

June 4th, 2018

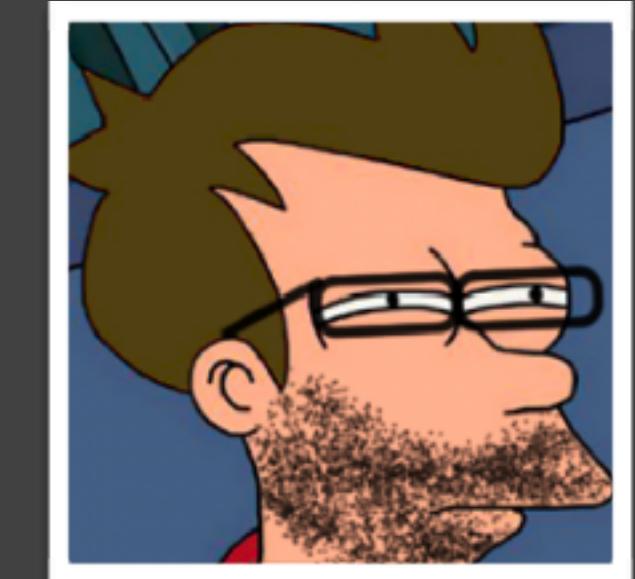
**Matthias Bussonnier**

bussonniermatthias@gmail.com

GitHub: @carreau

Twitter: @mbussonn

# Jupyter & Education



# About Me



# Matthias Bussonnier

- A Physicist/Bio-Physicist
- Core contributor of IPython/Jupyter since 2012
  - Co-founder, and Steering Council member
- Post doctoral Scholar on Jupyter at BIDS

# A short history



- 2001: Fernando Pérez Wrote “**IPython**”
- 2012: Birth of IPython **Notebook** (6th prototype)
  - Make IPython “network enabled”
  - Made possible by mature web tech.
- 2013: First non-Python (**Julia**) kernel
- 2014: We **renamed** the Python-Agnostic part to **Jupyter**.
- 2018:
  - Several millions users
  - Jupyter increasingly used in Education.



# Contributors



- 500+ Open source contributors
- Organisation with Open Governance



# Sponsors



ALFRED P. SLOAN  
FOUNDATION

NUMFOCUS  
OPEN CODE = BETTER SCIENCE

THE LEONA M. AND HARRY B.  
HELMSLEY  
CHARITABLE TRUST

GORDON AND BETTY  
MOORE  
FOUNDATION



SIMONS FOUNDATION

U.S. DEPARTMENT OF  
ENERGY

Google



CONTINUUM  
ANALYTICS

IBM

ENTHOUGHT  
SCIENTIFIC COMPUTING SOLUTIONS

Bloomberg

POWERED BY  
rackspace®  
the open cloud company

Microsoft

# How Jupyter came to be



# Life cycle of a Scientific Idea

- Individual exploratory work
- Collaborative development
- Parallel production runs (HPC, cloud, ...)
- Publication & communication (reproducibly!)
- Education
- Goto 1



**“The purpose of computing is insight,  
not numbers”**

*-Hamming'62*

# Life cycle of a Scientific Idea

- Individual exploratory work (Repl, Scripts)
- Collaborative development (Dropbox/ Google Doc / emails / git )
- Parallel production runs (MPI, rewrite C++, batch jobs)
- Publication & communication (Word, Latex, ppt...)
- Education
- Goto 1



# Rise of Jupyter

- An increasing number of discipline have a fast growing amount of data
- Technology is **a tool** that should
  - Empower **the User**
  - Amplify **Domain Knowledge and Expertise**
  - Facilitate **Sharing and Collaboration**

Jupyter provide a framework that can be use in all the step in the cycle of a scientific idea



# What is Jupyter

Individual,Collaboration,Parallel,Publication,Education

- Mainly Known for **The Notebook**

- Web server, a web app, containing code, narrative,

math and results.

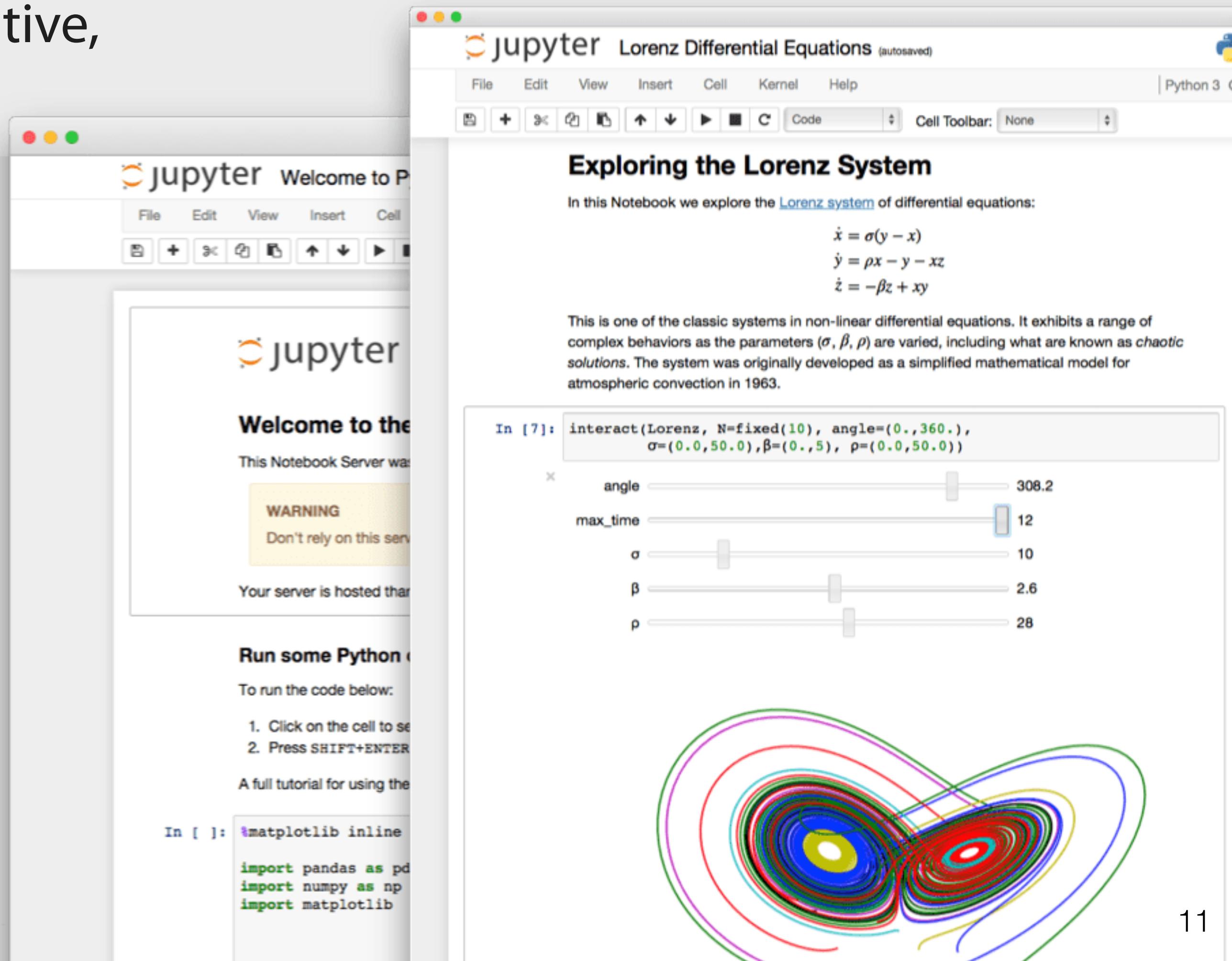
- Attached to a **Kernel** doing computation.

