In [1]:
```
#updated module 4 lectures and assignment

#Let's try BeautifulSoup library which is another solution for processing data from HTM
#reganizing HTML tags and extracting data between tags
from bs4 import BeautifulSoup

#BeautifulSoup usually works with request package hand-in-hand. Requests will open a we
#for processing.
import requests
import math

#The following page contains the system vulnerabilities reported by NATIONAL VULNERABIL
#Let's take the content of the page out for processing
#https://nvd.nist.gov/vuln/search

# convert string to time
# reference: https://docs.python.org/2/library/datetime.html
from datetime import datetime as dt
```

In [3]:
```
# Let's observe the url formation and create a query

#Ask user to type in a search term
#Reference : https://www.w3schools.com/python/python_user_input.asp
#Reference: https://docs.python.org/2/library/datetime.html


##################################################
##################################################
###### Complete the code here ####################

#Ask user to input a search term, e.g. "splunk"
search_term = input(f"Enter a Search Term: ")

#Ask user to type in a minimum severity , e.g. "7.4"
min_severity = float(input(f"Type in a minimum severity value (1-10): "))

#Ask user to type in a start date in a predefined format, e.g. "10-02-2017"
start_date = input(f"Type in a start publish date (MM-DD-YYYY): ")

#Ask user to type in an end date in a predefined format, e.g. "12-31-2018"
end_date = input(f"Type in a start publish date (MM-DD-YYYY): ")

#Don't forget to convert the input strings to the right date types that will be used
#by the rest of the program.



##################################################
##################################################
```

```
Enter a Search Term: splunk
Type in a minimum severity value (1-10): 7.5
Type in a start publish date (MM-DD-YYYY): 10-21-2015
Type in a start publish date (MM-DD-YYYY): 12-20-2019
```

In [4]:
```
#Formatting date inputs
format_str = '%m-%d-%Y'
start_date = dt.strptime(start_date, format_str)
end_date = dt.strptime(end_date, format_str)
```

In [5]:
```
#let's explore the result based on the query
url = 'https://nvd.nist.gov/vuln/search/results?form_type=Basic&results_type=overview&q

# Request content from web page
response = requests.get(url)
content = response.content

soup = BeautifulSoup(content, 'lxml')


# # #Observe the content of soup
# print(soup)
```

In [6]:
```
# Use a browser to open or download the page
# I used firebug, a FireFox/chrome plug in to analyze the page tag structure
# locate the division where the table is located in


#find the total number of the result
total = soup.find('strong', {"data-testid": "vuln-matching-records-count"})

#show the number of results
#calculate the number of pages of the result. The default page layout is 20 results per
####################################################
####################################################
###### Complete the code here ####################

#remove the comma: reference https://www.w3schools.com/python/ref_string_replace.asp

total = total.text
pages= math.ceil(int(total) / 20)
print(f"The search returned {total} results. Use the following criteria to refine your
print(f"There are {pages} pages.")
```

```
The search returned 75 results. Use the following criteria to refine your search.
There are 4 pages.
```

In [8]:
```
#the next step is to extract data out from the query
#and store the data in variables or save them as a datafile or in a database
#We will store them in variables for now


#create lists to store retrieved data

#vulnerability IDs
vul_IDs =[]

#vulnerability summaries
summaries=[]
```

```
#severity levels
severities = []

#publish dates
publish_dates = []

#the urls of individual vulnerbility description pages
#we don't see them from the survey page yet
urls = []
```

In [9]:

```
# go throught the pages and populate the lists

for page in range(pages):

    ###################################################
    ###################################################
    ###### Complete the code here ####################

    table = soup.find("table", {'class': 'table table-striped table-hover'})
    url = 'https://nvd.nist.gov/vuln/search/results?form_type=Basic&results_type=overvi
    response = requests.get(url)
    content = response.content
    soup = BeautifulSoup(content, 'lxml')


    rows = table.findAll('tr')

    print(f"We are on Page # {page}")
    print(url)

    #in each row
    for tr in rows[1:]: #from 2nd row
        #find table header
        table_head = tr.find('th')

        #append row id to the id array
        vul_IDs_ = table_head.find(text = True)
        vul_IDs.append(vul_IDs_)

        #append the severity score to the severity array
        sev = tr.find("td", {'nowrap': 'nowrap'})
        sevLevel = sev.findAll('a')
        if sevLevel == []:
            severities.append(10.0)
        else:
            sevNumber = sevLevel[0].find(text = True)
            splitSev = sevNumber.split(" ")
            severities.append(float(splitSev[0]))

        ##get publish dates into the dates array
        publish_ = tr.find('span')
        publish_dates_ = publish_.find(text = True)
        publish_dates.append(publish_dates_)

        #get urls into the url array
        url = table_head.find('a')
        urls.append("https://nvd.nist.gov" + url['href'])
```

```
        # some results may not contain a severity score
        # They show "not available" as the output
        # Find a way to bypass it. For example:
        # If the score is not available, assign the score to be 10
        # otherwise, take the real score

    ################################################
    ################################################
```

```
We are on Page # 0
https://nvd.nist.gov/vuln/search/results?form_type=Basic&results_type=overview&query=spl
unk&search_type=all&startIndex=0
We are on Page # 1
https://nvd.nist.gov/vuln/search/results?form_type=Basic&results_type=overview&query=spl
unk&search_type=all&startIndex=20
We are on Page # 2
https://nvd.nist.gov/vuln/search/results?form_type=Basic&results_type=overview&query=spl
unk&search_type=all&startIndex=40
We are on Page # 3
https://nvd.nist.gov/vuln/search/results?form_type=Basic&results_type=overview&query=spl
unk&search_type=all&startIndex=60
```

