniqueira e Águas Claras, estão entre as dez regiões administrativas com maior densidade de empresas, conforme apresentado na Tabela 3.

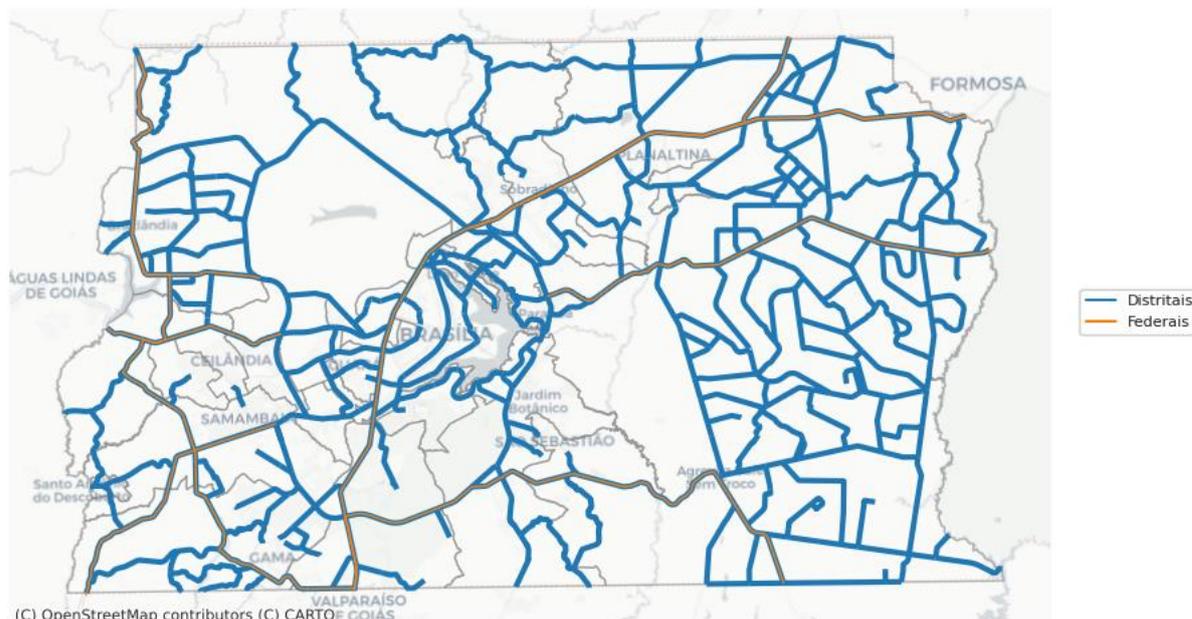
O *cluster* LL é formado pelas regiões de Sobradinho, Planaltina, Paranoá, São Sebastião, Sobradinho II, Jardim Botânico, Itapoã, Fercal e Arapoanga (Figura 7). Essas regiões apresentam valores baixos de densidade de empresas e estão localizadas próximas umas das outras. Cinco delas, Planaltina, Paranoá, Jardim Botânico, Fercal e São Sebastião, estão dentre as dez regiões com menor densidade de empresas, conforme apresentado na Tabela 3.

Em resumo, a análise da distribuição espacial das instalações logísticas nas diferentes regiões administrativas do DF revelou um grau positivo e moderado de autocorrelação espacial para a variável Densidade de empresas (Figura 6). Os indicadores locais indicaram a presença de dois *clusters*, um HH (*High-High*) e um LL (*Low-Low*), localizados, respectivamente, ao sudoeste e leste do DF (Figura 7). Existem outras regiões próximas das regiões que compõem o *cluster* HH que têm uma maior densidade de empresas (Figura 4 e Figura 5), porém, elas não são estatisticamente significativas. Além disso, as regiões que compõem o *cluster* HH representam uma parte minoritária da área total do DF, enquanto o *cluster* LL representa mais da metade (Tabela 3 e Figura 7), o que reforça a concentração das empresas em uma parte pequena do DF.

4.3. Disponibilidade de infraestrutura de transporte no Distrito Federal

Para analisar a disponibilidade de infraestrutura de transporte na região, foram desenhadas as rodovias distritais e federais presentes no DF (Figura 8) e foi calculada, em km, a extensão da rede viária presente em cada região administrativa (Tabela 4). Com isso, foi possível obter uma visão geral sobre o sistema viário do DF, desenhado na Figura 8.

Figura 8 – Sistema viário do DF



Fonte: elaborado pelo autor (2023)

A Figura 8 apresenta as rodovias federais (em azul) e distritais (em laranja) presentes no DF. Ao todo, são aproximadamente 2.456 km de rodovias por todo DF, 1944 km de rodovias distritais e 512 km de rodovias distritais, presentes em todas as regiões administrativas do DF, com as regiões de Planaltina, Paranoá, Brazlândia, Gama e Sobradinho, representando 63,94% de todas as rodovias. (Tabela 4). As regiões com maior extensão de rodovias são algumas das maiores regiões do DF, sendo Planaltina a maior de todas e a que tem a maior extensão de rodovias. A Tabela 4 apresenta a extensão de rodovias por região administrativa.

Tabela 4 – Extensão de rodovias em km por RA (continua)

RA	Distritais	Federais	Total	Participação relativa
PLANALTINA	628,54	164,22	792,77	32,28%
PARANOÁ	270,47	12,67	283,15	11,53%
BRAZLÂNDIA	178,28	69	247,28	10,07%
GAMA	111,09	34,38	145,47	5,92%
SOBRADINHO	63,82	38,07	101,89	4,15%
PLANO PILOTO	70,56	9,64	80,21	3,27%
SÃO SEBASTIÃO	49,01	28,82	77,84	3,17%
SANTA MARIA	45,4	31,93	77,33	3,15%
CEILÂNDIA	45,3	31,88	77,18	3,14%
JARDIM BOTÂNICO	57,83	9,82	67,65	2,75%
PARK WAY	37,67	23,88	61,55	2,51%
RECANTO DAS EMAS	39,22	19,18	58,41	2,38%
LAGO NORTE	47,68	6,19	53,87	2,19%

Tabela 4 – Extensão de rodovias em km por RA (conclusão)

RA	Distritais	Federais	Total	Participação relativa
FERCAL	48,6	0	48,6	1,98%
SOBRADINHO II	43,32	0	43,32	1,76%
LAGO SUL	37,05	0	37,05	1,51%
SAMAMBAIA	21,49	6,9	28,39	1,16%
TAGUATINGA	18,33	6,32	24,65	1,00%
ITAPOÃ	13,42	10,88	24,3	0,99%
VICENTE PIRES	20,32	0,01	20,33	0,83%
SIA	16,8	0	16,8	0,68%
GUARÁ	13,08	0	13,08	0,53%
ARAPOANGA	9,24	0	9,24	0,38%
RIACHO FUNDO II	8,32	0	8,32	0,34%
CANDANGOLÂNDIA	4,32	3,81	8,13	0,33%
SCIA	7,76	0	7,76	0,32%
ÁGUA QUENTE	7,61	0	7,61	0,31%
NÚCLEO BANDEIRANTE	7,11	0	7,11	0,29%
CRUZEIRO	2,62	2,67	5,28	0,22%
SOL NASCENTE E PÔR DO SOL	5,18	0	5,18	0,21%
SUDOESTE/OCTOGONAL	3,82	1,13	4,95	0,20%
ÁGUAS CLARAS	4,36	0	4,36	0,18%
VARJÃO	1,77	0,98	2,76	0,11%
ARNIQUEIRA	2,74	0	2,74	0,11%
RIACHO FUNDO	1,63	0	1,63	0,07%
Total	1943,78	512,37	2456,2	100,00%

Fonte: elaborado pelo autor (2023)

Como as regiões variam muito em termos de área, também é importante analisar a extensão de rodovias em relação a área de cada região administrativa, o que foi feito por meio da variável Densidade da rede viária, que é a razão entre a extensão de rodovias (km) e a área de uma RA (km²). A Tabela 5 apresenta a área, a extensão de rodovias e a densidade de rede viária de cada RA.

Tabela 5 – Rodovias, área e densidade da rede viária (continua)

RA	Rodovias (km)	Área (km ²)	Densidade da rede viária (km/km ²)
VARJÃO	2,76	1,63	1,69
CRUZEIRO	5,28	3,19	1,66
NÚCLEO BANDEIRANTE	7,11	5,05	1,41
CANDANGOLÂNDIA	8,13	6,60	1,23
SCIA	7,76	7,39	1,05
ÁGUA QUENTE	7,61	9,51	0,80
SUDOESTE/OCTOGONAL	4,95	6,47	0,77
LAGO NORTE	53,87	75,49	0,71

Tabela 5 – Rodovias, área e densidade da rede viária (conclusão)

