Fact-Checkable Spreadsheets

It's all about …

documentation

reproducibility

observability

Separation of Source Data and Analysis

- you should always have a reference to the source data that is unaltered
    - so you can compare
    - refer back to
    - verify
    - in the spreadsheet
    - zipped and stored securely
- you should always do analysis in a new worksheet so that you don't alter the source data
    - reference the source data, don't change it

Import Data Thoughtfully and Carefully

- don't let the software wag you
- be aware when your tool automatically does anything to your data
- explore your data in a non-changing way (CLI qsv)
- import as text, then convert columns to specific types of data
- be very careful of dates in spreadsheets
- use a common date format that is clear (ISO 8601)
- be very careful of numbers in spreadsheets
- be very careful of currencies in spreadsheets

Secure Source Data and Other Calculations From Accidents

- Use range or sheet protection
- Principle of least privilege

One Question One Sheet

- try and limit each worksheet to just one analysis or question
- don't undo what you've done to answer another question. Instead, create a new worksheet

Annotate Your Spreadsheets

- naming is framing – you won't remember obscure names in 6 months time
- name your columns well
- name your worksheets well
- add a data dictionary
- don't use color highlighting to encode information, but to make already encoded facets more humanly visible

Create Backups Of Your Work

- never underestimate the ability of technologies to fail
- back up often
- back up to a different location
- create a system for yourself that works