



RIO

PUC

Dissertação de Mestrado

Previsão da Movimentação de Lojas em Shoppings Centers Concorrentes:

Um Estudo de Caso em uma Administradora
de Shoppings

Bruna da Silva Pereira

Pontifícia Universidade Católica do Rio de Janeiro
Centro Técnico Científico
Departamento de Engenharia Industrial

Rio de Janeiro, 04 de setembro de 2025



Pontifícia
Universidade
Católica do
Rio de Janeiro

Dissertação de Mestrado

Previsão da Movimentação de Lojas em Shoppings Centers Concorrentes:

Um Estudo de Caso em uma
Administradora de Shoppings

Bruna da Silva Pereira

Orientação: Professora Paula Medina Maçaira Louro

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre em Logística (Opção Profissional) pelo programa de Pós-Graduação em Engenharia de Produção, no Departamento de Engenharia Industrial

Rio de Janeiro, 04 de setembro de 2025



Pontifícia
Universidade
Católica do
Rio de Janeiro

Previsão da Movimentação de Lojas em Shoppings Centers Concorrentes:

Um Estudo de Caso em uma Administradora de
Shoppings

Bruna da Silva Pereira

**Dissertação apresentada como requisito parcial para a obtenção do
grau de Mestre em Logística (Opção Profissional). Aprovada pela
Comissão examinadora abaixo:**

Professora Paula Medina Maçaira Louro
Departamento de Engenharia Industrial

Professor Igor Tona Peres
Departamento de Engenharia Industrial

Professor Pedro Gomes Andrade
Instituto de Pesquisa Econômica Aplicada

Rio de Janeiro, 04 de setembro de 2025



Pontifícia
Universidade
Católica do
Rio de Janeiro

Todos os direitos reservados. A reprodução, total ou parcial, do trabalho é proibida sem autorização da universidade, da autora e do orientador.

Bruna da Silva Pereira

Graduou-se em Engenharia de Produção pela Pontifícia Universidade Católica do Rio de Janeiro em 2023.

Ficha Catalográfica

Pereira, Bruna da Silva

Previsão da movimentação de lojas em shoppings centers concorrentes : um estudo de caso em uma administradora de shoppings / Bruna da Silva Pereira ; orientação: Paula Medina Maçaira Louro. – 2025.

69 f. ; 30 cm

Dissertação (mestrado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Industrial, 2025.

Inclui bibliografia

1. Engenharia Industrial – Teses. 2. Descarga de mercadorias. 3. Logística. 4. Modelagem preditiva. 5. Otimização de processos. 6. Previsão de séries temporais. I. Louro, Paula Medina Maçaira. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Industrial. III. Título.

CDD: 658.5

Agradecimentos

Concluir esta etapa do mestrado é uma imensa alegria e uma vitória que não seria possível sem o apoio de pessoas incríveis.

Em primeiro lugar, um agradecimento de coração aos meus colegas de sala. A união e o espírito de colaboração que vivemos foram fundamentais. As trocas de conhecimento e o apoio mútuo em cada trabalho e tarefa tornaram a jornada muito mais leve e produtiva.

Ao meu pai, minha eterna gratidão. Sua dedicação em me proporcionar as condições para estudar e trilhar este caminho acadêmico é a base de tudo que conquistei. Este título também é fruto do esforço dele e investimento no meu futuro.

E, por fim, ao meu noivo, que foi o meu maior incentivador ao longo destes dois anos. Sem o apoio dele seria tudo mais difícil.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001

Resumo

Pereira, Bruna da Silva; Louro, Paula Medina Maçaira (Orientadora). **Previsão da Movimentação de Lojas em Shoppings Centers Concorrentes: Um Estudo de Caso em uma Administradora de Shoppings.** Rio de Janeiro, 2025. 69p. Dissertação de Mestrado, Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

Esta dissertação investiga a dinâmica do segmento logístico no contexto de shopping centers brasileiros, um elemento estratégico para a eficácia operacional no transporte e distribuição de produtos. O problema central da pesquisa questiona os principais fatores influenciadores na rotatividade de lojas em shopping centers com base em séries temporais históricas. A hipótese de estudo estabelece que a movimentação de lojas é influenciada por fatores temporais, como sazonalidade e tendência, e pelo *tenant mix*. O objetivo geral do estudo consistiu em examinar a concorrência em shopping centers brasileiros a partir de uma base de dados reais, fornecida por uma administradora, contendo informações sobre composição e movimentações de lojas. Utilizou-se a aplicação de técnicas de análise descritiva, diagnóstica e preditiva, com o auxílio da linguagem Python. A análise descritiva revelou a predominância dos pilares Varejo, Entretenimento e Lazer. A análise diagnóstica confirmou a sazonalidade, com picos de entradas em maio, outubro e dezembro, e picos de saídas em janeiro e julho, e indicou que o Sudeste é o principal polo de movimentação e de maior rotatividade. A modelagem preditiva obteve alta acurácia na previsão de entradas e saídas de lojas, sendo as variáveis temporais (mês e ano) e de *lag* as mais importantes. Conclui-se que os resultados fornecem insights valiosos para a tomada de decisão estratégica de administradoras de shopping centers, permitindo antecipar tendências e otimizar o mix de lojas em um cenário competitivo.

Palavras-chave

Descarga de mercadorias; Logística; Modelagem preditiva; Otimização de processos; Previsão de séries temporais.

Abstract

Pereira, Bruna da Silva; Louro, Paula Medina Maçaira (Orientadora). **Forecasting Store Traffic in Shopping Malls Considering Competitors: A Case Study in a Shopping Mall Management Company.** Rio de Janeiro, 2025. 69p. Dissertação de Mestrado, Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

This dissertation investigates the dynamics of the logistics segment within the context of Brazilian shopping centers, recognizing it as a strategic element for operational effectiveness in product transportation and distribution. The central research problem questions the main factors influencing store turnover in shopping centers based on historical time series data. The study hypothesis states that store movement is influenced by temporal factors, such as seasonality and trend, and by the tenant mix. The general objective of the study was to examine competition in Brazilian shopping centers using a real-world database, provided by an administrative company, containing information on store composition and movements of entry and exit. Descriptive, diagnostic, and predictive analysis techniques were applied, aided by the Python language. Descriptive analysis revealed the predominance of the Retail, Entertainment, and Leisure pillars. Diagnostic analysis confirmed seasonality, with entry peaks in May, October, and December, and exit peaks in January and July, and indicated that the Southeast is the main pole of movement and of highest turnover. Predictive modeling achieved high accuracy in forecasting store entries and exits, with temporal variables (month and year) and lag variables being the most important. It is concluded that the results provide valuable insights for the strategic decision-making of shopping center administrators, allowing them to anticipate trends and optimize the store mix in a competitive scenario.

Keywords

Logistics; Predictive modeling; Process optimization; Time series forecasting; Unloading of goods.

Sumário

1. Introdução	13
2. Revisão da Literatura	16
2.1. Shopping Center: Definição e Complexidade Operacional	16
2.2. O Tenant Mix e a Rotatividade de Lojas como Fator de Desempenho	17
2.3. Análise e Previsão de Séries Temporais: O Framework Metodológico	18
2.3.1. Técnicas de decomposição e pré-processamento	19
2.3.2. Tratamento de anomalias e variáveis externas	20
2.3.3. Modelagem preditiva e avaliação	21
2.3.4. Conclusão da revisão de séries temporais	21
3. Metodologia	23
3.1. Desenho da Pesquisa e Unidade de Análise	23
3.2. Materiais: Base de Dados e Detalhes da Coleta	23
3.3. Métodos: Framework de Análises e Modelagem	24
3.3.1. Detalhamento do framework metodológico	24
3.3.2. Pré-processamento e feature engineering	25
3.3.3. Justificativa e escolha dos modelos preditivos	26
3.3.4. Treinamento e avaliação dos modelos	27
4. Resultados	29
4.1. Caracterização da Empresa	29
4.2. Análise Descritiva	29
4.2.1. Técnicas de análise descritiva	30
4.2.1.1. Estatísticas descritivas	30
4.2.1.2. Análise de frequência	30
4.2.1.3. Visualização dos dados	33
4.3. Análise Diagnóstica	39
4.3.1. Relação entre movimento e ramo	39
4.3.2. Relação entre movimento e administradoras	42

4.3.3.	Sazonalidade das movimentações	43
4.3.4.	Impacto de fatores geográficos na rotatividade	45
4.3.4.1.	Análise por região do Brasil	46
4.3.4.2.	Análise por estados do Brasil	48
4.4.	Análise Preditiva e Interpretação dos Resultados	49
4.4.1.	Avaliação, métricas e hiperparâmetros dos modelos	50
4.4.2.	Interpretação dos resultados preditivos	55
5.	Conclusão	59
6.	Referências Bibliográficas	62

Lista de Figuras

Figura 1 - Número de Lojas Únicas por ano	33
Figura 2 - Série temporal de entradas e saídas de lojas	34
Figura 3 - Gráfico ABC de Administradoras	35
Figura 4 - Quantidade de movimento por competencia	36
Figura 5 - Proporção de pilar ao longo dos anos	36
Figura 6 - Percentual de movimentações por região	38
Figura 7 - Percentual de movimentações por estado	39
Figura 8 - Top 10 ramos por taxa de entrada	40
Figura 9 - Top 10 ramos por taxa de saída	41
Figura 10 - Top 10 administradoras por taxa de entrada	42
Figura 11 - Top 10 administradoras por taxa de saída	43
Figura 12 - Entradas por mês	44
Figura 13 - Saídas por mês	45
Figura 14 - Taxa de entrada de lojas por região	46
Figura 15 - Taxa de saída de lojas por região	47
Figura 16 - Taxa de entrada de lojas por estado	48
Figura 17 - Taxa de saída de lojas por estado	49
Figura 18 - Ajuste da Regressão Linear no treino para entradas	51
Figura 19 - Ajuste do Random Forest no treino para entradas	51
Figura 20 - Ajuste do Gradient Boosting no treino para entradas	52
Figura 21 - Previsão de entradas	53
Figura 22 - Ajuste da Regressão Linear no treino para saídas	53
Figura 23 - Ajuste do Random Forest no treino para saídas	54
Figura 24 - Ajuste do Gradient Boosting no treino para saídas	54
Figura 25 - Previsão de saídas	55
Figura 26 - Série temporal de entradas + previsão	57
Figura 27 - Série temporal de saídas + previsão	58

Lista de Tabelas

Tabela 1 - Medidas	30
Tabela 2 - Distribuição da variável pilar	31
Tabela 3 - Distribuição da variável ramo	31
Tabela 4 - Distribuição da variável tipo	32
Tabela 5 - Métricas dos modelos de previsão para entradas	51
Tabela 6 - Previsão para entradas	52
Tabela 7 - Métricas dos modelos de previsão para saídas	53
Tabela 8 - Previsão para saídas	54

Lista de Quadros

Quadro 1 – Detalhes da Coleta	23
Quadro 2 – Modelos de previsão	26

1. Introdução

Nos cenários metropolitanos contemporâneos, os shopping centers se consolidaram como componentes cruciais na expansão do comércio varejista, assumindo funções que transcendem a simples transação comercial. Estes complexos se distinguem pela oferta concentrada de múltiplos atrativos em um único ambiente, como diversidade de lojas, praça de alimentação, entretenimento e percepção de segurança, o que lhes confere uma vantagem competitiva frente aos formatos tradicionais de varejo (Silva; Leme; Santos, 2018; Parente; Barki, 2014). Caracterizado por um planejamento, implantação, propriedade e gestão unificada, shopping center é um polo que integra opções de comércio, gastronomia e convívio social, influenciando o cotidiano de uma significativa parcela da população e atuando como núcleo dinamizador da economia e das interações sociais (Sales; Reis, 2017).

A instalação de operações varejistas em shopping centers confere aos lojistas o benefício do fluxo contínuo de clientes e o apoio estratégico da administração, mas impõe desafios operacionais singulares. Devido às margens de lucro tipicamente estreitas e às restrições físicas características desses ambientes, a gestão logística se destaca como um elemento-chave para a viabilidade econômica e a competitividade dos negócios (Benjamin; Oliveira, 2020; Silva, 2012). A logística em shopping center abrange o planejamento e execução de processos como abastecimento, recebimento, armazenagem e reposição de produtos, demandando técnicas eficientes para garantir o fluxo ininterrupto de mercadorias. O controle rigoroso desses processos não só auxilia na redução de custos, como também impacta diretamente a satisfação do cliente e o desempenho do negócio (Cabral Filho, 2023).

Para atender às expectativas diversificadas dos clientes e assegurar a atratividade do complexo, os SCs exigem um tenant mix estrategicamente estruturado. A composição de lojas, que inclui grandes varejistas (âncoras) e unidades menores (satélites), é um fator crucial no potencial de atração de clientes (Morigi, 2018). Dessa forma, a administração do shopping realiza avaliações periódicas para mensurar indicadores econômicos e operacionais, visando a sinergia entre varejistas e a direção do empreendimento (Almeida Júnior, 2020).

Neste contexto, a rotatividade de lojas (abertura e fechamento) surge como um indicador setorial relevante, pois afeta diretamente a atratividade, o equilíbrio do tenant mix e a estabilidade do ecossistema comercial. Uma alta vacância ou rotatividade descontrolada pode comprometer a perenidade e a competitividade do varejo no shopping (Santos, 2013).

Diante da complexidade gerencial e da relevância da rotatividade, a análise da movimentação de lojas em shopping centers se justifica como tema essencial para aprimorar a gestão do tenant mix e reduzir a vacância. A presente investigação propõe a utilização de previsão de séries temporais nesse contexto, uma ferramenta pertinente para identificar padrões recorrentes, tendências de longo prazo e variações que influenciam o desempenho varejista. A capacidade de prever essas variações é crucial para responder às transformações do mercado e fundamenta a elaboração de estratégias operacionais mais eficientes, alinhadas com a logística de abastecimento e a retenção de lojistas estratégicos. Ao utilizar dados reais fornecidos por administradoras de shoppings e ferramentas tecnológicas avançadas, esta pesquisa visa fornecer subsídios significativos para otimizar os processos gerenciais, promovendo um planejamento mais preciso da distribuição de lojas.

A partir deste contexto, o problema de pesquisa é apresentado a seguir: Quais são os principais fatores influenciadores na movimentação de lojas em shopping centers com base em séries temporais históricas? A hipótese de pesquisa estabelece que a movimentação de lojas é significativamente influenciada por fatores temporais (sazonalidade e tendência de longo prazo) e pelo *tenant mix*, sendo que a aplicação de modelos de séries temporais baseados em dados históricos pode prever com acurácia a rotatividade futura, fornecendo uma ferramenta estratégica para a gestão do ecossistema logístico e comercial dos empreendimentos. O objetivo geral é examinar a concorrência em shopping centers brasileiros a partir de uma base de dados com informações sobre composição de lojas nos shopping centers, incluindo as movimentações de entrada e saída, além dos segmentos de atuação e localização dos shoppings. Por meio da aplicação de técnicas de análise descritiva, diagnóstica e preditiva, com o auxílio da linguagem Python e suas bibliotecas, serão investigados padrões e tendências que afetam o cenário competitivo desse mercado.

Nesta pesquisa, definiram-se dois objetivos específicos: (i) verificar a relevância do mix de lojas para a criação de um ecossistema logístico eficiente em shopping centers; (ii) coletar dados históricos de entrada e saída de lojas em shoppings, e identificar padrões de movimentação de lojas nesse contexto, com o apoio de um estudo de caso.

Para alcançar os objetivos propostos, este trabalho foi estruturado da seguinte forma: o capítulo 2 apresenta a revisão da literatura; o capítulo 3 expõe a metodologia utilizada; o capítulo 4 desenvolve o estudo de caso e o capítulo 5 encerra com discussões e conclusões.

2. Revisão da Literatura

O presente capítulo estabelece a fundamentação teórica que suporta a análise da rotatividade de lojas em shopping centers brasileiros utilizando séries temporais. Para tal, a Seção 2.1 define o shopping center como um complexo operacional e logístico, e discute os desafios de gestão do tenant mix. A Seção 2.2 aprofunda o papel da rotatividade (entrada e saída de lojistas) como reflexo direto do desempenho e da satisfação do consumidor, justificando a necessidade de previsibilidade. Por fim, a Seção 2.3 apresenta o arcabouço teórico e metodológico da Previsão de Séries Temporais, detalhando o framework de análise, pré-processamento e modelagem preditiva que será aplicado nesta pesquisa.

2.1. Shopping Center: Definição e Complexidade Operacional

O shopping center é um centro comercial planejado, caracterizado pela arquitetura específica, pela reunião de estabelecimentos de comércio e serviços, e por uma administração única e centralizada. Sua distinção reside na oferta concentrada de múltiplos atrativos e serviços, como lojas, restaurantes, cinemas, consultórios médicos e academias, que atendem a necessidades de consumo, lazer e promoção de saúde. O objetivo central é a conveniência integrada para os usuários (Pereira, 2011). Caracteristicamente, o empreendimento opera como um modelo de negócio imobiliário (Williams, 2009), onde o investidor busca lucratividade através da locação, sujeitando os locatários a normas contratuais padronizadas para manter o equilíbrio da oferta e a funcionalidade do conjunto (Teller et al., 2008).

A natureza multifuncional do shopping center exige uma operação logística e gerencial complexa. Seu planejamento, que envolve a definição estratégica do mix de lojas e a otimização de fluxos e acessos, visa garantir a viabilidade financeira dos lojistas, cujo desempenho está intrinsecamente vinculado ao faturamento do complexo. Por atuarem como equipamentos autossuficientes, os shopping centers são agentes com grande capacidade de produzir vantagens locacionais, influenciando o tecido urbano e gerando sua própria zona de influência comercial (Maraschin; Avozani, 2020).

A expressão *Tenant Mix* refere-se à combinação planejada e estratégica de lojistas e operações comerciais dentro de um shopping center. Essa seleção visa

maximizar a sinergia entre os estabelecimentos, atrair um público diversificado e garantir o equilíbrio da oferta, sendo um elemento crucial para a performance do empreendimento.

Os aspectos que impulsionam o crescimento do shopping estão, portanto, ligados à adaptação contínua às expectativas do público e à manutenção de uma infraestrutura que suporte a complexidade logística, desde vagas de estacionamento (Schulz, 2014) até o eficiente sistema de abastecimento e gestão de mercadorias dos lojistas.

2.2. O Tenant Mix e a Rotatividade de Lojas como Fator de Desempenho

O diferencial essencial do shopping center reside no minucioso planejamento estratégico, que inclui a seleção cuidadosa do tenant mix — o arranjo de lojistas e a distribuição espacial dos estabelecimentos — para criar sinergias econômicas benéficas. O monitoramento contínuo do mercado e das expectativas dos clientes é fundamental, o que justifica a constante renovação e substituição de lojas observada nesses complexos (Almeida Júnior, 2020).

