

Structure of Rmd files

Second Training on quality checks
15/04/2024

Walter Zupa



COISPA



COISPA



CNR
IRBIM
Istituto Nazionale
per lo Studio e la
Ricerca in
Bioscienze Marine



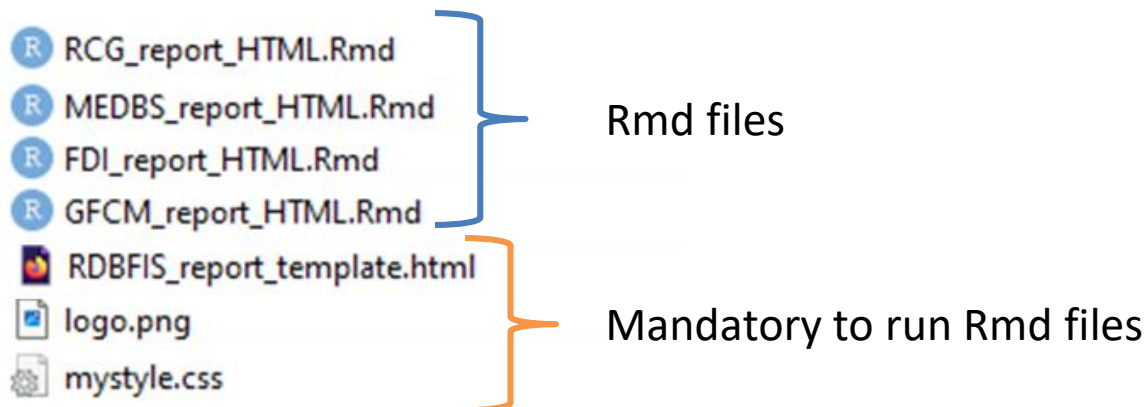
Markdown is a lightweight formatting language for authoring documents in **HTML** and other different file formats, such as **PDF**, and **MS Word**.

This language was created as text-to-HTML conversion tool for web writers.

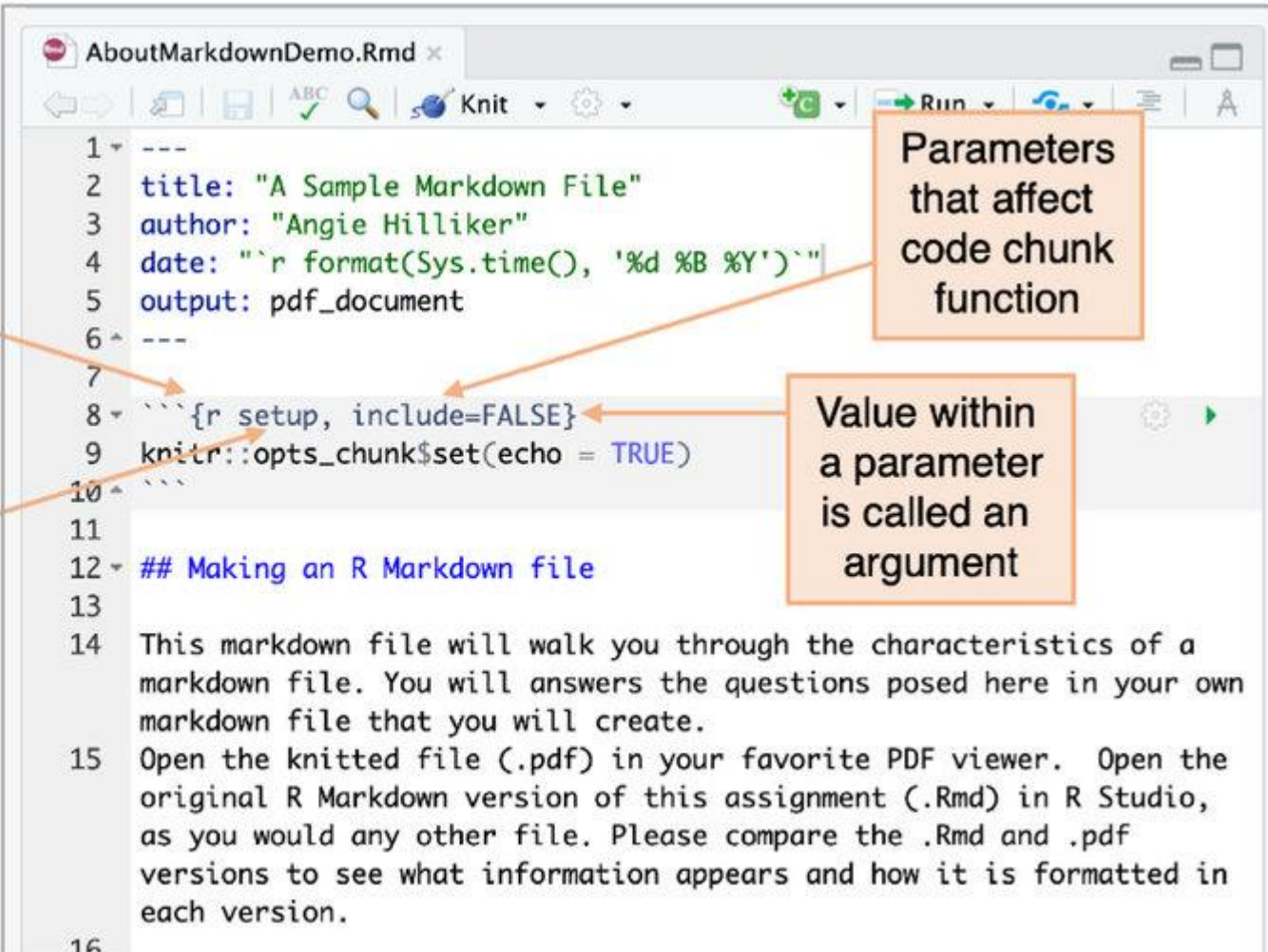
R markdown files are plain text fully reproducible documents with the extension “.Rmd” in which code and text are integrated in the same document. Rmd is used to create dynamic documents with R language in which tables and graphs could be included.