- Results can be:

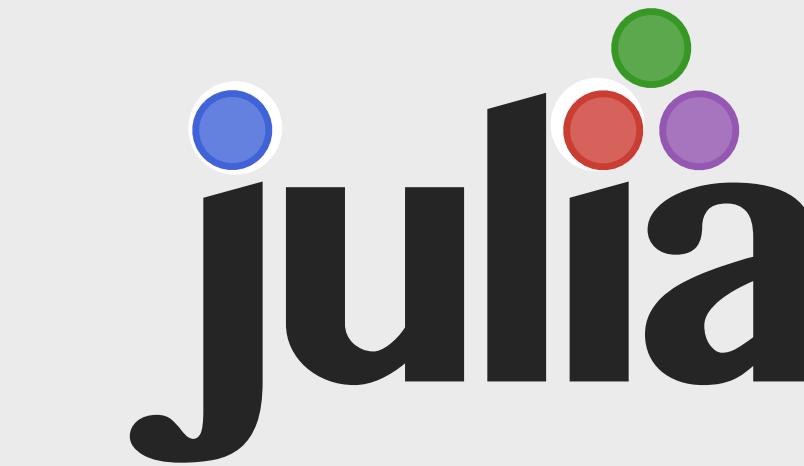
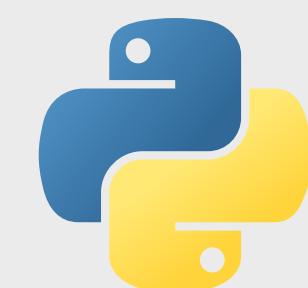
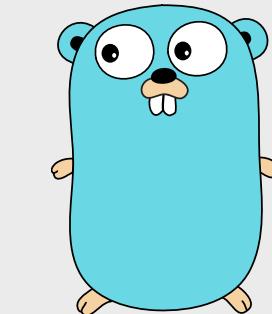
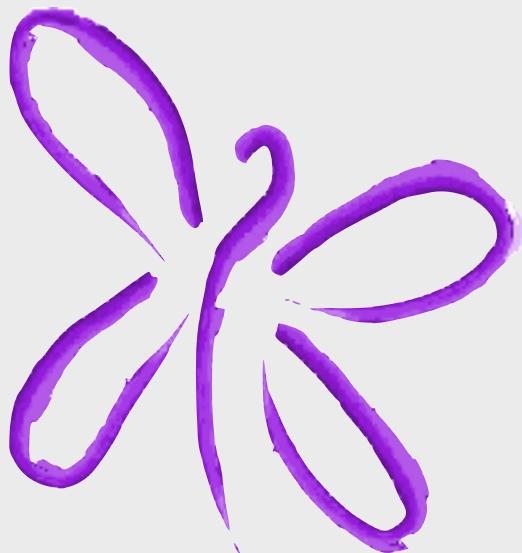
- Static (Image)

- Interactive (client-side scroll/pan/brush)

- Dynamic (Call back into Kernel)



# Many languages



...

An article about computational science in a scientific publication is **not** the scholarship itself, it is merely **advertising** of the scholarship. The **actual scholarship** is the complete software development environment and the complete set of instructions which generated the figures.

*Buckheit and Donoho, WaveLab and Reproducible Research, 1995*

MENU ▾ nature microbiology

Altmetric: 202 Views: 823 More detail >

Letter

# Dog and human inflammatory bowel disease rely on overlapping yet distinct dysbiosis networks

Yoshiki Vázquez-Baeza, Embriette R. Hyde, Jan S. Suchodolski & Rob Knight ✉

Nature Microbiology 1, Article number: 16177 (2016) doi:10.1038/nmicrobiol.2016.177 Download Citation

