# Homework 3

New Attempt

**Due** Nov 27 by 11:59pm **Points** 100 **Submitting** a file upload

Your datasets module has two XML files: bill.xml and usfs.html (obtained from <a href="https://data.fs.usda.gov/geodata/edw/datasets.php">https://data.fs.usda.gov/geodata/edw/datasets.php</a>]. For the following problems go to <a href="http://xpather.com/">http://xpather.com/</a>] <a href="http://xpather.com/">http://xpather.com/</a>] and copy the XML file into the textbox on the left. Specify XML queries corresponding to the following questions. Then send a screenshot of the highlight results and submit it as the response to the specific question. Please write the question with the screenshot.

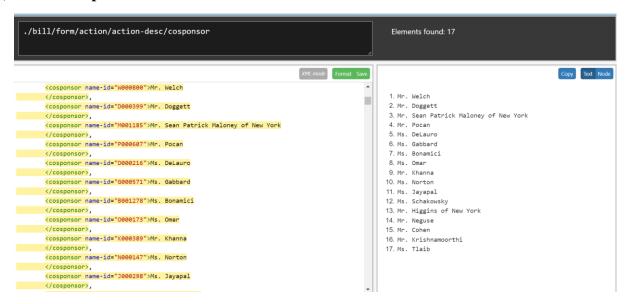
### bill.xml

- (a) List the cosponsors of the bill
- (b) Find the header elements that has the string "negotiation"
- (c) Find the text of the headers of subparagraphs whose text subelements have the string "rebate"
- (d) Find the headers of paragraphs where the text under the paragraphs contains the string "drug" and the header itself has the string "Negotiation" usfs.html
- (a) From the table that has the header "Feature Classes" and "Abstract", get the abstract of the row containing "**Activity Silviculture Timber Stand Improvement**" in the first column.
- (b) For the same row, get the URL of the anchor called "metadata" in the second column
- (c) From the left panel, extract the list of items "under" the header "Special Purpose FS Data".

# **Assignment-03**

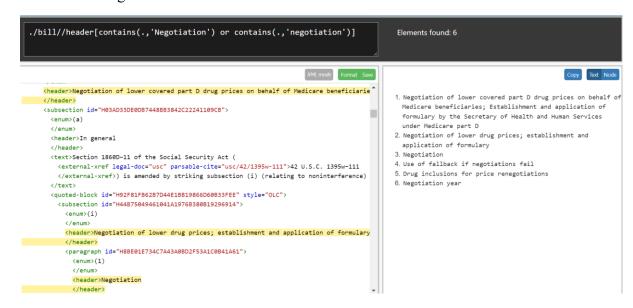
## Bill.xml

## (a) List the cosponsors of the bill

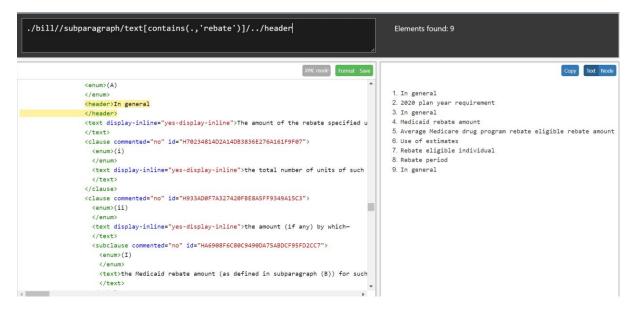


# (b) Find the header elements that has the string "negotiation"

Assumption: Taking both the cases of 'Negotiation' and 'negotiation' into consideration. Also assuming that the string negotiation can be a subpart of another word like 'renegotiation'.



