Introduction to R Markdown & knitr

{% include toc title="In This Lesson" icon="file-text" %}

In this tutorial we will use the `knitr` and `rmarkdown` packages in `RStudio` to create a report that links our analysis, results and associated data.

Learning Objectives

At the end of this activity, you will be able to:

- List benefits of using R Markdown to create reports.
- Explain how R Markdown is a useful tool in Open Science approaches.
- Explain one way that R Markdown can benefit your research.

What you need

Before you start this tutorial, be sure that you have R and R studio setup on your computer. Also we recommend have you have an `earth-analytics` directory setup on your computer with a `/data` directory with it. Follow the links below for help getting R, R studio and your data directory setup.

- How to Setup R / R Studio
- Setup your working directory

Why open science

Open science in a nutshell is about making scientific methods, data and outcomes available to everyone. It can be broken down into several parts (Gezelter 2009):

- Transparency in experimental methodology, observation, and collection of data.
- Public availability and reusability of scientific data.
- Public accessibility and transparency of scientific communication.
- Using web-based tools to facilitate scientific collaboration.

In this tutorial, we are not going to focus on all aspects of open science as listed above. However, we will introduce one tool that can be used to make our workflows:

1. More transparent and
2. More available and accesible to the public and our colleages.

In this tutorial, we will learn how to document our work - by connecting data, methods and outputs in one or more reports or documents. We will introduce the `R Markdown` file format which can be used to generate reports that connect our data, code (methods used to process the data) and outputs. We will use the `rmarkdown` and `knitr` package to write R Markdown files in `Rstudio` and publish them in different formats (html, pdf, etc).

Open Science Slideshow

Click through the slideshow below to learn more about open science. View Slideshow: Share, Publish & Archive Code & Data

About R Markdown

Simply put, `.Rmd` is a text based file format that allows you to include both descriptive text, code blocks and code output. You can run the code in `R` and using a package called `knitr` (which we will talk about next) you can export the text formated .Rmd file to a nicely rendered, shareable format like pdf or html. When you knit (or use `knitr`) the code is run and so your code outputs including plots, and other figures appear in the rendered document.

> "R Markdown (.Rmd) is an authoring format that enables easy creation of dynamic documents, presentations, and reports from R. It combines the core syntax of markdown (an easy to write plain text format) with embedded R code chunks that are run so their output can be included in the final document. R Markdown documents are fully reproducible (they can be automatically regenerated whenever underlying R code or data changes)." – RStudio documentation.

We use R Markdown (.Rmd) files to document workflows and to share data processing, analysis and visualization code & outputs.

Why R Markdown?

There are many advantages to using R Markdown in your work:

- **Human readable:** it's much easier to read a web page or a report containing text and figures.
- **Simple syntax:** markdown and `rmd` can be learned quickly.
- **A Reminder for Your Future Self** All components of your work are clearly documented. You don't have to rememberwhat steps, assumptions, tests were used.
- **Easy to Modify:** You can easily extend or refine analyses by modifying existing or adding new code blocks.
- **Flexible export formats:** Analysis results can be disseminated in various formats including HTML, PDF, slide shows and more.
- **Easy to share:** Code and data can be shared with a colleague to replicate the workflow.

Data Tip: RPubs is a one way to share and publish code online. {: .notice}

RMD is beneficial to your colleagues

The link between data, code and results make `.Rmd` powerful. You can share your entire workflow with your colleagues and they can quickly see your process. You can also write reports using `.Rmd` files which contain code and data analysis results. To enrich the document, you can add text, just like you would in a word document that describes your workflow, discusses your results and presents your conclusions - along side your analysis results.

RMD is beneficial to you & your future self

R Markdown as a format is an efficient tool. If you need to make changes to your workflow, you can simply modify the report and re-render (or knit) the report. This creates an efficient workflow. Your future self will appreciate it too. R Markdown provides documentation for you to see what code you used to create a figure or to analyze the data.

Data Tip: Many of the Earth Lab lessons- including this one - were created using R Markdown! {: .notice}

Use Knitr to convert .Rmd to .html

We use the `R knitr` package to render our markdown and create easy to read documents from `.Rmd` files. We will cover how to use `knitr` later in this lesson series.

```
<a href="{{ site.baseurl }}/images/course-materials/earth-analytics/week-1/intro-knitr-rmd/rmd-file.png
<img src="{{ site.baseurl }}/images/course-materials/earth-analytics/week-1/intro-knitr-rmd/rmd-file.png
</a>
<a href="{{ site.baseurl }}/images/course-materials/earth-analytics/week-1/intro-knitr-rmd/knitr-output
<img src="{{ site.baseurl }}/images/course-materials/earth-analytics/week-1/intro-knitr-rmd/knitr-output
</a>
<figcaption>R Markdown script (left) and the HTML produced from the knit R
Markdown script (right).
</figcaption>
```