**Bobino Environmental Data Best Practices Module**

June 5, 2018 Taylor Brockman

## Abstract

This educational module is written for educators and students who are ready to learn best practices for analyzing, visualizing, and interpreting environmental data sets using modern technology tools. This module is designed to extend the Bobino lesson plan. Completion of the module requires either an internet connection and web browser if online, or a spreadsheet application (such as Microsoft Excel) and copy of the NOAA data if offline. If possible, each team of students should have a dedicated computer or laptop to follow along with the instructor and complete each of the steps in order to gain experience working with real data sets and analysis tools.

# 1. Questions to Ask Your Data

(EV to help fill in content!)

? Are you comparing or contrasting two things?

? Are you trending changes over time?

? Do you know how your data was collected?

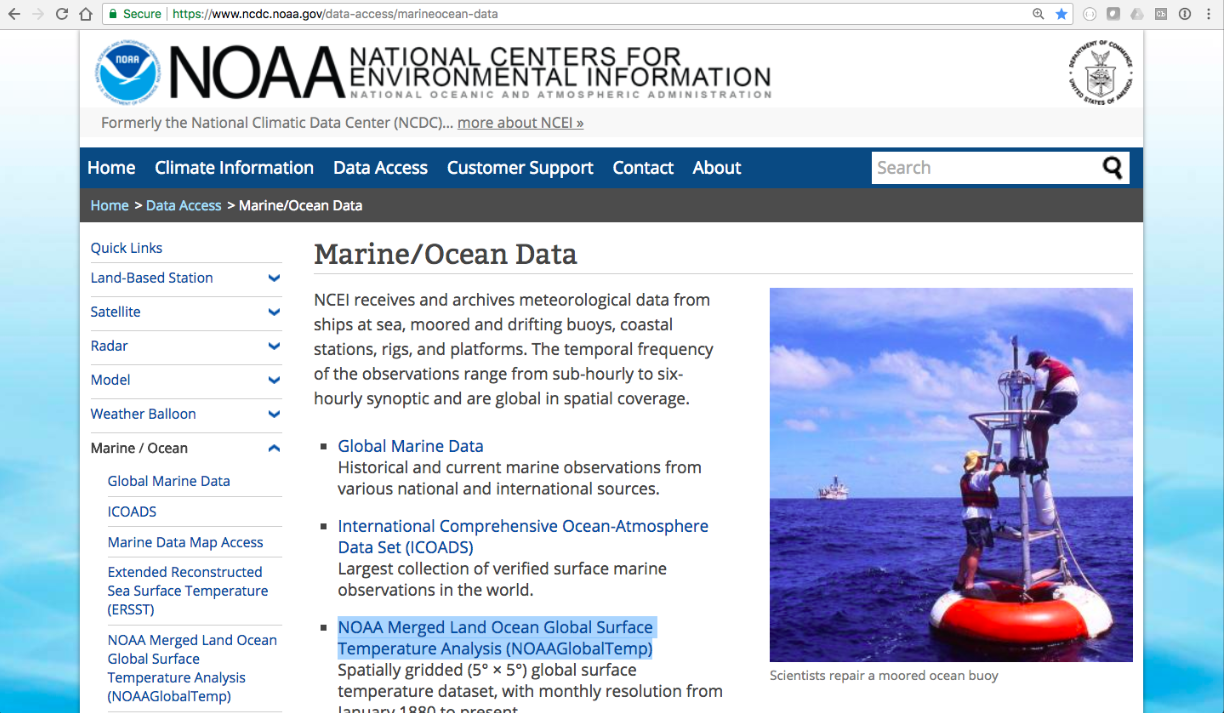
? Is your data pre-filtered or pre-edited or averaged or are they raw data points?

? How could measurement error be introduced into your data set and is it enough to possibly affect your conclustions?

