Danilo Vicensotto Bernardo

Aluno de Doutorado do Departamento de Genética e Biologia Evolutiva do Instituto de Biociências da Universidade de São Paulo. Minha pesquisa é direcionada ao estudo da diversidade morfocraniana humana e seus padrões microevolutivos. Meus interesses englobam ainda Evolução Humana, Arqueologia e Antropologia Biológica.

meus exercícios

exercicio1.txt

exercicio2.txt

exercicio3.txt

exercicio4.txt

exercicio5.txt

exercicio6.txt

exercicio7.txt

exer7.txt

exercicio8.txt

exercicio9.txt

trabalho final

Proposta A

"É muito comum a presença de valores faltantes (*missing values*) em bancos de dados extensos, que prejudicam a aplicação de determinadas técnicas analíticas, como a Análise de Componentes Principais, por exemplo, à esses bancos. Para contornar esse problema, sem eliminar as observações que apresentam NAs, normalmente se procede a substituição desses valores por um número estimado. A estimativa desse novo número pode ser feita por diferentes critérios, como a média geral de todas as observações, a média de um determinado grupo da amostra, um valor regredido para toda a amostra ou um valor regredido para o grupo da amostra". **A proposta consiste em desenvolver uma função que opere um data.frame qualquer, destaque os NAs presentes e faça a substituição desses valores por critérios escolhidos no argumento da função.**

Comentários PI

Execelente problema! Muito bem dimensionado e claramente apresentado.

Help da função

limpa.missing {base} R Documentation

```
Remove NA's de um data.frame
```

Descrição:

limpa.missing remove valores NA presentes em data frame, substituindo-os pela média global de cada variável do frame. Retorna os valores do data frame e produz um arquivo ".txt" com a nova matriz.

Uso:

limpa.missing(…)

Argumentos:

… O Objeto (data.frame) que deve ter os valores faltantes removidos.

Detalhes:

O data.frame deve ser composto por nomes (ID's) na primeira coluna, seguido pelas variáveis nas colunas seguintes.

Função em desenvolvimento!

Author:

Danilo V Bernardo danvb@ib.usp.br

Referencias:

Crawley, M.J. (2007). The R book. Wiley, Wiley, Imperial College London, UK.

Exemplo:

limpa.missing(dados)

Código da função

limpa.missing ← function(x) {

```
x -> mx      ### cria o objeto a ser manipulado
n=dim(mx)[2]        ### descobre o número de colunas do objeto
s=seq(2,(dim(mx)[2]))       ### cria uma sequência de 2 até o número final
de colunas do objeto
```

```
medias_g=c("médias_geral", round(mean(mx[,s],na.rm=TRUE), 3))     ### cria
um vetor contendo as médias das variáveis (colunas) independente dos grupos
(geral), excluindo apenas os NAs
mx -> MX
for (i in s)
{
    replace(mx[i], is.na(mx[i]), medias_g[i]) -> MX[i]
    MX[i]
}
```

return(MX) write.table(MX, “dados.txt”, sep=“\t”) }

arquivos

Arquivo do help help.txt

Arquivo do código limpamissing_v1.0_.txt

Arquivo para teste crandatar.txt

Proposta B

Comentários PI

Cadê o plano B?

Professores: A cabeça esvaziou (sim, sim, síndrome do Pink) mas prometo que até amanhã posto alguma coisa para o plano B, pode ser? Penso em fazer alguma função para simulação que se aplique à genética de populações.

pensamentos

“Quem mexe com R não precisa de playstation”

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

Permanent link:
http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2010:alunos:trabalho_final:danvb:start

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