

PROJECT CODE:

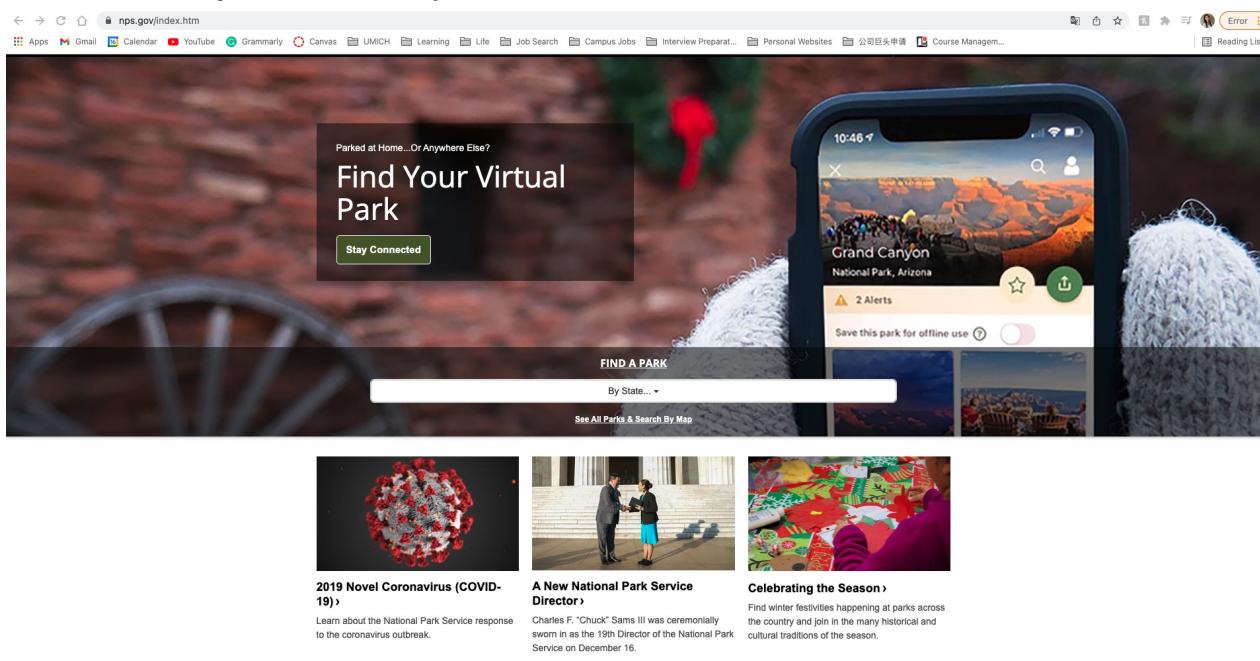
- Link to github repo: https://github.com/yaoqi617/SI507_FinalProject
- The short version of the README file has been attached in the github repo. More detailed explanations are contained in this document.
- Packages required for the program: BeautifulSoup / Requests / Pandas / Plotly.express / Numpy

