

# Astro Data Science Seminar

## Lab 5 - 2016 May 17

### Data visualization and making "good" plots

In today's installment we have only one goal: make great looking plot that is publication ready! Making "good" plots is 90% thinking about the story or point you're trying to tell, and 10% graphic design or art. The former takes years of work to learn (and, incidentally, is why companies like hiring scientists). Some people never grok the latter, but we can at least teach you about the tools needed to design visually pleasing figures.

Thankfully Python and the default style of Matplotlib make functional plots that you can read and share (typically). However, publishing and sharing results in science means your figures need to work in many different environments. These include:

- printed in color and black and white
- printed on a big paper poster
- in a PDF on a computer or tablet
- projected in powerpoint
- on a blog/website/tweet/etc...

All of these mediums have specific subtleties that must be appreciated, both in the artistic presentation and the content. For example, when advertising your work on a blog post it probably is better to make your figures more "flashy", maybe interactive, using lots of color, and removing (or hiding) any data or trends that don't convey the story you're telling. Similar rules apply to putting figures in a slideshow, except you must be careful about color and contrast with projectors that wash everything out.

Here are some of my guiding principles when making data visualizations:

1. **A pretty plot won't hide a poorly thought out idea**
  - but it might help to promote your really good idea!
2. It doesn't matter what software or package you use, as long as the result is good.
  - All programs can make good looking plots! Say NO to tool snobbery!
3. Contrast is key - people have to be able to read/discern what you're showing
4. Color should be used with careful intention. Avoid common colorblindness traps
5. If the viewer is trying to decode what your artistic choices are, they aren't thinking about your science
6. Treat every aspect of a plot with care, like a painter to a canvass.

In the example code today I will give you a bunch of hints and samples. You are going to have to Google or search the Python & Matplotlib documentation for specifics. You also will have to just guess and check some things to compare them.

**Goal: Recreate (as closely as you can) the plot I have provided!**

Today's data comes from the very famous HIPPARCOS mission