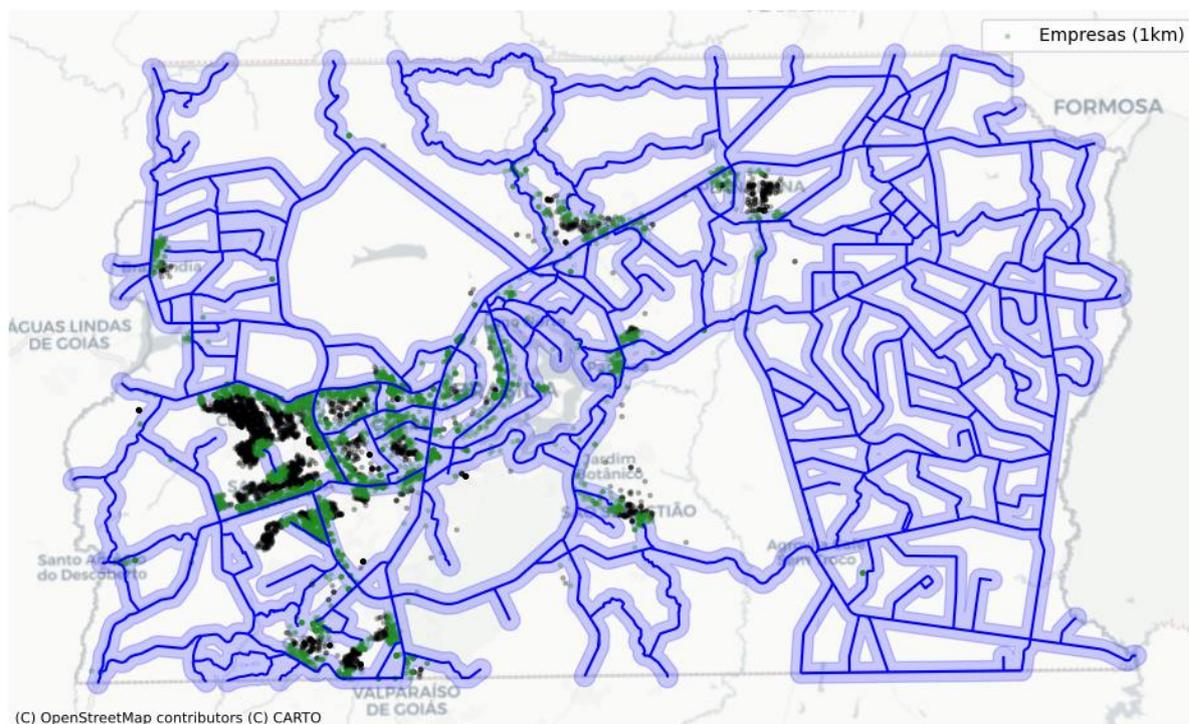
RA	Rodovias (km)	Área (km ²)	Densidade da rede viária (km/km ²)
ITAPOÃ	24,30	34,30	0,71
RECANTO DAS EMAS	58,41	93,11	0,63
SANTA MARIA	77,33	130,89	0,59
SIA	16,80	29,05	0,58
GAMA	145,47	275,41	0,53
SOBRADINHO	101,89	193,53	0,53
BRAZLÂNDIA	247,28	470,18	0,53
PLANALTINA	792,77	1509,41	0,53
PARK WAY	61,55	118,41	0,52
GUARÁ	13,08	25,22	0,52
LAGO SUL	37,05	76,05	0,49
ÁGUAS CLARAS	4,36	9,19	0,47
VICENTE PIRES	20,33	42,88	0,47
ARAPOANGA	9,24	21,99	0,42
CEILÂNDIA	77,18	192,66	0,40
TAGUATINGA	24,65	63,04	0,39
PARANOÁ	283,15	830,25	0,34
FERCAL	48,60	154,38	0,31
SÃO SEBASTIÃO	77,84	262,07	0,30
SAMAMBAIA	28,39	99,99	0,28
SOBRADINHO II	43,32	181,35	0,24
JARDIM BOTÂNICO	67,65	290,61	0,23
RIACHO FUNDO II	8,32	38,47	0,22
ARNIQUEIRA	2,74	13,36	0,21
PLANO PILOTO	80,21	437,20	0,18
SOL NASCENTE E PÔR DO SOL	5,18	40,49	0,13
RIACHO FUNDO	1,63	17,78	0,09
Total	2456,15	5766,59	-

Fonte: elaborado pelo autor (2023)

As regiões administrativas com maior densidade de rede viária são, respectivamente, Varjão, Cruzeiro, Núcleo Bandeirante, Candangolândia e SCIA. Juntas, essas regiões representam menos de 1% da área total do DF (Tabela 5). As regiões de Planaltina, Paranoá e Brazlândia são as maiores, tanto em área quanto em extensão de rodovias, o que faz com que elas não se destaquem em termos de densidade da rede viária. A Figura 9 apresenta a Densidade da rede viária para as diferentes RA.

rodovia. A análise da distribuição das empresas dentro desses *Buffers* ajuda a entender a influência da proximidade das rodovias em sua localização.

Figura 10 – Rodovias, *Buffers* e empresas



Fonte: elaborado pelo autor (2023)

Conforme é possível notar na Figura 10, a maioria das empresas está localizada a até 1 km de uma rodovia. Esse resultado também pode ser visto na Tabela 6, que apresenta a quantidade de empresas dentro dos *Buffers* das rodovias para cada região administrativa. No total, são 5016 pontos que estão localizados a até 1 km de uma rodovia, isto é, dentro dos *Buffers*, o que representa 60,24% do total de empresas.

Tabela 6 – Empresas localizadas a até 1 km de uma rodovia (continua)

RA	Total de empresas na RA	Empresas até 1 km de uma rodovia	% do total de empresas RA
PARANOÁ	97	97	100,00%
FERCAL	60	60	100,00%
ÁGUA QUENTE	31	31	100,00%
VARJÃO	24	24	100,00%
SIA	302	300	99,34%
CRUZEIRO	33	32	96,97%
CANDANGOLÂNDIA	62	60	96,77%
NÚCLEO BANDEIRANTE	118	113	95,76%
SCIA	97	91	93,81%

Tabela 6 – Empresas localizadas a até 1 km de uma rodovia (conclusão)

RA	Total de empresas na RA	Empresas até 1 km de uma rodovia	% do total de empresas RA
PLANO PILOTO	351	328	93,45%
LAGO NORTE	31	28	90,32%
BRAZLÂNDIA SUDOESTE/OCTOGONAL	139	116	83,45%
PARK WAY	43	35	81,40%
ITAPOÃ	55	44	80,00%
SOBRADINHO	81	63	77,78%
GAMA	163	126	77,30%
VICENTE PIRES	412	307	74,51%
RIACHO FUNDO	337	247	73,29%
ARNIQUEIRA	116	83	71,55%
RIACHO FUNDO II	215	152	70,70%
JARDIM BOTÂNICO	221	154	69,68%
TAGUATINGA	60	40	66,67%
RECANTO DAS EMAS	854	557	65,22%
GUARÁ	499	320	64,13%
SANTA MARIA	301	183	60,80%
SÃO SEBASTIÃO	471	275	58,39%
SOBRADINHO II	181	103	56,91%
SAMAMBAIA	159	86	54,09%
ÁGUAS CLARAS	905	483	53,37%
PLANALTINA	84	39	46,43%
LAGO SUL	250	114	45,60%
ARAPOANGA	40	18	45,00%
CEILÂNDIA SOL NASCENTE E PÔR DO SOL	97	42	43,30%
CEILÂNDIA SOL NASCENTE E PÔR DO SOL	1391	265	19,05%
DO SOL	47	0	0,00%
Total	8327	5016	60,24%

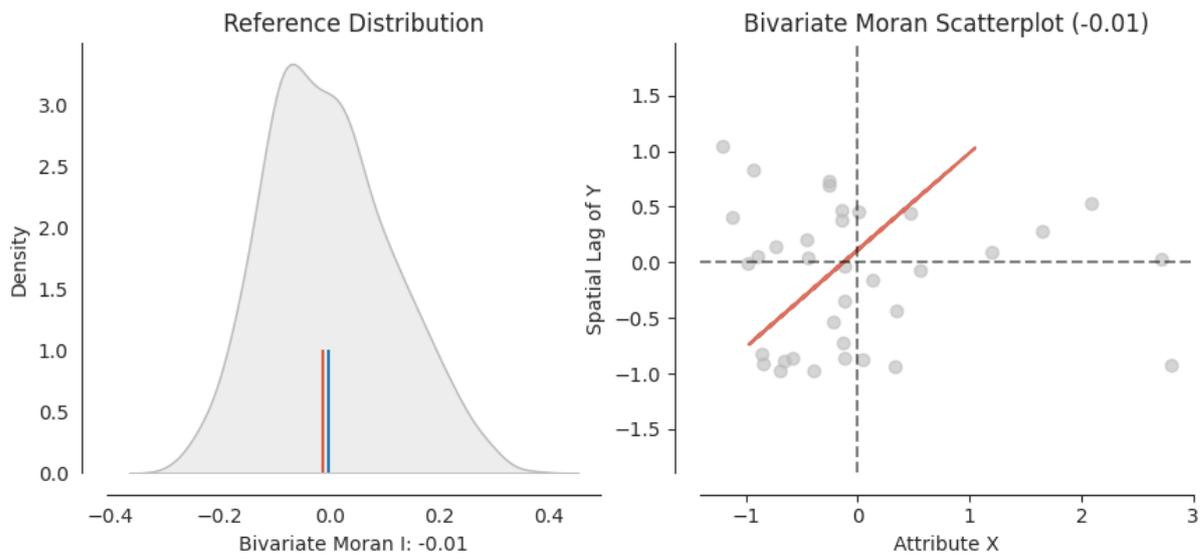
Fonte: elaborado pelo autor (2023)

Na maioria das regiões administrativas, a quantidade de empresas localizadas próximas de rodovias em relação ao total de empresas na RA é alta, destacando-se as regiões do Paranoá, Fercal, Água Quente, Varjão, SIA, Cruzeiro, Candangolândia e Núcleo Bandeirante, cujo o total de empresas próximas à rodovias supera 95% do total de empresas na RA. Essa relação fica abaixo de 50% apenas nas regiões de Águas Claras, Planaltina, Lago Sul, Arapoanga, Ceilândia e Sol Nascente e Pôr do Sol (Tabela 6).

Para analisar como densidades de empresas e de rede viária estão distribuídas nas diferentes RA, foram calculados os índices de Moran bivariados (global e

local). A Figura 11 apresenta o Diagrama de Dispersão de Moran Bivariado (*Bivariate Moran Scatterplot*) e a distribuição dos índices calculados nas permutações. A análise do *Moran's I* bivariado para as variáveis Densidade de empresas e Densidade da rede viária revelou um grau baixo de associação espacial, próximo de zero. O índice global foi de -0,01 com p-valor de 0,499, indicando que não há evidências suficientes para rejeitar a hipótese nula de aleatoriedade espacial.

