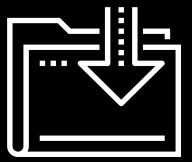
# Web Scraping With HTML & CSS

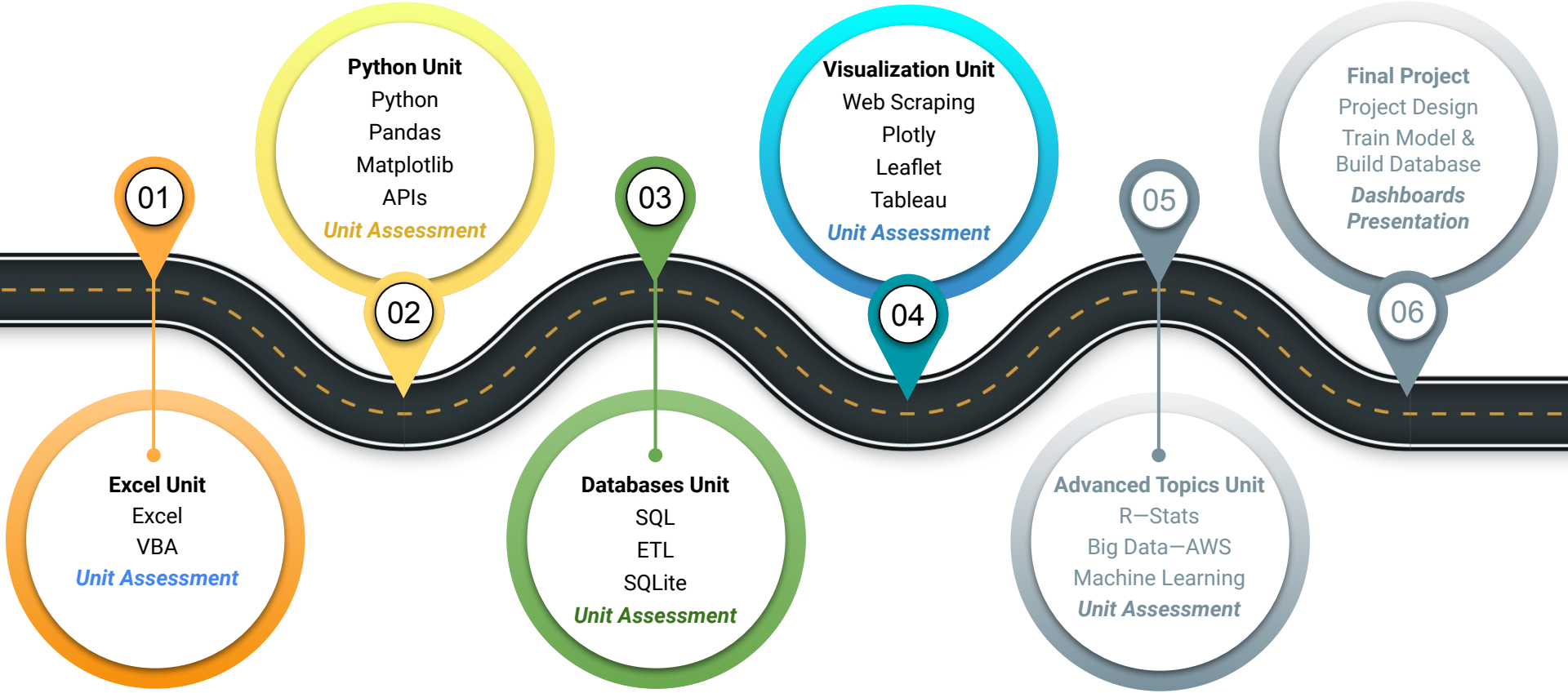
Data Boot Camp

Lesson 11.1



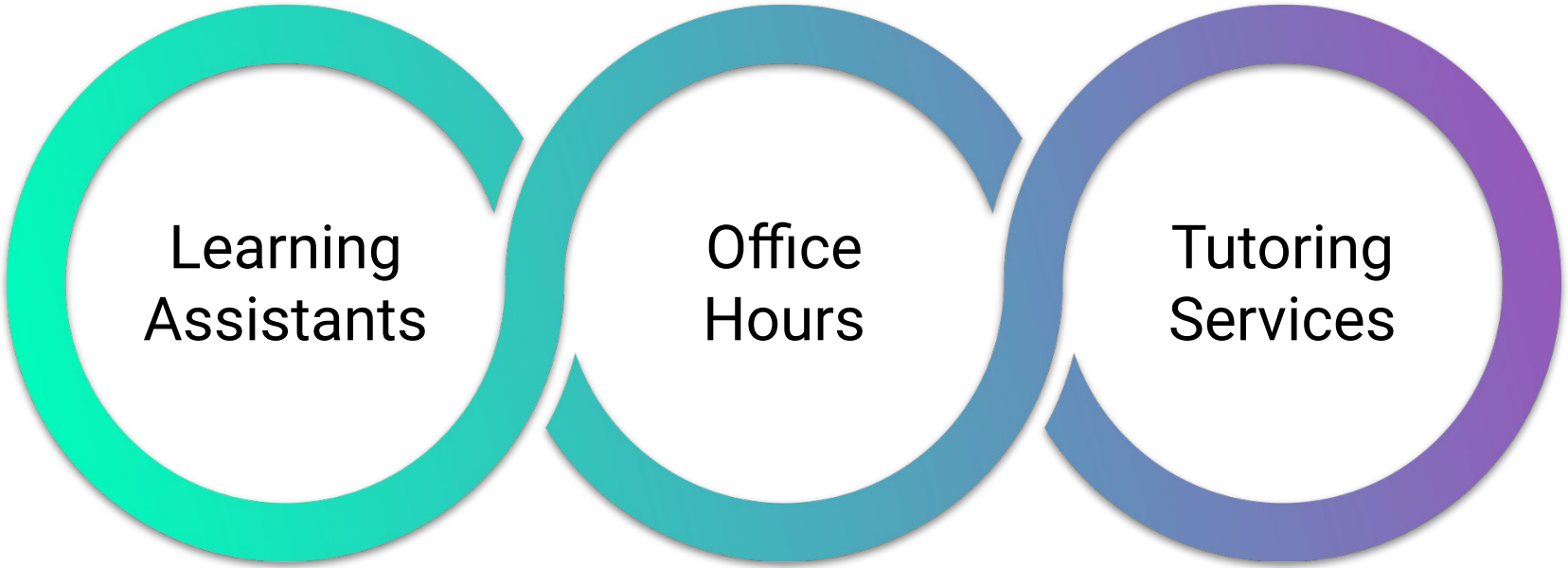
© 2022 edX Boot Camps LLC. Confidential and Proprietary. All Rights Reserved.

# The Big Picture



# Bootcamp Pointers

If you're struggling with this module, remember that you have the following resources to help you:



Module 11

# This Week:
# Web Scraping with HTML & CSS

# This Week: Web Scraping with HTML & CSS

By the end of this week, you'll know how to:

- Identify the basic structure of an HTML document
- Create a basic web page using HTML
- Use CSS selectors to style HTML elements
- Use BeautfulSoup to extract information from a web page
- Use Splinter to automate browser actions
- Store scraped information in a Python data structure

## This Week's Challenge

Using the skills learned throughout the week, scrape and analyze information from Mars websites.

## Career Connection

How will you use this module’s content in your career?

Module 11

# How to Succeed This Week

Module 11

# Today’s Agenda

# Today's Agenda

By completing today's activities, you'll learn the following skills:

01 Create a basic HTML document

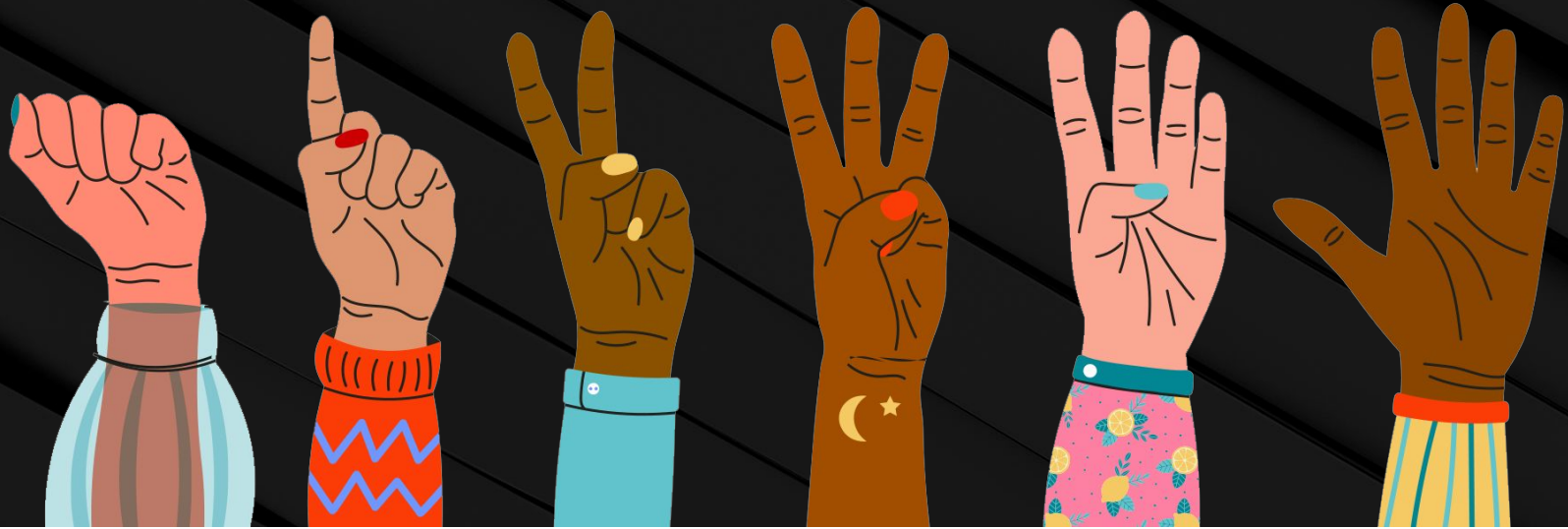
02 Scrape a HTML document with BeautifulSoup