There are some mandatory files that should be located in the same folder of the Rmd files.

Missing mandatory files will produce errors in the compilation of the Rmd files.



The Rmd files are structured in “chunks” containing R code



The screenshot shows an R Markdown file named 'AboutMarkdownDemo.Rmd' in a text editor. The file content is as follows:

```
1 ---
2 title: "A Sample Markdown File"
3 author: "Angie Hilliker"
4 date: "`r format(Sys.time(), '%d %B %Y')`"
5 output: pdf_document
6 ---
7
8 {r setup, include=FALSE}
9 knitr::opts_chunk$set(echo = TRUE)
10
11
12 ## Making an R Markdown file
13
14 This markdown file will walk you through the characteristics of a
15 markdown file. You will answers the questions posed here in your own
16 markdown file that you will create.
```

Annotations with arrows pointing to specific parts of the code:

- Chunk header:** Points to the first line of the chunk, which is a hyphen (`---`).
- Code chunk name:** Points to the text `{r setup` in line 8.
- Parameters that affect code chunk function:** Points to the parameter `include=FALSE` in line 8.
- Value within a parameter is called an argument:** Points to the value `TRUE` in line 9.

YAML section (YAML Ain't Markup Language)

```
---  
title: ·"![] (logo.png) {width=6in}<br>\\vspace{2in}\\n\\n\\n\\n`r ·format (params$term) `·  
Report\\vspace{0.1in} ·"  
·  
date: ·"Compiled ·on ·`r ·format (Sys.time(), ·'%d/%m/%Y, ·%H:%M') ·"  
subtitle: ·'·  
·  
output:  
· ·html_document:|  
· · ·template: ·RDBFIS_report_template.html  
· · ·toc: true  
· · ·toc-title: ·"Table of contents"  
· · ·toc_depth: 5  
· · ·number_sections: true  
· · ·collapsed: true  
· · ·df_print: paged  
params:  
· ·term: MEDBS ·datacall  
---
```

Environment initialization

```
<!-- INITIALIZATION -->
```

```
`{r·inizialization,·include=FALSE}~  
knitr::opts_chunk$set(~  
·echo = FALSE,~  
·message = FALSE,~  
·warning = FALSE~  
)~  
``~
```

```
`{r·libraries,·include=FALSE}~  
# loading needed libraries~  
lib<-c("RDBqc", "knitr", "kableExtra", "dplyr", "ggplot2", "rworldmap", "sp",  
"rworldxtra", "pander", "data.table", "grDevices", "magrittr", "tictoc", "tidyverse",  
"fishmethods", "tidyr", "gridExtra")~  
lapply(lib, require, character.only = TRUE)~  
``~
```

