МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования

НИЖЕГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ

УНИВЕРСИТЕТ им. Р.Е.АЛЕКСЕЕВА

Институт радиоэлектроники и информационных технологий

Кафедра информатики и систем управления

**Отчет по лабораторной работе №6**

по дисциплине

Управление данными

«Знакомство с XML»

РУКОВОДИТЕЛЬ:

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Балашова Т. И.

(подпись)

СТУДЕНТ:

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Демидова А.С

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(подпись)

17-АС

Работа защищена «\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_

С оценкой \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Нижний Новгород, 2020

**Цель лабораторной работы**:

Изучить работу с языком разметки XML.

**Задание**:

Часть1.

Составить свой корректный XML-документ, описывающий иерархическую базу данных.

Часть 2.

Ознакомится с языком запросов XPath. Написать пару запросов к своему XML документу созданном ранее.

**Выполнение работы:**

**Часть 1**

Описание документа:

**<?xml version=**"1.0" **encoding=**"UTF-8" **standalone=**"yes"**?>**

**<!DOCTYPE serial\_list**

**[**

**<!ELEMENT serial\_list (serial)>**

**<!ELEMENT serial (language | country)>**

**<!ATTLIST serial name CDATA #REQUIRED>**

**<!ATTLIST serial status CDATA #REQUIRED>**

**<!ATTLIST serial genres CDATA #REQUIRED>**

**<!ATTLIST serial season CDATA #REQUIRED>**

**<!ELEMENT country (country\_code | timezone)>**

**<!ELEMENT language (#PCDATA)>**

**<!ELEMENT country\_code (#PCDATA)>**

**<!ELEMENT timezone (#PCDATA)>**

**]>**

Корневой элемент serial\_list, содержащий вложенные элемент serial.  
Атрибутами serial являются: name, status, genre, season, кроме того, он будет содержать вложенные элементы language и country.   
Элемент country так же будет содержать вложенные элементы: country\_name и timezone

Составим XML – документ:

**<?xml version=**"1.0" **encoding=**"UTF-8" **standalone=**"yes"**?>**

**<!DOCTYPE serial\_list**

**[**

**<!ELEMENT serial\_list (serial)>**

**<!ELEMENT serial (language | country)>**

**<!ATTLIST serial name CDATA #REQUIRED>**

**<!ATTLIST serial status CDATA #REQUIRED>**

**<!ATTLIST serial genres CDATA #REQUIRED>**

**<!ATTLIST serial season CDATA #REQUIRED>**

**<!ELEMENT country (country\_code | timezone)>**

**<!ELEMENT language (#PCDATA)>**

**<!ELEMENT country\_code (#PCDATA)>**

**<!ELEMENT timezone (#PCDATA)>**

]>

**<serial\_list>**

**<serial** name**=**"Doctor Who"status**=**"Running"genres**=**"Fantastic, adventure, comedy, drama"season**=**"12"**>**

**<language>**English**</language>**

**<country>**UnitedKingdom

**<country\_code>**GB**</country\_code>**

**<timezone>**Europe/London**</timezone>**

**</country>**

**</serial>**

**<serial** name**=**"StarTrek"status**=**"Ended"genres**=**"Fantastic, action, adventure"season**=**"3"**>**

**<language>**English**</language>**

**<country>**USA

**<country\_code>**Us**</country\_code>**

**<timezone>**Europe/London**</timezone>**

**</country>**

**</serial>**

**<serial** name**=**"GameOfTrones"status**=**"Ended"genres**=**"Fantasy, action, drama, romance, adventure"season**=**"8"**>**

**<language>**English**</language>**

**<country>**USA

**<country\_code>**US**</country\_code>**

**<timezone>**America/New\_York**</timezone>**

**</country>**

**</serial>**

**<serial** name**=**"Supernatural"status**=**"Running"genres**=**"Fantasy, action, drama, romance, adventure"season**=**"15"**>**