# 2. Acquire Data Sets

NOAA National Centers for Environmental Information

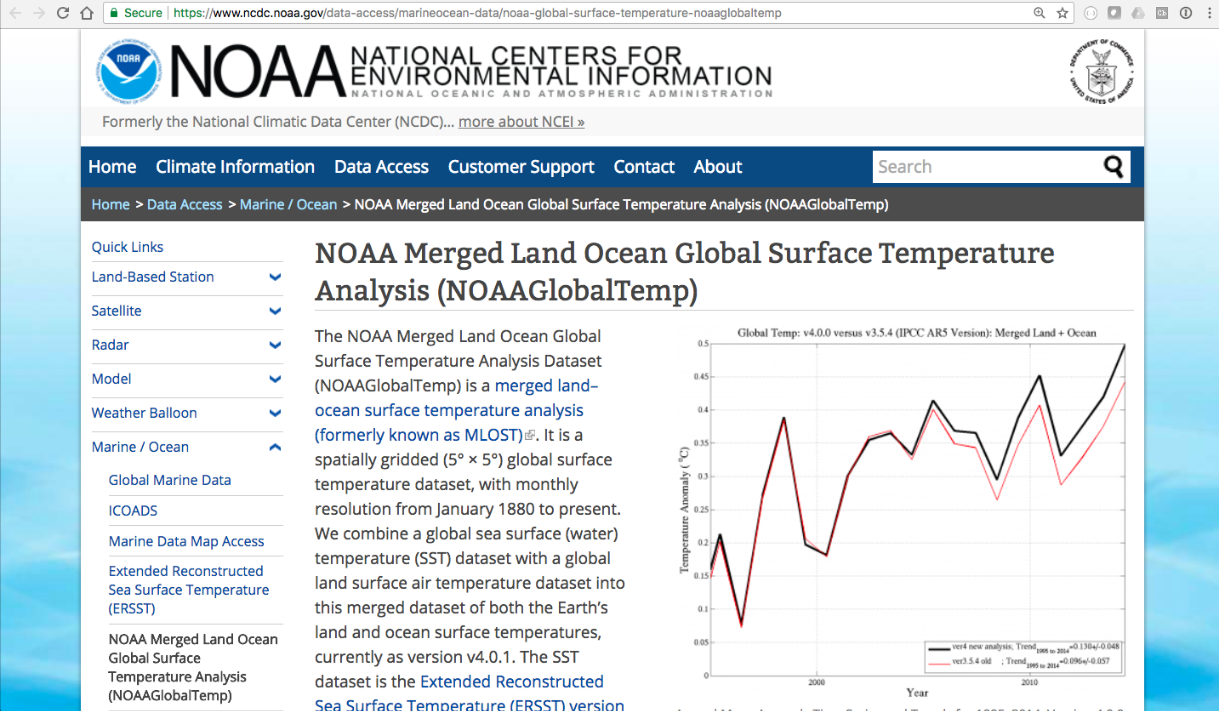
<https://www.ncdc.noaa.gov/data-access/marineocean-data>

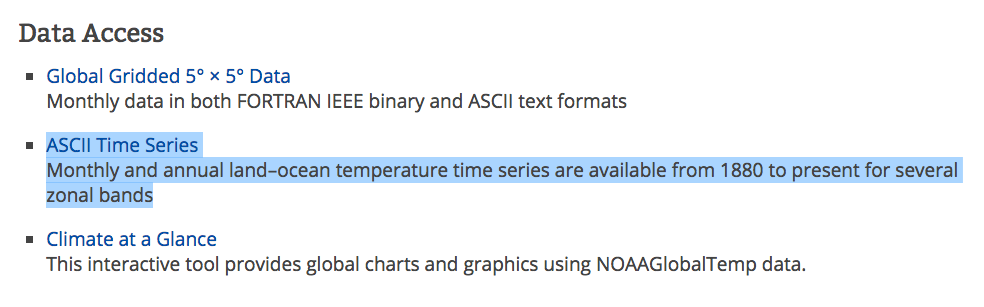


? What is Temporal Frequency?

NOAA Merged Land Ocean Global Surface Temperature Analysis

<https://www.ncdc.noaa.gov/data-access/marineocean-data/noaa-global-surface-temperature-noaaglobaltemp>





Data Access: ASCII Time Series

<ftp://ftp.ncdc.noaa.gov/pub/data/noaaglobaltemp/operational/>

? What is Time Series?