User initialization



Check the data separator set in the files!!!!

```
<!-- USER DEFINED SOURCE FILES -->
{r setup, include=FALSE}

# reading files
# UNCOMMENT THE FOLLOWING LINES FOR STAND-ALONE USE
# SELECT THE APPROPRIATE FILE PATHS ON THE LOCAL FOLDERS

Catch <- read.table("./catch.csv", sep = ",", header = TRUE)
Land <- read.table("./landings.csv", sep = ",", header = TRUE)
Disc <- read.table("./discards.csv", sep = ",", header = TRUE)
ML <- read.table("./ml.csv", sep = ",", header = TRUE)
MA <- read.table("./ma.csv", sep = ",", header = TRUE)
SL <- read.table("./srl.csv", sep = ",", header = TRUE)
SA <- read.table("./sra.csv", sep = ",", header = TRUE)
GP <- read.table("./gp.csv", sep = ",", header = TRUE)
ALK <- read.table("./alk.csv", sep = ",", header = TRUE)

SPs <- c("MUT")
MS <- "ITA"
GSAs <- "GSA_18"
end_year <- 2021
}
```

```
<!-- DO NOT MODIFY THE FOLLOWING CODE -->

```{r_setup, include=FALSE}
check the presence of user defined filter for member state (MS)
if (exists("MS")) {
 if (length(MS)==1 & is.na(MS[1])) {
 user_MS = FALSE
 } else {
 user_MS = TRUE
 }
} else {
 user_MS <- FALSE
}
}
```

### Specific checks for each table

```
\newpage
<!-- CATCH -->

```{r_check_catch_data, results = 'asis'}
#
if (exists("Catch")) {
  if (class(Catch) == "data.frame") {
    if (nrow(Catch) > 0) {
      check_Catch <- TRUE
      cat(paste0("\n# Catch table\n"))
      if (!user_SP) {
        SPs <- sort(unique(Catch$species))
      }
    }
  }
  if (!user_MS) {
```

Functions' chunks

```
```{r_catch_disaggregated_records, results = 'asis'}
if (check_Catch) {
 cat("\n## Presence of disaggregated records\n")
 cat("\nData included in the catch table were checked for the presence of disaggregated
records. The result of the check is here reported:\n")
}
disaggregated <- Catch[0,]
```



## MEDBS datacall Report

Compiled on 03/10/2023, 14:48

- 1 Catch table
  - 1.1 No replicated records in Catch data
  - 1.2 Presence of disaggregated records
  - 1.3 Check missing years in time series
  - 1.4 Catch data time series
    - 1.4.1 GSA 17
      - 1.4.1.1 MUT
  - 1.5 Check of Sum of Products (SOP)
    - 1.5.1 GSA 17
      - 1.5.1.1 MUT
- 2 Landing table
  - 2.1 No replicated records in Landing data
  - 2.2 Presence of disaggregated records
  - 2.3 Check missing years in time series
  - 2.4 Landing data time series
    - 2.4.1 GSA 17
      - 2.4.1.1 MUT
  - 2.5 Checks on landing volumes
    - 2.5.1 Comparison of landing volumes aggregated by quarter
      - 2.5.1.1 GSA 17
        - 2.5.1.1.1 MUT
    - 2.5.2 Comparison between landings in weight and volume
    - 2.5.3 GSA 17
      - 2.5.3.1 MUT

## 1.4 Catch data time series

## 1.4.1 GSA 17

## 1.4.1.1 MUT

Coverage of landing data for MUT in GSA 17

country	year	area	species	landings
ITA	2006	GSA 17	MUT	3101.000
ITA	2007	GSA 17	MUT	3298.000
ITA	2008	GSA 17	MUT	3158.000
ITA	2009	GSA 17	MUT	2433.000
ITA	2010	GSA 17	MUT	1796.000
ITA	2011	GSA 17	MUT	1890.000
ITA	2012	GSA 17	MUT	1626.000
ITA	2013	GSA 17	MUT	1979.000
ITA	2014	GSA 17	MUT	2399.000
ITA	2015	GSA 17	MUT	2220.000
ITA	2016	GSA 17	MUT	2042.000
ITA	2017	GSA 17	MUT	2672.000
ITA	2018	GSA 17	MUT	2517.000
ITA	2019	GSA 17	MUT	1698.000
ITA	2020	GSA 17	MUT	1261.773

## 2.13 Total landing time series by quarter

To identify possible inconsistencies in landing data, the species time series of landing volumes is here plotted by quarter.

## 2.13.1 GSA 17

## 2.13.1.1 MUT

Plot of total landings for MUT - ITA - GSA 17

MUT ITA GSA 17 Total landing by quarter