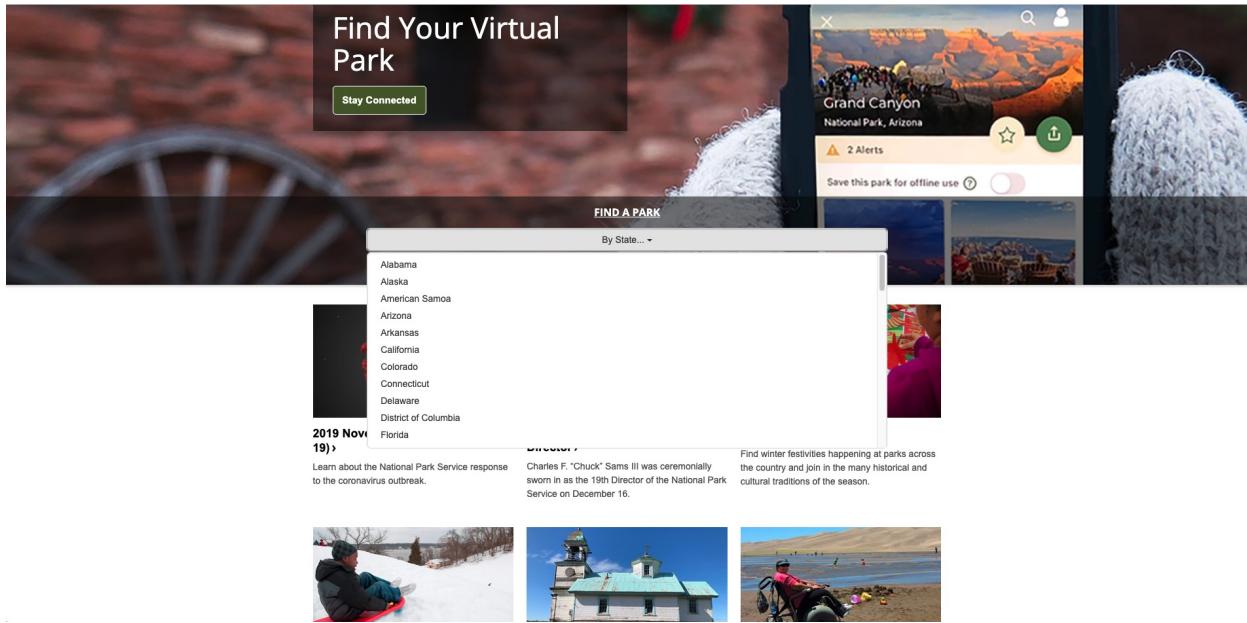
Data Source

Origin	Format	How I accessed the data	Summary of data
https://www.nps.gov/index.htm	JSON	Used beautifulsoup to scrape the web information. Note: Multiple pages scraped and Caching is used.	<p># records available: 400+</p> <p># records retrieved/cached: each state will have a different number of national parks. So the number of cached records depends on the number of national parks in that state.</p> <p>Description of the data: Each record is information about a specific national park. Some important fields include the type of national park(such as national monument, national historic trail, etc.), park name, a short description of the park,</p>

			street, city, state, zip, latitude, longitude of the park
https://gist.github.com/erichurst/7882666	CSV	Copied from GitHub Repo	<p># records available: 33,145</p> <p>Description of the data: this CSV file has contained three fields: zip, latitude, longitude</p>

For this project, I have scraped multiple pages in the nps.gov. As you can see from the parks.json file which has contained national parks in California(scraped from: <https://www.nps.gov/state/ca/index.htm>), Ohio(scraped from: <https://www.nps.gov/state/oh/index.htm>) and Michigan(scraped from: <https://www.nps.gov/state/mi/index.htm>). If you want to check the national parks in CA, OH or MI, then you don't have to scrape online again because I have submitted the json files and you could use the cached files.





nps.gov/state/al/index.htm

NPS.gov / Find A Park / Alabama

Alabama

PARKS

NATIONAL MONUMENT
Birmingham Civil Rights
AL

In 1963, images of snarling police dogs unleashed against non-violent protesters and of children being sprayed with high-pressure hoses appeared in print and television news around the world. These dramatic scenes of violent police aggression against civil rights protesters in Birmingham, Alabama were vivid examples of segregation and racial injustice in America.

NATIONAL MONUMENT
Freedom Riders
AL

In 1961, a small interracial band of "Freedom Riders" challenged discriminatory laws requiring separation of the races in interstate travel. They were attacked by white segregationists, who firebombed the bus. Images of the attack appeared in hundreds of newspapers, shocking the American public and spurring the Federal Government to issue regulations banning segregation in interstate travel.

NATIONAL MILITARY PARK
Horseshoe Bend
Davidson, AL

On 27 March 1814, Major General Andrew Jackson's army of 3,300 men attacked Chief Menawa's 1,000 Red Stick Creek warriors fortified in a horseshoe shaped bend of the Tallapoosa River. Over 800 Red Sticks died that day. The battle ended the Creek War, resulted in a land cession of 23,000,000 acres to the United States and created a national hero of Andrew Jackson.

