

# 03\_Como\_Submeter Projetos

James Hunter, Ph.D.

30 de maio de 2020

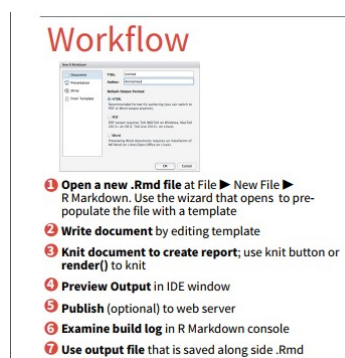
Este capítulo vai descrever como submeter os projetos para *Análise de Dados com R*. Eu quero os projetos num formato como um relatório. Eu não vou executar nenhum código seu. Os arquivos que vocês me enviam devem estar no formato *.docx* ou *.pdf*. Por necessidade, vão incluir blocos de texto e blocos de código no arquivo original que vocês preparam. Este arquivo original vai ser no mesmo formato que este documento aqui, *R Markdown* e vai ter a extensão *.rmd*. Mas, vocês não vão submeter o documento de *R Markdown*, só o resultado do processamento deste documento.

Este documento descreverá os passos para preparar os relatórios. Usarei um pequeno exemplo e anexarei o documento *R Markdown* e o pdf dele. Vocês podem usar este documento *R Markdown* como modelo para o seus.

Pode anotar que este documento, todos os capítulos dei para vocês e os slides 3b e 4 foram preparados usando o formato *R Markdown* no *RStudio*.

## R Markdown Cheatsheet

Também para ajudar vocês com como escrever texto e código neste formato, aqui é o link para o cheatsheet de *R Markdown*: <https://rstudio.com/resources/cheatsheets/> (<https://rstudio.com/resources/cheatsheets/>). Role para baixo até “*R Markdown Cheatsheet*” e clique no “*Download*”. Gostaria de destacar três regiões do Cheatsheet. A primeira mostra o fluxo de trabalho (*workflow*) para preparar um documento *R Markdown*.



R Markdown Fluxo de Trabalho

A segunda mostra um resumo de como integrar código no seu documento.

### Embed code with knitr syntax

#### INLINE CODE

Insert with ``r <code>``. Results appear as text without code.  
Built with ``r getRversion()`` ➔ Built with 3.2.3

#### CODE CHUNKS

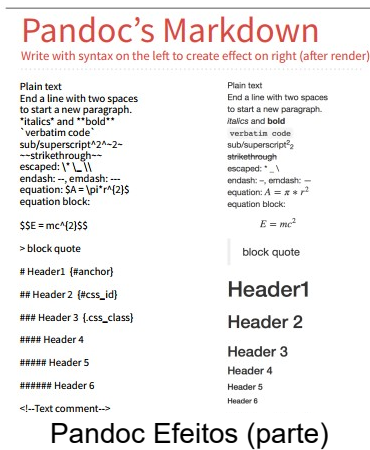
One or more lines surrounded with ````{r}` and `````. Place chunk options within curly braces, after `r`. Insert with `getRversion()` ➔

#### GLOBAL OPTIONS

Set with `knitr::opts_chunk$set()`, e.g.  
````{r include=FALSE}`  
`knitr::opts_chunk$set(echo = TRUE)`  
`````

R Markdown Código

A terceira mostra parte da lista das indicações no texto e os efeitos impressos que eles criam. Este parte do *R Markdown* está chamada *Pandoc* e serve para fazer os “*markups*” do texto. Esses efeitos não funcionam nos blocos de código.