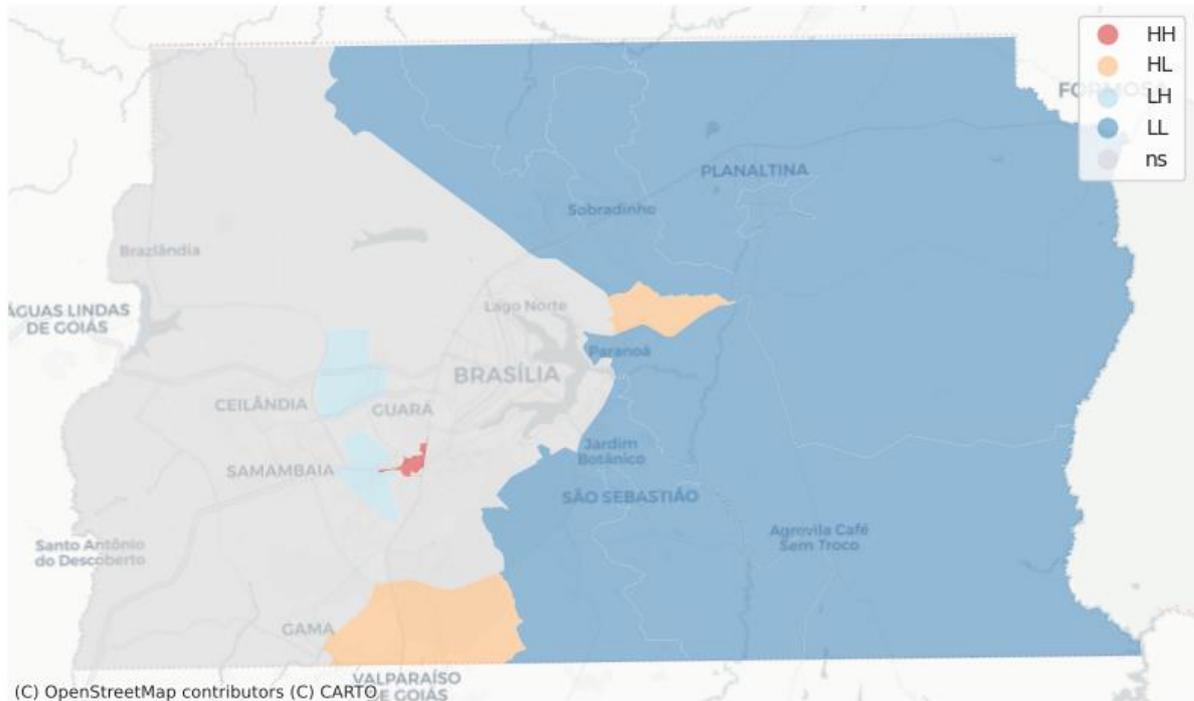
Figura 11 – Diagrama de Dispersão de Moran para as variáveis Densidade de empresas e Densidade da rede viária



Fonte: elaborado pelo autor (2023)

A análise do caso bivariado (Figura 11) é semelhante à do caso univariado (Figura 6). O Diagrama de Dispersão de Moran Bivariado apresenta a relação entre o valor de uma variável de análise na região (densidade da rede viária) e o seu *lag* espacial para outra variável (densidade de empresas). Assim, o índice bivariado permite verificar se regiões com altos ou baixos de densidade da rede viária tem vizinhos com altas ou baixos de densidade de empresas. Conforme apresentado na Figura 11, o *Moran's I* global bivariado é negativo e próximo de zero, indicando independência espacial entre densidade da rede viária e densidade de empresas. A Figura 12 apresenta os *clusters* espaciais para as variáveis densidade de empresas e densidade da rede viária, obtidos com base nos índices de Moran locais.

Figura 12 – Mapa de identificação de *clusters* e regiões isoladas para as variáveis Densidade da rede viária e Densidade de empresas



Fonte: elaborado pelo autor (2023)

Na Figura 12, é possível identificar uma região HH, duas regiões HL, três regiões LH e 8 regiões LL. As regiões que compõem esses grupos são apresentadas na Tabela 7. Efetivamente, há apenas um *cluster* LL, composto por 8 regiões, e um *cluster* LH, composto por 2 regiões. A terceira região LL e as regiões HL representam áreas isoladas (Tabela 7 e Figura 12).

Tabela 7 – *Clusters* e regiões isoladas para as variáveis densidade de rede viária e densidade de empresas

HH	HL	LH	LL
Núcleo Bandeirante	Itapoã	Riacho Fundo	Sobradinho
	Santa Maria	Vicente Pires	Planaltina
		Arniqueira	Paranoá
			São Sebastião
			Sobradinho II
			Jardim Botânico
			Fercal
			Arapoanga

Fonte: elaborado pelo autor (2023)

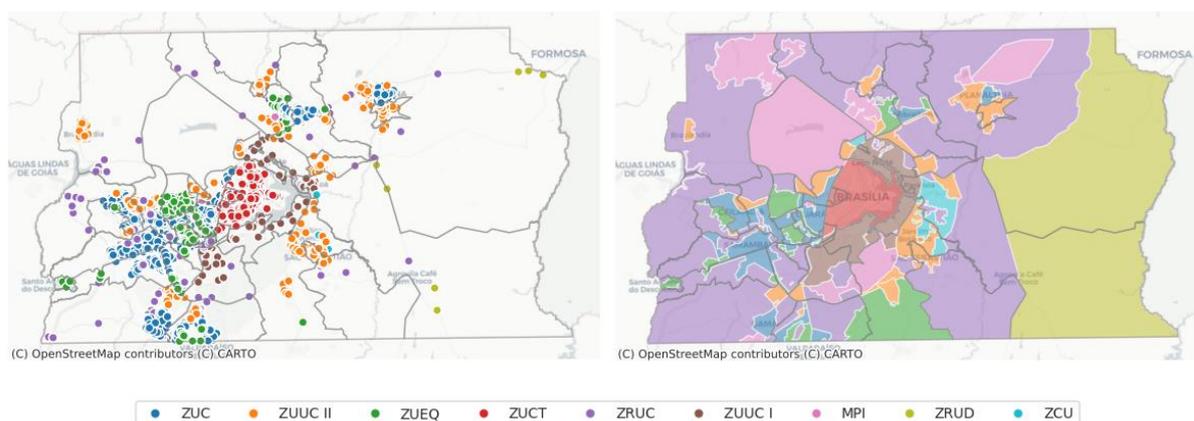
Em resumo, a análise das empresas em relação às rodovias mostrou que grande parte das empresas, 5016 pontos de um total de 8327 (Figura 10 e Tabela 6)

está localizada próximo a uma rodovia, indicando que há uma influência da proximidade das rodovias na distribuição espacial das empresas. No entanto, uma análise mais profunda com os índices de Moran não revelou associação espacial significativa entre a Densidade de empresas e Densidade da rede viária em nível global. Nos índices locais, destaca-se o *cluster* LL, no qual as regiões com valores baixos da primeira variável (Densidade da rede viária) estão cercadas por outras regiões com valores baixos da segunda variável (Densidade de empresas) (Figura 11 e Tabela 7).

4.5. Influência da política de zoneamento do DF na localização das instalações logísticas

O zoneamento do DF pode ajudar na compreensão dos resultados apresentados nas subseções anteriores. A Figura 13 apresenta a distribuição das empresas por tipo de zona, revelando que a maioria das empresas está concentrada nas regiões que compõem a Macrozona Urbana, dentre as quais se destacam a ZUC (Zona Urbana Consolidada) e a ZUUC II (Zona Urbana de Uso Controlado II) e a ZUEQ (Zona Urbana de Expansão e Qualificação).

Figura 13 – Empresas e zoneamento



Fonte: elaborado pelo autor (2023)

A Tabela 8 apresenta as frequências absoluta e relativa das empresas por tipo de zona. A ZUC, ZUUC II e ZUEQ concentram 88% das empresas, destacando-se a ZUC, que concentra, sozinha, 65%. Analisando a Figura 13 é possível verificar que a ZUC e a ZUEQ estão localizadas, principalmente, ao sudoeste do DF, enquanto a ZUUC II está mais dispersa no mapa. Além disso, as regiões com baixa concentração

de pontos são (Figura 3, Figura 4 e Figura 7), majoritariamente, compostas de zonas rurais (ZRUD e ZRUC), que são destinadas principalmente a atividades do setor primário da economia, em contraste com as zonas urbanas, destinadas predominantemente aos setores secundário e terciário (Distrito Federal, 2009).

Tabela 8 – Empresas por zona

Zona	Sigla	Frequência Absoluta (FA)	Frequência Relativa (FR)	FR Acum.
ZONA URBANA CONSOLIDADA	ZUC	5468	65,67%	65,67%
ZONA URBANA DE USO CONTROLADO II	ZUUC II	1236	14,84%	80,51%
ZONA URBANA DE EXPANSÃO E QUALIFICAÇÃO	ZUEQ	695	8,35%	88,86%
ZONA URBANA DO CONJUNTO TOMBADO	ZUCT	488	5,86%	94,72%
ZONA RURAL DE USO CONTROLADO	ZRUC	213	2,56%	97,27%
ZONA URBANA DE USO CONTROLADO I	ZUUC I	189	2,27%	99,54%
MACROZONA DE PROTEÇÃO INTEGRAL	MPI	17	0,20%	99,75%
ZONA RURAL DE USO DIVERSIFICADO	ZRUD	16	0,19%	99,94%
ZONA DE CONTENÇÃO URBANA	ZCU	5	0,06%	100,00%
Total	9	8327	100%	-

Fonte: elaborado pelo autor (2023)

Por fim, a Lei de Uso e Ocupação do Solo do DF (LUOS), define categorias de uso comercial levando em consideração a acessibilidade aos núcleos urbanos (Distrito Federal, 2019). Assim, espera-se que as atividades comerciais e de prestação de serviços dentro ou próximas das zonas urbanas, sendo que neste último caso as empresas podem estar próximas de rodovias, o que ajuda explicar a concentração de pontos próximos das rodovias do DF (Figura 10 e Tabela 6).

4.6. Discussão

A análise dos pontos que identificam as empresas do setor de transporte e armazenamento do DF evidencia a concentração desses pontos em áreas ao sudoeste da região. Assim, as empresas logísticas do DF estão espacialmente

concentradas, resultado similar ao dos estudos de Heuvel *et. al* (2013). Uma explicação possível para isso é o zoneamento do DF, no qual as zonas urbanas, destinadas à atividade comercial e de prestação de serviços, estão localizadas majoritariamente ao sudoeste do DF. Os trabalhos de Sakai, Beziat e Heitz (2020) e Giuliano e Kang (2018) apontam o zoneamento local como um dos fatores de influência na atividade de instalações de armazenamento e distribuição. Giuliano e Kang (2018) também apontam que a localização próxima ao mercado consumidor é preferível para estabelecimentos logísticos que fazem viagens frequentes. No caso do DF, esse os consumidores estão localizados, em grande parte, nas zonas urbanas que concentram o maior número de instalações logísticas.

