

# Practice Problems #2

*Alvin D. Jeffery, PhD, RN*

*Last Updated: 2018-05-02*

## Contents

|                                    |          |
|------------------------------------|----------|
| <b>Load and Review Data</b>        | <b>1</b> |
| Filter by Follow-Up Time . . . . . | 1        |
| <b>Working with paste()</b>        | <b>1</b> |
| <b>Filter by First Study Date</b>  | <b>1</b> |
| <b>Removing Duplicates</b>         | <b>2</b> |

## Load and Review Data

This is a dataset that Maddie had shared with us. (Maddie can provide overview.)  
Begin by reading in and exploring the data.

### Possible Approach

## Filter by Follow-Up Time

Let's say we want to only look at certain follow-up times.

### Possible Approach

## Working with paste()

Notice how long these file paths (directories) are. It can be a burden to have to continue typing them again and again. Additionally, it makes the code a little messy. Let's work on storing the directory in a variable and then calling it in future functions.

This could also help us quickly filter & save files for multiple follow-up times. Let's pretend we want to filter our data to each unique time point, and we want a separate file for each. Performing this task could be as simple as embedding 2 lines of code within a for-loop.

### Possible Approach

## Filter by First Study Date

In the other file, we want to subset each patient to only their first study day (i.e., find baseline information).

## Removing Duplicates

One last thing I had to do is sort out duplicate study ID's and also how to narrow down the list to only a specific set of patients. For example, the BRAIN data has 821 patients, but this idea density project only has ~240. How to get that 821 to only the ~240.