

# Nuts and Bolts of Jupyter

January 6, 2019  
Thomas Pinder



- Background of Jupyter.
- Why use Jupyter?
- Demonstration of Jupyter's capabilities.
- Exploratory analysis of PM<sub>2.5</sub> data in Jupyter.
- Limitations

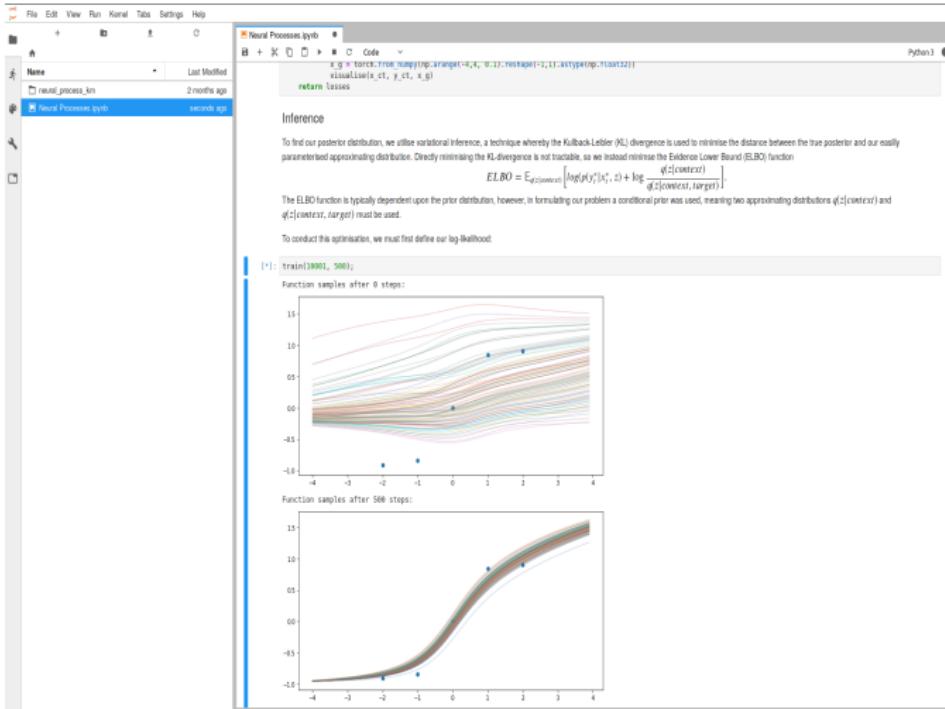
# Background

# Julia, Python, R $\implies$ Jupyter

Jupyter notebooks are comprised of cells that can contain either code or text.

This allows for extensive code commentary and inline figures.

# Readable Code



The screenshot shows a Jupyter notebook interface with the following details:

- Title Bar:** Neural Processes.ipynb
- File Menu:** File, Edit, View, Run, Kernel, Tabs, Settings, Help
- Code Cell:** Contains Python code:

```
x, g = torch.meshgrid(torch.arange(-4, 4, 0.2), torch.arange(2, 11, 1.1), indexing='ij')
visualise(x, y, ct, x_g)
```
- Text Cell:** Inference
- Text Content:** Describes variational inference and the Evidence Lower Bound (ELBO) function.
- Equation:** 
$$ELBO = \mathbb{E}_{q(z|context)} [\log p(y|z) + \log \frac{q(z|context)}{q(z|context, target)}]$$
- Text Content:** Notes that the ELBO function is typically dependent upon the prior distribution.
- Text Content:** Describes the conditional prior used for two approximating distributions.
- Text Content:** Notes that to conduct the optimisation, we must first define our log-likelihood.
- Text Cell:**

```
[*]: train[1000, 500];
```
- Text Content:** Function samples after 0 steps:
- Figure:** A scatter plot with multiple colored lines representing function samples. The x-axis ranges from -4 to 4, and the y-axis ranges from -1.0 to 1.5. Blue dots are scattered along the lines.
- Text Content:** Function samples after 500 steps:
- Figure:** A scatter plot with smooth, colored curves representing function samples. The x-axis ranges from -4 to 4, and the y-axis ranges from -1.0 to 1.5. Blue dots are clustered at the bottom left.

