

Práctica: Mapeado de Esquemas

Práctica

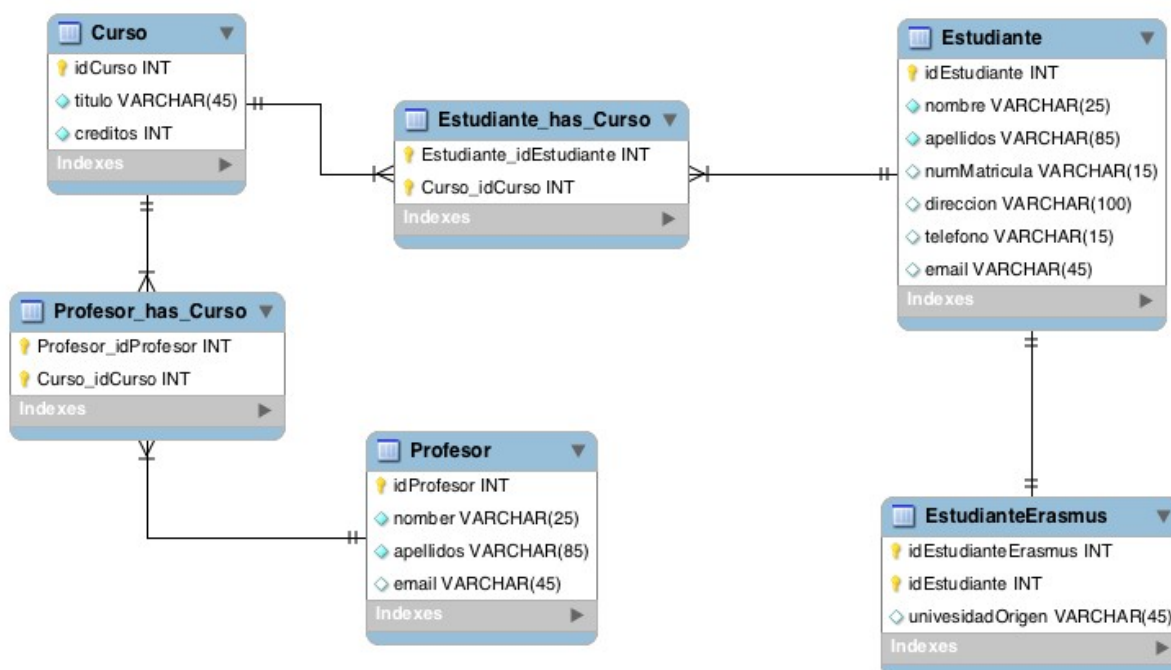
- Fecha de entrega: **15 de enero** hasta final del día; en el caso de que se entregue tarde se restará un punto de la nota de la práctica por día o fracción.
- Formato: Jupyter Notebook listo para ejecutar. Debe incluir tanto el código como explicaciones de las decisiones tomadas en celdas de comentarios.

Problema

Se tienen dos fuentes heterogéneas de datos con información sobre estudiantes internacionales (Erasmus) en dos universidades (UC y UoE). Se trata de crear un esquema global que contenga toda la información posible sobre dichos estudiantes para ambas fuentes.

