

ANÁLISE DE PREÇOS HEDÔNICOS NO SETOR IMOBILIÁRIO DE RESIDÊNCIAS NA CIDADE DE LONDRINA – PR

Thiago José Pinheiro¹
Amanda Thais Rocha²

RESUMO:

O preço de imóvel é definido como preço hedônico, composto de características que avaliam, além das propriedades do próprio bem, como número de quartos, vagas na garagem, entre outros além das características de localização desse imóvel. Este trabalho tem como objetivo mensurar a valorização dos preços de casas a venda na cidade de Londrina-PR, baseado na Teoria dos Preços Hedônicos e métodos e técnicas da econometria espacial como a Análise Exploratória de Dados Espaciais (AEDE) e o modelo de regressão espacial SAR, a base de dados apresenta 655 anúncios de residências à venda, coletadas de janeiro a fevereiro de 2021, o modelo verifica quais atributos que mais impactam no preço final dos imóveis. Através da AEDE constatou-se uma dependência espacial com correlação positiva para o preço das casas na cidade de Londrina-PR, o modelo SAR constatou que a quantidade de quartos apresentou o maior valor no preço final dos imóveis sendo de 20,44%, seguidas de um aumento de 10,03% a cada unidade adicional de banheiro, 7,93% para a vagas de garagem e de 0,16% para cada m² adicional na residência, impactando todas num aumento do preço final.

Palavras-chaves: Preços de casas em Londrina; Modelo de Preços Hedônicos; AEDE; Modelo de Regressão Espacial SAR.

ANALYSIS OF HEDONICAL PRICES IN THE RESIDENCE REAL ESTATE SECTOR IN THE CITY OF LONDRINA – PR

ABSTRACT:

The property price is defined as a hedonic price, composed of characteristics that evaluate, in addition to the properties of the property itself, such as the number of bedrooms, parking spaces, among others, in addition to the characteristics of the location where this property is located. This work aims to measure the valuation of real estate prices of houses for sale in the city of Londrina-PR, based on the Theory of Hedonic Prices and methods and techniques of spatial econometrics such as Exploratory Spatial Data Analysis (AEDE) and the SAR spatial regression model, the database presents 655 advertisements of homes for sale, collected from January to February 2021, the model checks which attributes have the greatest impact on the final price of properties. Through the AEDE, a spatial dependence was found with a positive correlation for the price of houses in the city of Londrina-PR, the SAR model found that the number of rooms had the highest value in the final price of real estate being 20.44%, followed by an increase of 10.03% for each additional bathroom unit, 7.93% for parking spaces and 0.16% for each additional m² in the residence, all impacting an increase in the final price.

Keywords: House prices in Londrina; Hedonic Price Model; Exploratory Spatial Data Analysis (ESDA); Spatial Autoregressive Model (SAR).

Data da submissão: 04-07-2023

Data do aceite: 31-10-2023

INTRODUÇÃO

Ao se deparar com a decisão de adquirir uma casa, os indivíduos levam em conta diversos fatores exógenos que podem afetar suas qualidades de vida, afinal o gasto com a habitação representa um gasto elevado dentro do orçamento familiar, principalmente se for o caso de um financiamento. Os imóveis são bens bastante heterogêneos, compostos por propriedades internas como quartos, salas, banheiros, quartos etc.,

¹ Mestrado em Economia Aplicada, Universidade Estadual de Ponta Grossa – UEPG. Contato: e-mail: pinheiro.tj@hotmail.com

² Mestranda no Programa profissional de pós-graduação em Economia da Universidade Federal do Paraná (UFPR).

mas também são compostos por características não mensuráveis, como localização, se estão localizados em bairro nobre, se ficam próximos de escolas, hospitais, comércios, entre outros fatores que impactam no valor do imóvel.

Portanto, os imóveis são bens decisivos para as famílias, pois eles representam uma elevada parcela dos gastos em seus orçamentos e, em muitos casos, representam boa parte de suas riquezas, conforme abordagem de Sheppard (1999). Ligados a este fato, os imóveis são importantes para a movimentação da atividade econômica em nível agregado, levando economistas a dispenderem seus esforços para entender melhor a estrutura de demanda e equilíbrio desse mercado.

A cidade de Londrina, localizada no norte do estado do Paraná, tem apresentado crescimento no setor imobiliário ao longo dos últimos anos. A revista EXAME publicou um estudo considerando a cidade paranaense como a 18ª melhor cidade para se investir em imóveis no Brasil, estando mais bem colocada dentre todas as cidades do Paraná, o estudo considerou apenas cidades com menos de 1 milhão de habitantes. Outros fatores como programas sociais, como o Casa Verde e Amarela e as universidades da cidade que participam indiretamente do mercado imobiliário através de moradia, transportes e serviços que são demandados por professores, alunos e funcionários, que muitas vezes provém de outras cidades.

Segundo reportagem da FOLHA DE LONDRINA (2020), a cidade vive hoje o fenômeno de verticalização, já conhecido em outras cidades como Curitiba. A cidade é a sexta colocada no ranking nacional em número de edifícios acima de 12 andares, considerando prédios/habitantes. Londrina também é destaque na qualidade dos imóveis disponíveis. As construções apresentadas são cada vez mais tecnológicas e sustentáveis. São diferenciais como acessibilidade, reaproveitamento de água da chuva, coleta de resíduos, geração de energia, características que até então estavam restritas a imóveis de luxo.

O objetivo do método de preços hedônicos é analisar quais as variáveis impactam no preço final das residências para uma determinada região de interesse. Nesse estudo analisaremos o mercado imobiliário de Londrina, com o objetivo principal de medir quais são as variáveis que impactam no preço das casas, como critério as residências domiciliares (do tipo casas) tanto como novas ou usadas, utilizando como embasamento teórico a Teoria dos Preços Hedônicos. A análise dos imóveis através da Teoria dos Preços hedônicos contribui para verificar quais as variáveis que influenciam os preços dos imóveis de Londrina, além de contribuir para a literatura, o estudo permite aos indivíduos verificar qual região da cidade possuem os imóveis mais caros.

Para integrar esse estudo, foram coletados através da internet mais de 655 anúncios de residências à venda, formando assim a base de dados para o estudo esperado. A metodologia empregada para estimar o modelo de preços hedônicos será o de mínimos quadrados ordinários MQO e o modelo de Regressão Espacial SAR.

Este trabalho está dividido em seis seções. A primeira apresenta uma breve introdução do trabalho. Na segunda seção, apresenta-se a revisão de literatura abordando autores que fizeram estudos analisando o mercado imobiliário utilizando a Teoria dos Preços Hedônicos. Na terceira seção encontra-se a metodologia utilizada para a elaboração do trabalho com a modelagem teórica, as variáveis que serão utilizadas e o modelo a ser empregado. A quarta seção aborda a descrição da base e a forma que os dados foram obtidos. Os resultados encontrados deste trabalho estão na quinta seção. E para finalizar a pesquisa são apresentadas as considerações obtidas através desse estudo na sexta seção .

