Interactive computing

CS594: Big Data Visualization & Analytics

Fabio Miranda

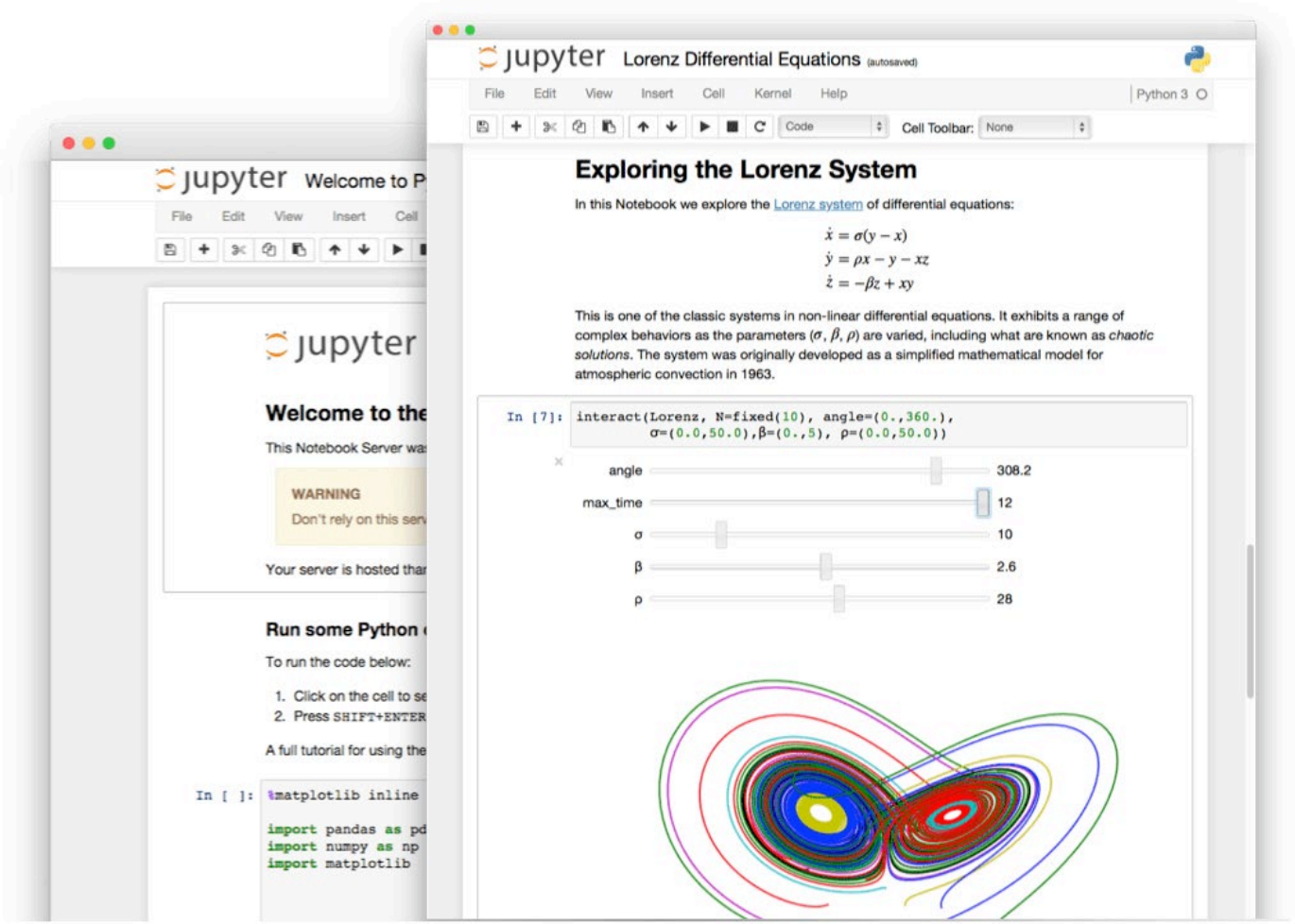
https://fmiranda.me

Interactive computing

- Software that accepts commands and output results right away.
  - Spreadsheet applications
  - Word processors
  - Jupyter Notebooks
  - Google Collab
  - Observable
  - …

Jupyter

- Interactive computing environment that mixes:
  - Code
  - Results of running code
  - Documentation



Data analysis with Jupyter

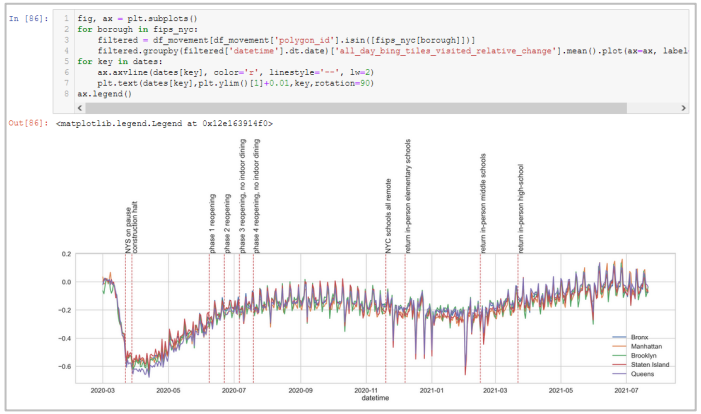
Data analysis with Jupyter

- Main use cases:
  - Development: write the different steps of an algorithm.
  - Document the development and thinking process that led to code or solution.
  - Initial data exploration: test hypotheses, find potentially interesting patterns.
    - Data is too large, we need to have a good understanding of the most important features, attributes, time slices, etc.

Data analysis with Jupyter

- How are you going to load the data?
- What operations are you going to perform?
- What libraries will you need to use?
  - Pandas
  - GeoPandas
  - Numpy
  - Keras
  - OpenCV
  - NLTK
  - Matplotlib
  - Scikit-learn
  - tdqm
  - …

Live examples



Spatial data exploration

OD data exploration

Image data exploration