



UFOP

Universidade Federal
de Ouro Preto

**Universidade Federal de Ouro Preto
Instituto de Ciências Exatas e Aplicadas
Departamento de Engenharia de Produção**

**Modelos de precificação e
caracterização da oferta de lotes em
uma cidade do interior de Minas
Gerais**

Roberta Pâmela Gomes de Araújo

**TRABALHO DE
CONCLUSÃO DE CURSO**

**ORIENTAÇÃO:
Paganini Barcellos de Oliveira**

**Janeiro, 2021
João Monlevade–MG**

Roberta Pâmela Gomes de Araújo

Modelos de precificação e caracterização da oferta de lotes em uma cidade do interior de Minas Gerais

Orientador: Paganini Barcellos de Oliveira

Monografia apresentada ao curso de Engenharia de Produção da Universidade Federal de Ouro Preto, como requisito parcial para obtenção de grau em Engenharia de Produção.

Universidade Federal de Ouro Preto

João Monlevade

Janeiro de 2021

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

A663m Araujo, Roberta Pamela Gomes de .

Modelos de precificação e caracterização da oferta de lotes em uma cidade do interior de Minas Gerais. [manuscrito] / Roberta Pamela Gomes de Araujo. - 2021.

48 f.: il.: color., tab..

Orientador: Prof. Dr. Paganini Barcellos de Oliveira.

Monografia (Bacharelado). Universidade Federal de Ouro Preto. Instituto de Ciências Exatas e Aplicadas. Graduação em Engenharia de Produção .

1. Análise de regressão . 2. Bens imóveis - Avaliação. 3. Bens imóveis - Preços . 4. Mercado imobiliário. 5. Modelos matemáticos . I. Oliveira, Paganini Barcellos de. II. Universidade Federal de Ouro Preto. III. Título.

CDU 332

Bibliotecário(a) Responsável: Flavia Reis - CRB6-2431



FOLHA DE APROVAÇÃO

Roberta Pâmela Gomes de Araújo

Modelos de precificação e caracterização da oferta de lotes em uma cidade do interior de Minas Gerais

Monografia apresentada ao Curso de Engenharia de Produção da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Bacharel em Engenharia de Produção

Aprovada em 29 de janeiro de 2021

Membros da banca

Doutor - Paganini Barcellos de Oliveira - Orientador - Universidade Federal de Ouro Preto (UFOP)
Doutor - Rafael Lucas Machado Pinto - Universidade Federal de Ouro Preto (UFOP)
Mestre - Diego Fernandes Pantuza Moura - Universidade Federal de Ouro Preto (UFOP)

Paganini Barcellos de Oliveira, orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 20/02/2021



Documento assinado eletronicamente por **Paganini Barcellos de Oliveira, PROFESSOR DE MAGISTERIO SUPERIOR**, em 20/02/2021, às 15:30, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0137182** e o código CRC **CFDDA906**.

Agradecimentos

Agradeço à Deus, pela vida, saúde e sabedoria. À minha mãe Marilene, a pessoa mais importante e inspiradora da minha vida, por todo amor e apoio. Agradeço aos demais familiares, pelo incentivo e amor. Um agradecimento especial ao Victor e aos amigos que fiz, por toda ajuda e carinho durante essa caminhada. Agradeço aos professores, pelos ensinamentos e atenção. Em especial ao professor Paganini, pela orientação e paciência durante o trabalho, e ao professor Diego, pela ajuda.

Resumo

O presente estudo tem como finalidade caracterizar a oferta dos terrenos comercializados no município de João Monlevade, definindo os preços médios dos terrenos, mensurando a relação entre suas características estruturais e geográficas, e comparando os preços médios dos imóveis e terrenos de mesma vizinhança (bairro). Propõe-se também a construção de modelos matemáticos de regressão, lineares e ajustados, para estimar o preço dos terrenos, e que incluam os fatores de maior relevância para caracterização dos preços. Ademais, são utilizadas técnicas de agrupamento, em função das características dos bairros, considerando os escores gerados para cada terreno. Como resultados, foram apresentados uma análise geral dos dados, uma descrição e comparação dos modelos de regressão formulados, e o agrupamento dos bairros em função do grau de similaridade.

Palavras-chaves: Mercado imobiliário, Precificação, Terrenos urbanos, Modelos de regressão.

Abstract

The present study aims to characterize the offer of urban lands in João Monlevade, defining average prices, measuring the relationship between structural and geographical characteristics, and comparing average property prices and lands in the same neighborhood. It is also proposed to develop mathematical regression models, linear and adjusted, to estimate the urban land prices, and that include the most relevant factors for their characterization. Grouping techniques based on the neighborhood characteristics are used, considering the scores generated for each land. As a result, a general data analysis, a description, and a comparison of the regression models were presented, and also the grouping of neighborhoods according to the similarity degree.

Key-words: Real estate market, Pricing, Urban land, Regression models.

Lista de ilustrações

Figura 1 – Unidades Residenciais Vendidas	13
Figura 2 – Ajuste cúbico para variável Área	29
Figura 3 – 2 Fatores - 85% Similaridade	36
Figura 4 – 3 Fatores - 85% Similaridade	37
Figura 5 – Ajuste quadrático para variável Área	43
Figura 6 – Ajuste quadrático para variável DC	43
Figura 7 – Ajuste cúbico para variável DC	44
Figura 8 – Ajuste quadrático para variável DAv	44
Figura 9 – Ajuste cúbico para variável DAv	44
Figura 10 – Valores médios - 85% Similaridade	46

Lista de tabelas

Tabela 1 – Atributos analisados	23
Tabela 2 – Quantidade de terrenos considerados por bairro	24
Tabela 3 – Descrição dos dados coletados	25
Tabela 4 – Normalização dos preços médios	26
Tabela 5 – Correlação das variáveis	27
Tabela 6 – Valor-p para correlação das variáveis	27
Tabela 7 – Significância do modelo de regressão linear simples	28
Tabela 8 – Significância dos ajustes de regressão	29
Tabela 9 – Significância do modelo de regressão múltipla	30
Tabela 10 – Coeficientes do modelo de regressão múltipla	31
Tabela 11 – Etapas do agrupamento para os valores médios	33
Tabela 12 – Coeficientes do modelo de regressão - Grupo 7	34
Tabela 13 – Cargas fatoriais	35
Tabela 14 – Variância e variabilidade dos dados	35
Tabela 15 – Dados sobre os preços	45
Tabela 16 – Preços médios anunciados	46

Lista de abreviaturas e siglas

Km - Quilômetro

CBIC - Câmara Brasileira da Indústria da Construção

m^2 - Metros Quadrados

MLGs - Modelos Lineares Generalizados

DLR - Distância em Linha Reta

DC - Distância de Carro

DP - Distância a Pé

DAv -Distância até Avenida/Rodovia

Sumário

1	INTRODUÇÃO	11
1.1	Objetivos	12
1.1.1	Objetivo Geral	12
1.1.2	Objetivos Específicos	12
1.2	Justificativa	13
1.3	Organização do trabalho	14
2	REFERENCIAL TEÓRICO	15
2.1	Mercado	15
2.2	Preço	15
2.3	Variáveis	16
2.3.1	Relação entre variáveis	16
2.4	Modelos Matemáticos de Regressão	17
2.5	Análise de Agrupamento	18
2.6	Análise Fatorial	19
2.7	Estudos Semelhantes	20
3	METODOLOGIA DE PESQUISA	22
3.1	Classificação da metodologia de pesquisa	22
3.2	Procedimentos metodológicos de pesquisa	22
4	ANÁLISE E RESULTADOS	25
4.1	Análise Geral dos Dados	25
4.2	Análise de Correlação	26
4.3	Análise de Regressão	27
4.3.1	Regressão Linear Múltipla	30
4.4	Análise de Agrupamento	32
4.5	Análise Fatorial	34
4.5.1	Critério de Kaiser	35
4.5.2	Critério da Variância Total	36
5	CONCLUSÃO	38
	REFERÊNCIAS	39
	APÊNDICE A – MATERIAIS ELABORADOS PELO AUTOR	43

1 Introdução

Em um mercado imobiliário é possível observar diversas semelhanças entre os imóveis, embora não haja um imóvel igual a outro. Sendo assim, cada imóvel terá pelo menos uma característica singular (PINTO; FERNANDES, 2019). De maneira geral, os imóveis possuem características que os equiparam ou diferenciam dos outros, e que são importantes na tomada de decisão dos agentes econômicos (RIBEIRO; ROSA; OLIVEIRA, 2017). Como todo imóvel tem necessariamente que se situar em um dado terreno, espera-se que este por si só seja um elemento base para a formação do preço dos imóveis comerciais e residenciais (CARMO, 2014).

Segundo Aguirre (2012), o setor imobiliário é caracterizado por ser dinâmico e complexo, em que fatores como renda, competição do mercado e tendências econômicas impactam em seus diversos campos. Analisando o mercado de vendas de terrenos urbanos (lotes), um fator que influencia no preço é o desenvolvimento econômico do país. De acordo com Carmo (2014 apud RIBEIRO; ROSA; OLIVEIRA, 2017), o desenvolvimento econômico impacta porque o crescimento da economia permite aos indivíduos adquirirem créditos imobiliários de maneira mais simplificada.

Em se tratando da economia local, referente à cidade foco deste estudo, sabe-se que o PIB per capita de João Monlevade em 2018 foi de R\$36.814,98 (IBGE, 2018a). Ademais, ao analisar os dados do Cadastro Central de Empresas, observa-se que João Monlevade contava, em 2018, com 2.416 unidades de empresas e outras organizações atuantes (IBGE, 2018b). Além do quantitativo considerável de empresas, também é importante destacar a existência de 3 unidades acadêmicas de ensino superior no município (RIBEIRO; ROSA; OLIVEIRA, 2017). Note que as particularidades destacadas mostram o potencial de movimentação do setor imobiliário na cidade.

Além do fator econômico, para Dantas, Magalhães e Vergolino (2007), a precificação dos imóveis também pode ser influenciada pelos fatores estruturais e locacionais. Para exemplificar esse tipo de comportamento, é fácil entender que a área de um lote, como um fator estrutural, tem relação direta com o seu preço, ou seja, quanto maior a área em m^2 maior o preço praticado (CARMO, 2014). Outro exemplo é notar que lotes que têm superfície plana em uma mesma região tendem a ser mais valorizados do que aqueles que tenham aclives ou declives acentuados, considerando que, em geral, estes implicarão em gastos mais elevados em uma eventual construção (FOLHA DE SÃO PAULO, 2002). No que se refere a localização, aspectos como proximidade do lote com centros comerciais, avenidas ou ruas pavimentadas, podem ocasionar na flutuação do preço (DANTAS; MAGALHÃES; VERGOLINO, 2007).

Uma forma de prever o preço de um bem ou serviço é traduzindo um conjunto de dados reais em modelos matemáticos e que incluem parâmetros e variáveis explicativas, ou seja, é possível formular equações aproximadas para descrever o valor de bens e serviços em função de suas características (AGUIRRE, 2012). Para estimar o preço de um lote, por exemplo, pode ser utilizado o método de regressão, que, segundo Rosa, Oliveira e Pinto (2018b), depende das informações referentes ao conjunto de imóveis e de uma lista contendo suas características físicas e elementos de atribuição de valor.

Quando agrupadas todas as informações dos lotes, ou seja, seus atributos físicos e os elementos de atribuição de valor, existe a possibilidade de haver similaridade entre as características urbanas de alguns bairros de um município. Essa particularidade possibilita que se tenha certo grau de semelhança entre os preços médios cobrados pelos imóveis (ROSA; OLIVEIRA; PINTO, 2018a). Tais semelhanças podem ser destacadas por meio da análise de similaridade e formação de *clusters* (agrupamento dos bairros), que por sua vez podem servir de base para identificação das variáveis determinantes na precificação dos lotes (ROSA; OLIVEIRA; PINTO, 2018a).

Neste contexto, considerando a importância de se estabelecer o preço dos bens e serviços que são transacionados no mercado imobiliário, em especial o valor dos terrenos, este estudo propõe definir quais os preços médios dos terrenos urbanos comercializados na cidade de João Monlevade, Minas Gerais, e o grau de correlação entre as características dos terrenos analisados. Além disso, este trabalho propõe a criação de modelos matemáticos de regressão, lineares e ajustados, para estimar o preço dos terrenos urbanos do município e identificar os atributos de valor que influenciam significativamente nos modelos. Por fim, este estudo também pretende avaliar a possibilidade de agrupamento de bairros, em função do grau de similaridade, e verificar se a formação de *clusters* permite aprimorar os modelos de regressão.

1.1 Objetivos

1.1.1 Objetivo Geral

O presente estudo objetiva avaliar o comportamento da oferta terrenos urbanos (lotes) no município de João Monlevade, no período que compreende os meses de setembro e outubro de 2019, assim como a proposição de modelos de regressão capazes de estimar os preços dos terrenos da cidade.

