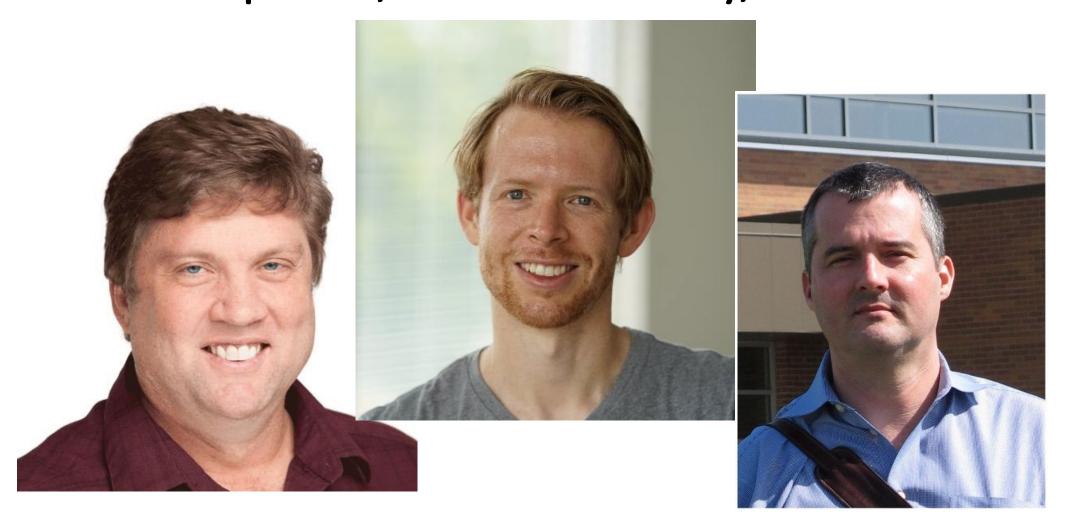# The Triumvirate of Data Science

ACE 592

Last time…

- We got the first taste of Python as a language.

- But we did everything in **pure Python**, and did not import any packages.

- Here we will go over three very important packages most data people use almost constantly:

The Python Boys:
Travis Oliphant, Wes McKinney, John Hunter

for Arrays and Matrices:

- Created by: **Travis Oliphant,** programmer and engineer by training.*

- The bedrock of most quantitative operations in Python.
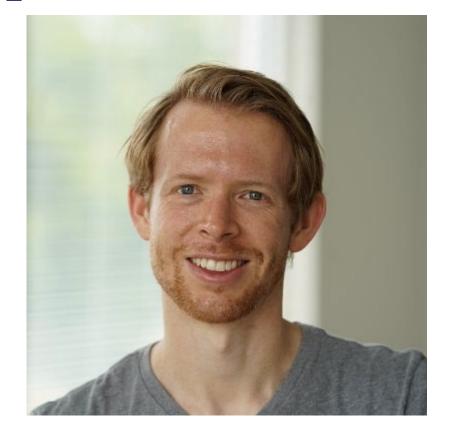
- The Key Object: **The Array**

* Also founded **Anaconda** and the package **scipy**

for Organizing Data:

- Created by: **Wes McKinney,** programmer and statistician by training.

- Name is actually derived from "Panel Data."

- The Key Object: **The DataFrame**

for Visualizing the Data:

- Created by: **John D. Hunter,** neurobiologist by training.

- The most important plotting library in all of Python.

- The Key Object: **The Plot**

Honorable Mention:

- Created by: **Travis Oliphant**

- All things optimization and statistics.

- The basic foundation for most of the statistical analysis in Python.

- The Key Object: **The Optimization Function**

Honorable Mention: statsmodels

- Created by: **Skipper Seabold**, economist by training.

- Closest package for standard econometric models.

- Still in development for some of the more complex models.

What are they used for?

	NumPy	pandas	matplotlib
Important objects	Arrays, matrices	DataFrames, Series	Plots
Important functions	Linear algebra, mathematic operations.	Data reading, data cleaning, data processing.	Various plot types.
Equivalent Software	Matlab	STATA, R	ggplot