



IIC2115 – Programación como Herramienta para la Ingeniería (II/2017)

Tarea 3 Parte A JOIN B

Objetivos

- Aplicar conocimientos para la correcta modelación de una base de datos relacional.
- Realizar consultas a la base de datos construida, usando SQL.
- Utilizar Python para realizar consultas y modificaciones a una base de datos.

Entrega

- **Lenguaje a utilizar:** Python 3.6.
- **Lugar:** GitHub.
- **Fecha:** 29 de octubre.
- **Hora:** 23:59.
- **Desarrollo Individual**

Recomendaciones

Lee la tarea completa antes de comenzar a trabajar. Fíjate en los sustantivos y verbos relacionados con el problema. Al terminar, escribe en la parte final de tu informe, un comentario, donde detalles tus opiniones y percepción de la tarea. Estos comentarios son útiles para mejorar el curso durante el mismo semestre, por lo que se sugiere realizarlos a conciencia. No habrá penalizaciones en casos donde el comentario sea negativo.

Introducción

Como cuerpo de ayudantes, nos quedamos sin ideas para tareas y actividades. Hans no nos paga lo suficiente para pensar en esto y solo nos dijo que estaban viendo SQL. Por esa razón, deberán implementar un programa para interactuar con una base de datos de cine del siglo XX. Lamentablemente, Bastian juntó toda la base de datos mientras hacía PUSH y Laika se encontraba trabajando cuando ocurrió aquel horrible accidente, así que no pudo arreglar el error de Bastian a tiempo. Es aquí donde entras tú para reparar tal desastre.

Descripción

En esta tarea, tendrás que modelar un diagrama de relaciones que permita una interacción eficiente con una base de datos de películas de cine. Para esto, deberás procesar los datos entregados, implementar las tablas de entidades y relaciones necesarias y generar un menú por consola que permita: agregar, modificar, consultar y eliminar dicha información.

DCCine.csv

El archivo *DCCine.csv* posee todos los ratings hechos por cada persona a cada película. El separador del archivo es “;” (punto y coma para los amigos) y las columnas del mismo son:

- `user_id`: identificador único para cada usuario.
- `user_age`: edad del usuario.
- `user_gender`: género del usuario. El dato se encuentra entre “” - comillas. Ej: “M”.
- `user_occupation`: ocupación del usuario. El dato se encuentra entre “” - comillas. Ej: “Technician”.
- `user_zip_code`: código postal del usuario.
- `movie_title`: título de la película. El dato se encuentra entre “” - comillas. Ej: “Star Trek: First Contact (1996)”.
- `movie_release_date`: fecha de lanzamiento de la película. El formato de la fecha es YYYY-MM-DD y se encuentra entre “” - comillas. Ej: “1994-12-13”.
- `user_rating`: rating dado por el usuario a la película.
- `movie_genre`: género de la película. El dato se encuentra entre “” - comillas. Ej: “Animation”.
- `rating_date`: fecha del rating. El formato de la fecha es YYYY-MM-DD y se encuentra entre “” - comillas. Ej: “1994-01-21”.

Modelación

Tu sofisticado sistema debe soportar a la perfección la **creación, consulta, actualización y destrucción** tanto de usuarios como de películas. Por esta razón, debes construir tu sistema basándote en un modelo adecuado para soportar cada una de las operaciones antes mencionadas. Se espera entonces que tu modelo incluya al menos **tres tablas distintas**. También **debes** respetar las restricciones de integridad.

Como podrás haber notado tras inspeccionar el archivo .csv entregado, ya existen usuarios registrados en el sistema. A ellos, deberás crearles un nombre de usuario y una contraseña. Los nombres de usuario son únicos y deben tener **más de tres caracteres**. Las contraseñas deben tener un **mínimo de 6 caracteres**, e incluir **al menos una mayúscula y un número**. Debes aplicar estas mismas restricciones a los nuevos usuarios.

Recuerda que los usuarios también poseen la siguiente información:

- Identificador numérico único¹.
- Edad.
- Género.
- Ocupación.
- Código postal.