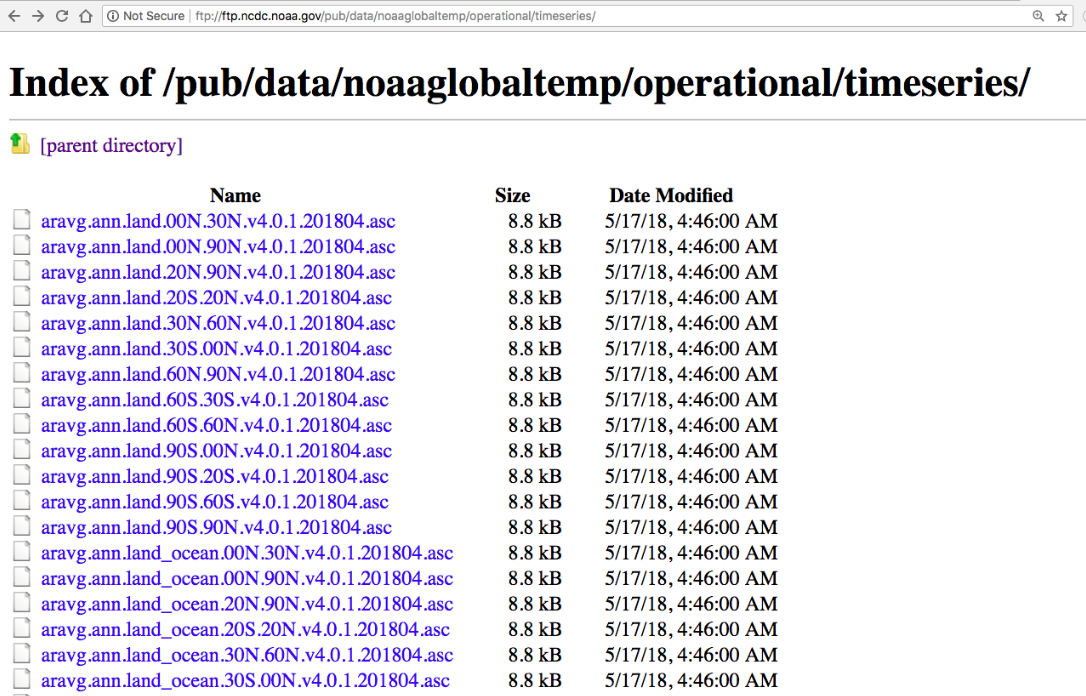
? What is Spatial?

Index of /pub/data/noaaglobaltemp/operational/

<ftp://ftp.ncdc.noaa.gov/pub/data/noaaglobaltemp/operational/>



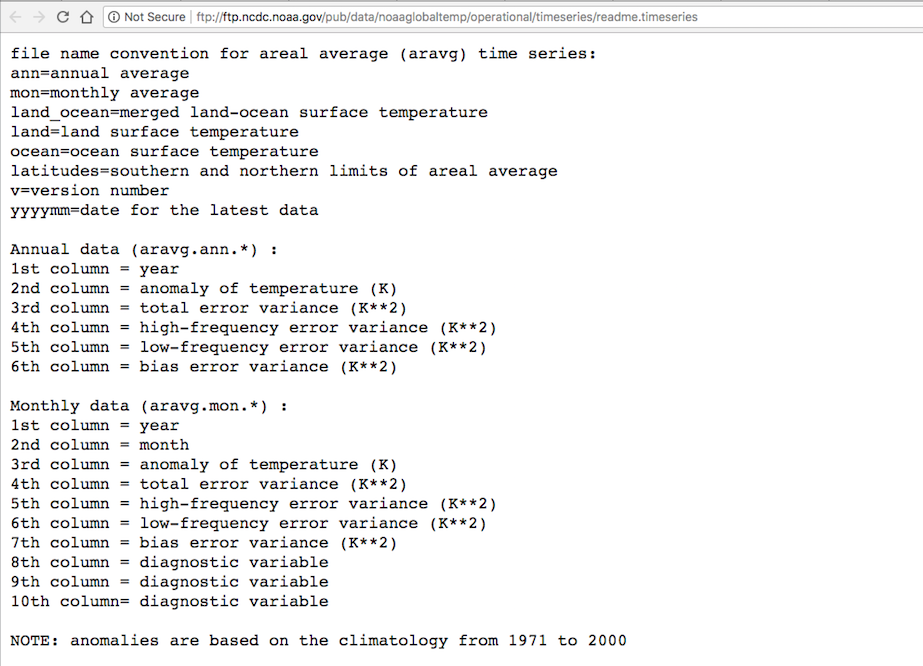
? What is FTP?