A proximidade entre empresas e rodovias pode indicar que esse é um fator de influência na localização dos estabelecimentos logísticos, conforme apontado no estudo de Jakubicek (2010). O grau de associação espacial entre a infraestrutura de transportes (densidade da rede viária) e localização das instalações logísticas (densidade de empresas) encontrado, negativo e baixo, que é diferente do valor encontrado no trabalho de Guerin *et al.* (2021) sobre a Região Metropolitana de São Paulo (RMSP), que é positivo e baixo. Há três possíveis explicações para isso: 1) mais empresas foram incluídas, ou seja, mais códigos CNAE foram utilizados, 2) as variáveis utilizadas neste trabalho foram diferentes das de Guerin *et al.* (2021) e 3) as duas regiões, DF e RMSP, são muito distintas. Assim, embora a análise de proximidade entre os pontos (empresas) e as rodovias tenha indicado uma relação entre os dois, não foi possível confirmar essa relação quando se passou para nível das regiões administrativas.

5. CONCLUSÕES E RECOMENDAÇÕES

O objetivo geral do trabalho foi analisar a distribuição espacial das instalações logísticas do Distrito Federal em relação à disponibilidade de infraestrutura de transporte na região.

Verificou-se que as instalações logísticas estão localizadas, majoritariamente, ao sudoeste do DF. A análise de pontos com o centro médio e SDE revelou que apenas 20 das 35 regiões administrativas do DF concentram 61% das empresas do setor

de transporte e armazenamento. Além disso, os pontos estão concentrados em uma área pequena em relação à área total do DF, reforçando a concentração espacial dos estabelecimentos logísticos. Dentre essas 20 regiões, destacam-se Ceilândia, Taguatinga e Samambaia, que têm maior ocorrência de empresas e representam a maioria dos pontos localizados na SDE.

Quando se analisa em termos relativos, isto é, a densidade de empresas, chega-se a um grupo menor de 5 regiões administrativas com maior concentração de empresas, que estão entre 20 regiões identificadas na análise dos pontos. Essas regiões, identificadas por meio dos índices de Moran, são: Núcleo Bandeirante, Riacho Fundo, Águas Claras, Vicente Pires e Arniqueira. Juntas, essas regiões constituem um *cluster* com alta densidade de empresas. Desse modo, tanto em termos absolutos (pontos) quanto relativos (densidade), pode-se identificar o sudoeste do DF como uma área de concentração logística (Heuvel et. al, 2013), em contraste com as regiões ao leste do DF, como Sobradinho, Planaltina, Paranoá, São Sebastião, Sobradinho II, Jardim Botânico, Itapoã, Fercal e Arapoanga, identificadas como de baixa concentração de empresas.

Por fim, verificou-se que grande parte das empresas está localizada próxima a uma rodovia, o que pode indicar que existe influência da proximidade das rodovias na distribuição espacial das empresas. Essa relação, no entanto, não é evidenciada nos índices de Moran calculados com variáveis Densidades de empresas e Densidade da rede viária, os quais não revelaram *clusters* de valores altos das duas variáveis (HH), apenas uma área isolada, o Núcleo Bandeirante. O leste do DF se destaca pelo *cluster* de valores baixos das duas variáveis (LL).

Esta pesquisa mostrou-se relevante pois trouxe elementos importantes que contribuem para gestores de empresas logísticas e formuladores de políticas públicas, que são: 1) a identificação de uma área de concentração logística no DF, 2) a proximidade dos estabelecimentos logísticos em relação às rodovias e 3) o zoneamento como um possível fator de influência na localização das instalações. Além disso, este trabalho explora um tema importante na logística na capital de um país emergente, onde estudos sobre a localização de estabelecimentos logísticos são menos frequentes, utilizando análise geoespacial.

Esta pesquisa apresenta algumas limitações, sobretudo quanto à identificação das instalações logísticas por meio do CNAE. Embora a utilização de uma da classificação de atividade esteja presente na literatura, sabe-se que ela tem falhas, sendo o

exemplo mais claro disso uma empresa com uma atividade primária cadastrada não relacionada à logística, mas que não reflete a real atividade econômica.

Pesquisas futuras podem ser conduzidas de modo a melhorar a identificação das instalações logísticas no DF, como o uso de imagens de satélite para verificar visualmente a existência das instalações. Além disso, a adição de outras variáveis relacionadas às características das instalações, como porte e tipo de instalação, pode fornecer outros *insights*, assim como outras características da região de estudo, como níveis menores de zoneamento e categorias de uso de solo nas regiões administrativas.

REFERÊNCIA

- ALLEN, J.; BROWNE, M.; CHERRETT, T. Investigating relationships between road freight transport, facility location, logistics management and urban form. **Journal of Transport Geography**, [s. l.], v. 24, p. 45–57, 2012.
- ANSELIN, Luc. Local indicators of spatial association—LISA. **Geographical analysis**, [s. l.], v. 27, n. 2, p. 93–115, 1995.
- ANSELIN, Luc; REY, Sergio J. Modern spatial econometrics in practice: A guide to GeoDa, GeoDaSpace and PySAL 2014.
- ANSELIN, Luc; SYABRI, Ibnu; SMIRNOV, Oleg. Visualizing multivariate spatial correlation with dynamically linked windows. *In*: 2002. **Proceedings, CSISS Workshop on New Tools for Spatial Data Analysis, Santa Barbara, CA**. [S. l.: s. n.], 2002.
- BALLOU, Ronald H. **Gerenciamento da Cadeia de Suprimentos-: Logística Empresarial**. [S. l.]: Bookman editora, 2009.
- BOWEN, John T. Moving places: the geography of warehousing in the US. **Journal of Transport Geography**, [s. l.], v. 16, n. 6, p. 379–387, 2008.
- BOWERSOX, Donald J *et al.* **Gestão logística da cadeia de suprimentos**. [S. l.]: AMGH Editora, 2013.
- BRASIL [Constituição (1988)]. **Constituição da República Federativa do Brasil de 1988**. Brasília: Senado Federal, 1988.
- BRASIL. Ministério da Fazenda. Secretaria Especial da Receita Federal do Brasil (RFB).
- BRASIL. Ministério da Infraestrutura. Departamento Nacional de Infraestrutura de Transportes (DNIT).
- BRASIL. Ministério da Infraestrutura. **Plano Nacional de Logística**. 2021.
- BRASÍLIA. Governo do Distrito Federal. **Lei de Uso e Ocupação do Solo do Distrito Federal**. Brasília, DF: Câmara Legislativa do Distrito Federal, 2009
- BRASÍLIA. Governo do Distrito Federal. **Plano Diretor de Ordenamento Territorial do Distrito Federal (PDOT-DF)**. Brasília, DF: Secretaria de Estado de Gestão do Território e Habitação, 2019.
- BULDEO RAI, Heleen *et al.* 'Proximity logistics': Characterizing the development of logistics facilities in dense, mixed-use urban areas around the world. **Transportation Research Part A: Policy and Practice**, [s. l.], v. 166, p. 41–61, 2022.

CERVO, Amado Luiz; BERVIAN, Pedro Alcino. Metodologia científica. *In*: METODOLOGIA CIENTÍFICA. [S. l.: s. n.], 1996.

CNT. **Anuário CNT do Transporte**. Confederação Nacional do Transporte. 2022

CNT. **Terminais de Carga do Brasil**. Confederação Nacional do Transporte. 2021

CHRISTENSEN, Larry B *et al.* Research methods, design, and analysis. [s. l.], 2011.

CROUCHER, Phil; BAKER, Peter; RUSHTON, Alan. **The handbook of logistics and distribution management: Understanding the supply chain**. [S. l.]: Kogan Page, 2014.

CUCHAVA ROCHA, Cynthia Vargas; GOUVEA CAMPOS, Vania Barcellos; DE MELLO BANDEIRA, Renata Albergaria. Mobility index for urban freight transportation. **Urbe**, [s. l.], v. 11, 2019.

DANTAS GABRIELE, Pedrita *et al.* Comparação internacional da eficiência ambiental dos modos de transporte rodoviário e ferroviário [The environmental efficiency of road and rail freight modes in the world]. **Journal of Transport Literature**, [s. l.], v. 7, n. 1, p. 212–229, 2013.

DE CASTRO, Newton. Mensuração de externalidades do transporte de carga brasileiro [Measurement of externalities of freight transport in Brazil]. **Journal of Transport Literature**, [s. l.], v. 7, n. 1, p. 163–181, 2013.

DE OLIVEIRA, Leise Kelli *et al.* Locational context for warehouse facilities in urban areas: A case study in Belo Horizonte (Brazil). 2020. **Transportation Research Procedia**. [S. l.]: Elsevier B.V., 2020. p. 401–415.

DE OLIVEIRA, Maxwell Ferreira. Metodologia científica: um manual para a realização de pesquisas em Administração. **Universidade Federal de Goiás**. Catalão–GO, [s. l.], 2011.

DE SMITH, M; GOODCHILD, M; LONGLEY, P. **Geospatial analysis: a comprehensive guide to principles, techniques and software tools sixth edition**. [S. l.]: Troubador Publishing LTD, 2021.

DISTRITO FEDERAL. Companhia de Planejamento do Distrito Federal. **Transporte de Carga no Distrito Federal: Questões e desafios**. 2018.

DISTRITO FEDERAL. Secretaria de Estado de Desenvolvimento Urbano e Habitação (SEDUH)

DORNELES, Ana. O Zoneamento e Sua Importância como um Instrumento de Planejamento Urbano. **Cadernos da Escola de Direito e Relações Internacionais**, [s. l.], v. 1, n. 13, p. 452–467, 2010.