Figure: Example screengrab of a Jupyter notebook

# colab

**Figure:** Google incorporated Jupyter into their online web services platform.



**Figure:** Summer 2018 saw Google allow a free K80 to be bolted onto an notebook.



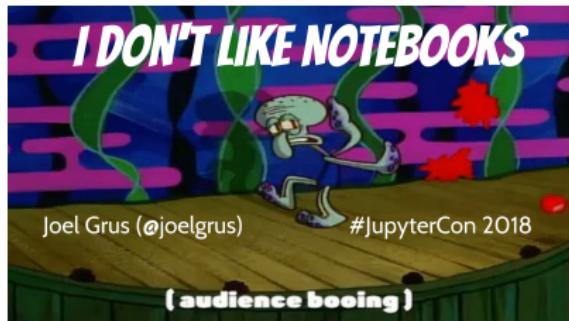
**Figure:** Subset of supported languages in Jupyter.

# Demo

## Launch Notebook

# Limitations

# Not Everyone Loves Jupyter



**Figure:** Author of Data Science from Scratch and developer Joel Grus gave a talk at JupyterCon 2018 of his dislike for Jupyter...

- Notebooks are *dangerous* unless each cell is run sequentially.
- Notebooks discourage good software development habits (e.g. modularity).
- Notebooks are not reproducible.
- Hard for beginners.

# Final Remarks

Recent development have allowed Jupyter notebooks to easily parallelised<sup>1</sup> through magic commands in .

Additionally, a Jupyter kernel is now available for Spark/PySpark<sup>2</sup>; an alternative way to accomplish parallel computing.

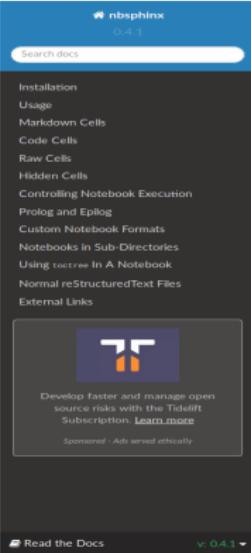


---

<sup>1</sup><https://github.com/ipython/ipyparallel>

<sup>2</sup><https://spark.apache.org/docs/0.9.0/python-programming-guide.html>

For those using Sphinx for documentation, there is now a Jupyter extension<sup>3</sup>.



The screenshot shows the `nbsphinx` documentation page on ReadTheDocs. The sidebar on the left contains links for Installation, Usage, Markdown Cells, Code Cells, Raw Cells, Hidden Cells, Controlling Notebook Execution, Prolog and Epilog, Custom Notebook Formats, Notebooks in Sub-Directories, Using `toctree` In A Notebook, Normal reStructuredText Files, and External Links. A sponsored advertisement for Tideflip Subscriptions is present, featuring a logo of two orange squares and the text: "Develop faster and manage open source risks with the Tideflip Subscription. [Learn more](#)". The main content area has a header "Docs > Jupyter Notebook Tools for Sphinx" and a "Edit on GitHub" button. Below the header, the title "Jupyter Notebook Tools for Sphinx" is displayed. A paragraph explains that `nbsphinx` is a Sphinx extension for parsing Jupyter Notebook files. It includes sections for "Quick Start", "Online documentation (and example of use)", "Source code repository (and issue tracker)", and "License". A note at the bottom states that most content was generated from Jupyter notebooks. The footer of the page shows "Read the Docs" and "v: 0.4.1".

**Jupyter Notebook Tools for Sphinx**

`nbsphinx` is a Sphinx extension that provides a source parser for `*.ipynb` files. Custom Sphinx directives are used to show Jupyter Notebook code cells (and of course their results) in both HTML and LaTeX output. Un-evaluated notebooks – i.e. notebooks without stored output cells – will be automatically executed during the Sphinx build process.

**Quick Start:**

1. Install `nbsphinx`
2. Edit your `conf.py` and add `'nbsphinx'` to `extensions`.
3. Edit your `index.rst` and add the names of your `*.ipynb` files to the `toctree`.
4. Run Sphinx!

**Online documentation (and example of use):**  
<http://nbsphinx.readthedocs.io/>

**Source code repository (and issue tracker):**  
<https://github.com/spatialaudio/nbsphinx>

**License:**  
MIT – see the file `LICENSE` for details.

All content shown below – except for the section `Normal reStructuredText Files` – was generated from Jupyter notebooks.

- Installation
  - `nbsphinx` Packages
  - `nbsphinx` Prerequisites
    - Python
    - Sphinx
    - pip
    - pandoc
    - Pygments Lexer for Syntax Highlighting
    - Juvopter Kernel

<sup>3</sup><https://nbsphinx.readthedocs.io/en/0.4.1/>

Thank you for listening

# Backup Slides