

# Especialização em Ciências de Dados - Lista 1

6 de outubro de 2018

1. Carregue a planilha "Consumo de carne.xlsx", e nomeie-a *Consumo*. Ela contém o consumo anual, preço no atacado e preço no varejo das carnes bovina, de frango e suína, no Reino Unido. O peso está em libras, e o preço em centavos por libra.

2. Observe que nossos dados apresentam a **evolução temporal** do consumo de carnes e seus preços. Nesse caso, o melhor tipo de gráfico para compreender como essa evolução temporal ocorre é o **gráfico de linhas**.

Para desenhar esse tipo de gráfico, utilizamos o comando `plot()`. Por exemplo, para compreender como o consumo de carne bovina evolui ao longo dos anos,

```
> plot(Consumo$ano, Consumo$consumo_bovino, type="l")
```

O comandos opcionais do `barplot()` também valem aqui, como `col` (alterar cores), `xlim` (limites do eixo x), `ylim` (limites do eixo y), `lwd` (grossura da linha), entre outros. Utilizamos o comando `type="l"` para informar que queremos um gráfico de linhas.

3. É interessante plotar, junto com o consumo de carne bovina, o consumo das demais carnes, para entendermos como foi a evolução entre os diversos tipos. Para isso, utilizamos o comando `plot()`, como anteriormente, e adicionamos os demais gráficos usando o comando `lines()`.

O comando `lines()` desenha por cima do gráfico já criado pelo `plot()`, então, se desejamos realizar ajustes, para tornar o gráfico mais fácil de compreender, devemos executar o `plot()` novamente, e refazer todos os `lines()`. Vamos fazer o exemplo, para entender melhor.

```
> plot(Consumo$ano, Consumo$consumo_bovino, type="l")
```

```
> lines(Consumo$ano, Consumo$consumo_frango)
```

Observe que tivemos alguns problemas:

- (a) Parte do gráfico do consumo de frangos ficou fora do gráfico.
- (b) Não fica claro qual linha pertence a qual dado.

Para corrigir isso,

- (a) Use o comando opcional para aumentar os limites do eixo y, de forma a aparecer todo o gráfico.
- (b) Troque as cores das linhas, usando "red" para carne bovina e "blue" para carne de frango.

Observe que o gráfico já ficou mais amigável. Tente deixar o gráfico o mais amigável possível, renomeando os eixos, aumento a grossura da linha para 2 e colocando um título. Exporte esse gráfico com o nome *gráfico1.png*.

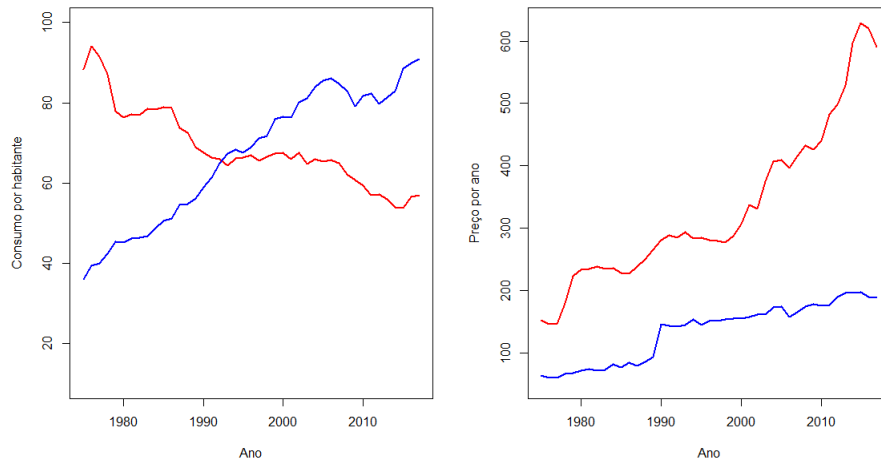
4. Baseado no item anterior, vamos tentar buscar respostas à seguinte pergunta: por que houve essa diferença no consumo de carne bovina e de frango?

Para isso, vamos plotar lado a lado dois gráficos: o primeiro, com o consumo bovino e de frango, como fizemos no item anterior, e outro com a evolução dos preço no varejo da carne bovina e de frango.

Para isso, dividimos nossa janela de gráficos, usando o comando

```
> par(mfrow=c(1,2))
```

Esse comando faz com, utilizando o comando `plot()` pela primeira vez, o gráfico é plotado à esquerda. De utilizarmos `lines()`, será desenhado no gráfico da esquerda. Ao utilizarmos um novo



`plot()`, ele passa a desenhar à direita da janela de gráficos. Um novo `plot()` ira sobrepor o desenho da esquerda, e assim vai acontecendo essa alternância.

Como o modelo acima, desenhe lado a lado os gráficos de consumo (à esquerda) e de preço no varejo (à direita), da carne bovina e de frango.

Exporte esse gráfico com o nome *gráfico2.png*.

5. Converta as unidades, de libra para quilogramas, e os preços de centavos por libra para euros por quilograma. Construa os mesmos gráficos, utilizando essas medidas. Exporte-os como *gráfico3.png* e *gráfico4.png*.
6. Faça um texto, explicando com suas palavras, quais os possíveis motivos que os gráficos nos induzem a conjecturar sobre esse aumento no consumo de frango e queda no consumo de carne bovina. Não é necessário termos técnicos.
7. Ao final das atividades, não esquece de digitar `> par(mfrow=c(1,1))` para os gráficos voltarem ao normal.
8. Você deve entregar
  - (a) Os quatro gráficos salvos.
  - (b) O texto criado.
  - (c) Os comandos utilizados.