O desempenho do shopping center depende diretamente da capacidade de sua gestão em atender às demandas do consumidor, que busca um espaço de convivência que transcende a mera praticidade operacional (Garg; Steyn, 2015). Fatores como a percepção de segurança (Parente et al., 2012), a ambientação (Hedhli et al., 2013) e o apelo comercial (Teller e Elms, 2012) exercem influência direta na decisão do consumidor de frequentar o local.

Análise da Rotatividade e Satisfação: A satisfação do cliente e sua predisposição à compra (Munaro; Martins; Kato, 2019) encontram-se intimamente associadas a fatores como a diversidade de estabelecimentos (tenant mix), a assistência disponibilizada e a organização (Hastreiter et al., 2000). Essa satisfação, por sua vez, está sujeita à teoria da discordância, onde a performance real da loja é comparada às expectativas do consumidor (Oliver; Swan, 1989). Como Mückenberger (2000) ressalta, o impacto da satisfação sobre a intenção de compra decai com o tempo, o que impõe à administração do shopping center a necessidade de monitoramento contínuo e ajustes estratégicos.

Nesse contexto, a rotatividade de lojistas (entrada e saída) não é apenas um indicador financeiro, mas o reflexo prático do sucesso ou fracasso da administração do tenant mix e da experiência de consumo oferecida. Como a intenção de compra, isoladamente, não garante a concretização da transação ou a permanência do lojista (Steffen et al., 2014), a alta rotatividade sinaliza uma disfunção no ecossistema logístico-comercial do SC. Por essa razão, a capacidade de prever a movimentação de lojas torna-se um imperativo estratégico para garantir a estabilidade do empreendimento e a perenidade do mix ideal.

2.3. Análise e Previsão de Séries Temporais: O Framework Metodológico

A previsão de cenários sempre ocupou posição central nos processos decisórios e de planejamento estratégico. A imprevisibilidade inerente ao porvir apresenta-se simultaneamente como motivadora e complexa, levando tanto indivíduos quanto corporações a buscarem minimizar ameaças e potencializar ganhos. A ampla gama de situações que demandam projeções futuras exige um repertório variado de técnicas preditivas capazes de responder aos desafios do ambiente contemporâneo. O domínio da predição registrou progressos substanciais nos âmbitos teórico e aplicado desde o marco representado pelo estudo de De Gooijer e Hyndman (2006). O aprimoramento tecnológico na área computacional viabilizou o processamento de conjuntos de dados mais extensos e complexos, fator que estimulou o crescimento do interesse pela análise de informações e pelo campo da ciência de dados (Hyndman; Athanasopoulos, 2021).

A ciência preditiva baseia-se no princípio de que informações passadas e presentes permitem antecipar eventos futuros. No âmbito das séries temporais, parte-se do pressuposto que padrões identificáveis em dados históricos podem ser projetados para estimar valores posteriores. Contudo, o propósito essencial não reside na obtenção de previsões exatas, mas sim na geração de diferentes representações do futuro, como valores médios esperados (previsões pontuais), faixas de previsão, percentis ou distribuições completas de probabilidade (Fiorucci; Pellegrini; Louzada; Petropoulos; Koehler, 2016).

Os registros ou informações coletadas em séries temporais apresentam uma diversidade de formatos que podem influenciar ou restringir a seleção de uma

técnica de previsão. A periodicidade das medições pode abranger diferentes intervalos, como minuto a minuto, hora a hora, semanal, mensal ou anual. Os conjuntos de dados podem variar desde uma única série temporal relevante até milhões delas, e as séries temporais frequentemente possuem relevância em múltiplos níveis, formando uma estrutura hierárquica. Em algumas situações, parte ou a totalidade dos valores pode ser nula, caracterizando uma série temporal esparsa. A variedade de formatos para os dados é praticamente ilimitada.

2.3.1. Técnicas de decomposição e pré-processamento

Antes da implementação de uma técnica de previsão, é crucial realizar um tratamento preliminar dos dados, o que inclui a confirmação da exatidão das informações, a identificação de lacunas nos registros e o ajuste sazonal. A análise decompositiva constitui elemento central em múltiplas metodologias preditivas, possibilitando expressar uma série temporal como combinação funcional de elementos constituintes. As estruturas aditiva e multiplicativa emergem como padrões predominantes, operacionalizadas respectivamente por relações aritméticas de soma e produto (Granger & Newbold, 1976).

A estrutura decompositiva mais básica para séries temporais com padrão sazonal único compreende três elementos fundamentais: tendência, sazonalidade e termo residual. O fundamento essencial reside na aplicação de médias móveis para atenuar oscilações sazonais e tendenciais (Macaulay, 1931). A evolução da decomposição tradicional levou ao surgimento de metodologias mais avançadas, como X-11, X-12-ARIMA e o método STL (Cleveland, Cleveland, McRae e Terpenning, 1990). O método STL, em particular, emprega um procedimento cíclico com suavização para aprimorar progressivamente as estimativas dos elementos de tendência e sazonalidade. Uma segunda categoria de métodos, como os modelos BATS e TBATS (De Livera, Hyndman, & Snyder, 2011), utiliza uma estrutura estatística unificada para realizar a decomposição, possibilitando a obtenção direta de intervalos de confiança e projeções.

Uma técnica frequentemente adotada em modelagem preditiva envolve a modificação da variável-alvo y , empregando a transformação de Box e Cox (1964). A finalidade primordial da transformação de Box-Cox consistia em assegurar que os dados atendessem aos pressupostos de normalidade e homocedasticidade. A

discussão metodológica sobre a seleção do parâmetro λ possui ampla tradição acadêmica, com Guerrero (1993) propondo uma abordagem amplamente adotada baseada na segmentação dos dados em subconjuntos para otimizar λ . As estimativas projetadas, no entanto, podem exibir distorções sistemáticas ao serem retro transformadas para a escala original (Granger e Newbold, 1976), exigindo procedimentos específicos de ajuste.

2.3.2. Tratamento de anomalias e variáveis externas

Informações temporais frequentemente enfrentam interferências inesperadas, gerando diferentes categorias de registros atípicos (outliers, picos, saltos). A omissão no tratamento dessas ocorrências pode levar a estimativas distorcidas dos coeficientes modelares e comprometer significativamente a acurácia das projeções (Bianco, García Ben, Martínez, & Yohai, 2001). A literatura especializada distingue entre anomalias aditivas (AO) e de inovação (IO) (Fox, 1972).

As técnicas para identificação de irregularidades podem ser classificadas em métodos fundamentados em modelagem (que comparam valores projetados e observados) e métodos baseados em atributos (que se fundamentam nas propriedades intrínsecas das séries temporais). A detecção de valores atípicos e pontos influentes exerce função crucial na otimização da capacidade preditiva dos modelos, sendo que a estimação robusta de parâmetros se configura como outra abordagem viável para elevar a precisão das previsões, dispensando ajustes ou exclusão de outliers (Sakata e White, 1998; Gelper, Fried e Croux, 2009).

Variáveis externas correspondem a elementos integrados em um sistema de projeção por seu potencial explicativo. Múltiplas abordagens de análise temporal dispõem de formulações ampliadas capazes de incorporar preditores externos, como os modelos autorregressivos com entrada exógena (ARX). Esses preditores podem apresentar natureza quantitativa ou qualitativa. Em situações com grande dimensionalidade, técnicas de compressão de características, como a Análise de Componentes Principais (Stock & Watson, 2002), podem ser empregadas para simplificar a estrutura do modelo preditivo, convertendo dados multidimensionais em representações de menor complexidade. O processamento de dados para fins preditivos tem adquirido importância crescente, acompanhando a expansão

continua do volume de informações disponíveis (Hastie; Tibshirani; Friedman, 2009).

2.3.3. Modelagem preditiva e avaliação

Os sistemas de previsão estruturam-se através de métodos e modelos específicos. Abordagens consolidadas como ARIMA e suavização exponencial preservam sua relevância, mas o aprimoramento tecnológico estimulou o uso de arquiteturas neurais e algoritmos de aprendizado automatizado (Hyndman; Athanasopoulos, 2021). A integração ou consolidação de previsões (forecast combination), embora constitua um conceito antigo, tem demonstrado eficácia comprovada, sendo que a fusão de modelos subjacentes gera, por si só, um novo modelo preditivo.

A natureza das variáveis envolvidas determina estratégias distintas de modelagem. Na previsão univariada, trabalha-se exclusivamente com a série histórica da variável-alvo, enquanto abordagens multivariadas incorporam outras séries temporais como preditores, analogamente aos modelos de regressão temporal. No contexto do varejo, a modelagem preditiva aplicada a séries temporais revela-se indispensável para otimizar o tenant mix em shopping centers, reforçando a importância da atitude estratégica na administração moderna.

A avaliação da eficácia de um modelo preditivo constitui uma etapa fundamental no processo de análise. A utilização das previsões em cenários reais fornece parâmetros para quantificar seu nível de precisão. Habitualmente, o foco principal reside na diferença entre o dado real e a estimativa pontual gerada, utilizando métricas de erro padronizadas. A qualidade de uma previsão pontual pode ser examinada mediante a utilização de intervalos preditivos e percentis, que oferecem uma perspectiva mais ampla sobre a confiabilidade da estimativa.

2.3.4. Conclusão da revisão de séries temporais

A ciência preditiva, amparada por dados cada vez mais granulares e o avanço contínuo da capacidade computacional, permanece em constante evolução. O desafio de capturar padrões complexos, como os que envolvem a rotatividade do varejo em shopping centers, exige a aplicação de um framework robusto de análise

e pré-processamento, aliado a modelos de alta capacidade preditiva, como os utilizados neste trabalho. O futuro da previsão aponta para a importância da interpretabilidade dos modelos (necessária para a gestão estratégica) e para o uso crescente de modelos pré-treinados e de grande escala, como os desenvolvidos por plataformas como Google e Amazon, indicando a fusão inevitável entre métodos estatísticos e técnicas avançadas de Machine Learning para a otimização de decisões em tempo real.

3. Metodologia

O presente trabalho adota a abordagem de estudo de caso, conforme delineado por Yin (2011), para investigar a dinâmica da concorrência no setor de shopping centers brasileiros. Tal método permite uma análise aprofundada de um fenômeno específico, valendo-se de múltiplas fontes de informação para alcançar uma compreensão detalhada. O processo investigativo é cíclico, partindo de um referencial teórico que embasa a formulação das questões e hipóteses, e culminando na aplicação de um framework de análise de dados.

3.1. Desenho da Pesquisa e Unidade de Análise

O fenômeno central a ser investigado é a dinâmica da concorrência no setor de shopping centers, com foco na movimentação de lojas (entradas e saídas). A unidade de análise consiste no conjunto de shoppings concorrentes para os quais a base de dados oferece informações detalhadas.

Em consonância com a perspectiva de Yin (2011), esta pesquisa se fundamenta em um referencial teórico que engloba a concorrência em mercados varejistas, a dinâmica do setor de shopping centers e as metodologias de previsão de séries temporais. As técnicas primárias de análise incluem a exploração estatística descritiva, a análise diagnóstica para identificar relações entre variáveis e, por fim, a modelagem de séries temporais para a previsão de tendências.

3.2. Materiais: Base de Dados e Detalhes da Coleta

A principal fonte de evidência para este estudo de caso é uma base de dados privada, fornecida por uma empresa administradora de shopping centers. Esta base detalha a movimentação histórica de lojistas nos shoppings concorrentes e do grupo de controle, detalhado no Quadro 1 abaixo.

Quadro 1 - Detalhes da coleta

Detalhe crucial	Especificação
Natureza da base	Dados privados e confidenciais, fornecidos sob acordo de sigilo por uma administradora líder do setor.

Janela temporal	Março de 2022 a abril de 2025, com exceção dos meses de março e abril de 2024, que estão ausentes.
Unidades envolvidas	Registros de movimentação de 14.543 lojas, distribuídas em 118 shopping centers de 25 administradoras diferentes.
Principais variáveis	Competência (ano/mês), Identificação do Shopping, Tipo de Movimento (Entrada, Base ou Saída), Ramo de Atividade, Categoria da Loja, Pilar de Segmento e Localização Geográfica (Cidade, Estado, Região).

A granularidade destas informações possibilita uma análise longitudinal e segmentada da dinâmica de entrada e saída de lojas. Para garantir a transparência da pesquisa, o processo de coleta e tratamento dos dados foi documentado de forma detalhada (cadeia de evidências).

3.3. Métodos: Framework de Análises e Modelagem

O framework metodológico desta pesquisa está estruturado em três fases principais: Análise Descritiva, Análise Diagnóstica e Análise Preditiva.

3.3.1. Detalhamento do framework metodológico

A análise seguiu um fluxo de trabalho em cinco etapas, conduzidas com o auxílio da linguagem Python e suas bibliotecas especializadas:

1. Pré-processamento e transformação de dados: Preparação e higienização da série temporal.
2. Criação de *features*: Formalização de variáveis preditivas.
3. Análise descritiva e diagnóstica: Visualização e investigação de padrões e correlações.
4. Modelagem preditiva: Treinamento e ajuste dos modelos de previsão.
5. Avaliação e projeção: Mensuração de desempenho e projeção para o horizonte futuro.

3.3.2. Pré-processamento e feature engineering

A fase de preparação dos dados para a modelagem preditiva contribui para a acurácia das previsões, conforme demonstram Petropoulos et al. (2022). O processo foi conduzido utilizando a linguagem Python e a biblioteca Pandas, abrangendo a higienização da base e a construção de variáveis preditivas (*feature engineering*).

Inicialmente, a higienização da base de dados histórica focou na garantia da qualidade e consistência dos registros. Iniciou-se com o tratamento de valores nulos, onde foram identificadas 13.551 ocorrências na coluna pilares_segmentos, além de 765 linhas nulas nas colunas ramo e tipo. Todos estes valores ausentes foram substituídos pela categoria “N/I” (Não Identificado), que se refere a lojas sem cadastro completo no sistema da empresa administradora.

Em relação à duplicidade, foram encontradas 202 linhas repetidas, sendo aplicado o critério de descarte para aquelas que não possuíam o movimento diferente de "base", mantendo-se apenas a primeira ocorrência válida. Por fim, realizou-se a correção de inconsistências textuais, unificando, por exemplo, o valor “NI” para “N/I” e padronizando “JÓIAS, RELÓGIOS E BIJOTERIAS” para “JÓIAS, RELÓGIOS E BIJOUTERIAS” na coluna ramo, e o tratamento de “OUTRS” para “OUTROS” na coluna tipo.

Para adaptar a base histórica ao formato que otimiza o desempenho dos algoritmos de *Machine Learning*, foi realizada a engenharia de *features*. A previsão de interesse concentra-se nas variáveis *target* num_entradas e num_saidas, tratando o problema como uma regressão de contagens (Petropoulos et al., 2022). A construção das *features* preditivas incluiu variáveis temporais, como o mês e o ano, que foram extraídas e transformadas via *One-Hot Encoding*. Essa transformação é necessária para permitir que o modelo capture tendências de longo prazo (ano) e a sazonalidade discreta (mês) da série.

Em linha com Petropoulos et al. (2022), a inclusão de variáveis autorregressivas (lags) é necessária para capturar a dependência temporal da série. Foram formalmente construídas *features* de lag1, que incluem as contagens de entradas (num_entradas_lag1) e saídas (num_saidas_lag1) do mês imediatamente anterior, bem como o número de lojas ativas (num_base_lojas_ativas_lag1) no

período precedente. Valores ausentes resultantes da criação dessas features foram preenchidos com zero, pressupondo a ausência de movimento anterior onde os dados não estavam disponíveis.

3.3.3. Justificativa e escolha dos modelos preditivos

A seleção de modelos para problemas de previsão de quantidades com séries temporais, como a contagem de entradas e saídas, é uma etapa importante para a subsequência do trabalho. Petropoulos et al. (2022) oferecem uma visão sobre as abordagens de previsão, destacando a crescente relevância de métodos baseados em *Machine Learning* para cenários com complexidade e relações não-lineares, os quais podem superar modelos estatísticos tradicionais.

Considerando a natureza do problema e as características dos dados, serão explorados diferentes tipos de modelos que se adequam à tarefa de previsão de contagens, presente no Quadro 2. Essa diversidade de abordagens é importante, pois não existe um único algoritmo universalmente superior para todos os conjuntos de dados e problemas (Petropoulos et al., 2022).

Quadro 2 - Modelos de previsão

Modelos	Explicação
Modelos Lineares	A Regressão Linear será utilizada como um benchmark inicial. Sua simplicidade permite estabelecer uma linha de base de desempenho para comparações. No entanto, Petropoulos et al. (2022) indicam que modelos lineares podem ter desempenho limitado em problemas com relações não-lineares ou interações complexas.
Modelos Baseados em Árvores (<i>Tree-based Models</i>)	Algoritmos como <i>Random Forest Regressor</i> e <i>Gradient Boosting Regressor</i> , são usados para capturar relações mais intrincadas e a alta dimensionalidade dos dados. Esses métodos são reconhecidos por sua robustez, por não exigirem escalonamento de features e por lidarem eficientemente com interações não-lineares nos dados.

Modelos de Contagem (Regressão de Poisson/Negativa Binomial)	Embora não sejam explicitamente implementados aqui via Scikit-learn devido à sua maior complexidade de ajuste neste contexto (comparado aos modelos baseados em árvore), é importante reconhecer que modelos de regressão de Poisson ou Negativa Binomial são teoricamente mais adequados para variáveis target que representam contagens. Contudo, a flexibilidade e a comprovada performance dos modelos baseados em árvores na prática os tornam escolhas eficazes mesmo sem a suposição explícita de uma distribuição de contagem.
--	--

Dessa forma, a análise preditiva inicial se concentrará na avaliação comparativa da Regressão Linear, do *Random Forest Regressor* e do *Gradient Boosting Regressor*. Essa seleção abrange um espectro de complexidade e robustez para a tarefa de previsão de contagens, permitindo identificar o modelo mais adequado para as características específicas dos dados de movimentação de lojas.

3.3.4. Treinamento e avaliação dos modelos

A avaliação da capacidade preditiva dos modelos se concentra em sua habilidade de generalizar para dados futuros, utilizando uma abordagem de treinamento e teste específica para séries temporais. A base de dados foi dividida em conjuntos de treinamento e teste de forma cronológica, seguindo o princípio de que a previsão em séries temporais deve simular um cenário real (*hold-out* cronológico), prevenindo o *data leakage* (Petroopoulos et al., 2022).

