Exercise: Importing, Exporting, and Between

# **Overview**

For this exercise, you will import/read a data set as a data frame or tibble and inspect the its using different functions in order to understand its contents. You will also use a piping operator to pass the data frame to various functions from **{dplyr}** in order to reorder its columns, standardize variable names, etc. so that you start thinking about initial preparation of the file. You will then export/write that data frame as a file for access later.

Relevant Files:

* champ\_all\_waves\_practice\_subset.csv
* champ\_all\_waves\_practice\_subset.Rds
* gng\_wide.Rds and gng\_long.Rds
* symmspan\_wide.Rds and symmspan\_long.Rds

# **Functions and Libraries**

Load the **{dplyr}** library and all functions that are in your /r/functions directory.

# **Setting up Directories**

If you don’t have them, create:

1. /r/gng/: for go/no-go task code
2. /r/symmspan/: for symmetry span task code
3. /r/demog/: for demographic data code
4. /data/gng/: for go/no-go task data
5. /data/symmspan/: for symmetry span data
6. /data/demog/: for demographic data
7. When you work on other elements of your project, decide as a team what new directories should be named.

# **Importing Data Files to Data Frames**

Do not hard code paths. Use **{here}**.

1. Import "champ\_all\_waves\_practice\_subset.Rds" from your team project and assign it to object CHAMP.
2. Import "gng\_long.Rds" from your team project and assign it to object GNG.
3. Import the data file "symmspan\_long.Rds" from your team project and assign it to object SSPAN.

# **Importing Data Files to Data Frames**

Import sample data file for the *go/no-go* data and the *symmetry span task*. The wide task files will let you know the general structure of how you will subset the file from the all waves data. You will need to pivot this file later. The long file has already been transformed from wide to long format, a wrangling task that you will work on for the project.

# **Inspecting the Data Frames**

1. Identify the number of rows and columns in the CHAMP.
2. Identify the names of the columns in the CHAMP.
3. Identify the variables that are id variables for subject, school, and wave. You will need to use these for another step.

# **Fixing Variable Names and Locations in the Data Frame**

For CHAMP:

1. Rename the variables so that all are in lower case.
2. Move the id variables (subject, wave, school) to the first three columns in the data frame.

# **Fixing Character Vector Variables**

For CHAMP:

1. Remove the spaces, if any, in the subject id variable.
2. Remove the spaces, if any, in the school id school.

# **Exporting the Data Frame**

For CHAMP, use the correct exporting functions to:

1. Write out the data frame as: "class\_practice\_all\_data" as a .csv file.
2. Write out the data frame as: "class\_practice\_all\_data" as a .Rds file.