1.1.2 Objetivos Específicos

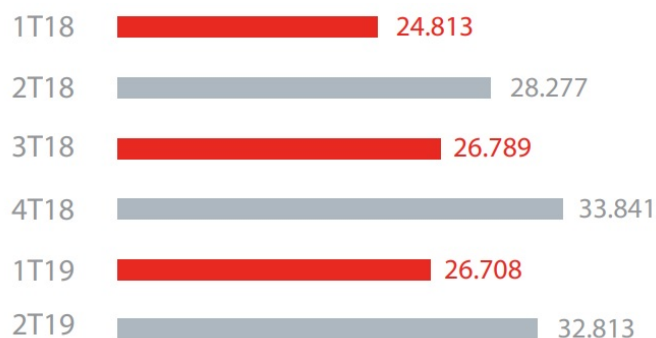
- Levantar um conjunto de dados vinculados aos anúncios online de lotes à venda na cidade de João Monlevade;

- Construir e comparar modelos de regressão para estimar os preços dos lotes;
- Analisar o grau de correlação entre as variáveis utilizadas nos modelos de regressão;
- Identificar o grau de similaridade entre os bairros considerando os atributos dos terrenos;
- Realizar uma análise comparativa entre os preços dos imóveis residenciais de 2017, disponíveis no trabalho de Rosa (2018), e os preços dos lotes da cidade em 2019.

1.2 Justificativa

Segundo a CBIC (2019), o mercado imobiliário brasileiro tem apresentado sinais de recuperação e um significativo crescimento durante o ano de 2019. A Figura 1 apresenta um demonstrativo trimestral das unidades residenciais vendidas entre 2018 e 2019 no Brasil. Quando comparado o segundo trimestre de 2018 com o mesmo período de 2019, é possível observar que a venda de imóveis em todo território brasileiro aumentou de 28.277 para 32.813, apresentando um crescimento de 16%. Somente a região sudeste do Brasil, localização do município foco deste estudo, foi responsável pela venda de 19.233 unidades residenciais no segundo trimestre de 2019, aproximadamente 59% do total (CBIC, 2019).

Figura 1 – Unidades Residenciais Vendidas



Fonte: CBIC (2019)

Assim, considerando a recuperação do mercado imobiliário brasileiro e a sua importância econômica, pode-se dizer que, trabalhos que envolvam esse tipo de conteúdo engrandecem a literatura que trata sobre o assunto. Em pesquisas recentes realizadas na cidade de João Monlevade, foram desenvolvidos alguns trabalhos que investigaram o perfil econômico dos mercados de venda e locação de imóveis residenciais na cidade (ROSA, 2018). Entretanto, não foi identificado até então nenhum trabalho que envolva análises econômicas referentes aos preços dos terrenos urbanos da cidade, ou seja, trata-se de uma lacuna de pesquisa que ainda pode ser explorada.

Portanto, este estudo se justifica em razão da importância do mercado imobiliário e da ausência de estudos sobre os lotes do município. Assim, desenvolver um trabalho envolvendo essa temática se torna fundamental para complementar as pesquisas já realizadas na localidade e para auxiliar agentes interessados em adquirir ou comercializar imóveis na cidade.

1.3 Organização do trabalho

O presente trabalho está estruturado em cinco capítulos. O Capítulo 1 é responsável por introduzir o tema, apresentando a contextualização do problema, descrição dos objetivos, justificativa e a organização do estudo. No Capítulo 2 encontra-se o referencial teórico sobre os temas Mercado, Preço, Variáveis, Modelos Matemáticos de Regressão, Análise de Agrupamento, Análise Fatorial e Estudos Semelhantes. O Capítulo 3 é destinado a metodologia de pesquisa. O Capítulo 4 inclui as análises e os resultados obtidos com o desenvolvimento do estudo sobre os terrenos urbanos. Por fim, o Capítulo 5 apresenta as conclusões e as sugestões para trabalhos futuros.

2 Referencial teórico

2.1 Mercado

Conforme [Pindyck e Rubinfeld \(2010\)](#), o mercado é formado pelo agrupamento de vendedores (ofertas) e compradores (demanda), que por meio de interações determinam o preço de cada bem. Ainda no entendimento de [Pindyck e Rubinfeld \(2010\)](#), sua definição consiste no processo de identificar os compradores, vendedores e produtos envolvidos no mercado. Sendo assim, o mercado imobiliário pode ser definido pela presença de vendedores que surgem em razão da expectativa de rendimentos dos imóveis, e de compradores que surgem da necessidade de aquisição de um ambiente habitacional ([RIBEIRO; ROSA; OLIVEIRA, 2017](#)).

O mercado imobiliário faz parte de um dos componentes urbanos que envolve um volume suficientemente grande de interações potenciais entre um conjunto de agentes econômicos que incluem os proprietários, locadores, locatários, intermediários (imobiliárias) e investidores, que objetivam explorar os espaços geográficos de uma cidade ([MATOS; BARTKIW, 2013](#)). Neste setor, cabe às imobiliárias a viabilização, facilitação e realização de boa parte do processo de intermediação comercial para compra, venda e locação de imóveis ([MATOS; BARTKIW, 2013](#)). Visto que o mercado habitacional é um setor chave para a economia, gerando empregos e renda para o sistema econômico, um estudo microeconômico desse mercado é voltado para o entendimento do comportamento de cada indivíduo, onde modelos explicam como o consumidor toma sua decisão no processo de seleção da habitação ([DANTAS; MAGALHÃES; VERGOLINO, 2010](#)).

2.2 Preço

De acordo com [Pindyck e Rubinfeld \(2010\)](#), o preço é definido pela interação entre consumidores, empresas e trabalhadores. Segundo [Almeida \(2011\)](#), ele possui a capacidade de influenciar na decisão sobre a quantidade que o consumidor adquire de um bem. Visto que preço é um influenciador na tomada de decisão, o entendimento do valor que o consumidor está disposto a pagar por um bem em função de seus atributos, também nomeado preço implícito ou hedônico, é de grande interesse ([DANTAS; MAGALHÃES; VERGOLINO, 2007](#)).

Para [Hermann e Haddad \(2005\)](#), a análise do preço hedônico é caracterizada pela estimativa de uma função preço, na qual o preço fica de um lado e suas variáveis explicativas (características) de outro. Dessa forma, estudos que visam analisar o mercado de venda

e locação de imóveis normalmente utilizam o método hedônico (PINTO; FERNANDES, 2019).

2.3 Variáveis

Segundo Levine, Stephan e Szabat (2017), as variáveis podem ser classificadas como categóricas ou numéricas, em que as variáveis categóricas são qualitativas, e as numéricas são quantitativas que podem ser segmentadas em discreta ou contínua. Montgomery e Runger (2018) definem que as variáveis discretas apresentam uma faixa que é finita, ou infinita contável, e as contínuas possuem um intervalo, finito ou infinito, de números reais para a faixa. Assim, as discretas surgem da técnica de contagem, e as contínuas surgem do procedimento de mensuração (LEVINE; STEPHAN; SZABAT, 2017).

Torna-se interessante representar numericamente as características dos dados (MONTGOMERY; RUNGER, 2018). Um conjunto de variáveis numéricas pode ser caracterizada considerando a sua variação. Um exemplo de medida de variação é o desvio-padrão, que considera como os dados estão distribuídos e mensura a dispersão média dos valores em relação a média (LEVINE; STEPHAN; SZABAT, 2017). Para Pinto e Fernandes (2019), um valor alto para o desvio-padrão, por exemplo, indica que os dados estão localizados distantes da média. Outrossim, segundo Montgomery e Runger (2018), os dados também podem ser descritos em relação a sua tendência central mediante a média aritmética.

Uma análise de regressão, segundo Levine, Stephan e Szabat (2017), permite a definição de modelos para estimar os valores de uma variável numérica a partir de outras variáveis, ou seja, em modelos de regressão existem as variáveis dependentes e as independentes. A variável dependente corresponde a característica que pretende estimar, ou seja, ocorre a previsão de seu valor através de um conjunto de outras variáveis explicativas (KASZNAR; GONÇALVES, 2011). Consequentemente, as variáveis explicativas são empregadas para prever ou descrever a variável dependente, e algumas das variáveis podem não ser fundamentais para o problema e podem ser excluídas do estudo (BRÁULIO, 2005). Para exemplificar, em uma avaliação de imóveis a variável que pretende estimar é o preço, e as variáveis independentes (explicativas) correspondem aos atributos físicos e locacionais dos imóveis que influenciam no valor final (GAZOLA, 2002).

2.3.1 Relação entre variáveis

Para avaliar a relação entre duas variáveis numéricas, pode ser utilizado o método do coeficiente de correlação (LEVINE; STEPHAN; SZABAT, 2017). Segundo Figueiredo Filho e Silva Júnior (2009), o coeficiente de correlação de Pearson flutua de -1 (forte correlação negativa) até 1 (forte correlação positiva), ou seja, o valor indica a força e o sinal a direção do relacionamento. Já um valor de correlação entre 0,30 e 0,70, por

exemplo, corresponde a uma relação média entre as variáveis (ALVES, 2005). Em resumo, o coeficiente de correlação corresponde a uma medida de relação linear entre as variáveis (ROSA; OLIVEIRA; PINTO, 2018b). Levine, Stephan e Szabat (2017) ressaltam que o coeficiente de correlação indica uma tendência dos dados, e não que a variação de uma variável causa a variação de outra.

2.4 Modelos Matemáticos de Regressão

Na análise de regressão é estudado o relacionamento entre a variável dependente e as independentes, e este relacionamento é representado por uma equação que associa essas variáveis (HENRIQUES, 2011). Esse modelo matemático pode ter características lineares e/ou não lineares. Para Montgomery e Runger (2018), um modelo linear considera que o valor esperado da variável dependente seja uma função linear da variável independente. Como exemplo de regressão linear, é possível listar a regressão simples, em que apenas uma variável independente é utilizada para explicar a variável dependente, e a regressão múltipla, na qual duas ou mais variáveis são utilizadas para explicar a variável dependente (LEVINE; STEPHAN; SZABAT, 2017). Ainda de acordo com Levine, Stephan e Szabat (2017), é possível fazer estimativas mais precisas quando é utilizada mais de uma variável explicativa.

A Equação (2.1) define o modelo de regressão linear simples e a Equação (2.2) corresponde ao modelo de regressão linear múltipla com duas variáveis independentes. Seguindo as descrições dos modelos utilizados por Levine, Stephan e Szabat (2017) e Montgomery e Runger (2018), o parâmetro Y é o que pretende estimar, M_0 corresponde ao intercepto de Y , e E_1 diz respeito a inclinação de Y em relação à primeira variável R_1 . Já o termo E_2 corresponde a inclinação de Y em relação à segunda variável R_2 . Além disso, *Erro* se refere ao erro aleatório em Y . A estratégia mais utilizada para determinar o intercepto e as inclinações das equações de regressão é o método dos mínimos quadrados (LEVINE; STEPHAN; SZABAT, 2017). Neste método, de acordo com o Suporte ao Minitab 18 (2019c), é formulado um modelo que minimiza a soma do quadrado das distâncias entre os dados coletados e os valores estimados pela equação, ou seja, o intercepto e as inclinações são estipuladas de forma que o somatório dos quadrados dos resíduos, ou erro, seja a menor possível (FÁVERO; BELFIORE, 2017).

$$Y = M_0 + E_1R_1 + Erro \quad (2.1)$$

$$Y = M_0 + E_1R_1 + E_2R_2 + Erro \quad (2.2)$$

Em relação aos modelos não lineares, a função quadrática é uma das mais frequentes. O modelo de regressão quadrático é similar ao linear múltiplo, entretanto a segunda variável

independente corresponde ao quadrado da primeira variável independente (LEVINE; STEPHAN; SZABAT, 2017). A Equação (2.3) define o modelo de regressão quadrático (não linear). De forma similar, a variável Y é o que objetiva determinar, M_0 refere ao intercepto de Y , E_1 diz respeito ao coeficiente do efeito linear em Y e E_2 é o coeficiente do efeito quadrático em Y . A variável R_1^2 equivale ao quadrado da variável R_1 , e *Erro* se refere ao erro aleatório em Y .

$$Y = M_0 + E_1R_1 + E_2R_1^2 + Erro \quad (2.3)$$

Referente a avaliação do modelo de regressão, é importante observar o coeficiente de determinação, R^2 , que mensura o quanto da variabilidade da variável dependente é explicada pelas variáveis independentes, e o coeficiente de determinação ajustado, $R^2(\text{adj})$, que pode ser utilizado na comparação entre modelos que estimam a mesma variável mas com o número de variáveis independentes diferentes (LEVINE; STEPHAN; SZABAT, 2017). O valor do coeficiente de determinação flutua entre 0 e 1, e um coeficiente igual a 0,50, por exemplo, indica que o modelo consegue explicar 50% da variação dos dados (MONTGOMERY; RUNGER, 2018). Como o coeficiente de determinação mensura a qualidade do ajuste, um valor próximo de 1, ou 100%, sinaliza que o ajuste é bom (BRÁULIO, 2005).

