

## Week 2: Getting Started with Jupyter and Google Collaboratory

### ▼ Student Name: Muhammad Rakibul Islam

- Date:
- Instructor: Lisa Rhody
- Assignment due:
- Methods of Text Analysis
- MA in DH at The Graduate Center, CUNY

Click in the following cell and then click on the "play" icon to run it.

```
1 + 5 * 2 - 3
```

8

If the code is running correctly, a green check mark will appear to the left of the `1`. Inside the `1` a number will appear. This number represents the "process number"—which is to say the order in which the code you executed was executed within the notebook. Also, you will see the result of the math equation below the cell where the equation is. The answer should be 8.

### ▼ Adding a Markdown / Text cell

At the top of your notebook and just under the menu (the list of words that begin File, Edit, View, etc...), there should be two actions you can perform. One is "+ Code" and the other is "+ Text." These will allow you to add either a code cell or a text cell in the area beneath whichever cell is currently active.

For your next trick, click on this cell to make it "active." Then move your cursor to the top of the page and click the "+Text" action. A new cell should open up. Click inside the new cell and type an answer to one of the following questions:

- What is your favorite comfort food?
- What is your favorite book?
- Do you have a pet? If so, what kind? and what is its name?
- Ice cream.
- The Idiot.
- Yes! A cat named Snorlax.

### ▼ Importing NLTK and text functions

Next, click in the cell below. This cell will tell the notebook to pull into its working memory a library called NLTK (the Natural Language Toolkit) which is something we will use quite a lot this semester. It will also download the book corpora.

```
import nltk
nltk.download("book")
```

Show hidden output

If this cell has executed properly, it will have lots of rows that begin `[nltk_data]` and at the end say "True." We're not going to use much of NLTK this week, but it's helpful to practice because we'll need to do this a lot in the future. Please create a new text cell below.

**EXERCISE:** If you had any problems, please explain. If not, just say "All good."

### ▼ Response:

All good.

Finally, we're just going to play around with what we installed a little bit to make sure it's working for you. No worries if you have trouble with it... that's why we're doing this. Just to practice.

**EXERCISE Part 1:** First, run the cell below. This is a sentence taken from the following article: Fabricant, Florence. "Porcini Mushrooms for Those With Fungi on the Brain." The New York Times, 30 Jan. 2023. NYTimes.com, <https://www.nytimes.com/2023/01/30/dining/mushrooms-porcini-italian.html>.

The output should include 2 rows. The first row will cut the text up into units called "tokens." The second row will tag each word with a label for the part of speech.

```
sentence = ""Midway upon the journey of our life I found myself within a forest dark,For the straightforward pathway had been los

tokens = nltk.word_tokenize(sentence)

print(tokens)

tagged = nltk.pos_tag(tokens)

print(tagged)
```

```
[('Midway', 'NNP'), ('upon', 'IN'), ('the', 'DT'), ('journey', 'NN'), ('of', 'IN'), ('our', 'PRP$'), ('life', 'NN'), ('I', 'I'), ('found', 'VBD'), ('myself', 'PRP'), ('within', 'IN'), ('a', 'DT'), ('forest', 'NN'), ('dark', 'JJ'), (',', ','), ('For', 'IN'), ('the', 'DT'), ('straightforward', 'JJ'), ('pathway', 'NN'), ('had', 'VBD'), ('been', 'VBN'), ('lost', 'VBN')]
```

**EXERCISE Part Two** Click into the code cell above and replace the sentence beginning with Porcini all the way to the word purveyors with your own sentence or two. Be sure to leave in all 3 sets of quotation marks. Then run the code.

## Congratulations!!! You've completed your first assignment.

Click File, and in the drop down menu, select "Save." Make sure that your notebook file saves in a folder that is shared with my gmail address: [lisarhody.gc@gmail.com](mailto:lisarhody.gc@gmail.com).

If you are someone who would like to push further, You could get started with the NLTK Book Chapter 1 here: <https://www.nltk.org/book/ch01.html>. You can create a new notebook to track your work and submit that the same way. If you do, please save the notebook with the same naming strategy: lastname-week.ipynb.

Congratulations! You're well on your way!