Activities

We are going to use the website of the Milwaukee Police Department. **It is very important that you do not make a large number of requests to this site, especially not from outside the US. You can and should access once and then save the source for all development purposes.**

(1) Go to `https://itmdapps.milwaukee.gov/MPDCallData/` and save the file as mpd.html.

(2) Open up the file using a browser such as chrome. Because we are accessing a file, the javascript scripts and the style sheets are not available, however, we can still make out the contents.

(3) Open the file in Beautiful Soup and print out a prettified version of it. The output is large, but usable. Identify the elements that have call data in them.

(4) As you will have seen, there is only one table in the whole file and a row of the table starts with the <tr> tag. Use the find_all method to print out all tags <tr> in the police file.

(5) As you see, the first row is not useful as it contains some navigation information. Modify your function to return all <tr> tags but the first one.

(6) Within each <tr> tag, we are interested into the individual cells. Use a find_all method applied on each <tr> tag that extracts the <td> tags and print them out. For each line in the table, you should get

```
<td style="border: 1px solid black; border-collapse: collapse;">200111499</td>
<td style="border: 1px solid black; border-collapse: collapse;">01/11/2020 04:39:29 PM</td>
<td style="border: 1px solid black; border-collapse: collapse;">2500-BLK W KILBOURN AV,MKE</
td>
<td style="border: 1px solid black; border-collapse: collapse; text-align: center;">3</td>
<td style="border: 1px solid black; border-collapse: collapse;">WELFARE CITIZEN</td>
<td style="border: 1px solid black; border-collapse: collapse;">Service in Progress</td>
```

(7) For future processing, we want to only get the contents. We can do that by using the contents method, which will return a list with exactly one element in our case. For each row in the table, we want to create a dictionary with keys 'Id', 'datetime', 'address', 'district', 'description', 'status'. The datetime needs to be converted into a Python datetime, which you can do importing parse from dateutil.parser.