Outrossim, é importante determinar se a relação entre as variáveis independentes e a variável dependente é significativa. Neste caso, utiliza-se o teste F geral, em que, um valor F maior que o valor crítico da distribuição F, que é tabulado considerando o nível de significância e os graus de liberdade, ou um valor-p menor que o nível de significância, indica que a variável independente está relacionada com a variável resposta (LEVINE; STEPHAN; SZABAT, 2017).

2.5 Análise de Agrupamento

A análise multivariada refere-se às técnicas estatísticas que objetivam analisar, de maneira conjunta, múltiplas medidas sobre os objetos estudados (HAIR et al., 2009). Um exemplo de técnica multivariada é a análise de agrupamento, ou análise de *cluster*, que objetiva separar os elementos de estudo em diferentes grupos, de maneira que os componentes de um mesmo grupo sejam semelhantes entre si (MINGOTI, 2005). Resumidamente, a análise de agrupamento visa formar grupos homogêneos de elementos (BARROSO; ARTES, 2003). Na formulação dos grupos, utilizando os métodos estatísticos, são aplicados o conceito de similaridade. A análise de similaridade é obtida através de coeficientes, que, dependendo do seu valor absoluto, podem indicar maior ou menor grau de semelhança entre os grupos (FREI, 2006). Sendo assim, quanto menor o índice de similaridade, maior

é a dissimilaridade entre os objetos analisados (BRÁULIO, 2005). Assim, a similaridade entre as observações pode ser mensurada pelo método da distância, ou seja, a proximidade entre os elementos (HAIR et al., 2009).

Referente a formulação dos grupos, pode ser utilizado o método hierárquico. Neste método, cada elemento inicia como um grupo e em cada etapa são reunidos dois grupos, até que na última etapa observa-se somente um conglomerado com todos os elementos (HAIR et al., 2009). Para a reunião dos grupos, é considerado o tipo de ligação entre os elementos (BRÁULIO, 2005), ou seja, a formação dos grupos pode ser baseada na distância euclidiana e pela ligação entre os grupos parciais (STEINER et al., 2008). Segundo Bráulio (2005), os grupos são estabelecidos através do cálculo das distâncias entre os vetores (formados pelos atributos), pela construção de uma matriz de distâncias e pela união dos itens conforme suas semelhanças.

A técnica de ligação simples e da distância euclidiana podem ser aplicadas para agrupar bairros de um município em razão dos atributos dos imóveis (ROSA; OLIVEIRA; PINTO, 2018a). Ou seja, em cada etapa é realizado o agrupamento reunindo os dois grupos que apresentaram a menor distância (BRÁULIO, 2005). É uma forma de retratar as etapas do agrupamento é por meio de um dendrograma, um gráfico na forma de árvore em que o eixo vertical assinala o grau de similaridade entre os grupos (MINGOTI, 2005).

2.6 Análise Fatorial

Outro exemplo do método multivariado é a análise fatorial, cuja técnica é aplicada em estudos que apresentam grande número de variáveis inter-relacionadas, e objetiva identificar um número reduzido de novas variáveis alternativas que resumem as informações das variáveis iniciais (MINGOTI, 2005). Ou seja, visa entender a correlação entre um conjunto grande de variáveis em termos de poucas variáveis aleatórias (fatores) (BAPTISTELLA; STEINER; NETO, 2006). Como resultado da análise fatorial, é possível obter grupos que são compostos por variáveis altamente correlacionadas entre si, mas que apresentam um nível de correlação baixa com as variáveis dos outros grupos (BAKKE; LEITE; SILVA, 2008).

Inicialmente, em uma análise fatorial são definidas as variáveis de estudo e a matriz de correlação, e posteriormente são definidos o método de extração e o número de fatores necessários para explicar os dados (HAIR et al., 2009). Ainda no entendimento de Hair et al. (2009), para extrair os fatores são utilizados alguns critérios como o de percentagem da variância, que visa garantir que os fatores selecionados expliquem um valor especificado da variância total, e o critério da raiz latente (critério de Kaiser), em que somente os fatores que possuem raízes latentes, ou autovalores, superior a 1 são significativos para o estudo. Em análises posteriores, é possível representar os fatores utilizando os escores. Para

calcular o escore de cada observação, são considerados os pesos dos fatores e os valores das variáveis iniciais (HAIR et al., 2009), em outras palavras, inicialmente os dados das variáveis são escalados, centralizados e subtraídos das médias, e depois são multiplicados pelos coeficientes dos fatores (SUPORTE AO MINITAB 18, 2019b).

2.7 Estudos Semelhantes

Almeida (2011) desenvolveu uma pesquisa no Distrito Federal visando entender os fatores determinantes na formação do preço, e definir um modelo de regressão linear que explicasse os preços praticados. O resultado obtido foi positivo, ou seja, com o auxílio das variáveis selecionadas o autor conseguiu estruturar um modelo de regressão que explicasse os preços no mercado. Analogamente, Aguirre (2012) realizou um estudo sobre a precificação de apartamentos em diferentes regiões do Distrito Federal. O trabalho visou compreender os fatores determinantes nos preços dos imóveis, além de mensurar seus impactos e de estruturar modelos de precificação. O autor estimou cinco modelos lineares generalizados (MLGs), um para cada região, e como resultado conseguiu analisar os impactos dos atributos nos preços e observar como o peso de cada característica sofre variação de uma região para outra.

Pinto e Fernandes (2019) trabalharam com uma análise de preços hedônicos no mercado residencial, casas e apartamentos, de Conselheiro Lafaiete, em Minas Gerais. Os autores estruturaram um modelo de regressão linear múltipla, e como resultado observaram que um conjunto de variáveis analisadas apresentaram uma correlação positiva com a variável dependente preço, indicando que esses atributos contribuem para o aumento do preço de mercado. Bráulio (2005) realizou um trabalho na cidade de Campo Mourão, Paraná, para analisar os preços dos imóveis (apartamentos, casas e terrenos). Na pesquisa foi apresentado o processo de construção do modelo de regressão linear múltipla com o auxílio da técnica de análise de agrupamento (*Clusters Analysis*).

Outro problema similar, sobre identificação de determinantes nos preços dos imóveis, foi abordado por Carmo (2014). Os imóveis avaliados foram os terrenos urbanos da cidade de Uberlândia, Minas Gerais, e o método utilizado na análise de regressão linear múltipla foi o *stepwise*. Neste método ocorre a inclusão ou exclusão de cada uma das variáveis explicativas, auxiliando na identificação das características que melhor explicam o modelo. Como resultado, somente quatro, dos nove atributos inicialmente analisados, foram determinantes nos preços. Já Fávero, Belfiore e Lima (2008) procuraram compreender os pacotes de características, intrínsecas e extrínsecas, que apresentam maior representatividade para cada submercado (demanda e oferta) da Região Metropolitana de São Paulo. Neste caso, a análise dos imóveis residenciais foi realizada através dos métodos de análise fatorial e de especificação logarítmica.

No município de João Monlevade, onde o presente estudo foi desenvolvido, foram realizadas quatro pesquisas relacionadas ao mercado imobiliário. [Ribeiro, Rosa e Oliveira \(2017\)](#) explorou o tema sobre as variáveis que impactam na formação do valor de compra, venda e locação de imóveis. Como resultado da pesquisa, notou-se que as variáveis que impactam significativamente no valor médio foram localização e o índice de criminalidade dos bairros. Com auxílio dos dados coletados e das análises elaboradas por [Ribeiro, Rosa e Oliveira \(2017\)](#), foram desenvolvidos outros trabalhos. No primeiro estudo, desenvolvido por [Rosa, Oliveira e Pinto \(2018a\)](#), foi realizada uma análise de *clusters*. O segundo, elaborado por [Rosa, Oliveira e Pinto \(2018b\)](#), desenvolveu modelos de precificação para venda e locação dos imóveis, utilizando os métodos de regressão linear múltipla. Os resultados observados foram positivos, ou seja, próximos da realidade do município. Outrossim, foi elaborado um trabalho de monografia por [Rosa \(2018\)](#), objetivando reunir todas as informações coletadas e realizar um estudo robusto sobre o mercado imobiliário de João Monlevade.

3 Metodologia de pesquisa

3.1 Classificação da metodologia de pesquisa

Para [Turrioni e Mello \(2012\)](#), uma pesquisa científica desenvolvida no curso de Engenharia de Produção pode ser classificada em quatro esferas: natureza, objetivos, abordagem e método. Em relação a natureza da pesquisa, o estudo corresponde a uma pesquisa aplicada, que, segundo [Rodrigues \(2007\)](#), tem o interesse de comprovar hipóteses indicadas em modelos teóricos. Quanto aos objetivos, o estudo se classifica como explicativo e descritivo, porque visa definir aspectos que determinam a ocorrência de fenômenos (explicativa) e posteriormente estabelecer relações entre variáveis (descritiva) ([TURRIONI; MELLO, 2012](#)).

Segundo [Miguel et al. \(2012\)](#), a principal característica da abordagem quantitativa é a mensuração das variáveis. Sendo assim, a pesquisa é classificada como quantitativa porque objetiva quantificar variáveis e utilizar recursos estatísticos. No que se refere ao método, a pesquisa pode ser classificada como experimento porque visa analisar como algumas variáveis influenciam no objeto de estudo.

3.2 Procedimentos metodológicos de pesquisa

Os procedimentos de coleta dos dados foram semelhantes aos realizados e apresentados por [Rosa \(2018\)](#). Sendo assim, objetivando o entendimento sobre os métodos de precificação, foi realizada uma pesquisa sobre os temas abordados e estudos semelhantes. Foram encontradas diversas pesquisas sobre análises de preços do mercado imobiliário, que foram utilizadas como orientação para o presente estudo. Ademais, foram coletados os dados dos lotes urbanos em uma pesquisa na *web*, considerando os seguintes atributos de valor: preço anunciado, área (m^2) e a localização geográfica (bairro e suas referências).

A coleta dos dados ocorreu durante os meses de setembro e início de outubro de 2019, em *websites* especializados em anunciar imóveis à venda e aluguéis em João Monlevade, Minas Gerais. Para auxiliar na coleta, foi utilizado o termo “Lote” como filtro nos *websites* consultados. Como resultado, 16 plataformas da *web* apresentaram anúncios sobre vendas de lotes, sendo 14 de imobiliárias da cidade de João Monlevade (São José, Martins, Casa Forte, Sollar, Janete Lage, Logos, Senna, Remo, Casa Nova, Delci Couto, Locar, Roberto, Pedrosa, Bruno Lopes), 1 do Facebook (imobiliária Pontual) e 1 de anúncios *online* (OLX).

Como resultado, foram observados e coletados 285 anúncios sobre vendas de terrenos

na cidade. Importante ressaltar que em alguns anúncios as informações sobre os bairros eram relativos aos loteamentos do município, assim a localização foi ajustada para o bairro mais próximo. Outrossim, foram encontrados anúncios em que a quantidade de lote ofertado era superior a 1, e as informações disponíveis eram referentes a quantidade total dos lotes. Somente em alguns casos eram informados a área e o preço dos lotes separados. Visando uma padronização, foram considerados os dados relacionados à quantidade total dos terrenos anunciados.

Para complementar os dados, foram definidas as distâncias dos lotes com o centro comercial de João Monlevade e dos lotes com a rodovia ou avenida mais próxima. Foi estabelecido como ponto de referência no centro comercial do município o Banco Bradesco ou, de forma alternativa, a Praça do Lindinho. A rodovia utilizada no estudo foi a BR-381, e as avenidas Gentil Bicalho, Wilson Alvarenga, Getúlio Vargas, Alberto Lima e Armando Fajardo. As cinco avenidas foram priorizadas porque juntas atravessam o município, ou seja, são consideradas estratégicas. Para determinar as distâncias, em quilômetros (Km), foi utilizado o *Google Maps*. Para as distâncias entre os lotes e o centro comercial da cidade foram selecionadas as opções de distâncias em linha reta, carro e a pé. Já nas distâncias em relação à avenida ou rodovia foi utilizada a opção a pé. Para as opções de distâncias de carro e a pé foram coletados os menores valores. A Tabela 1 demonstra os atributos analisados no estudo e suas unidades de medidas (Unid.).

Tabela 1 – Atributos analisados

Atributos	Unid.
Preço ofertado	R\$
Área	m ²
Distância do lote ao centro comercial em linha reta (DLR)	Km
Distância do lote ao centro comercial de carro (DC)	Km
Distância do lote ao centro comercial a pé (DP)	Km
Distância do lote em relação à avenida ou rodovia mais próxima (DAv)	Km

Fonte: O autor (2020)

Como resultado final da coleta, apenas 189 dos 285 anúncios encontrados, aproximadamente 66%, apresentaram nas descrições todas as informações necessárias para o desenvolvimento do estudo, ou seja, dados sobre preço, área e localização (endereço, ponto de referência ou imagens que possibilitaram a localização do lote). Foram encontrados anúncios em 37 bairros de João Monlevade. Do total, 27 bairros apresentaram no mínimo dois anúncios e 10 apresentaram anúncio único. A Tabela 2 apresenta a quantidade de anúncio consideradas por bairro e a Tabela 15 (Apêndice A) contém os preços máximo, mínimo e médio de cada bairro. Para a organização e armazenamento dessas informações coletadas durante o estudo, foi utilizado o *software* Excel 2016.

O *software* Minitab 17 foi utilizado nas etapas de desenvolvimento e análise dos

Tabela 2 – Quantidade de terrenos considerados por bairro

Bairro	Quantidade	Bairro	Quantidade
Aclimação	3	Nova Aclimação	6
Alvorada	4	Nova Esperança	3
Bau	1	Novo Cruzeiro	4
Belmonte	2	Novo Horizonte	1
Boa Vista	2	Paineiras	4
Campos Elísios	7	Petrópolis	2
Carneirinhos	6	Recanto Paraíso	7
Cidade Nova	13	República	3
Cruzeiro Celeste	3	Rosário	5
Industrial	1	Santa Bárbara	8
Ipiranga	1	Santa Cecília	2
Jacuí	1	São João	1
JK	3	Satélite	2
José Elói	6	Santo Hipólito	18
Loanda	14	Sion	35
Lourdes	6	Tanquinho	1
Lucília	1	Teresópolis	4
Mangabeiras	7	Vera Cruz	1
Metalúrgico	1		

Fonte: O autor (2020)

resultados, auxiliando na elaboração de informações, modelos e gráficos que facilitaram a visualização e o entendimento do estudo. A princípio, foi analisado o comportamento das variáveis, como o preço e a área média dos lotes anunciados em João Monlevade, e realizada uma breve comparação entre os valores médios dos terrenos e dos imóveis residenciais do município. Também foi mensurado o relacionamento entre as características analisadas no estudo, determinando a direção e o nível de força.

Em seguida, visando estimar os preços dos terrenos, foram formulados e avaliados diferentes modelos de regressão. Inicialmente foram formulados os modelos simples, considerando somente uma variável explicativa. Ao observar os resultados alcançados com os modelos, foram realizados diferentes ajustes que visavam melhorar o nível de explicação. Ao considerar que uma variável explicativa não era suficiente para estimar um modelo que atendesse todo município, foi estimado um modelo múltiplo, considerando três variáveis. Por fim, com os resultados obtidos com o modelo múltiplo, foram elaborados modelos que considerassem as semelhanças entre os bairros.

4 Análise e Resultados

4.1 Análise Geral dos Dados

Visando entender o comportamento das variáveis selecionadas no estudo, foram calculados os valores médio, mínimo, máximo, a mediana e o desvio padrão de cada atributo. A Tabela 3 retrata os valores encontrados para as características e o número de observações analisadas. Em relação aos valores máximos e mínimos, por exemplo, o lote mais caro encontrado no município foi R\$3.000.000,00, e o menor terreno anunciado tinha $165m^2$. Nota-se que um lote anunciado no município de João Monlevade possui um preço médio de R\$261.607,00 e uma área média de $467,30m^2$. Em relação às distâncias com o centro comercial da cidade, o lote fica aproximadamente 3 Km em linha reta (DLR), 4 Km de carro (DC) e 3 Km a pé (DP). Outrossim, o lote fica cerca de 0,60 Km de alguma avenida ou rodovia importante da cidade (DAv). Para analisar a variação dos dados coletados, foi utilizado o desvio-padrão. Observou-se que a maioria das características apresentaram um valor considerável para o desvio-padrão. Para exemplificar, o atributo “Área” está flutuando $334,80m^2$ em relação a média ($467,30m^2$), ou seja, entre $802,10m^2$ e $132,50m^2$.

Tabela 3 – Descrição dos dados coletados

Atributo	N	Média	Desvio padrão	Mínimo	Mediana	Máximo
Preço (R\$)	189	261.607,00	368.848,00	30.000,00	150.000,00	3.000.000,00
Área (m^2)	189	467,30	334,80	165,00	360,00	3510,00
DLR (Km)	189	2,63	1,42	0,19	2,47	7,70
DC (Km)	189	3,87	1,96	0,60	3,70	10,00
DP (Km)	189	3,32	1,79	0,25	3,10	9,60
DAv (Km)	189	0,60	0,46	0,00	0,50	3,60

Fonte: O autor (2020)

Outrossim, o estudo procurou realizar uma breve comparação entre os preços médios dos terrenos e dos imóveis residenciais do município. As informações dos imóveis utilizados na comparação foram agrupados e disponibilizados por Rosa (2018). É importante destacar alguns aspectos, como a diferença de 3 anos entre as coletas, o número diferente de anúncios comparados, e o número reduzido de bairros analisados, já que foram considerados somente os bairros que apresentaram as informações para os três aspectos (terreno, apartamento e casa). É possível observar na Tabela 16 (Apêndice A) os valores médios anunciados para cada segmento.

Para realizar a comparação, os valores foram normalizados entre 0 (menor preço) e

1 (maior preço), e posteriormente foram arredondados considerando duas casas decimais. Nota-se na Tabela 4 que o bairro Aclimação apresentou os maiores valores médios para o segmento apartamento e casa. Em relação ao terreno, ele possui o terceiro maior valor (0,38), entretanto é possível notar uma diferença significativa ao comparar com os dois primeiros, Santa Bárbara (1,00) e Carneirinhos (0,96).

Tabela 4 – Normalização dos preços médios

Bairro	Terreno	Apartamento	Casa
Aclimação	0,38	1,00	1,00
Carneirinhos	0,96	0,74	0,60
José Elói	0,09	0,27	0,00
Lourdes	0,24	0,29	0,35
Novo Horizonte	0,24	0,71	0,39
Paineiras	0,31	0,31	0,36
Santa Bárbara	1,00	0,60	0,33
Satélite	0,00	0,00	0,12
Sion	0,04	0,18	0,02

Fonte: O autor (2020)

Três bairros são responsáveis pelos menores valores nos três segmentos analisados. O bairro Satélite apresentou os menores valores para terreno e apartamento, e o terceiro menor em casa. Por outro lado, José Elói demonstrou o menor valor para casa e o terceiro menor para terreno e apartamento. O bairro Sion ocupou a segunda posição nos três segmentos. A localização dos três bairros é relativamente distante do centro comercial, ou seja, a proximidade com o centro comercial do município pode ser um requisito que impacta na escolha de um terreno ou imóvel.

4.2 Análise de Correlação

Objetivando identificar a relação entre os atributos, foi calculado o coeficiente de correlação. Como o foco do estudo é entender a precificação dos lotes, inicialmente foi elaborada uma análise considerando o relacionamento do preço com as demais variáveis. Observa-se na Tabela 5 que o atributo Preço tem uma correlação positiva de 0,533 com a variável Área. Por ser um coeficiente positivo, existe a tendência de que o aumento da área do lote gere um aumento no preço praticado. Em relação ao preço com as distâncias, todos os valores foram negativos. Isso pode indicar que à medida que o lote se distancia dos pontos de referência do estudo, o preço é reduzido. A correlação mais forte entre as distâncias até o centro comercial foi a distância de carro (DC), -0,385, seguido por distância a pé (DP), -0,376, e em linha reta (DLR), -0,354. Já a correlação do preço com a avenida/rodovia mais próxima foi -0,330. Esses valores de correlação indicam uma relação média entre as variáveis com o preço dos terrenos.

Tabela 5 – Correlação das variáveis

	Preço	Área	DLR	DC	DP
Área	0,533				
DLR	-0,354	0,026			
DC	-0,385	-0,006	0,978		
DP	-0,376	0,014	0,988	0,991	
DAv	-0,330	-0,168	0,451	0,434	0,453

Fonte: O autor (2020)

Outrossim, as variáveis DLR, DC e DP, apresentaram forte correlação positiva entre elas. Isso já era esperado, dado que as três medem a mesma característica, a distância do terreno até o centro comercial do município. Entretanto, apresentaram coeficientes quase nulos com a variável Área, ou seja, uma variável não está relacionada à outra. O atributo DAv também apresentou uma correlação próxima de 0 com Área, e valores médios para as demais distâncias.

Visando definir se os coeficientes de correlação entre as variáveis analisadas são significativos, foram calculados os valores-p. O nível de significância selecionado para a comparação foi de 0,05 (SUPORTE AO MINITAB 18, 2019a). Ao observar os resultados relacionados à variável Preço e cada variável explicativa (Área, DLR, DC, DP e DAv), Tabela 6, nota-se que o valor-p foi igual a 0. Como os valores são menores que 0,05, significa que todos os coeficientes de correlação, entre Preço e as variáveis explicativas, são significativos. Já o valor-p entre Área e as variáveis DLR (0,724), DC (0,938) e DP (0,844), foram maiores que 0,05. Ou seja, a correlação entre essas variáveis não é significativa.

Tabela 6 – Valor-p para correlação das variáveis

	Preço	Área	DLR	DC	DP
Área	0,000				
DLR	0,000	0,724			
DC	0,000	0,938	0,000		
DP	0,000	0,844	0,000	0,000	
DAv	0,000	0,021	0,000	0,000	0,000

Fonte: O autor (2021)

4.3 Análise de Regressão

Inicialmente, foram desenvolvidas equações que objetivam estimar o preço do terreno considerando, de forma isolada, cada variável explicativa. Esse modelo é o de regressão linear simples, que estima o preço de um terreno no município considerando somente uma variável independente. Para formular os modelos, o nível de confiança para

os intervalos foi de 95%. Foram elaboradas três equações, em que a Equação (4.1) tem como variável explicativa a área do terreno, a Equação (4.2) utiliza a distância do lote até o centro comercial e a Equação (4.3) utiliza a distância da avenida/rodovia mais próxima. Como descrição das equações, o primeiro valor corresponde ao intercepto da variável preço, e o segundo valor indica a inclinação do preço em relação às variáveis explicativas.

$$\text{Preço} = -12.773,00 + 587,20 \text{ Área} \quad (4.1)$$

$$\text{Preço} = 542.210,00 - 72.550,00 \text{ DC} \quad (4.2)$$

$$\text{Preço} = 421.788,00 - 264.406,00 \text{ DA} \quad (4.3)$$

Em relação às equações envolvendo as distâncias dos terrenos até o centro comercial, foram desenvolvidos modelos para as três variáveis e observados os coeficientes de determinação R^2 e $R^2(\text{adj})$. Desta forma, a variável DC foi selecionada como representante do aspecto locacional (distância do terreno urbano até o centro comercial), visto que apresentou um maior $R^2(\text{adj})$. Entretanto, como é possível observar na Tabela 7, todas as equações apresentaram um coeficiente de determinação, R^2 , e um coeficiente de determinação ajustado, $R^2(\text{adj})$, baixos. Uma explicação é que, em alguns casos, os dados não apresentam uma relação linear, então o modelo precisa ser ajustado (BLOG DA MINITAB, 2019a). Sendo assim, os dados foram ajustados para modelos quadráticos e cúbicos.

