

## **File Organization Worksheet**

Now that we have reviewed replication, efficiency and elegance in combination with how to plan, organize, and document our work, let's go through an exercise where you will plan out and organize your own data management filing system.

### **First subdivision: Preparing primary data or analyzing data**

Before beginning your plan, there are a few segregations we need to make between filing systems. For example, many STATA users will only ever use secondary data files while others use both secondary data and need to collect and prepare primary data. The project type is the first logical distinction between filing systems. The two types entail very different filing structures.

### **Second subdivision: Project folders**

The central unit of any system is the project folder. Scott Long suggests using a mnemonic for each project folder, such as *cwh* for a project on cohort, work and health (Long 2009: 20-21). Though it doesn't have to be an acronym, use a project title that is short, simple, and unique that aids in your recall of what the project entailed.

For data that is already cleaned in a general sense either by you or by other researchers, I suggest you separate folders by the anticipated deliverable (such as a paper or research report) and not by the dataset itself that is used. You may work with one dataset in many different ways for over a dozen papers. Download a clean copy of the data into each project folder so you're able to separate the project and its individual goals effectively.

### **Subfolders**

Whether in primary or secondary data, users will want to divide their project folder into various subfolders meant to maximize organizational efficiency.

For a large project where you will perform data collection, your project folder will have extensive subfolders regarding administration, etc. **Brainstorm below a list of what types of files you will likely have and come up with a structure for them.** Too many or too little folders may impede efficiency because you will spend more time to find your items. Example subfolders many require include: administration, grants, training, access databases for data entry, codebooks, surveys, etc. Although you may not personally have conducted your own survey, this same filing system may be necessary for a qualitative project or in merging datasets together.

For data that is ready to be analyzed, whether primary or secondary, subfolders are equally important. The subfolder structure here will likely vary from the subfolder structure for data collection projects. Use what you learned about planning a project here and work backward: if your final deliverable for this project is an academic paper accepted for publication at a journal, what steps will you need to take to accomplish this? A researcher will normally require subfolders such as: literature, data, analysis, paper (including drafts), and other more specialized folders. **Brainstorm what a normal project might look like for you here:**

For this second type of organization, some researchers like Jenn Earl will organize their project folders into superordinate folders based on the project status. Earl breaks her projects down into: in development, in progress, in review, in revisions, and in production. Once a paper is published, she moves the project folder into an archive. This way of organizing is ideal for an academic who is working on a lot of different projects simultaneously so they are able to immediately see the status of the project.

Now that you have your general structures planned, create an empty set of folders that you can use as a template for your new system. When creating these folders, did anything feel out of place? Go ahead and rearrange it now. You may want to duplicate a folder for a project you have been working on or completed and organize them within your subfolders. Now is the perfect time to add an additional subfolder, such as a subfolder for documentation, correspondence, or a folder specific to a submission to a journal or conference. Note below an additional folders you added.

The most important thing in creating your filing system is to use something that works for you. If it doesn't work well for you, you will not use it and it may cause more disorganization. Don't be afraid to let your system evolve with you as your types of projects and demands on the system change. Remember to always plan, organize, and document your work in order to maintain replicability, efficiency, and elegance in your academic endeavors.

References:

Long, Scott. 2009. *The Workflow of Data Analysis Using Stata*. College Station, Texas: STATA Press.