O conjunto de treinamento compreendeu todo o período histórico disponível, de março de 2022 a abril de 2025, utilizado para calibrar os parâmetros do modelo. O período subsequente, de maio a dezembro de 2025, constitui o horizonte de projeção final da pesquisa, onde as previsões são geradas na ausência de dados reais.

Para problemas de regressão, a avaliação foi focada na acurácia da previsão dos valores numéricos. As métricas selecionadas refletem a importância de mensurar

tanto a magnitude média dos erros quanto a proporção da variância explicada. Foram utilizadas as seguintes métricas de desempenho:

- Erro Quadrático Médio (RMSE): Medida da magnitude média dos erros. É a raiz quadrada da média dos erros quadráticos.
- Erro Médio Absoluto (MAE): Medida da magnitude média dos erros, sem considerar sua direção. É menos sensível a *outliers* que o RMSE.
- R-quadrado: Indica a proporção da variância na variável dependente que é previsível a partir das variáveis independentes. Um R-quadrado mais próximo de 1 indica um modelo que explica mais a variabilidade dos dados.

Por fim, em respeito aos acordos de confidencialidade firmados com a empresa administradora de shopping centers, os dados brutos utilizados e os códigos de pré-processamento não serão disponibilizados publicamente. A metodologia completa e os scripts utilizados na modelagem, no entanto, podem ser fornecidos mediante solicitação e aprovação da empresa, resguardando a segurança e privacidade das informações.

4. Resultados

Este capítulo apresenta o estudo de caso que visa analisar e prever a movimentação de lojas em shopping centers concorrentes à empresa de estudo.

4.1. Caracterização da Empresa

A empresa em foco neste estudo de caso é uma das principais administradoras de shopping centers no Brasil, com uma trajetória de mais de cinco décadas no mercado. Pioneira no setor, sua história remonta ao início da década de 1970, marcando sua entrada como uma das responsáveis pelo desenvolvimento de um dos primeiros shopping centers do país. Ao longo de sua existência, a empresa expandiu significativamente sua atuação, estabelecendo-se como uma das maiores do setor em termos de área bruta locável administrada e número de empreendimentos em seu portfólio.

Com presença nas cinco regiões do Brasil, a empresa demonstra um alcance nacional e uma compreensão das particularidades de cada mercado regional. Seu portfólio diversificado abrange uma variedade de shoppings, desde empreendimentos consolidados em grandes centros urbanos até shoppings em cidades de médio porte, totalizando uma área bruta locável que se aproxima a um milhão de metros quadrados. Essa abrangência permite uma análise rica e multifacetada da dinâmica concorrencial em diferentes contextos geográficos e econômicos.

A empresa se destaca não apenas pela sua dimensão, mas também pela sua experiência e conhecimento do mercado de shopping centers brasileiro. Ao longo dos anos, estabeleceu parcerias estratégicas e demonstrou capacidade de adaptação às mudanças do setor, buscando inovação e aprimoramento na gestão de seus empreendimentos. Sua atuação engloba desde a concepção e o desenvolvimento de novos shoppings até a administração e a comercialização de espaços nos empreendimentos existentes.

4.2. Análise Descritiva

Esta seção dedica-se à descrição da base de dados trabalhada neste trabalho, incluindo suas características essenciais. Aborda os tipos de dados disponíveis, o

período temporal abrangido pela análise, a cardinalidade de cada variável e as técnicas estatísticas empregadas.

4.2.1. Técnicas de análise descritiva

Para analisar e descrever os dados, foram utilizadas técnicas como estatísticas descritivas, análise de frequência e visualização dos dados.

4.2.1.1. Estatísticas descritivas

Para caracterizar a coluna “movimento”, utilizou-se a média como medida de tendência central, complementada pela identificação dos valores mínimo e máximo das movimentações, com o objetivo de compreender seu comportamento em relação à “competência”, que é a coluna de período utilizada na análise. As medidas descritas estão representadas na Tabela 1.

Tabela 1 - Medidas

Métrica	Quantidade
Média de entradas	304
Média de saídas	307
Mínimo de entradas	71
Mínimo de saídas	110
Máximo de entradas	604
Máximo de saídas	551

De acordo com a Tabela 1, a média de saídas por competência é ligeiramente superior à de entradas (307 vs 304). O mínimo de entradas de lojas na base completa é 71, e ocorreu na competência 202206. O mínimo de saídas é de 110, ocorrendo na competência 202204. Já os valores máximos são, 604 para entradas, e 551 para saídas, ocorrendo nas competências 202405 e 202308, respectivamente.

4.2.1.2. Análise de frequência

Dando sequência à análise, esta seção dedica-se à análise de frequência das variáveis categóricas das lojas, tendo como foco o último período disponível na base de dados: abril de 2025. O objetivo principal é identificar as categorias com

maior e menor ocorrência nesse retrato mais recente da base. Para entender a composição atual da base em termos das variáveis categóricas, serão considerados apenas os registros com “movimento” como “base” e “entrada”, excluindo-se as “saídas”.

As informações sobre a natureza das lojas são detalhadas através das colunas “pilares_segmentos”, “ramo” e “tipo”. A coluna “ramo” especifica um segmento mais granular dentro da classificação geral fornecida por “pilares_segmentos”. Por sua vez, a coluna “tipo” oferece uma categorização ainda mais específica, inserida no contexto de um determinado “ramo”. Essa organização reflete uma hierarquia onde “pilares_segmentos” abrange “ramo”, que por sua vez contém “tipo”, como no exemplo: Varejo -> Vestuário -> Vestuário Feminino.

Tabela 2 - Distribuição da variável pilar

Pilar	Distribuição
VAREJO	46,41%
ENTRETENIMENTO E LAZER	22,68%
SERVIÇO	12,08%
BEM-ESTAR	9,63%
N/I	9,2%

De acordo com a Tabela 2, a variável pilar, no último retrato da base, revela uma concentração no pilar VAREJO, que representa a maior parcela, com 46,41% das ocorrências. Em seguida, destaca-se o pilar ENTRETENIMENTO E LAZER, com 22,68%, indicando sua relevância na composição da base. O pilar SERVIÇO também possui uma participação considerável, correspondendo a 12,08%. Complementam a distribuição os pilares BEM-ESTAR, com 9,63%, e a categoria N/I, com 9,2%, representando uma parcela de informações não identificadas dentro dessa classificação. Esses percentuais oferecem um panorama da distribuição dos principais segmentos que compõem a base de dados no período analisado.

Tabela 3 - Distribuição da variável ramo

Ramo	Distribuição
ALIMENTAÇÃO	19,92%
VESTUÁRIO	18,99%
N/I	9,35%
CALÇADOS	7,07%

ARTIGOS DIVERSOS	6,82%
PERFUMARIA E COSMÉTICOS	5,21%
JÓIAS, RELÓGIOS E BIJOUTERIAS	5,19%
TELEFONIA E ACESSÓRIOS	3,44%
ARTIGOS DO LAR	2,85%
ÓTICAS	2,75%

A Tabela 3 apresenta as 10 principais classificações de ramos em termos de percentual no último retrato da base. Observa-se que o ramo de alimentação se destaca como o dominante nos shoppings, representando 19,92% da composição das lojas da base. Em seguida, o setor de vestuário figura com 18,99%.

Tabela 4 - Distribuição da variável tipo

Tipo	Distribuição
N/I	9,35%
OUTROS (SORVETE, BOMBONIÈRE, DOCETERIA, ETC)	7,38%
FAST-FOOD	6,48%
VESTUÁRIO FEMININO	4,6%
RESTAURANTE COM SERVIÇOS	3,71%
TELEFONIA CELULAR	3,42%
CALÇADOS FEMININOS	3,18%
OUTROS	3,17%
JOALHERIAS	3,13%
MODA JOVEM (UNISSEX)	3,04%

A Tabela 4, acima, discrimina as classificações de tipo com os maiores percentuais no último retrato da base. Observa-se que a categoria "N/I" apresenta a maior representatividade, com 9,35%. Em seguida, destaca-se o agrupamento "OUTROS (SORVETE, BOMBONIÈRE, DOCETERIA, ETC)" com 7,38%, seguido de perto pela categoria "FAST-FOOD", com 6,48%.

É importante destacar que, dentre as três categorias de tipo dominantes identificadas, a segunda e a terceira estão diretamente relacionadas ao setor de alimentação. Essa constatação reforça a dominância do ramo alimentício, previamente observada na análise da variável ramo, na composição das lojas que integram a base no período mais recente.

4.2.1.3. Visualização dos dados

Traduzir informações em formatos gráficos, permite uma facilitação na identificação de relações, distribuições e evoluções temporais, tornando a análise mais acessível e intuitiva. No contexto da concorrência em shoppings, a representação visual de dados como o número de lojas, a dinâmica de entradas e saídas, a distribuição geográfica e a composição setorial ao longo do tempo é o foco desta seção. As figuras que se seguem ilustram esses aspectos da base de dados.

A Figura 1 apresenta a evolução do número de lojas únicas nos shoppings da base de dados entre os anos de 2022 e 2025. Observa-se um aumento no número de lojas de 2022 para 2023, seguido por um declínio em 2024 e uma leve queda adicional em 2025. O pico no número de lojas únicas ocorreu em 2023. Essa dinâmica sugere um período de expansão do número de estabelecimentos nos shoppings da amostra até 2023, seguido por uma fase de retração nos anos subsequentes.

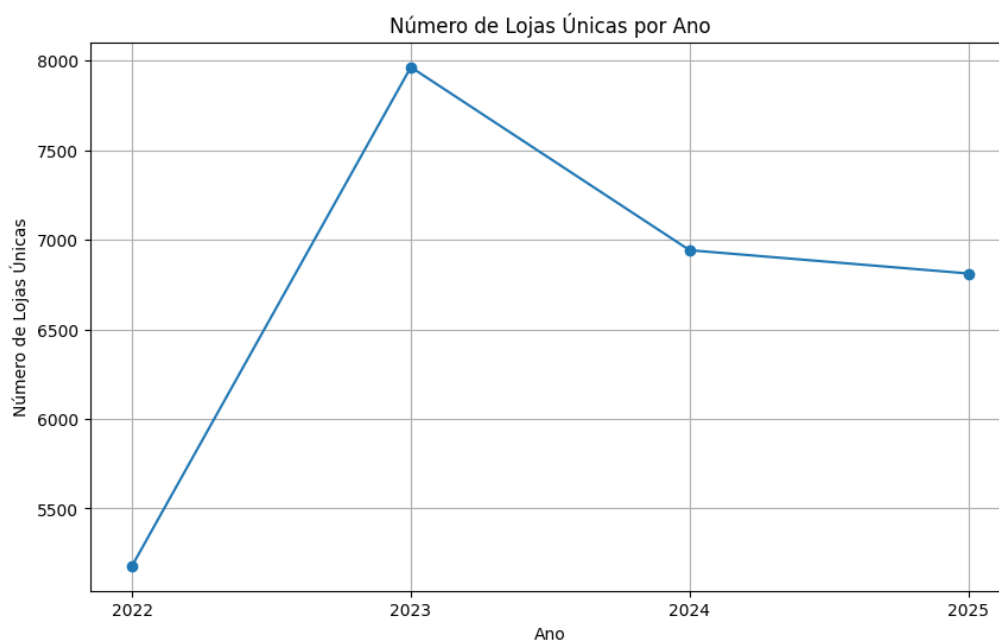


Figura 1 - Número de Lojas Únicas por ano

A Figura 2, que apresenta a série temporal de entradas e saídas de lojas por competência, permite a comparação direta da dinâmica de atração e perda de lojistas. Observa-se que, em alguns momentos, as linhas de entradas (verde) e saídas (vermelha) seguem tendências semelhantes, com picos e vales próximos, sugerindo

alta atividade de troca de lojistas, como em agosto de 2023. Contudo, há também períodos em que uma das linhas se sobressai, como em fevereiro de 2025, onde as saídas superaram as entradas, indicando uma retração. Essa análise é crucial para identificar ciclos de expansão ou retração e a volatilidade da ocupação ao longo do tempo.

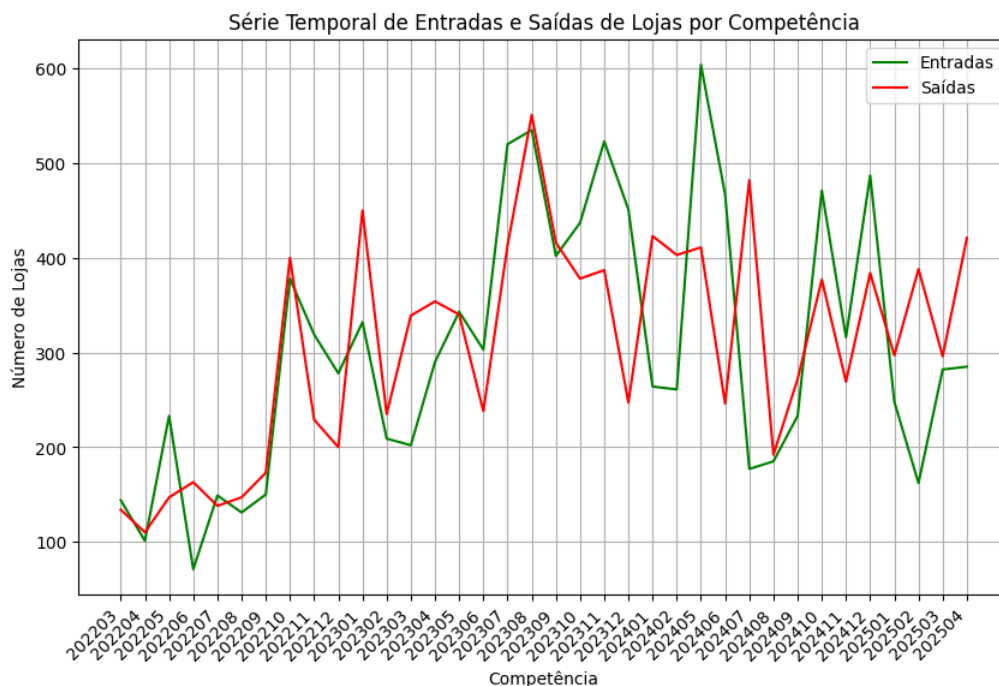


Figura 2 - Série temporal de entradas e saídas de lojas

A Figura 3 apresenta a curva ABC baseada no percentual acumulado de lojas por administradora. Esta curva demonstra uma concentração de lojas nas mãos de poucas administradoras, com as três maiores – Allos, Multiplan e JCPM – somando 61,1% do total de lojas, e as sete primeiras já respondendo por 81,0% dos estabelecimentos. Essa análise ABC confirma que um grupo seleto de empresas detém a maior parte do mercado em termos de volume de estabelecimentos, evidenciando a heterogeneidade no tamanho e na participação das administradoras e a maior influência de poucos players dominantes no setor.

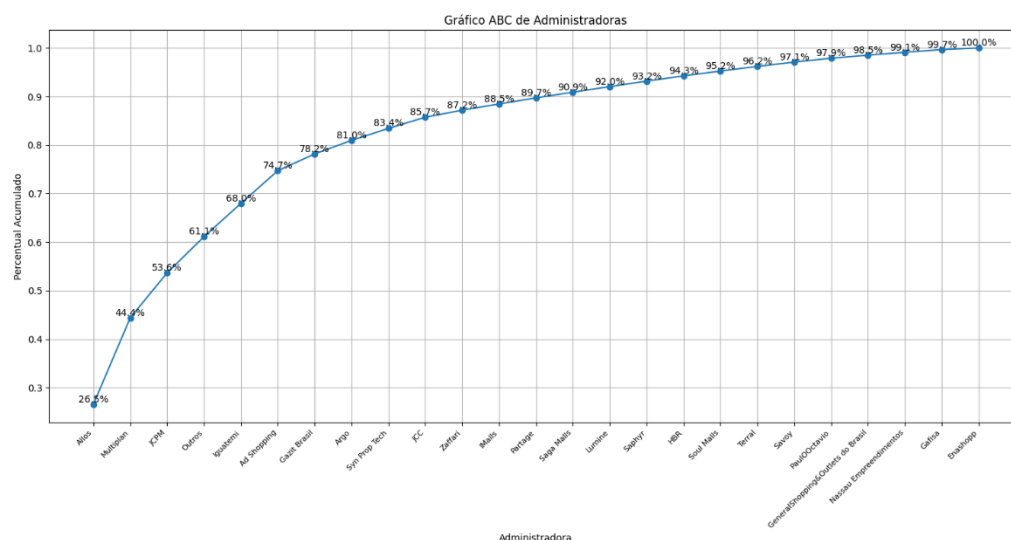


Figura 3 - Gráfico ABC de Administradoras

Com relação à Figura 4, percebe-se uma variação na quantidade de movimento ao longo do tempo. O gráfico apresenta picos, indicando períodos de alta atividade de abertura e fechamento de lojas. Por exemplo, há um pico expressivo em agosto de 2023, com mais de 1000 movimentos, e outro pico em maio de 2024, que ultrapassa a marca dos 1000 movimentos. Em contrapartida, há meses com menor atividade de movimento, como abril de 2022 e agosto de 2024. A natureza flutuante da linha sugere que o dinamismo da ocupação dos espaços nos shoppings está sujeito a influências temporais, que podem incluir fatores sazonais, econômicos ou estratégias de gestão do mix por parte das administradoras. Essa figura é fundamental para compreender a volatilidade e os momentos de maior e menor efervescência no mercado de locação de lojas em shopping centers.

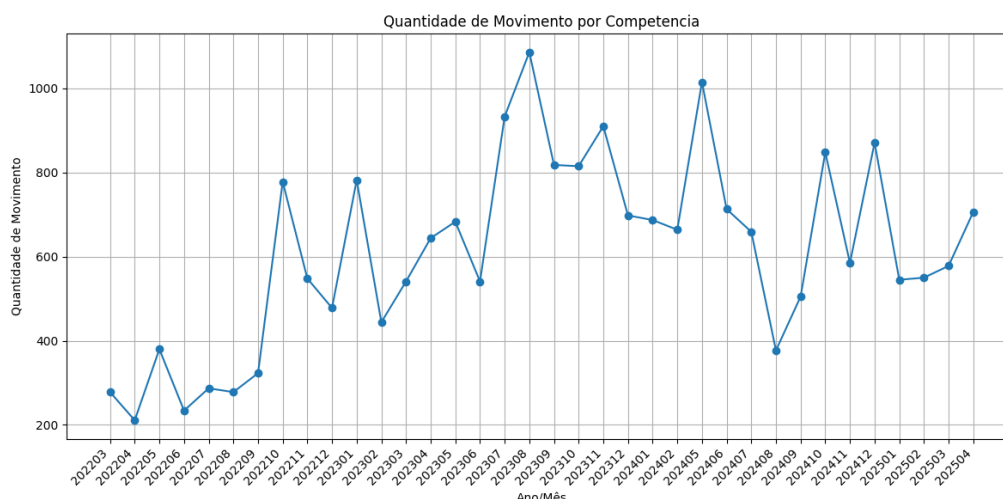


Figura 4 - Quantidade de movimento por competencia

A Figura 5 mostra a evolução percentual dos pilares de segmentos de lojas nos shoppings analisados entre 2022 e 2025. Observa-se uma tendência de queda na representatividade do pilar "VAREJO", que passa de 53,8% em 2022 para 46,3% em 2025. "ENTRETENIMENTO E LAZER" mantém participação estável, variando de 21,9% para 22,9% no mesmo período. O pilar "SERVIÇO" também apresenta leve redução, de 13,0% para 12,1%. Já "BEM-ESTAR" cai de 11,3% para 9,6%, enquanto o pilar "N/I", inexistente até 2023, surge com 2,0% em 2024 e alcança 9,1% em 2025, indicando a entrada de lojas não classificadas nas categorias anteriores.