2. REVISÃO DE LITERATURA

2.1 TEORIA DOS PREÇOS HEDÔNICOS

Cada vez mais nos deparamos com a realidade de um ambiente composto de recursos mais escassos, o que faz com que busquemos a maximização da utilidade. Sendo assim, alguns economistas buscaram desenvolver uma teoria para captar os atributos que se encontram nas decisões de compras dos consumidores, mas tais atributos não são possíveis de serem obtidos através de valores monetários, por exemplo, a satisfação do indivíduo ao adquirir um determinado imóvel próximo a uma universidade, ou de morar próximo a parques ou em lugares que apresentam baixos índices de criminalidade. Dessa forma, a Teoria dos Preços Hedônicos busca estimar os preços marginais desses atributos, que estão associados ao bem heterogêneo, de acordo com

Hermann e Haddad (2005 apud JOHN, p. 06, 2014) “o preço marginal dos atributos é estimado através da regressão do preço do bem heterogêneo em relação ao número de atributos associados”.

O pioneiro nas pesquisas voltadas para a área de preços hedônicos foi Kelvin J. Lancaster em 1966, desde então esse método vem sendo muito utilizado nas análises de bens heterogêneos, principalmente bens imobiliários. A metodologia de preços hedônicos se tornou muito atrativa, porque diferente de outras abordagens ela pode ser aplicada na avaliação do custo de atributos para os quais não existe um mercado explícito, como no caso dos bens públicos. Já no setor da demanda tal metodologia coopera para medir a disposição a pagar dos indivíduos dado os atributos internos das residências, como o número de quartos, banheiros, garagens etc., e das amenidades, como índice de criminalidade, proximidade do imóvel a serviços públicos, áreas de lazer etc.

Os preços hedônicos não revelam uma teoria composta de atributos que sejam relevantes de um bem, portanto, estima-se modelos de características físicas dos imóveis, suas amenidades e localização. Na teoria dos preços hedônicos, o preço de um imóvel está relacionado com a localidade e suas amenidades, ou seja, a maior parte do interesse é obtida através das características de onde se está localizado o imóvel, do número de escolas, hospitais, índices de criminalidade, áreas de lazer próximas etc.

Diversos são os estudos que utilizam a teoria de preços hedônicos. Um estudo publicado por Souza, Almeida, Martines e Galvão (2019) estima o diferencial do custo de vida das unidades federativas do Brasil a partir de um modelo de preços hedônicos. Foram utilizadas informações do PNAD divulgadas pelo IBGE. O modelo apresenta uma abordagem de regressão quantílica e tem como variável dependente os gastos com aluguel, considerados uma proxy para o custo de vida e variáveis relacionadas às características dos imóveis. Os resultados obtidos revelam divergências significativas nos custos de se viver nos diferentes locais.

Dantas, Magalhães e Vergolino (2007) publicaram um estudo sobre a importância dos vizinhos, analisando a cidade de Recife. O estudo reconhece a importância da questão espacial e mostra como se pode diagnosticar e incorporar os efeitos espaciais nos modelos hedônicos. Os resultados mostram que as negociações de compra dos apartamentos não ocorrem de forma independente, mas que existe uma verdadeira interação entre esses preços, de forma que uma negociação por um preço elevado tenderá a elevar o preço dos imóveis vizinhos.

Favero (2011) estudou os preços hedônicos no mercado imobiliário de São Paulo utilizando uma abordagem da modelagem multinível com classificação cruzada com a justificativa que os modelos de preços hedônicos frequentemente não consideram a natureza hierárquica com classificação cruzada existente no mercado de aluguel de escritórios. Como conclusão encontrou representatividade das características das propriedades, dos distritos e das zonas comerciais, em caráter multinível, para a diferenciação das médias dos preços de aluguel dos imóveis.

Uma das formas mais utilizadas para aplicar a teoria de preços hedônicos é a forma Log-Linear, essa forma de especificação foi utilizada por John e Porsse (2016) buscando analisar o mercado imobiliário de apartamentos em Curitiba. Os autores buscaram analisar os efeitos das amenidades urbanas e da localização sobre os preços dos imóveis. Verificaram que a cidade apresentava um modelo do tipo monocêntrico, essa configuração é dada devido à expansão territorial e a presença de amenidades que podem impactar na modificação desta estrutura, surgindo em grandes metrópoles de formatos que são multicêntricos. Ressaltamos que grande parte dos trabalhos relacionados a preços hedônicos do mercado de imóveis apresentam o formato monocêntrico.

Can (1998 apud JOHN, 2014, p. 07), destaca que “existe um grande desenvolvimento em relação à consciência da importância do desempenho da localização de um imóvel na escolha dos indivíduos, cujos comportamentos se refletem na sociedade e no setor de imóveis”. O autor apresenta as relações do impacto pelas variáveis e/ou externalidades devido a localização que são denominadas de “efeitos vizinhança”, que afetam como um todo o setor imobiliário.

2.2 REGIÃO E MERCADO IMOBILIÁRIO DE LONDRINA

A cidade de Londrina está localizada no norte do estado do Paraná, distante cerca de 381 Km da capital Curitiba, possui aproximadamente uma população de 575.377 habitantes, segundo estimativas do IBGE para

o ano de 2020. É o segundo município mais populoso do estado do Paraná, também possui o quarto maior PIB do estado, atrás apenas de Curitiba, São José dos Pinhais e Araucárias.

Um estudo realizado pelo Sinduscon e Sebrae e divulgado pelo CRECI-PR mostra que o índice de velocidade nas vendas de imóveis em Londrina é alto. As obras são entregues em média 30 meses após o lançamento, com cerca de 80% a 90% das unidades vendidas, a cidade não enfrenta dificuldades para vender as unidades que são produzidas. Entretanto, o desafio está em aumentar o número de unidades e o preço de venda, que está abaixo da média nacional. O metro quadrado das unidades em Londrina está em média 14% abaixo dos preços encontrados em Curitiba, mas os resultados mostram um equilíbrio entre a oferta e a procura de novos produtos no mercado. O diagnóstico por bairros utilizou os critérios do IPPUL – Instituto de Pesquisa e Planejamento Urbano de Londrina, que divide o município em cinco macro regiões: Norte, Sul, Leste, Oeste e Centro.