? What do the different filenames mean?

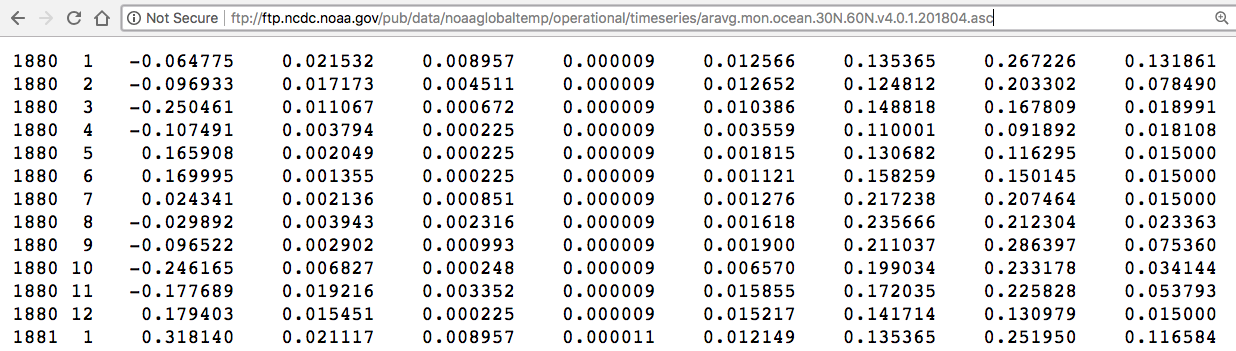
Readme file!

<ftp://ftp.ncdc.noaa.gov/pub/data/noaaglobaltemp/operational/timeseries/readme.timeseries>



? What do the columns mean?

<ftp://ftp.ncdc.noaa.gov/pub/data/noaaglobaltemp/operational/timeseries/aravg.mon.ocean.30N.60N.v4.0.1.201804.asc>



? Which datasets should I download locally?

# 3. Data Set First Impressions

? What are the boundaries of my data? What is in the ‘domain’ and what ‘ranges’ of values are available?

? What format is the data in? CSV, Tab Separated, Column Headers, Binary?

? Do you see any gaps in time?

? Do you see any obvious outliers?

? Is your data “rolled up”, pre-analyzed or are you working with raw data points?

