

Hello World

July 10, 2016

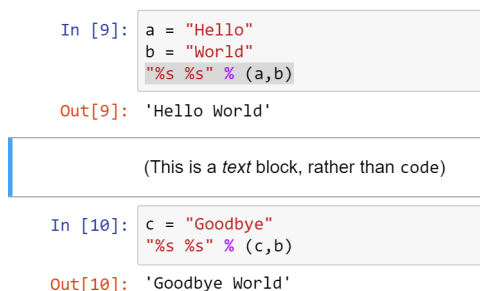
0.1 Download Lesson Files

The lesson files for this workshop can be downloaded at <https://github.com/htrc/ef-workshop/releases>. Download them and extract to an easy to access folder.

0.2 Start a Notebook

Using Python the traditional way – writing a script to a file and running it – can become clunky for text analysis, where the ability to look at and interact with data is invaluable. This lesson uses an alternative approach: Jupyter notebooks.

Jupyter gives you an interactive version of Python (called IPython) that you can access in a “notebook” format in your web browser. This format has many benefits. The interactivity means that you don’t need to re-run an entire script each time: you can run or re-run blocks of code as you go along, without losing your environment (i.e. the variables and code that are already loaded). The notebook format makes it easier to examine bits of information as you go along, and allows for text blocks to intersperse a narrative.



```
In [9]: a = "Hello"
        b = "World"
        "%s %s" % (a,b)

Out[9]: 'Hello World'
```

(This is a *text* block, rather than code)

```
In [10]: c = "Goodbye"
         "%s %s" % (c,b)

Out[10]: 'Goodbye World'
```

Figure 1: Jupyter Code Blocks

From the Start Menu (Windows) or Applications directory (Mac OS), open “Jupyter notebook”. This will start Jupyter on your computer and open a browser window. Keep the console window in the background, the browser is where the magic happens.

If your web browser does not open automatically, Jupyter can be accessed by going to the address “localhost:8888” - or a different port number, which is noted in the console (“The Jupyter Notebook is running at...”):

Jupyter is now showing a directory structure from your home folder. Navigate to the lesson folder where you unzipped **lesson-files.zip**.

In the lesson folder, open **Start Here.pynb**: your first notebook!

Here there are instructions for editing a cell of text or code, and running it. Try editing and running a cell, and notice that it only affects itself. Here are a few tips for using the notebook as the lesson continues:

- New cells are created with the Plus button in the toolbar. When not editing, this can be done by pressing ‘b’ on your keyboard.

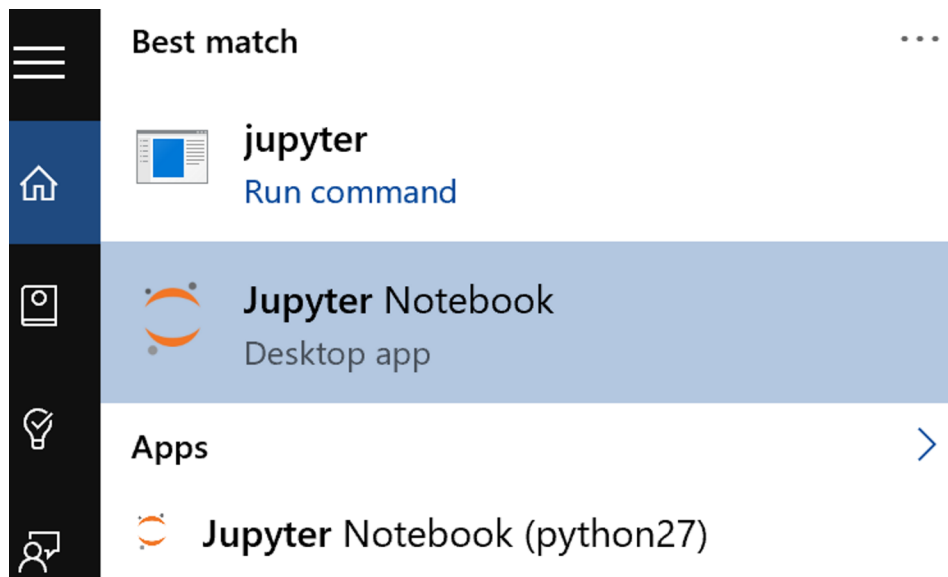


Figure 2: Jupyter in the Start Menu (Windows)