# Elementos do Projeto

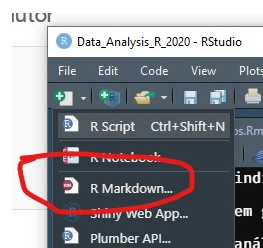
Todos os projetos devem conter os seguintes elementos:

- Importação dos dados em R
- Importar com as funções de `readr` ou `readxl` os dados de um spreadsheet ou criar um tibble dos dados se os dados vêm de dentro de R ou seus pacotes.
- Limpeza dos dados
- Preparar os dados para análise, prestando atenção às lições da Aula 2
- Análise exploratória dos dados
  - Resumos dos dados univariados e multivariados (e.g., correlação)
- Visualizações dos dados
- Análise dos dados
  - Projeto individual: regressão simples linear
  - Projeto em grupo: regressão múltipla ou outro modelo de ML

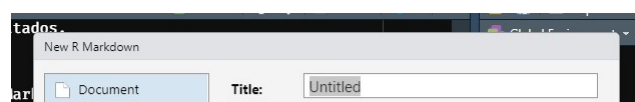
Cada fase da análise deve ser acompanhada por texto que explica ou que você fez e qual é sua análise em escrita sobre os resultados.

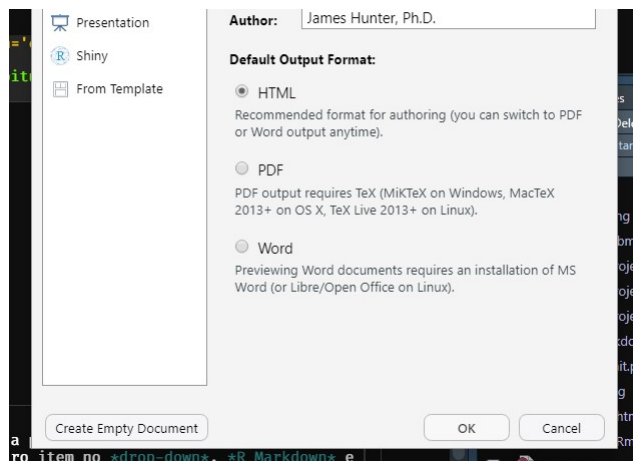
## Fluxo do Trabalho do Projeto

### Passo 1: Abrir um Novo Documento R Markdown



Clique no primeiro ícone a esquerda sob a palavra “File” no canto da tela de RStudio. Um *drop-down* vai abrir. Clique no terceiro item no *drop-down*, *R Markdown* e uma caixa de “New R Markdown” vai abrir.

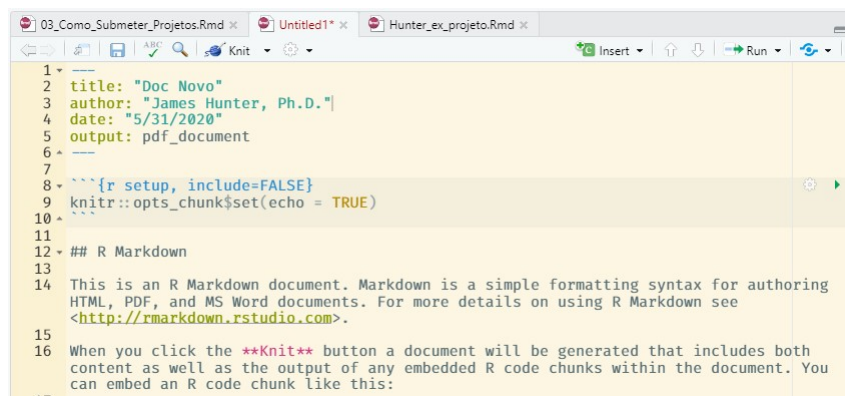




Nesta caixa, você quer ter o item *Document* selecionado na lista a esquerda. Você deve criar um título e RStudio teria lembrado seu nome. Para “Default Output Format:”, você pode selecionar “PDF” diretamente se você tem Latex (*L<sup>A</sup>T<sub>E</sub>X*) instalado no seu computador. Senão, ou se nem sabe o que é Latex, pode escolher a opção “Word”, que vai produzir um documento “.docx”.

**VSS:** Pode submeter os projetos em um desses 2 formatos ou outro.

Quando você clique em “OK”, o novo documento vai abrir como o seguinte.



Esta imagem do documento mostra três elementos importantes: o YAML, um bloco de código e texto com um *header*.

