

Assignment 1: Submit a data set & readme.txt file

Due: end of day, **Tuesday, September 3rd**

For week 2, you will need to have a data set you would like to practice with. You will need to have this saved as a text-readable file (e.g., .csv or .txt).

The assignment is to submit a sample dataset as well as a "readme.txt" file. Here are the criteria:

1. Prepare a sample dataset in saved as comma-separated or tab-delimited file (.csv or .txt).

You can provide a truncated dataset if you would like. Try to keep the dataset manageable, with up to 10 columns or so. This sample dataset should be shareable--i.e., don't include data that you want to keep confidential. This does not need to be the data set you work with for your independent project (though it can be), but it may be helpful if you choose data that is in a similar format.

2. Write up a "readme.txt" file that explains what is in the dataset.

This document should briefly explain what the data is, how it was collected (generally), and information on each column of data. The idea here is that a person who did not take part in a project (i.e., me) can understand your dataset completely. This must be a text file (.txt)—i.e., NOT a Microsoft word document.

3. Upload the two files in a single folder (or compressed file) with your last name on it.

Submit both files into the "Assignment 1" subfolder within your folder.

**Why .csv and .txt file formats? The main reason is that they are not proprietary file formats, so anyone with a computer can open these files without specialized software (like Microsoft Office). It is also easier to read-in text files into R and other programming languages. This facilitates sharing of data and makes for good practice in reproducible research.*