Creating a directory structure:

The main breakdown for a working file directory structure is to set the difference between files that are to be Posted (i.e. fixed, not changed), and those that you are 'working in,' that may be altered. As the first principle of reproducible research, the goal is to move files from the working to the posted categories to the point that all files are posted and the entire project can be reproduced easily without confusion.

- `Project Name` (e.g. for a paper on Rural Mathematic Gaps = RMG)
    - `Documentation` (Log File*, Codebooks*, IRB, outlines, etc)
    - `Working` (Files you are working on. You may further use sub directories or use filenames to differentiate the main categories below)
        - `ToDo`. A list of tasks to work on, with directions to which files.
        - `Syntax` – The following subsections can be used and can be differentiated with file names.
            - Data import
            - Cleaning
            - Variable Construction
            - Descriptives
            - Analysis
            - Scratch – This can be used for experiments. Note that if you use a portion, you must add it to the relevant syntax before it is moved to the posted section
        - `DataDerived`. Note derived here because the original source data must be archival and will be in the posted file
        - `Figures/Tables`. Export results here
    - `Posted` (Files that are finished/archival, can run without change from start to finish, or have been submitted)
        - `Syntax`
        - `Data`
            - `Derived`
            - `Source`
        - `Figures/Tables`

`* Log File:` The log file is a record of every decision made with data and towards which goal. It should have the date you worked on something, which file you used, and what decisions were made. Think of this as a journal of your decisions. You should be able to retrace your steps here.

* `Codebook`: Similarly, if you are creating a lot of variables and need a place to sort them and describe them, use a spreadsheet to name each variable and add comments on 1) how it was calculated, 2) What it is used for, and 3) what is meant by null/na if necessary

Project Example				
	Level 1	**Level 2**	**Level 3**	**Purpose**
/ RMG				
	/Docs			Documentation Folder
		Log.docx		Log file of data decisions and progress
		Codebook.xlsx		Variable Names, how calculated, and purpose
		/IRB		IRB Materials
		/Draft		Written Draft of Paper
	/ Working			Folder for Items being worked on
		ToDo.docx		List of tasks to do. Good to work backwards from research questions
		/Syntax		Folder for Syntax
			RMG_Import.do	
			RMG_Clean1.do	
			RMG_Vars1.do	
			RMG_Descript1.do	
			RMG_Analysis1.do	
			RMG_Scratch.do	
		/DataDeriv		
			RMGFull.dta	Data derived from the main set.
			RMG_Sc_Collapse.dta	Data collapsed at the school level for certain analyses
		/Figures_Tables		
			AllFormatted.xlsx	All Tables Formatted for manuscript
			Descriptives1.xlsx	
			Models1-5.xlsx	
	/Posted			Folder for Fixed/Finished files
		/Dataset		
		/Syntax		
		/Figures_Tables		Final Tables/Figures
		/ Submission1_EAQ		Submission to Journal
			Manuscript.docx	
			Cover Letter.docx	
		/Submission2_ER		
			Manuscript.docx	
			Cover Letter.docx	