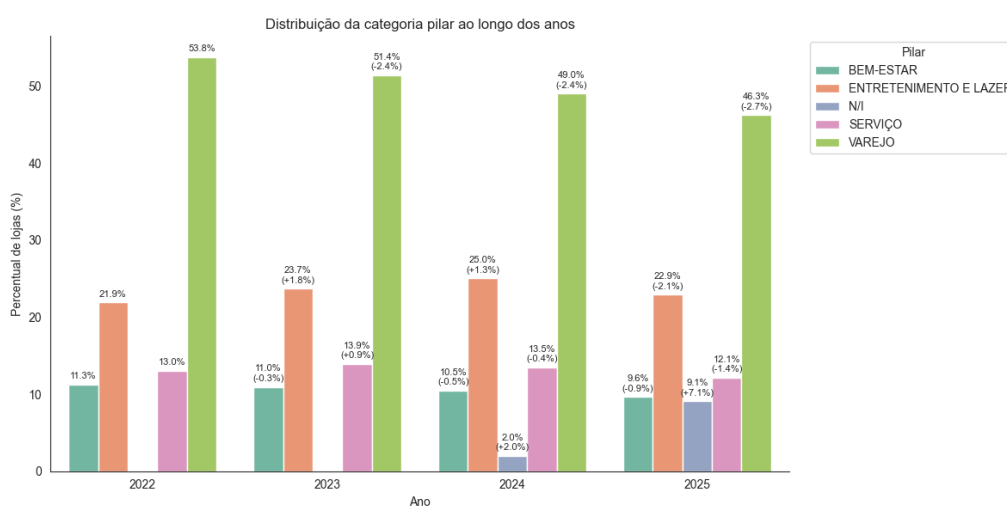


Figura 5 - Proporção de pilar ao longo dos anos

Essa mudança na composição do mix de lojas sugere uma adaptação estratégica dos shoppings às novas demandas do mercado e aos padrões de consumo, que podem estar migrando de um foco exclusivo no varejo de produtos para uma maior valorização de experiências e serviços. A diminuição da fatia do varejo, mesmo que gradual, e a manutenção da relevância dos demais pilares indicam uma busca por maior diversificação e por um modelo de negócio mais atrativo para os consumidores que buscam os shoppings para desfrutar de experiências, e não somente compras.

A Figura 6 apresenta um mapa de calor em formato de mapa do Brasil, ilustrando o percentual de movimentações de lojas por região do país. A intensidade das cores, variando do azul (menor movimentação) ao amarelo (maior movimentação), passando por roxo, rosa e laranja, permite identificar as áreas de maior dinamismo no setor de shopping centers. A Região Sudeste destaca-se com uma concentração maior, registrando 64,1% das movimentações, o que a posiciona como o principal polo de atividade no mercado de shoppings. Em seguida, a Região Nordeste aparece com 17,3% das movimentações, indicando um volume intermediário de transações. As regiões Centro-Oeste, Sul e Norte apresentam volumes menores, com a Região Norte sendo a que registrou a menor atividade, com 4,3% das movimentações. Essa distribuição geográfica reflete as disparidades regionais no desenvolvimento econômico e na infraestrutura de shopping centers no Brasil, com uma concentração no Sudeste.

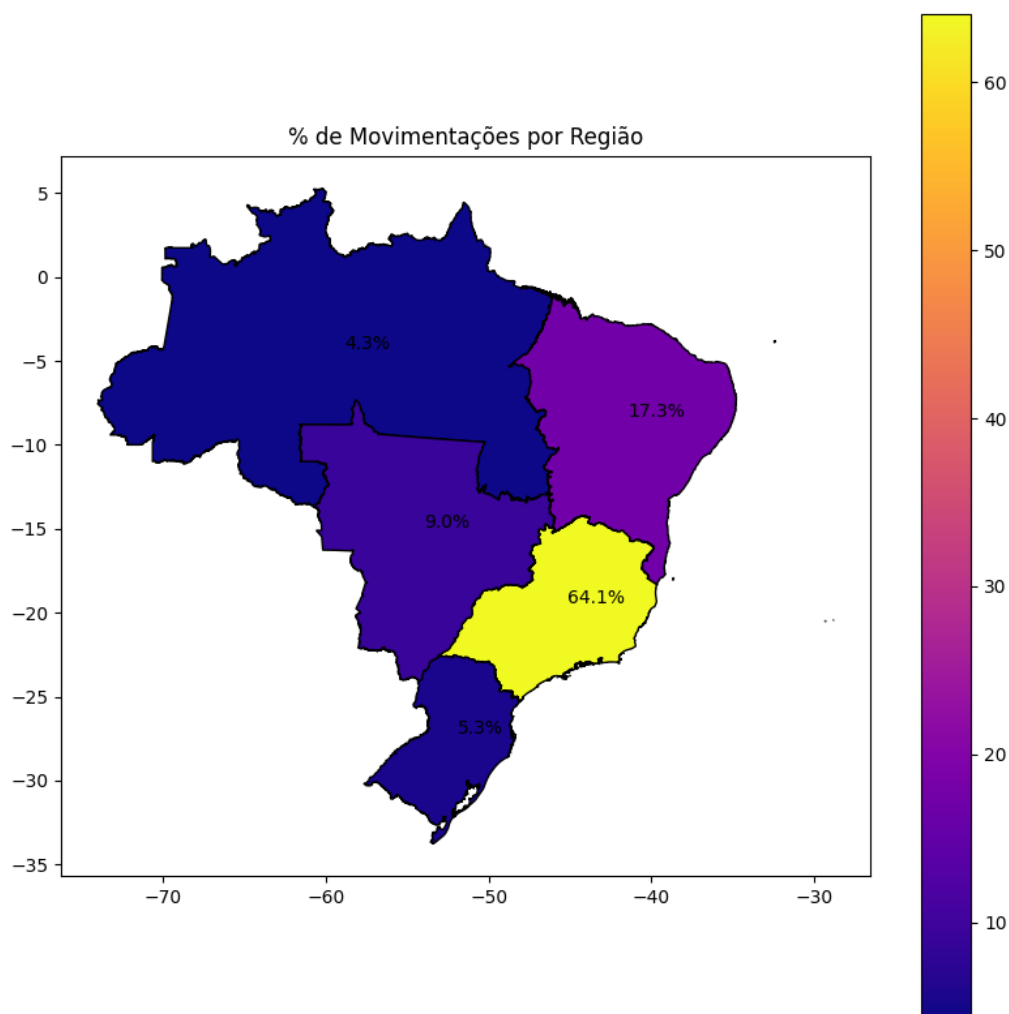


Figura 6 - Percentual de movimentações por região

A Figura 7 apresenta um mapa de calor do Brasil que ilustra a proporção percentual do total de movimentações de lojas por estado, complementando a visão anterior, sobre regiões. A intensidade da cor no mapa reflete a concentração dessas movimentações, e são destacados com rótulos os estados cuja participação supera 2%. São Paulo se sobressai com uma proporção de 39,9% das movimentações, o que o estabelece como o epicentro da dinâmica de lojas de shopping no país. O Rio de Janeiro segue como o segundo estado mais relevante, com 22% das movimentações, reforçando a concentração no Sudeste. Ceará, com 9,7%, e Rio Grande do Sul, com 5,3%, também demonstram uma participação significativa. Outros estados com proporções acima de 2% incluem Pernambuco (3,4%), Goiás e Distrito Federal (ambos com 2,7%), Mato Grosso (2,4%), Minas Gerais (2,2%) e Bahia (2,1%).

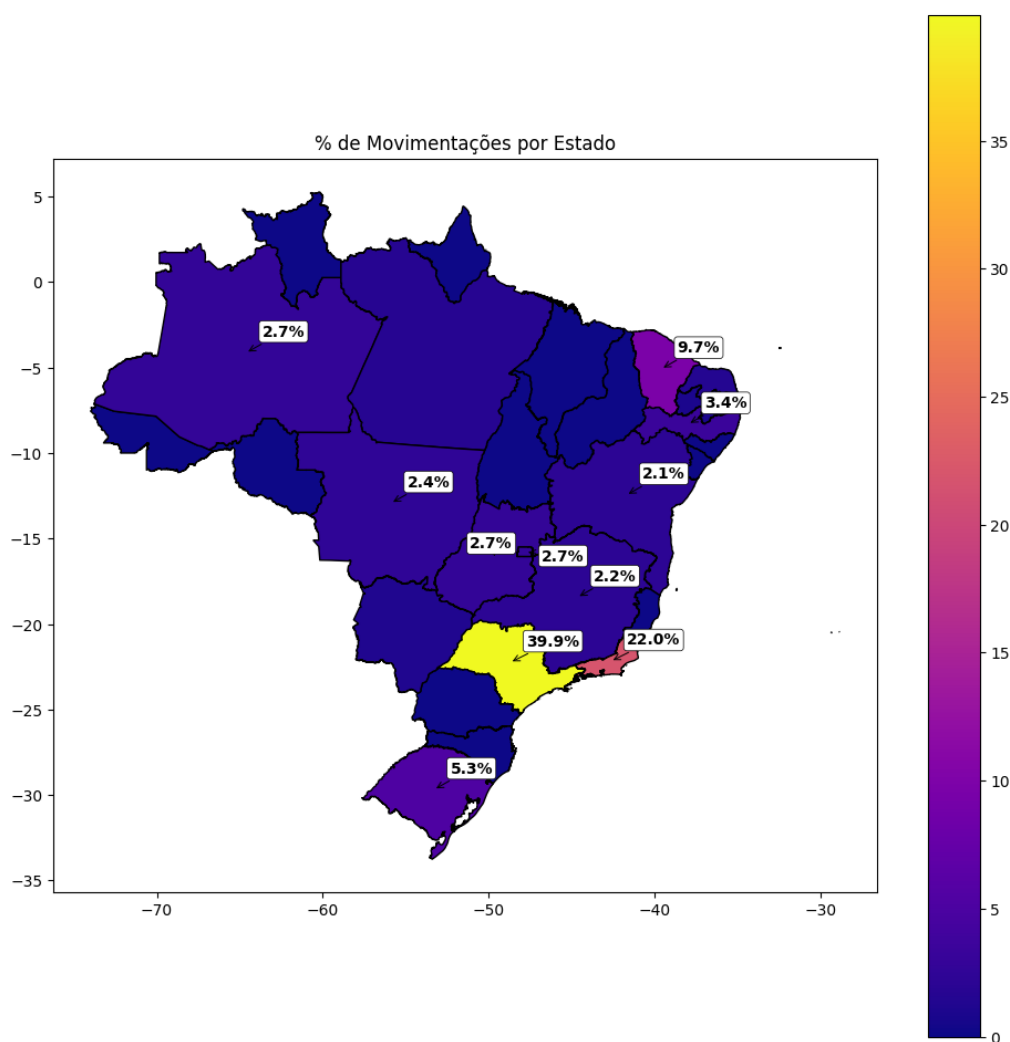


Figura 7 - Percentual de movimentações por estado

4.3. Análise Diagnóstica

Após a análise descritiva apresentada na seção anterior, que permitiu identificar padrões e tendências na dinâmica de concorrência dos shoppings, a presente seção avança para a Análise Diagnóstica. Esta etapa tem como objetivo primordial investigar as causas e as relações subjacentes que influenciam a movimentação de lojas e a estrutura do mix comercial. A partir dos insights levantados na fase descritiva, serão formuladas e testadas hipóteses que visam desvendar os fatores determinantes para a entrada e saída de lojas, bem como as variações na composição setorial dos shoppings ao longo do tempo.

4.3.1. Relação entre movimento e ramo

Um dos primeiros pontos de investigação diagnóstica reside na compreensão da relação entre a movimentação de lojas e suas classificações por ramo e pilar de segmento. A análise descritiva já indicou a predominância de certos segmentos no mix de lojas, como Varejo e Entretenimento e Lazer. É fundamental verificar como essa representatividade se relaciona com as taxas de entrada e saída, ou seja, se os ramos e pilares dominantes na base apresentam maior estabilidade (baixa rotatividade) ou se são, paradoxalmente, áreas de alta movimentação. Busca-se responder: Quais ramos apresentam maior taxa de entrada e saída e como isso se relaciona com sua participação na base de lojas?

Para investigar essa relação, será calculada a taxa de entrada e saída por ramo, permitindo uma comparação da dinâmica de cada categoria.

A Figura 8, apresenta os dez ramos que demonstraram as maiores taxas de entrada de lojas nos shoppings da base de dados, representadas pela cor verde. Ramos como "ARTIGOS ESPORTIVOS" e "MÓVEIS/DECORAÇÃO/CAMA, MESA E BANHO" exibem taxas de entrada elevadas, indicando uma alta frequência de novas aberturas ou renovações de estabelecimentos nesses setores. Outros ramos como "LAN HOUSES", "VINHOS" e "AGÊNCIA DE VIAGENS" também figuram entre os de maior taxa de entrada.

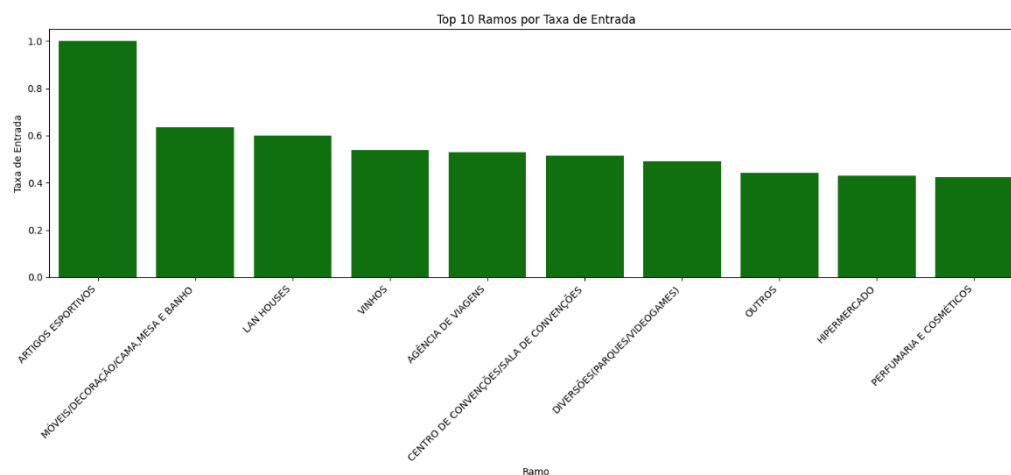


Figura 8 - Top 10 ramos por taxa de entrada

Em contrapartida, a Figura 9, "Top 10 Ramos por Taxa de Saída", ilustra os dez ramos com as maiores taxas de saída de lojas. Ramos como "VINHOS" e "HOTEL" apresentam as maiores taxas de saída, indicando maior volatilidade. "CENTRO DE CONVENÇÕES/SALA DE CONVENÇÕES", "CONSTRUÇÃO E

DECORAÇÃO (NOVO)" e "MÓVEIS/DECORAÇÃO/CAMA, MESA E BANHO" também se destacam com taxas de saída elevadas.

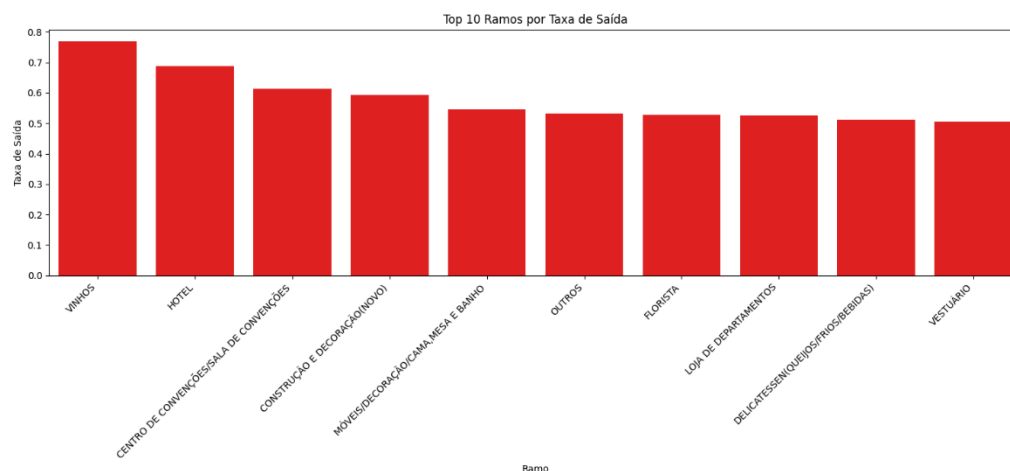


Figura 9 - Top 10 ramos por taxa de saída

A análise conjunta das Figuras 8 e 9, em correlação com a participação percentual dos ramos na base total de lojas, permite estabelecer relações concretas. Observa-se que ramos dominantes na base, como "ALIMENTAÇÃO" e "VESTUÁRIO", que representam, respectivamente, 19,92% e 18,99% das lojas ativas, não figuram entre os de maior taxa de entrada ou saída. A Figura 9, por exemplo, mostra "VESTUÁRIO" como o décimo ramo com maior taxa de saída, indicando uma rotatividade presente, mas não tão acentuada quanto a de outros ramos de menor expressão na base.

Por outro lado, ramos que possuem menor representatividade na base total demonstram as maiores taxas de movimentação. Por exemplo, "VINHOS" e "MÓVEIS/DECORAÇÃO/CAMA, MESA E BANHO" aparecem consistentemente com altas taxas tanto de entrada quanto de saída. Este padrão sugere que, embora não sejam os maiores em volume absoluto na base, esses segmentos são caracterizados por um alto dinamismo e renovação, com novas lojas entrando frequentemente e outras saindo, possivelmente devido a um mercado mais competitivo, nichos de alta volatilidade ou modelos de negócio com maior taxa de insucesso. A prevalência de ramos menores nos rankings de rotatividade indica que a dominância na base não necessariamente se traduz em alta volatilidade, e que

segmentos com menor volume podem experimentar uma dinâmica de entrada e saída proporcionalmente mais intensa.