- DUARTE, Fábio. **Planejamento urbano**. [S. l.]: Editora Ibpex, 2009.
- ERKAN, Birol. The importance and determinants of logistics performance of selected countries. **Journal of Emerging Issues in Economics, Finance and Banking**, v. 3, n. 6, p. 1237-1254, 2014.
- FRANCOIS, Joseph; MANCHIN, Miriam. Institutions, Infrastructure, and Trade. **World Development**, [s. l.], v. 46, p. 165–175, 2013.
- GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social**. [S. l.]: 6. ed. Editora Atlas SA, 2008.
- GIULIANO, Genevieve; KANG, Sanggyun. Spatial dynamics of the logistics industry: Evidence from California. **Journal of Transport Geography**, [s. l.], v. 66, p. 248–258, 2018.
- GUERIN, Leonardo *et al.* The geography of warehouses in the São Paulo Metropolitan Region and contributing factors to this spatial distribution. **Journal of Transport Geography**, [s. l.], v. 91, 2021.
- HEITZ, Adeline *et al.* Spatial patterns of logistics facilities in Gothenburg, Sweden. **Journal of Transport Geography**, [s. l.], v. 88, 2020.
- HEITZ, Adeline. The logistics dualization in question: Evidence from the Paris metropolitan area. **Cities**, [s. l.], v. 119, 2021.
- HESSE, Markus; RODRIGUE, Jean-Paul. The transport geography of logistics and freight distribution. **Journal of transport geography**, [s. l.], v. 12, n. 3, p. 171–184, 2004.
- JAKUBICEK, Paul. **Understanding the location choices of logistics firms**. 2010. Dissertação de Mestrado. University of Waterloo.
- KANG, Sanggyun. Warehouse location choice: A case study in Los Angeles, CA. **Journal of Transport Geography**, [s. l.], v. 88, 2020.
- KIN, Bram; VERLINDE, Sara; MACHARIS, Cathy. **Sustainable urban freight transport in megacities in emerging markets**. [S. l.]: Elsevier Ltd, 2017.
- LAVILLE, Christian; DIONNE, Jean. A construção do saber. **Belo Horizonte: UFMG**, [s. l.], v. 340, p. 1990, 1999.
- LEVY, John M. **Contemporary urban planning**. [S. l.]: Taylor & Francis, 2016.
- LONGLEY, Paul A *et al.* **Geographic information science and systems**. [S. l.]: John Wiley & Sons, 2015.

- MEERSMAN, Hilde; NAZEMZADEH, Marzieh. The contribution of transport infrastructure to economic activity: The case of Belgium. **Case Studies on Transport Policy**, [s. l.], v. 5, n. 2, p. 316–324, 2017.
- PARK, Soonchan. Quality of transport infrastructure and logistics as source of comparative advantage. **Transport Policy**, [s. l.], v. 99, p. 54–62, 2020.
- PRUD'HOMME, Remy. **Infrastructure and development**. [S. l.]: World Bank Washington, DC, 2004.
- RAHMAN, Imran Ur *et al.* Infrastructure and Trade: An Empirical Study Based on China and Selected Asian Economies. **SAGE Open**, [s. l.], v. 11, n. 3, 2021.
- RODRIGUE, Jean-Paul. **The geography of transport systems**. [S. l.]: Routledge, 2020.
- ROUMBOUTSOS, Athena; KAPROS, Seraphim; VANELSLANDER, Thierry. Green city logistics: Systems of Innovation to assess the potential of E-vehicles. **Research in Transportation Business and Management**, [s. l.], v. 11, p. 43–52, 2014.
- ŞAHAN, Duygu; TUNA, Okan. Policy Implications on Transport Infrastructure–Trade Dynamics: Case of Turkey. **Logistics**, [s. l.], v. 5, n. 3, 2021.
- SAKAI, Takanori; BEZIAT, Adrien; HEITZ, Adeline. Location factors for logistics facilities: Location choice modeling considering activity categories. **Journal of Transport Geography**, [s. l.], v. 85, 2020.
- VAN DEN HEUVEL, Frank P. *et al.* Spatial concentration and location dynamics in logistics: The case of a Dutch province. **Journal of Transport Geography**, [s. l.], v. 28, p. 39–48, 2013.
- VIEIRA, Yesus Emmanuel Medeiros *et al.* Government measures for urban freight transport: Perceptions on Brazilian scenario. **Urbe**, [s. l.], v. 13, 2021
- VLAHINIĆ LENZ, Nela; PAVLIĆ SKENDER, Helga; MIRKOVIĆ, Petra Adelajda. The macroeconomic effects of transport infrastructure on economic growth: the case of Central and Eastern E.U. member states. **Economic Research-Ekonomska Istrazivanja**, [s. l.], v. 31, n. 1, p. 1953–1964, 2018.
- WALLIMAN, Nicholas. **Research methods: The basics**. [S. l.]: Routledge, 2021.
- WANG, Chao *et al.* Railway and road infrastructure in the Belt and Road Initiative countries: Estimating the impact of transport infrastructure on economic growth. **Transportation Research Part A: Policy and Practice**, [s. l.], v. 134, p. 288–307, 2020.
- WOUDSMA, Clarence; JAKUBICEK, Paul. Logistics land use patterns in metropolitan Canada. **Journal of Transport Geography**, [s. l.], v. 88, 2020.

YUILL, Robert S. The *Standard Deviation Ellipse*; an updated tool for spatial description. **Geografiska Annaler: Series B, Human Geography**, [s. l.], v. 53, n. 1, p. 28–39, 1971.

ZHANG, Yijia; CHENG, Lu. The role of transport infrastructure in economic growth: Empirical evidence in the UK. **Transport Policy**, [s. l.], v. 133, p. 223–233, 2023.

APÊNDICE

Código em linguagem Python

```
# -*- coding: utf-8 -*-
"""monografia.ipynb

Automatically generated by Colaboratory.

Original file is located at
https://colab.research.google.com/drive/13hdEecwRLHCFWiX8twJq6ZwjKPix5YBm

## Download de bibliotecas
"""

!pip install pointpats
!pip install libpysal
!pip install splot
!pip install esda
!pip install contextily
!pip install shapely

"""## Importar bibliotecas"""

# Commented out IPython magic to ensure Python compatibility.
# %matplotlib inline
import matplotlib.pyplot as plt
import matplotlib.ticker as mtick
import pointpats as pp
from matplotlib.lines import Line2D
from splot.libpysal import plot_spatial_weights
from libpysal.weights.contiguity import Queen
from libpysal.weights import lag_spatial
from esda.moran import Moran
from splot.esda import plot_moran
from esda.moran import Moran_Local
from splot.esda import lisa_cluster
from esda.moran import Moran_BV, Moran_Local_BV
from splot.esda import plot_moran_bv_simulation, plot_moran_bv
import numpy as np
import seaborn as sns
import pandas as pd
import geopandas as gpd
import contextily as ctx
from shapely.geometry.point import Point
from shapely.plotting import plot_line, plot_polygon
from shapely import LinearRing, Polygon, MultiPolygon
import shapely.affinity

sns.set_style('white')
sns.set_palette('tab10')

"""## Leitura dos dados

### Empresas (RFB)
```

```

"""

# 30 colunas no total
colunas = list(range(0, 30))

# colunas de interesse
colunas_mantidas = [0, 1, 2, 5, 11, 14, 17, 18, 19]

# Abrindo arquivos zip das tabelas (base do dia 08/10/2023)
lista1 = []
for i in range(0, 10):
    chunk = pd.read_csv(f'.../data/Estabelecimentos{i}.zip', header=None,
                        delimiter=';', encoding = 'latin1',
                        on_bad_lines='skip', chunksize=1000000,
                        low_memory=False, names=colunas, usecols=colunas_mantidas)
    lista1.append(chunk)

# Renomeando colunas e filtrando dados (UF == DF e Situação Cadastral == 2,
ativa)
lista2 = []
for i in range(0, 10):
    empresas_df = pd.concat(lista1[i])

    # Renomear colunas
    empresas_df.rename(columns={0:'CNPJ BÁSICO', 1:'CNPJ ORDEM', 2:'CNPJ DV',
5:'SITUAÇÃO CADASTRAL',
11:'CNAE FISCAL PRINCIPAL', 14:'LOGRADOURO',
17:'BAIRRO',
18:'CEP', 19:'UF'}, inplace = True)

    # Manter somente cadastros ativos do DF
    empresas_df = empresas_df.loc[(empresas_df['SITUAÇÃO CADASTRAL']==2) &
(empresas_df['UF']=='DF')]

    lista2.append(empresas_df)

# Criar DataFrame
empresas_df = pd.concat(lista2)

# Empresas do DF
empresas_df.info()

lista1.clear()
lista2.clear()

# Formatar componentes do CNPJ
empresas_df['CNPJ BÁSICO'] = empresas_df['CNPJ BÁSICO'].apply(lambda x:
str('{0:0>8}'.format(x))) # Dígitos 1-8

empresas_df['CNPJ ORDEM'] = empresas_df['CNPJ ORDEM'].apply(lambda x:
str('{0:0>4}'.format(x))) # Dígitos 9-12

empresas_df['CNPJ DV'] = empresas_df['CNPJ DV'].apply(lambda x:
str('{0:0>2}'.format(x))) # Dígitos 13-14

# Criar coluna CNPJ

```

```

empresas_df['CNPJ'] = empresas_df['CNPJ BÁSICO'] + empresas_df['CNPJ
ORDEM'] + empresas_df['CNPJ DV']

# Retirar colunas antigas
empresas_df = empresas_df.loc[:, ~empresas_df.columns.isin(['CNPJ BÁSICO',
'CNPJ ORDEM', 'CNPJ DV'])]

empresas_df.head()

# Removendo linhas com valores nulos
empresas_df.dropna(axis=0, inplace=True)
empresas_df.isnull().sum()

empresas_df.info()

# Verificar CNPJ duplicados
empresas_df.drop_duplicates(subset=['CNPJ'], inplace=True)

empresas_df.info()

# Formatar CEP
empresas_df['CEP'] = empresas_df['CEP'].apply(lambda x: str(int(x))) #
tornar string

empresas_df['CEP'] = empresas_df['CEP'].apply(lambda x: x[:5] + '-' +
x[5:]) # formato XXXXX-XXX

empresas_df.head()

"""#### Filtrar empresas por tipo de atividade (CNAE)

Códigos CNAE mantidos:
* 4930-2/01 Transporte rodoviário de carga, exceto produtos perigosos e
mudanças, municipal;
* 4930-2/02 Transporte rodoviário de carga, exceto produtos perigosos e
mudanças, intermunicipal, interestadual e internacional;
* 4930-2/03 Transporte rodoviário de produtos perigosos;
* 5211-7/01 Armazéns gerais - emissão de warrant;
* 5211-7/99 Depósitos de mercadorias para terceiros, exceto armazéns gerais
e guarda-móveis;
* 5212-5/00 Carga e descarga;
* 5250-8/03 Agenciamento de cargas, exceto para o transporte marítimo;
* 5250-8/04 Organização logística do transporte de carga;
* 5250-8/05 Operador de transporte multimodal - OTM.
"""

# Códigos CNAE
cnaes = [4930201, 4930202, 4930203, 5211701, 5211799,
         5212500, 5250803, 5250804, 5250805]

# Filtrar empresas por CNAE
empresas_df_final = empresas_df[empresas_df['CNAE FISCAL
PRINCIPAL'].isin(cnaes)].copy()

empresas_df_final.info()

