

COMS BC1016

Introduction to Computational Thinking and Data Science

## Lecture 2: Introduction to Python

Sep. 3, 2025

Copyright © 2025 Barnard College

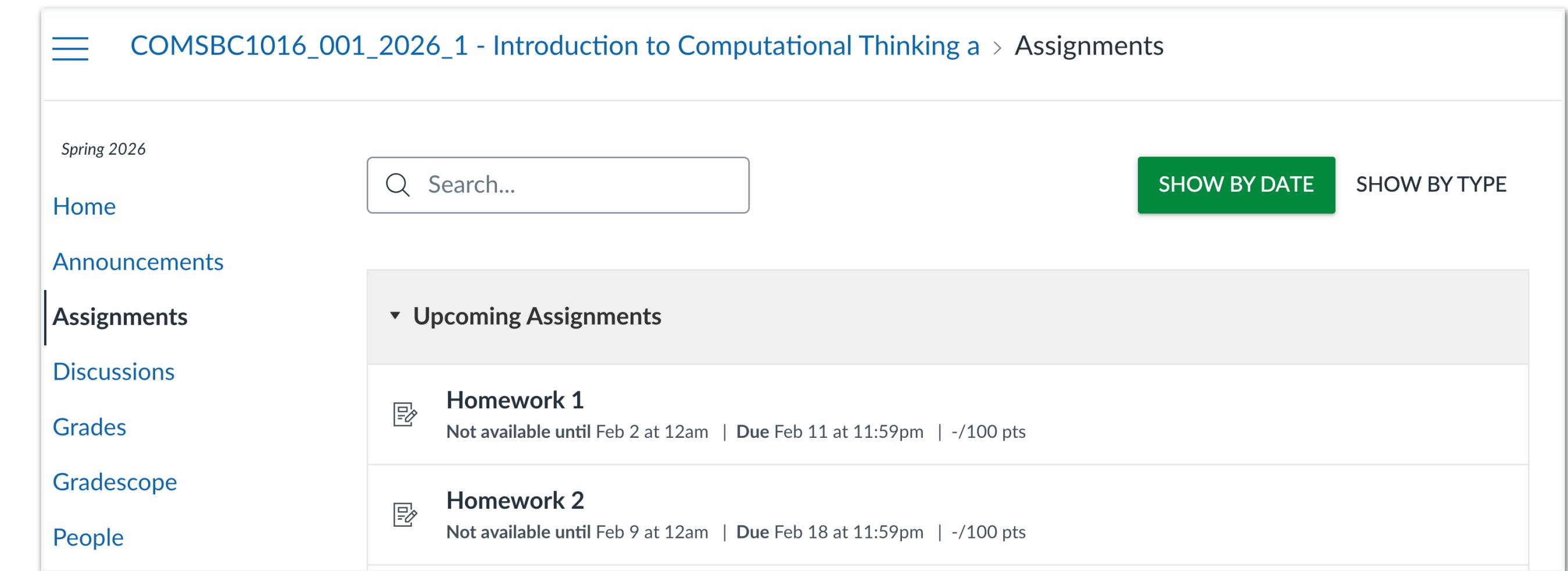
January 26, 2025

# Office Hours

- Office hours begin next week (first week of February)
- **Monday:**
  - Elena Lukac, 1:30-3pm in Milstein 503
  - Eysa Lee, 3-5pm in Milstein 512
- **Tuesday:** Nami Jain, 4-5:30pm in Milstein 503
- **Wednesday:** Madeline Gutierrez, 5:30-7pm in Milsten 503
- **Thursday:** Sathya Raman, 4-5:30pm in Milstein 503

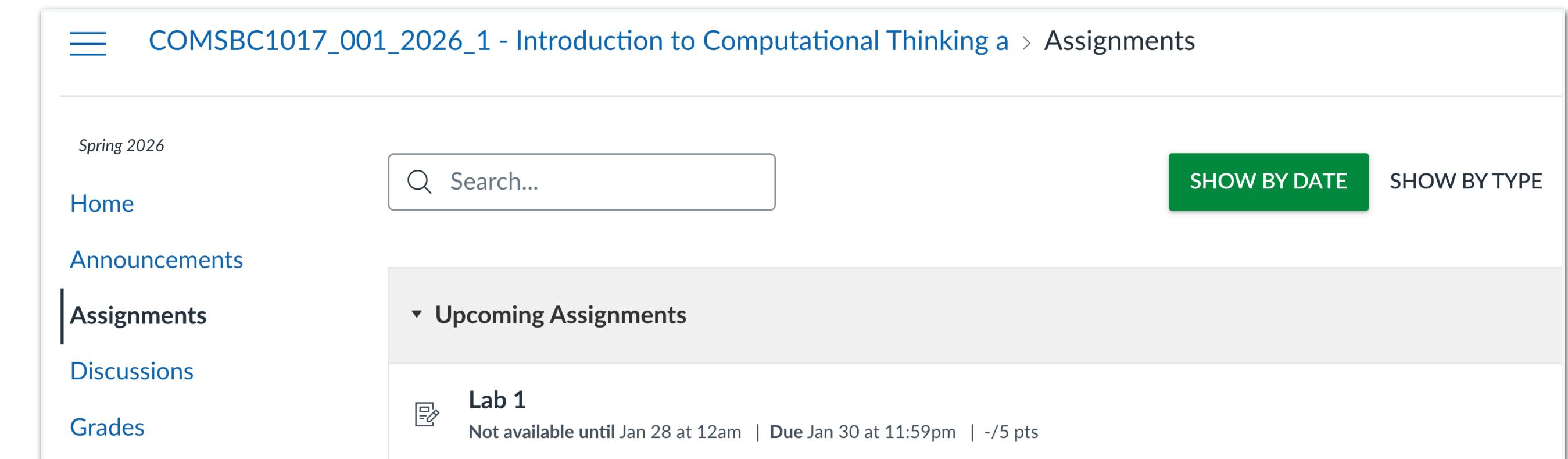
# Assignments

- Labs and HWs located on Courseworks
  - Download the files from the assignment page
  - HWs posted to 1016, Labs to 1017
- Submit HWs as .ipynb
- Submit Labs as PDFs



