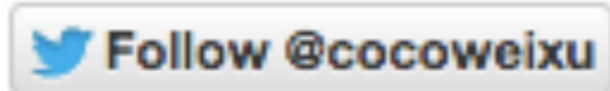


Social Media & Text Analysis

part 1 - Intro to Python



Ohio State University

Instructor: Wei Xu

Website: socialmedia-class.org



python



Why Python?

- Python is an object-oriented and high level programming language (first released in 1991).
- Very beginner-friendly!
 - shorter code needed for the same task.
- Very powerful!
 - many well-maintained libraries (e.g. numpy, Matplotlib, Scikit-learn, PyTorch, TensorFlow,)
 - a popular programming language in AI and machine learning research

Simplicity of Python

- Create a list of integers:

- Python

```
nums = [45, 23, 51, 32, 5]
```

- Java, in contrast:

```
List<Integer> nums =  
    Arrays.asList(new Integer[] {45, 23, 51, 32, 5});
```

Simplicity of Python

- Create a list of integers, and print them out:

- Python

```
nums = [45, 23, 51, 32, 5]
for idx, num in enumerate(nums):
    print idx, num
```

- Java, in contrast:

```
List<Integer> nums =
    Arrays.asList(new Integer[] {45, 23, 51, 32, 5});

for (int i = 0; i < nums.size(); i++) {
    String number = nums.get(i);
    System.out.println(i + " " + number);
}
```

Simplicity of Python

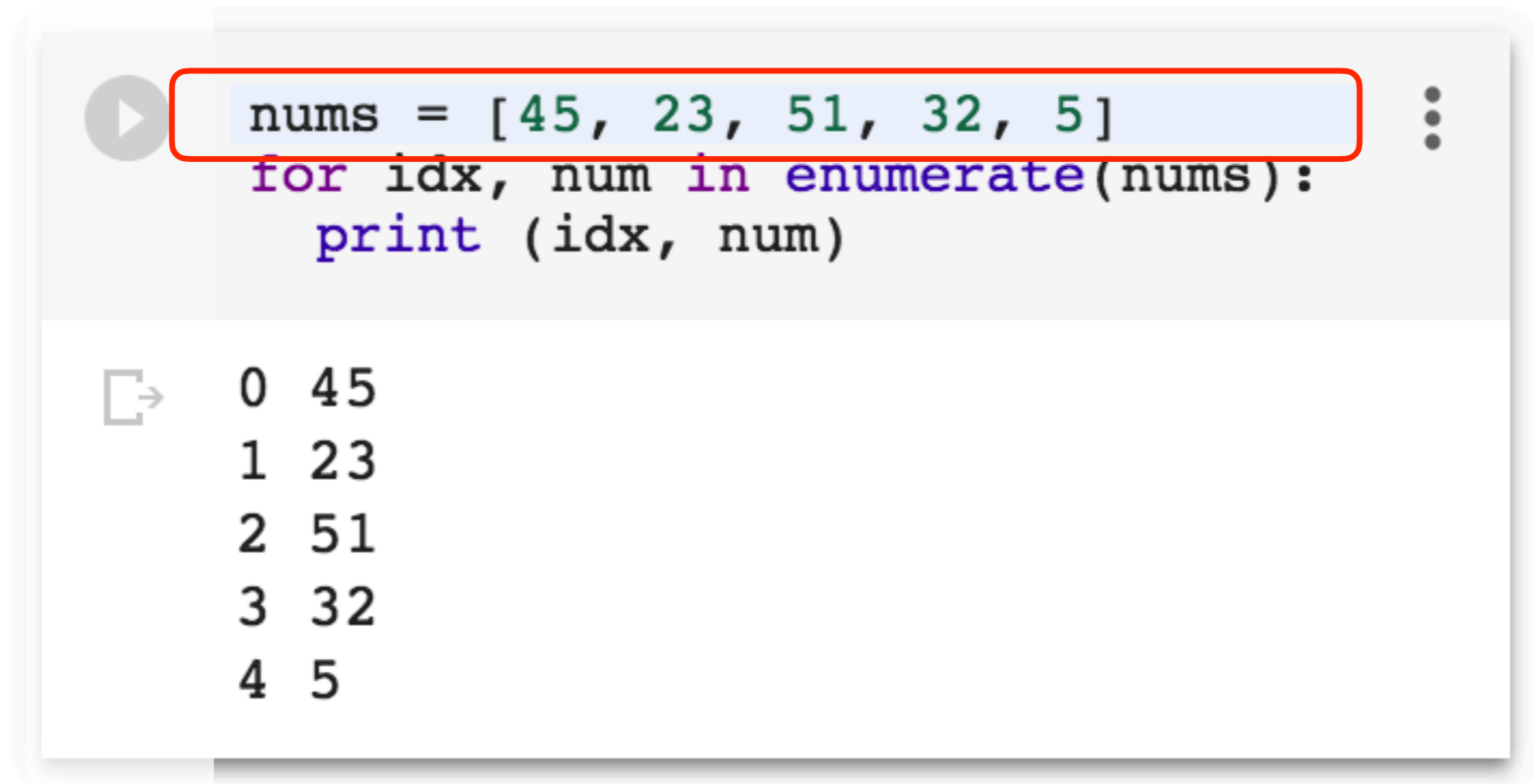
- Create a list of integers, and print them out:

```
▶ nums = [45, 23, 51, 32, 5]
  for idx, num in enumerate(nums):
    print (idx, num)
```

