



How to ensure research data quality

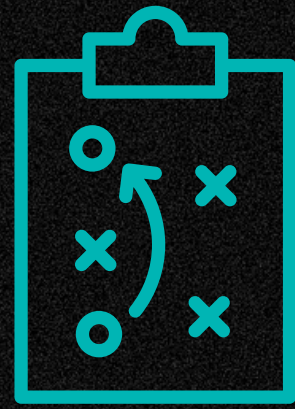
Anusha Pandey and Nishant Pokhrel, 2022 Data Jam Mentors



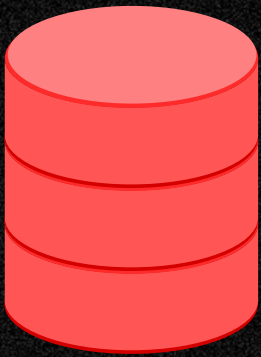
1. Plan properly

An effective way to run a project is by being on top of every single detail.

Everyone in the team should understand the objectives of research questions. This will help team members identify accurate data sources and understand target audience for surveys.



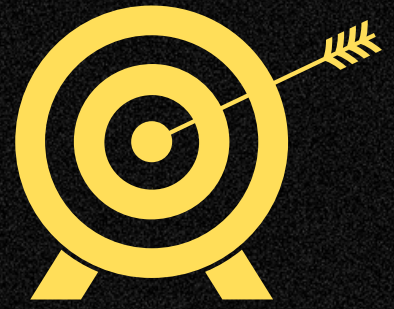
2. Maintain good data storage



It is very easy to lose track of datasets that you come across during research. Make sure your team has a centralized database before starting out so you do not misplace or lose any potential source. You can use Google Sheet/Docs, Noodle tools, Excel Sheets or whatever is convenient for your team.

3. Be consistent in entering datasets

What is your data entry process looking like? Are you copy/pasting your data from sources or are you manually entering values? Maybe you are surveying? Make sure you are consistent while entering your data and that everyone in your team follows the same method.



4. Cite Data Sources Correctly



Data analysis requires citations to help acknowledge original contributors and to help others get to the same resources as you. There are various ways to cite data sources. Make sure you stick with a consistent format throughout your research.

5. Document your data

You might have to revisit data sources in the future. It will be much easier to do so, if you document your rationale for choosing and using particular data sets.

It is also a good idea to make sure your thought processes are captured so that your team members can validate your reasonings or logics behind it as well.



6. Ask for feedback



Your mentors and teachers are always there to help you. If you have questions about any data source, make sure you ask them before analyzing the data; faulty datasets give faulty outcomes. Your data sources should not be biased. They should always be credible.

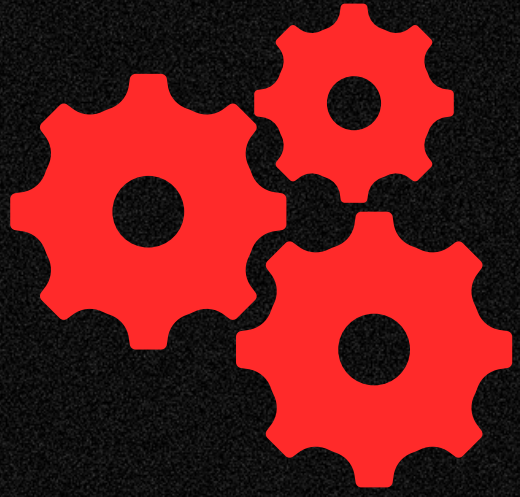
Resources/ Tips

Citation Resources

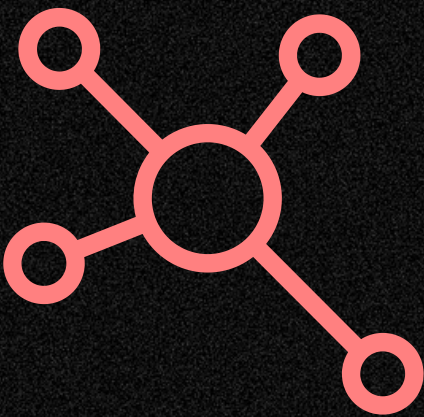
[NoodleTools](#)

[Purdue Owl](#)

[Scribrr](#)



Locating Sources



[Connected Papers](#)

[Credo References](#)

Statistics Databases and Websites

[Statista](#)

[Data USA](#)

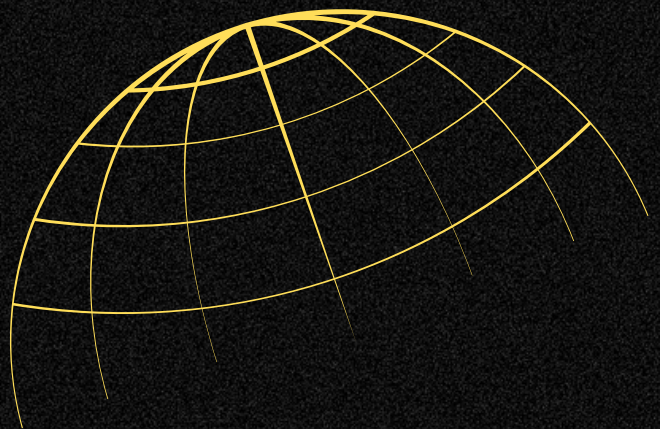
[UN Data](#)

[US Census Bureau](#)

[World Bank Open Data](#)

[World FactBook](#)

[UN Demographic Yearbook](#)



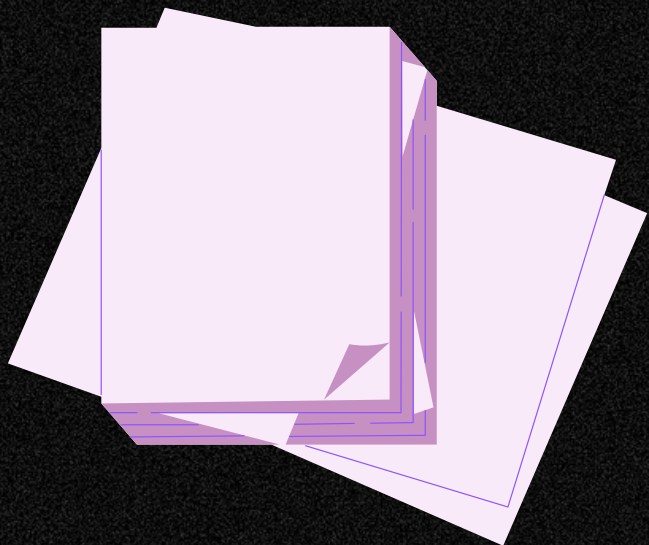
Newsletter

[FlowingData](#)

[The Pudding](#)

[Kaggle](#)

[Dataworld](#)



Subreddits

r/datasets

r/dataisbeautiful