4.3.2. Relação entre movimento e administradoras

A análise diagnóstica prossegue investigando se a administradora do shopping exerce influência sobre a rotatividade de lojas. A seção descritiva já evidenciou a concentração de lojas em determinadas administradoras. É pertinente verificar se essa escala se relaciona com a capacidade de retenção ou atração de lojistas, ou se algumas administradoras apresentam taxas de rotatividade significativamente distintas das demais.

A Figura 10 ilustra as administradoras com as maiores taxas de entrada de lojas. As maiores taxas de entrada são observadas em administradoras como JCC e Multiplan, seguidas por Allos, JCPM, Saphyr, Iguatemi, IMalls, Ad Shopping, HBR e Argo.

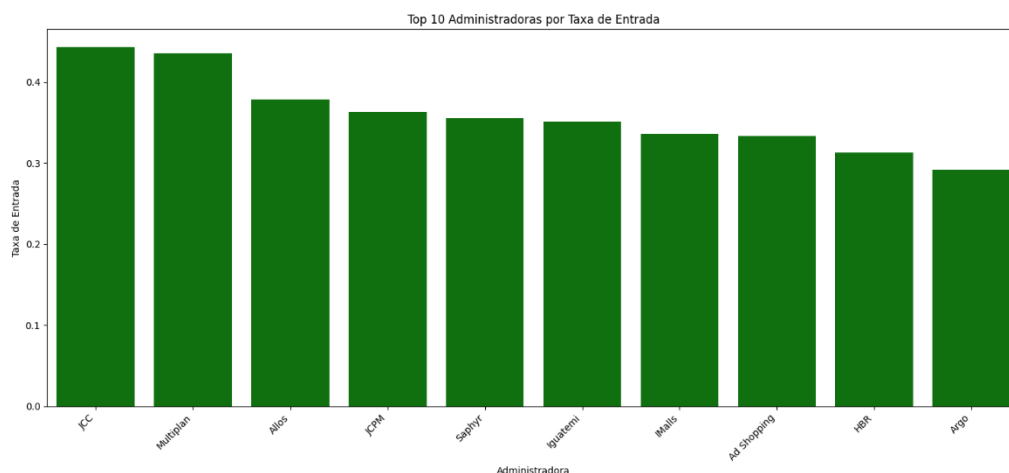


Figura 10 - Top 10 administradoras por taxa de entrada

A Figura 11, por sua vez, apresenta as administradoras com as maiores taxas de saída de lojas. As maiores taxas de saída são observadas em Multiplan e Gazit Brasil, seguidas por Allos, Ad Shopping, JCPM, Iguatemi, Argo, Partage, Saphyr e Terral.

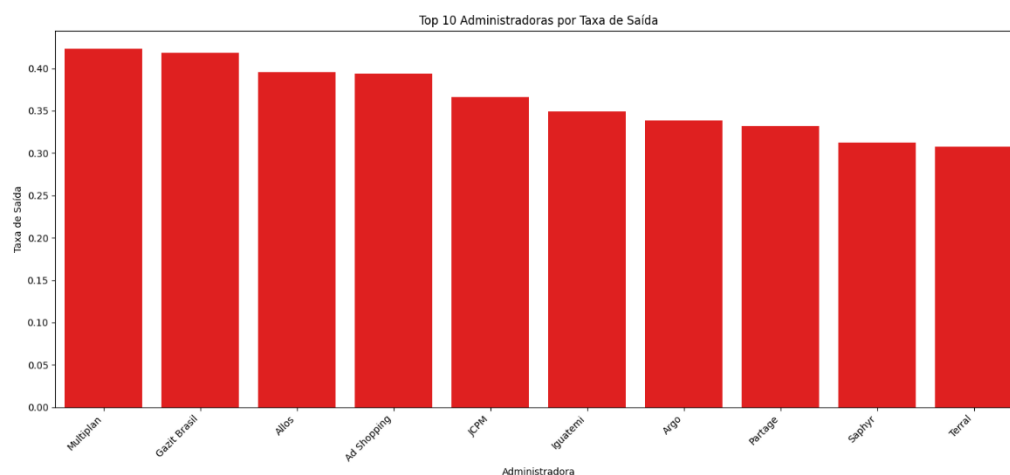


Figura 11 - Top 10 administradoras por taxa de saída

Ao correlacionar esses resultados com a dominância na base de dados, que mostra que a "Allos" e "Multiplan" detêm a maior parte das lojas na base, percebe-se uma dinâmica interessante. A "Multiplan", embora seja uma das administradoras mais relevantes em número total de lojas, figura entre as de maior taxa de entrada e, especialmente, entre as de maior taxa de saída. Isso sugere que, mesmo sendo um grande player, a Multiplan pode operar em um ambiente de maior renovação do mix de lojas ou em shoppings com maior fluxo de entrada e saída de locatários. A "Allos", outra gigante em volume de lojas, também aparece entre as de maior taxa de entrada e saída, embora com taxas menores que a Multiplan. Administradoras como "JCC" e "Gazit Brasil", apesar de talvez não estarem no topo em volume absoluto de lojas na base, mostram-se com taxas de movimentação significativas, com JCC liderando as entradas e Gazit Brasil se destacando nas saídas.

Essa análise diagnóstica revela que a dimensão da administradora não é o único fator que determina a estabilidade ou a volatilidade do mix de lojas. Administradoras com grandes portfólios podem experimentar alta rotatividade, o que pode ser uma estratégia de gestão ativa para renovação de portfólio ou um reflexo de condições de mercado nos shoppings sob sua gestão.

4.3.3. Sazonalidade das movimentações

A análise descritiva indicou flutuações mensais nas entradas e saídas de lojas, levantando a questão sobre a existência de padrões sazonais. Esta seção aprofundará essa observação, buscando confirmar se há períodos do ano com maior

ou menor intensidade de movimentação, o que poderia estar relacionado a fatores como datas comemorativas, períodos de férias, ciclos de renovação de contratos ou condições econômicas cíclicas.

Para investigar a sazonalidade, serão analisados os dados de entrada e saída de lojas agrupados por mês, ao longo de todo o período da base de dados, para identificar tendências recorrentes.

A Figura 12 ilustra o número total de lojas de entrada agrupado por mês do ano. Observa-se que a atividade de entradas apresenta um padrão sazonal, com picos nos meses de maio, outubro e dezembro. Em contrapartida, os meses de fevereiro e abril registram os menores volumes de entradas, com valores abaixo de 600.

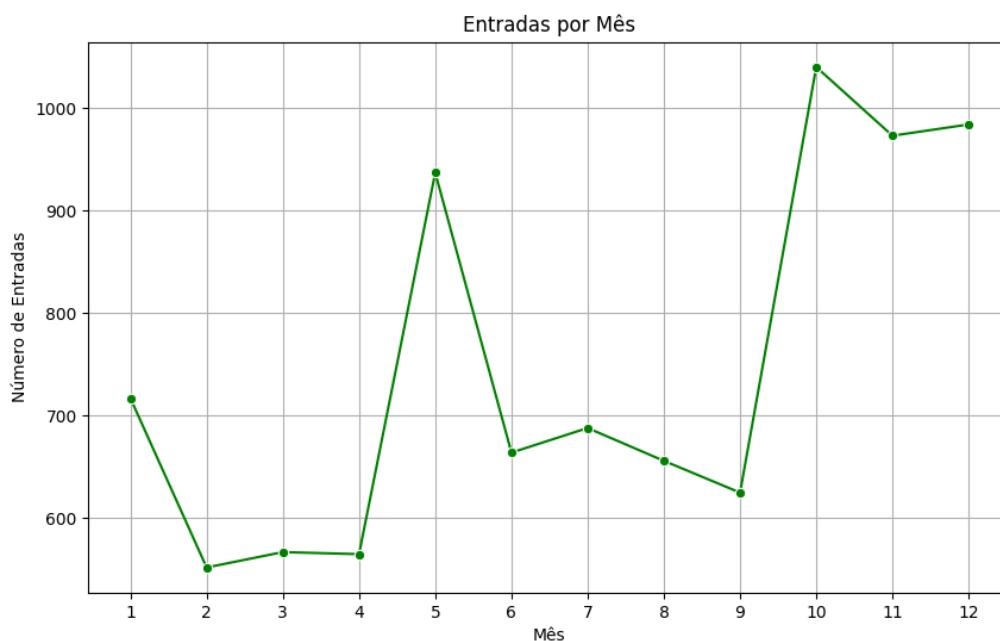


Figura 12 - Entradas por mês

A Figura 13 exibe o número total de lojas de saída agrupado por mês do ano. Similarmente, o gráfico de saídas também demonstra uma sazonalidade distinta. Os maiores picos de saídas são observados nos meses de janeiro, julho e outubro. Os meses de março e junho apresentam os menores volumes de saídas.

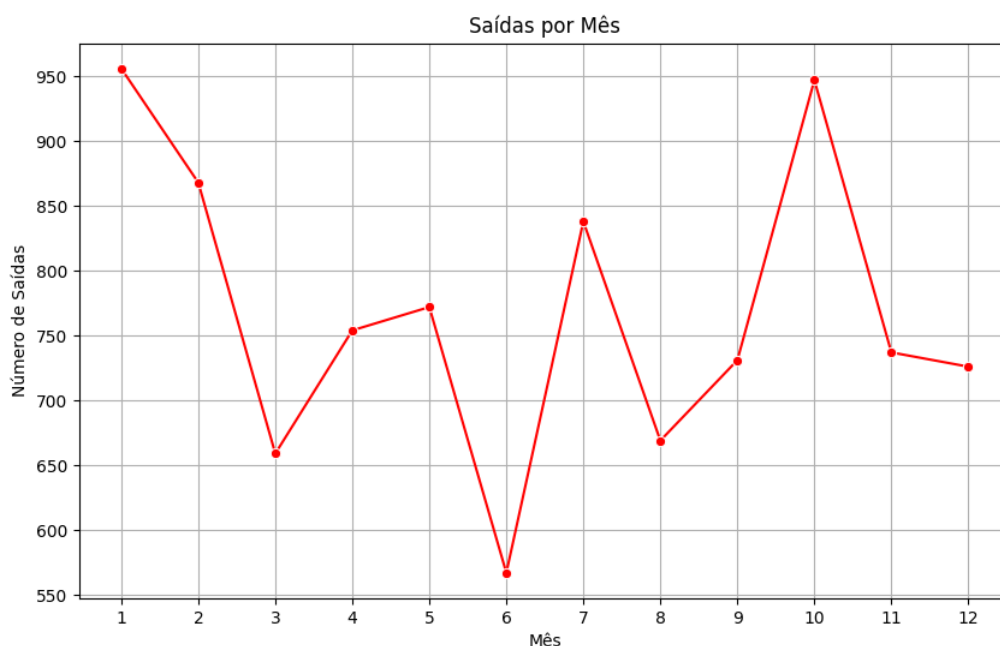


Figura 13 - Saídas por mês

A comparação das Figuras 12 e 13 revela que a sazonalidade no movimento de lojas é presente, mas com padrões distintos para entradas e saídas. Picos de entradas em maio, outubro e dezembro podem estar relacionados a períodos de aquecimento comercial, como o planejamento para o Dia das Mães, Black Friday e Natal, estimulando novas aberturas. Já as saídas, com picos em janeiro e julho, podem refletir o encerramento de operações após o período de festas de fim de ano ou o ajuste de contratos ao final de semestres. Os meses de menor movimento de entradas (fevereiro e abril) e de saídas (março e junho) indicam períodos de menor dinamismo ou reajuste de mercado. A compreensão desses ciclos sazonais é fundamental para administradoras de shoppings planejarem suas estratégias de comercialização e gestão de vacâncias.

4.3.4. Impacto de fatores geográficos na rotatividade

A análise descritiva já revelou uma concentração de movimentações de lojas em determinadas regiões e estados do Brasil, com o Sudeste e, em particular, São Paulo e Rio de Janeiro, destacando-se em volume absoluto. Esta seção diagnóstica aprofundará essa observação, buscando determinar se essa alta concentração de movimentações se traduz em taxas de rotatividade distintas entre as regiões e

estados. A hipótese é que regiões com maior volume de mercado podem apresentar dinâmicas de rotatividade diferentes das regiões com menor volume.

Para investigar isso, serão calculadas as taxas de entrada e saída por região e por estado, permitindo comparar a volatilidade do mix de lojas em diferentes contextos geográficos.

4.3.4.1. Análise por região do Brasil

A Figura 14 apresenta as taxas de entrada de lojas para as diferentes regiões do Brasil. A Região Sul se destaca com a maior taxa de entrada, seguida pelo Sudeste, Nordeste, Centro-Oeste e, por fim, a Região Norte com a menor taxa. Já a Figura 15, ilustra as taxas de saída de lojas por região. A Região Sudeste exibe a maior taxa de saída, seguida pelo Nordeste, Centro-Oeste, Sul e, por último, a Região Norte com a menor taxa de saída.

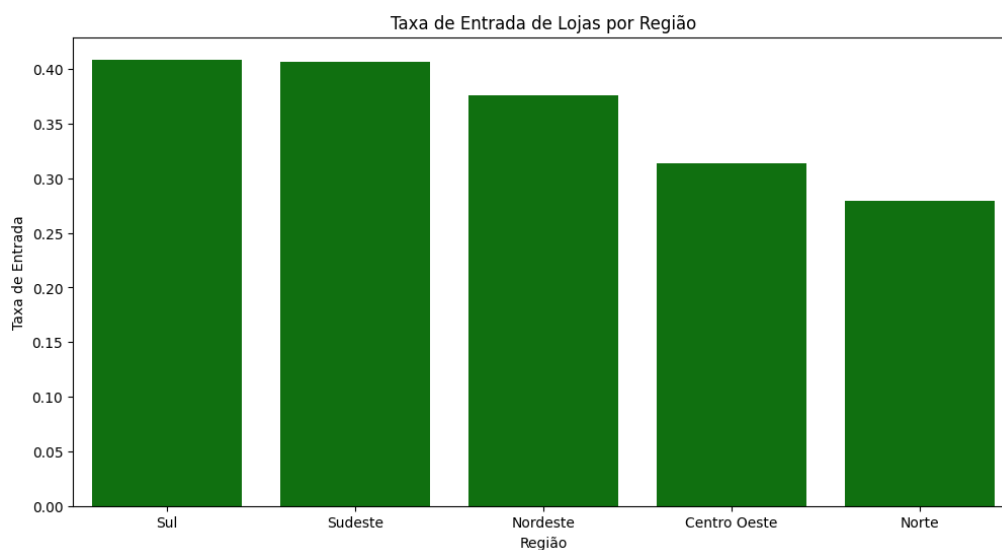


Figura 14 - Taxa de entrada de lojas por região

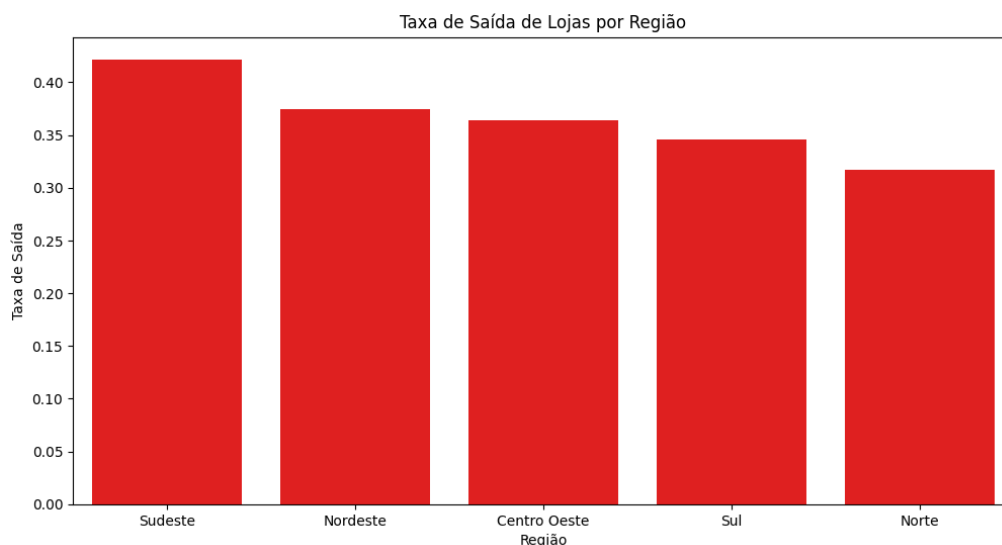


Figura 15 - Taxa de saída de lojas por região

A análise comparativa entre as Figuras 14 e 15, e em relação aos volumes absolutos de movimentação por região, permite inferir relações importantes. A Região Sudeste, que já demonstrava o maior volume absoluto de movimentações (64,1% das movimentações totais), também apresenta a maior taxa de saída, indicando um mercado de alta rotatividade. Surpreendentemente, a Região Sul, que tem um volume absoluto de movimentações muito inferior ao Sudeste (4,3% das movimentações totais), lidera em taxa de entrada, mas apresenta uma das menores taxas de saída, sugerindo um crescimento líquido mais expressivo ou maior estabilidade em sua base de lojas.

Já a Região Nordeste, com o segundo maior volume absoluto de movimentações (17,3%), figura em segundo lugar tanto nas taxas de entrada quanto de saída, indicando um mercado dinâmico e em crescimento. As regiões Centro-Oeste e Norte, que possuem os menores volumes absolutos de movimentação, também exibem as menores taxas de entrada e saída, sugerindo mercados menos efervescentes.

Essa análise revela que o volume absoluto de movimentação não é diretamente proporcional à taxa de entrada, mas há uma correlação mais forte com a taxa de saída nas regiões de maior volume, apontando para mercados mais desenvolvidos que, embora atraíam muitas novas lojas, também lidam com uma alta taxa de desocupação ou renovação.

4.3.4.2. Análise por estados do Brasil

Para uma compreensão mais granular do impacto geográfico na rotatividade, a análise se aprofunda no nível estadual. Serão calculadas e comparadas as taxas de entrada e saída de lojas para os estados com maior representatividade na base de dados, permitindo identificar padrões e anomalias que não são visíveis na agregação por região.

A Figura 16 apresenta os estados presentes na base e suas taxas de entrada de lojas. Os estados que lideram essa taxa são Rio de Janeiro (RJ) e Ceará (CE). O Rio Grande do Sul (RS) e São Paulo (SP) também figuram entre os estados com altas taxas de entrada.