↳ 0 45
1 23
2 51 ← output
3 32
4 5

Simplicity of Python

- Create a list of integers, and print them out:



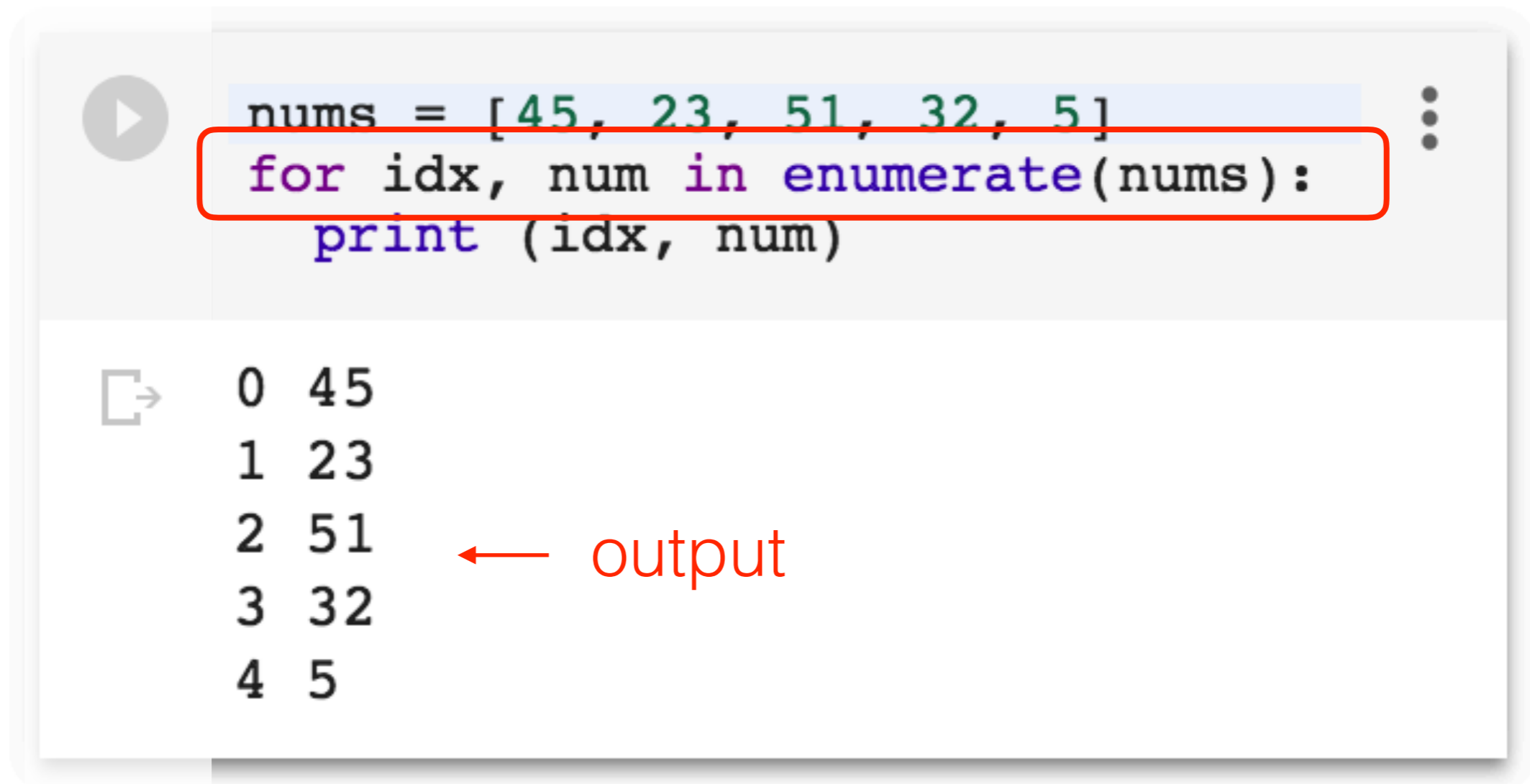
```
▶ nums = [45, 23, 51, 32, 5]
  for idx, num in enumerate(nums):
    print (idx, num)
```