The screenshot shows the 'Assignments' section of the Courseworks platform for the course 'COMSBC1016\_001\_2026\_1 - Introduction to Computational Thinking a'. The sidebar on the left includes links for Home, Announcements, Assignments (which is selected), Discussions, Grades, Gradescope, and People. The main content area displays 'Upcoming Assignments' with two entries: 'Homework 1' and 'Homework 2'. Both assignments have a due date of Feb 11 at 11:59pm and are worth 100 points.

Assignment	Due Date	Points
Homework 1	Feb 11 at 11:59pm	100 pts
Homework 2	Feb 18 at 11:59pm	100 pts



The screenshot shows the 'Assignments' section of the Courseworks platform for the course 'COMSBC1017\_001\_2026\_1 - Introduction to Computational Thinking a'. The sidebar on the left includes links for Home, Announcements, Assignments (selected), Discussions, and Grades. The main content area displays 'Upcoming Assignments' with one entry: 'Lab 1'. The lab is due on Jan 30 at 11:59pm and is worth 5 points.

Assignment	Due Date	Points
Lab 1	Jan 30 at 11:59pm	5 pts

# More on Homeworks

- Homeworks will be released on Mondays and due the following Wednesday
  - HW 1 will be released next week Monday
  - Autograder tests for common mistakes and type errors
  - We will be checking more than just what tests are in the autograder!
  - True/False, short answer, and multiply choice are all manually graded
- Homeworks are submitted on Gradescope via Courseworks

# Lab Reminders

- **Reminder: You must be enrolled in a 1017 section!**
  - As of Sunday night, there are more people enrolled in 1016 than 1017
  - Labs begin this week
- **Email your TA if you'll be late or missing!**
  - 50% of your lab grade is attendance!
  - One unexcused absence + lowest lab dropped

# Reminder: Midterm Exam

- Paper exam happens during class **Wednesday, March 11, 2026**
  - This is the week before Spring Recess
- You will be allowed a note sheet to use as a reference during the exam
  - It will be submitted along with your exam
- If you need particular accommodations, please contact CARDS

# Course Website

Slides, emails, and helpful links are on the course website:

<https://www.eysalee.com/courses/s26/bc1016.html>

## Course Links

**Jupyter Hub:** [Link](#) (login required)

**Class Discussion Forum:** [EdStem](#) (login required)

**Courseworks:** [Link](#)

**Syllabus:** [Link](#)

## Resources

### Python Resources:

Data8 Python Reference: <https://www.data8.org/fa24/reference/>

DataScience Python Library Developer Documentation: <https://www.data8.org/datascience/>

