

Data Analysis Best Practices

- 1) **Always make a copy** of your original data for safekeeping.
- 2) **Double-check your work.** Make sure your results pass the "common sense" test. If it seems too good to be true, it might be wrong.
- 3) **Never assume** you know what a field means in a database or what a code means. Ask the person you got the data from to go through each of the fields and explain how it is used and what it means (if this information isn't already spelled out in a record layout and/or codesheet)
- 4) When making a public records request for data, always ask for all **record layouts and codesheets** that go with the data. This should be a standard statement that you include in every FOI request.
- 5) **Share your analysis results** (at least the overarching results) with two or more of your sources as a way to check your findings. Do they match with what your sources have seen or would expect? On some occasions you might want to show them very detailed results or show them maps, charts or other information that you might publish. Don't wait until the last minute before publication/broadcast to do this, though, because your sources might point out flaws in the data or underlying factors that you weren't aware of that might have affected your analysis.
- 6) **Document your analysis trail.** There are many reasons why you would want to create a sort of paper trail, documenting each step of your analysis. This is especially important when you're a beginner. Here are some reasons why it's worth it:
 - 1) Sometimes you work on something over the course of many weeks, if not months. It's hard to remember what you did a couple days ago, let alone a couple weeks ago.
 - 2) Writing down all the steps you take in your analysis can help you (or someone else) go back and vet your work or look for where something went wrong.
 - 3) You might need to explain your work to an editor or a colleague at some point down the road. Having it all written down will save you some time.
 - 4) It can serve as a place to consolidate all your findings, making it easier to pull out pieces for your story.
 - 5) Writing down what you did and why will also help you remember and understand the new concepts you're learning.

What to include?

I'd recommend being as detailed as possible. You can make it like a journal where you list the steps you took and then explain why you did it and what you found.

Also include a note at the top about what data you're using, where you got it from (might be helpful to include name/phone number of person you got it from or the URL where you

downloaded it from, etc). Also include a note about the data contents (what years does it cover? What 'universe' does it cover? Are there any notable things that might limit your findings or trip you up?)

You can write down the formulas you used in Excel or the SQL queries in Access, etc. Be sure to write about any data cleaning you did or make notations about fields you might have added (i.e. a field to store just the year portion of the date, or new fields to separate the firstname and lastname of a person).

MaryJo Webster
Mjwebster71@gmail.com
@MaryJoWebster
February 2015