

LABORATORIO 6

XML Y SQL

ALUMNOS:

Ignacio Andrés Castillo Rendon

Anderson Fabian Garcia Nieto

**Laboratorio-Modelos de bases de datos
2024-2**

DOCENTE:

MARIA IRMA DIAZ ROZO

Bogotá D.C-Colombia

PUNTO UNO. Sólo XML. CATÁLOGO DE CURSOS (courses.xml)

El caso es uno de los propuestos en el curso de bases de datos de Stanford.

A. Explorando

1. Explore el archivo usando un navegador y un editor.

¿Cuáles son las ventajas de uno u otro?

RTA: Al usar el navegador para ver el archivo XML, pueden mostrar el contenido del XML de manera estructurada y fácil de leer, con la posibilidad de expandir y contraer nodos, sin embargo es solo de visualización, es decir, que no se va a poder editar el archivo ni validar con un DTD. Mientras que con un editor, en primer lugar, se modifica el contenido del XML directamente, además de que dependiendo del editor (en nuestro caso Visual Studio Code) ofrece características avanzadas como validación, autocompletado, y transformación XSLT.

2. Describa la información que encuentra en el archivo

RTA: Es una base de datos del catálogo de cursos de alguna entidad de Educación Superior en donde se divide por departamentos (dirigidos por algún profesor) las clases de la “Universidad” y cada clase (curso) tiene un código asignado, la capacidad de inscripción, la descripción de la clase, uno o más instructores para cada clase y dependiendo de la clase, tiene o no prerrequisitos.

B. Consultas iniciales

Implemente las siguientes consultas:

Para evaluar use la herramienta que aparece en moodle.

1. ¿De cuáles cursos se tiene información?

[a. Nombre con etiqueta]

-XPath Result-

Copy

Save

```
Element='<Title>Programming Methodology</Title>'
Element='<Title>Programming Abstractions</Title>'
Element='<Title>Computer Organization and Systems</Title>'
Element='<Title>Introduction to Probability for Computer Scientists</Title>'
Element='<Title>From Languages to Information</Title>'
Element='<Title>Compilers</Title>'
Element='<Title>Introduction to Databases</Title>'
Element='<Title>Artificial Intelligence: Principles and Techniques</Title>'
Element='<Title>Structured Probabilistic Models: Principles and Techniques</Title>'
Element='<Title>Machine Learning</Title>'
Element='<Title>Digital Systems I</Title>'
Element='<Title>Digital Systems II</Title>'
Element='<Title>From Languages to Information</Title>'
```

Course_Catalog/Departement/Course/Title

[b. Nombre sin etiqueta]

```

Text='Programming Methodology'
Text='Programming Abstractions'
Text='Computer Organization and Systems'
Text='Introduction to Probability for Computer Scientists'
Text='From Languages to Information'
Text='Compilers'
Text='Introduction to Databases'
Text='Artificial Intelligence: Principles and Techniques'
Text='Structured Probabilistic Models: Principles and Techniques'
Text='Machine Learning'
Text='Digital Systems I'
Text='Digital Systems II'
Text='From Languages to Information'

```

Copy

Save

Course_Catalog/Department/Course/Title/text()

[c. Número con atributo]

```

Attribute='Number=CS106A'
Attribute='Number=CS106B'
Attribute='Number=CS107'
Attribute='Number=CS109'
Attribute='Number=CS124'
Attribute='Number=CS143'
Attribute='Number=CS145'
Attribute='Number=CS221'
Attribute='Number=CS228'
Attribute='Number=CS229'
Attribute='Number=EE108A'
Attribute='Number=EE108B'
Attribute='Number=LING180'

```

Copy

Save

Course_Catalog/Department/Course/@Number

[d. Número sin atributo]

[e. Número total de cursos]

```
Int64='13'
```

Copy

Save

count(Course_Catalog/Department/Course)

2. ¿Cuál es la información del curso Machine Learning?

[a. Toda la información]

```
Element='<Course Enrollment="320" Number="CS229">
  <Title>Machine Learning</Title>
  <Description>A broad introduction to machine learning and statistical pattern recognition.</Description>
  <Instructors>
    <Professor>
      <First_Name>Andrew</First_Name>
      <Last_Name>Ng</Last_Name>
    </Professor>
  </Instructors>
</Course>'
```

Copy Save

Course_Catalog/Department/Course[@Number='CS229']

[b. Todos los atributos]

```
Attribute='Enrollment=320'
Attribute='Number=CS229'
```

Copy Save

Course_Catalog/Department/Course[@Number='CS229']/@*

[c. Todas las etiquetas]

```
Element='<Title>Machine Learning</Title>'
Element='<Description>A broad introduction to machine learning and statistical pattern recognition.</Description>'
Element='<Instructors>
  <Professor>
    <First_Name>Andrew</First_Name>
    <Last_Name>Ng</Last_Name>
  </Professor>
</Instructors>'
```

