**Goal**

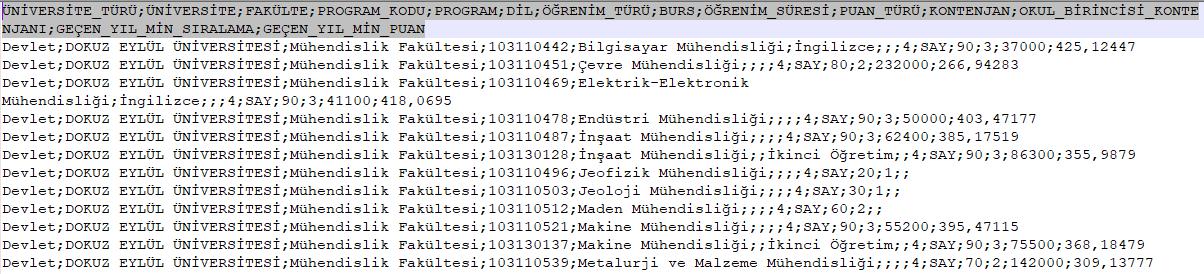
In this assignment, you are asked to develop a tool to convert files among CSV, XML and JSON formats in Python. In addition, the tool must be able to validate your XML file via the corresponding XSD file.

**Implementation Details & Requirements**

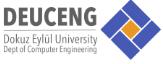
* The tool takes command line arguments according to the formats you want to convert between them. A typical command line usage is as follows:

python <filename> <input file> <output file/xsd file> <type>

* The first argument, <filename> is the python file for conversion operations, <input file> refers to the source filename which will be converted and the third one, <output file>, refer to the target filename, or XSD file. The last argument, <type>, defines conversion type (1=CSV to XML, 2=XML to CSV, 3=XML to JSON, 4=JSON to XML, 5=CSV to JSON, 6=JSON to CSV,7=XML validates with XSD)
* The sample command line usage converting from XML to JSON as follows:: python student\_id.py test.xml test.json 3
* For XML operations, you should use xml.etree.ElementTree and for JSON operations you should use json library of python. XSD validation should be implemented using etree library under lxml.Any other libraries for file types other than these are not allowed.
* We give a sample CSV to test your program.



1/4

CME 2202 Data Organization and Management *2019-2020 Spring*

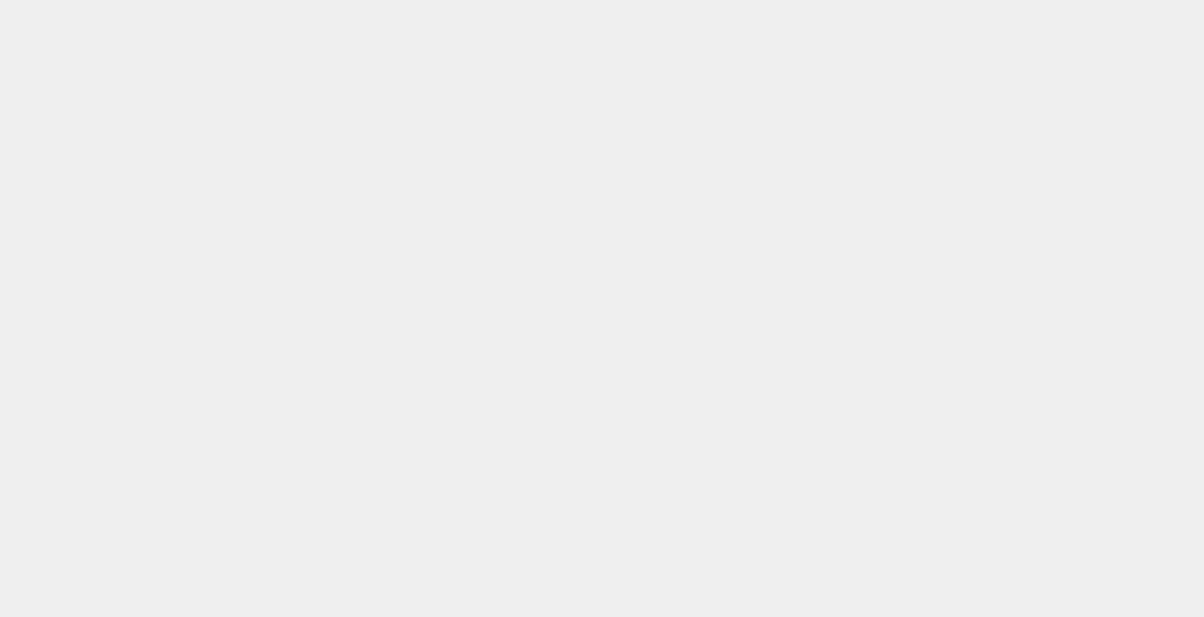
In this file, the separator character is semicolon (;). It has 14 columns and their details are given below:

|  |  |
| --- | --- |
| **Header Title** | **Header Description** |
|  |  |
| ÜNİVERSİTE\_TÜRÜ | Displays the type of university and may have 2 values: “*Devlet”*, “*Vakıf”* |
|  |  |
| ÜNİVERSİTE | Displays the name of university. |
|  |  |
| FAKÜLTE | Displays the name of faculty. |
|  |  |
| PROGRAM\_KODU | Displays the department ID and has 9 digits. |
|  |  |
| PROGRAM | Displays the name of program. |
|  |  |
| DİL | Displays the language of department and may have 2 values: “*İngilizce”* |
|  | *(en)* or *empty*. “*empty*” value means “*Türkçe*” (tr). |
|  |  |
| ÖĞRENİM\_TÜRÜ | Displays the type of department and may have 2 values: “İkinci Öğretim” |
|  | (*iö*) or *empty*. “*empty*” value means “*Örgün Öğretim*” (*öö*). |
|  |  |
| BURS | If the department gives scholarship, the field shows its percentage and |
|  | may be these values: 25, 50, 100. It may be empty if it doesn’t give. |
|  |  |
| ÖĞRENİM\_SÜRESİ | Displays the education period of department and it must be numeric value. |
|  |  |
| PUAN\_TÜRÜ | Displays the score type of department and may be 4 values: “*SAY”*, |
|  | “*EA*”,”*SÖZ*”,”*DİL*” |
|  |  |
| KONTENJAN | Displays the quota of department and it must be numeric value. |
|  |  |
| OKUL\_BİRİNCİSİ\_KO | Displays the quota that students who graduate as the top of the school |
| NTENJANI | can prefer and it may be numeric value or empty (0). |
|  |  |
| GEÇEN\_YIL\_MİN\_SI | Displays the score of the person who has the minimum success ranking |
| RALAMA | among those who preferred last year and it may be numeric value or |
|  | empty (0). |
|  |  |
| GEÇEN\_YIL\_MİN\_PU | Displays the score of the person who has the minimum score of those who |
| AN | preferred last year and and it must be decimal value or empty (0). |
|  |  |

* XSD file you will prepare for the sample CSV file should include all properties including restrictions.

2/4

* The general XML format should be as follows (JSON file will use the same structure):



<departments>

<university name=”DOKUZ EYLÜL ÜNİVERSİTESİ” uType=”Devlet”> <item id=”103110442” faculty=”Mühendislik Fakültesi”>

<name lang=”en” second=”No”>Bilgisayar Mühendisliği</name>

<period>4</period>

<quota spec=”3”>90</quota>

<field>SAY</field>

<last\_min\_score

order=”37000”>425,12447</last\_min\_score> <grant/>

</item>

<item ...>

…

</item>

</university>

</departments>

**Documentation**

In this assignment, in line documentation is expected, as well as good coding practices such as consistent naming, proper usage of indentation and high readability of code.

3/4

**Grading policy**

* Conversion Processes
  + CSV <-> XML 20%
  + CSV <-> JSON 20%
  + XML <-> JSON 20%
  + XSD validation 20%
* Documentation/coding style 20%

4/4