```

```

# Frequência
empresas_df_final.groupby('CNAE FISCAL
PRINCIPAL').size().sort_values(ascending=False)

# Resetar índice
empresas_df_final.reset_index(drop=True, inplace=True)

# Criar coluna com o endereço completo
empresas_df_final['ENDERECO'] = empresas_df_final['LOGRADOURO'].copy() + ',
' + empresas_df_final['BAIRRO'].copy() + ', ' + 'Brasília - DF' + ', ' +
empresas_df_final['CEP'].copy()

# Excluindo colunas com componentes do endereço
empresas_df_final = empresas_df_final.loc[:,
~empresas_df_final.columns.isin(['LOGRADOURO', 'BAIRRO'])]

print(empresas_df_final['ENDERECO'][500]) #

"""### Limites das Regiões Administrativas (RA)"""

ras = gpd.read_file('../data/regioes_administrativas (1).zip',
encoding='utf-8')

ras.info()

# Mostrar sistema de coordenadas
ras.crs

"""### Geocodificar endereços

Provedor (consulta) = ArcGIS
"""

# Geocodificar endereços
# enderecos_gdf = gpd.tools.geocode(empresas_df_final['ENDERECO'],
provider='arcgis')

# Salvar arquivo no Google Drive
# enderecos_gdf.to_file('../data/enderecos_gdf.gpkg', driver='GPKG')

enderecos_gdf = gpd.read_file('../data/enderecos_gdf.gpkg', driver='GPKG')

enderecos_gdf.head()

# Mostrar sistema de coordenadas
enderecos_gdf.crs

enderecos_gdf.iloc[5678]

empresas_df_final.iloc[5678]

# Concatenar dados: empresas (RFB) + Endereços (coordenadas)
empresas_gdf = pd.concat([empresas_df_final, enderecos_gdf],
axis='columns')
empresas_gdf.head()

```

```

type(empresas_gdf)

# Criar GeoDataFrame
empresas_gdf = gpd.GeoDataFrame(empresas_gdf, geometry='geometry',
crs=enderecos_gdf.crs.to_string())

empresas_gdf.info()

# Converter sistema de coordenadas
empresas_gdf = empresas_gdf.to_crs(ras.crs.to_string())

empresas_gdf.crs

enderecos_ol = gpd.overlay(empresas_gdf, ras, how='intersection')
enderecos_ol.head()

# 8327 empresas com endereços válidos
enderecos_ol.info()

"""### Zoneamento do DF"""

# Zoneamento do DF
zoneamento = gpd.read_file('../data/zoneamento_do_distrito_federal.zip',
encoding='utf-8')

zoneamento.head()

# Empresas + RA + Zonas
dados_empresas_final = gpd.overlay(enderecos_ol, zoneamento,
how='intersection')
dados_empresas_final.head()

dados_empresas_final.info()

"""### Sistema viário: rodovias federais e distritais"""

# Rodovias estaduais (Brasil)
estadual = gpd.read_file('../data/cide_2021_.zip', encoding='utf-8')

# Mostrar sistema de coordenadas
estadual.crs

# Rodovias federais (Brasil)
federal = gpd.read_file('../data/SNV_202308A.zip', encoding='utf-8')

# Mostrar sistema de coordenadas
federal.crs

# Rodovias distritais (DF)
distrital_local = estadual.loc[(estadual['Unidade_Fe']=='DF')]

# Converter crs
distrital_local = distrital_local.to_crs(ras.crs.to_string())

# Rodovias federais (DF)
federal_local = federal.loc[(federal['sg_uf']=='DF')]

```

```

# Converter crs
federal_local = federal_local.to_crs(ras.crs.to_string())

# Rodovias distritais + RA
distrital_local_ras_ol = gpd.overlay(distrital_local, ras,
how='intersection')

# Calcular a extensão das rodovias distritais em km
distrital_local_ras_ol['extensao_distrital (km)'] =
distrital_local_ras_ol['geometry'].length/1000

# DataFrame extensão por RA
extensao_distrital_por_ra =
distrital_local_ras_ol.groupby(['ra_nome']).sum(numeric_only=True)

extensao_distrital_por_ra.head()

# Extensão total das rodovias distritais
round(extensao_distrital_por_ra['extensao_distrital (km)'].sum(), 2)

# Rodovias federais + RA
federal_local_ras_ol = gpd.overlay(federal_local, ras, how='intersection')

# Calcular a extensão das rodovias federais em km
federal_local_ras_ol['extensao_federal (km)'] =
federal_local_ras_ol['geometry'].length/1000

# DataFrame extensão por RA
extensao_federal_por_ra =
federal_local_ras_ol.groupby(['ra_nome']).sum(numeric_only=True)

extensao_federal_por_ra.head()

# Extensão total das rodovias federais
round(extensao_federal_por_ra['extensao_federal (km)'].sum(), 2)

extensao_distrital_por_ra.keys()

# Retirar colunas
retirar = ['ogc_fid', 'Quilometra', 'Quilom', 'Extensao', 'ra_cira',
          'st_area_sh', 'st_length_']

extensao_distrital_por_ra = extensao_distrital_por_ra.loc[:,
~extensao_distrital_por_ra.columns.isin(retirar)]

extensao_federal_por_ra.keys()

# Retirar colunas
retirar = ['ogc_fid', 'id_trecho_', 'vl_km_inic', 'vl_km_fina',
          'vl_extensa', 'ds_ato_leg', 'id_versao', 'ra_cira', 'st_area_sh',
          'st_length_',]

extensao_federal_por_ra = extensao_federal_por_ra.loc[:,
~extensao_federal_por_ra.columns.isin(retirar)]

# Extensão por RA: distrital e federal

```

```

extensao_por_ra = pd.concat([extensao_distrital_por_ra,
extensao_federal_por_ra], axis=1)

# Preencher nulos com zero (0)
extensao_por_ra = extensao_por_ra.fillna(0)

round(extensao_por_ra['extensao_distrital (km)'].sum(), 2)

round(extensao_por_ra['extensao_federal (km)'].sum(), 2)

# Criar coluna com extensão total
extensao_por_ra['extensao_total (km)'] =
extensao_por_ra['extensao_distrital (km)'] +
extensao_por_ra['extensao_federal (km)']

extensao_por_ra.head()

temp_df = ras[['ra_cira', 'ra_nome']].sort_values(by=['ra_nome'])

# Adicionar coluna
v_ra_cira = temp_df['ra_cira'].values
extensao_por_ra.insert(loc=0, column='CIRA', value=v_ra_cira)

# Adicionar coluna
v_ra_nome = temp_df['ra_nome'].values
extensao_por_ra.insert(loc=1, column='RA', value=v_ra_nome)

# Novo índice
extensao_por_ra.sort_values(by='CIRA', inplace=True)
extensao_por_ra.reset_index(drop=True, inplace=True)
extensao_por_ra.head()

a = ras[['ra_cira',
'ra_nome']].sort_values(by=['ra_cira'])['ra_nome'].values
b = extensao_por_ra['RA'].values
a == b

# Reordenar
ras.sort_values(by='ra_cira', inplace=True)
ras.reset_index(drop=True, inplace=True)
ras.head()

temp_df = ras[['ra_cira', 'ra_nome']].sort_values(by=['ra_nome'])

# Frequência de empresas por RA
a = dados_empresas_final.groupby(['ra_nome']).count()

# Adicionar coluna
v_ra_cira = temp_df['ra_cira'].values
a.insert(loc=0, column='CIRA', value=v_ra_cira)

a.sort_values(by='CIRA', inplace=True)
# a.reset_index(drop=True, inplace=True)

a = a[['CIRA', 'CNPJ']]
a.reset_index(inplace=True)
a

```

```

ras['ra_nome'].values == extensao_por_ra['RA'].values

#a.to_excel('.../data/empresas_por_ra.xlsx')

dados_ra_final = extensao_por_ra.copy()
dados_ra_final['area_ra'] = ras['geometry'].area / 1000000
dados_ra_final['empresas_count'] = a['CNPJ']
dados_ra_final['densidade_rede_viaria'] = dados_ra_final['extensao_total
(km)'] / dados_ra_final['area_ra']
dados_ra_final['empresas_por_km2'] = dados_ra_final['empresas_count'] /
dados_ra_final['area_ra']
dados_ra_final['geometry'] = ras['geometry']
dados_ra_final = gpd.GeoDataFrame(dados_ra_final, geometry='geometry',
crs=ras.crs.to_string())

dados_ra_final.head()

ras['ra_codigo'].values

dados_ra_final['CIRA'].values

#dados_ra_final.to_excel('.../data/dados_ra_final.xlsx')

"""## Mapeamento das instalações logísticas no Distrito Federal

### Localização das empresas no DF
"""

# Coordenadas
x_coord = dados_empresas_final['geometry'].x
y_coord = dados_empresas_final['geometry'].y

# Renomear colunas
dados_empresas_final.rename(columns={'macrozona': 'MACROZONA',
'macroarea': 'ZONA'}, inplace=True)

# Adicionando colunas
dados_empresas_final.insert(loc=0, column='Xcoord', value=x_coord)
dados_empresas_final.insert(loc=1, column='Ycoord', value=y_coord)

# 5 RA com menor ocorrência de empresas
ultimas =
dados_ra_final[['RA', 'empresas_count']].sort_values(by='empresas_count')['R
A'].values[0:5]
ultimas

# 5 RA com maior ocorrência de empresas
primeiras =
dados_ra_final[['RA', 'empresas_count']].sort_values(by='empresas_count')['R
A'].values[30:]
primeiras

# Configurar figura e eixo

fig, ax = plt.subplots(1, figsize=(8, 6))