**<language>**English**</language>**

**<country>**USA

**<country\_code>**US**</country\_code>**

**<timezone>**America/New\_York**</timezone>**

**</country>**

**</serial>**

**<serial** name**=**"TheBigBangTheory"status**=**"Ended"genres**=**"Romance, comedy"season**=**"12"**>**

**<language>**English**</language>**

**<country>**USA

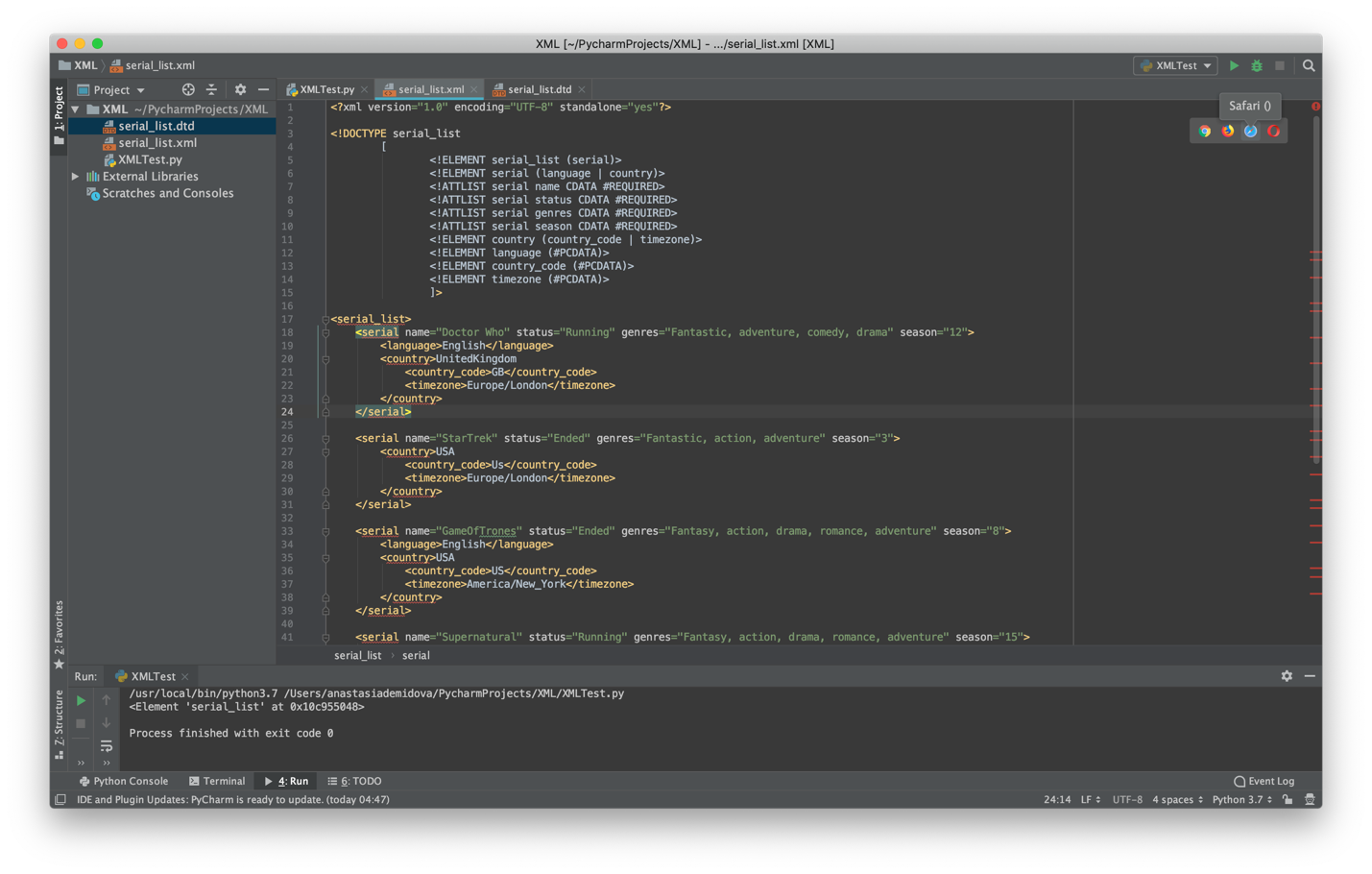
**<country\_code>**US**</country\_code>**

