

MBA⁺

**ARTIFICIAL INTELLIGENCE
& MACHINE LEARNING**

MBA⁺

PROGRAMANDO IA COM R

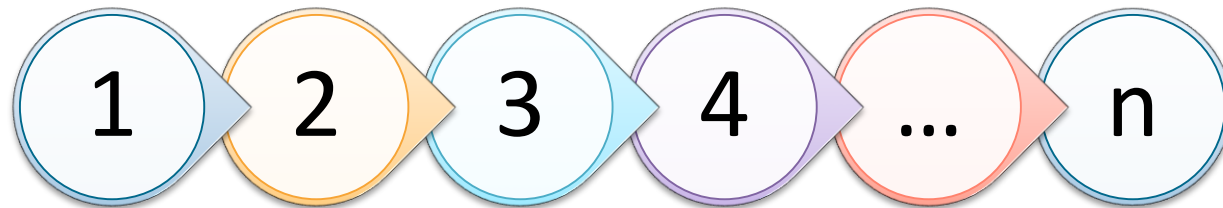
Prof. Elthon Manhas de Freitas
elthon@usp.br

2018

Revisão da última aula

- O que vimos na aula passada?

Seqüências



- Dados sequenciais podem ser usados rapidamente das seguintes formas:

– : (isso mesmo, dois pontos)

- 1:9
- 80:32
- 4:-2
- 3:3
- 1:0

– função `seq` (e suas variações)

- `seq.int`

- ★ • `seq_along`

- ★ • `seq_len`

presente no help:

- `seq(0, 1, length.out = 11)`
- `seq(stats::rnorm(20))`
- `seq(1, 9, by = 2)`
- `seq(1, 9, by = pi)` # stays below 'end'
- `seq(1, 6, by = 3)`
- `seq(1.575, 5.125, by = 0.05)`
- `seq(17)` # same as `1:17`, or even better `seq_len(17)`

- Comando que “repete” um valor n vezes:
 - `rep(0, times = 40)`
 - `rep(c(0, 1, 2), times = 10)`
 - `rep(c(0, 1, 2), each = 10)`

Exercícios com seqüências

- Crie uma variável `my_seq` com 30 valores entre 5 e 10
- Veja o help da função ``:``
- Qual a diferença das instruções
 - `pi:10`
 - `10:pi`
- Como consultar o tamanho do vetor `my_seq`
- Como fazer uma sequencia que acompanhe o tamanho do vetor `my_seq`?

```
Resposta FIAP
# my_seq <- seq(5, 10, length=30)
# pi:10
# 10:pi

# length(my_seq)
# seq_along(with = my_seq)
# seq_along(my_seq)
```



VOID

MISSING VALUES:

NA, NaN e NULL

SPECIAL VALUE:

Inf

OUTROS TIPOS DE
DADOS (DATA E
SEQUENCIAS)

Valores especiais

- NA
 - Missing genérico, para representar Not Available
- NaN (Not a Number)
 - Resultados de contas que causam erros aritméticos
 - Exemplo: $\frac{0}{0}$ ou $\sqrt{-4}$
- Inf e -Inf
 - Números infinitos
 - Exemplo: $\frac{19}{0}$
- NULL
 - Null representa o vazio, o nada!
 - Até NA tem um tipo, mas NULL não!
 - Não é possível colocar NULL em um vetor



Exercícios sobre Special Values

- Criar 5 variáveis. Cada uma com um tipo dos valores especiais (+Inf e -Inf)
- Tentar criar um vetor com as 5 variáveis.
 - Foi possível?
- Verificar os valores das variáveis e vetor através das funções
 - is.na()
 - is.nan()
 - is.null()
 - is.infinite()
 - is.finite()

Resposta do exercício

```
v1 = NA
v2 = NaN
v3 = NULL
v4 = Inf
v5 = -Inf

vall = c(v1, v2, v3, v4, v5)
is.na(vall)
is.nan(vall)
is.null(vall)
is.null(vall)
is.infinite(vall)
is.finite(vall)

is.na(v2)
is.nan(v2)
is.null(v2)
is.infinite(v2)
is.finite(v2)
```



Filtrando dados das estruturas – Subsets simples

Filtrando dados das estruturas (subset)

- Os dados são filtrados em diversos tipos de estruturas através do uso de colchetes.

```
• mtcars
• mtcars[2, 5]
• mtcars[2:4, 5]
• mtcars[2:4, 5:7]
• mtcars[8:9, c(1, 2, 4)]
• mtcars[2:4, c("mpg", "wt")]
```