Copy Save

Course_Catalog/Department/Course[@Number='CS229']/*

3. ¿Cuántos departamentos tenemos?

```
Int64='3'
```

CopySave

count(Course_Catalog/Department)

¿Cuáles son?[Nombre]

```
Text='Computer Science'
Text='Electrical Engineering'
Text='Linguistics'
```

CopySave

Course_Catalog/Department/Title/text()

4. ¿A qué departamento pertenece el curso Machine Learning? [Codigo]

```
Attribute='Code=CS'
```

CopySave

Course_Catalog/Department/Course[@Number='CS229']/../@Code

5. ¿Cuáles cursos son iniciales (no tienen prerequisites)? [Nombre]

```
Text='Programming Methodology'
Text='Artificial Intelligence: Principles and Techniques'
Text='Structured Probabilistic Models: Principles and Techniques'
Text='Machine Learning'
Text='Digital Systems I'
```

CopySave

Course_Catalog/Department/Course[not(Prerequisites)]/Title/text()

6. ¿Cuáles cursos son finales (no son prerequisites de ninguno)? [Nombre]

```
Text='Programming Abstractions'  
Text='Computer Organization and Systems'  
Text='Introduction to Probability for Computer Scientists'  
Text='From Languages to Information'  
Text='Compilers'  
Text='Introduction to Databases'  
Text='Digital Systems II'  
Text='From Languages to Information'
```

Copy

Save

Course_Catalog/Department/Course[(Prerequisites)]/Title/text()

7. ¿Cuáles directores de departamento son profesores de cursos del departamento?[La información del director] (Los profesores se pueden identificar por su apellido)

```
Element='<Professor>  
  <First_Name>Jennifer</First_Name>  
  <Last_Name>Widom</Last_Name>  
</Professor>'
```

Copy

Save

**/Course_Catalog/Department[Chair/Professor/Last_Name =
Course/Instructors/Professor/Last_Name] /Chair/Professor**

8. ¿Cuál es el total de inscripciones de los cursos?

```
Double='3420'
```

Copy

Save

sum(Course_Catalog/Department/Course/@Enrollment)

9. ¿De cuáles cursos no se conoce el número de inscritos? [Nombre]

Text='Digital Systems I'
Text='Digital Systems II'
Copy Save

Course_Catalog/Department/Course[not(@Enrollment)]/Title/text()

10. ¿Cuál curso tiene el mayor número de inscritos? [Nombre]

Element='<Title>Programming Methodology</Title>'
Copy Save

/Course_Catalog/Department/Course[@Enrollment="1070"]/Title

B. Consultas quiz Stanford

1. Return all Title elements (of both departments and courses). (16)

Element='<Title>Computer Science</Title>'
Element='<Title>Programming Methodology</Title>'
Element='<Title>Programming Abstractions</Title>'
Element='<Title>Computer Organization and Systems</Title>'
Element='<Title>Introduction to Probability for Computer Scientists</Title>'
Element='<Title>From Languages to Information</Title>'
Element='<Title>Compilers</Title>'
Element='<Title>Introduction to Databases</Title>'
Element='<Title>Artificial Intelligence: Principles and Techniques</Title>'
Element='<Title>Structured Probabilistic Models: Principles and Techniques</Title>'
Element='<Title>Machine Learning</Title>'
Element='<Title>Electrical Engineering</Title>'
Element='<Title>Digital Systems I</Title>'
Element='<Title>Digital Systems II</Title>'
Element='<Title>Linguistics</Title>'
Element='<Title>From Languages to Information</Title>'
Copy Save

//Title

2. Return last names of all department chairs. (3)

```
Element='<Last_Name>Widom</Last_Name>'
Element='<Last_Name>Horowitz</Last_Name>'
Element='<Last_Name>Levin</Last_Name>'
```

Copy

Save

//Chair//Last_Name

3. Return titles of courses with enrollment greater than 500. (2)

```
Element='<Title>Programming Methodology</Title>'
Element='<Title>Programming Abstractions</Title>'
```

Copy

Save

//Course[@Enrollment > 500]/Title

4. Return titles of departments that have some course that takes "CS106B" as a prerequisite. (2)

```
Element='<Title>Computer Science</Title>'
Element='<Title>Electrical Engineering</Title>'
```

Copy

Save

//Department[Course/Prerequisites/Prereq = 'CS106B']/Title

5. Return last names of all professors or lecturers who use a middle initial. Don't worry about eliminating duplicates. (5)

```
Element='<Last_Name>Cain</Last_Name>'
Element='<Last_Name>Cain</Last_Name>'
Element='<Last_Name>Aiken</Last_Name>'
Element='<Last_Name>Horowitz</Last_Name>'
Element='<Last_Name>Dally</Last_Name>'
```