BY THE NUMBERS

Alerts & Conditions »	Basic Information »
Calendar »	Maps »

9 National Parks	1,279,753 Visitors to National Parks
\$78,200,000 Economic Benefit from National Park Tourism	\$71,385,656 of Land & Water Conservation Fund Appropriated for Projects (since 1965)
\$10,807,739 in Historic Preservation Grants (since 1969)	34 Certified Local Governments
17,118 Hours Donated by Volunteers	1 National Heritage Area
3 National Trails Administered by NPS	1,329 National Register of Historic Places Listings
38 National Historic Landmarks	7 National Natural Landmarks
1,254 Places Recorded by Heritage Documentation Programs	433,301 Objects in National Park Museum Collections
6,124 Archaeological Sites in National Parks	5 Teaching with Historic Places Lessons Plans

Data Structure:

- Describing data structure below:

The data structure used in this project is binary search trees. BST is used to sort and store the geographical distance from the user's zip to each national park. I

I created a BST class with the insert method and in order retrieval. The Insert method allows me to loop through all the national parks in a state, calculate the distance between each park to the user's location, then insert the distance to the BST; the In-order method allows me to retrieve distance from the smallest to the largest distance.

- A python file that constructs the binary tree: `binary_tree.py`
 - A sample tree has been stored in: `treeFile.txt`
 - A python file that reads the json of the tree: `load_tree.py`
 - Screenshot of the data and data structures:

JSON	Raw Data	Headers
Save	Copy	Collapse All
Expand All (slow)	Filter JSON	
▶ STATES_PARKS_HTTPS://WWW.NPS.GOV/STATE/MI/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/ISRO/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/KEWE/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/NOCO/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/PIRO/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/RIBA/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/SLBE/INDEX.HTM.JSON:	{...}	
▶ STATES_PARKS_HTTPS://WWW.NPS.GOV/STATE/CA/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/ALCA/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/CABR/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/CALI/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/CAHO/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/CECH/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/CHHS/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/DEVA/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/DEPO/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/EUON/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/FOPO/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/GOGA/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/JOMU/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/JOTR/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/JUBA/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/LAVO/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/LABE/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/MANZ/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/MOJA/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/MUNO/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/OLSP/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/PIMV/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/PORE/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/POEX/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/POCH/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/PRSF/INDEX.HTM.JSON:	{...}	
▶ PARKINFO_ADDRESS_HTTPS://WWW.NPS.GOV/REDW/INDEX.HTM.JSON:	{...}	

Zip_Lat_Long

ZIP	LAT	LNG
601	18.180555	-66.749961
602	18.361945	-67.175597
603	18.455183	-67.119887
606	18.158345	-66.932911
610	18.295366	-67.125135
612	18.402253	-66.711397
616	18.420412	-66.671979
617	18.445147	-66.559696
622	17.991245	-67.153993
623	18.083361	-67.153897
624	18.064919	-66.716683
627	18.4126	-66.863926
631	18.190607	-66.832041
637	18.076713	-66.947389
638	18.295913	-66.515588
641	18.263085	-66.712985
646	18.43315	-66.285875
647	17.963613	-66.947127
650	18.349416	-66.578079
652	18.448452	-66.594127
653	17.985033	-66.886536
656	18.053539	-66.792931
659	18.407226	-66.808999
660	18.134695	-67.116199

Interaction and Presentation Options:

- **High-level summary of the program:** This program allows users to input a state to check the national parks and show those parks on the map.
- **Run the program and follow the instructions that pop up on the screen (command line prompts).** Below are showing four different options users could choose:
 - a. If the user inputs '1', the program will ask the user to input a state abbreviation (example: MI -> Enter). Then the program will output all national parks in that state.
 - b. If the user inputs '2', the program will ask the user to input both zip code and state abbreviation (Example: type in 48104 -> Enter -> OH -> Enter). The program will output all national parks in the state which the user-provided. And the program will automatically sort all the returned national parks by the distance from the park to the user-entered zip code in ascending order. However, one thing to keep in mind is that the distance is geographical distance, not driving distance.
 - c. If the user inputs '3', the program will ask the user to input a state abbreviation (example: OH -> Enter). The program will then plot all national parks in the state user provided on an interactive map, with park names and park coordinates shown on the map.
 - d. If the user inputs '4', the program will exit.

Demo Link:

<https://drive.google.com/file/d/1oMMc0CmLhKcJeeCG9h3ksnEwD2gOh9HP/view?usp=sharing>