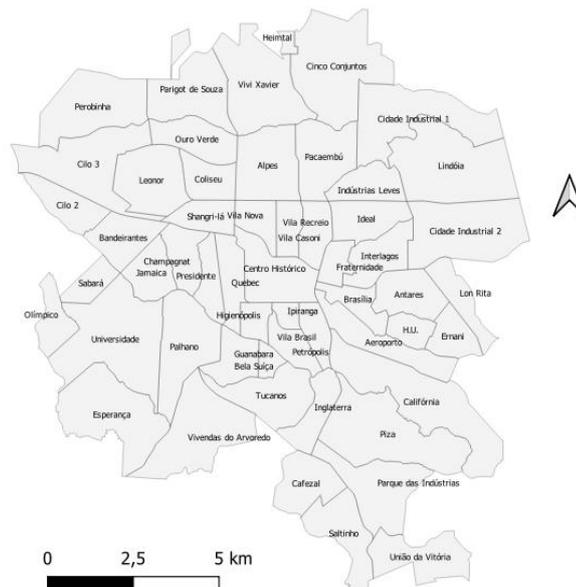


Figura 1. Bairros que compõem a cidade de Londrina – PR.

Nota: Utilizado o software Q-GIS com o arquivo shapefile da prefeitura de Londrina-PR.

Fonte: Elaborado pelos autores (2021).

De acordo com o levantamento, a região Norte concentra os bairros com população de menor poder aquisitivo e com ótimo potencial para a construção de moradias populares. Londrina possui 55 bairros e os 20 mais populosos concentram 67,58% do total da população. A pesquisa revela também que a região central é a que possui a maior concentração de renda do município, enquanto a região Norte é a que possui a menor concentração de renda.

3. METODOLOGIA

3.1 REGRESSÃO ECONOMETRICA

O termo regressão foi criado pelo inglês Francis Galton³ e mencionado em seu artigo “Family likeness in stature”⁴, no ano de 1886. Gujarati (2011) destaca que Galton (1886) verificou que apesar de existir uma tendência de que pais altos tenham filhos altos e pais baixos tenham filhos baixos, a altura média das crianças nascidas de pais com uma dada altura tendia a se mover ou “regredir” à altura média da população geral. A lei de regressão universal idealizada por Galton foi confirmada por seu amigo Karl Pearson⁵. Segundo Gujarati (2011), a análise de regressão consiste em analisar a relação de dependência de uma variável dependente em relação a uma ou mais variáveis, que são as variáveis explicativas, isso é realizado para

³ Francis Galton é antropólogo, meteorologista, matemático e estatístico inglês. Em estatística, ele formulou conceitos de regressão enquanto estudava semelhanças familiares e desenvolveu questionários e métodos de pesquisa para coleta de dados.

⁴ Artigo publicado no ano de 1886 por Francis Galton como tema “Semelhança familiar em estatura”.

⁵ Karl Pearson foi um grande contribuinte para o desenvolvimento da estatística como uma disciplina científica séria e independente.

verificar quais os efeitos que as variáveis explicativas têm sobre determinada variável dependente (também chamada e fenômeno), com isso estima-se o valor médio (da população) do fenômeno.

O Método dos Mínimos Quadrados Ordinários – MQO⁶ é atribuído ao alemão Carl Friedrich Gauss⁷. O MQO tem algumas propriedades estatísticas bastante atraentes que o tornaram um dos métodos de análise de regressão mais utilizados. Os estimadores de MQO são expressos em termos de quantidades observáveis (amostras), como X e Y. Sendo assim, podem ser facilmente calculados. São estimadores pontuais, isto é, dada a amostra, cada estimador proporciona apenas um único valor (ponto) do parâmetro populacional relevante. Depois de calculadas as estimativas de MQO para os dados amostrais, a linha de regressão amostral pode ser obtida.

3.2 MODELO DE PREÇOS HEDÔNICOS

Não há uma forma correta de se estimar os dados no modelo de preços hedônicos. Macedo (1996 apud FERREIRA NETO, 2002, p. 35) diz que “a teoria econômica ainda não desenvolveu um critério de forma funcional, a maioria dos pesquisadores vê a escolha como uma questão empírica para ser decidida de acordo com a melhor adequação aos dados”. Alguns autores descrevem que as particularidades relacionadas à localização tendem a ser mais consideráveis do que as particularidades físicas das residências na avaliação do seu valor. Cheshire e Sheppard (2004 apud ALVES et al., 2011, p.170) dizem que “o valor de qualquer imóvel varia sistemática e substancialmente com sua localização”.

Sendo assim, a satisfação dos indivíduos relaciona-se com as características da sua localidade, como proximidade às escolas, baixo índice de criminalidade, etc. Segundo Hermann e Haddad (2005, p.12), “a vizinhança interfere na qualidade do imóvel, afetando seu preço de locação”. O autor neste trabalho utilizou uma equação hedônica para estimar o preço implícito das amenidades a partir de dados de imóveis do município de São Paulo, verificou que a presença das áreas verdes, a proximidade das estações de trem e a área residencial impactam positivamente no valor do imóvel e o índice de criminalidade impacta reduzindo o seu valor.

Outros autores analisam a forma linear, dupla logarítmica e semi logarítmica. Ferreira Neto (2002) buscou verificar a disposição a pagar dos indivíduos pelas características físicas como quantidade de quartos, número de elevadores do prédio, tamanho do imóvel etc., e pela vizinhança do imóvel, número de supermercados, de escolas, índice de assaltos, entre outros fatores para cidade do Rio de Janeiro no ano de 2002, o autor chegou à conclusão de que a forma dupla logarítmica foi a que mais se ajustou ao seu trabalho.

3.3 ECONOMETRIA ESPACIAL

A econometria espacial se difere da econometria tradicional pois considera os chamados efeitos espaciais na análise, especificação, estimação, no teste de hipótese e previsão de modelos, podendo ter dados do tipo *cross-section*⁸ ou com painel de dados (ANSELIN, 1988; ANSELIN, 2003).

A econometria espacial é utilizada para reduzir os vieses dos resultados, pois leva-se em consideração que um modelo econométrico de regressão linear tradicional tem a limitação de não ser capaz de controlar efeitos espaciais. A dependência espacial é uma propriedade de funções de densidade conjunta. Sendo assim, é difícil de observar na prática. Logo, procura-se avaliar a dependência espacial através da autocorrelação espacial, a qual devemos estimar e testar. Toda a análise da dependência espacial será feita por intermédio do conceito de autocorrelação espacial que “apresenta existência de coincidência de similaridade de valores de um atributo com a similaridade de localização desse atributo” (ALMEIDA, 2012, p.36).