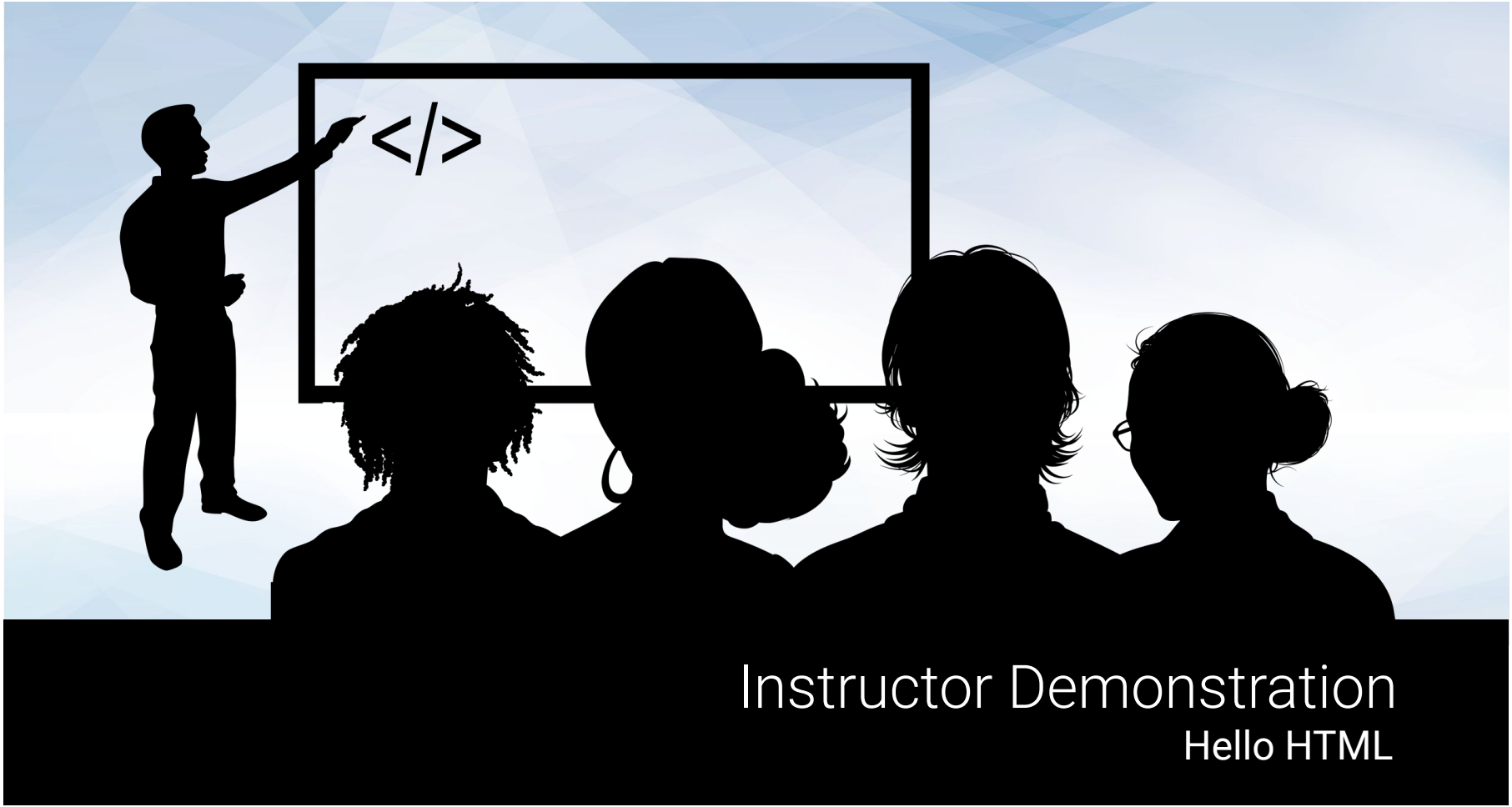
03 Style an HTML document with CSS

**Make sure you've downloaded any relevant class files!**

## FIST TO FIVE:

How comfortable do you feel with this topic?






# Instructor Demonstration

## Hello HTML

# Hello HTML

## HTML5

- HTML is one of the three base languages behind every website.
- It defines all the basic content and a bit of formatting.

# Hello HTML

HTML elements are rendered by the browser as visible parts of a webpage

# Hello HTML

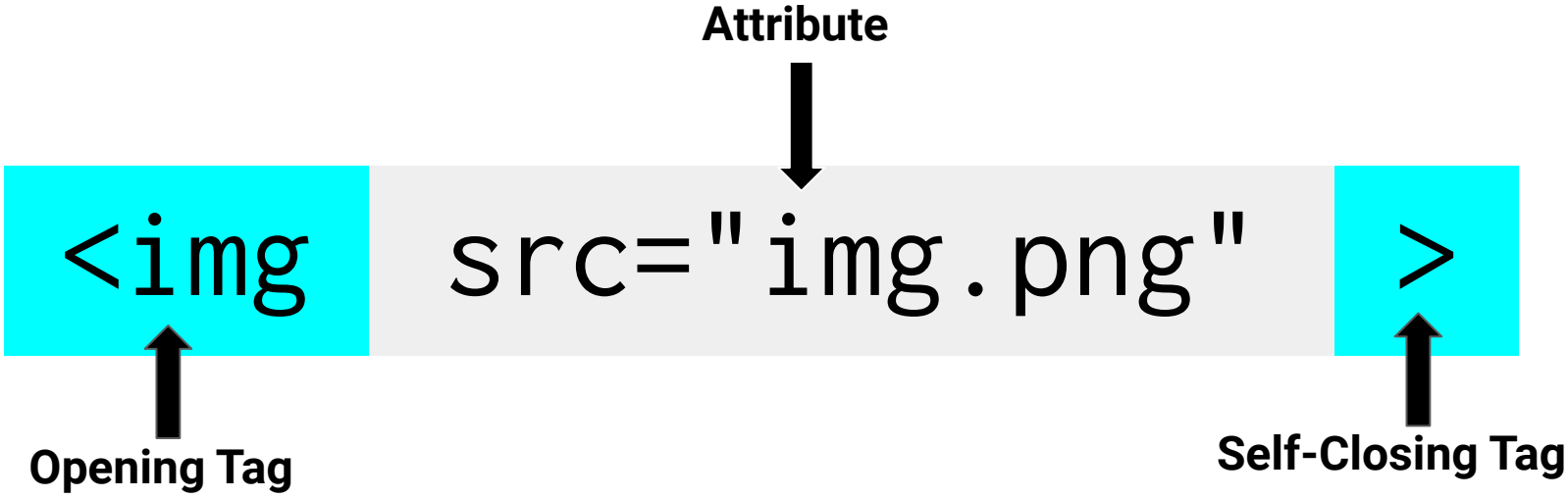
## HTML Syntax (Basic)

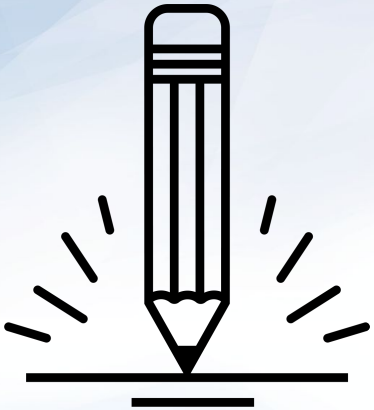
# Hello HTML

## HTML Syntax (with Attribute)

# Hello HTML

## Tricky Tags (Self-Closing)





# Activity: My first HTML

In this activity, you will create your first web page using HTML.

**Suggested Time:**
20 minutes

# My First HTML Instructions

In this activity, you will create a web page containing a header, an image, a link, and lists. Below is an example:

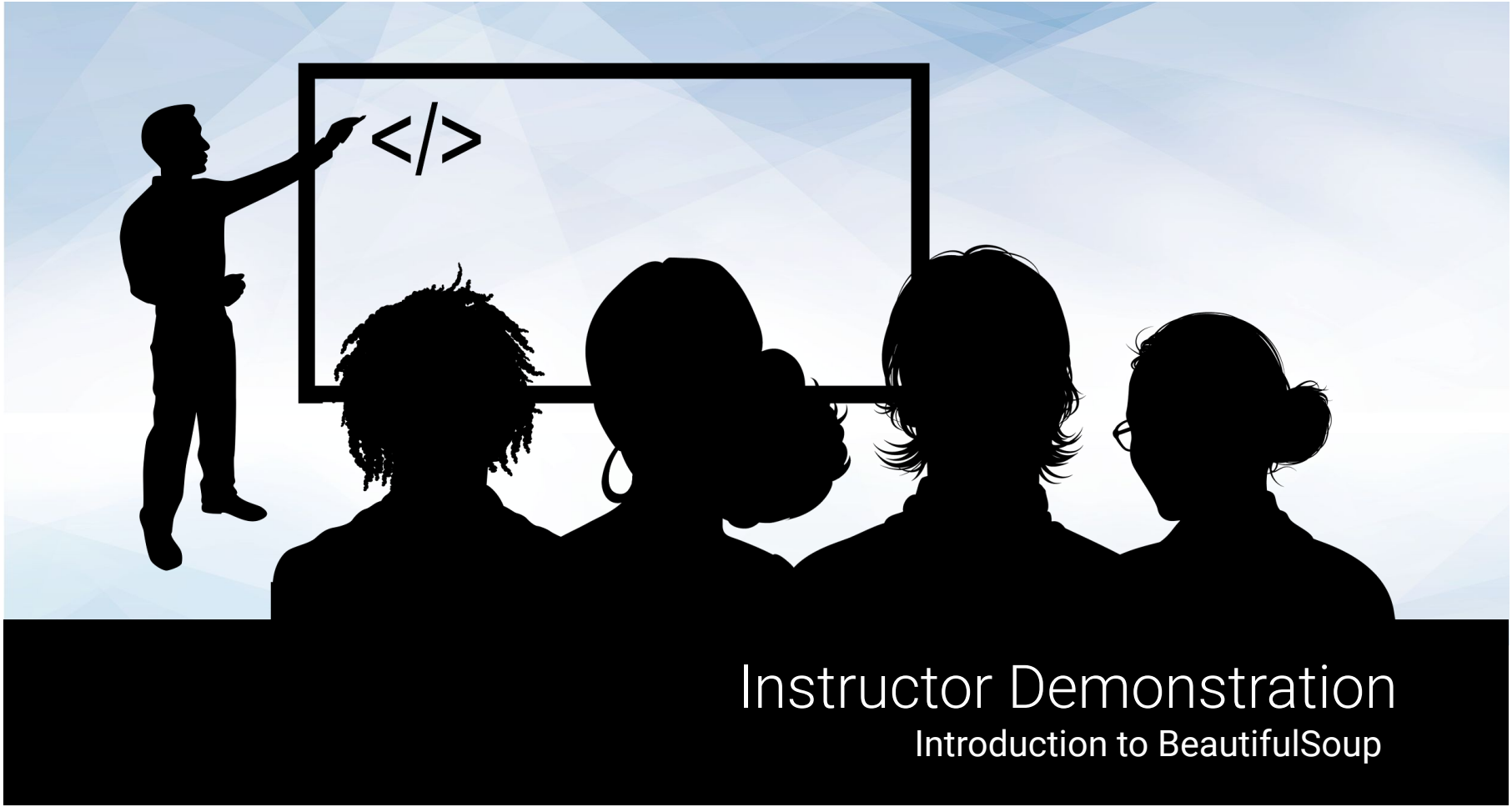
**Hello World!**

Google

1. Visit Grand Canyon
2. Hike the trails
3. Take photos

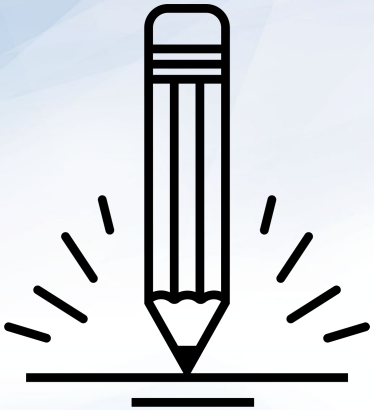
- Bach
- Mozart
- Beethoven
- Adele

# Let’s Review




# Instructor Demonstration

## Introduction to BeautifulSoup



# Activity: From Soup to Nuts

In this activity, you will perform basic HTML scraping with BeautifulSoup

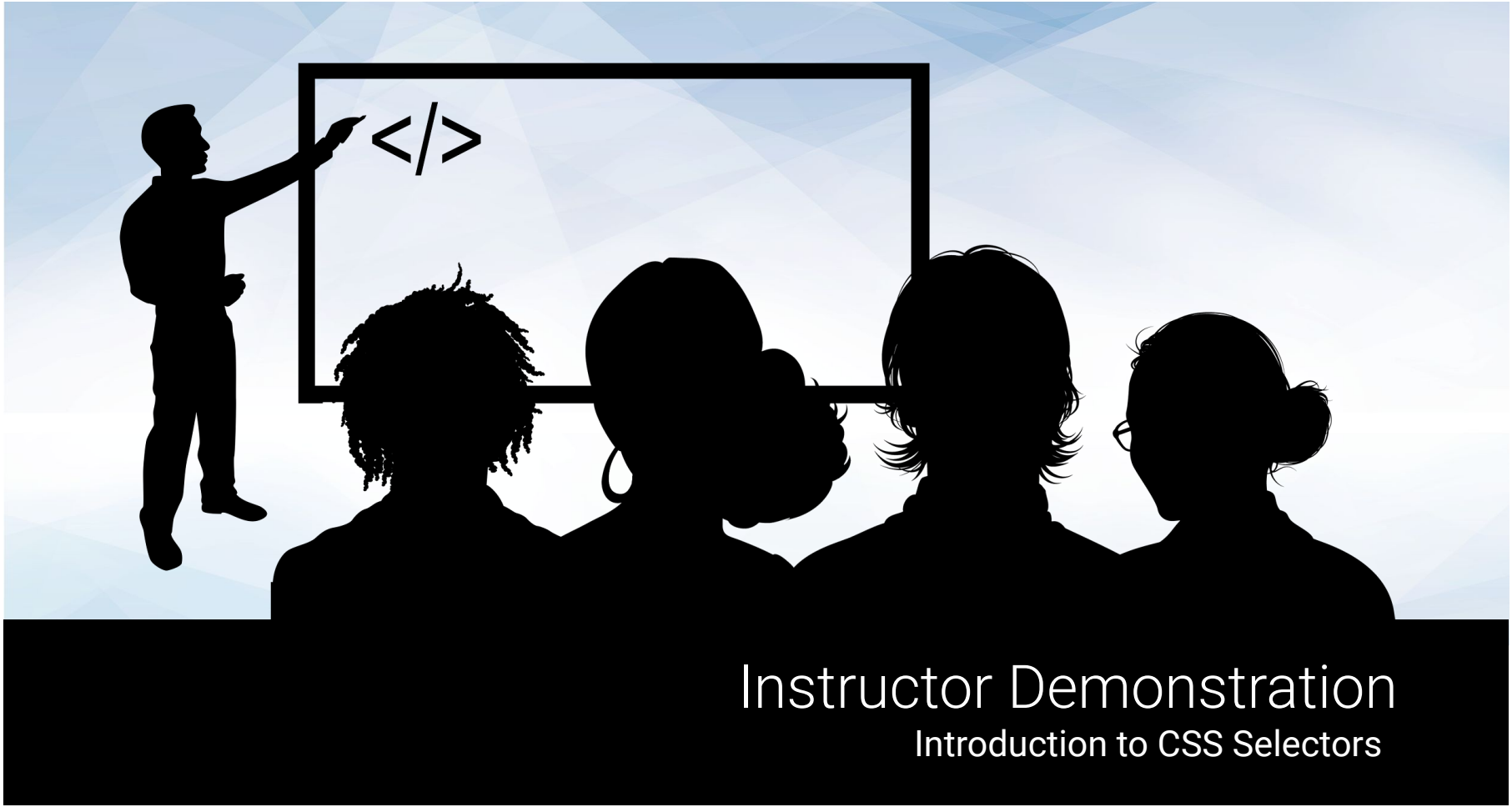
**Suggested Time:**
20 minutes

# From Soup to Nuts Instructions

In this activity, you will use BeautifulSoup to extract the following information from an HTML document:

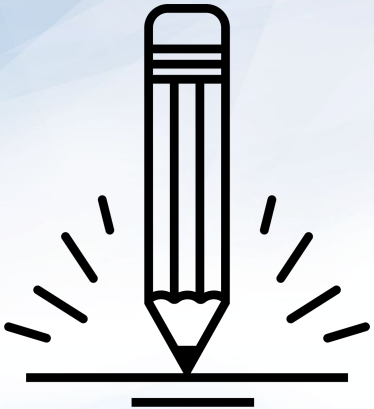
- The <head> element
- The first <h1> element, then its text
- The first <h2> element's text
- The first anchor (<a>), then its href attribute.
- The first <ul> element, and its first list item (<li>), as well as the list item's text

# Let's Review




# Instructor Demonstration

## Introduction to CSS Selectors



# Activity: CSS My List

In this activity, you will create your first web page using HTML.

**Suggested Time:**
20 minutes

# My First HTML Instructions

In this activity, you will create a web page containing a header, an image, a link, and lists. Below is an example:

**Hello World!**

Google

1. Visit Grand Canyon
2. Hike the trails
3. Take photos

- Bach
- Mozart
- Beethoven
- Adele

# Let's Review