- Nas matrizes, a primeira dimensão é a linha e a segunda é a coluna;
- Pode-se filtrar pelo número ou pelo nome, quando a dimensão for nomeada;
- Ao sub-conjunto de dados se dá o nome de “subset”

Sub-sets simples: Vetores

- Vetores, por terem apenas uma dimensão, possuem o subset mais simples de se obter
- Uso dos colchetes simples:

`-vetor[filtro] -> resultado`

(subset)

nome do vetor
(set)

Diversos modos de
aplicar filtros
(subscript)

R-Markdown:
Subsets

Sub-sets simples: Matrizes

- Similar ao subset dos vetores, porém com 2 dimensões

- Uso dos colchetes simples:

```
-matriz[linha,coluna] -> resultado
```

R-Markdown:
Subsets

Sub-sets simples: Listas

- Para listas, há mais opções de filtros:

- `lista[filtro]` -> sub-lista

- `lista[[filtro]]` -> elemento

- `lista$elemento` -> elemento

R-Markdown:
Subsets

- Data é representada pela classe Date
 - São armazenados os dias a partir de 01-01-1970
- Hora pode ser representada por:
 - POSIXct
 - Número de segundos a partir de 01-01-1970
 - POSIXlt
 - Lista de ano, mês, dia, hora, minuto, segundo, etc.

Experimento com Data e Hora

- Façam o seguinte experimento, analisem e comentem o resultado:

```
• dia_texto <- "28/09/2017 T 18:51:30"
• dia_date <- as.Date(dia_texto,format="%d/%m/%Y T
%H:%M:%S",tz="America/Sao_Paulo")
• dia.time1 <- as.POSIXct(dia_texto,format="%d/%m/%Y T
%H:%M:%S",tz="America/Sao_Paulo")
• dia.time2 <- as.POSIXlt(dia_texto,format="%d/%m/%Y T
%H:%M:%S",tz="America/Sao_Paulo")

• dia_date
• dia.time1
• dia.time2
• unclass(dia_date)
• unclass(dia.time1)
• unclass(dia.time2)

• dia.time1$year
• dia.time2$year
```

- O que a função “unclass” faz?
- Qual a diferença entre os objetos dia.time1 e dia.time2?

Funções especiais com data e hora

- Para obter a data e hora atual

- `Sys.time()`
- `Sys.Date()`
- `Sys.timezone()`

- Converter um texto em data

- `as.Date()`
- `as.POSIXct()`
- `as.POSIXlt()`
- `strptime("Janeiro 10, 2012 10:40", "%B %d, %Y %H:%M")`



DATA E HORA : LUBRIDATE

- Trata-se de um pacote facilitador para manipular tempo:
 - Um instante no tempo
 - Uma data (ou data/hora) específica
 - Intervalo
 - Um período entre um “Início e um Fim”, uma duração de tempo
 - Função `duration()`

- Conversão de datas com facilitador de formatos

- `library(lubridate)`
- `ymd("20110604")`
- `mdy("06-04-2011")`
- `dmy("04/06/2011")`
- `ymd_hms("2018-04-23T19:02:13")`
- `dmy_hms("23/04/2018 19:03:14")`

- Extração de componentes da data

- Intervalos de datas

- dia, hora, segundos, etc

- Tratamento de fuso horário

- Operações com datas

- somar dias, minutos, etc

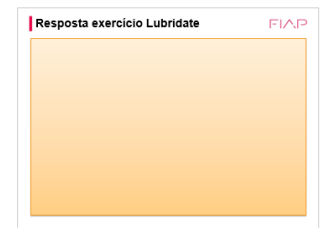
Abrir Rmarkdown:

Aula 02 - Lubridate

- Criar uma `duration` com 260 segundos
- Criar uma `duration` com 260 minutos e divida-a por 60
- Criar uma `duration` de 1 dia e subtraia uma `duration` de 60 minutos
- Criar uma `duration` de 12 dias, 5 horas e 10 minutos
- Calcular a Black Friday de 2018:

O Black Friday acontece 1 dia depois do dia de ação de graças americano, que é celebrado na 4ª quinta-feira de novembro.

