

Avaliação por regressão linear múltipla para valor de imóveis residenciais

Wilson S. Mateus^{1†}, Ana Lúcia S. S. Mateus², Júlio S. S. Bueno Filho³

¹Universidade Federal de Lavras - UFLA.

²Universidade Federal de Santa Maria - UFSM. E-mail: analucia.stat@gmail.com.

³Universidade Federal de Lavras - UFLA. E-mail: jssbueno@ufla.br.

Resumo: Este artigo tem por objetivo encontrar um modelo de regressão linear capaz de identificar variáveis que tem influência sobre o valor de imóveis residenciais. Para estimação do modelo foi considerado como variável dependente o valor do imóvel enquanto que as variáveis independentes constituíram na área do terreno, idade do imóvel, distância até o marco central da cidade, número de cômodos e número de comércios próximos ao imóvel. Os dados foram obtidos em uma imobiliária que chamaremos de “maria corporation” na cidade de Curitiba. observa-se uma correlação positiva e forte entre o valor do imóvel e a área do terreno (0,77), positiva e moderada entre área e número de cômodos (0,54) e positiva e fraca entre número de cômodos próximo ao imóvel e seu valor (0,37). Esta análise pode contribuir para que o futuro proprietário avalie se o preço praticado é um preço justo de venda.

Palavras-chave: Inferência estatística; regressão; mercado imobiliário.

Abstract: This paper aims to find a linear regression model capable of identifying variables that influence the value of residential real estate. To estimate the model, the value of the property was considered as the dependent variable, while the independent variables constituted the land area, age of the property, distance to the city center, number of rooms and number of trades near the property. The data were obtained from a real estate company that we will call “maria corporation” in the city of Curitiba. There is a positive and strong correlation between the value of the property and the land area (0.77), positive and moderate between area and number of rooms (0.54) and positive and weak correlation between number of rooms near the property and its value (0.37). This analysis can help the future owner to evaluate if the price charged is a fair selling price.

Keywords: Statistical Inference; regression; real estate market.

1. Introdução

Existem inúmeros métodos de se obter o valor de um imóvel, entretanto, segundo a norma NBR 14.653-2/2011 os métodos avaliatórios se divide em: método involutivo, método da capitalização e renda, método comparativo direto de custos, método comparativo direto de dados de mercado e método de quantificação de custos. Entre todos os métodos propostos pela NBR14.653/2011, citaremos o método comparativo direto de dados de mercado, que é aquele em que o valor do bem é estimado por meio da comparação de dados de mercado assemelhados. É condição fundamental para aplicação deste método a existência de um conjunto de dados que possa ser tomado estatisticamente como amostra do mercado imobiliário (MENDONÇA, 2001). A finalidade de avaliações de preços não se resume apenas em estimar e atribuir valor a um bem, ela deve responder questões importantes como: Quais as preferências do mercado? Quais as variáveis que interferem mais fortemente na formação do preço? (DANTAS, 1998; FIKER, 2001).

†Autor correspondente: sancheswm78@gmail.com.

Caracteriza como variável uma medida que assume valores diferentes em diferentes pontos de observação. Existem variáveis fáceis de identificar, como o número de dormitórios e a sua área. Mas existem outras mais complexas, como a localização. A NBR 14653-2011, afirma que, as variáveis podem ser do tipo dependente ou independente, onde a dependente é afetada pelas variáveis independentes.

Com isso, o presente trabalho tem como objetivo encontrar um modelo de regressão linear capaz de identificar as variáveis que tem influência sobre o valor de imóveis residenciais. Espera-se oferecer uma contribuição na área de avaliações, mais especificamente na avaliação de imóveis na cidade de Curitiba-PR, definindo algumas das variáveis que podem explicar o valor de um imóvel.

2. Material e Métodos

Os dados obtidos para a realização do trabalho são provenientes de uma imobiliária que chamaremos de “*maria corporation*” na cidade de Curitiba/PR. Para estimação do modelo foi considerado como variável dependente o valor do imóvel enquanto que as variáveis independentes constituíram na área do terreno, idade do imóvel, distância até o marco central da cidade, número de cômodos e número de comércios próximos ao imóvel.

Os modelos de regressão linear múltipla envolvem três ou mais variáveis, onde uma delas é chamada de variável dependente e as demais são chamadas de variáveis independentes. O objetivo está em obter-se um a equação do modelo que relacione o valor da variável dependente y , com todas as demais variáveis independentes x_i . A equação de regressão tem a forma:

$$y = a + b_1x_1 + b_2x_2 + \dots + b_kx_k$$

em que: a e b são denominados coeficientes de regressão.

Aplicou-se o procedimento “*stepwise*” com objetivo de selecionar as covariáveis que tem maior poder explicativo sobre o modelo. O modelo começa com todas as variáveis do conjunto e remove de forma gradativa as que são estatisticamente menos significantes. Esse processo ocorre até que as variáveis restantes sejam todas importantes (estatisticamente relevantes), ou seja, até que não haja melhora no desempenho do modelo ou não haja variáveis a serem retiradas (CHARNET et al.,1999). As pressuposições básicas do modelo clássico de regressão: normalidade, homocedasticidade e independência dos erros, foram avaliadas utilizando-se os testes de Lilliefors, Breusch-Pagan e Durbin Watson, respectivamente. Todas as análises estatísticas foram realizadas empregando as funções desenvolvidas no software livre R (R CORE TEAM, 2017).

