

# Exercício 4

## Descrição

No arquivo **dboston.dat** é apresentado parte do conjunto de dados de uma amostra aleatória de 506 setores censitários de 96 distritos da cidade de Boston (USA) em 1970. O objetivo principal do estudo é tentar explicar a relação entre o preço mediano das residências ocupadas pelos proprietários em cada setor censitário com algumas variáveis explicativas. Vamos considerar apenas quatro variáveis explicativas que estão colocadas na seguinte ordem:

- dist: distância ponderada do distrito a cinco centros de emprego de Boston (em milhas)
- tax: imposto distrital anual do imóvel (por 10 mil USD)
- ptratio: relação aluno-professor no distrito
- lstat: porcentagem da população com baixa renda
- medv: preço mediano das residências ocupadas pelos proprietários (em mil USD)

## Enunciado

- Faça inicialmente uma análise descritiva construindo por exemplo boxplots e diagramas de dispersão de cada variável explicativa contra a variável resposta.
- Apresente também a densidade da variável resposta.
- Para cada ligação (logarítmica, identidade e recíproca) proponha um modelo com resposta gama e selecione as variáveis explicativas usando o método AIC.
- Através de procedimentos de diagnóstico escolha um modelo.
- interprete os parâmetros do modelo escolhido.

## Leitura dos dados

```
dboston <- data.frame(scan("dados/dboston.dat",  
                           list(dist=0, tax=0, ptratio=0, lstat=0, medv=0)))
```

	dist	tax	ptratio	lstat	medv
1	4.09	296.00	15.30	4.98	24.00
2	4.97	242.00	17.80	9.14	21.60
3	4.97	242.00	17.80	4.03	34.70
4	6.06	222.00	18.70	2.94	33.40
5	6.06	222.00	18.70	5.33	36.20
6	6.06	222.00	18.70	5.21	28.70
7	5.56	311.00	15.20	12.43	22.90
8	5.95	311.00	15.20	19.15	27.10
9	6.08	311.00	15.20	29.93	16.50
10	6.59	311.00	15.20	17.10	18.90
11	6.35	311.00	15.20	20.45	15.00
12	6.23	311.00	15.20	13.27	18.90

## Análise descritiva

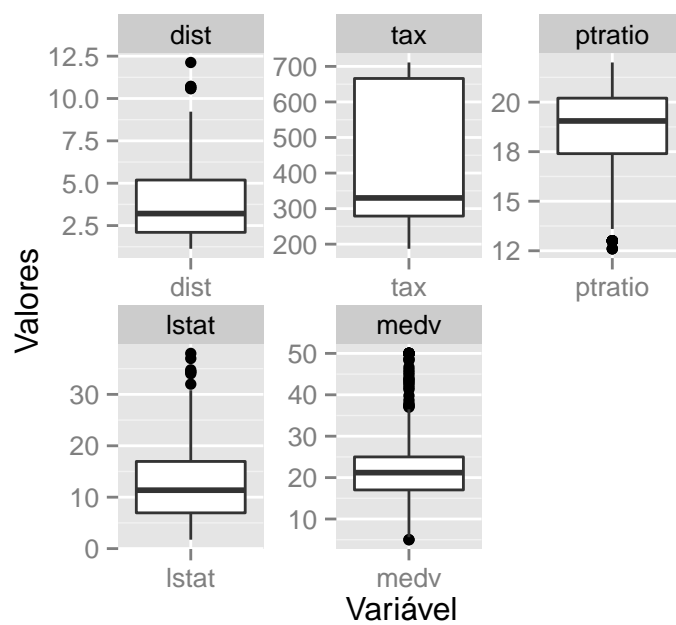


Figure 1: Boxplots das variáveis do banco de dados

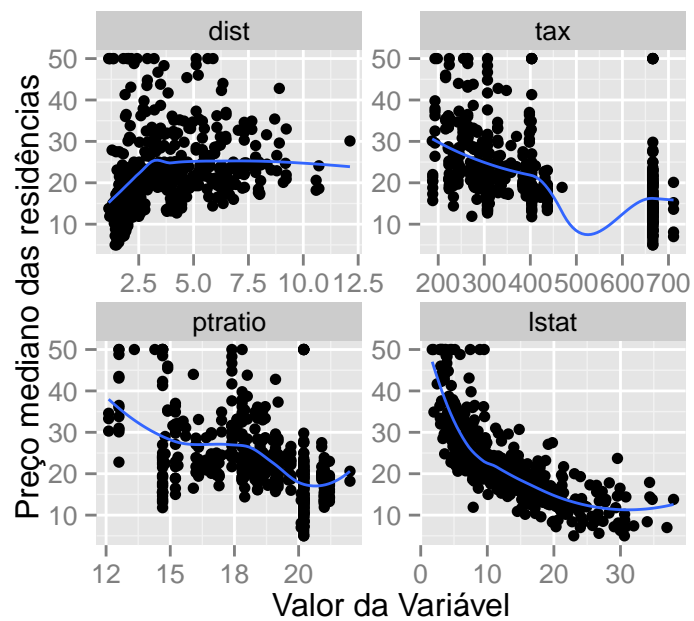


Figure 2: Gráficos de dispersão das variáveis explicativas pela variável resposta

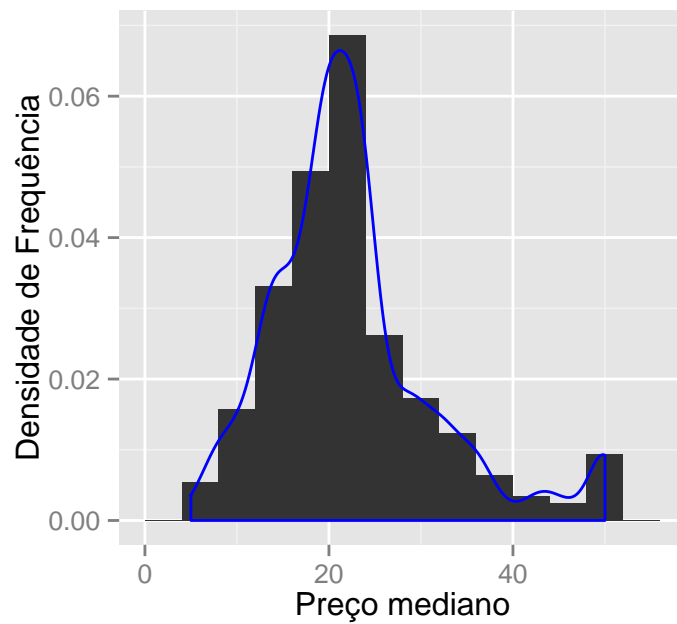


Figure 3: Gráficos de dispersão das variáveis explicativas pela variável resposta