Tabela 7 – Significância do modelo de regressão linear simples

Equação	S	R^2	$R^2(\text{adj})$
(4.1)	312.931,00	28,40%	28,00%
(4.2)	341.249,00	14,90%	14,40%
(4.3)	349.119,00	10,90%	10,40%

Fonte: O autor (2020)

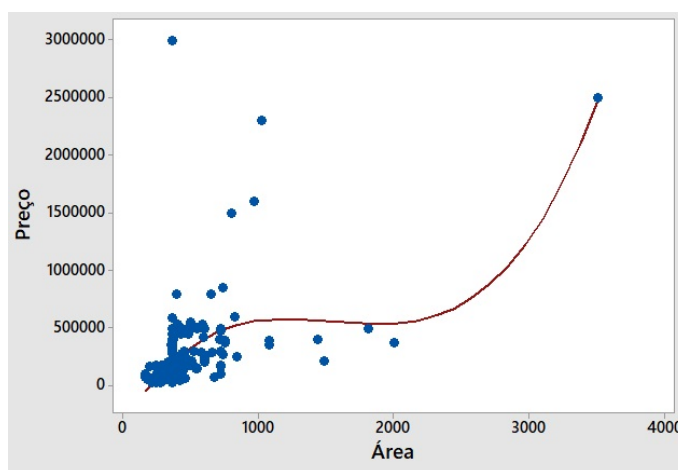
É válido destacar que o termo “S”, presente na Tabela 7, corresponde ao erro padrão da regressão, ou seja, o termo indica a distância média que os dados estão da linha de regressão (BLOG DA MINITAB, 2019b). Outrossim, descrevendo os modelos de regressão ajustados para a variável Área, que apresentou o maior $R^2(\text{adj})$ entre as variáveis explicativas, tem-se na Equação (4.4) um valor de 13.669,00 para o intercepto da variável preço, um coeficiente de efeito linear de 510,40, e de 0,02863 para o efeito quadrático. Além disso, Área^2 corresponde ao quadrado da variável Área. Na Equação (4.5), o intercepto da variável preço é - 331.839,00, e os coeficientes são 1.915,00 (linear), - 1,297 (quadrático) e 0,000279 (cúbico).

$$\text{Preço} = 13.669,00 + 510,40 \text{ Área} + 28,63 \times 10^{-3} \text{ Área}^2 \quad (4.4)$$

$$\text{Preço} = -331.839,00 + 1.915,00 \text{ Área} - 1,30 \text{ Área}^2 + 2,79 \times 10^{-4} \text{ Área}^3 \quad (4.5)$$

A Figura 2 mostra o comportamento cúbico da variável Área. Ao comparar os modelos, o ajuste cúbico proporcionou um aumento do coeficiente $R^2(\text{adj})$, conseguindo explicar 32,80% das alterações dos preços. Esse aumento pode estar relacionado ao fato do ajuste cúbico conseguir descrever os picos e vales dos dados (SUPORTE AO MINITAB 19, 2020). O método de ajuste também foi aplicado para as demais variáveis independentes, entretanto, mesmo com a melhora considerável dos coeficientes, como observado na Tabela 8, as variáveis DC e DAv não apresentaram resultados melhores que a variável Área. Sendo assim, não foi realizada a descrição das equações. As Figuras de 5 a 9 no Apêndice A mostram curvas geradas pelo ajuste quadrático e cúbico das demais variáveis.

Figura 2 – Ajuste cúbico para variável Área



Fonte: O autor (2020)

Tabela 8 – Significância dos ajustes de regressão

Ajuste	S	R^2	$R^2(\text{adj})$
Quadrático (Área)	313.557,00	28,50%	27,70%
Cúbico (Área)	302.363,00	33,90%	32,80%
Quadrático (DC)	339.256,00	16,30%	15,40%
Cúbico (DC)	339.990,00	16,40%	15,00%
Quadrático (DAv)	336.697,00	17,60%	16,70%
Cúbico (DAv)	327.600,00	22,40%	21,10%

Fonte: O autor (2020)

Verificando a aplicação dos modelos, foi adicionado nas equações a informação sobre a área de um dos lotes coletados no estudo. O terreno teve preço anunciado de

R\$450.000,00 e 483m² de área (Área = 483). O preço encontrado na Equação (4.1) foi de R\$270.844,60, de R\$266.871,26 na (4.4), e de R\$321.967,50 na (4.5). Percebe-se que o modelo ajustado apresentou um valor mais próximo ao preço praticado. Entretanto, existem terrenos que possuem o mesmo valor para área mas com preços diferentes. Para exemplificar, considere um terreno padrão com 360m² de área. Para os 10 primeiros anúncios, foram encontrados 3 terrenos com esse valor de área. Porém, os preços praticados foram de R\$500.000,00, R\$350.000,00 e R\$45.000,00. Ou seja, só a variável área pode não ser suficiente na elaboração de um modelo que explique a precificação dos terrenos. Sendo assim, um diferencial no modelo de precificação dos lotes pode ser o acréscimo das variáveis relacionadas à localização.

4.3.1 Regressão Linear Múltipla

Visando a construção de um modelo que explique melhor as alterações dos preços dos terrenos, foi desenvolvida uma equação considerando as variáveis independentes analisadas anteriormente e que apresentaram os maiores coeficientes de determinação (R^2 e $R^2(\text{adj})$), ou seja, os atributos área, distância de carro até o centro comercial e distância da avenida/rodovia mais próxima. O nível de confiança para os intervalos permaneceu em 95%. A equação visa estimar o preço de um lote localizado em qualquer ponto do município de João Monlevade, representada pela Equação (4.6).

$$\text{Preço} = 291.038,00 + 567,40 \text{ Área} - 64.209,00 \text{ DC} - 76.297,00 \text{ DAv} \quad (4.6)$$

Como descrição da equação, tem-se que o valor do intercepto da variável preço é 291.038,00 e o valor da inclinação do preço em relação à Área é de 567,40. Já os parâmetros relacionados às distâncias, a inclinação do preço são 64.209,00 e 76.297,00 para DC e DAv, respectivamente. Nota-se, pela Tabela 9, que os valores dos coeficientes de determinação foram de 43,74% (R^2) e 42,83% ($R^2(\text{adj})$), ou seja, o modelo consegue explicar 42,83% das alterações dos preços dos loteamentos. Realizando uma comparação entre os modelos, observou-se que o modelo múltiplo apresentou o maior $R^2(\text{adj})$. Sendo assim, o modelo de regressão linear com três variáveis independentes apresentou uma capacidade de explicação maior que os modelos lineares simples e ajustados.

Tabela 9 – Significância do modelo de regressão múltipla

S	R^2	$R^2(\text{adj})$	$R^2(\text{pred})$
278.892,00	43,74%	42,83%	39,33%

Fonte: O autor (2020)

Uma vez construído o modelo, é importante identificar as variáveis significativas para a equação. Para realizar o teste, foi considerado o nível de significância de 0,05. Ao

observar o valor-p na Tabela 10, nota-se que somente DAv (0,128) apresentou um valor maior que 0,05, ou seja, a variável DAv não é significativa para o modelo representado pela Equação (4.6).

Tabela 10 – Coeficientes do modelo de regressão múltipla

Termo	Coeficiente	SE Coeficiente	valor-t	valor-p	VIF
Constante	291.038,00	55.629,00	5,23	0,000	-
Área	567,40	61,80	9,18	0,000	1,04
DC	-64.209,00	11.556,00	-5,56	0,000	1,24
DAv	-76.297,00	49.909,00	-1,53	0,128	1,28

Fonte: O autor (2020)

Objetivando verificar a aplicação do modelo, considerou-se as informações do terreno utilizado no exemplo do modelo simples e ajustado. O lote está localizado a 1,90 Km do centro comercial da cidade (DC = 1,90) e 0,40 Km de uma avenida/rodovia (DAv = 0,40). Ele tem uma área de 483m² (Área = 483) e o seu valor é R\$450.000,00. Ao inserir os dados do terreno na Equação (4.6), o modelo estimou o valor R\$412.576,30. Outrossim, foi considerado um outro lote, com um preço anunciado de R\$130.000,00, 360m² de área, e localizado a 5,40 Km do centro comercial da cidade e 0,50 Km de uma avenida/rodovia. O modelo retornou o valor de R\$110.424,90. Ou seja, os valores retornados pelo modelo proposto no estudo são aproximados aos preços anunciados.

Outro método utilizado para testar a aplicação do modelo foi observado na pesquisa de Rosa, Oliveira e Pinto (2018b), em que se definiram uma configuração para um imóvel, estimaram um valor médio, e posteriormente observaram os bairros que apresentaram valores próximos. Sendo assim, foram considerados valores aleatórios para características utilizadas na precificação do terreno. O terreno fictício possui as seguintes características: localizado a 0,50 Km do centro comercial da cidade e a 0,20 Km da avenida, e possui uma área de 500,00m². Utilizando o modelo (4.6), o valor estimado pela equação foi R\$527.374,10. Analisando somente os valores médios de cada bairro, o lote fictício apresentou similaridade com os bairros Alvorada (R\$487.500,00), JK (R\$466.667,00) e Aclimação (R\$463.333,00). Observando os bairros de maneira detalhada, observa-se que no bairro Alvorada a área média do lote é de 455,00m², com DC médio de 1,14 Km e DAv médio de 0,30 Km. No bairro JK, a área média é de 436,20m², DC de 0,97 Km e DAv de 0,25 Km. Já o lote no bairro Aclimação possui uma área média de 531,50m², DC 1,90 Km e DAv 0,39 Km. Em resumo, considerando todos os valores, nota-se que o lote fictício assemelha-se a um lote localizado no bairro Alvorada.

4.4 Análise de Agrupamento

Partindo da hipótese de que o agrupamento dos bairros que apresentam determinada semelhança pode auxiliar na formulação de equações que explicam melhor a variação dos preços dos terrenos, foi realizada uma etapa de clusterização. Assim como os modelos propostos na etapa de Análise de Regressão, o nível de confiança para os intervalos permaneceu em 95%. Para a etapa de agrupamento, foi definido como parâmetros a distância euclidiana e o método de ligação simples. Outrossim, foram gerados agrupamentos considerando três níveis de similaridade, 85%, 90% e 95%. Como a amostra do estudo engloba 189 terrenos distribuídos por 37 bairros, o agrupamento foi definido em função dos valores médios do preço, área, distância até o centro comercial de carro e distância até a avenida/rodovia mais próxima.

Observa-se, na Tabela 11, que em cada etapa são reunidos os grupos que apresentaram a menor distância (ROSA; OLIVEIRA; PINTO, 2018a). Note que alterações dos níveis de similaridade não proporcionaram uma diferença significativa no número de agrupamentos. Considerando 95% de similaridade, etapa 33 do agrupamento, são identificados 4 conglomerados. Já no agrupamento, considerando o nível de 90%, observa-se que o número de *clusters* formados reduz para 3.

É possível notar também que nos dois níveis, o Grupo 27 (República) permaneceu isolado, ou seja, considerando determinadas características e o grau de similaridade de 95% e 90%, o bairro República não apresentou similaridade mínima com os outros bairros do município. Ao formular a equação para o Grupo 27 considerando todas as variáveis, observou-se que não foram disponibilizadas informações sobre o valor-p pelo *software*, dificultando a identificação das variáveis significativas para o modelo. Considerando tais informações, a equação formulada para o Grupo 27 não foi apresentada no estudo. Sendo assim, foram formuladas e apresentadas somente as equações para estimar o preço dos terrenos para os conglomerados formados no nível de semelhança de 85%.

No nível de similaridade de 85%, foram necessárias 35 etapas para que todos os 37 bairros estivessem agrupados em 2 novos *clusters*. O Grupo 1 e o Grupo 27 se aglomeraram e formaram um novo *clusters* com 35 bairros. Os elementos 7 e 29 correspondem aos outros 2 bairros da cidade, Carneirinhos e Santa Bárbara respectivamente, e integram o Grupo 7. Observou-se que mesmo realizando as alterações dos níveis de similaridade, os elementos 7 e 29 permaneceram agrupados. O agrupamento reflete como o preço médio dos membros do Grupo 7, Carneirinhos (R\$1.035.000,00) e Santa Bárbara (R\$1.071.875,00), são consideravelmente maiores quando comparados aos outros grupos. Reforçando essa informação, o nível de similaridade necessário para o agrupamento do Grupo 7 com o Grupo 1 é de aproximadamente 59%, como pode ser observado na Figura 10 (Apêndice (A)). Isso comprova a diferença entre esses dois bairros com os demais.