3. Resultados e Discussão

Pela Figura 1, observa-se uma correlação positiva e forte entre o valor do imóvel e a área do terreno (0,77), positiva e moderada entre área e número de cômodos (0,54) e positiva e fraca entre número de cômodos próximo ao imóvel e seu valor (0,37). O restante das correlações é desprezível.

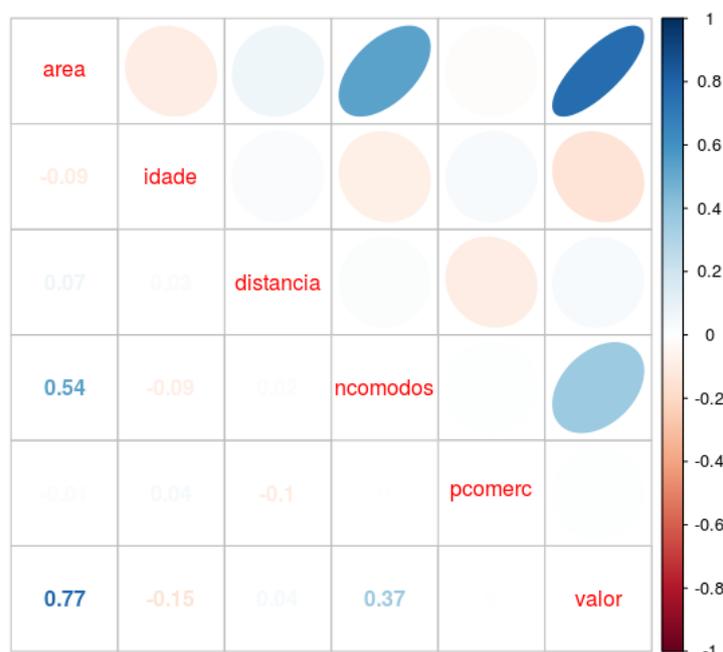


Figura 1: Matriz de coeficiente de correlação

Pelo procedimento “*stepwise*”, verificou-se que as covariáveis que exercem influência sobre a variável dependente foram: área, idade e número de cômodos do imóvel. Os resultados da análise estão na Tabela 1. Para as covariáveis idade, distância e número de cômodos obtiveram-se valores negativos para os coeficientes, o que significa que quanto maior o valor destas covariáveis, menor será o valor do imóvel. Para as covariáveis área e número de comércios próximos ao imóvel verificou-se uma relação direta com a variável dependente.

O modelo apresenta um coeficiente de determinação de cerca de 60,28%, o que significa que, 60,28% da variabilidade que ocorre nos valores dos imóveis, é explicado pela área, idade e número de cômodos do imóvel. Pela Tabela 2 observa-se que as pressuposições básicas do modelo clássico de regressão não foram violadas.

Tabela 1: Estimativas dos parâmetros no modelo clássico de regressão

Covariáveis	Coefficiente	Erro Padrão	t	Valor-p
Intercepto	48,5587	8,6622	5,606	$3,45 \times 10^{-8}$ ***
Área	1.4960	0.0631	23.689	2×10^{-16} ***
Idade	-0.7131	0.2480	-2.875	0.00422**
Distância	-0.1357	0.2973	-0.457	0.64817 ^{ns}
Ncomodos	-0.91784	0.4544	-2.020	0.04397*
Pcomerc	0,18177	0,33323	0,545	0,58568 ^{ns}

(* - significativo a 5%, ** - significativo a 1%, *** - significativo a 0,1%, ns - não significativo)

(Área - idade do terreno; Idade - idade do imóvel; Distância - distância até o marco central da cidade; Ncomodos - número de cômodos; Pcomerc - número de comércios próximos ao imóvel.

Tabela 2: Estimativas dos parâmetros no modelo clássico de regressão

Pressupostos	Nome do teste	Estatística do teste	Valor-p
Normalidade	Lilliefors	0,0238	0,7013
Homocedasticidade	Breusch-Pagan	1,7537	0,8821
Independência	Durbin Watson	1,9843	0,4287

4. Conclusão

Com o ajuste do modelo de regressão linear múltipla percebe-se que a variável valor de imóvel sofre uma influência significativa em relação às covariáveis área do terreno, idade do imóvel e número de cômodos.

Devido à gama possibilidade de demonstrar as variáveis explicativas e da amplitude de características que podem ser consideradas na pesquisa, pois podemos envolver outras inúmeras variáveis sobre este estudo, compreendeu-se que este estudo não esgota o assunto, pelo contrário, demonstra apenas uma possibilidade de tratamento para tema. Sendo assim outros modelos poderão ser elaborados se utilizando desta metodologia, com emprego de mais variáveis ou de diferentes formas de mensuração das mesmas.

5. Referências Bibliográficas

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. NBR 14653 – 2 / 2011.

CHARNET, R. et al. Análise de modelos de regressão linear com aplicações. Campinas: Unicamp, 1999. 351 p.

DANTAS, Rubens Alves. Engenharia de Avaliações: uma introdução à metodologia científica. 2. ed. São Paulo: PINI, 2005

MENDONÇA, Marcelo Corrêa; et al. Fundamentos de Avaliação e Perícias de engenharia. São Paulo: PINI, 1998.

FIKER, J. Manual de avaliações e perícias em imóveis urbanos. 3. ed. São Paulo: Pini, 2008. 157p.

R CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna: R Foundation for Statistical Computing, 2017. Available at <<http://www.r-project.org>>. Access in: 13 mar. 2019.