```

```

fig.subplots_adjust(right=0.8)

# Limites das RA
ras.boundary.plot(ax=ax, linewidth=1, alpha=0.2, color='grey')

# Maior e menor ocorrência de empresas
ras[ras['ra_nome'].isin(primeiras)].copy().plot(ax=ax, linewidth=1,
alpha=0.5, color='red', legend=True)
ras[ras['ra_nome'].isin(ultimas)].copy().plot(ax=ax, linewidth=1,
alpha=0.5, color='blue', legend=True)

# Empresas (pontos)
ax.scatter(x=dados_empresas_final['Xcoord'],
y=dados_empresas_final['Ycoord'], c='slategrey', s=10, alpha=0.5, zorder=1)

# Mapa base
ctx.add_basemap(
    ax,
    crs=dados_empresas_final.crs.to_string(),
    source=ctx.providers.CartoDB.Positron)

# Legendas
custom_points = [Line2D([0], [0], marker='o', linestyle='none',
markersize=5, color=color) for color in ['slategrey', 'red', 'blue']]

leg_points = ax.legend(custom_points, ['Empresas', 'Regiões com maior nº de
empresas', 'Regiões com menor nº de empresas'],
                        loc='center left',
                        bbox_to_anchor=(0.8, 0.5),
                        bbox_transform=fig.transFigure,
                        fontsize='8')

# Retirar eixos
ax.set_axis_off()

# Mostrar
plt.show()

"""### Centro médio e SDE"""

# Centro médio
mean_center = pp.centrography.mean_center(dados_empresas_final[['Xcoord',
'Ycoord']])

# Standard Deviational Ellipse (SDE)
major, minor, rotation =
pp.centrography.ellipse(dados_empresas_final[['Xcoord', 'Ycoord']])

# 1º elem = coordenadas (x,y) do centro
# 2º elem = os dois valores do semi-eixo (ao longo de x, ao longo de y)
# 3º elem = ângulo em graus entre o eixo x da base cartesiana e o semi-eixo
correspondente
ellipse = (mean_center, (major, minor), np.rad2deg(rotation))

# Criar um círculo de raio 1 em torno do ponto central:
circ = shapely.geometry.Point(ellipse[0]).buffer(1)

# Criar a elipse ao longo de x e y:

```

```

ell = shapely.affinity.scale(circ, int(ellipse[1][0]), int(ellipse[1][1]))

# Girar a elipse (sentido horário, eixo x apontando para a direita):
ellr = shapely.affinity.rotate(ell,ellipse[2])

type(ellr)

pontos = [] # pontos dentro da SDE
pontos_out = [] # pontos fora da SDE
for point in dados_empresas_final['geometry']:
    if point.within(ellr):
        pontos.append(point)
    else:
        pontos_out.append(point)

# Coordenadas dos pontos da SDE
x_coords = [list(point.coords)[0][0] for point in pontos]
y_coords = [list(point.coords)[0][1] for point in pontos]

# Coordenadas dos pontos fora da SDE
x_coords_out = [list(point.coords)[0][0] for point in pontos_out]
y_coords_out = [list(point.coords)[0][1] for point in pontos_out]

# Configurar figura e eixo
f, ax = plt.subplots(1, figsize=(12, 6))

ras.boundary.plot(ax=ax, linewidth=1, alpha=0.25, color='grey')

# Pontos
ax.scatter(x_coords, y_coords, c='darkblue', s=10, zorder=1, alpha=0.5,
label='Pontos dentro da SDE')
ax.scatter(x_coords_out, y_coords_out, c='slategrey', s=10, zorder=1,
alpha=0.5)

# Centro médio
ax.scatter(*mean_center, color='red', marker='x', label='Centro Médio')

# SDE
ax.plot(ellr.exterior.coords.xy[0], ellr.exterior.coords.xy[1],
color='red', linestyle='--', label='Std. Ellipse')

# Legenda
ax.legend()

# Adicionar mapa base
ctx.add_basemap(
    ax,
    crs=dados_empresas_final.crs.to_string(),
    source=ctx.providers.CartoDB.Positron)

# Mostrar
ax.set_axis_off()
plt.show()

round(sum(ras.area)/1000000, 2)

# Área da SDE

```

```

round(ellr.area/1000000, 2)

# Total de pontos na SDE
len(pontos)

# Porcentagem de pontos na SDE
round(len(pontos)/len(dados_empresas_final['geometry']) * 100, 2)

# Criar GeoDataFrame
pontos_gdf = gpd.GeoDataFrame({'geometry':pontos},
                              crs=ras.crs.to_string(),
                              geometry='geometry')

# Regiões dos pontos da SDE
regioes_pontos_in = gpd.overlay(pontos_gdf, ras, how='intersection')

[x.title() for x in list(regioes_pontos_in['ra_nome'].unique())]

regioes_pontos_sde =
ras[ras['ra_nome'].isin(list(regioes_pontos_in['ra_nome'].unique()))].copy(
)

# Configurar figura e eixo
f, ax = plt.subplots(1, figsize=(12, 6))

ras.boundary.plot(ax=ax, linewidth=1, alpha=0.25, color='grey')

regioes_pontos_sde.plot(ax=ax, alpha=0.5, color='lightblue')

ax.scatter(x_coords, y_coords, c='darkblue', s=10, zorder=1, alpha=0.5,
label='Pontos dentro da SDE')

# Centro médio
ax.scatter(*mean_center, color='red', marker='x', label='Centro Médio')

# SDE
ax.plot(ellr.exterior.coords.xy[0], ellr.exterior.coords.xy[1],
color='red', linestyle='--', label='Std. Ellipse')

# Legenda
ax.legend()

# Adicionar mapa base
ctx.add_basemap(
    ax,
    crs=dados_empresas_final.crs.to_string(),
    source=ctx.providers.CartoDB.Positron)

# Mostrar
ax.set_axis_off()
plt.show()

# Quantidade de pontos na SDE: frequência por RA
tab_freq =
regioes_pontos_in.groupby('ra_nome').size().sort_values(ascending=False)

list(tab_freq.keys())

```

```

tab_freq.values

tab_freq = pd.DataFrame({'RA':list(tab_freq.keys()),
                        'Frequência Absoluta (FA)':list(tab_freq.values)})

tab_freq

tab_freq['Frequência Absoluta (FA)'].sum()

"""## Distribuição espacial das instalações logísticas em diferentes
regiões administrativas

### Densidade de empresas por RA
"""

dados_ra_final.columns

# Empresas/km2

fig, ax = plt.subplots(1, figsize=(12, 6))

dados_ra_final.plot(ax=ax, column='empresas_por_km2', scheme='quantiles',
                    cmap='Reds', alpha=0.5, legend=True)

ctx.add_basemap(
    ax,
    crs=dados_ra_final.crs.to_string(),
    source=ctx.providers.CartoDB.Positron)

ax.set_axis_off()

plt.show()

"""### Autocorrelação"""

y = dados_ra_final['empresas_por_km2'].values # variável/atributo
w = Queen.from_dataframe(dados_ra_final) # Queen Weights (Pesos)
w.transform = 'r' #row standardization

# Y spatial lag
np.random.seed(12345)
ylag = lag_spatial(w, y)

# Spatial Lag Densidade de Empresas
fig, ax = plt.subplots(figsize=(12, 6))

dados_ra_final.assign(cl=ylag).plot(column='cl', scheme='quantiles', \
                                   k=5, cmap='Reds', linewidth=0.1, ax=ax, \
                                   edgecolor='white', alpha=0.5, legend=True)

ctx.add_basemap(ax, crs=dados_ra_final.crs.to_string(),
                source=ctx.providers.CartoDB.Positron)

ax.axis('off')

plt.show()

```

```

np.random.seed(12345)
mi = Moran(y, w)
round(mi.I, 2)

# Permutações
mi.permutations

# Diagrama de Dispersão de Moran
plot_moran(mi, zstandard=True, figsize=(10, 5))
plt.show()

# p-valor
mi.p_sim

mi.p_sim < 0.05

# Calcular Moran_Local (Local Moran's I)
np.random.seed(12345)
moran_loc = Moran_Local(y, w)

fig, ax = plt.subplots(1, figsize = (12, 6))

lisa_cluster(moran_loc, dados_ra_final, p=0.05, alpha=0.5, ax=ax)

ctx.add_basemap(ax, crs=dados_ra_final.crs.to_string(),
source=ctx.providers.CartoDB.Positron)

plt.show()

# Identificar clusters e regiões
sig = moran_loc.p_sim < 0.05
hotspot = sig * moran_loc.q==1
coldspot = sig * moran_loc.q==3
doughnut = sig * moran_loc.q==2
diamond = sig * moran_loc.q==4

# Quantidade de HH
sum(hotspot)

# Quantidade de LL
sum(coldspot)

# Regiões HH
spots = ['n.sig.', 'hot spot']
labels = [spots[i] for i in hotspot*1]

t = dados_ra_final.copy()
t['cl'] = labels
t['p_sim'] = moran_loc.p_sim

t.loc[(t['cl']=='hot spot')]['RA'].values

# Regiões LL
spots = ['n.sig.', 'cold spot']
labels = [spots[i] for i in coldspot*1]

```

```

t = dados_ra_final.copy()
t['cl'] = labels
t['p_sim'] = moran_loc.p_sim

t.loc[(t['cl']=='cold spot)]['RA'].values

"""## Disponibilidade de infraestrutura de transporte no Distrito
Federal"""

# Paleta de cores
palette = list(sns.color_palette('tab10').as_hex())

# Tirar cor 7
palette.pop(7)

fig, ax = plt.subplots(1, figsize=(12, 6))

fig.subplots_adjust(right=0.8)

ras.boundary.plot(ax=ax, linewidth=1, alpha=0.5, color='grey')

distrital_local.plot(ax=ax, linewidth=3, color=palette[0])
federal_local.plot(ax=ax, linewidth=1, color=palette[1])

# contexto
ctx.add_basemap(
    ax,
    crs=dados_empresas_final.crs.to_string(),
    source=ctx.providers.CartoDB.Positron)

custom_points = [Line2D([0], [0], marker='none', linestyle='-',
markersize=5, color=color) for color in palette[0:2]]

leg_points = ax.legend(custom_points, ['Distritais', 'Federais'],
                        loc='center left',
                        bbox_to_anchor=(0.8, 0.5),
                        bbox_transform=fig.transFigure,
                        fontsize='8')

ax.add_artist(leg_points)

# Eixos
ax.set_axis_off()
plt.show()

# Densidade rede viária

fig, ax = plt.subplots(1, figsize=(8, 6))

dados_ra_final.plot(ax=ax, column='densidade_rede_viaria',
scheme='quantiles', cmap='Reds', alpha=0.5, legend=True)

ctx.add_basemap(
    ax,
    crs=dados_ra_final.crs.to_string(),
    source=ctx.providers.CartoDB.Positron)