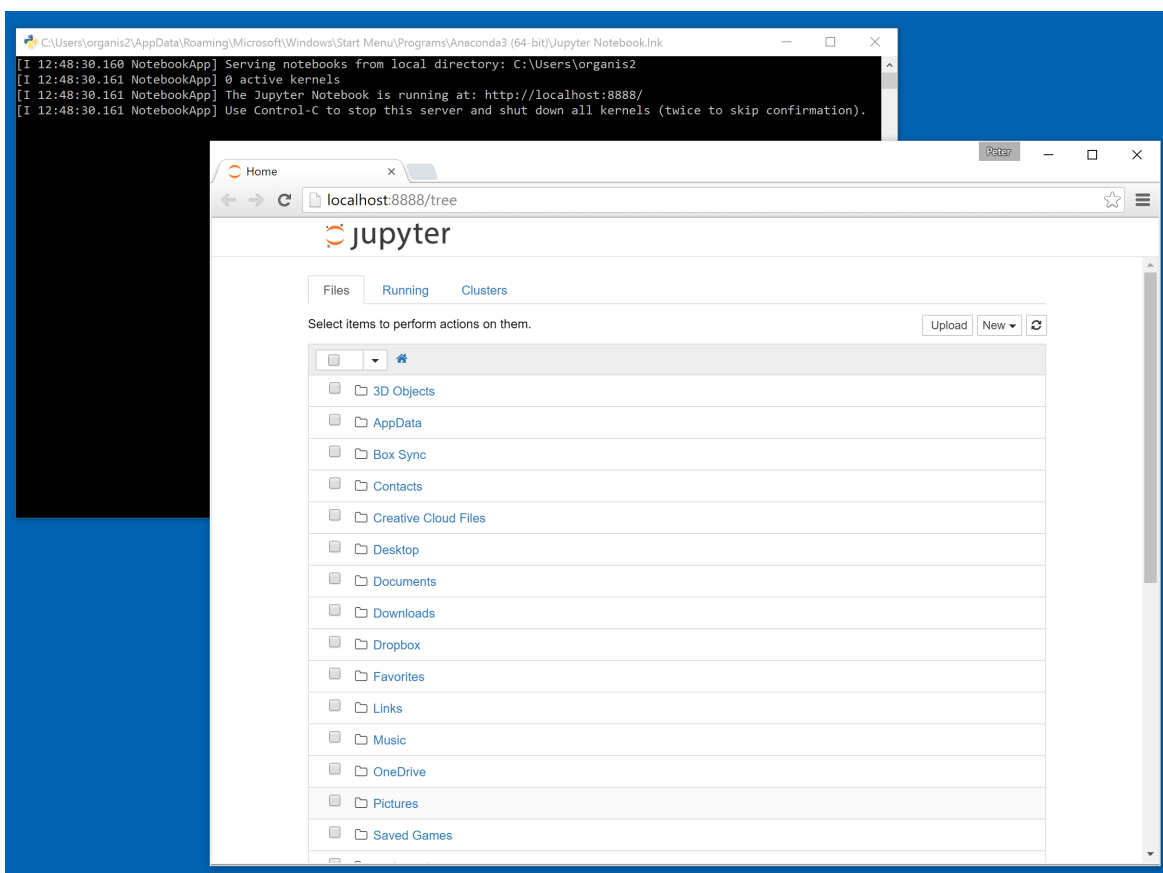


Figure 3: Jupyter in the browser after being started

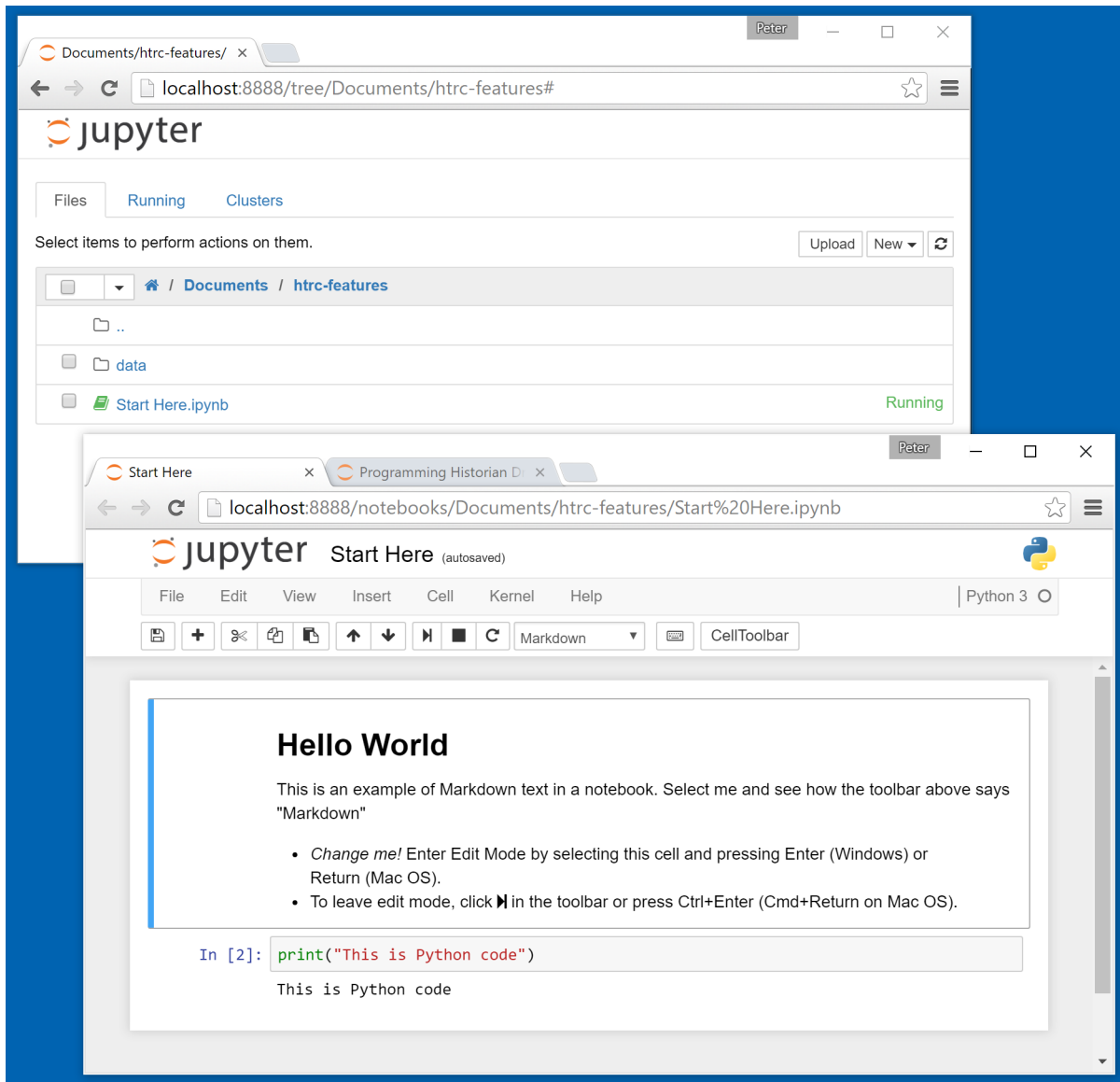


Figure 4: Opening the Start Here notebook

- New cells are “code” cells by default, but can be changed to “Markdown” (a type of text input) in a dropdown menu on the toolbar. In edit mode, you can paste in code from this lesson or type it yourself.
- Switching a cell to edit mode is done by pressing Enter.
- Running a cell is done by clicking Play in the toolbar, or with **Ctrl+Enter** (Cmd+Return on Mac OS). To run a cell and immediately move forward, use **Shift+Enter** instead.

An example of a full-fledged notebook is included with the lesson files in `example/Lesson Draft.ipynb`.

Before continuing, click on the title to change it to something more descriptive than “Start Here”.

0.3 Hello World

To test that the Feature Reader installed correctly and to try running code in a notebook, we’ll try running some code.

1. Create a new cell in your notebook with the plus button.
2. Make sure the description of the cell in the toolbar dropdown above is ‘Code’.
3. Enter the code below into the cell. We’ll explain shortly what it does, for now we want to make sure it works. (Tip: If you don’t want to type out the file path to `data/sample-file1.basic.json.bz2`, you can press TAB on your keyboard to auto-complete once you’ve started typing.)
4. Run the code by pressing **Shift+Enter** or pushing the play button in the toolbar. we are expecting it to print the title of a book called June.

```
In [1]: from htrc_features import FeatureReader
        # Remember to use '/' for windows paths, '\' for Mac or Windows
        fr = FeatureReader('data/sample-file1.basic.json.bz2')
        fr.first().title
```

```
Out[1]: 'June / by Edith Barnard Delano ; with illustrations.'
```