3.4 ANÁLISE EXPLORATÓRIA DE DADOS ESPACIAIS (AEDE)

Esse tipo de análise busca abordar os efeitos da dependência espacial e da heterogeneidade espacial. De acordo com Le Gallo e Erthur (2003, apud HOECKEL, CASAGRANDE, DOS SANTOS, 2014, p. 9):

⁶ Técnica de otimização matemática que procura encontrar o melhor ajuste para um agrupamento de dados tentando minimizar a soma dos quadrados das diferenças entre o valor estimado e os dados observados.

⁷ Matemático alemão, astrônomo e físico, que contribuiu muito em diversas áreas da ciência, dentre elas a teoria dos números, estatística, análise matemática, geometria diferencial, geofísica, eletroestática, astronomia e óptica.

⁸ Dados utilizados para designar informações de várias unidades amostrais (indivíduos, empresas, etc.) acompanhadas, em geral, ao longo do tempo.

A AEDE é um conjunto de técnicas voltadas a descrever e visualizar distribuições espaciais, localizações atípicas e/ou outliers espaciais, além de detectar padrões de associação espacial e clusters e sugerir regimes espaciais ou outras formas de heterogeneidade espacial.

Segundo Silva, Borges e Parré (2013, p.41), “o estudo de econometria espacial começa com a representação dos dados espaciais na forma matricial, em que os vizinhos de cada localidade são especificados por meio da matriz de pesos espaciais (W)”. Essa matriz é uma matriz quadrada ($n \times n$) em que os membros apresentam o grau de conexão espacial entre casas de Londrina seguindo algum critério de proximidade (tipo torre, rainha e de vizinho mais próximo).

Tal matriz tem por objetivo encontrar a melhor forma de avaliar as iterações obtidas, por exemplo, indicar na abordagem de múltiplos fenômenos as regiões que apresentam influência maior entre elas em relação a regiões não contíguas, sendo assim, regiões afastadas entre si tendem a ter uma interação mais baixa, nesse caso, onde a distância entre as regiões é importante na definição da força da interação, estabelecemos uma matriz W fundamentada na distância inversa entre as regiões para obter tal conjunto espacial da interação” (ALMEIDA; PEROBELLI; FERREIRA, 2008; p.39). Sendo assim, o elemento W_{ij} da matriz de ponderação espacial “ W ” terá valor de um quando os bairros de Londrina forem contíguos e zero caso contrário.

3.4.1 Indicador I de Moran

Almeida (2012, apud JOHN, 2014, p. 14) apresenta o Indicador I de Moran como:

$$I = \frac{n z' W z}{S_0 z' z} \quad (1)$$

Onde n representa a quantidade de amostras espaciais, z é a variável de interesse padronizada, Wz representa os valores médios da variável de importância padronizada nos vizinhos definida conforme a matriz de ponderação espacial W , já S_0 é a soma dos elementos da matriz de pesos espaciais (W).

O I de Moran apresenta valores entre 1 e -1, caso o I de Moran venha a apresentar um valor positivo e próximo de 1, a sua concentração é mais forte, já se apresentar valor negativo e próximo de -1, significa que os dados estão mais espalhados. Já os “valores de I de Moran próximos à 0 indicam falta de um modelo espacial nos dados, visto que o I de Moran tem um valor esperado de $-[1/(n-1)]$ ” (ALMEIDA; PEROBELLI; FERREIRA, 2008, p.43).

Para encontrarmos o grau de significância do I de Moran, testamos a hipótese nula $I=0$ (ausência de autocorrelação espacial) contra a hipótese alternativa $I \neq 0$ (presença de autocorrelação espacial).

No diagrama de dispersão de Moran significa, se a reta de regressão for inclinada positivamente há evidências de uma autocorrelação espacial positiva, já se a reta de regressão for negativamente inclinada há evidências de autocorrelação negativa. O diagrama oferece quadrantes onde as casas localizadas nos quadrantes Alto-Alto (AA) e Baixo-Baixo (BB) apresentam uma autocorrelação espacial positiva, ou seja, possuem valores similares e os quadrantes Baixo-Alto (BA) e Alto-Baixo (AB) apresentam autocorrelação espacial negativa, ou seja, estas regiões apresentam dissimilaridades entre os valores (ALMEIDA; PEROBELLI; FERREIRA, 2008).

Segundo Almeida (2012, p.126) “o coeficiente I_i de Moran local univariado gera uma distribuição do indicador global de autocorrelação na contribuição local de cada valor observado em quatro categorias (AA, BB, AB e BA), cada uma dessas categorias indica um quadrante no diagrama de dispersão de Moran”. Uma forma de exibir as regiões com estatísticas do I de Moran local univariado é pelo mapa de clusters *Local Indicator of Spatial Association (LISA)*, “pois ele combina a informação do diagrama de dispersão de Moran e a informação do mapa de significância das medidas de associação local I_i ”.

3.4.2 Modelo SAR

O modelo SAR (*Spatial Autoregressive Model*, tradução: Modelo autorregressivo espacial) é um modelo de defasagem espacial, de acordo com Hoeckel; Casagrande e Dos Santos (2014, p.12) “os modelos SAR utilizam a mesma ideia dos modelos AR (Auto Regressivos) em séries temporais, incorporando um termo de lag entre os repressores da equação”. Segundo Silva, Borges e Parré (2013) o SAR é representado pela seguinte fórmula:

$$Y = \rho W_1 Y + \beta_1 + \beta_2 X + \varepsilon \quad (2)$$

A $W_1 Y$ é a variável dependente defasada pela matriz de pesos espaciais, ρ é o coeficiente auto regressivo espacial que capta a influência média dos vizinhos sobre a observação no vetor y , indicando a proporção da variação total em y explicada pela dependência espacial. Para decidirmos se o modelo é o mais apropriado, avaliamos o teste do tipo Multiplicador de Lagrange (ML), onde o teste $ML\rho$ considera a hipótese nula $\rho=0$ (existe presença de autocorrelação espacial na defasagem espacial).

3.4.3 Modelo SEM

Da mesma forma que os modelos SAR partem da mesma ideia dos modelos AR, os modelos SEM (*Spatial Error Models*) partem da especificação de modelos MA (médias móveis) para observações de tempo. Os modelos SEM possuem a seguinte configuração:

$$Y = X \beta + u_i \quad (3)$$

Os resíduos da equação possuem uma estrutura auto regressiva da forma:

$$u = \lambda W u + \varepsilon \quad (4)$$

Nota-se que, diferente do modelo SAR, os modelos SEM não apresentam a variável resposta como uma função direta dos seus lag espaciais. A autocorrelação dos modelos SEM aparece nos termos de erro. Combinando as expressões (3) e (4) temos:

$$Y = X \beta + (1 - \lambda W)^{-1} \varepsilon \quad (5)$$

4. DADOS

4.1 COLETA DOS DADOS

Os dados para o estudo proposto foram coletados nos meses de janeiro e fevereiro de 2021, onde a coleta foi realizada através da pesquisa na internet no site zap imóveis⁹, onde as informações utilizadas para o parâmetro de busca constituíram primeiro pela área de análise, ou seja, a cidade de Londrina-PR; e o segundo pela unidade de corte, que são os tipos de imóveis, que para o caso desse estudo proposto são residenciais do tipo casas a venda na cidade de Londrina-PR. Com posse desses filtros de pesquisa, obteve-se as amostras necessárias, com as informações de: valor do imóvel em R\$, número de quartos, metragem da área construída em metros quadrados (m²), número de banheiros, quantidade de vagas de garagem, além do endereço completo da residência.