## Passo 2A - Blocos de Texto

Aqui você escreve o texto. Pode usar ornamentos como **negrito** (2 asteriscos em volta da expressão) ou *itálico* (com 1 asterisco) como você quiser. Pode criar uma estrutura com *headers* de diferente graus (aumentando ou diminuindo o número de *hashtags* “#” como quiser). Pode incluir *links* ao internet com a imagem mostra.

**VSS:** Quando você mudar de um bloco de texto a um bloco de código ou vice-versa, deixa duas linhas entre os blocos.

## Passo 2B - Blocos de Código

```
7
8- {r setup, include=FALSE}
9 knitr::opts_chunk$set(echo = TRUE)
10
```

Tem algum cuidado que precisa exercer com um bloco de código. Precisa iniciar um bloco de código com os caracteres três acentos graves (ou *backticks* em inglês) seguido com “{r}” numa linha sozinho e termina o bloco com os três acentos graves (“```”) numa linha sozinha, como mostra o bloco acima. Você pode incluir na linha inicial um nome do bloco (altamente recomendado) e vários argumentos. Você pode achar

os argumentos no R Markdown Cheat Sheet.

Dentro desse início e fim do bloco pode escrever o código que quer usar. Como mostra meu projeto exemplar, eu uso um número de blocos com a regra que cada bloco deve cumprir uma tarefa para eu possa entender o que estou fazendo.

**VSS:** No primeiro bloco de `setup`, você deve acertar que tem a linha de código

`"knitr::opts_chunk$set(echo = TRUE)"`. Essa linha instrui o RStudio para pôr no relatório final as linhas de código. Não posso saber o que vocês fizeram sem ver o código.

## Passo 2C - YAML

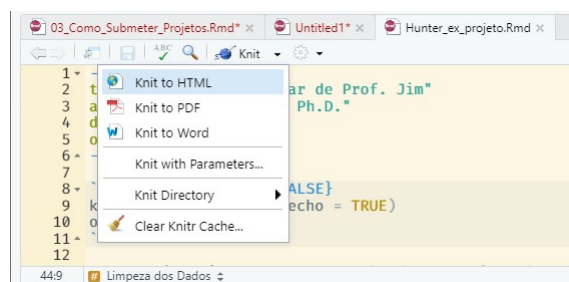
Aquela parte de novo documento que começa e termina com linhas compostas de três hífen ("—") é a parte introdutória do documento e fornece a RStudio e `knitr`, o motor que vai tornar o R Markdown em um produto final, instruções globais sobre o documento. "YAML" quer dizer "*Yet Another Markup Language*" (acho porque precisou um nome e ninguém podia pensar em algo melhor). Aí são o título do documento, o seu nome, a data que você iniciou o documento e um ou mais tipos de *output* que você quer produzir. O campo de `output` segue a opção que você clicou no campo de "*Default Output Format*".

Você pode mudar a vontade qualquer texto dentro desses campos. A única restrição é que você deve manter as aspas onde R colocou elas.

Agora você tem as ferramentas necessárias para preparar seus estudos—Item # 2 na imagem de fluxo de trabalho acima.

## Passo 3 - Como Produzir o Documento Final

Agora, você decidiu que tudo está pronto para imprimir o documento final. Fazemos isso utilizando o botão "Knit" em cima na tela. Quando você clica no botão, um *drop-down* vai aparecer.



Você escolha uma dessas opções, provavelmente aquela opção que você indicou quando você iniciou o documento (PDF ou Word). Uma nova janela vai abrir no lugar do *Console* que vai mostrar o progresso na execução do processo de produzir o produto. Também, se o processo parar por causa de um erro, vai mostrar o que era o erro e onde no seu arquivo ocorreu. Quando termina o processo, RStudio vai mostrar uma cópia do resultado na sua tela no formato que você pediu.

## Lembrete – O Que Precisa Enviar

Você vai enviar 2 arquivos:

1. O resultado em formato de PDF ou in Word (.docx)
2. O arquivo .rmd