

Exercícios de Leitura e Manipulação de Dados

Distância entre cidades

Abaixo as distâncias por estradas entre quatro cidades da Europa, em quilômetros:

- Atenas a Madri: 3949
- Atenas a Paris: 3000
- Atenas a Estocolmo: 3927
- Madri a Paris: 1273
- Madri a Estocolmo: 3188
- Paris a Estocolmo: 1827

1. Construa uma matriz de distâncias com esses valores.
2. Compare sua matriz com o objeto `eurodist`, disponível no pacote `datasets`.

Criação de um data frame

Imagine um experimento em que hamsters de dois fenótipos (claros e escuros) recebem três tipos diferentes de dieta, e no qual as diferenças dos pesos (g) entre o fim e o início do experimento sejam:

	DIETA A	DIETA B	DIETA C
CLAROS	0.1 , 1.1 , 3.7	5.7, -1.2, -1.5	3.0, -0.4, 0.6
ESCUROS	1.5, -0.1, 2.0	0.6, -3.0, -0.3	-0.2, 0.3, 1.5

1. Crie um *data frame* com esses dados, na qual cada hamster seja uma linha, e as colunas sejam as variáveis `cor`, `dieta` e `variação do peso`.
2. Qual a media por dieta e por fenótipo?

DICA: Use as funções de gerar repetições para criar os vetores dos tratamentos.

OPCIONAL: depois de fazer este exercício com as funções de repetição, experimente repeti-lo com a função `expand.grid`.

Criando uma Matriz

1. Crie um objeto da classe `matriz` com 5 colunas e 3 linhas contendo uma amostra de uma distribuição normal de média 10 e variância 3,6.
2. Mude os nomes das linhas para "L1" a "L3" e das colunas para "C1" a "C5". **Dica:** A função `paste` combinada com funções de repetições podem lhe poupar trabalho.
3. Calcule a média e a variância por colunas e depois por linhas. Guarde os resultados em um *dataframe* que tenha a variância e a média em vetores separados.

Lendo e Salvando seus dados

1. Leia um arquivo de dados de sua autoria e faça um sumário das variáveis independentes. Caso tenha variáveis qualitativas faça tabelas de contagem para cada uma e duas a duas. Caso não tenha dados, poderá ir a campo ou utilize alguns dos dados da área [Arquivos de dados](#) do nosso wiki 😊. **Dica:** veja o help da função `read.table()`
2. Faça uma seleção de um estado de atributo de um vetor fator dos seus dados e salve em um objeto (uma espécie, uma localidade...), mantendo o objeto original intacto. Em seguida selecione apenas três colunas no objeto resultante e salve em um arquivo de formato .txt, separado por tabulação e sem nome de linhas. Lembre-se de olhar o help antes de surtar! Para verificar se o arquivo está correto abra-o no EXCEL, *URG!*

Classes de Objetos

O pacote “datasets” contém vários conjuntos de dados para uso em treinamento com a linguagem R.

O conjunto “iris” é distribuído de duas formas diferentes, nos objetos `iris` e `iris3`. São quatro medidas de flores de três espécies de *Iris* (Iridaceae).

1. Quais são as classes desses dois objetos?
2. Calcule a média de cada uma das quatro medidas por espécie, dos dois objetos. **DICA:** as funções da família `apply` irão te ajudar.
3. Os nomes das variáveis estão em inglês. Mude-os para português no objeto `iris`. **DICA:** Como tudo mais no R, os resultados da função `names` podem ser armazenados em um objeto.

Acrescentando Dados de Síntese

1. Crie um objeto com os dados do arquivo-texto [Biomassa de Árvores de Eucalyptus saligna](#).
2. Verifique o conteúdo do objeto resultante, com a função `summary`.
3. Acrescente uma nova coluna ao data frame resultante, com a soma das biomassas de folhas e do tronco de cada árvore.
4. Acrescente outra coluna, com o valor da área basal de cada árvore.
5. Calcule a área basal total por talhão e salve em um objeto
6. Calcule a média da área basal por talhão
7. Selecione apenas os dados relativos às árvores com mais de 10 cm de diâmetro
8. Salve em um arquivo `txt`, separado por tabulação.

UFA! 😊

From:
<http://ecor.ib.usp.br/> - **ecoR**

Permanent link:
http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2010:exercicios3

Last update: **2020/08/12 06:04**