Interacciones - Parte B

Esta parte de la tarea consiste en la elaboración de un menú por consola que permitirá realizar toda interacción con la base de datos. Los requisitos mínimos con los que debe cumplir tu sofisticado sistema son:

- Tener un sistema de registro de nuevos usuarios. La consola debe contar con una opción para registrar nuevos usuarios (solicitando todos los datos necesarios) o ingresar al sistema con solo dar el nombre de usuario y contraseña.
- Diferenciar entre usuarios comunes y administradores. Al momento de registrarse, deben poder diferenciar si registran o no a un administrador, en ese caso, ese usuario tiene la capacidad de añadir nuevas películas, modificar datos de otros usuarios que no sean administradores, modificar información de películas y eliminar a usuarios que no sean administradores. Si es un usuario normal, este solo puede

¹Se espera que **no** le pidas esto al usuario en el formulario de registro, sino que tu sistema le asigne uno automáticamente.

dar un puntaje a una película o realizar consultas básicas para obtener información de la base de datos, sin modificarla.²

- Permitir la adición de nuevas películas, con toda su información relevante (solo por parte de los administradores).³
- Permitir la modificación de datos del usuario o de una película (solo por parte de los administradores).
- Eliminar personas de forma permanente de un sistema (solo por parte de los administradores). En este caso, se debe eliminar toda información que exista de esa persona en la base de datos, eso implica eliminar también sus ratings.

Las consultas mínimas que debe soportar su sistema⁴ son:

1. Todas las consultas necesarias para cumplir con los requisitos anteriores (*e.g.* la(s) consulta(s) para cumplir con el requisito de adición de nuevas películas).
2. Dado el nombre de una película y un puntaje, buscar todos los usuarios que hayan comentado (*rateado*) la película con una nota igual o mayor al puntaje señalado.
3. Dada una lista de géneros, entregar el puntaje/*rating* promedio de las películas de cada género de la lista.
4. Dada una ocupación, entregar el género más *rateado* por personas con esa ocupación.
5. Dadas dos fechas, mostrar la película (que no haya sido estrenada entre las fechas) más *rateada* por mujeres entre las fechas con *ratings* sobre el promedio de *ratings* de mujeres entre las fechas y la película que cumpla las mismas condiciones, pero para hombres.

Informe

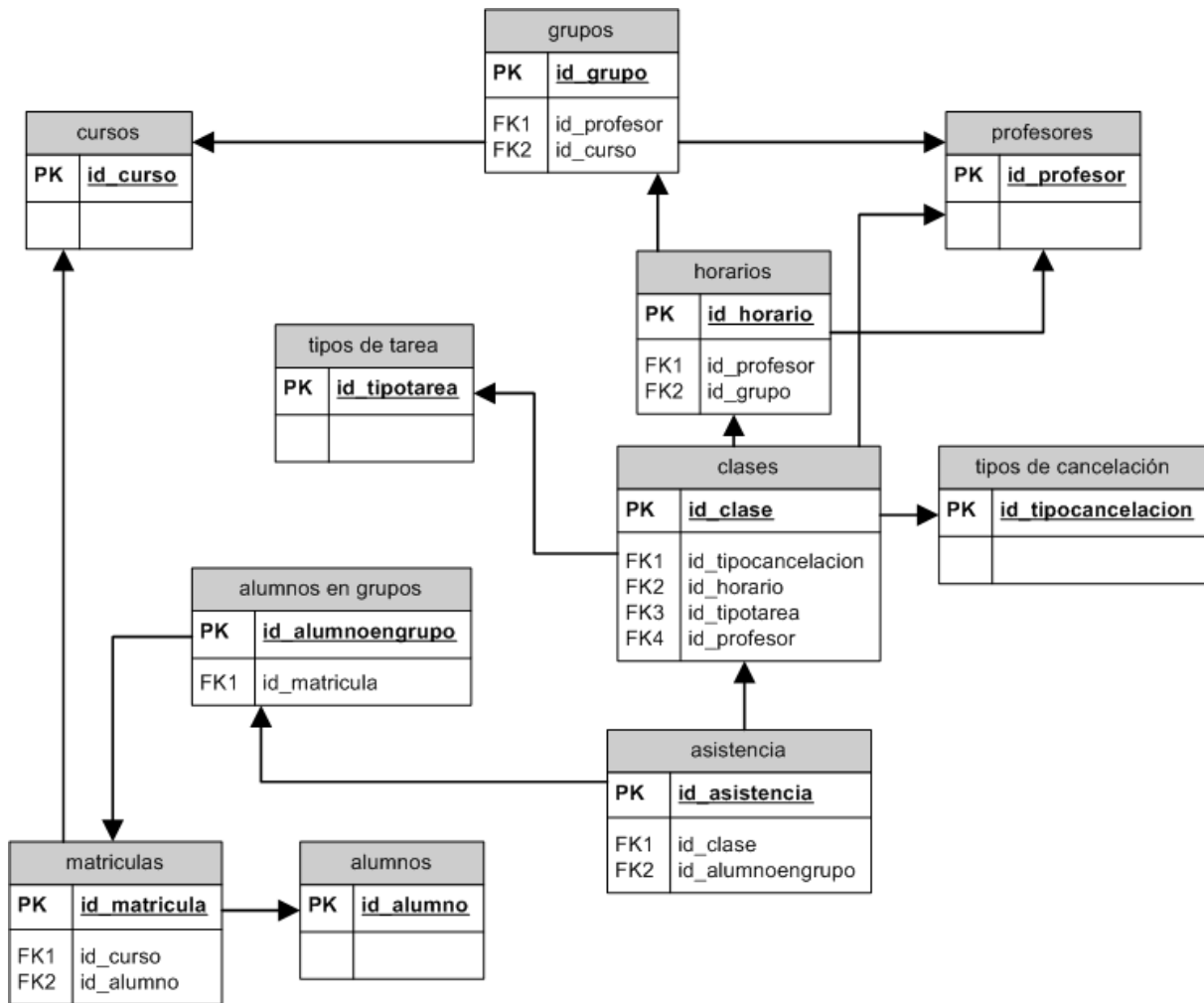
Para esta parte⁵ deberás entregar un diagrama de como el siguiente:

²Esto significa que los usuarios que no sean administradores solo pueden modificar la base de datos evaluando una película con un rating.

³Para esta funcionalidad, revisa los atributos relativos a películas en el archivo adjunto a esta tarea.

⁴Que su sistema soporte las consultas, se refiere a que provea la posibilidad de ejecutarlas mediante una interacción por consola

⁵Si, en la parte B pediremos otras cosas en el informe.



El diagrama debe especificar:

- Cada tabla creada para su base de datos.
- Cada atributo de las tablas.
- Destacar los atributos utilizados para conectar la información de una tabla con otra.
- Flechas que indiquen las relaciones de una tabla, es decir, dado una fila de una tabla, a que otra tabla se puede acceder.

Recomendamos fuertemente utilizar la siguiente página la cual puede exportar el diagrama en PDF y agregarlo al latex. Puedes, por supuesto, utilizar otras páginas o herramientas para este diagrama, según lo estimes conveniente.

Además, deberás explicitar y explicar la secuencia de pasos que gatilla cada una de tus consultas en tu base de datos, incluyendo la generación de tablas intermedias o parciales por medio de *joins* o selecciones. Es muy importante que tus explicaciones sean lo más detalladas, claras, precisas y concisas posible. Por

ejemplo, un paso en una consulta podría ser: "selección de todas las filas que tengan cierto atributo mayor a 8. Este paso genera un grupo de filas que puede ser más pequeño que la tabla sobre la cual se ejecuta la selección."

¡OJO! No se aceptará el uso de *pantallazos* para mostrar aspectos de su consola o el diagrama. Estos deben estar mostrados con alguna herramienta proporcionada por L^AT_EX ⁶

Para su corrección **debe** subir todos los archivos necesarios para la compilación del pdf, es decir, los *.tex*, cualquier imagen, pdf, etc.

⁶Sugerencia: investigar `listing` y/o `begin{verbatim}`.

Política de Integridad Académica

Los alumnos de la Escuela de Ingeniería deben mantener un comportamiento acorde al Código de Honor de la Universidad:

“Como miembro de la comunidad de la Pontificia Universidad Católica de Chile me comprometo a respetar los principios y normativas que la rigen. Asimismo, prometo actuar con rectitud y honestidad en las relaciones con los demás integrantes de la comunidad y en la realización de todo trabajo, particularmente en aquellas actividades vinculadas a la docencia, el aprendizaje y la creación, difusión y transferencia del conocimiento. Además, velaré por la integridad de las personas y cuidaré los bienes de la Universidad.”

En particular, se espera que mantengan altos estándares de honestidad académica. Cualquier acto deshonesto o fraude académico está prohibido; los alumnos que incurran en este tipo de acciones se exponen a un procedimiento sumario. Ejemplos de actos deshonestos son la copia, el uso de material o equipos no permitidos en las evaluaciones, el plagio, o la falsificación de identidad, entre otros. Específicamente, para los cursos del Departamento de Ciencia de la Computación, rige obligatoriamente la siguiente política de integridad académica en relación a copia y plagio: Todo trabajo presentado por un alumno (grupo) para los efectos de la evaluación de un curso debe ser hecho individualmente por el alumno (grupo), sin apoyo en material de terceros. Si un alumno (grupo) copia un trabajo, se le calificará con nota 1.0 en dicha evaluación y dependiendo de la gravedad de sus acciones podrá tener un 1.0 en todo ese ítem de evaluaciones o un 1.1 en el curso. Además, los antecedentes serán enviados a la Dirección de Docencia de la Escuela de Ingeniería para evaluar posteriores sanciones en conjunto con la Universidad, las que pueden incluir un procedimiento sumario. Por “copia” o “plagio” se entiende incluir en el trabajo presentado como propio, partes desarrolladas por otra persona. Está permitido usar material disponible públicamente, por ejemplo, libros o contenidos tomados de Internet, siempre y cuando se incluya la cita correspondiente.