



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 <https://data.fs.usda.gov/geodata/edw/datasets.php>  [\(https://data.fs.usda.gov/geodata/edw/datasets.php\)](https://data.fs.usda.gov/geodata/edw/datasets.php)). For the following problems go to <http://xpather.com/>  [\(http://xpather.com/\)](http://xpather.com/) 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

./bill/form/action/action-desc/cosponsor

Elements found: 17

XML mode Format Save

Copy Text Node

```
<cosponsor name-id="W000800">Mr. Welch</cosponsor>
<cosponsor name-id="D000399">Mr. Doggett</cosponsor>
<cosponsor name-id="M001185">Mr. Sean Patrick Maloney of New York</cosponsor>
<cosponsor name-id="P000607">Mr. Pocan</cosponsor>
<cosponsor name-id="D000216">Ms. DeLauro</cosponsor>
<cosponsor name-id="G000571">Ms. Gabbard</cosponsor>
<cosponsor name-id="B001278">Ms. Bonamici</cosponsor>
<cosponsor name-id="O000173">Ms. Omar</cosponsor>
<cosponsor name-id="K000389">Mr. Khanna</cosponsor>
<cosponsor name-id="N000147">Ms. Norton</cosponsor>
<cosponsor name-id="J000298">Ms. Jayapal</cosponsor>
```

1. Mr. Welch
2. Mr. Doggett
3. Mr. Sean Patrick Maloney of New York
4. Mr. Pocan
5. Ms. DeLauro
6. Ms. Gabbard
7. Ms. Bonamici
8. Ms. Omar
9. Mr. Khanna
10. Ms. Norton
11. Ms. Jayapal
12. Ms. Schakowsky
13. Mr. Higgins of New York
14. Mr. Neguse
15. Mr. Cohen
16. Mr. Krishnamoorthi
17. Ms. Tlaib

(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'.

./bill//header[contains(., 'Negotiation') or contains(., 'negotiation')]

Elements found: 6

XML mode Format Save

Copy Text Node

```
<header>Negotiation of lower covered part D drug prices on behalf of Medicare beneficiaries</header>
<subsection id="H03AD33DE00B74488B3842C22241109CB">
  <enum>(a)</enum>
  <header>In general</header>
  <text>Section 1860D-11 of the Social Security Act (
    <external-xref legal-doc="usc" parsable-cite="usc/42/1395w-111">42 U.S.C. 1395w-111
    </external-xref>) is amended by striking subsection (i) (relating to noninterference)
  </text>
  <quoted-block id="H92F81F86287D44E18B19866D60833FEE" style="OLC">
    <subsection id="H44875049461041A19768380819296914">
      <enum>(i)</enum>
      <header>Negotiation of lower drug prices; establishment and application of formulary</header>
      <paragraph id="H88E01E734C7A43A08D2F53A1C0B41A61">
        <enum>(1)</enum>
        <header>Negotiation</header>
      </paragraph>
    </subsection>
  </quoted-block>
```

1. Negotiation of lower covered part D drug prices on behalf of Medicare beneficiaries; Establishment and application of formulary by the Secretary of Health and Human Services under Medicare part D
2. Negotiation of lower drug prices; establishment and application of formulary
3. Negotiation
4. Use of fallback if negotiations fail
5. Drug inclusions for price renegotiations
6. Negotiation year

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

XML query: `./bill//subparagraph/text[contains(., 'rebate')]/../header` Elements found: 9

XML mode | Format | Save | Copy | Text | Node

```

<enum>(A)
</enum>
<header>In general
</header>
<text display-inline="yes-display-inline">The amount of the rebate specified u
</text>
<clause commented="no" id="H70234814D2A14DB3836E276A161F9F07">
  <enum>(i)
  </enum>
  <text display-inline="yes-display-inline">the total number of units of such
  </text>
</clause>
<clause commented="no" id="H933AD0F7A327420FBE8A5FF9349A15C3">
  <enum>(ii)
  </enum>
  <text display-inline="yes-display-inline">the amount (if any) by which-
  </text>
  <subclause commented="no" id="HA6908F6C80C9490DA75ABDCF95FD2CC7">
    <enum>(I)
    </enum>
    <text>the Medicaid rebate amount (as defined in subparagraph (B)) for such
    </text>
  </subclause>
</clause>

```

1. In general
2. 2020 plan year requirement
3. In general
4. Medicaid rebate amount
5. Average Medicare drug program rebate eligible rebate amount
6. Use of estimates
7. Rebate eligible individual
8. Rebate period
9. In general

(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)

XML query: `./bill//paragraph//text[contains(., 'drug')]/ancestor::paragraph/header[contains(., 'Negotiation')]` Elements found: 1

XML mode | Format | Save | Copy | Text | Node

```

</header>
<paragraph id="H88E01E734C7A43A08D2F53A1C0841A61">
  <enum>(1)
  </enum>
  <header>Negotiation
  </header>
  <subparagraph id="H2C7CC84E5B9443019E887FBD2E83B002">
    <enum>(A)
    </enum>
    <header>In general
    </header>
    <text>Notwithstanding any other provision of law, subject to subparagraph (B), t
    </text>
    <clause id="HC04E34347CA54184891B273ABF6A64DF">
      <enum>(i)
      </enum>
      <text>during the negotiation year (as defined in such subparagraph) for such p
      </text>
    </clause>
    <clause id="HED14A1433A7A4C1593F7D4E3C7DF4495">
      <enum>(ii)
      </enum>
      <text>complete such negotiations not less than 30 days before the first day of
    </text>
  </subparagraph>

```

1. 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.

➔ `//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][contains(., "Activity Silviculture Timber Stand Improvement")]]/td[last()]')
print(nodes)
get_result(nodes)
```

[<selenium.webdriver.remote.webelement.WebElement (session="2f0cbb5bb753456ee4c694c8b7ec7c20", element="04d7f98a-cccd-4e41-b130-8f244807067b")>]

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 Fertilization). The Activities data set portrays the areas where activities are accomplished... [see more]

parent dataset: Activities

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-b4d8-51918122de48")>]

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 Fertilization). 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, '//table[tbody[tr[1]][td[.="Feature Classes"] and td[.="Abstract"]]]/tbody/tr[td[1][contains(., "Activity Silviculture Timber Stand Improvement")]]/td[2]/a[.="metadata"]')
print(nodes)
get_result(nodes)
```

[<selenium.webdriver.remote.webelement.WebElement (session="2f0cbb5bb753456ee4c694c8b7ec7c20", element="b7632aa4-910a-48e6-b77c-f7631aa3944e")>]

Tag: a

Text: metadata

or

```
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-8977-82aff2e5aa98")>]
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-b0ac-77a868603b2f")>]
Tag: ul
Text: State and Private Forestry
Wilderness & WSR & WSA
Research and Development
Puerto Rico Gap Analysis Project
```

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-832d-d6f46e7d6e9b")>]
Tag: ul
Text: State and Private Forestry
Wilderness & WSR & WSA
Research and Development
Puerto Rico Gap Analysis Project
```