**Data8 Textbook:** <https://inferentialthinking.com/chapters/intro.html>

## Lecture Schedule

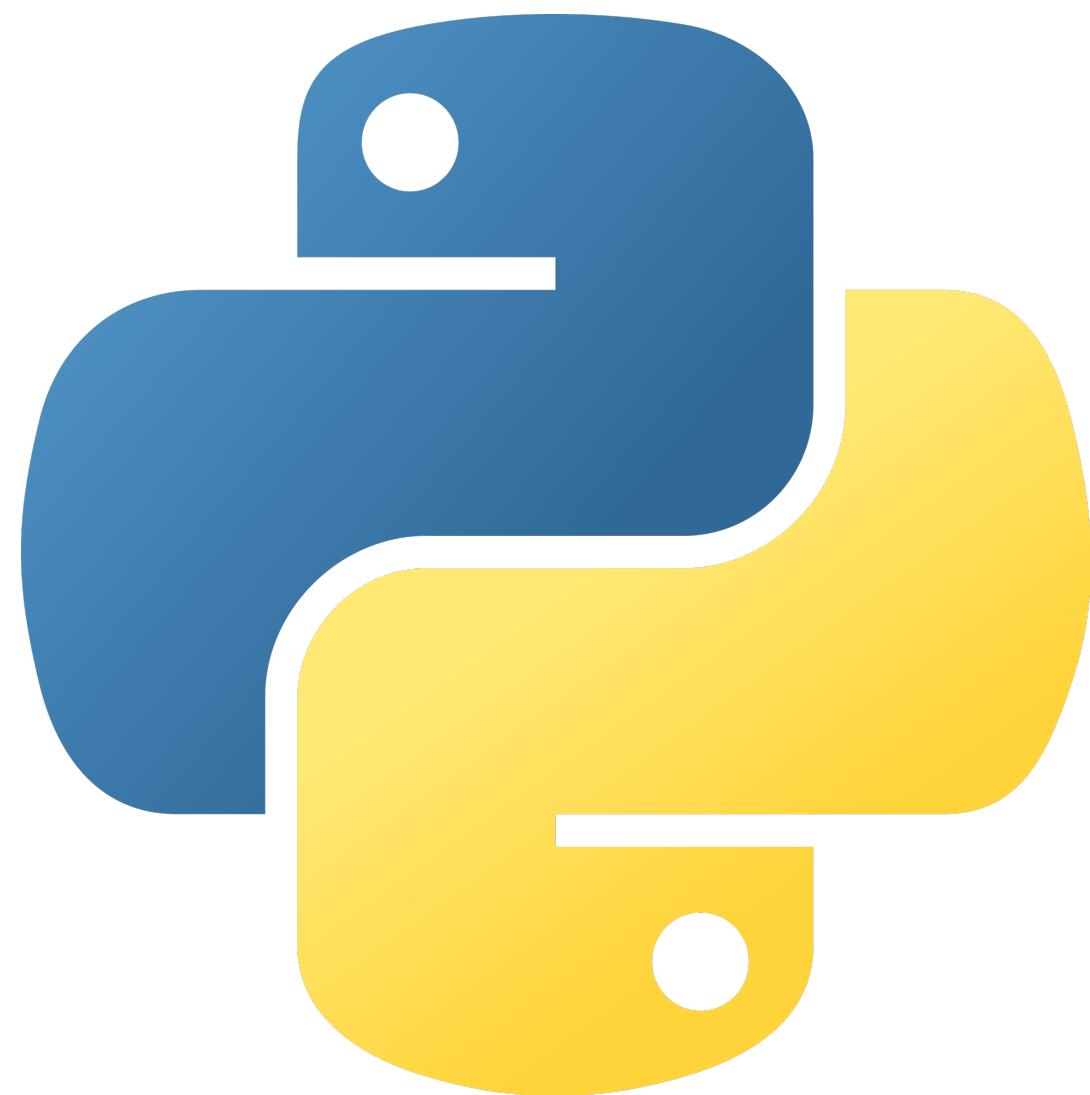
The schedule below will be updated as the course progresses.

Week	Date	Topic	Lab	Assignment
1	1/21	1 - Introduction [ <a href="#">Slides</a> ]	No Lab	
2	1/26	2 - Introduction to Python [ <a href="#">Slides</a> ] (Remote - Snow Day)		

# Python

# Python Intro

- Popular programming for software development
- Especially popular for data science
- Learning programming is about learning how to think computational & transfers to other languages