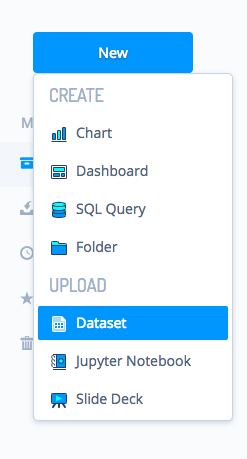
# 4. Online Data Visualization

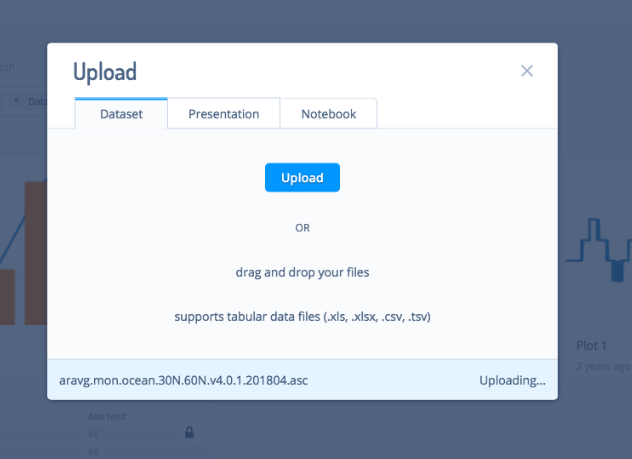
<https://plot.ly/>



Sign up for Free Account

Upload



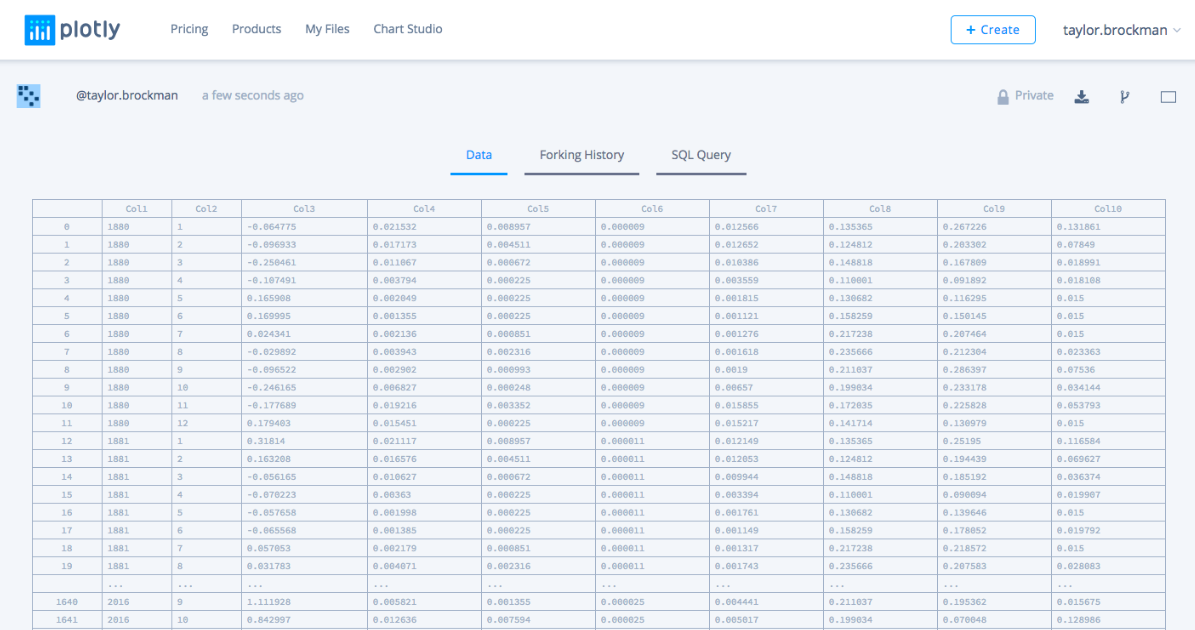


Uploaded aravg.mon.ocean.30N.60N.v4.0.1.201804.asc with no modifications

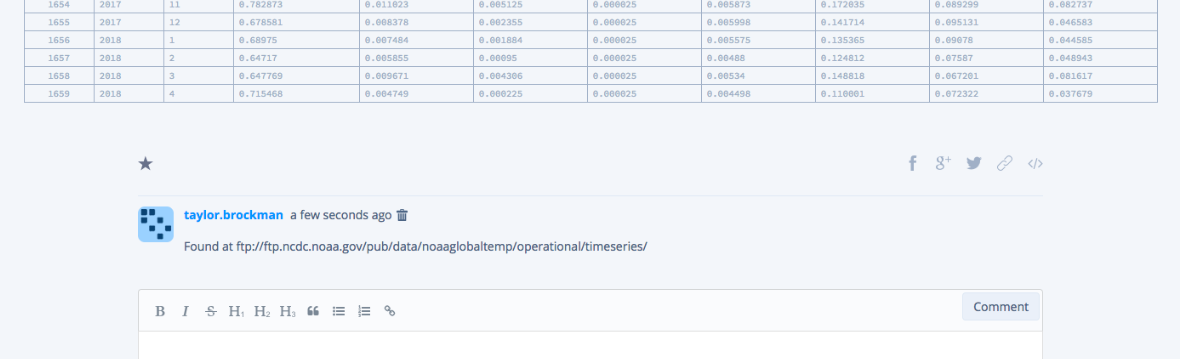