Com posse dos dados dos endereços completos de cada residência a venda, foi possível então realizar a coletada das informações referentes a localidade, onde então foram coletados através do Software Google Earth as coordenadas Euclidiana Universal Transversa de Mercator (UTM), tendo posse dessas informações foi possível obter as coordenadas geográficas para cada amostra de residência a venda, e então conseguir realizar a elaboração dos mapas com as distribuições das unidades a venda por seus respectivos bairros e zonas para a cidade de Londrina-PR e em seguida foi construída a camada de pontos, o arquivo para a camada do mapa foi coletado o shapefile disponível no site da prefeitura municipal de Londrina e então realizado a combinação das duas camadas de mapa de polígonos e mapa de pontos, toda a elaboração do arquivo de mapa e o arquivo shapefile foi realizada com auxílio do Software Q-GIS, utilizando a versão 3.16.3.

Foram coletadas ao todo 655 amostras de residências a venda na cidade de Londrina-PR, distribuídas pelos 55 bairros da cidade e suas 5 regiões, cada amostra se refere a uma unidade de casa a venda, a Figura 2 apresenta a distribuição das amostras das residências a venda, ao longo dos bairros e regiões da cidade de Londrina-PR, sendo cada ponto uma residência com suas coordenadas UTM, compondo assim o mapa de pontos.

⁹ Disponível em: www.zapimoveis.com.br

5. RESULTADOS

5.1 RESULTADOS DE REGRESSÃO HEDÔNICA POR MQO

Conforme exposto por Gujarati (2011), o método de Mínimos Quadrados Ordinários (MQO), no inglês conhecido como Ordinary Least Squares (OLS), consiste em encontrar os estimadores ($\beta_0, \beta_1, \dots, \beta_k$) que minimizam a soma dos quadrados dos resíduos (μ_i) da função de regressão amostral.

As variáveis representadas na sequência, portanto, demonstram que o preço dos imóveis na sua forma logarítmica natural ($\ln Y_i$) é uma função das características intrínsecas dos imóveis. A adoção da forma logarítmica da equação (1) provê a informação da variação percentual dos preços dos imóveis em relação a variação de cada variável explicativa. Em outras palavras, podemos afirmar que a utilização do modelo log-linear se traduz, conforme exposto por Gujarati (2011), numa medição de uma semi elasticidade do preço dos imóveis quando ocorrem variações dos atributos intrínsecos as residências e da variação da localização, ou seja, obtemos a semi elasticidade (ou efeito marginal) de Y com relação as variáveis dependentes.

No modelo de preços hedônicos não há uma forma correta para se estimar os dados, de acordo com FERREIRA NETO (pág. 35; 2002) apud MACEDO (1996) “a teoria econômica ainda não desenvolveu um critério de escolha da forma funcional, dessa forma a maioria dos pesquisadores vê a escolha como uma questão empírica para ser decidida pela melhor adequação aos dados.”. Assim, neste trabalho a forma log linear ou semi logarítmica será utilizada para adaptação dos dados ($\ln \text{Preço}_i = \alpha + \beta X_i + \mu_i$).

Portanto, a equação da regressão com as variáveis utilizadas neste estudo será evidenciada da seguinte maneira:

$$\ln \text{Preço}_i = \beta_0 + \beta_1 \text{Área} + \beta_2 \text{Quarto} + \beta_3 \text{Banheiro} + \beta_4 \text{Garagem} + \mu_i \quad (6)$$

Onde:

$\ln \text{Preço}_i$ = Variável Dependente do Logaritmo Neperiano do Preço do Imóvel;

β_0 = Constante;

$\beta_1 \text{Área}$ = Variável explicativa Área Construída;

$\beta_2 \text{Quarto}$ = Variável explicativa quantidade de Quartos;

$\beta_3 \text{Banheiro}$ = Variável explicativa quantidade de Banheiros;

$\beta_4 \text{Garagem}$ = Variável explicativa quantidade de Garagem;

μ_i = Termo de erro estocástico.

Obtidos os resultados do modelo estimado em MQO para a equação (6), com o auxílio do Software GeoDaSpace, obteve os seguintes coeficientes para cada variável da equação (6), que serão apresentados na Tabela 2, obteve-se para o coeficiente de determinação¹⁰ (R^2) 0,6922, ou seja, 69,22% dos preços dos imóveis podem ser explicados pelas variáveis explicativas presentes no modelo. Destaca-se também que todas as variáveis analisadas tiveram seus coeficientes estatisticamente significativos a nível de 1% (p-valor), ou seja, todas as variáveis possuem relevância estatística sobre o preço das residências analisados nesta pesquisa.

A variável explicativa que apresentou o maior valor de coeficiente foi quarto tendo um valor encontrado de 0,2261553, ou seja, uma unidade de quarto a mais na casa impacta positivamente em 22,62% ao preço dessa residência, ou seja, para cada quarto que a residência possuir o preço desse tipo de imóvel tende a aumentar em 22,62%. O impacto de aumento no preço final das residências é visto para todas as variáveis explicativas: Quarto 22,62% (0,2261553), Banheiro 10,55% (0,1055191), Garagem 9,22% (0,0922443) e Área Construída do imóvel em m² 0,17% (0,0017847).

Analisando os resultados dos testes apresentados na Tabela 2, podemos inferir que o modelo apresenta o valor de multicolinearidade¹¹ menor do que 30 e o sinal dos coeficientes estimados das variáveis são consistentes com a expectativa teórica, em princípio, não há forte evidência empírica de violação do suposto de ausência de multicolinearidade elevada, o teste de normalidade aplicado foi o Jarque-Bera¹² disponível no GeoDaSpace, a estatística do teste de Jarque-Bera, apresentou um valor de 12,962, com p-valor tendo significância estatística ao nível de 1%, resultando assim na rejeição da hipótese de normalidade dos erros da

¹⁰ Valor varia entre 0 e 1, indicando, em percentagem, o quanto o modelo consegue explicar os valores observados. Quanto maior o R^2 , mais explicativo é o modelo, melhor ele se ajusta à amostra.