# How to Approach Programming in BC1016

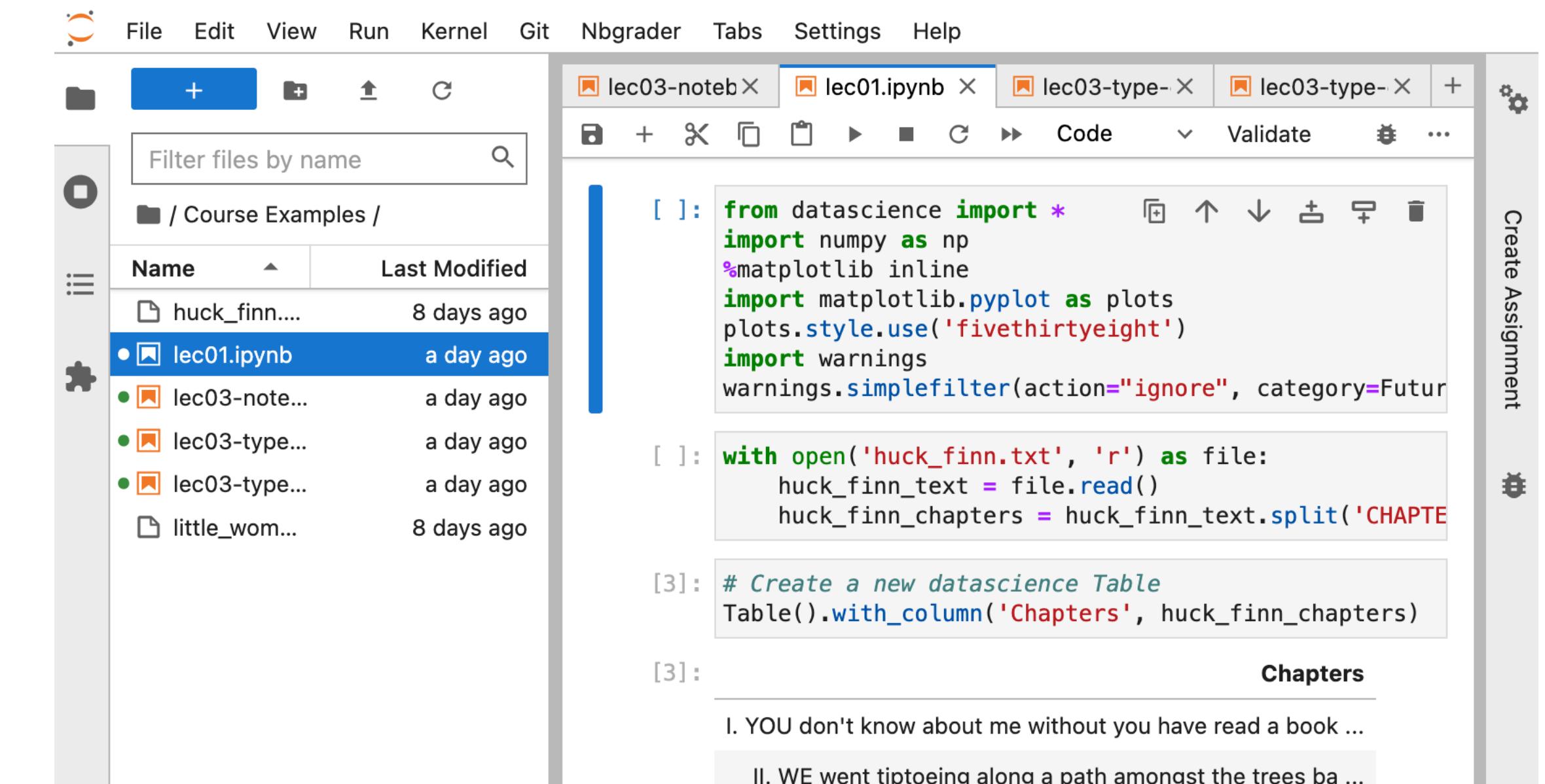
Programming is all about practice

- Class: Demos introduce terms & rules
- Labs & Homework: Try out programming yourself

Goal: writing your own code that can solve new problems!

# Jupyter Notebooks

- Can be run locally or in cloud-based environments
  - For this class: JupyterHub Server
- Benefits of cloud-based
  - Access anywhere
  - Expandable compute resources



The screenshot shows a Jupyter Notebook interface. On the left, there's a sidebar with a file browser titled 'Course Examples'. It lists several files: 'huck\_finn....' (8 days ago), 'lec01.ipynb' (a day ago, currently selected), 'lec03-note...', 'lec03-type...', 'lec03-type...', and 'little\_wom...'. The main area has tabs for 'lec03-notebook', 'lec01.ipynb', 'lec03-type...', and 'lec03-type...'. Below the tabs, there's a toolbar with icons for file operations like '+' and 'C'. The code editor shows a snippet of Python code:

```
[1]: from datascience import *
import numpy as np
%matplotlib inline
import matplotlib.pyplot as plots
plots.style.use('fivethirtyeight')
import warnings
warnings.simplefilter(action="ignore", category=FutureWarning)

[2]: with open('huck_finn.txt', 'r') as file:
    huck_finn_text = file.read()
    huck_finn_chapters = huck_finn_text.split('CHAPTER')

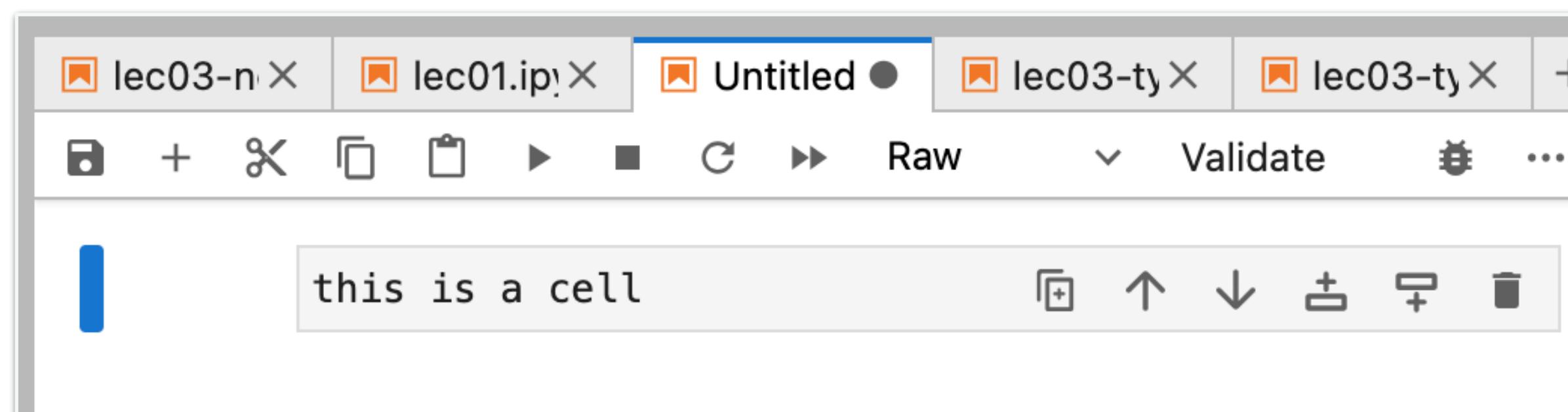
[3]: # Create a new datascience Table
Table().with_column('Chapters', huck_finn_chapters)

[3]: Chapters
I. YOU don't know about me without you have read a book ...
II. WE went tiptoeing along a path amongst the trees ba ...
```

Let's poke around in Jupyter...