Dysbiosis Inflammatory bowel disease

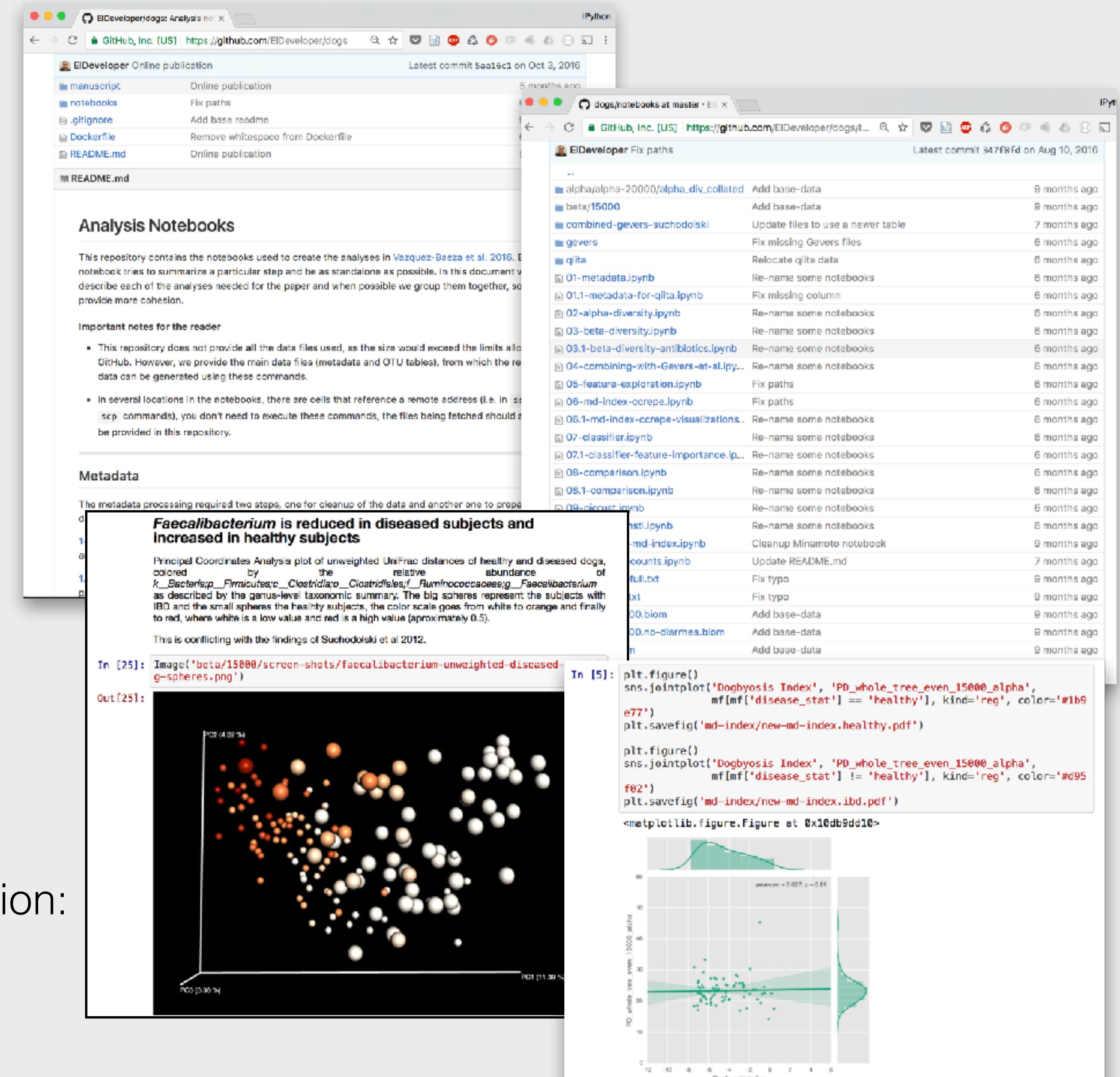
Microbiome Molecular medicine



Advertising

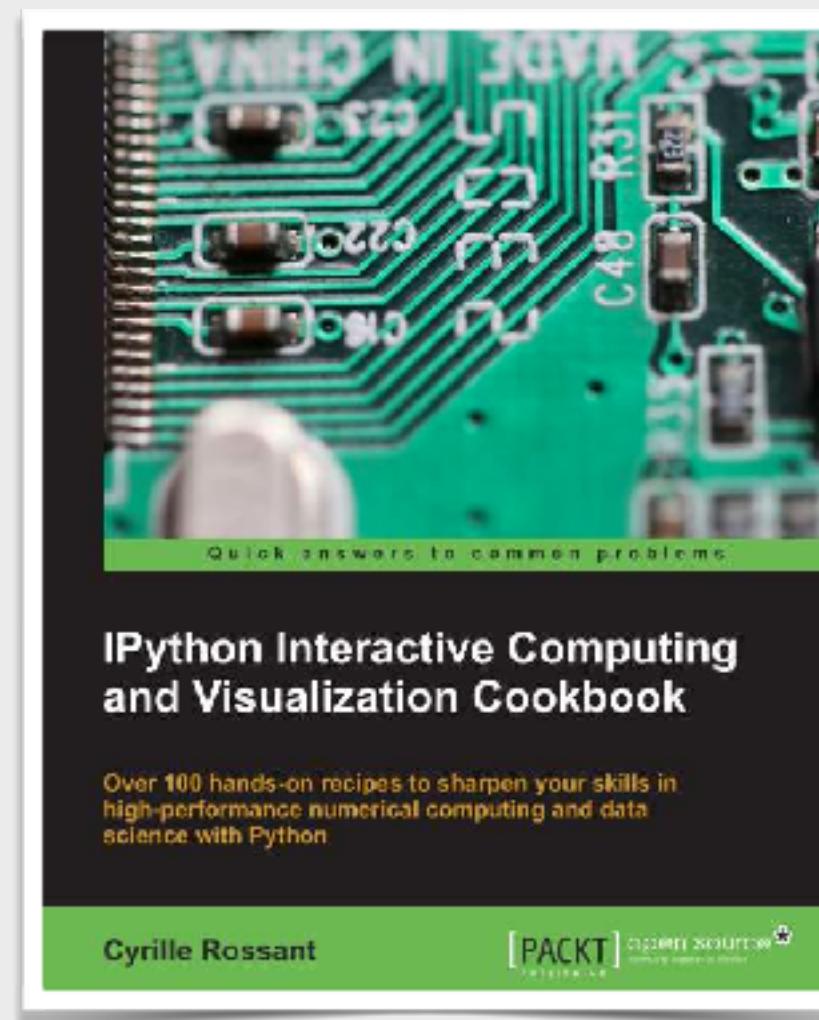