Dataset Appears on your Home Screen



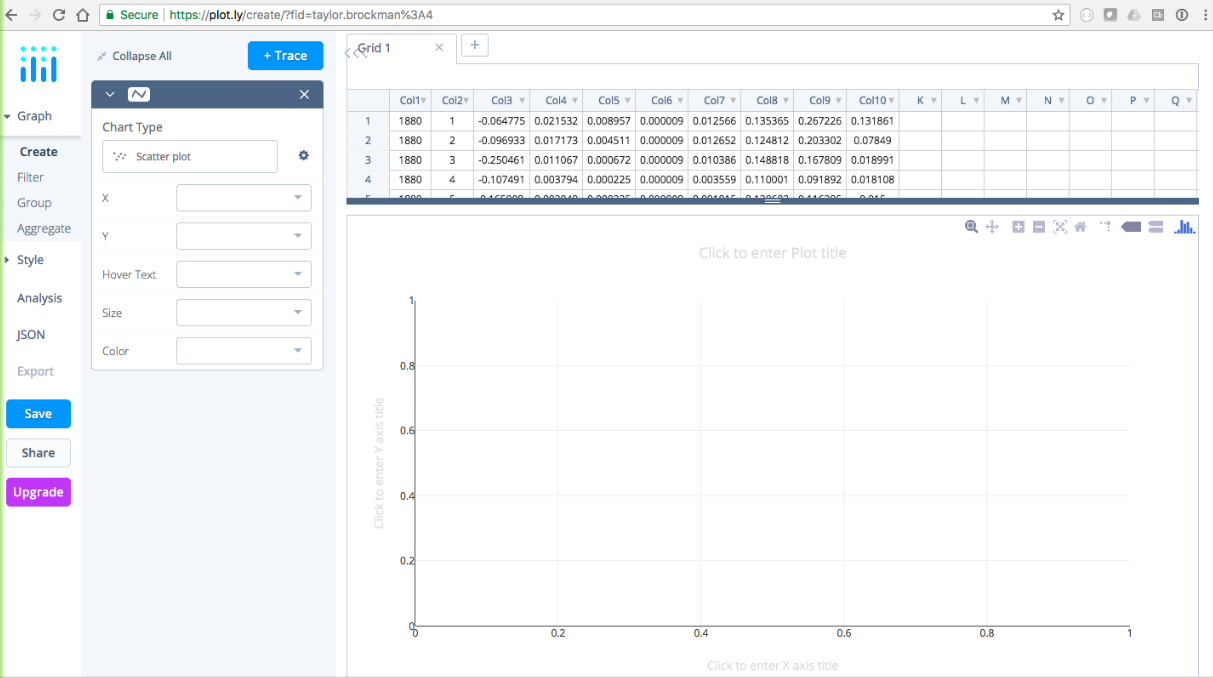
Browsing the Data Set in Detail



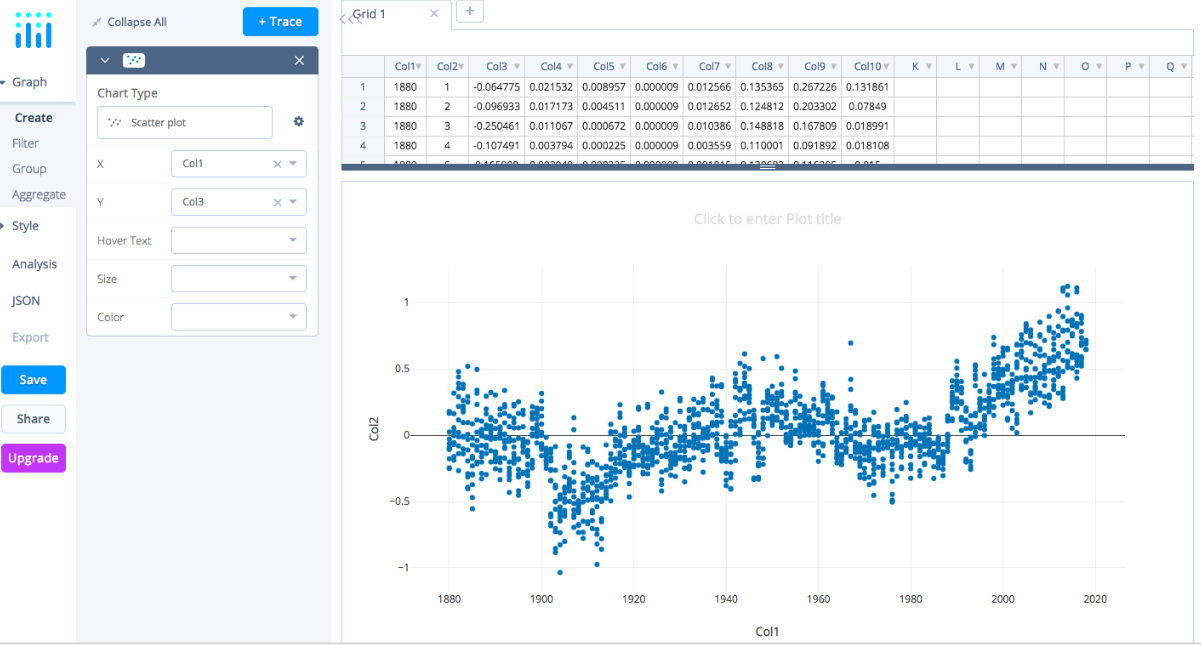
Commenting on the Data Set



Create Visualization



Set Columns into X and Y



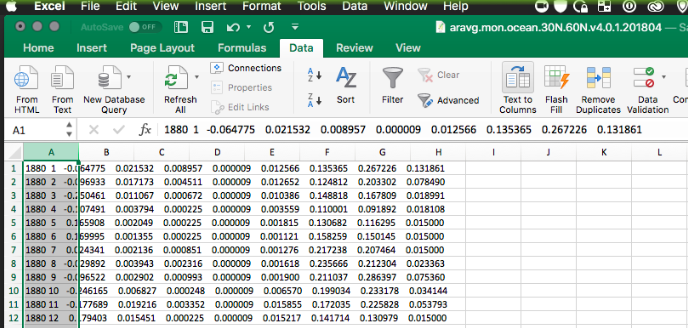
? Could you add a 2nd File as another data series in a different color to serve as a comparison?