¹¹ Problema em que variáveis independentes possuem relações lineares exatas ou aproximadamente exatas.

¹² Em termos estatísticos, o teste Jarque-Bera é um teste de ajuste, se os dados da amostra têm a afinidade e a curtose correspondentes à distribuição normal.

regressão por meio de MQO, ou seja, não normalidade dos resíduos, assim os testes para verificar a presença de heterocedasticidade, apresentados na Tabela 4, foram os testes de Breusch-Pagan e Koenker-Bassett, como o modelo apresentou não normalidade devemos observar o teste de Koenker-Bassett, os quais ambos apresentaram valores estatisticamente significativos, indicando a presença de heterocedasticidade. Apesar dos sinais dos coeficientes estimados por MQO serem compatíveis com o modelo teórico de preços de imóveis, eles podem estar viesados e inconsistentes devido a violação dos pressupostos do modelo de regressão linear múltipla: normalidade e homoscedasticidade de resíduos e não autocorrelação espacial dos resíduos e das variáveis dependentes. Com base nisso, houve a necessidade de corrigir a presença de dependência espacial através da estimação do modelo de preços hedônicos e buscar definir qual o melhor modelo, para corrigir falta de normalidade foi utilizado o método dos modelos generalizados (GMM), e pelo método de erros padrões robustos de White foi corrigido a presença de heterocedasticidade.

Tabela 2. Resultados da regressão em MQO.

Variável	Coefficiente	erro-padrão	t-valor	p-valor
CONSTANT	11.3639967	0.0486832	233.4273696	0.0000000
AREA	0.0017847	0.0001560	11.4383168	0.0000000
BANHEIRO	0.1055191	0.0128681	8.2000379	0.0000000
GARAGEM	0.0922443	0.0098144	9.3988478	0.0000000
QUARTO	0.2261553	0.0195414	11.5731631	0.0000000
TESTES				
Observações	655	Critério Schwarz	-181.315	
Variáveis	5	Multicolinearidade	19.584	
R2	0.6922	Jarque-Bera	12.962	0.0015
Teste F	365.3832	Breusch-Pagan	80.442	0.0000
Probabilidade (F)	1.14E-164	Koenker-Bassett	60.887	0.0000
Critério Akaike	-203.738	White	84.897	0.0000

Nota: Resultados da pesquisa com auxílio Software GeoDaSpace (2021).

Fonte: Elaborado pelo autor (2021).

5.2 RESULTADOS AEDE

5.2.1 I de Moran Global

Obtidos os valores para os indicadores de I de Moran Local e Global, tendo como base a regressão estimada por MQO, podemos descrever as informações dos resultados obtidos. Na Figura 3, a inclinação de forma positiva da reta e o valor positivo de I de Moran: 0,373 sugerem a presença de autocorrelação positiva, ou seja, residências com valor maior que a média tem ao seu redor vizinhos com valor maior que a média, e as casas com valor menor em relação à média tem ao seu redor vizinhos com média menor.

Além disso são apresentados os quatro tipos de associação linear espacial (alto-alto, baixo-baixo, alto-baixo e baixo-alto), no Gráfico 3 os valores são padronizados tanto para o eixo X como eixo Y, sendo zero o valor médio.

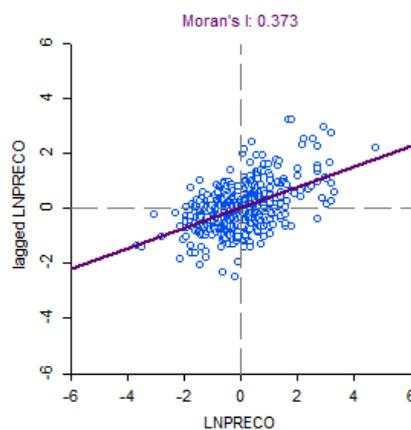


Figura 3. I De Moran Global Univariado.

Nota: Resultados da pesquisa com auxílio Software GeoDa (2021).

Fonte: Elaborado pelo autor (2021).

5.2.2 I de Moran Local

Para analisar o I de Moran Local, foi obtido o Mapa LISA apresentado na Figura 4, que auxilia na localização dos clusters, onde é possível verificar as residências com dependência espacial, onde os pontos de cor vermelha alto-alto (Q1) apresentou 55 casas que em seu entorno que impactam no aumento do preço das casas, neste grupo vermelho estão presentes as casas que possuem um valor médio acima do preço médio da cidade rodeadas por outros vizinhos com valor acima da média, os pontos de cor azul apresentam uma correlação do tipo baixo-baixo (Q3), com 41 casas que estão abaixo da média da cidade com seus vizinhos que também apresentam valores abaixo da média, também observa-se que temos os pontos com cor vermelha mais clara 8 residências associadas como alto-baixo, ou seja, residências de alto valor rodeadas por vizinhos com baixo valor além disso os pontos em azul mais claro com 6 residências são do tipo baixo-alto, nesse caso essas residências tem um valor abaixo da média rodeados por vizinhos com valor acima da média.

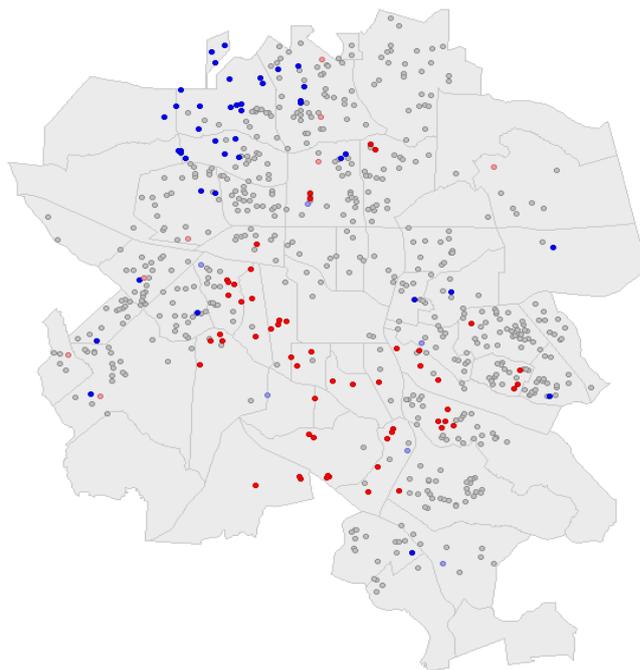


Figura 4. Mapa de Clusters do Lnpreço e Londrina Paraná em 2021.

Nota: Resultados da pesquisa com auxílio Software GeoDa (2021).

Fonte: Elaborado pelo autor (2021).

