

Statistics and Data Science Workshop Series

Friday March 3





**Today's topic: how to
manage your project's
files/repository**



Motivation

Unless you are working alone, you need a system with which to share files with your team.

Some common file-sharing platforms include:

- GitHub
- Google Drive
- Posit Cloud



Good practices

Agree upon naming conventions for files/variables with your team. Here are some common ones:

- Snakecase: `variable_one`, `variable_two`
- Kebabcase: `variable-one`, `variable-two`
- Pascalcase: `VariableOne`, `VariableTwo`
- Camelcase: `variableOne`, `variableTwo`
- Hungarian Notation—describes variable type or purpose at the beginning
 - `arrDistributeGroup` // Array called "Distribute Group"
 - `sUserName` // String called "User Name"
 - `iRandomSeed` // Integer called "Random Seed"



When referencing other files, do not use your computer's specific filepath. Instead, assume that the main repository file is your working directory. For example, instead of:

```
data <-  
read.csv("~/Users/anyamac/Desktop/GitHub/sdsw-project/data/data.csv")
```

You could write

```
data <- read.csv("~/data/data.csv")
```

The goal is to reduce confusion and make it easier to reproduce your work on other computers.



Miscellaneous tips

- If you're unsure what to do, Google it!
- If you are using randomly generated numbers or data, make sure you are setting a seed for reproducibility.
- Have a groupchat on some platform so that no one works on the same file at the same time. Also just communicate.
- COMMENT YOUR CODE!
- Be consistent with tabbing and whitespace conventions.



README.md files

We will talk about this more later, but your repository should at minimum have a README file. It should give any user a general overview of the purpose and what your project is doing.

The background is a solid orange color. In the top-left corner, there are three vertical bars of varying heights, each composed of several overlapping semi-transparent orange circles. In the bottom-right corner, there are four vertical bars of varying heights, also composed of overlapping semi-transparent orange circles.

Demo: RMarkdown and GitHub