Copy

Save

//(Professor|Lecturer)[Middle_Initial]/Last_Name

6. Return the title of courses that have a cross-listed course (i.e., that have "Cross-listed" in their description). (2)

```
Element='<Title>From Languages to Information</Title>'
Element='<Title>From Languages to Information</Title>'
```

CopySave

//Course[contains(Description, 'Cross-listed')]/Title

7. Return the enrollment of all courses in the CS department. (10)

```
Attribute='Enrollment=1070'
Attribute='Enrollment=620'
Attribute='Enrollment=500'
Attribute='Enrollment=280'
Attribute='Enrollment=60'
Attribute='Enrollment=90'
Attribute='Enrollment=130'
Attribute='Enrollment=180'
Attribute='Enrollment=110'
Attribute='Enrollment=320'
```

CopySave

//Department[@Code = 'CS']/Course/@Enrollment

8. Return last names of instructors teaching at least one course that has "system" in its description and enrollment greater than 100. (2)

```
Element='<Last_Name>Zelenski</Last_Name>'
Element='<Last_Name>Widom</Last_Name>'
```

CopySave

//Course[contains(Description, 'system') and @Enrollment > 100]/Instructors/*/Last_Name

9. Return the title of the course with more than 2 instructors (*) (1)

```
Element='<Title>Programming Methodology</Title>'
```

CopySave

```
//Course[count(Instructors/*) > 2]/Title
```

10. Return course numbers of courses that have the same title as some other course. (Hint: You might want to use the "preceding" and "following" navigation axes for this query, which were not covered in the video or our demo script; they match any preceding or following node, not just siblings.) (2)

```
Attribute='Number=CS124'  
Attribute='Number=LING180'
```

Copy Save

```
//Course[Title = preceding::Course/Title or Title = following::Course/Title]/@Number
```

11. Return the course numbers of courses that have no lecturers as instructors. (*) (10)

```
Attribute='Number=CS109'  
Attribute='Number=CS124'  
Attribute='Number=CS143'  
Attribute='Number=CS145'  
Attribute='Number=CS221'  
Attribute='Number=CS228'  
Attribute='Number=CS229'  
Attribute='Number=EE108A'  
Attribute='Number=EE108B'  
Attribute='Number=LING180'
```

Copy Save

```
//Course[not(Instructors/Lecturer)]/@Number
```

12. Return titles of courses taught by the chair of a department. For this question, you may assume that all professors have distinct last names. (1)

```
Element='<Title>Introduction to Databases</Title>'
```

Copy Save

```
//Course[Instructors/Professor/Last_Name=//Chair/Professor/Last_Name]/Title
```

13. Return titles of courses taught by a professor with the last name "Ng" and by a professor with the last name "Thrun". (*) (1)

```
Element='<Title>Artificial Intelligence: Principles and Techniques</Title>'
Element='<Title>Machine Learning</Title>'
```

Copy

Save

```
//Course[Instructors/Professor/Last_Name='Ng' or
Instructors/Professor/Last_Name='Thrun']/Title
```

14. Return course numbers of courses that have a course taught by Eric Roberts as a prerequisite. (4)

```
Attribute='Number=CS106B'
Attribute='Number=CS107'
Attribute='Number=CS109'
Attribute='Number=EE108B'
```

Copy

Save

```
//Course[Prerequisites/Prereq = //Course[Instructors/Professor[First_Name='Eric'
and Last_Name='Roberts']]/@Number]/@Number
```

15. Return last names of instructors teaching at least one course that has system in its description and enrollment greater than 100.(2)

```
Element='<Last_Name>Zelenski</Last_Name>'
Element='<Last_Name>Widom</Last_Name>'
```

Copy

Save

```
//Course[contains(Description, 'system') and @Enrollment >
100]/Instructors/*/Last_Name
```

C. Consultas propias

Propongan e implementen cinco consultas propias

- 1. Retornar los profesores del course 'Programming Methodology'**

```
String='Jerry Cain'  
String='Eric Roberts'  
String='Mehran Sahami'
```

Copy

Save

```
//Course[Title = 'Programming Methodology']/Instructors/*[First_Name and  
Last_Name]/concat(First_Name, ' ', Last_Name)
```

2. Retornar todas las clases que tiene Eric Roberts

```
String='Programming Methodology - Eric Roberts'  
String='Programming Abstractions - Eric Roberts'
```