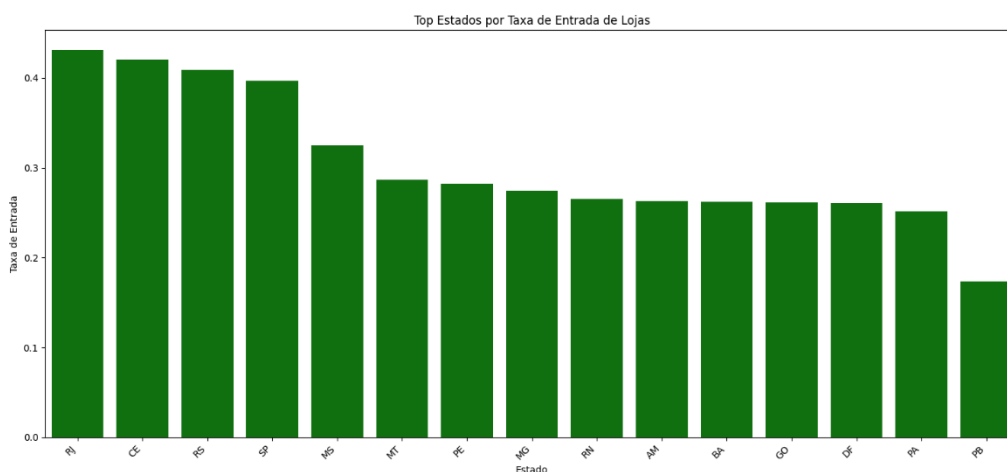


Figura 16 - Taxa de entrada de lojas por estado

Na Figura 17, que ilustra as taxas de saída de lojas por estado, os maiores valores são observados no Rio de Janeiro (RJ) e em São Paulo (SP), seguidos de perto por Ceará (CE) e Goiás (GO). O Rio Grande do Sul (RS) também apresenta uma taxa de saída expressiva.

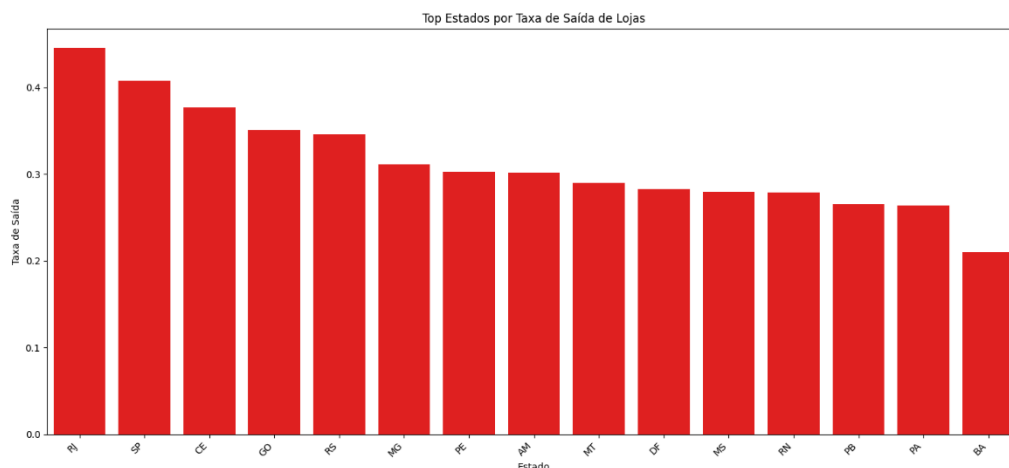


Figura 17 - Taxa de saída de lojas por estado

A análise conjunta dessas taxas de entrada e saída por estado, em comparação com a porcentagem de movimentações totais por estado, revela padrões importantes. Rio de Janeiro (RJ) e São Paulo (SP), que concentram a maior parte das movimentações absolutas do país (22% e 39,9% respectivamente), também se destacam com as maiores taxas de saída. Isso indica que, embora sejam os maiores mercados em volume de movimentação, são ambientes de altíssima rotatividade de lojas, com um fluxo constante de fechamentos, sugerindo um mercado dinâmico e competitivo. O Ceará (CE) e o Rio Grande do Sul (RS) também exibem essa dualidade de altas taxas de entrada e saída, reforçando a ideia de mercados dinâmicos e com intensa renovação do mix de lojas. Essa granularidade por estado permite identificar que a efervescência do mercado e a concorrência não se restringem apenas aos grandes polos em volume absoluto, mas também se manifestam em estados com menor volume, onde a taxa de rotatividade pode ser um indicador crucial da dinâmica competitiva local.

4.4. Análise Preditiva e Interpretação dos Resultados

A transição da análise de dados históricos para a capacidade de antecipar eventos futuros constitui um pilar fundamental para a tomada de decisão estratégica em ambientes competitivos. Após a conclusão das análises descritiva e diagnóstica, que permitiram mapear a dinâmica da concorrência e identificar os fatores que influenciam a movimentação de lojas em shopping centers, a Análise Preditiva emerge como o próximo passo crucial desta pesquisa. O objetivo central aqui é

desenvolver modelos capazes de prever a ocorrência de entradas e saídas de lojas, oferecendo às administradoras de shopping centers uma ferramenta para antecipar mudanças no mix, otimizar estratégias de comercialização, mitigar riscos de vacância e identificar oportunidades de atração de novos lojistas.

Conforme Petropoulos et al. (2022), a previsão é um campo em constante evolução, onde a integração de métodos estatísticos tradicionais e abordagens de Machine Learning tem se mostrado cada vez mais relevante. Este trabalho se baseia na estrutura de previsão de Petropoulos et al. (2022), que enfatiza a necessidade de uma metodologia robusta que abranja a preparação dos dados, a seleção criteriosa de modelos, a avaliação rigorosa de desempenho e a interpretabilidade dos resultados.

4.4.1. Avaliação, métricas e hiperparâmetros dos modelos

A modelagem utilizou como targets a quantidade de entradas e saídas, sendo o conjunto de treinamento o período de março de 2022 a abril de 2025. O *Random Forest Regressor* e o *Gradient Boosting Regressor* foram comparados com benchmarks estatísticos.

O treinamento dos modelos de Machine Learning envolveu a otimização de hiperparâmetros para equilibrar o ajuste e a capacidade de generalização. Por exemplo, o *Random Forest Regressor* utilizou parâmetros como o número de árvores (*n_estimators*) e a profundidade máxima (*max_depth*), enquanto o *Gradient Boosting Regressor* utilizou a taxa de aprendizado (*learning_rate*) e regularizações, visando o melhor desempenho preditivo.

Os resultados da execução dos modelos de previsão no conjunto de treinamento e as projeções para o futuro são apresentados nas Tabelas 5 e 6, respectivamente, para entradas, e Tabelas 7 e 8, para saídas. Os ajustes de cada modelo ao período de teste também são apresentados, representados pelas Figuras 18, 19 e 20 para entradas, sendo a Regressão Linear, *Random Forest Regressor* e *Gradiente Boosting Regressor*, respectivamente, e Figuras 22, 23 e 24 para saídas. Além das representações em gráfico, das previsões de entradas (Figura 21) e previsões de saídas (Figura 25).

Tabela 5 - Métricas dos modelos de previsão para entradas

Entradas	Regressão Linear	<i>Random Forest Regressor</i>	<i>Gradient Boosting Regressor</i>
RMSE	77,73	43,79	4,35
MAE	59,26	35,71	3,45
R-quadrado	0,68	0,90	1

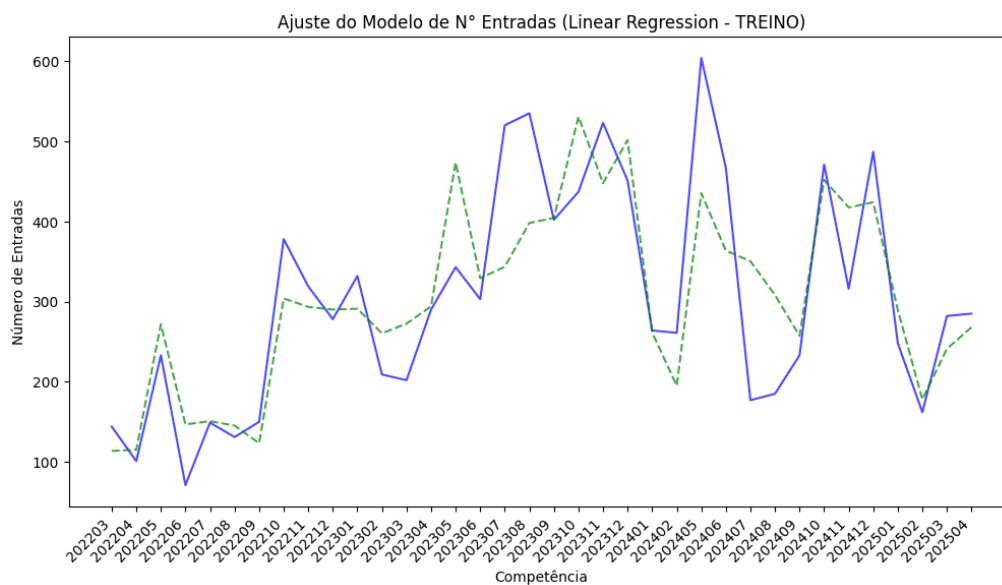


Figura 18 - Ajuste da Regressão Linear no treino para entradas

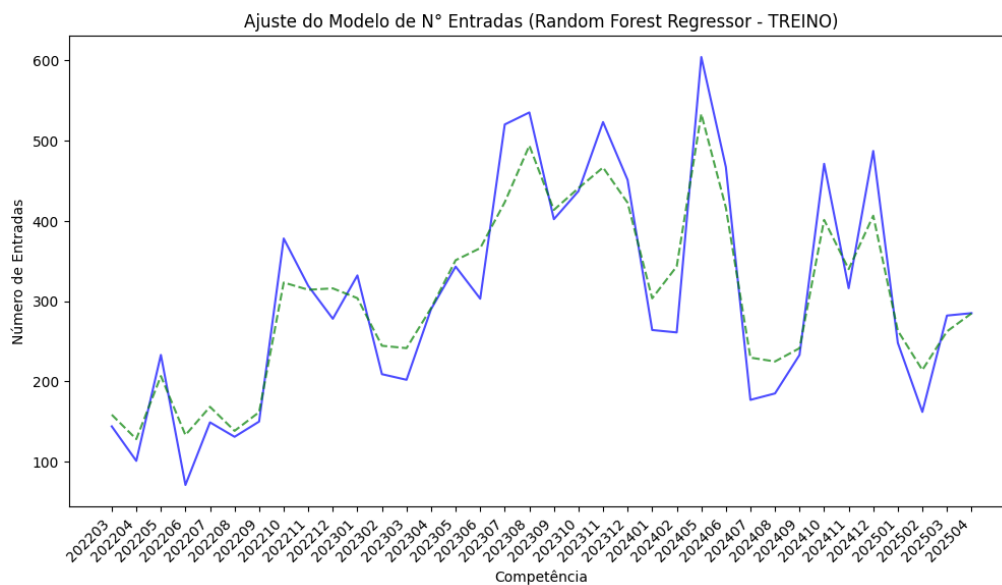


Figura 19 - Ajuste do Random Forest no treino para entradas

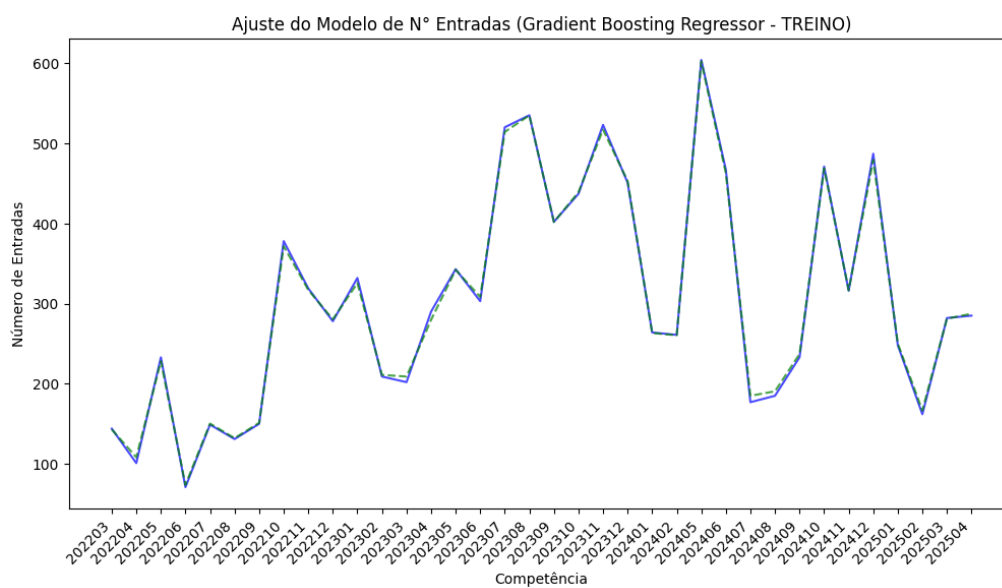


Figura 20 - Ajuste do Gradient Boosting no treino para entradas

Tabela 6 - Previsão para entradas

Competência	Previsão de entradas
202505	299
202506	280
202507	278
202508	271
202509	277
202510	350
202511	286
202512	324

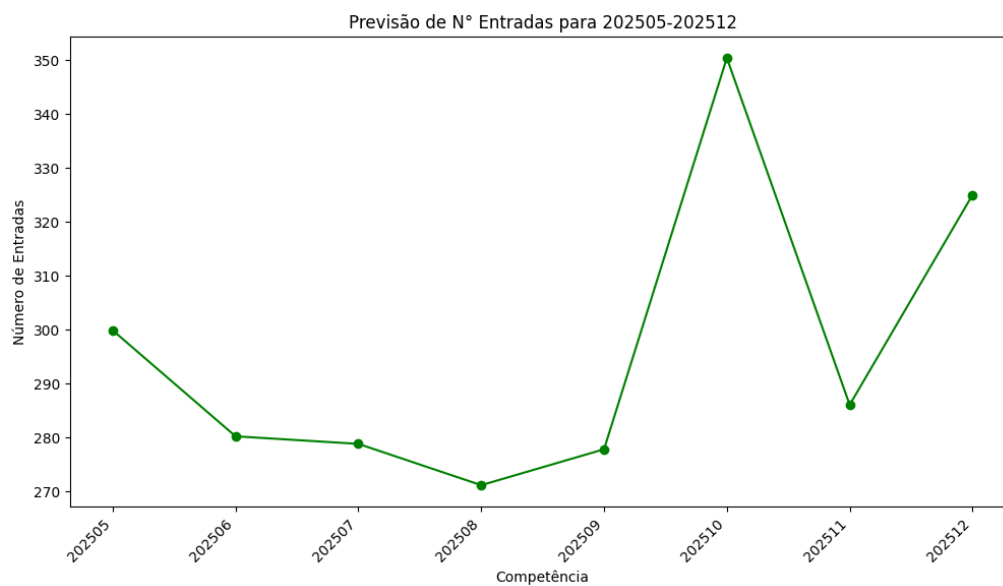


Figura 21 - Previsão de entradas

Tabela 7 - Métricas dos modelos de previsão para saídas

Saídas	Regressão Linear	<i>Random Forest Regressor</i>	<i>Gradient Boosting Regressor</i>
RMSE	66,19	34,21	2,89
MAE	55,92	26,21	2,33
R-quadrado	0,66	0,91	1

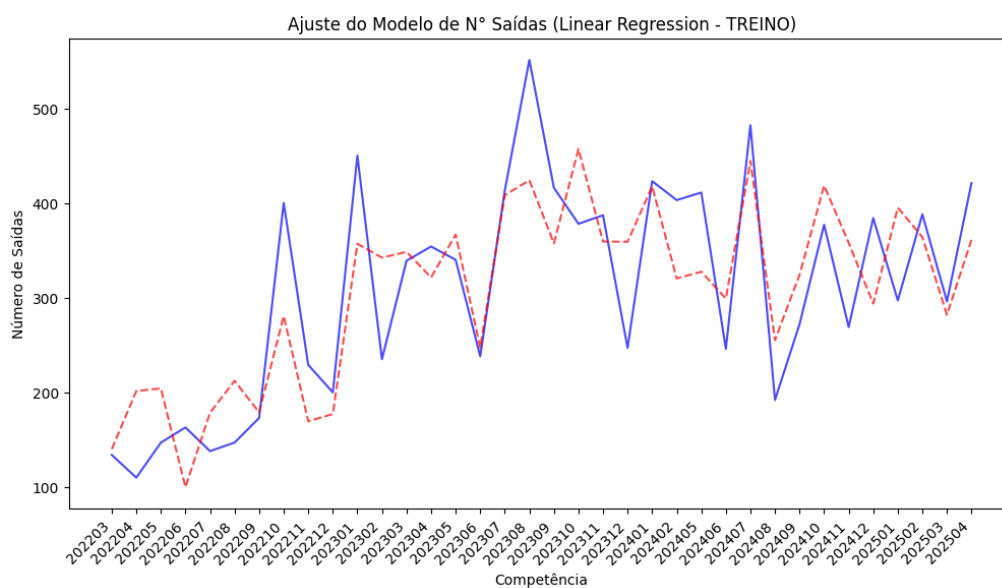


Figura 22 - Ajuste da Regressão Linear no treino para saídas

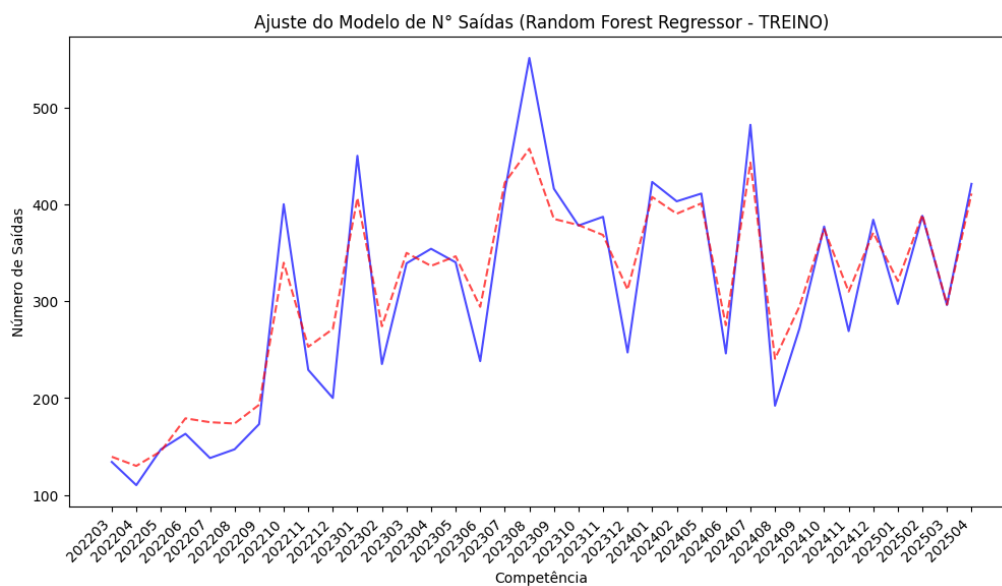


Figura 23 - Ajuste do Random Forest no treino para saídas

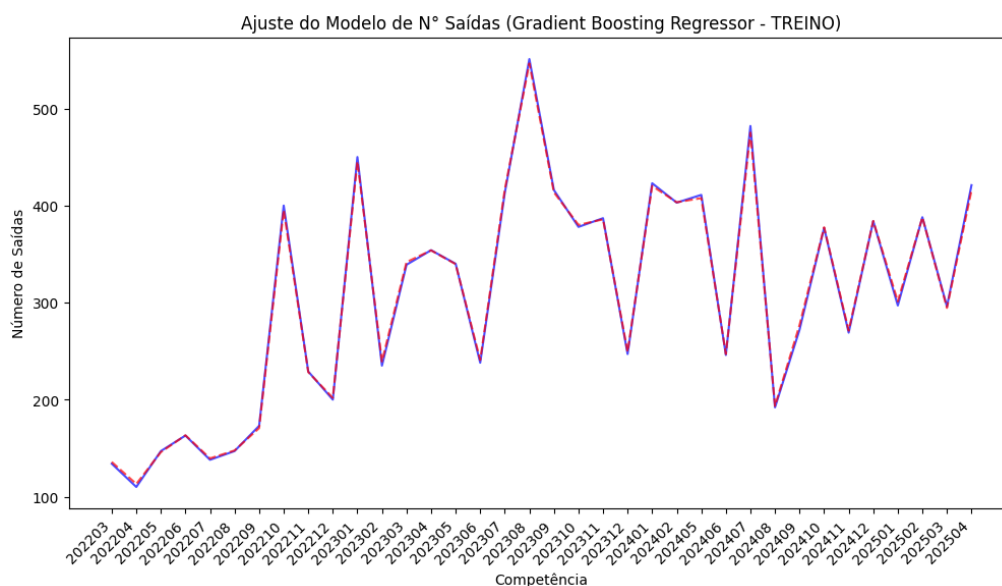


Figura 24 - Ajuste do Gradient Boosting no treino para saídas

Tabela 8 - Previsão para saídas

Competencia	Previsão de saídas
202505	327
202506	302
202507	331
202508	326
202509	311
202510	328