↳ 0 45
1 23
2 51
3 32
4 5

Lists are declared using [item1, item2, item3, ...]

Simplicity of Python

- Create a list of integers, and print them out:



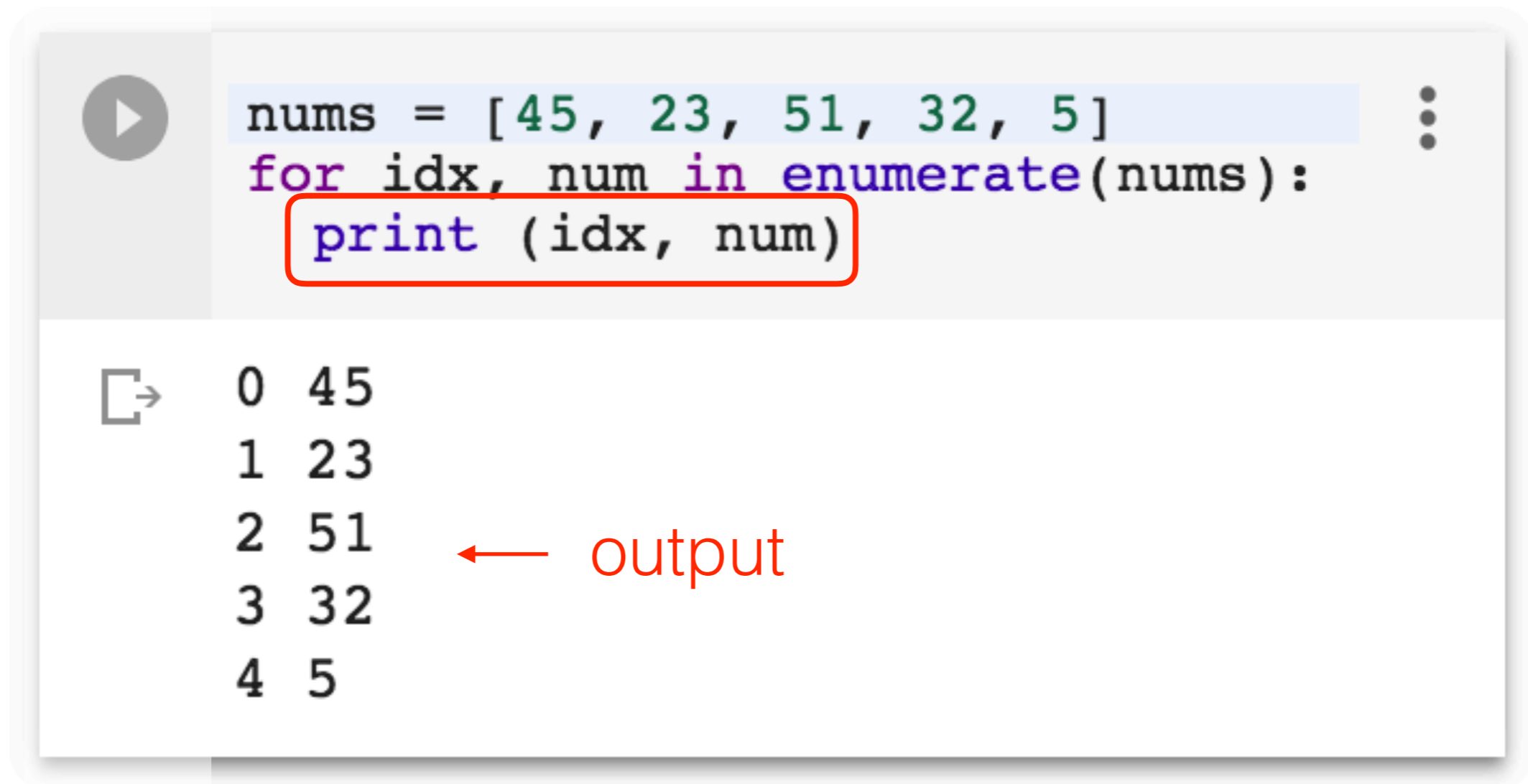
```
nums = [45, 23, 51, 32, 5]
for idx, num in enumerate(nums):
    print (idx, num)
```