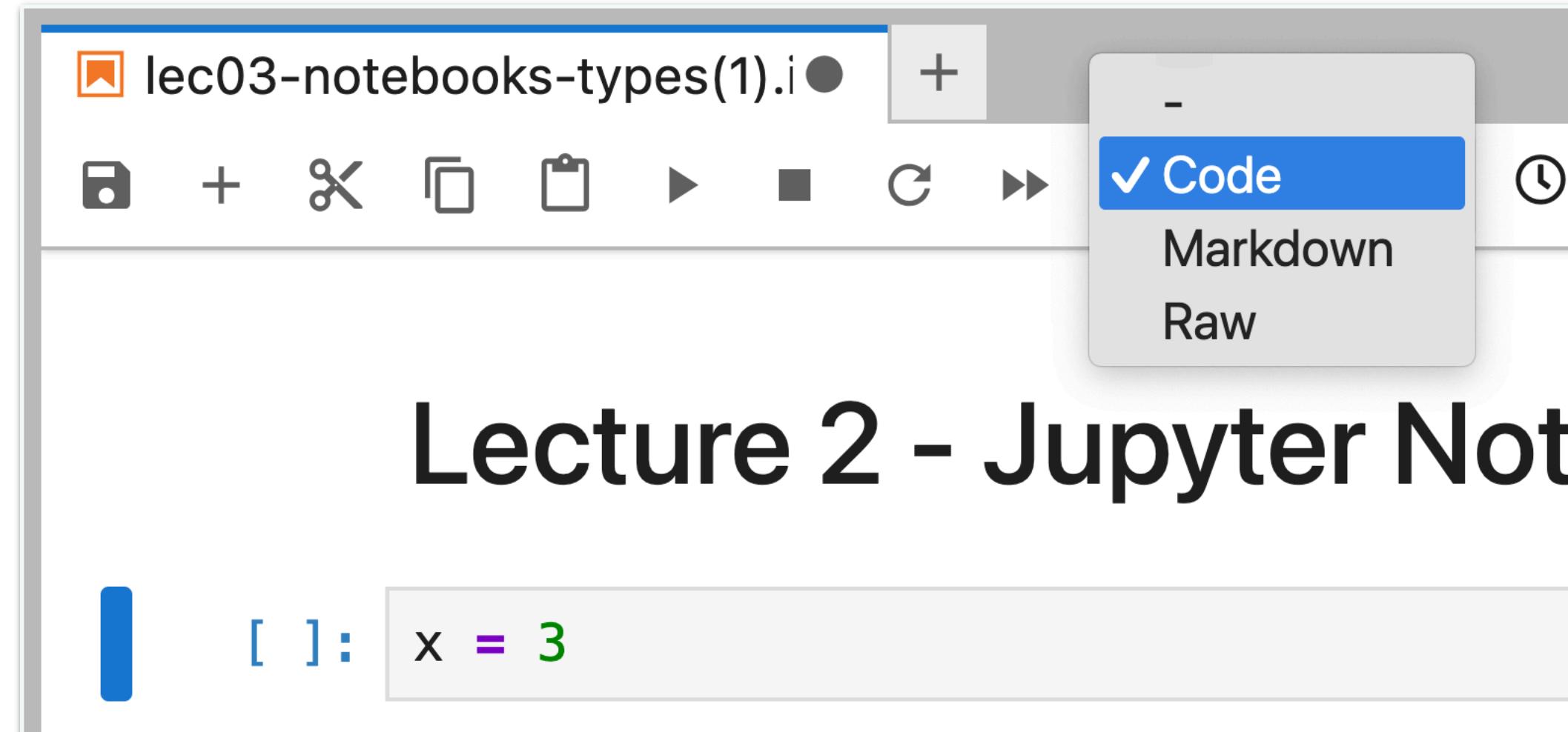
# Jupyter Notebooks: Terms

- **Cells:** block or section for writing code / notes / etc
- **Kernel:** executes code in cells (when you run a “cell”)

A screenshot of a Jupyter Notebook interface showing a cell execution. The cell contains the text 'this is a cell'. Below it, a code cell is shown with the command '[1]: print("Hello World")'. The output of the cell is 'Hello World', displayed in black text. The interface includes standard Jupyter controls for managing cells.

# Jupyter Notebooks: Cell Types

- **Code:** For writing and running Python code
- **Markdown:** Markup language for formatting plaintext
  - You've likely come across it on the internet before (# for headings, \*text\* for italics, \*\*text\*\* for bold, ...)
- **Raw**



A screenshot of a Jupyter Notebook interface showing two cells. The top cell is a Markdown cell containing the text "# This is a Heading". The bottom cell is also a Markdown cell containing the text "This is a heading". Both cells have a blue vertical bar on the left. At the top of the interface, there's a toolbar with icons and a dropdown menu set to "Markdown".

# Data Types

# Numbers

- **Integers:** Whole numbers
  - e.g., 3, -10, 25
- **Floats:** Anything with decimals
  - e.g., 3.1, -10.2, 2.0
- Basic calculations
  - e.g., +, -, \*, /

# Numbers

- **Integers:** Whole numbers
  - e.g., 3, -10, 25
- **Floats:** Anything with decimals
  - e.g., 3.1, -10.2, 2.0
- Basic calculations
  - e.g., +, -, \*, /

Some questions:

1. What happens if we add different number types?
  - a. Multiply or divide?
2. What if I wanted 2 raised to the power of 4?

# Strings

- Text in python
- You can specify a string using either single quotes or double quotes
  - "a"
  - 'This is a sentence'
  - "This is another sentence. Wow!"