**<timezone>**America/New\_York**</timezone>**

**</country>**

**</serial>**

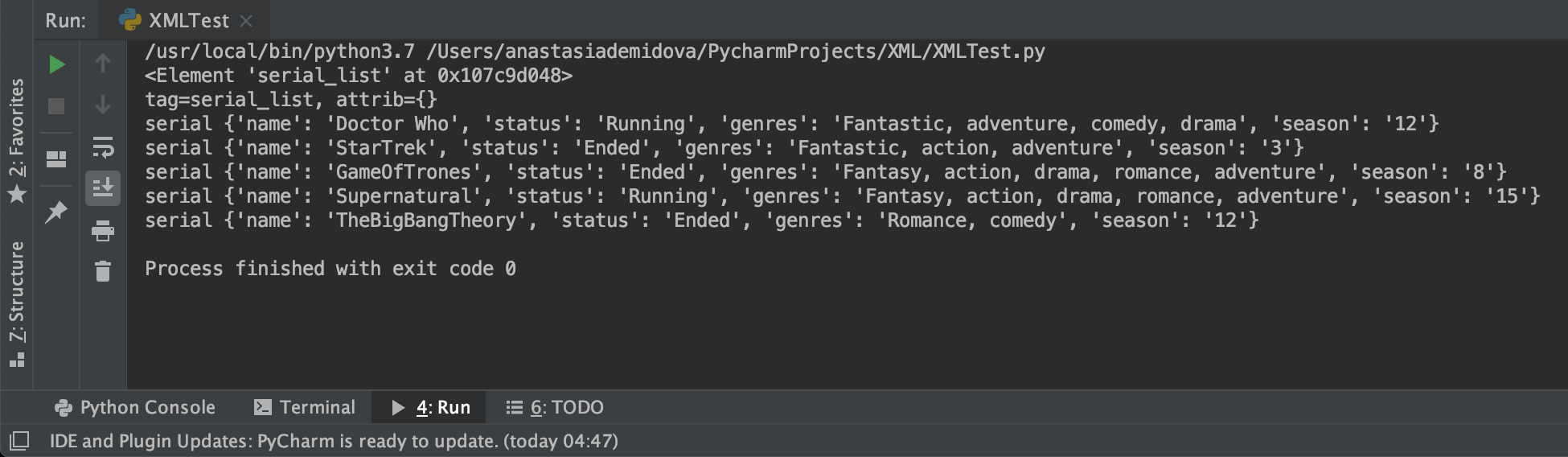
**</serial\_list>**

В PyCharm есть возможность просмотреть составленный XML документ в браузере через localhost:  
  
  
Документ открывается, значит ошибок нет

Часть 2

Для начала распарсим наш XML документ, используем python 3.7 с помощью модулю ElementTree

