

Naming Conventions

Emilio M. Bruna

1/19/2021

Introduction

The first (and easiest) way to get started organizing your data is to use a simple, clear, and consistent system for naming your files (also called a “naming convention”). Using a well-thought out naming convention has a number of important benefits:

- Files will be easier to find
- You won’t have to open files to see what they are
- Files are easier to sort
- Files are easier to share with collaborators (and for collaborators to use)
- It helps prevent accidentally overwriting or deleting files

How to name your files:

Your file names should:

1. **Tell you about the contents of the file and allow you to uniquely identify it.** For instance (compare 'data 1998' vs 'survey data' vs 'survey data 1998' vs 'small mammal survey data 1998'). You can use things like the an acronym for the project, abbreviation for the study, location where collected, the name of the investigator, the year or month when collected, the type of data type, etc.
2. **Be as short as possible.** Many database systems (like MS One Drive) have limits on file name length. Use a max of 25 characters, and aim for less than that.
3. **Avoid using special characters and spaces** Don’t use characters like \$ % ^ & # | : . This is really important because software, automated data processing tools, web browsers, and computer operating systems use spaces and special characters for dividing up (parsing) strings of text, as well as other processes. Instead of a space, use underscore (_) or dashes (-) to separate meaningful parts of file names.
4. **Start with letters, never numbers.** . Using numbers at the start of file names can lead to problems when they are read into stats software and other programs. Instead of '2020_survey_responses' use 'survey_responses_2020'
5. **Help you quickly sort files chronologically or numerically.** Dates are very useful in file names because they can help identify the most recent files easily. You can *not* rely on the file date associated with the file on the computer! However, it is really important to use a clear format for dates, especially if collaborating internationally. I learned the hard way that not everyone reads 'census_data_3-4-2018' the same way I do, so I switched to 'census_data_4march2018'. However, even this isn’t ideal because of how it sorts things in a folder. Now I now use this format: 'file_name_yearmonthday' ('census_data_20180403'). You can also use file_name_year_month_day. In addition to the

benefit of sorting more easily, you'll never have to use the the dreaded '`_____.final.docx`' and '`_____.actual_final.docx`' in your file names ever again...you'll always know which is the last one. If the files don't require a date, but you want to maintain a sequence, use leading zeros to maintain sort order. 7 and 701 don'ty sort the way you expect them to, so use 007 and 701.

6. **Use a consistent method of dealing with spaces.** . For a the spreadsheet with responses to a survey by 5 different people, don't use '`survey responses maria`'. Instead use (i) '`survey_responses_maria`' or (ii) '`SurveyResponsesMaria`'. Try to avoid mixing the two (e.g., '`survey_responses_Maria`'). By the way, (i) is known as `pothole` and (ii) is known as `camel`. Personally, I prefer `pothole` name is better because of it is easier to read and you never to remember what words get capitalized (because none of the do).

Once you have a system, write it down, print it out, and stick it up somewhere. Remember -whenever possible make the names simple, informative, and unique. If you collected survey data and behavioral data in 2017, don't call them both `data_2017` even if they are in different folders. Call one '`behavior_obs_2017`' and the other '`survey_responses_2017`'. Simple names also help you avoid problems with “system and software portability” - you might be working on a Mac but your collaborators are working on Linux or Windows, and the software made for each can read file names in different ways. File names that are simple, lower-case, and have no special characters are less software- and platform-dependent.

Tools & Resources

1. **Bulk Renaming of files**
 - Adobe Bridge
 - Bulk Rename
 - Renamer 5 (macOS)
 - PSRenamer
 - Batch Renaming of Files in Windows
 - Batch Renaming of Files in MacOS