Por fim para os pontos de cor cinza é referente ao preço das casas que não tem dependência espacial estatisticamente, ou seja, nesse caso 545 casas tem o valor próximo da média da cidade de Londrina-PR.

5.3 MODELO DE PREÇO HEDÔNICOS ESPACIAL SAR

De modo geral, podemos deduzir que a amostra apresentada e a regressão em MQO obtida são significativas para o estudo proposto, porém, como foi realizado teste para verificar a presença de dependência ou autocorrelação espacial, ou seja, se o preço pode estar correlacionados com os vizinhos mais próximos, buscou então com base na AEDE, descobrir se os dados estão distribuídos aleatoriamente ou se eles seguem um determinado padrão espacial, após a realização dos testes para verificar se consta a presença de dependência espacial, por meio do Índice de Moran (I de Moran), onde foram determinados valores de 1 a 6 vizinhos como critério de escolha, a Tabela 3 retorna os valores obtidos pelo teste I de Moran, sendo H_0 : Não há dependência espacial e H_1 : Há dependência espacial. O melhor resultado obtido foi de 03 vizinhos, sendo o maior valor de I de Moran de 0,1471 e com p-valor significativo a 1%, resultando assim na rejeição da hipótese nula e aceitando a hipótese alternativa de que há dependência espacial no modelo.

Para buscar corrigir a dependência espacial positiva nos preços de casas, foi avaliado qual o melhor modelo deveria ser utilizado e considerando a matriz de pesos espaciais ao nível de 03 vizinhos mais próximos que apresentou maior I de Moran conforme apresentado anteriormente, com base nos resultados obtidos que

são apresentados na Tabela 4 foi possível verificar que o melhor modelo de regressão espacial é o modelo de Lag Espacial mais conhecido como modelo SAR¹³, o qual apresentou o valor de 64,519 sendo estatisticamente significativo ao nível de 1%, uma vez que os modelos de erro espacial (SEM) e lag espacial (SARMA), não tiveram significância estatística quando analisado ao nível de robusto.

Tabela 3. I de Moran modelo MQO para matriz de peso espacial.

Matriz	I de Moran	Value	p-Valor
1 Vizinho	0,12170	2,4840	0,0130
2 Vizinhos	0,14500	4,1240	0,0000
3 Vizinhos	0,14710	5,0610	0,0000
4 Vizinhos	0,13630	5,3960	0,0000
5 Vizinhos	0,12970	5,7080	0,0000
6 Vizinhos	0,12940	6,2210	0,0000

Nota: Resultados da pesquisa com auxílio Software GeoDaSpace (2021).

Fonte: Elaborado pelo autor (2021).

Tabela 4. Diagnóstico para dependência espacial.

Test	MI/DF	VALUE	PROB
Moran's I (error)	0,1471	5,061	0,0000
Lagrange Multiplier (lag)	1	64,519	0,0000
Robust LM (lag)	1	39,820	0,0000
Lagrange Multiplier (error)	1	24,700	0,0000
Robust LM (error)	1	0,000	0,9855
Lagrange Multiplier (SARMA)	2	64,520	0,0000

Nota: Resultados da pesquisa com auxílio Software GeoDaSpace (2021).

Fonte: Elaborado pelo autor (2021).

Atribuído a distância aos vizinhos mais próximos no nível de 3 vizinhos o modelo pode então ser descrito como:

$$\ln Preço_i = \beta_0 + \beta_1 \text{Área} + \beta_2 \text{Quarto} + \beta_3 \text{Banheiro} + \beta_4 \text{Garagem} + \beta_5 W_ \ln Preço_i + \mu_i \quad (7)$$

Em que a variável $\beta_5 W_ \ln Preço_i$ representa a dependência espacial nos preços de casas com seus 3 vizinhos mais próximos.

Na Tabela 5 reporta-se os resultados obtidos através do modelo de regressão espacial SAR, onde se pode observar que todos os valores são estatisticamente significativos ao nível de 1%, sendo que variável que capta a dependência espacial de preços de casas de seus Vizinhos mais próximos $\beta_5 W_ \ln Preço_i$, apresentou o maior valor de coeficiente 24,03% (0,243012) estatisticamente significativa ao nível de 1%, indicando que os preços das casas ao redor da residência impactam no preço final do imóvel.

Ao analisar as variáveis individualmente, verifica se que a variável com maior poder explicativo sobre o preço dos imóveis é a variável quarto, apresentando um valor de 0,2044214, ou seja, que a cada quarto incluído na residência o seu preço final tende a aumentar em cerca de 20,44%, o mesmo acontece para as demais variáveis, seguidas de um aumento de 10,03% a cada unidade adicional de banheiro na residência, acompanhados de 7,93% de acréscimo para a vagas de garagem e de 0,16% para cada m² adicional na residência.

A estatística de teste de Anselin-Kelejian apresentado na Tabela 5 permite inferir que as estimativas de preços hedônicos de casas de Londrina-PR por meio do modelo de regressão com lag defasagem espacial da variável dependente (SAR) corrigiu a presença de dependência espacial positiva nos preços das residências de Londrina-PR, onde o teste de Anselin-Kelejian mostra que os modelos SAR estimados com uma matriz de pesos espaciais de 3 vizinhos mais próximos não rejeitam a hipótese nula de ausência de dependência espacial ao nível de significância de 10% apresentando o probabilidade empírica 26,01% (0,2601), ou seja, podemos inferir que a dependência espacial foi controlada.

¹³ modelo a ser usado quando a dependência espacial emerge devido a lag ou defasagem de preços de casas de vizinhos mais próximos.

Tabela 5. Resultados da regressão do modelo SAR espacial.

Variável	Coefficiente	erro-padrão	t-valor	p-valor
CONSTANT	8,4367280	0,4224221	19,9722705	0,0000000
AREA	0,0016611	0,0002414	6,8812629	0,0000000
BANHEIRO	0,1003278	0,0148341	6,7633123	0,0000000
GARAGEM	0,0793407	0,0098133	8,0850180	0,0000000
QUARTO	0,2044214	0,0217070	9,4172849	0,0000000
W_LNPRECO	0,2403012	0,3466370	6,9323661	0,0000000
TEST	MI/DF	VALUE	PROB	
Anselin-Kelejian	1	1	0,2601	

Nota: Resultados da pesquisa com auxílio Software GeoDaSpace (2021).

Fonte: Elaborado pelos autores (2021).

6. CONSIDERAÇÕES FINAIS

O objetivo deste trabalho foi de procurar estimar os determinantes dos preços das casas em Londrina-PR, através da utilização da metodologia de preços hedônicos, e empregar conceitos e métodos da econometria espacial como a Análise Exploratória de Dados Espaciais (AEDE). Com a utilização da AEDE através do indicador de I de Moran Global Univariado, verificou-se que a cidade de Londrina-PR apresenta uma dependência espacial do tipo positiva entre o preço das casas e de seus vizinhos.