Tabela 11 – Etapas do agrupamento para os valores médios

Etapa	N	Similaridade	Distância	Cluster Agrupado	Novo Cluster	Nº Obs.
1	36	99,9970	30	31 - 37	31	2
2	35	99,8955	1073	24 - 26	24	2
3	34	99,8904	1125	12 - 21	12	2
4	33	99,8756	1278	9 - 34	9	2
5	32	99,8127	1923	8 - 36	8	2
6	31	99,7072	3007	5 - 32	5	2
7	30	99,6752	3335	1 - 13	1	2
8	29	99,6429	3667	9 - 28	9	3
9	28	99,5915	4195	5 - 33	5	3
10	27	99,5131	5000	4 - 31	4	3
11	26	99,5131	5000	11 - 30	11	2
12	25	99,5131	5000	4 - 11	4	5
13	24	99,5130	5000	3 - 22	3	2
14	23	99,4026	6135	8 - 9	8	5
15	22	99,2687	7509	5 - 25	5	4
16	21	99,2396	7808	4 - 5	4	9
17	20	99,2001	8214	8 - 15	8	6
18	19	99,1925	8292	6 - 18	6	2
19	18	99,1885	8333	16 - 23	16	2
20	17	99,1072	9168	14 - 17	14	2
21	16	99,0262	10000	3 - 4	3	11
22	15	99,0257	10005	6 - 19	6	3
23	14	99,0200	10064	10 - 12	10	3
24	13	98,3941	16490	24 - 35	24	3
25	12	98,3722	16716	6 - 20	6	4
26	11	97,9712	20833	1 - 2	1	3
27	10	97,8898	21669	6 - 16	6	6
28	9	97,8783	21787	8 - 14	8	8
29	8	97,1271	29501	3 - 8	3	19
30	7	96,5916	35000	6 - 10	6	9
31	6	96,4087	36878	7 - 29	7	2
32	5	96,1133	39912	3 - 6	3	28
33	4	95,1282	50027	3 - 24	3	31
34	3	93,5890	65833	1 - 3	1	34
35	2	87,7460	125833	1 - 27	1	35
36	1	58,9369	421667	1 - 7	1	37

Fonte: O autor (2020)

A Equação (4.7) representa o modelo reformulado para o Grupo 7, alcançando os valores de 42,23% para R^2 e 24,90% para $R^2(\text{adj})$. Ao analisar o valor-p, Tabela 12, observou-se que todas as variáveis apresentaram um valor maior que 0,05, ou seja, Área, DC e DAv, não são significativas para o modelo. Em relação ao Grupo 1, composto pelos demais bairros do município, foi estimada a Equação (4.8). Como resultado, o modelo apresentou

um coeficiente de determinação de 50,77%, e um coeficiente de determinação ajustado de 49,91%. Em relação ao valor-p dos coeficientes, somente DAv (0,653) apresentou um valor superior a 0,05, ou seja, é a única variável não significativa para o modelo. Sendo assim, as variáveis Área e DC conseguem explicar 49,91% da variação dos preços para os bairros do Grupo 1.

$$\text{Preço} = 1.055.740,00 + 590 \text{ Área} - 171.373,00 \text{ DC} - 1.594.967 \text{ DAv} \quad (4.7)$$

$$\text{Preço} = 233.630,00 + 324,90 \text{ Área} - 42.932,00 \text{ DC} - 9.205,00 \text{ DAv} \quad (4.8)$$

Tabela 12 – Coeficientes do modelo de regressão - Grupo 7

Termo	Coeficiente	SE Coeficiente	valor-t	valor-p	VIF
Constante	1.055.740,00	681.471,00	1,55	0,152	-
Área	590,00	316,00	1,86	0,092	1,30
DC	-171.373,00	448.623,00	-0,38	0,710	1,55
DAv	-1.594.967,00	1.514.520,00	-1,05	0,317	1,46

Fonte: O autor (2020)

Realizando um comparativo entre os modelos estimados na etapa de agrupamento e o modelo de regressão múltipla (Equação (4.6)), nota-se que a equação proposta para o Grupo 1 apresentou um aumento dos coeficientes de determinação (R^2) e de determinação ajustado ($R^2(\text{adj})$). Entretanto, os resultados obtidos para o Grupo 7 apresentaram uma redução dos coeficientes, além de indicar que as variáveis analisadas não são significativas para o modelo. Considerando esses aspectos, o modelo de regressão linear múltipla, representado pela Equação (4.6), torna-se mais adequado para estimar o preço de um lote localizado no município.

4.5 Análise Fatorial

Como última etapa do estudo, uma análise fatorial para estimar novos valores para o agrupamento dos bairros foi realizada. Para essa etapa, foi considerada todas as variáveis coletadas no estudo, ou seja, preço, área, distância do terreno até o centro comercial de carro (DC), a pé (DP), em linha reta (DLR), e distância até avenida/rodovia mais próxima (DAv). A Tabela 13 apresenta os 6 fatores para determinar as cargas fatoriais. Outrossim, foi considerado o método dos componentes principais, o qual a variância é igual o autovalor (SUPPORTE AO MINITAB 18, 2019b).

Também foi calculado a variância e a variabilidade (%) dos dados, disponibilizados na Tabela 14. Essas informações são importantes, e foram utilizadas como critérios na definição do número de fatores necessários para explicar todas as variáveis. O Fator 1, por

Tabela 13 – Cargas fatoriais

Atributo	Fator 1	Fator 2	Fator 3	Fator 4	Fator 5	Fator 6
Preço	-0,549	-0,679	-0,182	0,453	-0,003	0,001
Área	-0,147	-0,900	-0,115	-0,395	-0,002	-0,001
DLR	0,958	-0,231	0,107	0,075	0,108	-0,025
DC	0,962	-0,199	0,141	0,071	-0,094	-0,038
DP	0,965	-0,216	0,116	0,062	-0,013	0,064
DAv	0,606	0,194	-0,771	-0,015	-0,003	-0,001

Fonte: O autor (2020)

exemplo, apresentou um autovalor de 3,4633 e uma variação dos dados de 57,70%. Já o Fator 6 consegue explicar somente 0,10% da variação dos atributos.

Tabela 14 – Variância e variabilidade dos dados

	Fator 1	Fator 2	Fator 3	Fator 4	Fator 5	Fator 6
Variância	3,4633	1,4471	0,6866	0,3760	0,0208	0,0062
Variabilidade (%)	0,5770	0,2410	0,1140	0,0630	0,0030	0,0010

Fonte: O autor (2020)

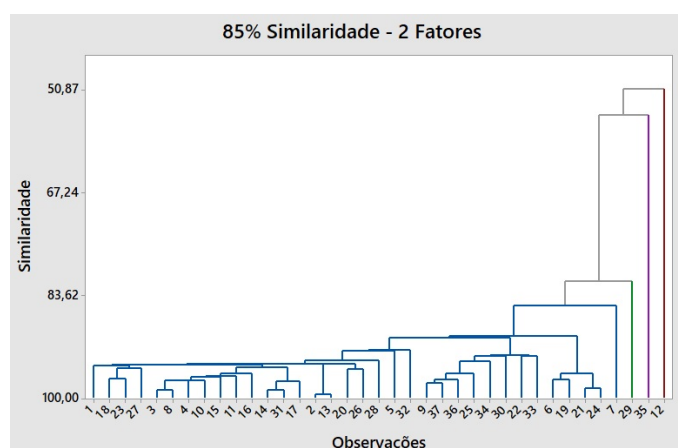
4.5.1 Critério de Kaiser

Para determinar o número de fatores utilizados no agrupamento, o primeiro critério seguido foi o de Kaiser. Sendo assim, foram selecionados os fatores que apresentaram um autovalor superior a 1, o Fator 1 (3,4633) e Fator 2 (1,4471). Como é possível observar na Tabela 13, as cargas fatoriais indicam a correlação do fator com o atributo. Nota-se que o Fator 1 é referente a localização do terreno, porque apresentou uma forte correlação com as variáveis DLR (0,958), DC (0,962) e DP (0,965). Já o Fator 2 está mais relacionado com a característica estrutural do terreno, visto que apresenta uma forte correlação negativa com a área (-0,900), e também com o preço (-0,679).

Em seguida, foram calculados pelo *software* os escores dos dois fatores para cada terreno. Para realizar o agrupamento, conforme ilustrado pela Figura 3, foram selecionados os valores dos escores médios, do Fator 1 e Fator 2, para cada bairro, e os parâmetros da distância euclidiana e do método de ligação simples. Ademais, foi considerado o nível de significância de 85%, visto que foi o grau de similaridade utilizado na formulação das equações na etapa de Análise de Agrupamento.

Observa-se na Figura 3 que o número de *clusters* formados, considerando os 2 fatores, foi de 4 conglomerados (Grupo 1, Grupo 29, Grupo 35 e Grupo 12). O Grupo 12 (Jacuí) é o que apresenta menor similaridade com os demais grupos, com um valor alto para o Fator 1 (3,8595). Ao analisar a informação do único lote anunciado no bairro, notou-se

Figura 3 – 2 Fatores - 85% Similaridade



Fonte: O autor (2021)

que ele está localizado distante do centro comercial do município, com os valores de 7,70 Km (DLR), 10,00 Km (DC) e 9,60 Km (DP). O terreno possui uma área notável, entretanto o seu valor é baixo considerando a área total. Outro elemento isolado no agrupamento é o bairro Santa Bárbara (Grupo 29), um dos bairros centrais do município. O bairro apresentou como valores médios os escores -1,2830 (Fator 1) e -1,7500 (Fator 2). A medida do Fator 1 indica que o bairro, em média, apresentou valor baixo para localização, ou seja, os terrenos não estão distantes do centro comercial. O valor médio da variável DC para o bairro, por exemplo, foi de 2,03 Km.

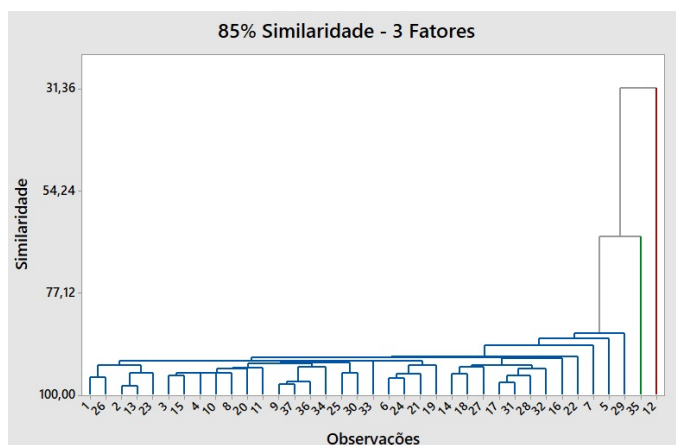
4.5.2 Critério da Variância Total

Outro critério seguido no estudo foi a seleção de fatores que explicam um valor da variância total. Foi especificado o valor mínimo de 90%, portanto foram selecionados o Fator 1 (57,70%), Fator 2 (24,10%) e Fator 3 (11,40%), que juntos explicam 93,20% da variação. Como foi explicado anteriormente, o Fator 1 tem relação com a localização e o Fator 2 com o aspecto estrutural. O Fator 3, assim como o Fator 1, está relacionado ao aspecto da localização, já que apresenta uma forte correlação negativa com a variável DAv (-0,771).

Assim como no agrupamento considerando o critério de Kaiser, foram calculados os escores dos três fatores para cada terreno e selecionado os valores médios. Observa-se, na Figura 4, que o acréscimo do Fator 3 gerou uma alteração mínima na formação dos *clusters*, que foi a associação do Grupo 1 com o Grupo 29 (Santa Bárbara). Sendo assim, foram formados 3 conglomerados considerando o nível de significância de 85%.

Para facilitar a análise do Grupo 1, composto por 35 bairros, foi observado o valor máximo e mínimo dos escores médios. Ou seja, foram selecionados os escores médios dos fatores (Fator 1, Fator 2 e Fator 3) de cada bairro presente no Grupo 1, e posteriormente

Figura 4 – 3 Fatores - 85% Similaridade



Fonte: O autor (2021)

foram calculados os valores máximo e mínimo de cada fator. Por exemplo, o menor escore médio encontrado para o Fator 1 foi -1,6280 (Carneirinhos), e o maior foi 1,8632 (Novo Cruzeiro). O Fator 1 está relacionado à localização dos bairros até o centro comercial, sendo assim, observou-se que os bairros ficam, em média, 1,32 Km (Carneirinho) e 7,47 Km (Novo Cruzeiro) do centro comercial de carro (DC). Com relação à distância até avenida/rodovia mais próxima, Fator 3, o menor valor encontrado foi do bairro Lourdes (-1,8640), e o maior foi Teresópolis (0,9954). Os bairros ficam, em média, 1,13 Km (Lourdes) e 0,4875 Km (Teresópolis) de alguma avenida/rodovia.

Vale destacar que ocorreu um distanciamento entre o bairro Jacuí (Grupo 12) e os demais grupos analisados, visto que ele só é agrupado ao considerar um nível de semelhança baixo, de 31,36%. O outro conglomerado formado, Grupo 35 (Tanquinho), por exemplo, apresentou um escore médio de 0,6581 para o Fator 1. Existe somente um lote anunciado no bairro, como ocorreu no bairro Jacuí (Grupo 12), e ele está localizado distante do centro comercial do município, com os valores de 4,62 Km (DLR), 6,30 Km (DC) e 5,90 Km (DP).

5 Conclusão

O presente trabalho objetivou uma análise e compreensão sobre a oferta de terrenos urbanos (lotes) dentro do mercado imobiliário da cidade João Monlevade. Foram identificados os fatores que impactam nos preços dos lotes e proposto diferentes modelos de regressão que utilizassem tais informações. Ademais, foram realizadas diferentes análises de agrupamento visando a melhoria da qualidade dos modelos.

Como resultado foi verificado que a variável Área apresentou o maior grau de correlação positiva com o preço dos lotes. Por outro lado, DAV apresentou o menor grau de correlação. No que se refere aos modelos, observou-se que os ajustes realizados foram positivos e proporcionaram uma melhora na estimativa dos preços. Entretanto, foi constatado que o uso de somente uma variável explicativa para predição do preço não é adequada, uma vez que o modelo de regressão múltipla possibilitou a obtenção de um coeficiente de significância maior. Ademais, verificou também que a distância do terreno até uma avenida/rodovia não foi uma característica significativa nos modelos.