```

```

ax.set_axis_off()

plt.show()

"""## Relação entre a localização das instalações e infraestrutura de
transporte

### Buffering
"""

rodovias = gpd.read_file('.../data/rodovia.zip')

rodovias.crs

# Criar buffer de 1km
buffers_gdf = gpd.GeoDataFrame({'geometry':rodovias.buffer(1000).values},
                               geometry='geometry',
                               crs=ras.crs.to_string())

# União
u = buffers_gdf['geometry'].unary_union

# GeoDataFrame
gdf = gpd.GeoDataFrame({'id':[1, 2],
                       'geometry':u},
                       crs=ras.crs.to_string(),
                       geometry='geometry')

# Overlay
ol = gpd.overlay(dados_empresas_final, gdf, how='intersection')

# Remover duplicados
ol.drop_duplicates(subset=['CNPJ'], inplace=True)

ol.info()

point_in = [] # próximo
ra_in = []
point_out = [] # distante
ra_out = []
for i in range(0, len(dados_empresas_final['geometry'])):
    point = dados_empresas_final['geometry'][i]
    ra = dados_empresas_final['ra_nome'][i]
    if point.within(u):
        point_in.append(point)
        ra_in.append(ra)
    else:
        point_out.append(point)
        ra_out.append(ra)

pi = gpd.GeoDataFrame({'RA':ra_in, 'geometry':point_in},
                      crs=ras.crs.to_string(),
                      geometry='geometry')

# Pontos próximos a uma rodovia
len(point_in)

```

```

ol['ra_nome'].values

temp_df0 = dados_ra_final[['RA', 'empresas_count']].sort_values(by=['RA'])

temp_df = temp_df0[temp_df0['RA'].isin(ol['ra_nome'].values)].copy()

# Pontos próximos de rodovias
pontos_buffers = ol[['ra_nome', 'CNPJ']].groupby('ra_nome').count()

# Criar coluna com total na RA
v_total_ra = temp_df['empresas_count'].values
pontos_buffers.insert(loc=0, column='Total RA', value=v_total_ra)

# Resetar index
pontos_buffers.reset_index(inplace=True)

# Criar coluna Frequencia Relativa
pontos_buffers['Freq Rel'] = (pontos_buffers['CNPJ']/pontos_buffers['Total
RA'])*100

# Ordenar por percentual próximo as rodovias
pontos_buffers.sort_values(by='Freq Rel', ascending=False, inplace=True)

# Resetar index
pontos_buffers.reset_index(inplace=True, drop=True)

pontos_buffers

pontos_buffers.sort_values(by='CNPJ', ascending=False)['CNPJ'].values ==
pi.groupby('RA').size().sort_values(ascending=False).values

pontos_buffers['CNPJ'].sum()

len(pi['geometry'])

# % Pontos próximos a rodovias
round(100*pontos_buffers['CNPJ'].sum()/dados_ra_final['empresas_count'].sum
(), 2)

# Pontos +1km
pontos_out =
dados_empresas_final[~dados_empresas_final['CNPJ'].isin(ol['CNPJ'].values)]
.copy()

pontos_out.info()

# Configurar figura e eixo
fig, ax = plt.subplots(1, figsize=(12, 12))

# Limites das RA
plot_line(LinearRing(ras.unary_union.exterior), ax=ax, add_points=False,
color='slategrey', alpha=0.2)

# Buffers
plot_polygon(u, ax=ax, add_points=False, color='blue', alpha=0.2)

```

```

# Rodovias
rodovias.plot(ax=ax, color='blue')

# Pontos +1km
pontos_out.plot(ax=ax, color='k', markersize=5, alpha=0.25, zorder=1)

# Pontos 1km
ol.plot(ax=ax, color='forestgreen', markersize=5, alpha=0.25, zorder=1,
label='Empresas (1km)')

# Adicionar mapa base
ctx.add_basemap(ax=ax,
                crs=ras.crs.to_string(),
                source=ctx.providers.CartoDB.Positron)

# Legenda
ax.legend(loc='upper right')

# Retirar eixos
ax.set_axis_off()

# Mostrar
plt.show()

empresas_perto = gpd.GeoDataFrame({'id':list(range(0, len(point_in))),
                                  'geometry':point_in},
                                  crs=ras.crs.to_string(),
                                  geometry='geometry')

ol['geometry'].geom_equals_exact(empresas_perto['geometry'],
tolerance=0.05).sum()

empresas_longe = gpd.GeoDataFrame({'id':list(range(0, len(point_out))),
                                  'geometry':point_out},
                                  crs=ras.crs.to_string(),
                                  geometry='geometry')

fig, ax = plt.subplots(1, figsize=(12, 12))

plot_line(LinearRing(ras.unary_union.exterior), ax=ax, add_points=False,
color='slategrey', alpha=0.2)

plot_polygon(u, ax=ax, add_points=False, color='blue', alpha=0.2)

rodovias.plot(ax=ax, color='blue')

empresas_longe.plot(ax=ax, color='k', markersize=5, alpha=0.25, zorder=1)
empresas_perto.plot(ax=ax, color='forestgreen', markersize=5, alpha=0.25,
zorder=1, label='Empresas (1km)')

ctx.add_basemap(ax=ax,
                crs=ras.crs.to_string(),
                source=ctx.providers.CartoDB.Positron)

ax.legend(loc='upper right')
ax.set_axis_off()

```

```

plt.show()

"""### Correlação espacial

"""

y = dados_ra_final['empresas_por_km2'].values # variável/atributo
x = dados_ra_final['densidade_rede_viaria'] #variável/atributo
w = Queen.from_dataframe(dados_ra_final) # Queen Weights (Pesos)
w.transform = 'r' #row standardization

np.random.seed(12345)
moran = Moran(y, w) # Univariate Moran's I
moran_bv = Moran_BV(x, y, w) # Bivariate Moran's I
moran_loc = Moran_Local(y, w) # Univariate Local Moran's I
moran_loc_bv = Moran_Local_BV(x, y, w, seed=12345) # Bivariate Local
Moran's I

round(moran.I, 2)

plot_moran_bv(moran_bv) # Bivariate Moran's I
plt.show()

moran_bv.p_sim

moran_bv.p_sim < 0.05

fig, ax = plt.subplots(1, figsize = (12, 6))

lisa_cluster(moran_loc_bv, dados_ra_final, p=0.05, alpha=0.5, ax=ax)

ctx.add_basemap(ax, crs=dados_ra_final.crs.to_string(),
source=ctx.providers.CartoDB.Positron)

plt.show()

sig = moran_loc_bv.p_sim < 0.05
hotspot = sig * moran_loc_bv.q==1
coldspot = sig * moran_loc_bv.q==3
doughnut = sig * moran_loc_bv.q==2
diamond = sig * moran_loc_bv.q==4

spots = ['n.sig.', 'hot spot']
labels = [spots[i] for i in hotspot*1]

t = dados_ra_final.copy()
t['cl'] = labels

# Regiões HH
t.loc[(t['cl']=='hot spot')]['RA'].values

spots = ['n.sig.', 'cold spot']
labels = [spots[i] for i in coldspot*1]

t = dados_ra_final.copy()
t['cl'] = labels

```

```

# Regiões LL
t.loc[(t['cl']=='cold spot)][['RA']].values

spots = ['n.sig.', 'doughnut']
labels = [spots[i] for i in doughnut*1]

t = dados_ra_final.copy()
t['cl'] = labels

# Regiões LH
t.loc[(t['cl']=='doughnut)][['RA']].values

spots = ['n.sig.', 'diamond']
labels = [spots[i] for i in diamond*1]

t = dados_ra_final.copy()
t['cl'] = labels

# Regiões LH
t.loc[(t['cl']=='diamond)][['RA']].values

"""## Zoneamento do DF"""

# Renomear colunas
zoneamento.rename(columns={'macrozona':'MACROZONA', 'macroarea':'ZONA'},
inplace=True)

t =
dados_empresas_final.groupby('ZONA').size().sort_values(ascending=False)
siglas = ['ZUC', 'ZUUC II', 'ZUEQ', 'ZUCT', 'ZRUC', 'ZUUC I', 'MPI',
'ZRUD', 'ZCU']
zonas_df = pd.DataFrame({'NOME_ZONA':t.keys().values,
                        'SIGLA_ZONA':siglas,
                        'CONTAGEM':t.values,
                        'FREQUENCIA':t.values/(t.values.sum()),
                        'FREQUENCIA (pp)':(t.values/(t.values.sum())) *
100})
zonas_df

# Nomes das zonas
zonas = zonas_df['NOME_ZONA'].values
zonas_siglas = zonas_df['SIGLA_ZONA'].values

# Cores
color_dict = {}

for i in range(0, len(zonas)):
    color_dict[zonas[i]] = palette[i]

# Coluna cores
dados_empresas_final['colors'] =
dados_empresas_final['ZONA'].map(color_dict)
zoneamento['colors'] = zoneamento['ZONA'].map(color_dict)

fig, ax = plt.subplots(1, 2, figsize=(12, 6), tight_layout=True)

```

```

fig.subplots_adjust(bottom=0.2)

ras.boundary.plot(ax=ax[0], linewidth=1, alpha=0.5, color='grey')
ras.boundary.plot(ax=ax[1], linewidth=1, alpha=0.5, color='grey')

dados_empresas_final.plot(ax=ax[0], color=dados_empresas_final['colors'],
edgecolor='white')

zoneamento.plot(ax=ax[1], color=zoneamento['colors'], alpha=0.5)

# Adicionar mapa base
ctx.add_basemap(
    ax[0],
    crs=dados_empresas_final.crs.to_string(),
    source=ctx.providers.CartoDB.Positron)

ctx.add_basemap(
    ax[1],
    crs=dados_empresas_final.crs.to_string(),
    source=ctx.providers.CartoDB.Positron)

# Adicionar legenda
custom_points = [Line2D([0], [0], marker='o', linestyle='none',
markersize=5, color=color) for color in color_dict.values()]

fig.legend(custom_points, zonas_siglas,
           ncol=9,
           loc='upper center',
           bbox_to_anchor=(0.5, 0.16),
           bbox_transform=fig.transFigure)

# Eixos
ax[0].set_axis_off()
ax[1].set_axis_off()

plt.show()

```