In [10]:

```
################################################
################################################
###### Complete the code here ####################

#Show the result in the format as in the sample answer below

#number of records count
count = 1

#print records for the rows that match the criteria
for vul in vul_IDs:
    convert_date = dt.strptime(publish_dates[vul_IDs.index(vul)], "%B %d, %Y; %H:%M:%S
    published_converted = dt.isoformat(convert_date)
    published_converted = published_converted[0:10] + ' '+published_converted[11:19]
    if (dt.isoformat(start_date) <= published_converted <= dt.isoformat(end_date)):
        if severities[vul_IDs.index(vul)] >= min_severity:
            print(f"No. {count}")
            print(f"Vul_Id: {vul}")
            if isinstance(severities[vul_IDs.index(vul)],float):
                print(f"Severity: {str(severities[vul_IDs.index(vul)])}")
            else:
                print(f"Severity: {str('10')}")
            print(f"Publish Date: {str(published_converted)}")
            print(f"For more information, visit {urls[vul_IDs.index(vul)]}")
            print('-------------------------------------------------------------------
            count += 1
        else:
            continue
    else:
        continue

################################################
################################################
```

```
No. 1
Vul_Id: CVE-2019-10390
Severity: 8.8
Publish Date: 2019-08-28 12:15:11
```

```
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2019-10390
--------------------------------------------------------------------------------
No. 2
Vul_Id: CVE-2019-5729
Severity: 8.1
Publish Date: 2019-03-21 12:01:05
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2019-5729
--------------------------------------------------------------------------------
No. 3
Vul_Id: CVE-2019-0029
Severity: 7.8
Publish Date: 2019-01-15 04:29:01
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2019-0029
--------------------------------------------------------------------------------
No. 4
Vul_Id: CVE-2018-7432
Severity: 7.5
Publish Date: 2018-10-23 05:31:39
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2018-7432
--------------------------------------------------------------------------------
No. 5
Vul_Id: CVE-2018-7429
Severity: 7.5
Publish Date: 2018-10-23 05:31:39
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2018-7429
--------------------------------------------------------------------------------
No. 6
Vul_Id: CVE-2017-17067
Severity: 9.8
Publish Date: 2017-11-29 09:29:04
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2017-17067
--------------------------------------------------------------------------------
No. 7
Vul_Id: CVE-2015-4017
Severity: 7.5
Publish Date: 2017-08-25 02:29:00
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2015-4017
--------------------------------------------------------------------------------
No. 8
Vul_Id: CVE-2019-10390
Severity: 8.8
Publish Date: 2019-08-28 12:15:11
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2019-10390
--------------------------------------------------------------------------------
No. 9
Vul_Id: CVE-2019-5729
Severity: 8.1
Publish Date: 2019-03-21 12:01:05
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2019-5729
--------------------------------------------------------------------------------
No. 10
Vul_Id: CVE-2019-0029
Severity: 7.8
Publish Date: 2019-01-15 04:29:01
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2019-0029
--------------------------------------------------------------------------------
No. 11
Vul_Id: CVE-2018-7432
Severity: 7.5
Publish Date: 2018-10-23 05:31:39
```

```
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2018-7432
--------------------------------------------------------------------------------
No. 12
Vul_Id: CVE-2018-7429
Severity: 7.5
Publish Date: 2018-10-23 05:31:39
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2018-7429
--------------------------------------------------------------------------------
No. 13
Vul_Id: CVE-2017-17067
Severity: 9.8
Publish Date: 2017-11-29 09:29:04
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2017-17067
--------------------------------------------------------------------------------
No. 14
Vul_Id: CVE-2015-4017
Severity: 7.5
Publish Date: 2017-08-25 02:29:00
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2015-4017
--------------------------------------------------------------------------------
No. 15
Vul_Id: CVE-2017-7565
Severity: 8.8
Publish Date: 2017-04-06 11:59:00
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2017-7565
--------------------------------------------------------------------------------
No. 16
Vul_Id: CVE-2016-10126
Severity: 9.8
Publish Date: 2017-01-10 06:59:00
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2016-10126
--------------------------------------------------------------------------------
No. 17
Vul_Id: CVE-2016-6304
Severity: 7.5
Publish Date: 2016-09-26 03:59:00
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2016-6304
--------------------------------------------------------------------------------
No. 18
Vul_Id: CVE-2016-6302
Severity: 7.5
Publish Date: 2016-09-16 01:59:12
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2016-6302
--------------------------------------------------------------------------------
No. 19
Vul_Id: CVE-2016-2182
Severity: 9.8
Publish Date: 2016-09-16 01:59:02
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2016-2182
--------------------------------------------------------------------------------
No. 20
Vul_Id: CVE-2016-2181
Severity: 7.5
Publish Date: 2016-09-16 01:59:01
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2016-2181
--------------------------------------------------------------------------------
No. 21
Vul_Id: CVE-2016-2179
Severity: 7.5
Publish Date: 2016-09-16 01:59:00
```

```
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2016-2179
--------------------------------------------------------------------------------
No. 22
Vul_Id: CVE-2016-5636
Severity: 9.8
Publish Date: 2016-09-02 10:59:06
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2016-5636
--------------------------------------------------------------------------------
No. 23
Vul_Id: CVE-2016-2183
Severity: 7.5
Publish Date: 2016-08-31 08:59:00
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2016-2183
--------------------------------------------------------------------------------
No. 24
Vul_Id: CVE-2016-2180
Severity: 7.5
Publish Date: 2016-07-31 10:59:11
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2016-2180
--------------------------------------------------------------------------------
No. 25
Vul_Id: CVE-2016-2177
Severity: 9.8
Publish Date: 2016-06-19 09:59:02
For more information, visit https://nvd.nist.gov/vuln/detail/CVE-2016-2177
--------------------------------------------------------------------------------
```

In [ ]: