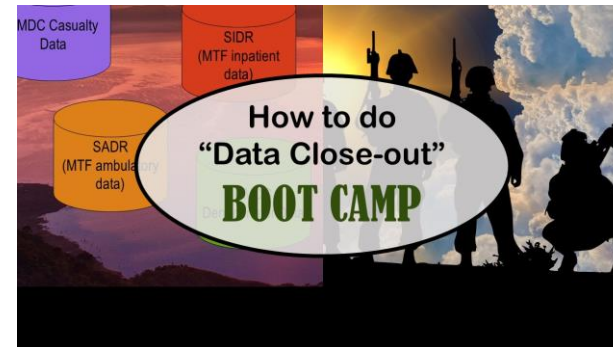




Data Close-out Tips and Tricks

From the online
course



Access course: <https://buff.ly/3Sw4Fsa>

This course is part of an online
data science mentoring program.

Here is a link to learn more: <https://buff.ly/3Wrmw6T>



Data Forensics Tips and Tricks

- 1. Take pictures** with your phone – of data entry, of lab equipment, of space layout, etc.
- 2. Take videos** that can help explain data output (e.g., how a treadmill is gathering data)
3. Save lab datasets, interview laboratorians, and then **reformat lab data** for statistical software (either manually or using programming)
- 4. Take screen shots** of output (such as lab reports), and then annotate them on a PowerPoint slide
- 5. Save code from queries** that extract data from the original data system
- 6. Manually record n's** that researchers report to you so that you can match these up with the number of rows you get in datasets



Tips for Documenting Native Variables from Raw Data

1. **Keep all datasets** you receive (or obtain). You might get them from various sources (software, someone e-mails you, etc.).
2. **Keep documentation** with all the datasets you receive to help you remember the context. E-mail – print to PDF.
3. **Keep** what you receive **in date order**. I use this naming convention as a prefix: 221015 means Oct. 15, 2022.
4. **Curate what you receive** as best you can.

LinkedIn Learning Course - Data Curation Foundations: <https://linkedin-learning.pxf.io/ZnNeg>
5. **Assemble necessary documentation** together to help you. Remember from earlier: grant application, ethics/IRB paperwork, study forms.
6. **Make your own documentation** based on what you learn from interviewing people.



Tips for Make Your Closed-Out Data Usable

- 1. Create transformed variables** that make everyone's life easier – but don't clutter up the dataset with ones no one uses.
- 2. Instruments:** Fully document, create scores, create classifications
- 3. Site-specific data:** Create site-specific documentation, include location information
- 4. Crosswalk files for decoding:** FIPS codes for state or county data, provider data based on ID
- 5. Operationalize** where you can: Have your documentation help users choose variables for certain measurement you know work well
- 6. Codebooks** with frequencies are useful
- 7. Extensive documentation** outside data dictionary (e.g., online, in a report) for complex transformations/data clean-up



Dashboard Development Do's and Don'ts

Do's

- **Do** determine exactly who you want to serve as an audience with your dashboard (a user group?)
- **Do** create static or prototypes designs, and get formal feedback from users
- **Do** create a budget and timeline before developing, use project management approach, and get a sponsor

Don't's

- **Don't** choose a consulting firm before doing the design work by yourself first and writing requirements
- **Don't** do a ton of development without choosing an audience and getting feedback from users
- **Don't** make the dashboard confusing or too busy. Better to have multiple dashboards, or sections of the dashboard