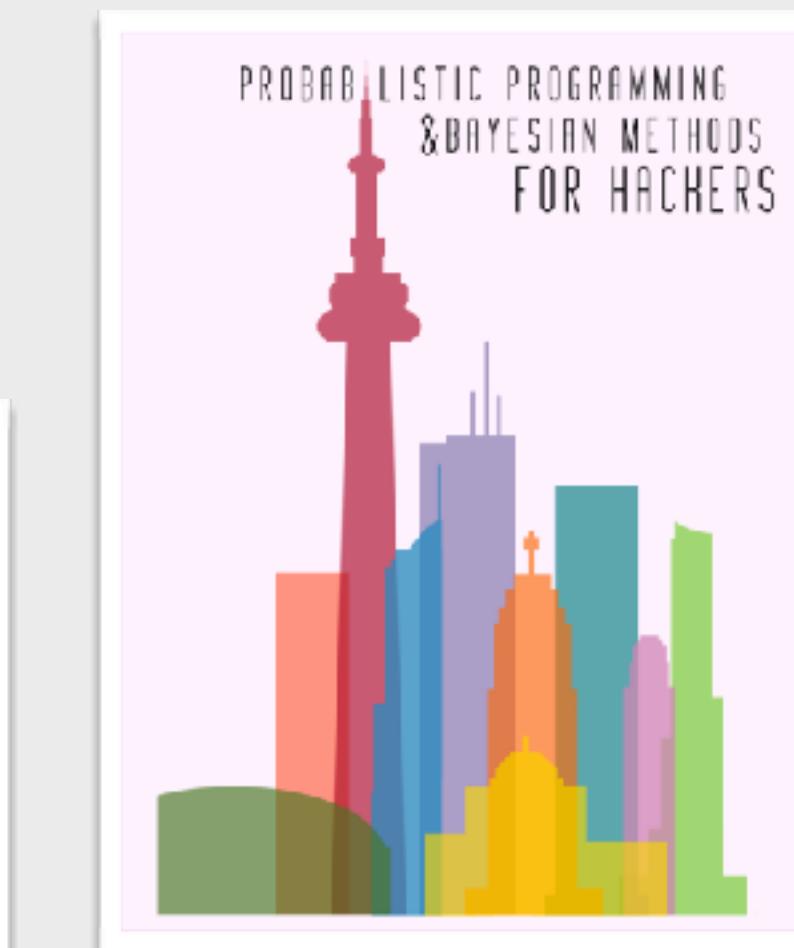
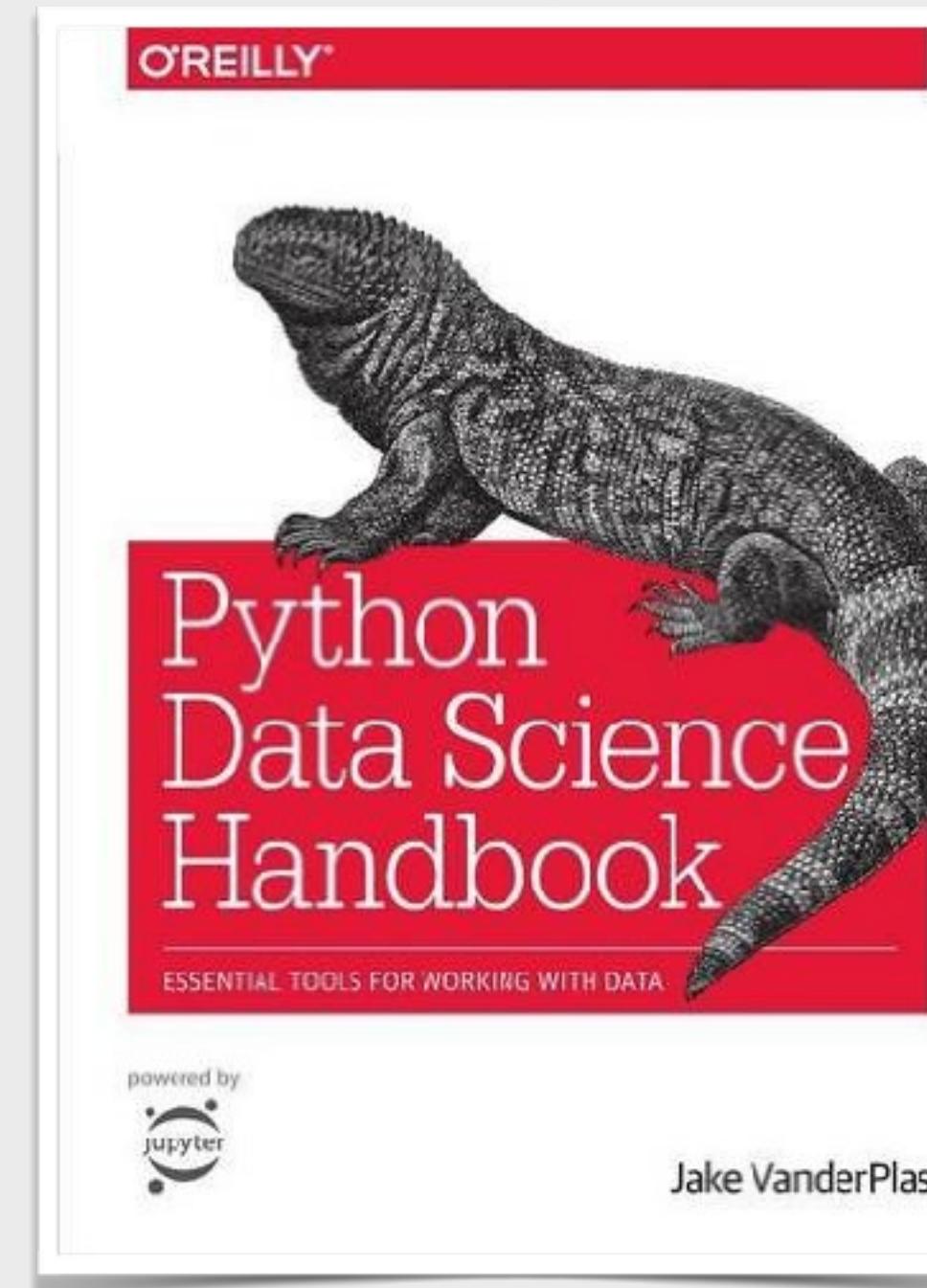
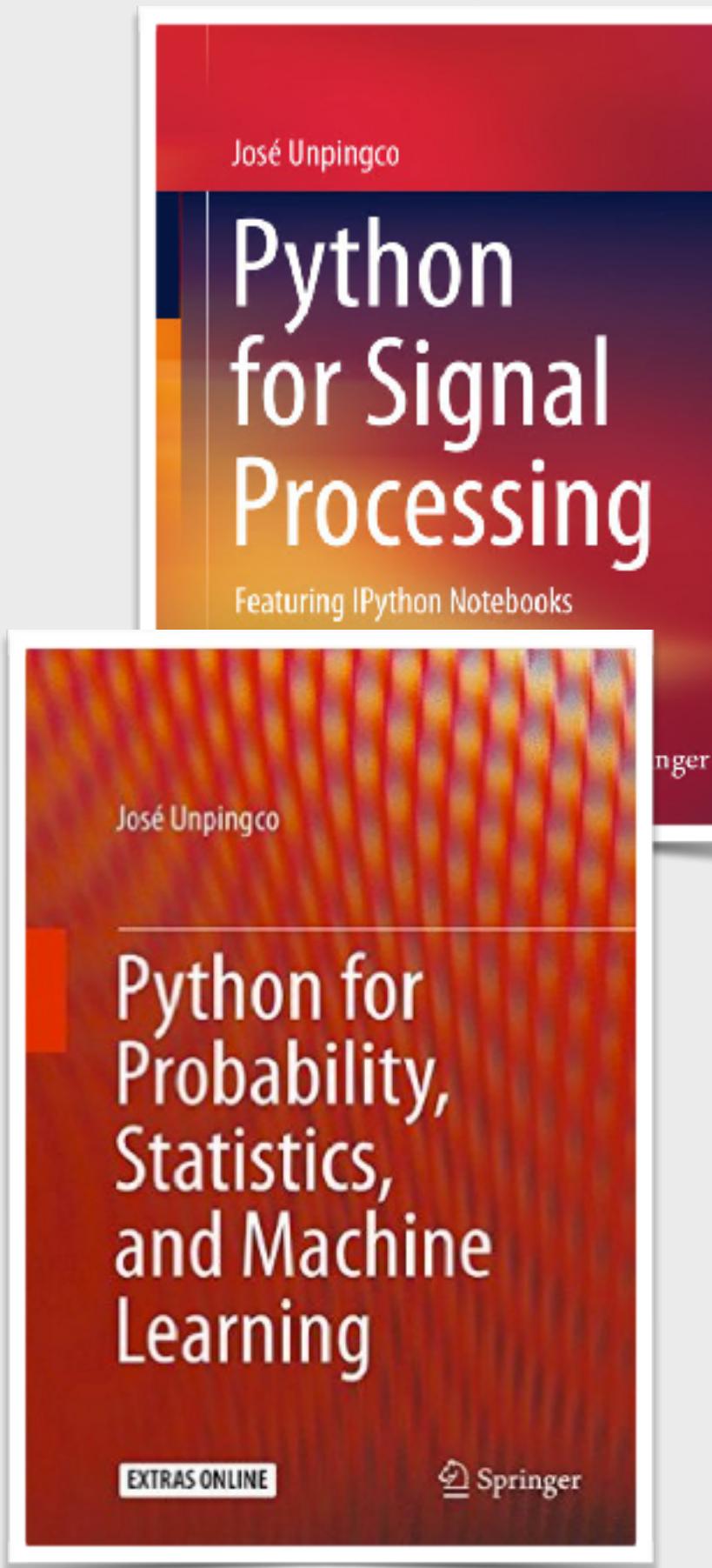
Full code and specific discussion:  
executable supplementary  
materials:

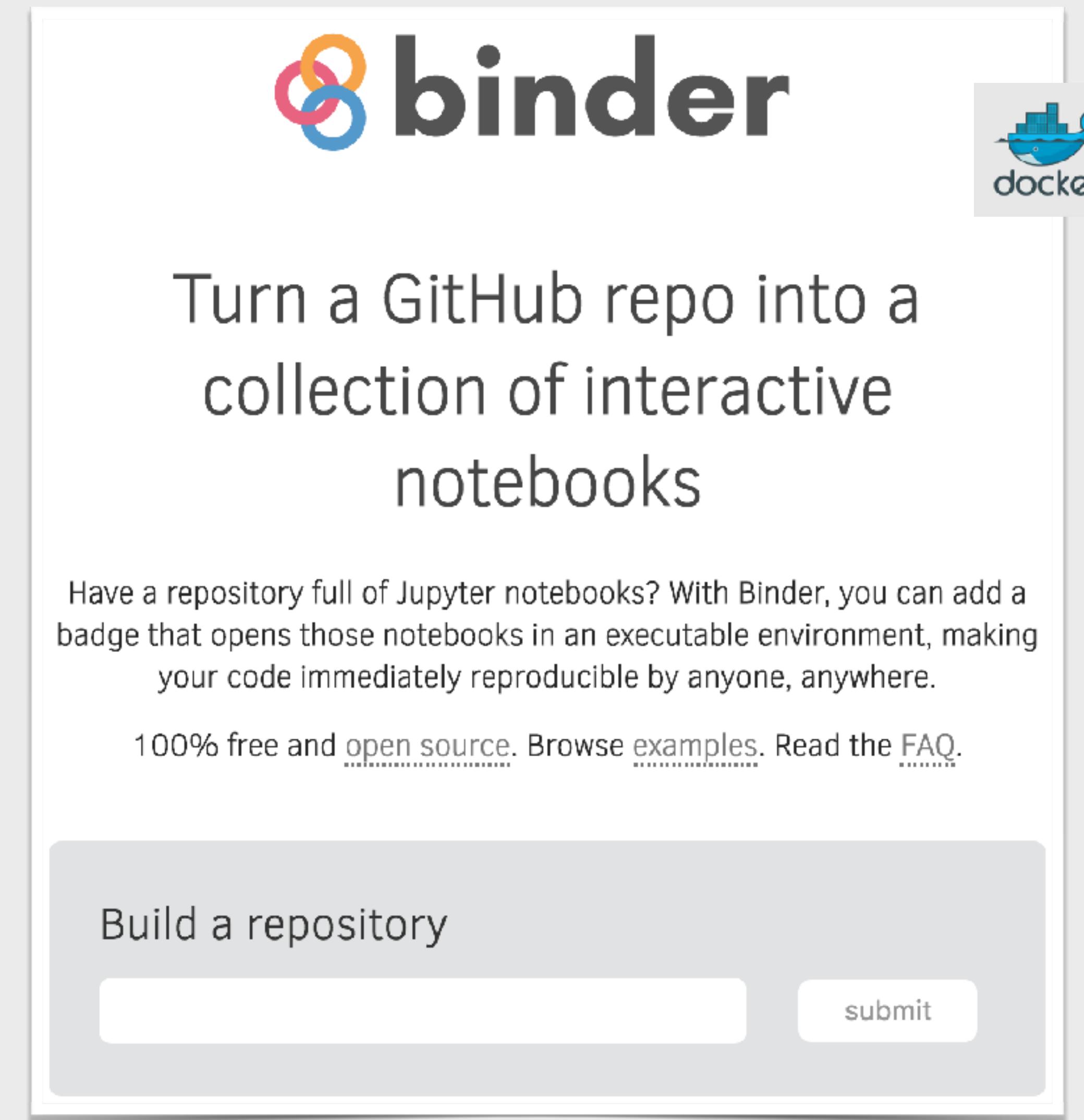
“computational companions”



<https://github.com/ELDeveloper/dogs>

# Executable books





The image shows the MyBinder.org landing page. At the top left is the MyBinder logo, which consists of three overlapping circles (orange, red, and blue) followed by the word "binder" in a bold, lowercase sans-serif font. To the right of the logo is a small square icon containing a blue Docker logo (a whale inside a container). Below the logo, the text reads: "Turn a GitHub repo into a collection of interactive notebooks". Underneath this, there is a paragraph: "Have a repository full of Jupyter notebooks? With Binder, you can add a badge that opens those notebooks in an executable environment, making your code immediately reproducible by anyone, anywhere." Below the paragraph, it says "100% free and [open source](#). Browse [examples](#). Read the [FAQ](#)." At the bottom of the page is a large button with the text "Build a repository" and a "submit" button.

binder

Turn a GitHub repo into a collection of interactive notebooks

Have a repository full of Jupyter notebooks? With Binder, you can add a badge that opens those notebooks in an executable environment, making your code immediately reproducible by anyone, anywhere.

