Aman Patel

October 19, 2018

UROC @ SICE

Week 3: Extracting Information from URLs

This week I was tasked with extracting information from Wikipedia API URLs using Python. To begin, I stored the URL in a variable named “url.” I then imported urllib.request and applied urllib.request.urlopen to “url,” saving this to a new variable called “page.” Applying .read() to “page” and printing showed the contents of the page in JSON format. One character had to be removed from the beginning of page.read() to allow it to be parsed. I then imported json and used json.loads on page.read(), saving this to a new variable called “j.” I then checked the type of j and found it to be a dictionary. Through each tier of the dictionary, I checked the keys and selected one.