Open Science Droplets 01
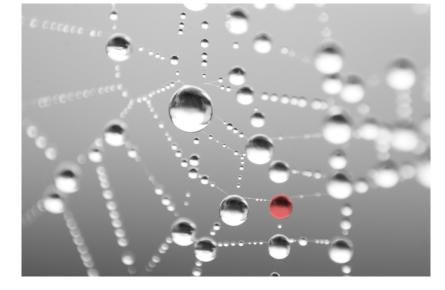
Jupyter notebook, the XXI century lab book

Javier Moldón

2020-03-02

Hi!

The other day my friend Ptolemy told me about a "nebulous mass in the breast of Cancer"[1]. I observed it with my new telescope and I could resolve about 40 individual stars! This is amazing, and it could be a good topic for a paper. Please, do a first exploratory analysis of the region and see what you can find.

R.A. 130.025 deg
Dec. 19.9833 deg

Best regards,

Galileo

[1] Claudius Ptolemy, Almagest, AD 100-170 (see V/61)

Objectives

- Jupyter notebooks as a dynamic tool for exploratory analysis
- Initialize a notebook
- Basic structure and syntax: cells

Additional information

- droplets resources
- fortran magic
- Server options

Summary (I)

- pros
  - Felixibility for exploratory data, training, sharing
  - Web app, accessible from anywhere (ssh, server)
  - Markdown + code + resuts. Program code is more "tangible"
  - Dashboards (also interactive), reports in different formats
  - Many extensions, and growing!

Summary (II)

- cons
  - Hidden state and out-of-order execution
  - Notebooks encourage bad habits (not ideal for software development)
  - In general, not as powerful as a stand-alone application or modules (not ideal for sharing good code)
  - Some difficulties to obtain diff

Other resources

- Jake VanderPlas youtube series on Reproducible data analysis with jupyter Youtube
- Try Jupyter in your browser link
- Quickview Notebook sharing the Gravitational Wave detection Notebook
- A Machine Learning course using Notebooks: Lecture 1: Density Est, Lecture 3: Classification and Lecture 4: Dimensionality Reduction.
- The full tutorial on an international Python conference: PyCon 2015 Scikit-learn Tutorial

Take away

- Easy to learn tool
- Interweave results, ideas, and hypotheses with the code
- Natural format to create a scientific narrative
- State of scripts is not linear, depends on user
- Excellent tools to share your research

Next session

- March 16
- Collaborative Jupyter notebooks through GitHub