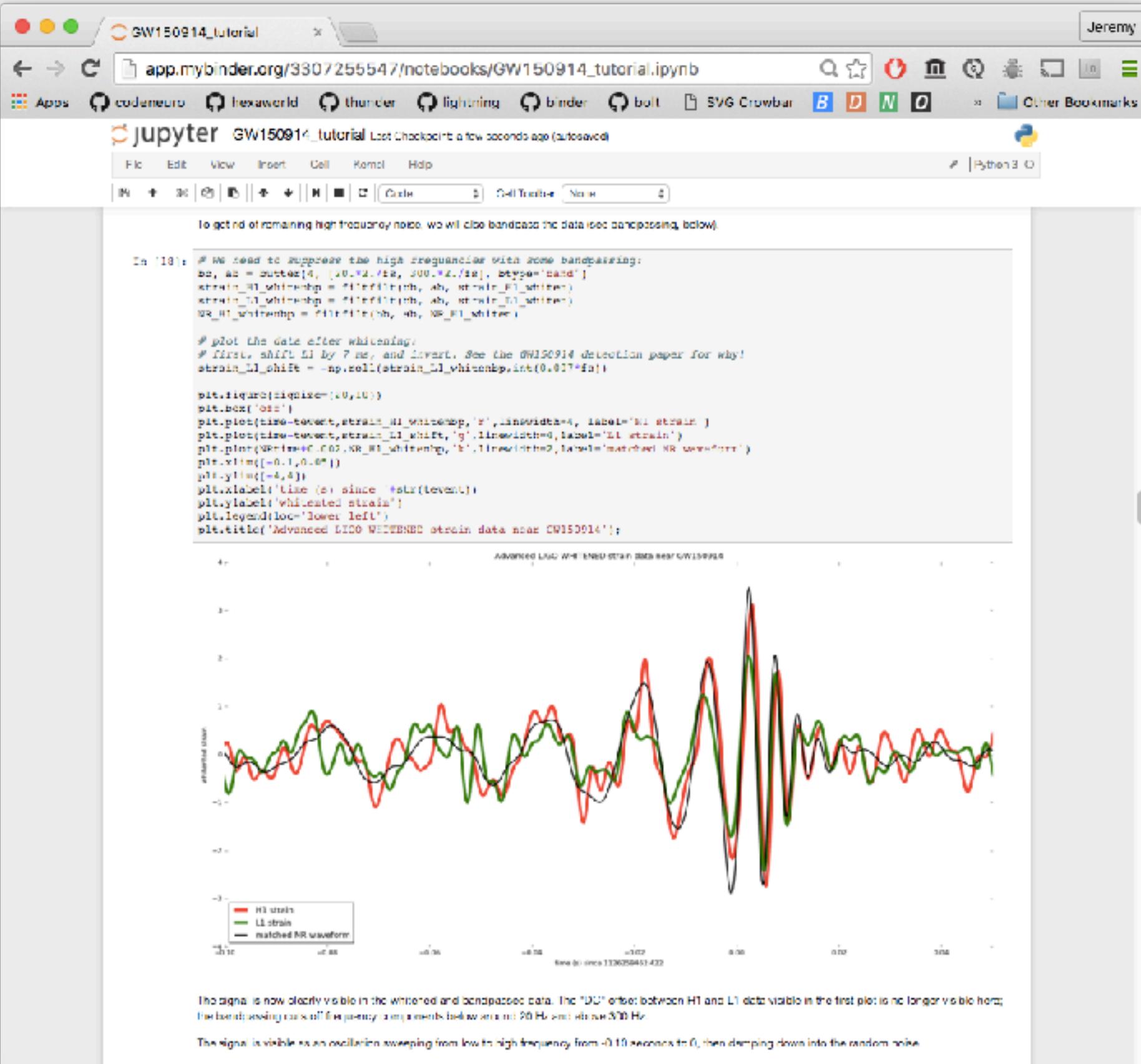
100% free and [open source](#). Browse [examples](#). Read the [FAQ](#).

Build a repository

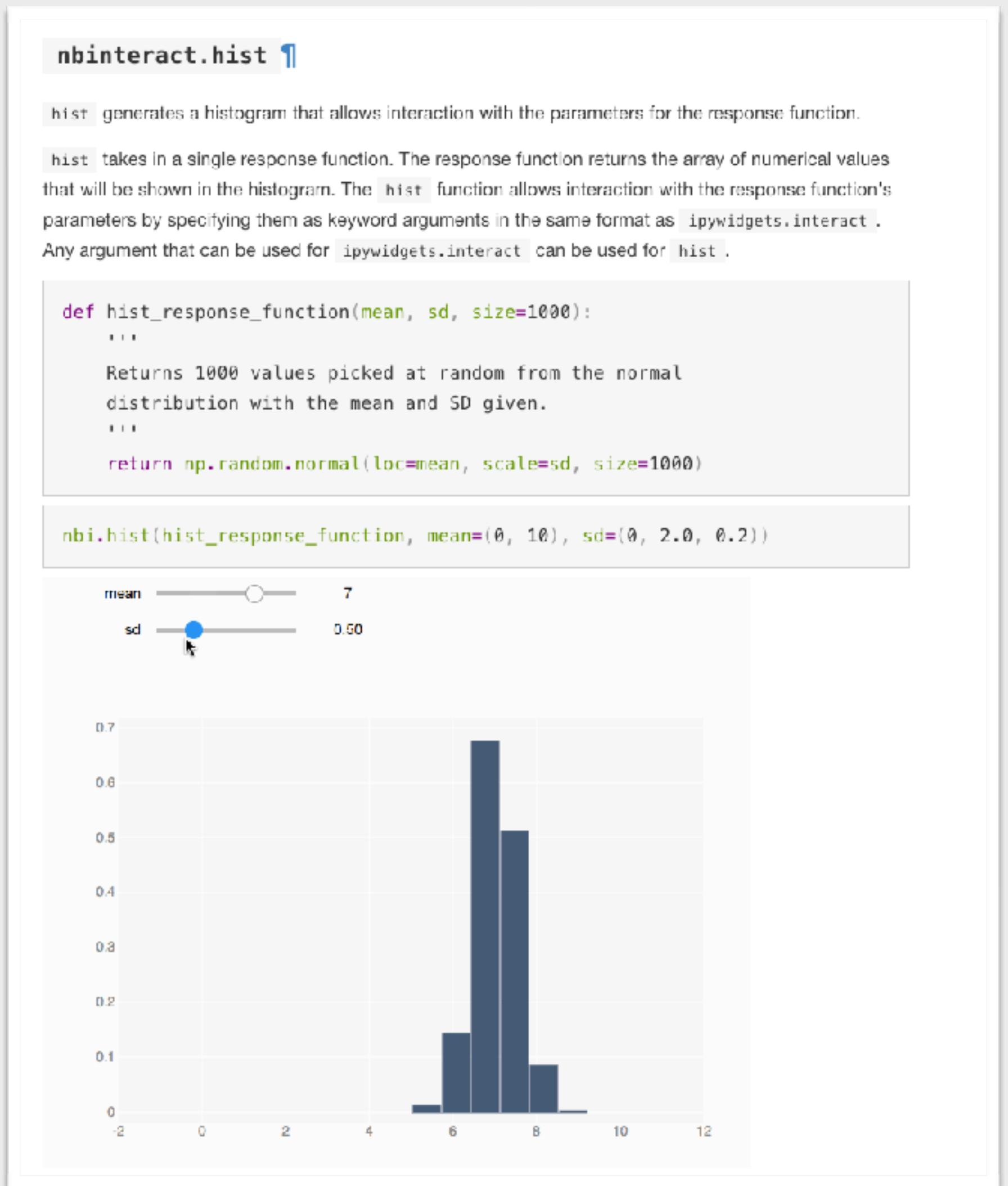
submit



# Ligo binder



# NbInteract



# In the classroom





# DataHub

[datahub.berkeley.edu](http://datahub.berkeley.edu)

