

Three main principles for reproducibility

- ▶ organize for reproducibility from the beginning
 - naming scheme for directories (folders)
 - naming and placement scheme for files
- ▶ explicitly link files
 - using the naming schemes in the scripts
- ▶ don't repeat yourself (DRY)
 - be alert for copy/paste of code
 - source materials should be read once only

I've planned the Project 1 file structure for you

Project directory

```
me497-project-1-YourLastName\  
|-- data\  
|-- reports\  
|-- resources\  
|-- results\  
|-- scripts\  
|  
|-- me497-project-1-YourLastName.Rproj  
|-- .Renvirom  
|-- .gitignore  
`-- README.md
```

Sub-directories

```
data\  
|-- 007_wide-data.csv  
|-- 01_calibr_data_active-report.csv  
`-- 02_calibr_data-tidy.csv
```

```
reports\  
|-- 06_calibr_report.docx  
`-- 06_calibr_report.Rmd  
  
resources\  
`-- load-cell-setup-786x989px.png  
  
results\  
|-- 01_calibr_data-wide.csv  
|-- 03_calibr_graph-draft.png  
|-- 04_calibr_outcomes.csv  
`-- 05_calibr_graph.png  
  
scripts\  
|-- 01_calibr_data-wide.Rmd  
|-- 02_calibr_data-tidy.Rmd  
|-- 03_calibr_graph-draft.Rmd  
|-- 04_calibr_regression.Rmd  
`-- 05_calibr_graph.Rmd
```

All this for a 3-page word.docx report!

Explicitly link files

For example, the first script

```
01_calibr_data-wide.Rmd
```

reads the raw data file

```
data_received <- read_csv('data/007_wide-data.csv')
```

and writes two files

```
# to data directory for subsequent scripts  
write_csv(data_received, "data/01_calibr_data_active-report.csv")  
  
# to results directory for final report  
write_csv(input_output_data, "results/01_calibr_data-wide.csv")
```

The reason there are so many scripts

- ▶ a script should perform one task
- ▶ short, between 60–100 lines
- ▶ generally produces one object written to file, e.g., CSV, PNG
- ▶ simplifies editing, testing, readability, debugging

Run the scripts in order to compile the report

01_calibr_data-wide.Rmd

02_calibr_data-tidy.Rmd

03_calibr_graph-draft.Rmd

04_calibr_regression.Rmd

05_calibr_graph.Rmd

06_calibr_report.Rmd

Compiling the scripts can be done

- ▶ manually (OK for small projects)
- ▶ automatically with an R script
- ▶ automatically using make

We'll discuss the automated methods later in the term (section 6.1 in the book).