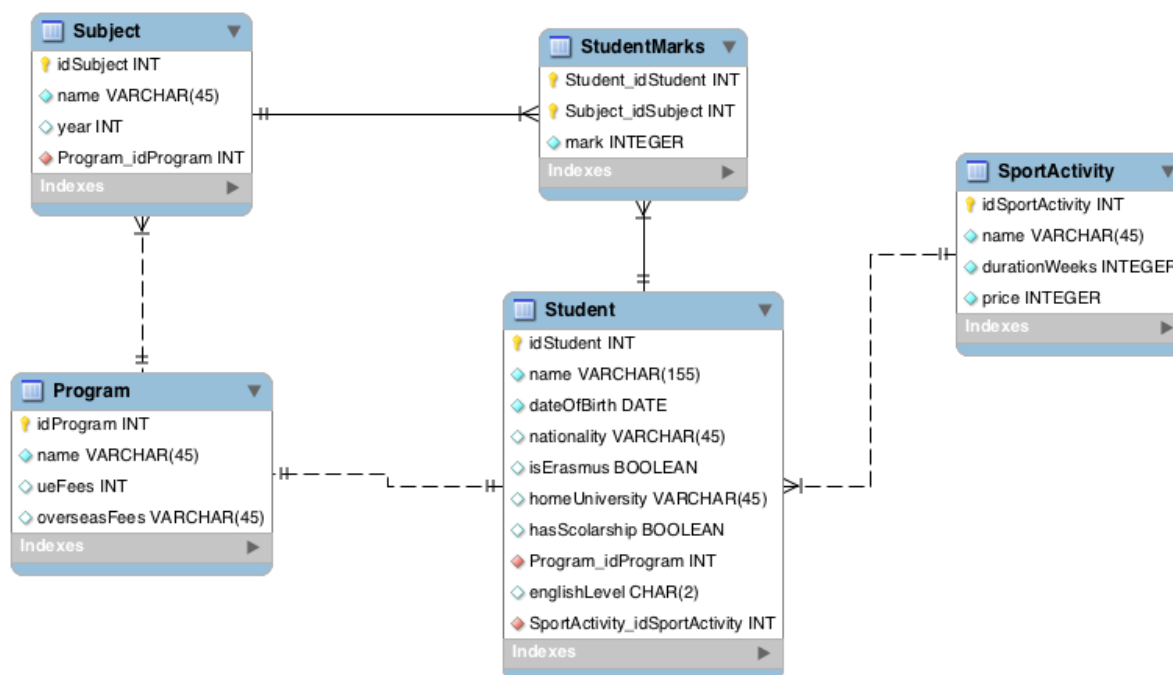
A continuación mostramos los esquemas a utilizar:

Base de datos UC



Base de datos UoE

Ciclo de vida de los datos 2023-2024



Tareas

1. Creación de las bases de datos UC y UoE siguiendo los esquemas especificados. Implementación preferentemente en sqlite3.
2. Inserción en dichas bases de datos de datos generados artificialmente, con al menos 100 estudiantes por base de datos y datos complementarios necesarios para cada esquema. Se sugiere usar Faker (visto en clase) para esta tarea.
3. Creación de un esquema objetivo e implementación en una base de datos que servirá para integrar los datos de las dos bases de datos originales (de nuevo preferentemente en sqlite3).
4. Finalmente, se procederá a extraer, transformar y cargar (ETL) los datos de las base de datos originales en la base de datos integrada.

Todo el código necesario para ejecutar el proceso se incluirá en el Jupyter notebook a entregar, y se discutirán las decisiones de diseño en celdas de comentarios.

Consejos y sugerencias

- El Jupyter Notebook debe de ser ejecutable de principio a fin sin necesidad de estar haciendo ajustes y correcciones. Comprobadlo antes de enviar el *entregable*.
- Si usáis paquetes no estándar (aparte de sqlite3 y Faker), bien instaladlos en el código, o por lo menos incluid una nota o comentario.
- Es una buena idea incluir al final del código alguna comprobación de que se ha hecho lo esperado, esto puede incluir consultas a la base de datos final.
- Durante el desarrollo os puede ser de utilidad incluir una instrucción al principio que borre las bases de datos que creáis para no tener que hacerlo a mano cada vez que ejecutáis.
- Los esquemas entidad-relación fueron creados usando MySQL Workbench, por si os sirve de ayuda.
 - Los atributos con relleno azul son “not null”.

Ciclo de vida de los datos 2023-2024

- A continuación los símbolos de la cardinalidad de las relaciones.



One



Many



One (and only one)



Zero or one



One or many



Zero or many