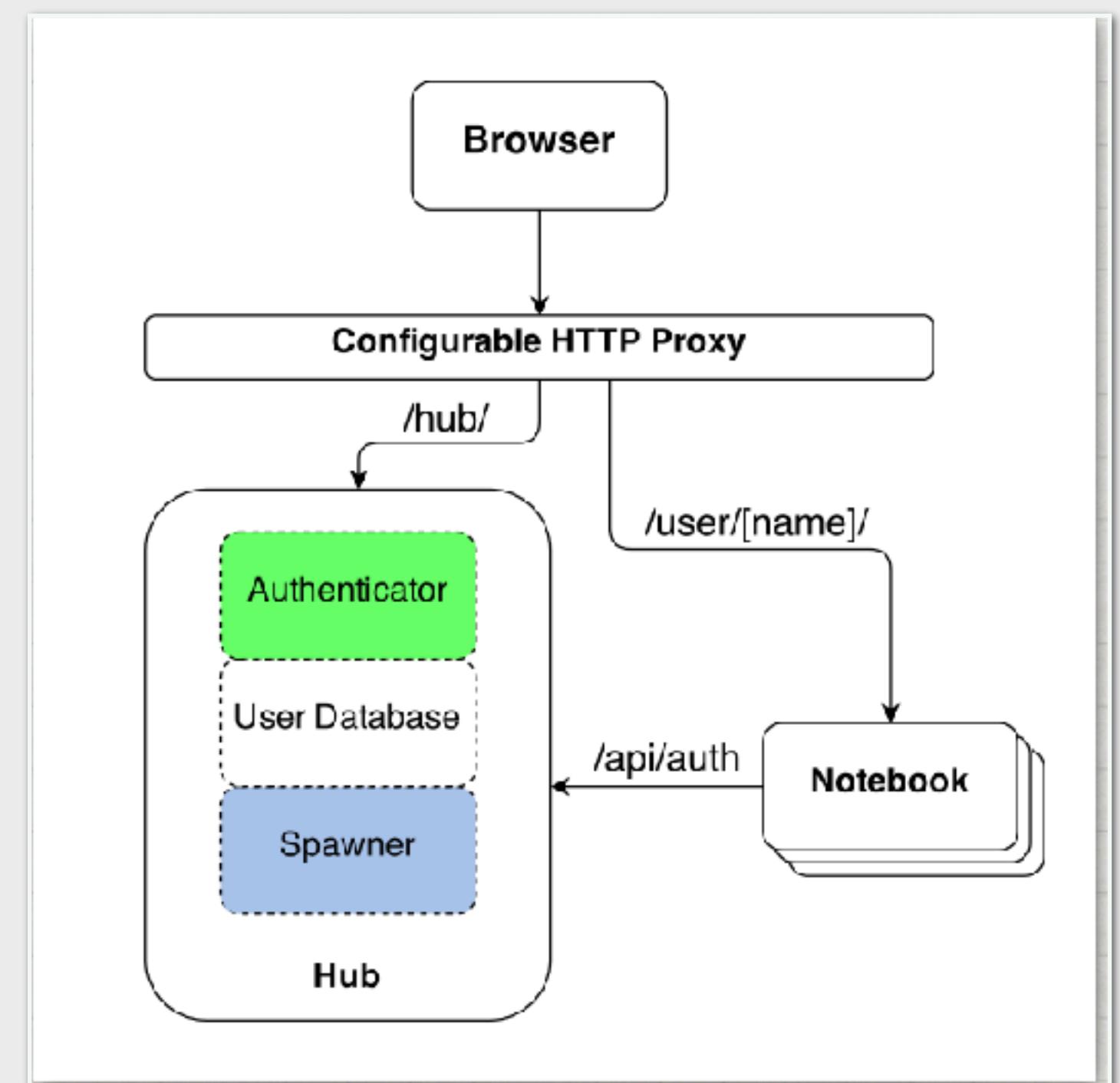


# Zero setup\*



- Campus Wide deployment
- Login with Cal ID
- Can focus on Domain Knowledge

Students can still optionally install  
Jupyter on their machine later on.



\* At least for students



# Autograding

- Via Nbgrader
- Generate “Students” version of the notebook
- Partial auto grading
  - Frees up a lo of TA time.



J. Hamrick  
Deep Mind



B. Granger  
Cal Poly



Fernando Pérez  
UC Berkeley

jupyter Problem 1 Last Checkpoint: a few seconds ago (autosaved) | Python 3

File Edit View Insert Cell Kernel Help

Cell Toolbar: None

**Part A (2 points)**

Write code to compute the mean of a list of numbers.

```
In [ ]: def mean(x):
    """Compute the mean of a list of numbers given in `x`."""
    ### BEGIN SOLUTION
    return sum(x) / len(x)
    ### END SOLUTION
```

```
In [ ]: """Check that the `mean` function is correct."""
assert mean([1]) == 1.0
assert mean([1, 2]) == 1.5
assert mean([5.5, 0, 2, 3.4]) == 2.725
assert mean(range(100)) == 49.5
assert mean(range(100, 0, -1)) == 50.5
```

**Part B (3 points)**

Describe the difference between an *arithmetic mean*, a *harmonic mean*, and a *geometric mean*.

Arithmetic mean:

$$\frac{1}{N} \sum x_i$$


# MOOCS



Ani Adhikari



John DeNero

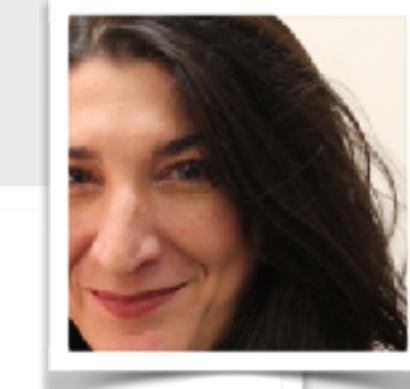


David Wagner

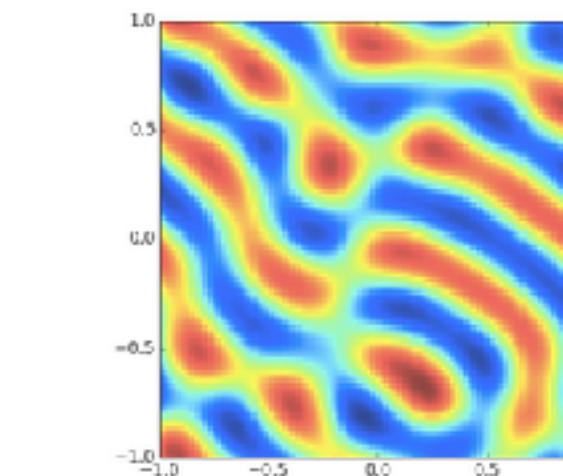
Foundation of Data Science  
(edX/ Berkeley)



## Practical Numerical Methods



Lorena Barba



Pattern formation:  
solution for a reaction-diffusion system like:

$$u_t = \delta D_1 \nabla^2 u + f(u, v)$$
$$v_t = \delta D_2 \nabla^2 v + g(u, v)$$

An example of the types of problems we will learn to solve in this course, among others governed by differential equations.