Some questions:

1. What happens if you start a string with a double quote and end it with a single quote?
2. What if I want to write the following sentence as a string using single quotes?
  - a. That's weird

# Booleans

- True or False values
- Useful for conditional statements
  - Usually of the form “if some statement is true, then execute some code. Otherwise, do something else.”
- Use == to check equivalence
- We will talk more about Booleans in a few weeks

# Lists

- Lists are sequences of values
- Start with brackets and each element is separated by commas
  - [2,4,6,8]
  - ['apples', 'bananas', 'oranges']
  - ['apples', 'bananas', 'apples', 'apples']
- Lists are zero-indexed
  - The first element is the 0th and the last is length-1

# datascience Arrays

- In this class we'll mostly use NumPy arrays
  - Can make arrays using `datascience.make_array` or `numpy.array`
- Elements of an array should have the same type
  - e.g., `["Mystery", "Abby", "Jinu", "Baby", "Romance"]` or `[1, 2, 3, 5]`
- First element is index 0 and the last is the length-1
- Can perform component-wise arithmetic
  - Note this only works for numpy arrays but not built-in Python lists!

```
from datascience import *
onetwothree = make_array(1,2,3)
onetwothree * 2
array([2, 4, 6])
```

```
from datascience import *
onetwothree = make_array(1,2,3)
twothreefour = make_array(2,3,4)
onetwothree + twothreefour
array([3, 5, 7])
```

~~onetwothree = [1,2,3]  
twothreefour = [2,3,4]  
onetwothree + twothreefour  
[1, 2, 3, 2, 3, 4]~~

# Functions

- Type of abstraction for pre-defined set of code or instructions
  - Commonly takes inputs, performs a computation, and produces output
  - Many useful things built in!
  - Some useful functions aren't built in, so we need to import them

```
abs(-1)  
1
```



```
max(4, 200, 7)  
Function      inputs separated  
200           by commas
```

```
import numpy  
numpy.sqrt(4)  
2.0
```

We tell Python this function is from here

# Built-in conversions

- You can convert values to a string using `str( . . . )`
  - `str(5)` becomes "5"
- You can convert strings of numbers to numbers
  - `int('12'), float('1.2')`

## Questions:

1. What happens if you try to convert a float to an integer?
2. What happens if you try to convert non-numbers to a number?

# Variables and Assignments

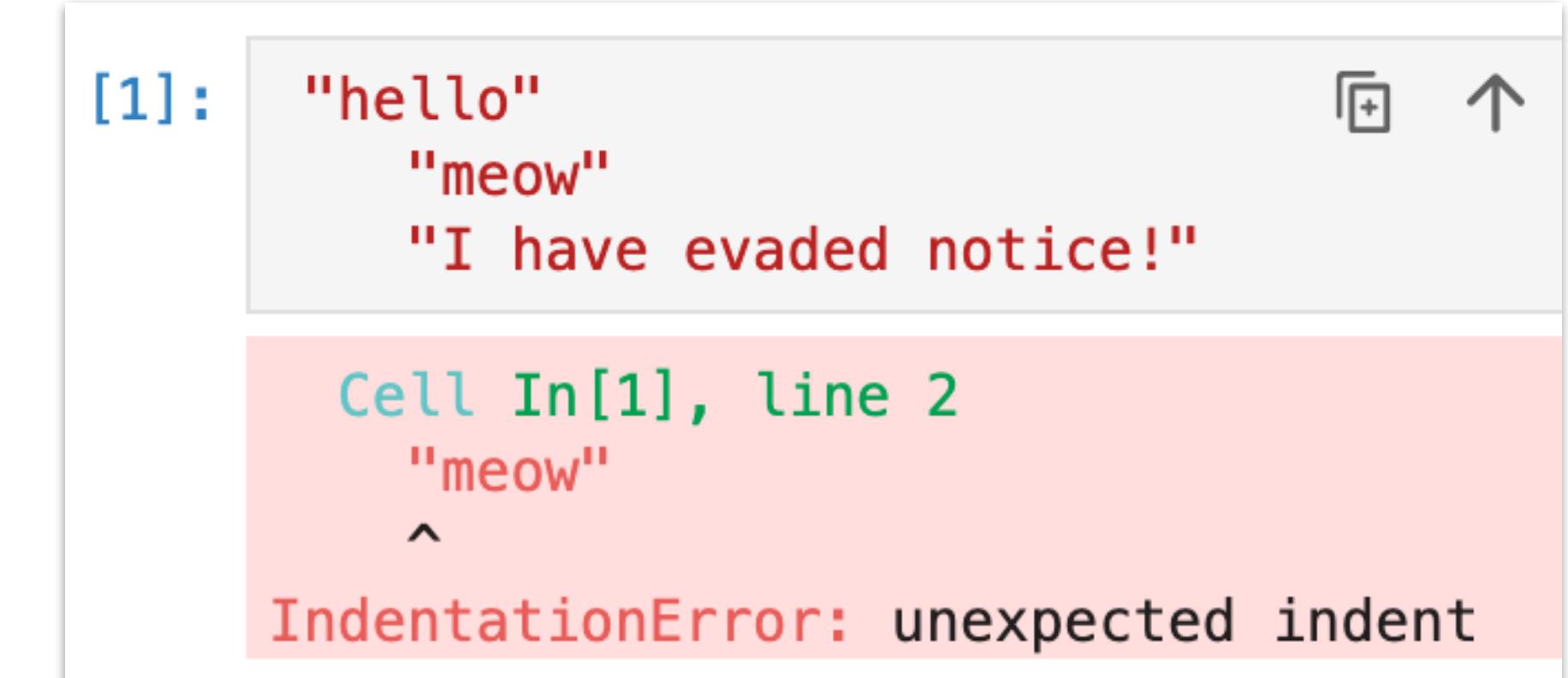
- **Assignments** change the meaning of the name to the left the = symbol
- **Variables** are values you can assign values to
  - “Variable” because they can change
- You can assign outputs of functions and operations to variables

```
max_of_list = max(4,200,7)  
max_of_list + 5
```