? How do we add the month level detail into the X domain?

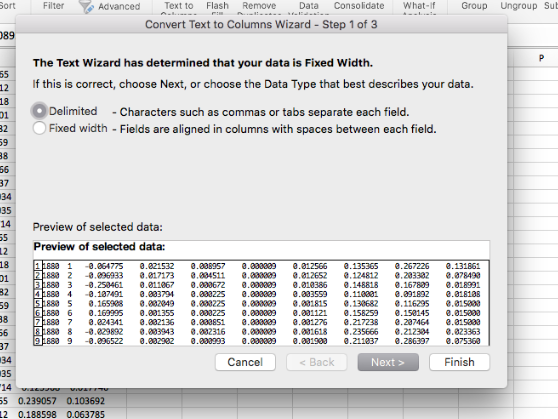
? What other visualization types help you understand the shape and the trend of this dataset?

# 5. Offline Data Visualization

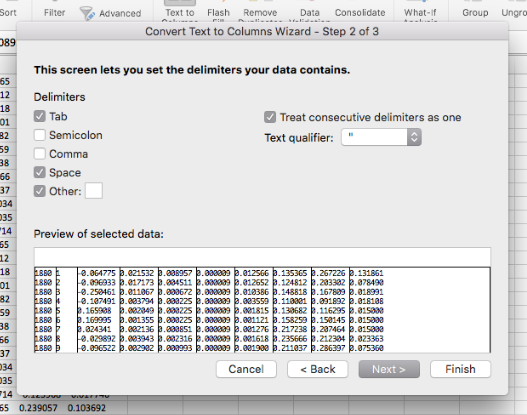
Using Microsoft Excel, the File opens as a single column per row:



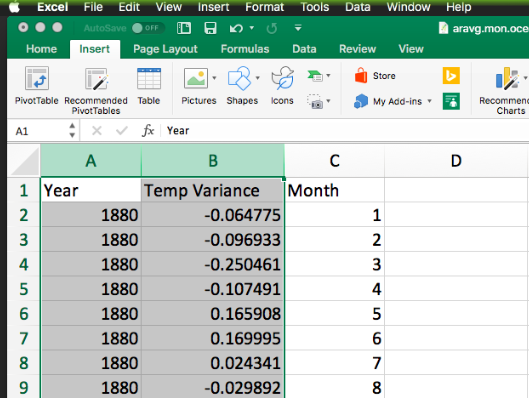
Convert Text To Columns Wizard



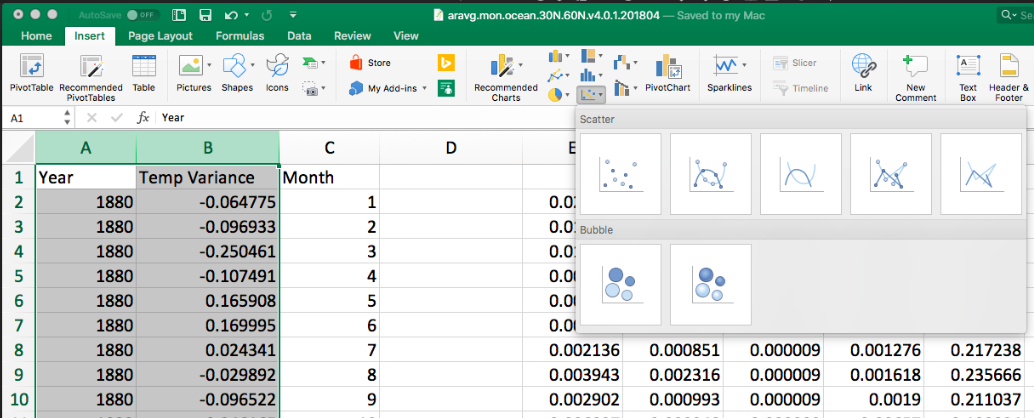
Delimiter Space:

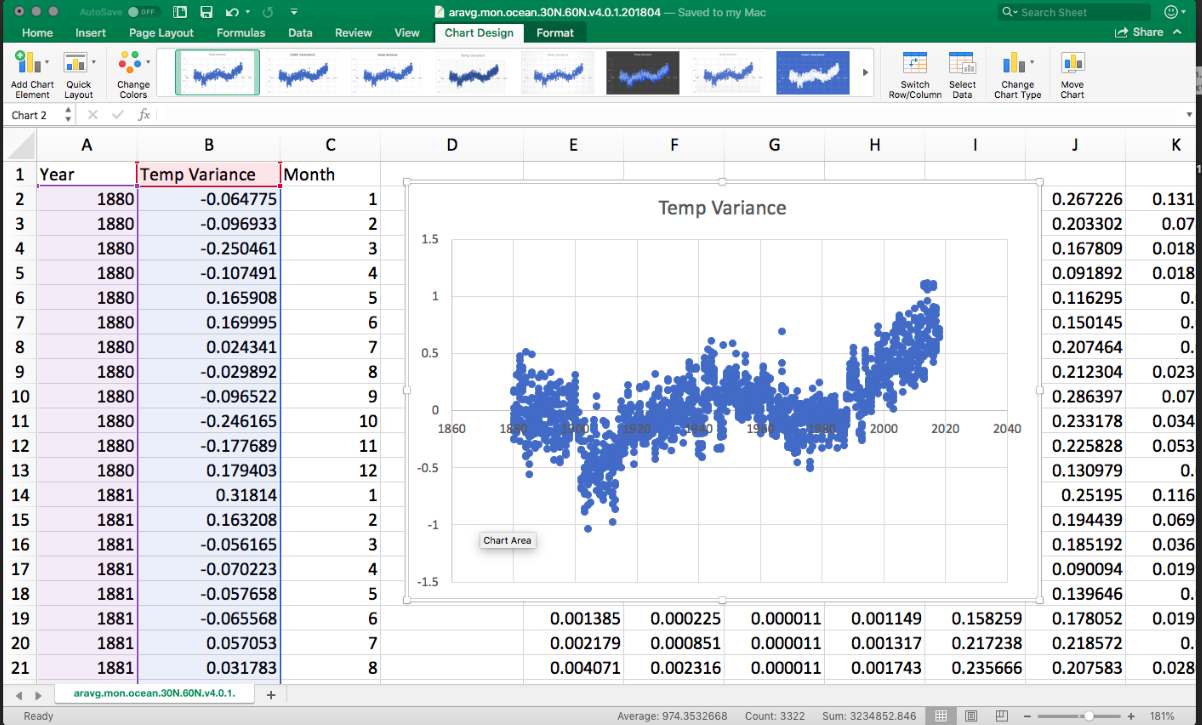


Add Column Headers To Label Data



Create new Scatter Plot





# 6. Communicating Your Conclusions

? Who is your audience?

? Will your conclusion inform a specific decision? Do you understand the implications of that decision?

? Will you be presenting the conclusions using an in person presentation, conference call, screen share, email or printed paper?

? Can you predict the challenging questions your audience may ask about your conclusions or analysis techniques?

? What are you list of data sources to list as references?