

## 2ª Lista de Exercícios

### Providencias preliminares:

1. Instale e importe a biblioteca “MASS”.
2. Instale e importe a biblioteca “AER”.
3. Instale e importe a biblioteca “lmtest”.
4. Carregue a base de dados “Boston”. Para tanto basta acionar o comando `data("Boston")` após já ter importado a biblioteca “MASS”.
5. Importe a base de dados contida no arquivo “despesas.csv” e a salve como “despesas”.
6. Importe a base de dados contida no arquivo "bankwages.csv" e a salve como “bankwages”.

Dica ao importar as bases de dados “.csv”: utilizem o comando “read.table” e fiquem atentos a qual caractere serve como separador entre as colunas. Se for o caso, abram o arquivo “.csv” no bloco de notas primeiro para investigar.

## QUESTÕES

1. Obtenha a descrição de cada variável na base de dados “Boston”. Dica: utilize o comando “?  
Boston” no RStudio.
2. Utilize o comando “summary” para obter algumas estatísticas descritivas da base de dados “Boston”. Utilize o RStudio para visualizar a planilha da base de dados. Sugere-se que essas providências sejam tomadas sempre que uma base de dados for importada. Isso é especialmente importante quando se importa a base de algum arquivo .xlsx, .csv, ou .txt pois é necessário certificar que os dados foram importados corretamente.
3. Estime um modelo de regressão linear simples que explique o valor mediano das casas nos subúrbios de Boston através do percentual de pessoas de baixo *status* socioeconômico e de uma constante. Salve o modelo como “bh\_mod”.
4. Utilizando a base de dados “Boston” estime um modelo de regressão linear que explique o valor mediano das casas nos subúrbios de Boston através: da proporção de unidades ocupadas pelos proprietários que foram construídas antes de 1940; da taxa de crime per capita; percentual de pessoas de baixo *status* socioeconômico; e de uma constante. Salve o modelo como “mod”.

5. Verifique, utilizando os operadores “>” ou “<”, se o  $R^2$  do modelo “mod” é maior que o do modelo “bh\_mod”.
6. Considerando que não é adequado utilizar o  $R^2$  para comparar modelos com quantidades diferentes de variáveis explicativas (regressores), verifique se o  $R^2$ -ajustado do modelo “mod” é maior que o do modelo “bh\_mod”.
7. Estime um modelo de regressão linear que explique o valor mediano das casas nos subúrbios de Boston através de todas as outras variáveis disponíveis na base de dados “Boston”. Salve o modelo como “full\_mod”.
8. Apresente os erros-padrão, as estatísticas t, e os p-valores da regressão do modelo “full\_mod”. Qual variável apresentou maior significância estatística?
9. O modelo “full\_mod” apresentou um  $R^2$ -ajustado maior que o modelo “mod”?
10. Salve os coeficientes do modelo “mod” como “coefs”.
11. Utilize métodos econométricos para estimar as taxas de crescimento discretas ( $g$ ) para as quatro variáveis de consumo na base de dados “despesas”. Apresente também o p-valor dessas estimativas e o  $R^2$  da regressão. Sugestão: utilize um *for-loop* para evitar a repetição de código.
12. Utilizando a base de dados “bankwages”, desenvolva e apresente o sumário (summary) de um modelo econométrico que explique a variável LOGSAL (log do salário anual) através de um intercepto e da variável EDUC (anos de estudo). Segundo esse modelo, em quantos por cento o **salário** de uma pessoa aumenta para cada ano adicional de educação?
13. Utilizando a base de dados “bankwages”, desenvolva e apresente o sumário (summary) de um modelo econométrico que explique a variável LOGSAL através de um intercepto, da variável EDUC, e da variável EDUC<sup>2</sup>. Com base nesse modelo, qual é o logaritmo do salário esperado de uma pessoa sem qualquer tempo de estudo? Com quantos anos de estudo a pessoa receberá o menor salário estimado? Apresente uma explicação para o fato de uma pessoa com mais anos de estudo receber menos que uma pessoa com menos anos de estudo.