205

# Python quirks

- Indentation and new lines matter
  - Be careful not to add extra spaces or indentations at the beginning of the line!
- Python runs line-by-line
  - It'll stop as soon as it runs into an issue and tell you what's wrong
- Lines starting with # are comments and are ignored



The screenshot shows a Jupyter Notebook cell with the following content:

```
[1]: "hello"
     "meow"
     "I have evaded notice!"
```

Below the code, a pink error message box contains:

Cell In[1], line 2  
"meow"  
^  
IndentationError: unexpected indent

# Types

Type	Example
Int	3
Float	3.0
String	'three'
Boolean	True
Arrays	[1, 2, 3, 4]
Functions	abs(-5)

# Type Exercise

Let's say you have defined the following variables in your notebook

```
x = 3
```

```
y = '4'
```

```
z = '5.6'
```

What would the source of the error in these examples?

How could you fix it?

1. x + y

2. x + int(y + z)

3. str(x) + int(y)

4. y + float(z)

# Type Exercise

Let's say you have defined the following variables in your notebook

```
x = 3
```

```
y = '4'
```

```
z = '5.6'
```

What would the source of the error in this example:

```
x + y
```

How could you fix it?

# Type Exercise

Let's say you have defined the following variables in your notebook

```
x = 3
```

```
y = '4'
```

```
z = '5.6'
```

What would the source of the error in this example:

```
x + int(y + z)
```

How could you fix it?

# Type Exercise

Let's say you have defined the following variables in your notebook

```
x = 3
```

```
y = '4'
```

```
z = '5.6'
```

What would the source of the error in this example:

```
str(x) + int(y)
```

How could you fix it?

# Type Exercise

Let's say you have defined the following variables in your notebook

```
x = 3
```

```
y = '4'
```

```
z = '5.6'
```

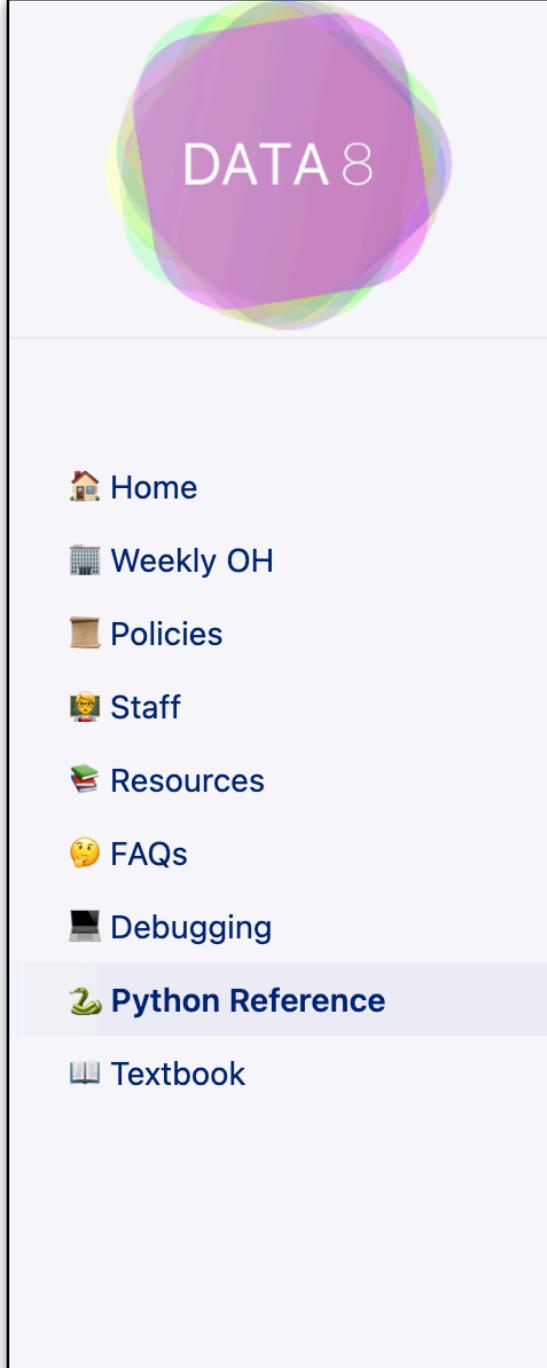
What would the source of the error in this example:

```
y + float(z)
```

How could you fix it?

# Python Reference

<https://www.data8.org/sp25/reference/>



The screenshot shows the Data 8 website interface. At the top right is a search bar labeled "Search Data 8" and navigation links for "OH Queue", "Textbook", and "Extensions". The main content area has a title "Detailed Python Reference Sheet". Below it is a note about contributions and a "TABLE OF CONTENTS" section. The sidebar on the left contains links: Home, Weekly OH, Policies, Staff, Resources, FAQs, Debugging, Python Reference (which is highlighted), and Textbook.

