

Aula 5 - Manipulação de dados, análise exploratória

Vitor Rios

11 de novembro de 2017

#Função subset() Divide um objeto em subconjuntos

x = objeto a ser dividido subset = condição lógica indicando o que deve ser mantido
select= indica as colunas a serem selecionadas em um dataframe

#Familia apply() Existem várias funções no R que aplicam funções sobre objetos, e elas variam principalmente no tipo de objeto que devolvem

##Função apply()

Aplica uma função a todas as linhas ou todas as colunas de um array (dataframe numérico ou matriz), retorna um vetor ou matriz *ATENÇÃO* : só funciona para dados numéricos, não serve para fatores apply(X, MARGIN, FUN, ...) x é array que se quer agrupar MARGIN 1 = linhas, 2 = colunas, c(1,2), ambos FUN é a função que se quer aplicar nas margens, sem parênteses ... são os argumentos que serão passados para a função FUN Por exemplo, se queremos a soma de cada coluna

#Função lapply() Aplica uma função a cada elemento de uma lista, e retorna uma lista. Pode lidar com qualquer tipo de dado, a depender de FUN lapply(X, FUN, ...) X um lista ou objeto que será convertido em lista (pode ser data.frame ou matriz) FUN é a função que se quer aplicar nos elementos de, sem parênteses ... são os argumentos que serão passados para a função FUN

```
## $dragoes_completo.peso
```

```
## [1] 10377.68
```

```
##
```

Para facilitar:

- ▶ `apply`: genérica: aplica uma função a linhas ou colunas de uma matriz (ou às dimensões de um array), retorna vetor ou matriz
- ▶ `lapply`: “list apply”. Age em uma lista ou vetor e retorna uma lista
- ▶ `sapply`: “simple lapply”. Igual a `lapply`, mas retorna um vetor ou matriz sempre que possível
- ▶ `tapply`: “tagged apply”. subconjuntos (tags) identificam os grupos nos quais a função será aplicada. Tipo de retorno depende da função, geralmente array
- ▶ `aggregate`: `tapply` que converte o resultado para dataframe

Na maioria dos casos você vai usar `aggregate` ou `tapply`