

Banco de dados: Boston House Prices

Fernando Bispo, Jeff Caponero

Sumário

| | |
|-------------------------------------|----------|
| Introdução | 2 |
| Metodologia | 3 |
| Objetivo | 3 |
| Sobre o conjunto de dados | 3 |
| Variáveis | 3 |
| Análise Descritiva | 4 |
| Running Code | 4 |

Introdução

A busca por melhor qualidade de vida é um dos principais objetivos de vida das pessoas em geral e a escolha de moradias que proporcionem isso não foge a esse preceito.

Com o objetivo de investigar a disposição das pessoas em pagar por imóveis localizados em regiões com maior concentração de ar limpo, imóveis esses pertencentes a região metropolitana de Boston, esta análise visa obter através das técnicas da análise de regressão uma visão sobre essa preferência.

Composto por 506 observações e 14 variáveis, o conjunto de dados foi publicado no *Jornal of Environmental Economics & Management*, vol.5, 81-102, 1978.

Metodologia

Objetivo

O objetivo deste trabalho será determinar, a partir de técnicas de regressão linear, o preço de casas em Boston com base nos dados fornecidos pelo banco de dados analisado.

Sobre o conjunto de dados

Os dados de preços de 506 casas em Boston publicados em Harrison, D. and Rubinfeld, D.L. *'Hedonic prices and the demand for clean air'*, J. Environ. Economics & Management, vol.5, 81-102, 1978.

Os dados podem ser acessados na plataforma para aprendizado de ciência de dados [Kaggle](https://www.kaggle.com/datasets/fedesoriano/the-boston-houseprice-data) através do link: <https://www.kaggle.com/datasets/fedesoriano/the-boston-houseprice-data>.

Variáveis

1. CRIM: Índice de criminalidade per capita por bairro.
2. ZN: Proporção de terreno residencial zoneada para lotes acima de 25.000 sq.ft. (Square feet – Pés quadrados)
3. INDUS: Proporção de hectares de negócios não varejistas por bairro.
4. CHAS: Margem do rio Charles (1 se o trecho margeia o rio; 0 caso contrário).
5. NOX: Concentração de óxidos nítricos (partes por 10 milhões. [partes/10M]).
6. RM: Número médio de cômodos por habitação.
7. AGE: Proporção de unidades próprias construídas antes de 1940.
8. DIS: Distâncias ponderadas para cinco centros de emprego de Boston.
9. RAD: Índice de acessibilidade às rodovias radiais.
10. TAX: Valor total do imposto predial por \$10.000 [\$10k].
11. PTRATIO: proporção aluno-professor por bairro.
12. B: O resultado da equação $B = 1000(Bk - 0,63)^2$ onde Bk é a proporção de negros por bairro.
13. LSTAT: % da população de “classe baixa”.

Análise Descritiva

Running Code

When you click the **Render** button a document will be generated that includes both content and the output of embedded code. You can embed code like this:

```
# 1 + 1
```

You can add options to executable code like this

```
[1] 4
```

The `echo: false` option disables the printing of code (only output is displayed).