File Organization 101

In the previous lessons, we setup R and RStudio. The last part of our setup is to setup our working directory. A **working directory** is an organized space or directory on our computer where we keep our data, scripts and outputs. It is important to think about the organization of that directory, to make our own future lives easier (so we can find things) and also to make it easier to collaborate with other people.

Setup your project

Project organization is integral to efficient research. A well organized project structure will allow you to more easily find components of your project AND make it easier for others you are working with to understand and find data, code, and results. In this tutorial, we will create a well-organized **working directory**.

Learning Objectives

At the end of this activity, you will:

- Be able to describe the key characteristics of a well structured project.
- Be able to summarize in 1-3 sentences why good project structure can make your work more efficient and make it easier to collaborate with colleagues.
- Be able to explain what a working directory is.

What you need

You will need the most current version of `R` and, preferably, `RStudio` loaded on your computer to complete this tutorial.

- How to Setup R / RStudio

Characteristics of a Well Structured Project / Working Directory

Please note that in this lesson, we will be using our project directory as our working directory. Thus these terms will be used intechangably throughout.

1. Organization - Files & Directories

When it comes to structuring the names of the files and folders that create your project, the more self explanatory, the better.

```
<a href="{{ site.baseurl }}/images/slide-shows/intro-rr/basmati-rice.png">
<img src="{{ site.baseurl }}/images/slide-shows/intro-rr/basmati-rice.png" alt="basmati rice label on c
<figcaption> A well structured project uses directory (folder) names that describe
```

the contents of the directory. Source: Jenny Bryan, Reproducible Science Curriculum.

A well structured project directory should:

- Utilize a naming convention that is:
- **Human readable** - use directory names that are easy to understand.
- **Machine readable** - avoid funky characters OR SPACES.

- **Support sorting** - If you have a list of input files, it's nice to be able to sort them to quickly see what's there and find What you need.
- **Preserve raw data so it's not modified:** We'll worry about this later.
- Have easy to read directory names that contains components of the project (e.g. code, data, outputs, figures, etc)

```
<a href="{{ site.baseurl }}/images/slide-shows/intro-rr/file-organization.png">
<img src="{{ site.baseurl }}/images/slide-shows/intro-rr/file-organization.png" alt="good file organiza
<figcaption> Example of a well-organized project directory. Source: Jenny Bryan, Reproducible Science C
</figcaption>
```

Which Filenames Are Most Self-explanatory?

Your goal when structuring a project directory is to use a naming convention that someone who is not familiar with your project can quickly understand. Case in point, have a look at the graphic below. Which list of file names are the most self explanatory? The ones on the LEFT? Or the ones of the RIGHT?

```
<a href="{{ site.baseurl }}/images/slide-shows/intro-rr/human-readable-jenny.png">
<img src="{{ site.baseurl }}/images/slide-shows/intro-rr/human-readable-jenny.png" alt="example of huma
<figcaption> Compare the list of file names on the LEFT to those on the right
```

which ones are easier to quickly understand? Source: Jenny Bryan, Reproducible Science Curriculum.

Consider the structure of your project as we build the project or working directory for our earth analytics tutorials in the next lesson.