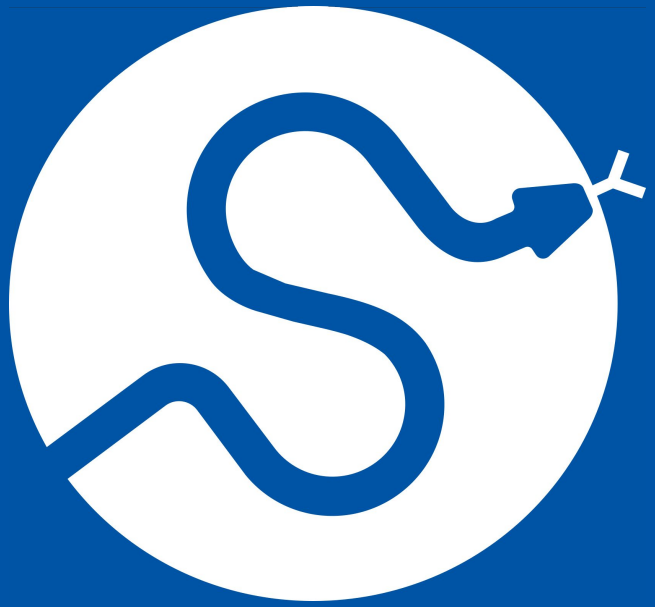
A Solid Foundation for Statistics with SciPy

Warren Weckesser
Berkeley Institute for Data Science

Matt Haberland
Assistant Professor, Cal Poly



SciPy – The Library

- Core numerical library in the Python scientific computing ecosystem.
- Fundamental building blocks for modeling and solving scientific problems.
- SciPy includes algorithms for:
  - linear algebra
  - sparse matrices
  - optimization
  - numerical integration
  - differential equation solvers
  - interpolation
  - **statistics**
  - and more.

Just published in *Nature Methods:*

"SciPy 1.0: fundamental algorithms for scientific computing in Python"
https://www.nature.com/articles/s41592-019-0686-2

CZI project: *improve the statistics library.*

A Solid Foundation for Statistics in SciPy

What are your goals and expectations for the coming year?

- Improve the statistical library in SciPy.
  - Increase the coverage of fundamental statistical tools in SciPy, so users have a more complete library.
  - Tackle some maintenance tasks that have been languishing for too long because of lack of developer time and interest.
- Improve diversity of SciPy contributors.
  - Engage students in open source development by hosting a coding sprint at Cal Poly in collaboration with clubs that support underrepresented groups.

SciPy CZI Project

A Solid Foundation for Statistics in SciPy

What do you hope to achieve by learning from or collaborating with other grantees?

- We would love to learn more about SciPy users.
  - What do you use?
  - What do you need?
- What else can we do to improve diversity of contributors?

What expertise do you have to share with other grantees?

- Warren Weckesser and Matt Haberland are long time SciPy developers, with experience developing the statistics, signal processing, optimization and linear algebra libraries.