Para os resultados econométricos obtidos pelo modelo de preços hedônicos estimado pela metodologia SAR que buscou controlar a dependência espacial, os resultados mostraram que todos os coeficientes estimados do modelo de regressão espacial SAR são consistentes e as variáveis relacionadas as características físicas que apresentaram maior impacto nos preços das residências são quarto e banheiro.

Pode se concluir que os resultados obtidos foram satisfatórios para o mercado imobiliário de Londrina-PR, mostrando que a cidade conta com uma vasta quantidade de residências disponíveis a venda e os indivíduos ao adquirir alguma casa tendem a buscar residências com dois ou mais quartos com pelo menos uma ou duas vagas de garagem.

Diante de todo o exposto, analisamos os resultados satisfatórios frente ao objetivo proposto inicialmente da aplicação da teoria dos preços hedônicos e de técnicas de econometria espacial para o segmento imobiliário da cidade de Londrina-PR. No entanto, percebemos que uma grande barreira para avançar nesse tipo de estudo reside na dificuldade de se obter informações mais precisas para a formação de uma base de dados mais aprimorada, como a inclusão da idade dos imóveis, dados de amenidades urbanas mais detalhadas, maior facilidade de acesso às informações de fontes oficiais como as secretarias do governo do estado e do município, entre outros. e expandindo o estudo para as regiões vizinhas de Londrina-PR, assim como para as outras cidades.

REFERÊNCIAS

ALMEIDA; E. **Econometria Espacial Aplicada**. Editora Alínea, Campinas, 2012.

ALMEIDA, E. S., PEROBELLI, F. S. e FERREIRA, P. G. C. **Existe convergência espacial da produtividade agrícola no Brasil?**. Revista de Economia e Sociologia Rural, Rio de Janeiro, vol. 46, n. 1, 2008.

ALVES, D. C. O. et al. **Modelagem dos preços de imóveis residenciais paulistanos**. Revista Brasileira de Finanças, Rio de Janeiro, v. 9, n. 2, p.167-187, jun. 2011.

ANSELIN, L.; BERA, A. **Spatial dependence in linear regression models with an introduction to spatial econometrics**. In: Ullah A. e Giles D. E. (eds.) Handbook of applied economic statistics, Marcel Dekker, New York, p. 237-289, 1998.

ANSELIN, L; SYABRI, I; SMIRNOV, O. **Visualizing multivariate spatial correlation with dynamically linked Windows**. Mimeo, University of Illinois, 2003.

CRECI-PR. **Pesquisa mostra que construções verticais aquecem o mercado imobiliário em Londrina**. Disponível em <http://www.crecipr.gov.br/novo/news/ultimas-noticias/302-pesquisa-mostra-que-construcoes-verticais-aquecem-mercado-imobiliario-em-londrina>

DANTAS, Rubens Alves; MAGALHAES, André Matos; VERGOLINO, José Raimundo de Oliveira. **Avaliação de imóveis: a importância dos vizinhos no caso de Recife**. Econ. Apl., Ribeirão Preto, v.11, n. 2, p. 231-251, Junho 2007. Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S141380502007000200004&lng=en&nrm=iso>. Acessado em 06 fevereiro de 2021. <https://doi.org/10.1590/S1413-80502007000200004>.

EXAME, **As 100 melhores cidades brasileiras para se investir em imóveis**. Disponível em <<https://exame.com/mercado-imobiliario/as-100-melhores-cidades-do-brasil-para-investir-em-imoveis/>> Acesso em 10 de fevereiro de 2021.

FAVERO, Luiz Paulo Lopes. **Preços hedônicos no mercado imobiliário comercial de São Paulo: a abordagem da modelagem multinível com classificação cruzada**. Estud. Econ., São Paulo, v. 41, n. 4, p. 777-810, Dec. 2011. disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-41612011000400005&lng=en&nrm=iso>. Acessado em 06 de fevereiro 2021. <https://doi.org/10.1590/S0101-41612011000400005>.

FERREIRA NETO, E. **Estimação do preço hedônico: uma aplicação para o mercado imobiliário da cidade do Rio de Janeiro**. Dissertação (Mestrado) - FGV, Rio de Janeiro, 2002. 48 p.

FOLHA DE LONDRINA. **O crescimento do mercado imobiliário de Londrina**. Disponível em <<https://www.folhadelondrina.com.br/cv-folha/o-crescimento-do-mercado-imobiliario-de-londrina-2979852e.html#:~:text=O%20panorama%20geral%20do%20mercado,para%20R%24%20500%2C00>> acesso em 10 de fevereiro de 2021.

GUJARATI, D. N.; PORTER, D. C. **Econometria Básica**. 5ª edição. AMGH, 2011.

HERMANN, B. M.; HADDAD, E. A. **Mercado imobiliário e amenidades urbanas: a view through the window**. Estudos Econômicos (São Paulo), v. 35, n. 2, p. 237-269, 2005.

JOHN, E. M. C. **Análise de Preços Hedônicos no Mercado Imobiliário de Apartamentos em Curitiba**. 2014. Dissertação (Mestrado em Economia) – Setor de Ciências Sociais Aplicadas, Universidade Federal do Paraná, Curitiba, 2014.

JOHN, E. M. C.; PORSSE, A. A. **Análise de Preços Hedônicos no Mercado Imobiliário de Apartamentos em Curitiba**. Revista Paranaense de Desenvolvimento-RPD, v. 37, n. 130, p. 99-115, 2016. Disponível em: <http://www.ipardes.pr.gov.br/ojs/index.php/revistaparanaense/article/view/765> >. Acesso em 14 nov. 2017.

LESAGE, J. P. e PACE, R. K. **Introduction to Spatial Econometrics**. CRC Press, Boca Raton, 2009.

PREFEITURA MUNICIPAL DE LONDRINA, **Arquivos Shapefile**. Disponível em: <<https://www.londrina.pr.gov.br/downloads-siglon>>. Acessado em 04 de fevereiro de 2021.

SOUZA JUNIOR, M. L. DE; ALMEIDA, A. N. DE; MARTINES FILHO, J. G.; GALVÃO, M. C. **Diferenciais de custos de vida entre os estados brasileiros: uma aplicação de preços hedônicos a partir dos dados da PNAD**. Revista Brasileira de Estudos Regionais e Urbanos, v. 13, n. 2, p. 223-242, 16 out. 2019.

ZAP IMÓVEIS, **Classificados de imóveis à venda em Londrina-PR**. Disponível em: <<https://www.zapimoveis.com.br>> Acessado entre 18 de janeiro de 2021 a 03 de fevereiro de 2021.