202511	316
202512	312



Figura 25 - Previsão de saídas

4.4.2. Interpretação dos resultados preditivos

Os resultados da modelagem preditiva oferecem uma visão sobre a capacidade dos modelos de estimar o número de entradas e saídas de lojas por competência.

A análise das métricas de desempenho no conjunto de treinamento revela que os modelos baseados em árvores (*Random Forest Regressor* e *Gradient Boosting Regressor*) superaram o modelo de Regressão Linear. Para a previsão de `num_entradas`, o *Gradient Boosting Regressor* obteve o melhor desempenho, com um R-quadrado de 0,9990, indicando que este modelo explica quase a totalidade da variância dos dados de entrada no período de treino.

O *Random Forest Regressor* também apresentou um R-quadrado robusto de 0,8970. Para `num_saídas`, o *Gradient Boosting Regressor* novamente se destacou com um R-quadrado de 0,9993, seguido pelo *Random Forest Regressor* com R-quadrado de 0,9084. A Regressão Linear, como esperado em cenários com relações não-lineares, teve um R-quadrado consideravelmente mais baixo, cerca de 0,67 para ambos os *targets*.

Este resultado está alinhado com as expectativas de Petropoulos et al. (2022), que afirmam que modelos de *Machine Learning* frequentemente

demonstram desempenho superior em cenários complexos com relações não-lineares, em comparação com modelos lineares, devido à sua capacidade de capturar padrões mais intrincados nos dados.

O R-quadrado extremamente alto no conjunto de treinamento para os modelos de *boosting* e *bagging* sugere uma compatibilidade com os dados históricos. Contudo, em uma avaliação preditiva real, seria crucial um conjunto de teste não visto pelo modelo para realmente confirmar a capacidade de generalização. Como não há um conjunto de teste real separado após a competência de treino, a avaliação aqui se refere mais ao poder de ajuste do modelo do que à sua capacidade de prever dados totalmente novos e futuros.

A visualização das Figuras 19, 20, 23 e 24, para os modelos de *Random Forest* e *Gradient Boosting*, demonstram que as linhas de previsão acompanham muito de perto as linhas dos valores reais. Isso confirma a alta capacidade desses modelos em aprender e reproduzir os padrões e a sazonalidade presentes nos dados históricos de treinamento. A Regressão Linear, presente nas Figuras 18 e 22, por sua vez, mostra um ajuste menos preciso, com desvios mais notáveis dos valores reais, especialmente em picos e vales.

Na etapa de previsão de volumes futuros, a escolha do *Random Forest Regressor* para gerar e plotar as projeções para 2025 foi uma decisão metodológica estratégica. Embora o *Gradient Boosting Regressor* tenha demonstrado um ajuste ligeiramente superior aos dados de treinamento, o *Random Forest Regressor* é frequentemente preferido para previsões em dados não vistos devido à sua inerente robustez contra o *overfitting*, ou seja, o excesso de ajuste aos dados históricos. Esta característica decorre de sua natureza de método *ensemble* que utiliza a técnica de *bagging*, onde múltiplas árvores de decisão são construídas de forma independente em subamostras dos dados, e suas previsões são então combinadas., se alinhando com as práticas recomendadas em previsão, conforme discutido por Petropoulos et al. (2022) ao abordar a generalização e a robustez de modelos em cenários dinâmicos.

As previsões de número de entradas para os meses restantes de 2025, presente na Figura 21, geradas pelo *Random Forest Regressor*, indicam uma média mensal em torno de 296 entradas, com um pico previsto em outubro (350 entradas) e os menores volumes em agosto (271 entradas). Esta projeção mantém a

sazonalidade observada na análise descritiva, onde outubro e dezembro historicamente apresentam picos de entrada, sugerindo que o modelo capturou esse padrão.

Para o número de saídas, representado pela Figura 25, geradas pelo *Random Forest Regressor*, as previsões para 2025 mostram uma média mensal em torno de 317 saídas, com picos esperados em julho (331 saídas) e outubro (328 saídas), e o menor volume em junho (302 saídas). Essa sazonalidade também se alinha com os picos de saída observados historicamente em janeiro, julho e outubro.

A Figura 26 oferece uma visão integrada da evolução histórica das entradas de lojas e as projeções geradas pelo modelo para os meses futuros. O gráfico exibe a série histórica de entradas reais, plotada em linha azul, abrangendo o período de março de 2022 até abril de 2025. Complementarmente, a linha verde tracejada representa as entradas previstas para o restante de 2025. Uma linha pontilhada vertical cinza demarca o ponto de transição entre os dados históricos observados e o início do horizonte de previsão, após abril de 2025. A visualização combinada permite observar a extensão dos padrões históricos de entradas para o futuro, com a linha de previsão seguindo a tendência e a sazonalidade aprendidas pelo modelo.

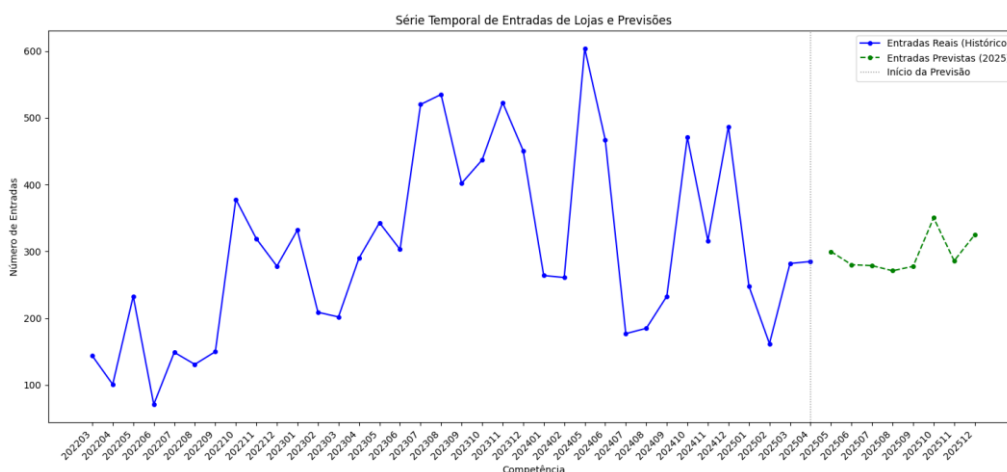


Figura 26 - Série temporal de entradas + previsão

De forma análoga, a Figura 27 apresenta a série temporal completa das saídas de lojas. A linha azul ilustra as saídas reais históricas, desde março de 2022 até abril de 2025. As projeções do modelo *Random Forest Regressor* para o restante de 2025 são representadas pela linha vermelha tracejada, com a mesma linha

pontilhada cinza indicando o início do período de previsão. Este gráfico é crucial para compreender a dinâmica histórica de saídas e como o modelo projeta sua continuidade, incluindo picos e vales sazonais, no futuro próximo.

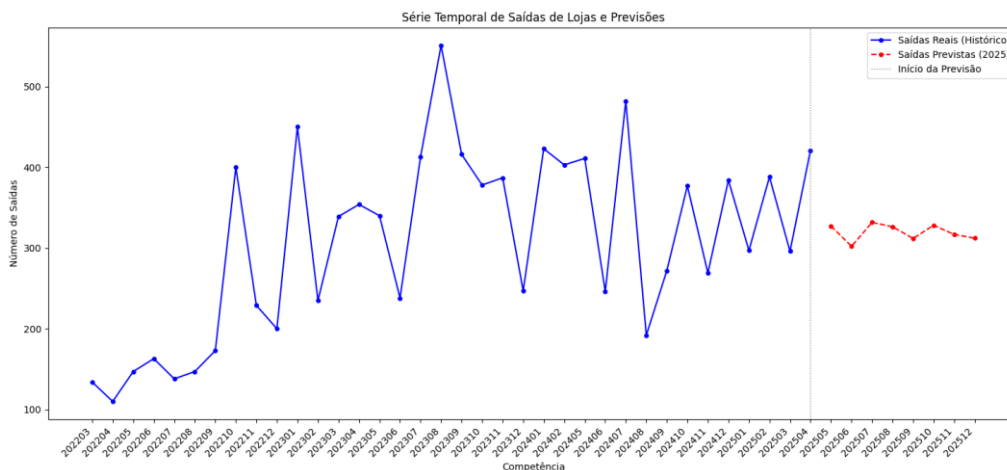


Figura 27 - Série temporal de saídas + previsão

A análise da importância das *features* é crucial para entender quais variáveis contribuem mais para a acurácia das previsões. Petropoulos et al. (2022) valorizam a interpretabilidade dos modelos como um aspecto que vai além da mera acurácia, auxiliando na tomada de decisão gerencial. É provável que as *features* de lag (e.g., *num_entradas_lag1*, *num_saidas_lag1*) e as *features* temporais (mes) se mostrem como as mais importantes, pois a movimentação de lojas é uma série temporal com forte dependência de seus valores passados e de padrões sazonais. Outras *features* categóricas de contexto (ramo, administradora, localização) podem ter importância secundária, mas significativa, validando os insights da análise diagnóstica sobre sua influência na rotatividade.

5. Conclusão

Este trabalho teve como objetivo central examinar a dinâmica da concorrência em shopping centers brasileiros, investigando os principais fatores que influenciam a movimentação de lojas com base em séries temporais históricas. A pesquisa utilizou uma base de dados privada de shoppings concorrentes, abrangendo um período de março de 2022 a abril de 2025, o que permitiu uma análise representativa do cenário pós-pandemia do varejo. Para isso, empregou-se uma abordagem metodológica tripartite, compreendendo análises descritiva, diagnóstica e preditiva, com o suporte da linguagem Python e suas bibliotecas. A relevância do estudo reside em sua contribuição para o aprimoramento da gestão estratégica desses empreendimentos, otimizando o mix de lojas e mitigando os riscos de vacância em um ambiente de mercado altamente competitivo.

A análise descritiva forneceu um panorama detalhado da base de dados, revelando que a predominância do pilar VAREJO, dos ramos de Alimentação e Vestuário, na composição dos shoppings, foi consistente ao longo do período analisado. A visualização das séries temporais de entradas e saídas de lojas por competência evidenciou picos e vales marcantes, confirmando a alta atividade de troca de lojistas e a presença de sazonalidade. A concentração de lojas em poucas administradoras, como Allos, Multiplan e JCPM, foi demonstrada, reforçando a ideia de que o mercado de shoppings no Brasil possui players dominantes, característica comum em ecossistemas de varejo planejados.

A análise diagnóstica aprofundou-se nos padrões observados. Ao investigar a relação entre movimento e ramo, constatou-se que segmentos com menor representatividade na base, como "Vinhos" e "Móveis/Decoração/Cama, Mesa e Banho", apresentaram as maiores taxas de entrada e saída, sugerindo uma alta renovação nesses nichos. A análise da relação entre movimento e administradoras revelou que, embora grandes players detenham a maioria das lojas, elas também apresentam altas taxas de rotatividade, o que pode ser interpretado como uma estratégia de gestão ativa para renovação de portfólio. A sazonalidade das movimentações foi confirmada com picos de entradas em maio, outubro e dezembro, e picos de saídas em janeiro e julho. Tais padrões podem ser diretamente relacionados a períodos de aquecimento comercial e encerramento de contratos semestrais, fornecendo insights valiosos para o planejamento do tenant mix. Por

fim, o impacto dos fatores geográficos demonstrou que a Região Sudeste, com o maior volume absoluto de movimentações, também é a de maior taxa de saída, apontando para um mercado desenvolvido e de alta concorrência.

A etapa de análise preditiva utilizou os modelos Random Forest Regressor e Gradient Boosting Regressor para prever o volume de entradas e saídas de lojas. O Gradient Boosting Regressor, em particular, obteve um desempenho superior nas métricas de ajuste (in-sample), explicando quase a totalidade da variância dos dados de entrada e saída no período de treinamento. As visualizações das projeções para os meses restantes de 2025 confirmaram a capacidade dos modelos de capturar os padrões sazonais históricos, projetando uma continuidade na dinâmica de movimentação de lojas. A análise da importância das features revelou que as features de lag (valores passados de entradas e saídas) e as features temporais (mês e ano) foram as mais importantes. Isso corrobora a natureza de série temporal da movimentação de lojas, com forte dependência de seus valores passados e de padrões sazonais.

Apesar dos resultados promissores, este estudo possui algumas limitações inerentes à natureza dos dados e da metodologia. Primeiramente, a ausência de um conjunto de teste *out-of-sample* real para avaliação da performance preditiva dos modelos após abril de 2025 impede uma validação mais robusta da capacidade de generalização para dados totalmente novos, sendo a validação da acurácia baseada na validação cruzada no período de treinamento. Além disso, a granularidade dos dados sobre o porte dos shoppings não permitiu uma investigação aprofundada da hipótese de que shoppings maiores possuem menor rotatividade. A análise também não incorporou explicitamente fatores econômicos externos mais amplos, como inflação ou taxas de juros, que podem impactar a dinâmica do varejo, conforme já discutido na literatura (Almeida Júnior, 2020; Abrasce, 2025).

Para pesquisas futuras, sugere-se a expansão da base de dados com informações adicionais, como a Área Bruta Locável (ABL) dos shoppings, para uma análise mais precisa da relação entre porte e rotatividade. A inclusão de dados macroeconômicos e indicadores de consumo permitiria uma modelagem preditiva mais abrangente e a identificação de correlações mais complexas. Adicionalmente, a aplicação de técnicas de modelagem de séries temporais mais avançadas, como as redes neurais recorrentes (RNNs), seria um caminho promissor. Por fim,

recomenda-se o uso de métodos de interpretabilidade de modelos avançados, como SHAP (*SHapley Additive exPlanations*), para aprofundar o entendimento sobre a contribuição individual de cada *feature* nas previsões. A criação de um sistema de alerta de vacância, baseado nas previsões de saída de lojas, pode se tornar uma ferramenta prática valiosa para as administradoras de shopping centers, auxiliando na proatividade e no planejamento estratégico.

6. Referências Bibliográficas

ABRAHAM, B.; BOX, G. E. P. Bayesian analysis of some outlier problems in time series. **Biometrika**, [s.l.], v. 66, n. 2, p. 229–236, 1979.

ABRAHAM, B.; CHUANG, A. Outlier detection and time series modeling. **Technometrics**, [s.l.], v. 31, n. 2, p. 241–248, 1989.

ABRASCE. **Censo brasileiro de shopping centers 2023/2024**. Abrasce - Associação Brasileira de Shopping Centers, Livraria da Associação Brasileira de Shopping Centers, 2025. Disponível em: <https://loja.abrasce.com.br/produto/censo-brasileiro-de-shopping-centers-2024/>. Acesso em: 6 maio 2025.

ALBON, C. **Python machine learning cookbook**. O'Reilly UK Ltd, 2018.

ALMEIDA JÚNIOR, Milton Fernandes de. **Tenant mix**: um estudo de caso de um shopping regional na Amazônia. 2020. 85 f. Dissertação (Mestrado em Gestão de Serviços) - Universidade do Porto, Faculdade de Economia, Porto, 2020. Disponível em: <https://repositorio-aberto.up.pt/bitstream/10216/130385/2/431617.pdf>. Acesso em: 6 maio 2025.

ANDERSON, V. O.; NOCHNAIS, U. The elimination of spurious correlation due to position in time or space. **Biometrika**, [s.l.], v. 10, n. 2/3, p. 269–279, 1914.

ATKINSON, A. C.; RIANI, M.; CORBELLINI, A. The Box-Cox transformation: Review and extensions. **Statistical Science**, [s.l.], v. 36, n. 2, p. 239–255, 2021.

BALKE, N. S. Detecting level shifts in time series. **Journal of Business & Economic Statistics**, [s.l.], v. 11, n. 1, p. 81–92, 1993.

BASS, F.; KRISHNAN, T.; JAIN, D. Why the bass model fits without decision variables. **Marketing Science**, [s.l.], v. 13, p. 203–223, 1994.