 - Crie uma variável representando 1-nov-2018
 - Verifique que dia da semana é 1-nov
 - Adicione a quantidade necessária para chegar na quinta feira e adicione em uma segunda variável
 - Adicione 3 semanas.
- Testar para 2019, 2020 e 2021



rmarkdown

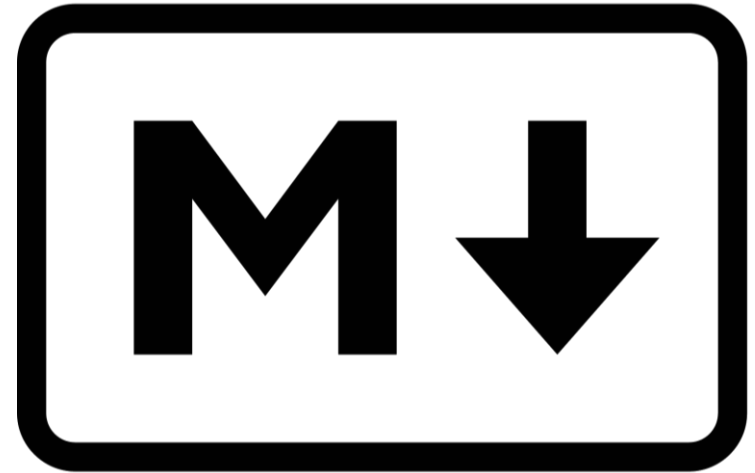
○ MARKDOWN DO R



O que é Markdown?

- Uma linguagem simples de marcação que converte textos simples em textos formatados

- HTML, Word, PDF, etc.



- Exemplos:

- Wikipedia

- LaTeX

- Notebooks (como o Jupyter, etc.)

L^AT_EX

O que é RMarkdown?

- Mescla expressões Markdown tradicional com códigos R através de TAG.
 - Texto sem compilação
 - Código R
 - Código Python (não será visto)
 - Código SQL (não será visto)



```
• ```{r [options]}
```

```
• <Código em R>
```

```
• ```
```


Para preparar o ambiente

Para iniciar com o R-Markdown, é necessário instalar os seguintes pacotes:

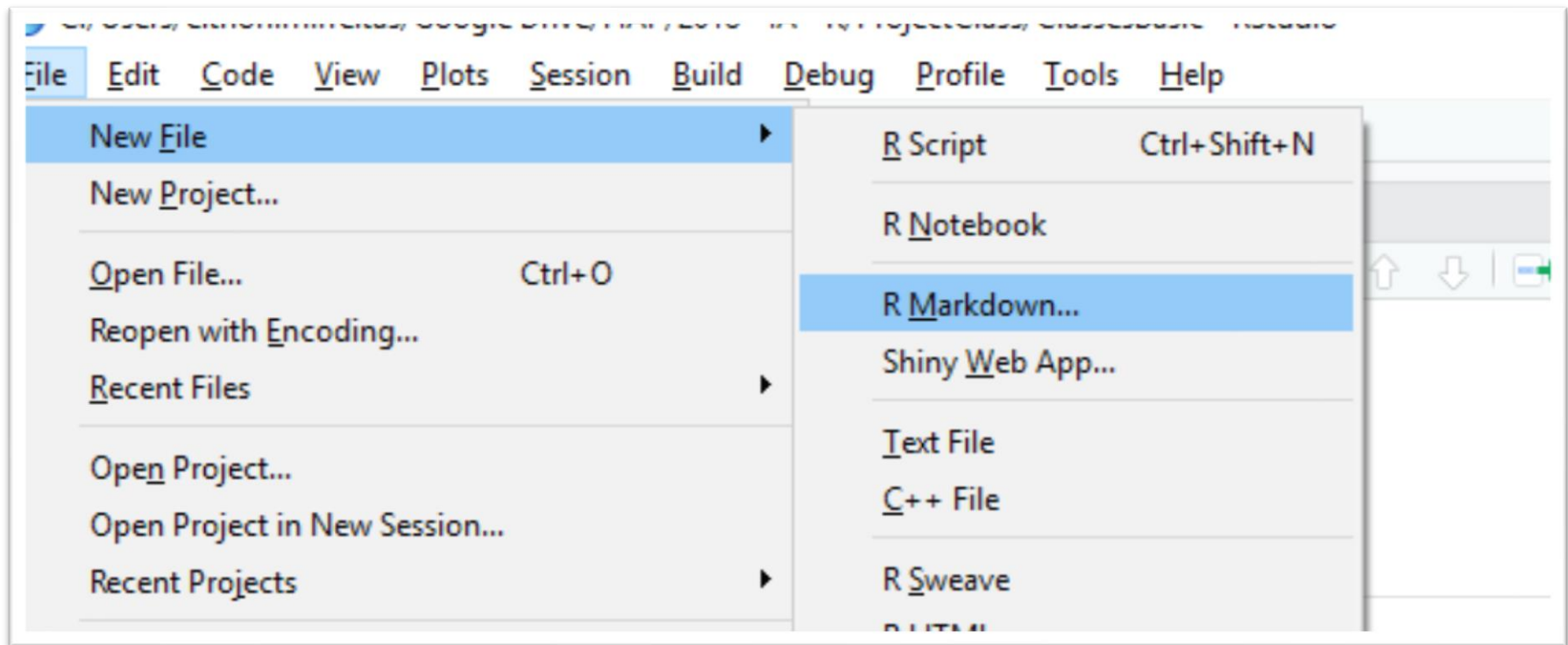
- evaluate
- digest
- highr
- markdown
- stringr
- yaml
- Rcpp
- htmltools
- caTools
- bitops
- knitr
- jsonlite
- base64enc
- rprojroot
- rmarkdown

