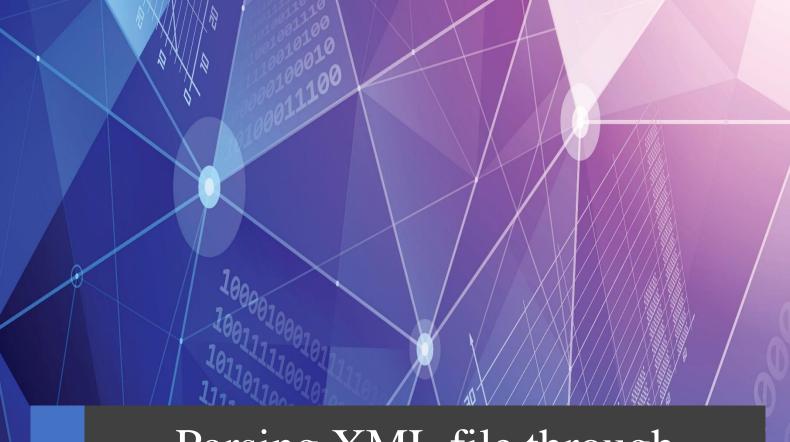# Compiler Construction
Assignment # 03

Parsing XML file through Python

Made by: Instructor: Ma'am Reeda

MUHAMMAD QALB E ALI
200901105 CS-01 B

1/14/23

Code Output:

```
Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.

Try the new cross-platform PowerShell https://aka.ms/pscore6

PS C:\Users\syed_\Desktop\CC>  & 'C:\Users\syed_\AppData\Local\Microsoft\WindowsApps\python3.10.exe' 'c:\Users\syed_\.vscode\extensions\ms-python.python-2022.20.2\py
thonFiles\lib\python\debugpy\adapter/../..\debugpy\launcher' '65353' '--' 'c:\Users\syed_\Desktop\CC\200901105_Assign_03.py'

Book Id: bk101
Author: Gambardella, Matthew
Title: XML Developer's Guide
Genre: Computer
Price: 44.95
Publish Date: 2000-10-01
Desc: An in-depth look at creating applications
     with XML.

Book Id: bk102
Author: Ralls, Kim
Title: Midnight Rain
Genre: Fantasy
Price: 5.95
Publish Date: 2000-12-16
Desc: A former architect battles corporate zombies,
     an evil sorceress, and her own childhood to become queen
     of the world.

Book Id: bk103
Author: Corets, Eva
Title: Maeve Ascendant
Genre: Fantasy
Price: 5.95
Publish Date: 2000-11-17
Desc: After the collapse of a nanotechnology
     society in England, the young survivors lay the
     foundation for a new society.

Book Id: bk104
Author: Corets, Eva
Title: Oberon's Legacy
Genre: Fantasy
Price: 5.95
Publish Date: 2001-03-10
Desc: In post-apocalypse England, the mysterious
     agent known only as Oberon helps to create a new life
     for the inhabitants of London. Sequel to Maeve
     Ascendant.

Book Id: bk105
Author: Corets, Eva
Title: The Sundered Grail
Genre: Fantasy
Price: 5.95
Publish Date: 2001-09-10
Desc: The two daughters of Maeve, half-sisters,
     battle one another for control of England. Sequel to
     Oberon's Legacy.

Book Id: bk106
Author: Randall, Cynthia
Title: Lover Birds
Genre: Romance
Price: 4.95
Publish Date: 2000-09-02
Desc: When Carla meets Paul at an ornithology
     conference, tempers fly as feathers get ruffled.

Book Id: bk107
Author: Thurman, Paula
Title: Splish Splash
Genre: Romance
Price: 4.95
Publish Date: 2000-11-02
Desc: A deep sea diver finds true love twenty
     thousand leagues beneath the sea.

Book Id: bk108
Author: Knorr, Stefan
Title: Creepy Crawlies
Genre: Horror
Price: 4.95
Publish Date: 2000-12-06
Desc: An anthology of horror stories about roaches,
     centipedes, scorpions  and other insects.

Book Id: bk109
Author: Kress, Peter
Title: Paradox Lost
Genre: Science Fiction
Price: 6.95
Publish Date: 2000-11-02
Desc: After an inadvertant trip through a Heisenberg
     Uncertainty Device, James Salway discovers the problems
     of being quantum.
```

```
Book Id: bk110
Author: O'Brien, Tim
Title: Microsoft .NET: The Programming Bible
Genre: Computer
Price: 36.95
Publish Date: 2000-12-09
Desc: Microsoft's .NET initiative is explored in
      detail in this deep programmer's reference.

Book Id: bk111
Author: O'Brien, Tim
Title: MSXML3: A Comprehensive Guide
Genre: Computer
Price: 36.95
Publish Date: 2000-12-01
Desc: The Microsoft MSXML3 parser is covered in
      detail, with attention to XML DOM interfaces, XSLT processing,
      SAX and more.

Book Id: bk112
Author: Galos, Mike
Title: Visual Studio 7: A Comprehensive Guide
Genre: Computer
Price: 49.95
Publish Date: 2001-04-16
Desc: Microsoft Visual Studio 7 is explored in depth,
      looking at how Visual Basic, Visual C++, C#, and ASP+ are
      integrated into a comprehensive development
      environment.
PS C:\Users\syed_\Desktop\CC>
```

Code Explanation:

This code uses Python's xml.etree.ElementTree library to parse an XML file called "compiler.xml". The XML file is loaded into a tree structure, with the root element accessible via the "root" variable.
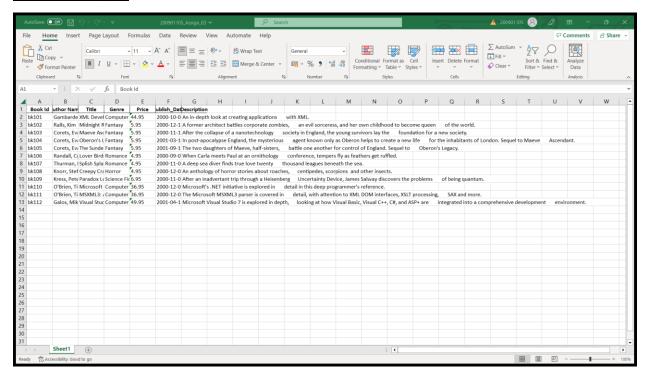
The script then iterates through all the "book" elements within the root element, using the findall() method. For each "book" element, it extracts the information from the child elements of "book" such as 'author','title','genre','price','publish_date','description' using the find() method and the text attribute of the element, and assigns them to variables (book_id, author, title, genre, price, publish_date, description).

The script also prints out the extracted data for each book in the XML file.

Then it appends the extracted data for each book in the list, named "data".

Finally, it creates a Pandas dataframe from the "data" list, and uses the to_excel() method to write the data to an excel file named "200901105_Assign_03.xlsx" with the columns "Book Id", "Author Name", "Title", "Genre", "Price", "Publish_Date", "Description" respectively. The index=False argument is used to prevent the creation of an index column in the excel file.

Excel File Output:

My GitHub: https://github.com/Qalb-E-Ali/IST-Sem-5-Parsing-XML-file-via-Python