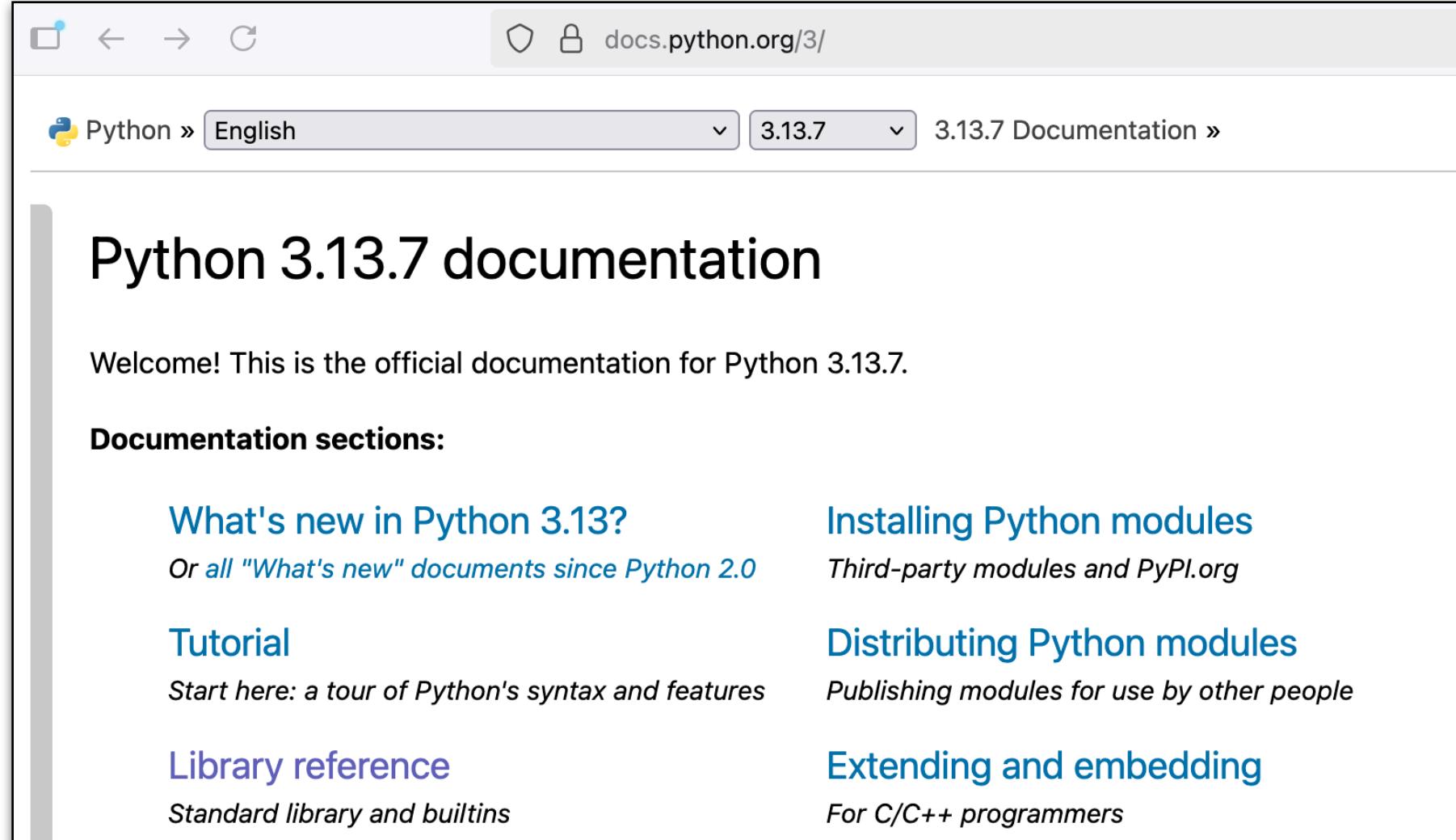
## Detailed Python Reference Sheet

Created by Nishant Kheterpal and Jessica Hu, Contributions by Jonathan Ferrari, Updated and Maintained by Edwin Vargas and Bing Concepcion

### TABLE OF CONTENTS

- 1 [Detailed Python Reference Sheet](#)
  - a [Abbreviated Reference Sheet](#)
  - b [Table Functions and Methods](#)
  - c [String Methods](#)
  - d [Array Functions and Methods](#)
  - e [Table Filtering Predicates](#)
  - f [Miscellaneous Functions](#)
  - g [JupyterHub Keyboard Shortcuts](#)

<https://docs.python.org/3/>



The screenshot shows the Python 3.13.7 documentation page. The browser address bar shows "docs.python.org/3/". The page title is "Python 3.13.7 documentation". A welcome message states "Welcome! This is the official documentation for Python 3.13.7." Below it is a "Documentation sections:" heading with several links: "What's new in Python 3.13?", "Tutorial", "Library reference", "Installing Python modules", "Distributing Python modules", and "Extending and embedding".

## Python 3.13.7 documentation

Welcome! This is the official documentation for Python 3.13.7.

**Documentation sections:**

<a href="#">What's new in Python 3.13?</a> <small>Or all "What's new" documents since Python 2.0</small>	<a href="#">Installing Python modules</a> <small>Third-party modules and PyPI.org</small>
<a href="#">Tutorial</a> <small>Start here: a tour of Python's syntax and features</small>	<a href="#">Distributing Python modules</a> <small>Publishing modules for use by other people</small>
<a href="#">Library reference</a> <small>Standard library and builtins</small>	<a href="#">Extending and embedding</a> <small>For C/C++ programmers</small>

# Next Class

- Today
  - Jupyter Notebooks
  - Data Types
- *Wednesday*
  - Tables
- Monday
  - Charts & Visualization