# Oriole: executable, video-narrated tutorials

The screenshot shows the homepage of Oriole Online Tutorials. At the top, there's a navigation bar with links for Ideas, Learning, Events, Shop, and YOUR ACCOUNT. Below the navigation is a banner featuring a bird (likely an oriole) perched on a branch. The main title "Oriole Online Tutorials" is displayed, along with the tagline "Learn alongside smart people solving hard problems". A text block explains that Oriole is a unique medium that blends code, data, text, and video. It highlights that lessons are narrated by brilliant minds like Peter Norvig and Fernando Perez. Below this, there's a section about the convenience of learning side-by-side with the instructor. At the bottom, three headshots are shown: Paco Nathan, Taylor Martin, and Andrew Odewahn.

Oriole is a unique new medium that blends code, data, text, and video into a narrated learning experience with executable content.

Led by some of the most brilliant minds in technology, each lesson is an easily digestible and engaging thought-by-thought tour of the instructor's approach to the problem in both narrative and executable code. No set-up or installation is necessary; Oriole Online Tutorials require nothing more than an internet connection and a laptop. You can write and run code within the environment. Make a mistake? Change it, and try again.

Oriole combines the expert insight and hands-on learning of in-person or online courses with the on-demand, at-your-own desk, back-up-and-run-it-again convenience of video training. You learn by doing, on your own schedule, and at your own pace.

Paco Nathan

Taylor Martin

Andrew Odewahn

The screenshot shows a specific tutorial titled "REGEX GOLF". The page includes a sidebar with a cartoon illustration of a person working at a desk. The main content area features a video player showing a man (Peter Norvig) speaking. Below the video, there's a text box with a snippet of Python code related to regex golf. A code editor window is also visible, showing the same code and some output results.

MEYA-REGEX GOLF:  
SO I WROTE A PROGRAM THAT PLAYS REGEX GOLF WITH ARBITRARY LISTS...

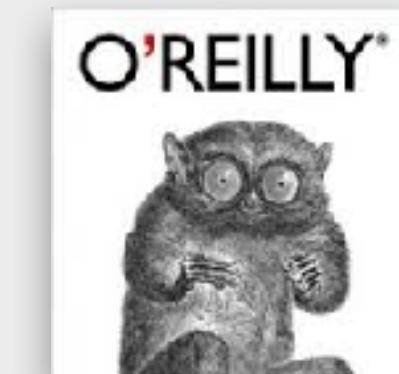
We have now fulfilled panel two of the strip.  
Let's try another example, separating the top ten best-selling drugs from the top 10 cities to visit:

```
1 drugs = words('tipitox nexium plavix advair albuterol se')
2 cities = words('paris trinidad capeTown riga zurich s')
3 report(drugs, cities)
```

Characters: 16, Parts: 6, Competitive ratio: 5.3,  
Kimmers: 10, Losers: 10

'e.\$|x|1e|q|b|en'

Run Again



[oreilly.com/learning/regex-golf-with-peter-norvig](https://oreilly.com/learning/regex-golf-with-peter-norvig)

# The Future ?



# JupyterLab

- Extends the notebook interface with text

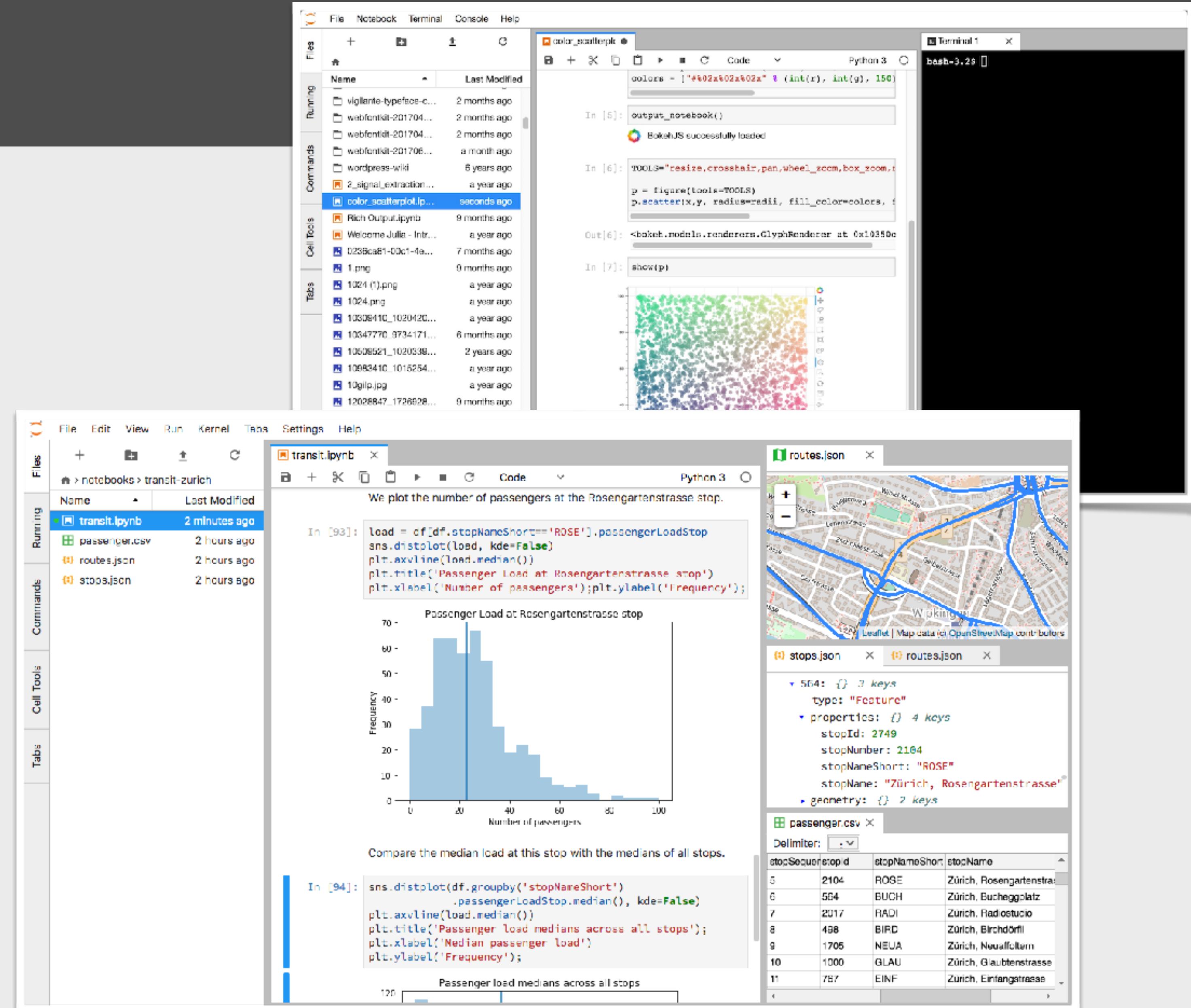
editor, shell, ...etc

- Plugins are first class citizen

- Custom Layouts (Education-oriented layout ?)

- Realtime collaboration

- Broadcast the teacher's notebook ?





jupyter-education@googlegroups.com



# Questions ?