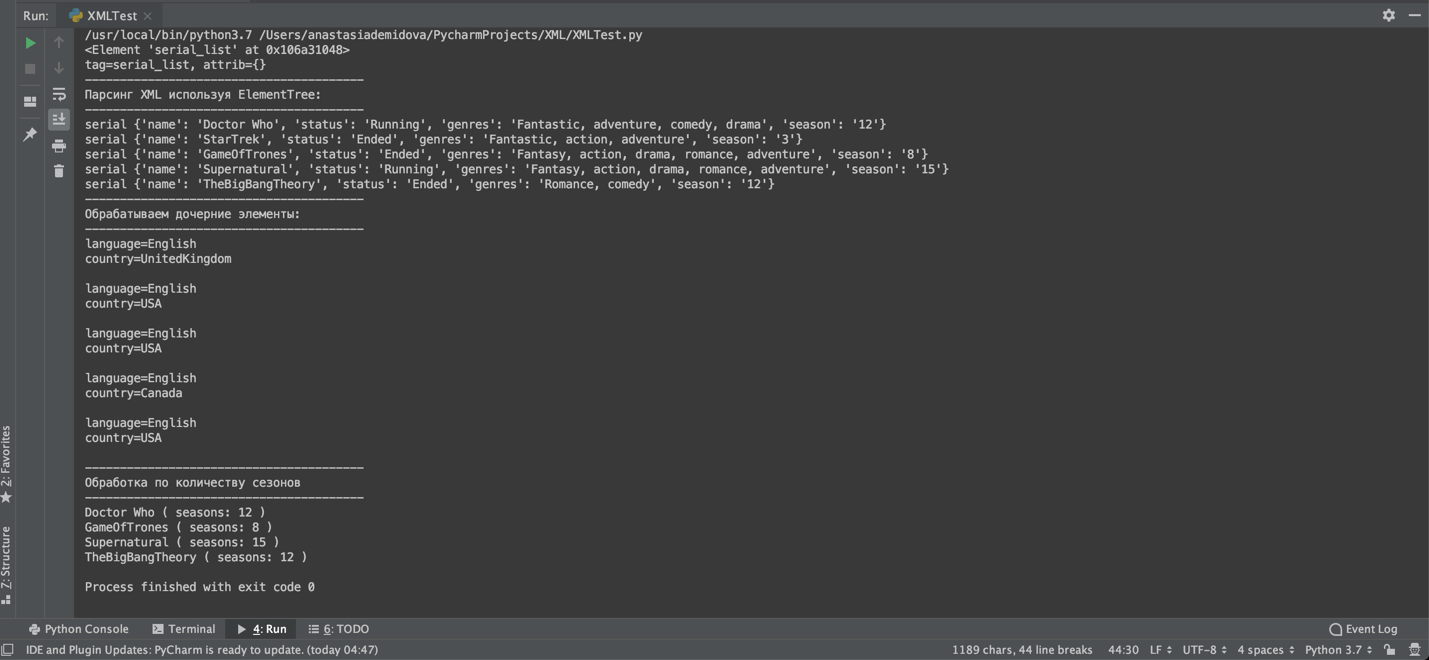
Результат парсинга, все верно



**Код:**

import xml.etree.cElementTree as xml  
  
def parseXML(xml\_file):  
 tree = xml.ElementTree(file=xml\_file)  
 print(tree.getroot())  
 root = tree.getroot()  
 print("tag=%s, attrib=%s" % (root.tag, root.attrib))  
  
 # выводим все данные  
 print("-" \* 40)  
 print("Парсинг XML используя ElementTree: ")  
 print("-" \* 40)  
  
 for child in root:  
 print(child.tag, child.attrib)  
 if child.tag == "appointment":  
 for step\_child in child:  
 print(step\_child.tag)  
  
 # получаем данные используя дочерние элементы.  
 print("-" \* 40)  
 print("Обрабатываем дочерние элементы: ")  
 print("-" \* 40)  
  
 for appointment in root.findall('serial'):  
 for appt\_child in appointment:  
 print("%s=%s" % (appt\_child.tag, appt\_child.text))

# вывод данных, где количество сезонов больше 3  
 print("-" \* 40)  
 print("Обработка по количеству сезонов")  
 print("-" \* 40)  
  
 for child in root.findall('serial'):  
 season = int(child.get('season'))  
 if season > 3:  
 name = child.get('name')  
 print(name, "( seasons:", season, ")")  
  
  
  
  
if \_\_name\_\_ == "\_\_main\_\_":  
 parseXML("serial\_list.xml")

**Результаты работы:  
  
**

**Вывод:**

В процессе выполнения лабораторной работы были изучены основы написания валидных xml-документов, а также получены базовые навыки в работе с языком запросов XPath.