(c) Find the text of the headers of subparagraphs whose text subelements have the string "rebate"



(d) Find the headers of paragraphs where the text under the paragraphs contains the string "drug" and the header itself has the string "Negotiation"

Assumption: Considering any text under the paragraphs that contain "drug" (like even in subparagraph, etc)



### **Usfs.html**

- (a) From the table that has the header "Feature Classes" and "Abstract", get the abstract of the row containing "Activity Silviculture Timber Stand Improvement" in the first column.
  - → //table[tbody[tr[1][td[.="Feature Classes"] and td[.="Abstract"]]]]/tbody/tr[td[1][contains(.,"Activity Silviculture Timber Stand Improvement")]]/td[last()]'

```
nodes = driver.find_elements(By.XPATH, '//table[tbody[tr[1][td[.="Feature Classes"] and td[.="Abstract"]]]]/tbody/tr[td[1][cc print(nodes)
get_result(nodes)

{

| (<selenium.webdriver.remote.webelement.WebElement (session="2f0cbb5bb753456ee4c694c8b7ec7c20", element="04d7f98a-cccd-4e41-b
130-8f244807067b")>]
Tag: td
Text: The SilvTSI (Silviculture Timber Stand Improvement) feature class represents activities associated with the following
performance measure: Forest Vegetation Improved (Release, Weeding, and Cleaning, Precommercial Thinning, Pruning and Fertili
zation). The Activities data set portrays the areas where activities are accomplished... [see more]
```

or

```
nodes = driver.find_elements(By.XPATH, '//*[@id="skipheader"]/section/table/tbody/tr[3]/td[3]')
print(nodes)
get_result(nodes)

[<selenium.webdriver.remote.webelement.WebElement (session="a7043d62750444d8712c71414e67ffc9", element="e341e036-865a-421c-b 4d8-51918122de48")>]
Tag: td
Text: The SilvTSI (Silviculture Timber Stand Improvement) feature class represents activities associated with the following performance measure: Forest Vegetation Improved (Release, Weeding, and Cleaning, Precommercial Thinning, Pruning and Fertili zation). The Activities data set portrays the areas where activities are accomplished... [see more]
parent dataset: Activities
```

- (b) For the same row, get the URL of the anchor called "metadata" in the second column
  - → //table[tbody[tr[1][td[.="Feature Classes"] and td[.="Abstract"]]]]/tbody/tr[td[1][contains(.,"Activity Silviculture Timber Stand Improvement")]]/td[2]/a[.="metadata"]

```
nodes = driver.find_elements(By.XPATH, '//*[@id="skipheader"]/section/table/tbody/tr[3]/td[2]/a[1]')
print(nodes)
get_result(nodes)

[<selenium.webdriver.remote.webelement.WebElement (session="a7043d62750444d8712c71414e67ffc9", element="89896617-c6d0-4e0a-8
977-82aff2e5aa98")>]
Tag: a
Text: metadata
```

# (c) From the left panel, extract the list of items "under" the header "Special Purpose FS Data".

→ //div[@id="leftcol"]//\*[preceding-sibling::h2[. = "Special Purpose FS Data"]]

```
nodes = driver.find_elements(By.XPATH, '//div[@id="leftcol"]//*[preceding-sibling::h2[. = "Special Purpose FS Data"]]')
print(nodes)
get_result(nodes)|

[<selenium.webdriver.remote.webelement.WebElement (session="2f0cbb5bb753456ee4c694c8b7ec7c20", element="6544ae70-072a-4332-b
0ac-77a868603b2f")>]
Tag: ul
Text: State and Private Forestry
Wilderness & WSR & WSA
Research and Development
Puerto Rico Gap Analysis Project
```

### $\mathbf{or}$

```
nodes = driver.find_elements(By.XPATH, '//*[@id="righttop"]/ul[4]')
print(nodes)
get_result(nodes)

[<selenium.webdriver.remote.webelement.WebElement (session="a7043d62750444d8712c71414e67ffc9", element="22f5e36f-c92a-492c-8
32d-d6f46e7d6e9b")>]
Tag: ul
Text: State and Private Forestry
Wilderness & WSR & WSR
Research and Development
Puerto Rico Gap Analysis Project
```