0 45
1 23
2 51 ← output
3 32
4 5

Iterates over each item in the list, yielding (index, value) tuples

Simplicity of Python

- Create a list of integers, and print them out:



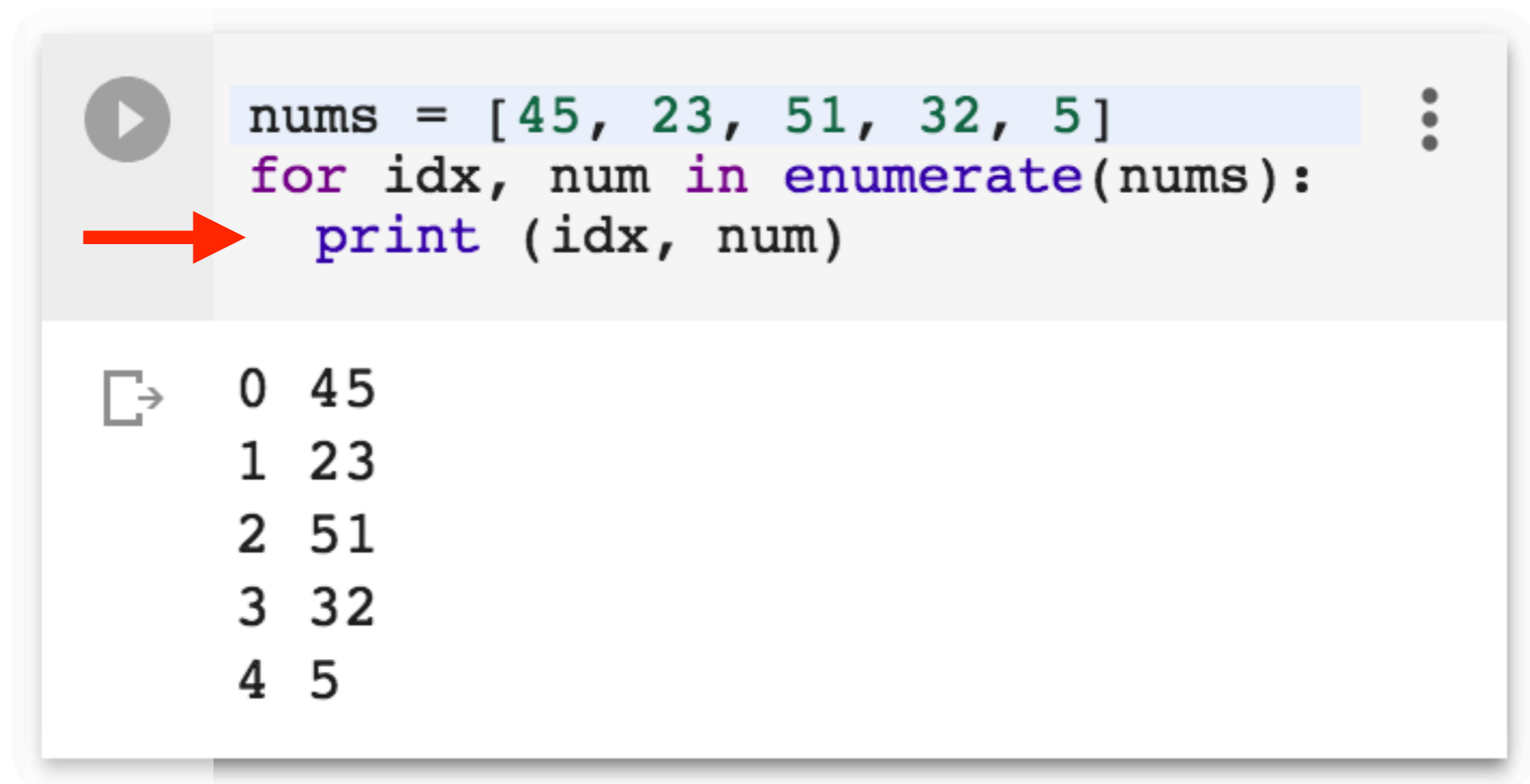
```
nums = [45, 23, 51, 32, 5]
for idx, num in enumerate(nums):
    print (idx, num)
```

0 45
1 23
2 51 ← output
3 32
4 5

simple and very useful print() function

Simplicity of Python

- Create a list of integers, and print them out:



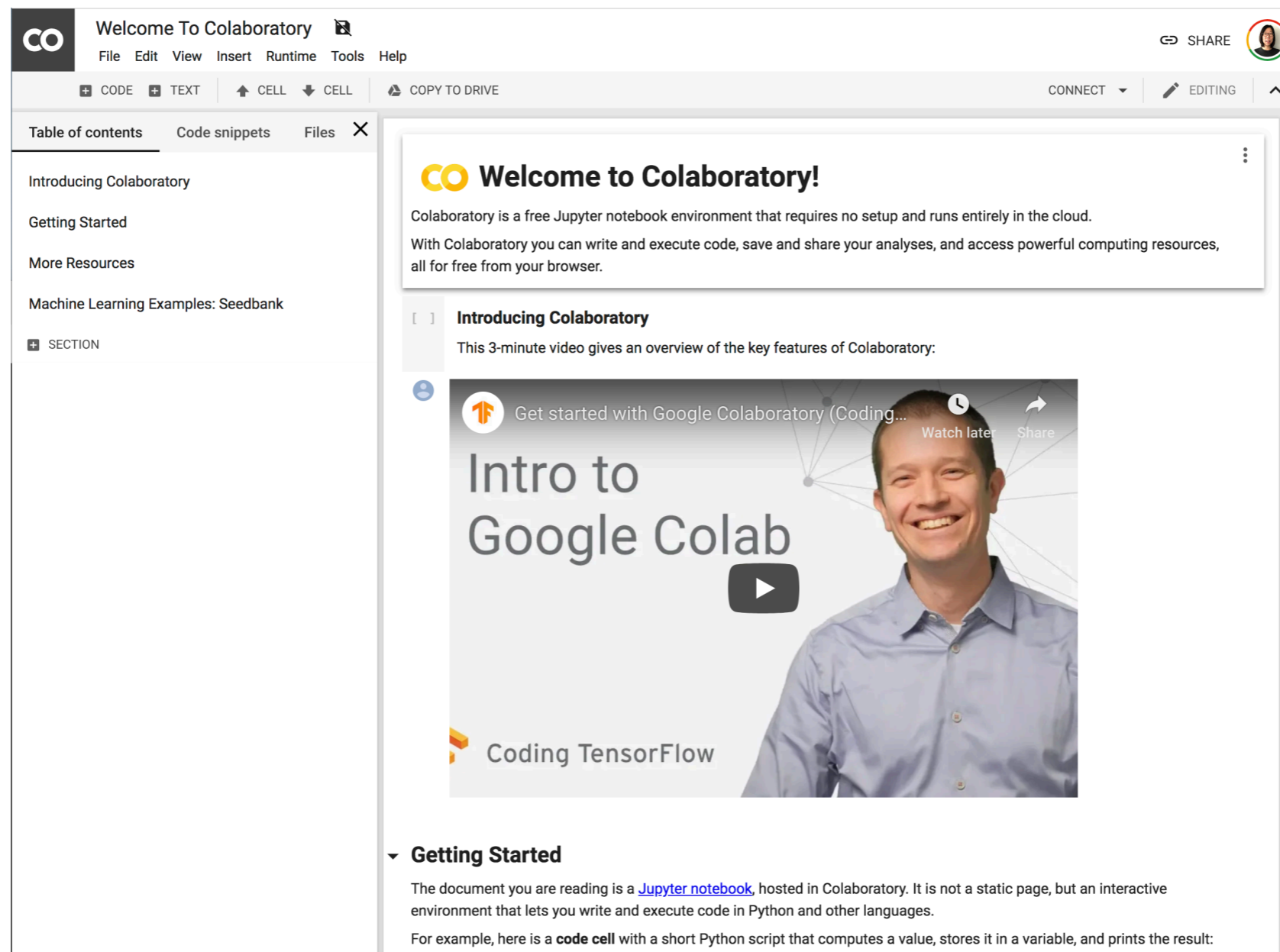
```
▶ nums = [45, 23, 51, 32, 5]
  for idx, num in enumerate(nums):
    print (idx, num)
```

↳ 0 45
1 23
2 51
3 32
4 5

Indentation is very important in Python!

Try it Out!

- We will use **Google's Colab** programming environment:



The screenshot displays the Google Colaboratory web interface. At the top, there is a navigation bar with the Colab logo, the text "Welcome To Colaboratory", and a menu with options: File, Edit, View, Insert, Runtime, Tools, and Help. On the right side of the top bar, there are "SHARE" and "CONNECT" buttons, along with a user profile icon. Below the top bar is a secondary toolbar with buttons for "+ CODE", "+ TEXT", "CELL", "COPY TO DRIVE", "CONNECT", "EDITING", and an upward arrow. A left sidebar contains a "Table of contents" with links to "Introducing Colaboratory", "Getting Started", "More Resources", and "Machine Learning Examples: Seedbank". The main content area features a large "Welcome to Colaboratory!" message, followed by a video player titled "Intro to Google Colab" with a play button. Below the video, there is a section titled "Getting Started" with introductory text about the Jupyter notebook environment.

Welcome To Colaboratory

File Edit View Insert Runtime Tools Help

+ CODE + TEXT CELL COPY TO DRIVE CONNECT EDITING

Table of contents Code snippets Files

Introducing Colaboratory

Getting Started

More Resources

Machine Learning Examples: Seedbank

SECTION

Welcome to Colaboratory!

Colaboratory is a free Jupyter notebook environment that requires no setup and runs entirely in the cloud. With Colaboratory you can write and execute code, save and share your analyses, and access powerful computing resources, all for free from your browser.