Opção:

```
path = "http://raw.githubusercontent.com/elthonf/fiap-mba-r/master/"  
source(file = paste0( path, 'geral_codes/inst_rmarkdown.r' ) )
```

Criar arquivo RMarkdown

- Arquivo extensão .Rmd

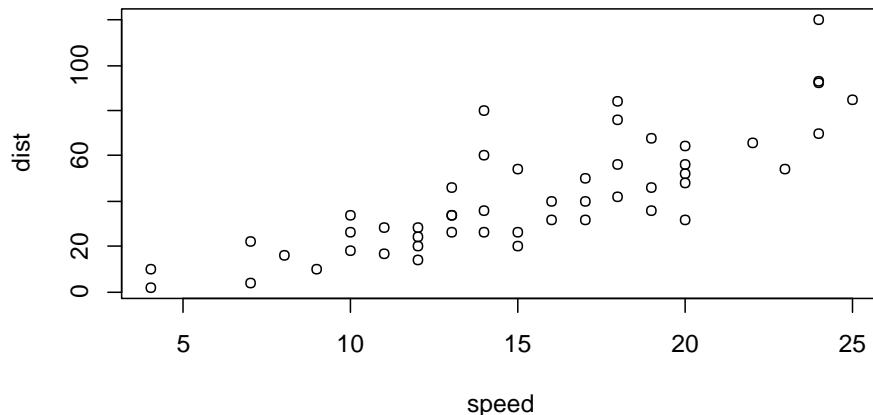


- Diversos tipos de saída: HTML, PDF (requer LaTeX), Word e Notebook

- Cabeçalho YAML
 - Configurações gerais sobre o arquivo
- Corpo
 - Texto Markdown
 - Código HTML
 - Fórmulas
 - Chunk com códigos a serem executados
 - As saídas dos Chunks são inseridas no texto final
 - Os códigos dos Chunks podem ser inseridos no texto final


Exercícios RMarkdown - 01

- Criar um arquivo .Rmd
- Configurar output para html ou notebook (que é html)
 - Chunk com todo o conteúdo da tabela 'cars'
 - Chunk com 5 primeiras linhas da tabela 'cars'
 - Chunk com 5 últimas linhas da tabela 'cars'
 - Chunk com o resumo da tabela 'cars'
 - Chunk com plot simples da tabela 'cars'
 - Texto explicativo antes de cada chunk



Utilize os comandos:

```
head  
tail  
summary  
plot
```

- Criar um arquivo .Rmd com
 - exercícios “Aula 01 A” e “Aula 01 B”. (estão com )
- Criar texto explicativo antes de cada Chunk.
- Dicas para melhorar o texto:
 - `*itálico*`
 - `**negrito**`
 - escrever as fórmulas entre `$$`
 - `$j \leftarrow a^2 + \sqrt{\frac{1}{100}}$`
 - escrever códigos demonstrativos “inline” com crases
 - ``a <- 123``



Trabalho em grupo

- Definição dos grupos
- Preencher formulário presente no site:
<https://sites.google.com/usp.br/programando-ia-com-r>
- Escolha dos Cases (próxima aula)



Exercícios individuais (Aprenda R no R)

MBA⁺

Copyright © **2018**

Prof. Elthon Manhas de Freitas

Todos direitos reservados. Reprodução ou divulgação total ou parcial deste documento é expressamente proibido sem o consentimento formal, por escrito, do Professor (autor).