Preparing your data to be visualized

Vanessa Raymond
Axiom Data Science
vanessa@axiomalaska.com
What stands out to you about the best data visualizations?
“Labels on the plot axes.”

Jessica Austin
Software Developer
Axiom Data Science

Image source:
https://portal.aoos.org/?sensor_version=v2cache#platform/045d6976-8686-5583-acbf-a2a28009e83e/v2?pid=13
“Links to explanation, data, additional context”

Chris Turner
Data Librarian
Axiom Data Science

Image source: https://mbon.ioos.us/#default-data/5
“Simple - trying to show one or a few things clearly.”

Chris Turner
Data Librarian
Axiom Data Science

Image source:
https://portal.aoos.org/?#metadata/75373/station
What are the most important components for building a really good data visualization?
“A clear vision of what you're trying to convey”

Will Koeppen
Staff Scientist
Axiom Data Science

Image source:
https://www.nugit.co/what-is-data-storytelling/
“A plan for how to handle missing (or erroneous) data.”

Will Koeppen  
Staff Scientist  
Axiom Data Science
“A group of people who can be a sounding board & a willingness to iterate.”

Will Koeppen
Staff Scientist
Axiom Data Science

Image source:
https://www.wiremedia.net/create-a-data-culture/
“Deciding between a general visualization (e.g., AOOS) and a customized, specially-tailored visualization that will be a standalone experience (e.g., NYTimes).”

Will Koeppen
Staff Scientist
Axiom Data Science

Image source:
Spotlight on a dataset

Historical gridded lightning, Alaska, AK NSF EPSCoR Fire and Ice, (1986 - 2017)
How did we visualize these data?

Historical gridded lightning, Alaska, AK NSF EPSCoR Fire and Ice, (1986 - 2017)
Format is king

- NetCDF-CF
- NetCDF
- .csv using standard headings
“If the metadata in the file is incorrect or non-standard, even in small ways, it can be a real pain.”

- Will Koeppen
NetCDF files are great!

1. Already have a standardized structure
2. Can be read by many modern (programming-language-based) data libraries,
3. Inherently store all metadata necessary to understand the data
Check for compliance

https://compliance.ioos.us/index.html
Cleaning the data

- Need "standard_name" attributes for latitude & longitude
- Impeccable metadata existed in the Research Workspace, but it needed to be added to the NetCDF-CF file
- Standardized the variables from the RW metadata
“There is something special about when you load the data [into the system] and it just works.”

- Will Koeppen
Checklist for data visualization

- Have a story you want to tell with the visualization
- Have a plan for missing or erroneous data
- Use standard headers
- Use a self-defining format (NetCDF-CF)
- Run a compliance checker
- Be ready to iterate to ensure the visualization is telling the right story
Thank you

vanessa@axiomalaska.com