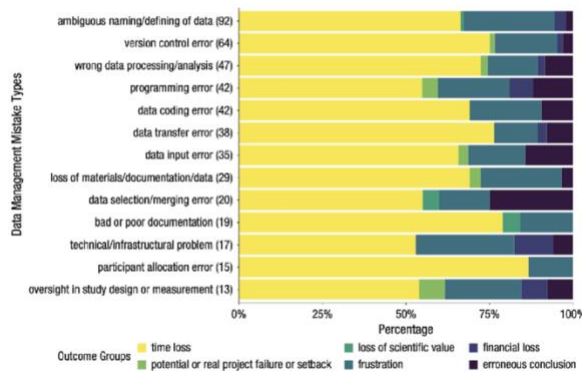


Week #2 Negative Academic Experiences: Preparing the Data for coding pt. 2:
Points one and two are adapted from Shelly Cooper's Lecture Slides for PSYCH 5067

I. Why take all this time to pre-process data?



- This figure came from a larger paper examining the type of errors made when scientists run analyses and the consequences of the errors.
- What surprised you most about the figure?

II. An illustrative example

https://twitter.com/myriam_sander/status/1509076244628971523

Myriam Sander, PhD
This happened 3/30/2022

The screenshot shows a Twitter thread by Myriam Sander, PhD. The main tweet is a 'MANUSCRIPT RETRACTION ALERT' stating that a serious error was discovered in her preprocessing pipeline, affecting two recent papers. The thread includes several replies from other scientists, such as Randy McIntosh and Elko Fried, who express understanding and support for her transparency. Myriam's replies indicate that she is working on the reanalysis and will provide a full report as soon as possible.

- In the immortal words of Hannah Montana "Everybody makes mistakes"

III. Your task: Preprocessing

- You might have noticed that there are some redundant columns. Now we are going to collapse across columns. The sample code is written below.

#collapsing across the "other" option

```
data$event_location <- paste(data$event_location, data$event_location_other, sep = ", ")
```

#deleting the unnecessary column

```
data$event_location_other <- NULL
```