Quanto aos agrupamentos dos bairros, observou-se um número reduzido de conglomerados mesmo com as alterações dos níveis de similaridade. Em geral, os bairros centrais, e que apresentavam preços de terrenos elevados, ficavam isolados ou agrupados entre si. A formulação de modelos para estimar os preços dos bairros considerando suas semelhanças não demonstrou ser adequada, visto que determinados grupos não possibilitaram melhorias nos modelos.

Uma limitação do estudo foi a ausência da análise dos pressupostos para as regressões, como o teste da normalidade e da variância constante. Por fim, para trabalhos futuros, sugere-se identificar e analisar outros atributos relacionados aos terrenos urbanos, como a superfície dos lotes e as medidas de largura e comprimento. A proposição e análise de outros modelos não lineares, assim como a realização de um trabalho que consiga analisar de maneira conjunta os terrenos e os imóveis residenciais do município também são alternativas para continuidade da pesquisa.

Referências

- AGUIRRE, L. d. L. Modelos de Precificação: uma aplicação no setor imobiliário do DF. *Odds&Actions Inteligência Analítica*, Brasília, p. 21, dez. 2012. Citado 3 vezes nas páginas 11, 12 e 20.
- ALMEIDA, P. H. R. d. *Fatores determinantes para a formação de preço no mercado imobiliário de Brasília*. 42 p. Monografia (Administração) — Universidade de Brasília, Brasília, 2011. Disponível em: <<http://bdm.unb.br/handle/10483/2122>>. Acesso em: 14 nov. 2019. Citado 2 vezes nas páginas 15 e 20.
- ALVES, V. *Avaliação de imóveis urbanos baseada em métodos estatísticos multivariados*. 133 p. Dissertação (Mestrado em Ciências) — Universidade Federal do Paraná, Campo Mourão, 2005. Disponível em: <<https://hdl.handle.net/1884/3234>>. Acesso em: 1 jan. 2020. Citado na página 17.
- BAKKE, H. A.; LEITE, A. S. de M.; SILVA, L. B. da. Estatística multivariada: aplicação da análise fatorial na engenharia de produção. *Revista Gestão Industrial*, Paraná, PR, v. 4, n. 4, p. 01–14, 2008. ISSN 1808-0448. Disponível em: <<https://core.ac.uk/reader/193036390>>. Acesso em: 02 jan. 2020. Citado na página 19.
- BAPTISTELLA, M.; STEINER, M. T. A.; NETO, A. C. O uso de redes neurais e regressão linear múltipla na engenharia de avaliações: Determinação dos valores venais de imóveis urbanos. In: SIMPÓSIO BRASILEIRO PESQUISA OPERACIONAL. *Anais do XXXVIII SBPO - Simpósio Brasileiro Pesquisa Operacional*. Goiânia, GO, 2006. Citado na página 19.
- BARROSO, L. P.; ARTES, R. *Análise multivariada*. Lavras, 2003. 151 p. Citado na página 18.
- BLOG DA MINITAB. *Ajuste de curva com regressão linear e não-linear*. 2019. Disponível em: <<https://blog.minitab.com/pt/ajuste-de-curva-com-regressao-linear-e-nao-linear>>. Acesso em: 24 nov. 2020. Citado na página 28.
- BLOG DA MINITAB. *Análise de regressão: Como interpretar S, o erro padrão da regressão*. 2019. Disponível em: <<https://blog.minitab.com/pt/analise-de-regressao-como-interpretar-s-o-erro-padrao-da-regressao>>. Acesso em: 10 fev. 2021. Citado na página 28.
- BRÁULIO, S. N. *Proposta de uma metodologia para avaliação de imóveis urbanos baseado em métodos estatísticos multivariados*. 159 p. Dissertação (Mestrado em Ciências) — Universidade Federal do Paraná, Curitiba, 2005. Disponível em: <<https://hdl.handle.net/1884/1895>>. Acesso em: 15 nov. 2019. Citado 4 vezes nas páginas 16, 18, 19 e 20.
- CÂMARA BRASILEIRA DA INDÚSTRIA DA CONSTRUÇÃO. *Indicadores Imobiliários Nacionais*. 2019. Disponível em: <<https://cbic.org.br/wp-content/uploads/2019/08/2%C2%BA-Trimestre-2019-Indicadores-Imobili%C3%A1rios-Nacionais-APRESENTA%C3%87%C3%83O.pdf>>. Acesso em: 07 nov. 2019. Citado na página 13.

- CARMO, C. R. S. Precificação imobiliária baseada em modelagem hedônica e externalidades: um estudo aplicado a terrenos urbanos. *Revista da Faculdade de Administração e Economia*, v. 5, n. 2, p. 2–23, 2014. ISSN 2176-9583. Disponível em: <<https://www.metodista.br/revistas/revistas-metodista/index.php/ReFAE/article/view/4896>>. Acesso em: 13 nov. 2019. Citado 2 vezes nas páginas 11 e 20.
- DANTAS, R. A.; MAGALHÃES, A. M.; VERGOLINO, J. R. d. O. Avaliação de imóveis: a importância dos vizinhos no caso de Recife. *Economia Aplicada*, Ribeirão Preto, v. 11, n. 2, p. 231–251, abr./jun. 2007. Disponível em: <<http://dx.doi.org/10.1590/S1413-80502007000200004>>. Acesso em: 14 nov. 2019. Citado 2 vezes nas páginas 11 e 15.
- DANTAS, R. A.; MAGALHÃES, A. M.; VERGOLINO, J. R. d. O. Um modelo espacial de demanda habitacional para a cidade do Recife. *Estudos Econômicos*, São Paulo, v. 40, n. 4, p. 891–916, out./dez. 2010. Disponível em: <<http://dx.doi.org/10.1590/S0101-41612010000400006>>. Acesso em: 14 nov. 2019. Citado na página 15.
- FÁVERO, L. P.; BELFIORE, P. *Manual de Análise de Dados - Estatística e Modelagem Multivariada com Excel[®], SPSS[®] e Stata[®]*. 1. ed. Rio de Janeiro: Elsevier, 2017. 498–595 p. Citado na página 17.
- FÁVERO, L. P. L.; BELFIORE, P. P.; LIMA, G. A. Modelos de precificação hedônica de imóveis residenciais na região metropolitana de São Paulo: uma abordagem sob as perspectivas da demanda e da oferta. *Estudos Econômicos*, São Paulo, v. 38, n. 1, p. 73–96, jan./mar. 2008. Disponível em: <<http://dx.doi.org/10.1590/S0101-41612008000100004>>. Acesso em: 14 nov. 2019. Citado na página 20.
- FIGUEIREDO FILHO, D. B.; SILVA JÚNIOR, J. A. d. Desvendando os mistérios do coeficiente de correlação de pearson (r). *Revista Política Hoje*, v. 18, n. 1, 2009. Citado na página 16.
- FOLHA DE SÃO PAULO. *Projetos arquitetônicos diferenciados são a recompensa pelos gastos maiores com fundações e nivelamento*. São Paulo, 2002. Disponível em: <<https://www1.folha.uol.com.br/fsp/construcao/cs2402200204.htm>>. Acesso em: 13 nov. 2019. Citado na página 11.
- FREI, F. *Introdução à análise de agrupamentos: Teoria e Prática*. [S.l.]: UNESP, 2006. Citado na página 18.
- GAZOLA, S. *Construção de um modelo de regressão para avaliação de imóveis*. 104 p. Dissertação (Mestrado em Engenharia de Produção.) — Universidade Federal de Santa Catarina, Florianópolis, 2002. Disponível em: <<http://repositorio.ufsc.br/xmlui/handle/123456789/82455>>. Acesso em: 23 nov. 2020. Citado na página 16.
- HAIR, J. F. et al. *Análise multivariada de dados*. 6. ed. São Paulo: Bookman Editora, 2009. Citado 3 vezes nas páginas 18, 19 e 20.
- HENRIQUES, C. *Análise de Regressão Linear Simples e Múltipla*. Portugal, 2011. Departamento de Matemática. Escola Superior de Tecnologia de Viseu. Citado na página 17.

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, abr./jun. 2005. Disponível em: <<http://dx.doi.org/10.1590/S0101-41612005000200001>>. Acesso em: 14 nov. 2019. Citado na página 15.

IBGE - INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *IBGE / Cidades@ / Minas Gerais / João Monlevade / Panorama*. 2018. Disponível em: <<https://cidades.ibge.gov.br/brasil/mg/joao-monlevade/panorama>>. Acesso em: 09 fev. 2021. Citado na página 11.

IBGE - INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *IBGE / Cidades@ / Minas Gerais / João Monlevade / Pesquisa / Cadastro Central de Empresas*. 2018. Disponível em: <<https://cidades.ibge.gov.br/brasil/mg/joao-monlevade/pesquisa/19/29761>>. Acesso em: 09 fev. 2021. Citado na página 11.

KASZNAR, I. K.; GONÇALVES, B. M. L. Regressão múltipla: uma digressão sobre seus usos. *IBCI*, Rio de Janeiro, RJ, 2011. Citado na página 16.

LEVINE, D. M.; STEPHAN, D. F.; SZABAT, K. A. *Estatística - teoria e aplicações usando o Microsoft® Excel em português*. 7. ed. Rio de Janeiro: LTC, 2017. 4–535 p. Citado 3 vezes nas páginas 16, 17 e 18.

MATOS, D.; BARTKIW, P. I. N. *Introdução ao mercado imobiliário*. Curitiba, 2013. e-Tec Brasil. Citado na página 15.

MIGUEL, P. A. C. et al. *Metodologia de Pesquisa em Engenharia de Produção e Gestão de Operações*. Rio de Janeiro: Campus. 3. ed. [S.l.]: Elsevier, 2012. Citado na página 22.

MINGOTI, S. A. *Análise de Dados Através de métodos de Estatística multivariada: uma abordagem aplicada*. Belo Horizonte: UFMG, 2005. 295 p. Citado 2 vezes nas páginas 18 e 19.

MONTGOMERY, D. C.; RUNGER, G. C. *Estatística Aplicada e Probabilidade Para Engenheiros*. 6. ed. Rio de Janeiro: LTC, 2018. 45–361 p. Citado 3 vezes nas páginas 16, 17 e 18.

PINDYCK, R. S.; RUBINFELD, D. L. *Microeconomia*. 7. ed. São Paulo: Pearson, 2010. 3–15 p. Citado na página 15.

PINTO, V. H. L.; FERNANDES, R. A. S. Análise de preços hedônicos no mercado imobiliário residencial de Conselheiro Lafaiete, MG. *Interações*, Campo Grande, v. 20, n. 2, p. 627–643, abr./jun. 2019. Disponível em: <<http://dx.doi.org/10.20435/inter.v20i2.1788>>. Acesso em: 14 nov. 2019. Citado 3 vezes nas páginas 11, 16 e 20.

RIBEIRO, W. C.; ROSA, V. S.; OLIVEIRA, P. B. d. Estudo de variáveis que impactam na formação dos preços de compra, venda e aluguéis residenciais em João Monlevade-MG. In: CONGRESSO BRASILEIRO DE ECONOMIA. *Anais do XXII CBE - Congresso Brasileiro de Economia*. Belo Horizonte, MG, 2017. Citado 3 vezes nas páginas 11, 15 e 21.

RODRIGUES, W. C. *Metodologia científica*. Paracambi, 2007. Faetec/IST. Citado na página 22.

ROSA, V. S. *Estudo sobre o mercado imobiliário na cidade de João Monlevade-MG*. 69 p. Monografia (Engenharia de Produção) — Universidade Federal de Ouro Preto, João Monlevade, 2018. Disponível em: <<http://www.monografias.ufop.br/handle/35400000/1710>>. Acesso em: 12 nov. 2019. Citado 4 vezes nas páginas 13, 21, 22 e 25.

ROSA, V. S.; OLIVEIRA, P. B. d.; PINTO, R. L. M. Aplicação de técnicas de análise cluster no setor imobiliário em uma cidade do interior de Minas Gerais. In: CONGRESSO BRASILEIRO DE ENGENHARIA DE PRODUÇÃO. *Anais do VIII ComBRepro - Congresso Brasileiro de Engenharia de Produção*. Ponta Grossa, PR, 2018. ISSN 2237-6143. Citado 4 vezes nas páginas 12, 19, 21 e 32.

ROSA, V. S.; OLIVEIRA, P. B. d.; PINTO, R. L. M. Modelos de precificação para locação e venda de imóveis residenciais na cidade de João Monlevade MG via regressão linear multivariada. In: SIMPÓSIO DE ENGENHARIA DE PRODUÇÃO. *Anais do XXV SIMPEP - Simpósio de Engenharia de Produção*. Bauru, SP, 2018. ISSN 1809-7189. Citado 4 vezes nas páginas 12, 17, 21 e 31.

STEINER, M. T. A. et al. Métodos estatísticos multivariados aplicados à engenharia de avaliações. *Gest. Prod.*, São Carlos, v. 15, p. 23–32, jan./abr. 2008. Disponível em: <<http://dx.doi.org/10.1590/S0104-530X2008000100004>>. Acesso em: 14 nov. 2019. Citado na página 19.

SUPORTE AO MINITAB 18. *Interpretar os principais resultados para Correlação*. 2019. Disponível em: <<https://support.minitab.com/pt-br/minitab/18/help-and-how-to/statistics/basic-statistics/how-to/correlation/interpret-the-results/key-results/#step-2-determine-whether-the-correlation-coefficient-is-significant>>. Acesso em: 10 fev. 2021. Citado na página 27.

SUPORTE AO MINITAB 18. *Interpretar todas as estatísticas e gráficos para Análise fatorial*. 2019. Disponível em: <<https://support.minitab.com/pt-br/minitab/18/help-and-how-to/modeling-statistics/multivariate/how-to/factor-analysis/interpret-the-results/all-statistics-and-graphs/#score-plot>>. Acesso em: 27 nov. 2020. Citado 2 vezes nas páginas 20 e 34.

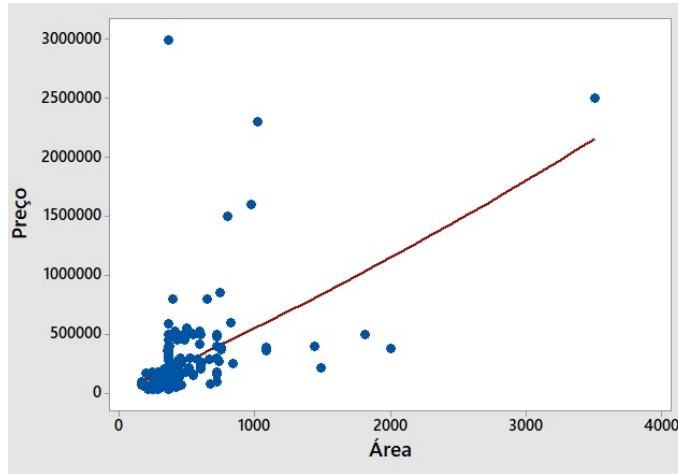
SUPORTE AO MINITAB 18. *Tipos de análises de regressão*. 2019. Disponível em: <<https://support.minitab.com/pt-br/minitab/18/help-and-how-to/modeling-statistics/regression/supporting-topics/basics/types-of-regression-analyses/>>. Acesso em: 08 fev. 2021. Citado na página 17.

SUPORTE AO MINITAB 19. *Linhas de regressão ajustadas*. 2020. Disponível em: <<https://support.minitab.com/pt-br/minitab/19/help-and-how-to/graphs/general-graph-options/labels-lines-and-annotations/fitted-regression-lines/>>. Acesso em: 28 nov. 2020. Citado na página 29.

TURRIONI, J. B.; MELLO, C. H. P. Metodologia de pesquisa em Engenharia de Produção: Estratégias, métodos e técnicas para condução de pesquisas quantitativas e qualitativas. Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Itajubá - UNIFEI, 2012. Citado na página 22.

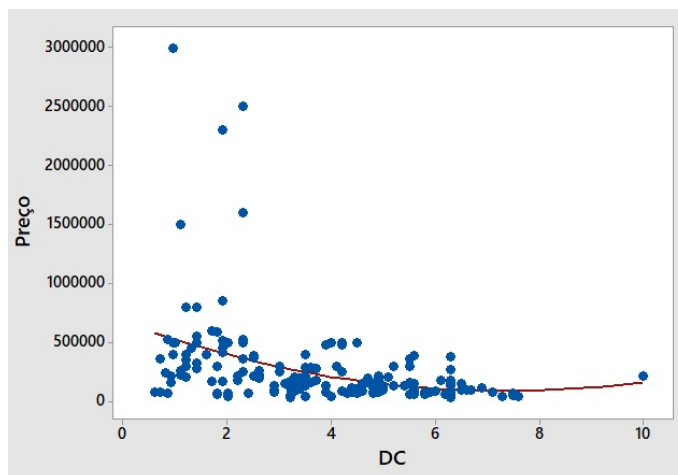
APÊNDICE A – Materiais elaborados pelo autor

Figura 5 – Ajuste quadrático para variável Área



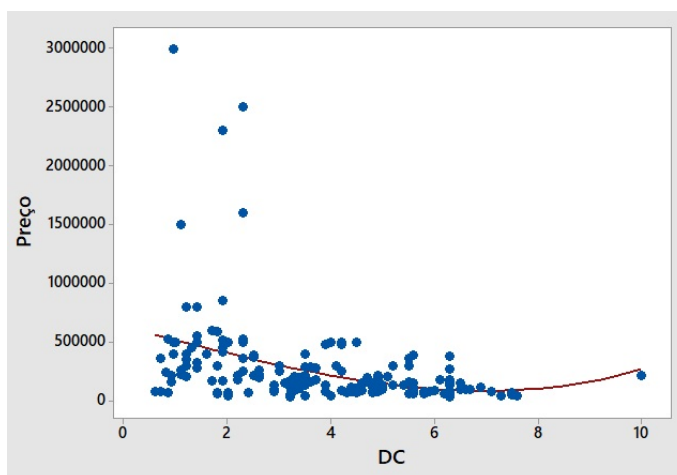
Fonte: O autor (2020)

Figura 6 – Ajuste quadrático para variável DC



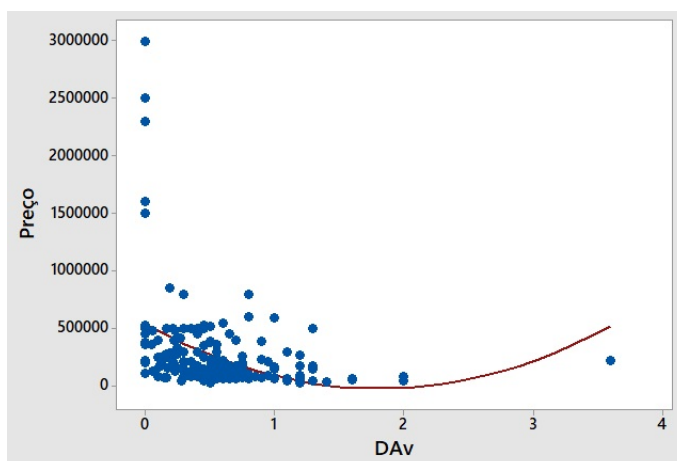
Fonte: O autor (2020)

Figura 7 – Ajuste cúbico para variável DC



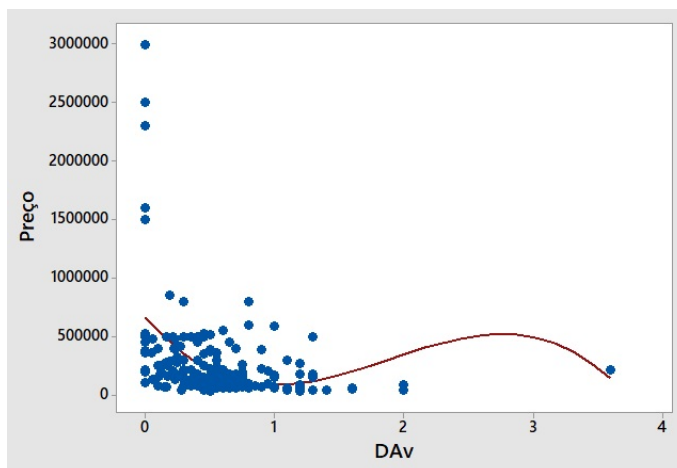
Fonte: O autor (2020)

Figura 8 – Ajuste quadrático para variável DAv



Fonte: O autor (2020)

Figura 9 – Ajuste cúbico para variável DAv



Fonte: O autor (2020)

Tabela 15 – Dados sobre os preços

Bairro	Preços (R\$)		
	Máximo	Médio	Mínimo
Aclimação	520.000,00	463.333,00	420.000,00
Alvorada	800.000,00	487.500,00	300.000,00
Bau	45.000,00	-	-
Belmonte	80.000,00	65.000,00	50.000,00
Boa Vista	130.000,00	87.000,00	44.000,00
Campos Elísios	500.000,00	281.714,00	100.000,00
Carneirinhos	3.000.000,00	1.035.000,00	360.000,00
Cidade Nova	260.000,00	138.077,00	30.000,00
Cruzeiro Celeste	180.000,00	130.667,00	95.000,00
Industrial	230.000,00	-	-
Ipiranga	70.000,00	-	-
Jacuí	220.000,00	-	-
JK	500.000,00	466.667,00	400.000,00
José Elói	550.000,00	179.167,00	45.000,00
Loanda	220.000,00	148.214,00	45.000,00
Lourdes	600.000,00	321.667,00	170.000,00
Lucília	170.000,00	-	-
Mangabeiras	500.000,00	290.000,00	210.000,00
Metalúrgico	300.000,00	-	-
Nova Aclimação	300.000,00	265.000,00	180.000,00
Nova Esperança	400.000,00	219.075,00	92.225,00
Novo Cruzeiro	60.000,00	50.000,00	40.000,00
Novo Horizonte	330.000,00	-	-
Paineiras	500.000,00	397.500,00	130.000,00
Petrópolis	100.000,00	97.500,00	95.000,00
Recanto Paraíso	530.000,00	396.429,00	200.000,00
República	800.000,00	613.333,00	450.000,00
Rosário	240.000,00	127.000,00	75.000,00
Santa Bárbara	2.500.000,00	1.071.875,00	215.000,00
Santa Cecília	80.000,00	75.000,00	70.000,00
São João	60.000,00	-	-
Satélite	100.000,00	90.000,00	80.000,00
Santo Hipólito	270.000,00	82.806,00	30.000,00
Sion	390.000,00	131.943,00	60.000,00
Tanquinho 1	380.000,00	-	-
Teresópolis	150.000,00	140.000,00	130.000,00
Vera Cruz	60.000,00	-	-

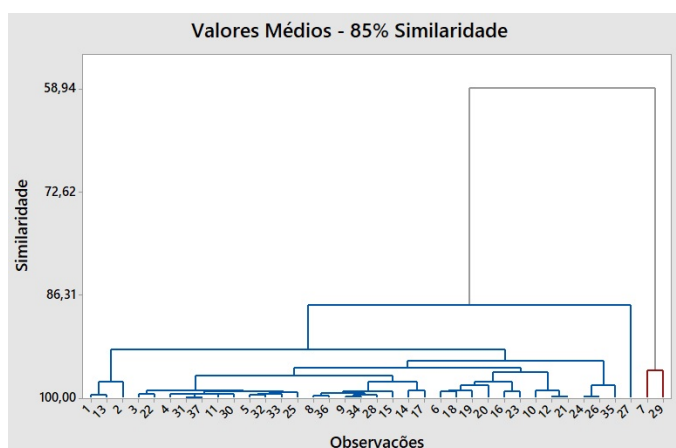
Fonte: O autor (2020)

Tabela 16 – Preços médios anunciados

Bairro	Preços (R\$)		
	Terreno	Apartamento	Casa
Aclimação	463.333,00	487.692,31	1.258.333,33
Carneirinhos	1.035.000,00	411.333,33	854.000,00
José Elói	179.167,00	270.000,00	239.250,00
Lourdes	321.667,00	276.666,67	598.157,89
Novo Horizonte	330.000,00	400.809,52	637.777,78
Paineiras	397.500,00	281.785,71	607.272,73
Santa Bárbara	1.071.875,00	370.386,96	579.888,89
Satélite	90.000,00	191.363,64	357.142,86
Sion	131.943,00	244.545,45	255.625,00

Fonte: O autor (2020)

Figura 10 – Valores médios - 85% Similaridade



Fonte: O autor (2020)



MINISTÉRIO DA EDUCAÇÃO
Universidade Federal de Ouro Preto – UFOP
Instituto de Ciências Exatas e Aplicadas - ICEA
Colegiado do Curso de Engenharia de Produção - COEP
Campus João Monlevade



TERMO DE RESPONSABILIDADE

O texto do Trabalho de Conclusão de Curso intitulado “Modelos de precificação e caracterização da oferta de lotes em uma cidade do interior de Minas Gerais” é de minha inteira responsabilidade. Declaro que não há utilização indevida de texto, material fotográfico ou qualquer outro material pertencente a terceiros sem o devido referenciamento ou consentimento dos referidos autores.

João Monlevade, 20 de fevereiro de 2021.

Roberta Pâmela Gomes de Araújo

Roberta Pâmela Gomes de Araújo