Copy

Save

```
//Course[Instructors/Professor[First_Name='Eric' and  
Last_Name='Roberts']]/concat(Title, ' - ', Instructors/Professor[First_Name='Eric' and  
Last_Name='Roberts']/First_Name, ' ', Instructors/Professor[First_Name='Eric' and  
Last_Name='Roberts']/Last_Name)
```

3. Retornar el course number de los cursos en el departamento Linguistics

```
Attribute='Number=LING180'
```

Copy

Save

```
//Department[@Code='LING']/Course/@Number
```

4. Retornar la cantidad de courses por cada departamento que hay

```
String='Computer Science: 10'  
String='Electrical Engineering: 2'  
String='Linguistics: 1'
```

Copy

Save

```
concat(//Department[@Code='CS']/Title, ': ', count(//Department[@Code='CS']/Course)),  
concat(//Department[@Code='EE']/Title, ': ', count(//Department[@Code='EE']/Course)),  
concat(//Department[@Code='LING']/Title, ': ',  
count(//Department[@Code='LING']/Course))
```

5. Retornar los profesores que son del departamento de Electrical Engineering

```

Element='<Professor>
  <First_Name>Mark</First_Name>
  <Middle_Initial>A.</Middle_Initial>
  <Last_Name>Horowitz</Last_Name>
</Professor>'
Element='<Professor>
  <First_Name>Subhasish</First_Name>
  <Last_Name>Mitra</Last_Name>
</Professor>'
Element='<Professor>
  <First_Name>William</First_Name>
  <Middle_Initial>J.</Middle_Initial>
  <Last_Name>Dally</Last_Name>
</Professor>'
Element='<Professor>
  <First_Name>Oyekunle</First_Name>
  <Last_Name>Olukotun</Last_Name>
</Professor>'

```

Copy Save

//Department[@Code='EE']//Professor

D. Esquema

Proponga un DTD para estos datos.

<!ELEMENT Course_Catalog (Department+)>

<!ELEMENT Department (Title, Chair, Course+)>

<!ATTLIST Department Code CDATA #REQUIRED>

<!ELEMENT Title (#PCDATA)>

<!ELEMENT Chair (Professor)>

<!ELEMENT Course (Title, Description?, Instructors, Prerequisites?)>

<!ATTLIST Course Number CDATA #REQUIRED>

<!ATTLIST Course Enrollment CDATA #IMPLIED>

<!ELEMENT Description (#PCDATA)>

<!ELEMENT Instructors (Lecturer*, Professor*)>

<!ELEMENT Prerequisites (Prereq*)>

<!ELEMENT Prereq (#PCDATA)>

<!ELEMENT Lecturer (First_Name, Middle_Initial?, Last_Name)>

<!ELEMENT Professor (First_Name, Middle_Initial?, Last_Name)>

<!ELEMENT First_Name (#PCDATA)>

<!ELEMENT Middle_Initial (#PCDATA)>

<!ELEMENT Last_Name (#PCDATA)>

NOTA: El validador brindado por la universidad botaba un tipo de error en donde marcaba la primera letra de cada línea.

Para verificar use la herramienta que aparece en moodle.

E. Nuevos datos

Incluya en el archivo la información de tres cursos que no existan en el archivo. Verifique que cumple el esquema definido anteriormente.

RTA: Esta implementación está adjunta en el archivo txt del .zip

RETROSPECTIVA

1. ¿Cuál fue el tiempo total invertido en el laboratorio por cada uno de ustedes? (Horas/Hombre)

Ignacio: Más o menos me demoré unas 4-5 horas para la solución de este laboratorio. En realidad era algo sencillo.

Anderson: Más o menos entre unas 5-6 horas, debido a que me gusto la parte de XML y quería lograr entenderla lo más que se pudiera

2. ¿Cuál es el estado actual del laboratorio? ¿Por qué?

RTA: El estado del laboratorio es finalizado, porque se dividió muy bien las partes de cada uno, y cada integrante cumplió con su entrega

3. ¿Cuál consideran fue el mayor logro? ¿Por qué?

RTA: Lograr sacar el laboratorio en la última semana de clases

4. ¿Cuál consideran que fue el mayor problema técnico? ¿Qué hicieron para resolverlo?

RTA: El mayor problema técnico fue el momento de usar el validador de DTD ya que no nos aceptaba el código. No pudimos resolverlo de esa manera, pero si se ejecuto correctamente en sql oracle, por lo cual dedujimos que el código servía

5. ¿Qué hicieron bien como equipo? ¿Qué se comprometen a hacer para mejorar los resultados?

RTA: Como equipo siento que a lo largo del semestre ambos integrantes tuvieron una gran comunicación, apoyo y un poco de ambición de querer hacer más. Tal vez para un futuro como POOB, si quedamos juntos, nos comprometemos a organizarnos mejor, para poder enviar los laboratorios con tiempo de anticipación