BENJAMIN, Adailton Medeiros; OLIVEIRA, Raquel da Silva. **Os indicadores de desempenho utilizados pela gestão de shopping centers e sua influência na tomada de decisão estratégica**. In: XVII Congresso USP de Iniciação Científica em Contabilidade. "A contabilidade como mecanismo de governança". São Paulo, 29 to 31 july 2020. Disponível em: <https://congressousp.fipecafi.org/anais/20UspInternational/ArtigosDownload/2717.pdf>. Acesso em: 6 maio 2025.

BEYAZTAS, U.; SHANG, H. L. Forecasting functional time series using weighted likelihood methodology. **Journal of Statistical Computation and Simulation**, [s.l.], v. 89, n. 16, p. 3046–3060, 2019.

BIANCO, A. M.; GARCÍA BEN, M.; MARTÍNEZ, E. J.; YOHAI, V. J. Outlier detection in regression models with ARIMA errors using robust estimates. **Journal of Forecasting**, [s.l.], v. 20, n. 8, p. 565–579, 2001.

BOX, G. E. P.; COX, D. R. An analysis of transformations. *Journal of the Royal Statistical Society. Series B. Statistical Methodology*, [s.l.], v. 26, n. 2, p. 211–243, 1964.

BROWELL, Jethro. Exogenous variables and feature engineering. In: PETROPOULOS, Fotios et al. (Eds.). *Forecasting: theory and practice. International Journal of Forecasting*, [s.l.], v. 38, n. 3, p. 705–871, jul./sep. 2022. DOI: 10.1016/j.ijforecast.2021.11.001. pp. 714-715.

BURRIDGE, P.; ROBERT TAYLOR, A. Additive outlier detection via extreme-value theory. *Journal of Time Series Analysis*, [s.l.], v. 27, n. 5, p. 685–701, 2006.

BUYS-BALLOT, C. H. D. **Les changements périodiques de temperature**. Utrecht: Kemink Et Fils, 1847.

CABRAL FILHO, Djalma Alves. **Gestão logística e tendências da logística 4.0**. Ponta Grossa: Atena, 2023. 121 p. ISBN: 978-65-258-1460-5. DOI: 10.22533/at.ed.605232405.

CATALÁN, B.; TRÍVEZ, F. J. Forecasting volatility in GARCH models with additive outliers. *Quantitative Finance*, v. 7, n. 6, p. 591–596, 2007.

CHANDOLA, V.; BANERJEE, A.; KUMAR, V. (2007). Outlier detection: A survey. *ACM Computing Surveys*, [s.l.], v. 14, p. 15, 2007.

CHANDOLA, V.; BANERJEE, A.; KUMAR, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys*, [s.l.], v. 41, n. 3, p. 1–58, 2009.

CHEN, C.; LIU, L.-M. (1993a). Forecasting time series with outliers. *Journal of Forecasting*, [s.l.], v. 12, n. 1, p. 13–35, 1993.

CHEN, C.; LIU, L.-M. (1993b). Joint estimation of model parameters and outlier effects in time series. *Journal of the American Statistical Association*, [s.l.], v. 88, n. 421, p. 284–297, 1993.

CHENG, G.; YANG, Y. Forecast combination with outlier protection. *International Journal of Forecasting*, [s.l.], v. 31, n. 2, p. 223–237, 2015.

CHIGBU, Uchendu Eugene; ATIKU, Sulaiman Olusegun; PLESSIS, Cherley C. Du. The science of literature reviews: searching, identifying, selecting, and synthesising. *Publications*, [s.l.], v. 11, n. 1, jan. 2023. DOI: 10.3390/publications11010002.

CLEVELAND, R. B.; CLEVELAND, W. S.; MCRAE, J. E.; TERPENNING, I. STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, [s.l.], v. 6, n. 1, p. 3–73, 1990.

COMMANDEUR, J. J. F.; KOOPMAN, S. J.; OOMS, M. Statistical software for state space methods. *Journal of Statistical Software*, [s.l.], v. 41, n. 1, p. 1–18, 2011.

COPELAND, M. T. Statistical indices of business conditions. **Quarterly Journal of Economics**, [s.l.], v. 29, n. 3, p. 522–562, 1915.

DAGUM, E. B. **The X11ARIMA/88 seasonal adjustment method: foundations and user's manual**. Statistics Canada, Time Series Research and Analysis Division, 1988.

DE GOOIJER, J. G.; HYNDMAN, R. J. 25 years of time series forecasting. **International Journal of Forecasting**, [s.l.], v. 22, p. 443–473, 2006.

DE LIVERA, A. M.; HYNDMAN, R. J.; SNYDER, R. D. Forecasting time series with complex seasonal patterns using exponential smoothing. **Journal of the American Statistical Association**, [s.l.], v. 106, n. 496, p. 1513–1527, 2011.

DIVITIIS, Marte de. **Abrasce divulga censo brasileiro de shopping centers 2024-2025**. FashionUnited, 12 fev. 2025. Disponível em: <https://fashionunited.com.br/news/business/abrasce-divulga-censo-brasileiro-de-shopping-centers-2024-2025/20250212510807>. Acesso em: 6 maio 2025.

DOKUMENTOV, A. **Smoothing, decomposition and forecasting of multidimensional and functional time series using regularisation**. Tese (Doutorado) – Monash University, 2017.

DOKUMENTOV, A.; HYNDMAN, R. J. **STR decomposition**. R package version 0.4, 2018. Disponível em: <https://cran.r-project.org/package=stR>(<https://cran.r-project.org/package=stR>). Acesso em: 6 maio 2025.

FERRARI, Rossella. Writing narrative style literature reviews. **Medical Writing**, [s.l.], v. 24, n. 4, pp. 230-235, dec. 2015. DOI: 10.1179/2047480615Z.000000000329.

FINDLEY, D. F. Some recent developments and directions in seasonal adjustment. **Journal of Official Statistics**, [s.l.], v. 21, n. 2, p. 343, 2005.

FINDLEY, D. F.; MONSELL, B. C.; BELL, W. R.; OTTO, M. C.; CHEN, B.-C. New capabilities and methods of the X-12-ARIMA seasonal-adjustment program. **Journal of Business & Economic Statistics**, [s.l.], v. 16, n. 2, p. 127–152, 1998.

FIORUCCI, J. A.; PELLEGRINI, T. R.; LOUZADA, F.; PETROPOULOS, F.; KOEHLER, A. B. Models for optimising the theta method and their relationship to state space models. **International Journal of Forecasting**, [s.l.], v. 32, n. 4, p. 1151–1161, 2016.

FOX, A. J. Outliers in time series. **Journal of the Royal Statistical Society. Series B. Statistical Methodology**, [s.l.], v. 34, n. 3, p. 350–363, 1972.

FRANSES, P. H.; GHIJSELS, H. Additive outliers, GARCH and forecasting volatility. **International Journal of Forecasting**, [s.l.], v. 15, n. 1, p. 1–9, 1999.

FULCHER, B. D.; JONES, N. S. Highly comparative feature-based time-series classification. **IEEE Transactions on Knowledge and Data Engineering**, [s.l.], v. 26, n. 12, p. 3026–3037, 2014.

GELPER, S.; FRIED, R.; CROUX, C. Robust forecasting with exponential and Holt-Winters smoothing. **Journal of Forecasting**, [s.l.], v. 11, 2009.

GIL, Antonio Carlos. **Como fazer pesquisa qualitativa**. Grupo GEN, Atlas, 2021. 190 p. ISBN-10: 6559770478. ISBN-13: 978-6559770472.

GRANGER, C. W. J.; NEWBOLD, P. Forecasting transformed series. **Journal of the Royal Statistical Society. Series B. Statistical Methodology**, [s.l.], v. 38, n. 2, p. 189–203, 1976.

GUERRERO, V. M. Time-series analysis supported by power transformations. **Journal of Forecasting**, [s.l.], v. 12, n. 1, p. 37–48, 1993.

GUPTA, M.; GAO, J.; AGGARWAL, C. C.; HAN, J. Outlier detection for temporal data: A survey. **IEEE Transactions on Knowledge and Data Engineering**, [s.l.], v. 26, n. 9, p. 2250–2267, 2013.

HAND, D. J. Mining the past to determine the future - problems and possibilities. **International Journal of Forecasting**, [s.l.], v. 25, n. 3, p. 441–451, 2009.

HARVEY, A. C. **Forecasting, structural time series models and the Kalman filter**. Cambridge: Cambridge University Press, 1990.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The elements of statistical learning**. Berlin: Springer-Verlag GmbH, 2009.

HOOKE, R. H. The suspension of the Berlin produce exchange and its effect upon corn prices. **Journal of the Royal Statistical Society**, [s.l.], v. 64, n. 4, p. 574–613, 1901.

HOSSIN, M.; SULAIMAN, M. A review on evaluation metrics for data classification evaluations. **International Journal of Data Mining & Knowledge Management Process**, [s.l.], v. 5, n. 2, p. 1–11, 2015.

HYNDMAN, Rob J.; ATHANASOPOULOS, George. **Forecasting: Principles and Practice**. 3. ed. Melbourne: OTexts, 2021. Disponível em: <https://otexts.com/fpp3/>. Acesso em: 23 de outubro de 2025.

KOEHLER, Anne B. Introduction to forecasting theory. In: PETROPOULOS, Fotios et al. (Eds.). Forecasting: theory and practice. **International Journal of Forecasting**, [s.l.], v. 38, n. 3, p. 705–711, jul./sep. 2022. DOI: 10.1016/j.ijforecast.2021.11.001. pp. 710–711.

KOTLER, Philip; KELLER, Kevin Lane. **Administração de marketing**. 15. ed. Pearson, 2019. 896 p. ISBN-10: 8543024951. ISBN-13: 978-8543024950.

KUHN, M.; JOHNSON, K. **Feature engineering and selection**. Boca Raton: Taylor & Francis Ltd, 2019.

LADIRAY, D.; QUENNEVILLE, B. **Lecture notes in statistics 158**, Seasonal adjustment with the X-11 method. New York: Springer, 2001.

LEDOLTER, J. (1989). The effect of additive outliers on the forecasts from ARIMA models. **International Journal of Forecasting**, [s.l.], v. 5, n. 2, p. 231–240, 1989.

LEDOLTER, J. (1991). **Outliers in time series analysis**: some comments on their impact and their detection. Image, 1991.

LEIGH, C. et al. A framework for automated anomaly detection in high frequency water-quality data from in situ sensors. **Science of the Total Environment**, [s.l.], v. 664, p. 885–898, 2019.

LUO, J.; HONG, T.; FANG, S.-C. Benchmarking robustness of load forecasting models under data integrity attacks. **International Journal of Forecasting**, [s.l.], v. 34, n. 1, p. 89–104, 2018a.

LUO, J.; HONG, T.; YUE, M. Real-time anomaly detection for very short-term load forecasting. **Journal of Modern Power Systems and Clean Energy**, [s.l.], v. 6, n. 2, p. 235–243, 2018.

MACAULAY, F. R. **The smoothing of time series**. In: NBER Books. National Bureau of Economic Research, Inc, 1931.

MARASCHIN, Clarice; AVOZANI, Waleska Andrzejewski. **Dinâmica do perfil populacional no entorno de shopping centers**. In: VII CINCCI - Colóquio Internacional sobre Comércio e Cidade. Fortaleza, 03 a 07 de Novembro de 2020. Disponível em: <https://lume.ufrgs.br/bitstream/handle/10183/216682/001119770.pdf>. Acesso em: 6 maio 2025.

MARCZAK, M.; PROIETTI, T. Outlier detection in structural time series models: The indicator saturation approach. **International Journal of Forecasting**, [s.l.], v. 32, n. 1, p. 180–202, 2016.

MARTINS, Maria de Fátima M. **Estudos de revisão de literatura**. Fio Cruz, Rio de Janeiro, set. 2018, 37 p. Disponível em: https://www.arca.fiocruz.br/bitstream/handle/icict/29213/Estudos_revisao.pdf;jsessionid=3E383C20F104CD577228383DCC6BA054. Acesso em: 6 maio 2025.

MONSELL, B.; ASTON, J.; KOOPMAN, S. **Toward x-13?** In: PROCEEDINGS OF THE AMERICAN STATISTICAL ASSOCIATION, SECTION ON BUSINESS AND ECONOMIC STATISTICS, 2003. p. 1–8. U.S. Census Bureau.

MORIGI, Josimari de Brito. Shopping centers como espaços de consumo e de lazer: o caso do Maringá Park e do Catuaí Shopping de Maringá, Paraná, Brasil Shopping centers as spaces of consumption and leisure: the case Maringá Park and Shopping

Catuaí Maringá, Paraná, Brazil. **Observatorium: Revista Eletrônica de Geografia**, [s.l.], v. 9 n. 1, p. 85-111, jan./abr. 2018. DOI: 10.14393/OREG-v9-n1-2018-6.

OLIVER, R. L.; SWAN, J. E. Consumer perceptions of interpersonal equity and satisfaction in transaction: a field survey approach. **Journal of Marketing**, Chicago, v. 53, n. 2, p. 21-35, 1989.

PATAH, Rodrigo; ABEL, Carol. **O que é pesquisa exploratória?** Veja como obter insights e ideias com ela. MindMiners, Blog, fev. 2022. Disponível em: <https://mindminers.com/blog/o-que-e-pesquisa-exploratoria/>. Acesso em: 6 maio 2025.

PARENTE, Juracy; BARKI, Edgard. **Varejo no Brasil: gestão e estratégia**. 2. ed. São Paulo: Atlas, 2014.

PARK, B. J. An outlier robust GARCH model and forecasting volatility of exchange rate returns. **Handbook of E: Journal of Forecasting**, [s.l.], v. 21, n. 5, p. 381–393, 2002.

PETROPOULOS, Fotios et al. Forecasting: theory and practice. **International Journal of Forecasting**, [s.l.], v. 38, n. 3, p. 705–871, jul./sep. 2022. DOI: 10.1016/j.ijforecast.2021.11.001.

PORTUGALI, Juval. **Handbook on cities and complexity**. Edward Elgar Pub, 2023. 456 p. ISBN-10: 103532525X. ISBN-13: 978-1035325252.

POYNTING, J. H. A comparison of the fluctuations in the price of wheat and in the cotton and silk imports into great britain. **Journal of the Statistical Society of London**, [s.l.], v. 47, n. 1, p. 34-74, 1884.

RANAWANA, R.; PALADE, V. **Optimized precision-a new measure for classifier performance evaluation**. In: IEEE INTERNATIONAL CONFERENCE ON EVOLUTIONARY COMPUTATION, 2006. p. 2254–2261. IEEE.

SAKATA, S.; WHITE, H. High breakdown point conditional dispersion estimation with application to S&P 500 daily returns volatility. **Econometrica**, [s.l.], v. 66, n. 3, p. 529–568, 1998.

SAKIA, R. M. The box-cox transformation technique: A review. **Journal of the Royal Statistical Society: Series D (the Statistician)**, [s.l.], v. 41, n. 2, p. 169–178, 1992.

SALES, Nathália Oliveira; REIS, Túlio Baita dos. Estudo do comércio varejista realizado nos shopping centers em Campos dos Goytacazes: uma análise sobre os critérios da Abrasce. **Revista Perspectiva Online: Humanas & Sociais Aplicadas**, [s.l.], v. 7, n. 20, p. 43-61, dez. 2017. ISSN: 2236-8876 (Online). DOI: 10.25242/887672020171113.

SANTOS, Givaldo Guilherme dos. **Lojas em shopping center**. 2013. 154 f. Dissertação (Mestrado em Administração das Micro e Pequenas Empresas) - Faculdade Campo Limpo Paulista, Campo Limpo Paulista, 2013. Disponível em: https://unifaccamp.edu.br/mestrado/administracao/arquivo/Documentos/producao_discente/givaldo_guilherme_dos_santos.pdf. Acesso em: 6 maio 2025.

SCHULZ, Daniel. **Shopping centers planning & design**. Design Media Pub Ltd, 2014. 331 p. ISBN-10: 9881296765. ISBN-13: 978-9881296764.

SHISHKIN, J.; YOUNG, A. H.; MUSGRAVE, J. C. **The x-11 variant of the census II method seasonal adjustment program**. Washington: Bureau of the Census, US Department of Commerce, 1967.

SILVA, Carlos Henrique Costa da; LEME, Jeferson Ricardo Jampietri; SANTOS, Joe Andrew Mateus. Cidade, comércio, consumo e lazer: uma análise a partir dos shopping centers. **Estudos Geográficos**, Rio Claro, v. 16, n. 2, p. 87-105, jul./dez. 2018. ISSN: 1678-698X. DOI: 10.5016/estgeo.v16i2.13163.

SILVA, Renata Rodrigues da. **Centros comerciais e shopping centers: transformações no espaço urbano de Uberlândia (MG)**. 2012. 182 f. Dissertação (Mestrado em Geografia) - Universidade Federal de Uberlândia, Uberlândia, 2012. Disponível em: <https://repositorio.ufu.br/bitstream/123456789/16154/1/CentrosComerciaisShopping.pdf>. Acesso em: 6 maio 2025.

SOBHANI, M.; HONG, T.; MARTIN, C. Temperature anomaly detection for electric load forecasting. **International Journal of Forecasting**, [s.l.], v. 36, n. 2, p. 324–333, 2020.

SPENCER, J. On the graduation of the rates of sickness and mortality presented by the experience of the Manchester Unity of Oddfellows during the period 1893–97. **Journal of the Institute of Actuaries**, [s.l.], v. 38, n. 4, p. 334–343, 1904.

STOCK, J. H.; WATSON, M. W. Forecasting using principal components from a large number of predictors. **Journal of the American Statistical Association**, [s.l.], v. 97, n. 460, p. 1167-1179, 2002.

TALAGALA, P. D.; HYNDMAN, R. J.; LEIGH, C.; MENGERSSEN, K.; SMITH-MILES, K. A feature-based procedure for detecting technical outliers in water-quality data from in situ sensors. **Water Resources Research**, [s.l.], v. 55, n. 11, p. 8547–8568, 2019.

TALAGALA, P. D.; HYNDMAN, R. J.; SMITH-MILES, K.; KANDANAARACHCHI, S.; MUNOZ, M. A. Anomaly detection in streaming nonstationary temporal data. **Journal of Computational and Graphical Statistics**, [s.l.], v. 29, n. 1, p. 13–27, 2020.

TSAY, R. S. Time series model specification in the presence of outliers. **Journal of the American Statistical Association**, [s.l.], v. 81, n. 393, p. 132–141, 1986.

WILLIAMS, John. Shopping center. **Estratégias e gestão**. Gouvêa de Souza, Abrasce, 2009. 192 p. ISBN-10: 8560949054. ISBN-13: 978-8560949052.

YIN, Robert K. Case Study Research: Design and Methods. 4. ed. California: SAGE Publications , 2009. 219 p. ISBN 9781412960991.