Introducing Colaboratory

This 3-minute video gives an overview of the key features of Colaboratory:

Get started with Google Colaboratory (Coding... Watch later Share

Intro to Google Colab

Coding TensorFlow

Getting Started

The document you are reading is a [Jupyter notebook](#), hosted in Colaboratory. It is not a static page, but an interactive environment that lets you write and execute code in Python and other languages.

For example, here is a **code cell** with a short Python script that computes a value, stores it in a variable, and prints the result:

Basic String Operations

- A Code Sample:

```
▶ sent1 = "Hello world!" | # A comment.
print (sent1)             # Another one.

print (sent1[4])         # The 5th char (index starts from 0)

l = len(sent1)           # The length (in number of characters)
print ("There are " + str(l) + " charasters.")

tokens = sent1.split()   # Split a string by space
print (tokens)

print (len(tokens))     # The length (in number of tokens/words)
```

```
↳ Hello world!
o
There are 12 charasters.
['Hello', 'world!']
2
```

Basic String Operations

- A Code Sample:

```
▶ sent1 = "Hello world!" | # A comment.
print (sent1)             # Another one.

print (sent1[4])         # The 5th char (index starts from 0)

l = len(sent1)           # The length (in number of characters)
print ("There are " + str(l) + " charasters.")

tokens = sent1.split()   # Split a string by space
print (tokens)

print (len(tokens))      # The length (in number of tokens/words)
```

↳ Hello world!
o
There are 12 charasters. ← output
['Hello', 'world!']
2

Characters in a string can be assessed using the [] syntax.

Basic String Operations

- A Code Sample:

```
▶ sent1 = "Hello world!" | # A comment.
print (sent1)             # Another one.

print (sent1[4])         # The 5th char (index starts from 0)

l = len(sent1)           # The length (in number of characters)
print ("There are " + str(l) + " charasters.")

tokens = sent1.split()   # Split a string by space
print (tokens)

print (len(tokens))     # The length (in number of tokens/words)
```

```
↳ Hello world!
○
There are 12 charasters.
['Hello', 'world!']
2
```

The `len(string)` function returns the length of a string.

Basic String Operations

- A Code Sample:

```
▶ sent1 = "Hello world!" | # A comment.
print (sent1)             # Another one.

print (sent1[4])         # The 5th char (index starts from 0)

l = len(sent1)           # The length (in number of characters)
print ("There are " + str(l) + " charasters.")

tokens = sent1.split()   # Split a string by space
print (tokens)

print (len(tokens))     # The length (in number of tokens/words)
```

```
↳ Hello world!
○
There are 12 charasters. ← output
['Hello', 'world!']
2
```

The `str()` function converts values to a string data type.

Basic String Operations

- A Code Sample:

```
▶ sent1 = "Hello world!" | # A comment.
print (sent1)             # Another one.

print (sent1[4])         # The 5th char (index starts from 0)

l = len(sent1)           # The length (in number of characters)
print ("There are " + str(l) + " charasters.")

tokens = sent1.split()   # Split a string by space
print (tokens)

print (len(tokens))     # The length (in number of tokens/words)
```

```
↳ Hello world!
○
There are 12 charasters.
['Hello', 'world!']
2
```

← output

The + operator can concatenate two strings.

Basic String Operations

- A Code Sample:

```
▶ sent1 = "Hello world!" | # A comment.
print (sent1)             # Another one.

print (sent1[4])         # The 5th char (index starts from 0)

l = len(sent1)           # The length (in number of characters)
print ("There are " + str(l) + " charasters.")

tokens = sent1.split()   # Split a string by space
print (tokens)

print (len(tokens))     # The length (in number of tokens/words)
```

```
↳ Hello world!
o
There are 12 charasters.
['Hello', 'world!'] ← output
2
```

The `split()` function returns a list of substrings.

More Resources

- Python :
 - Google's class: <https://developers.google.com/edu/python/>
 - Christophe Morisset's notebook: https://github.com/Morisset/Python-lectures-Notebooks/blob/master/Notebooks/intro_Python.pdf
 - and many others ...
- Got a Error Messages or questions?
 - Search on Google
 - StackOverflow