



Ingeniería de datos

Tarea: Analizando una base de datos con SQL III

¡Bienvenido(a)!

Te invitamos a realizar el primer trabajo.

- **Objetivo:** Elegir las consultas adecuadas para un caso de uso concreto.
- **Tipo de actividad:** Individual
- **Tipo de evaluación:** Sumativa (con calificación)
- **Ponderación:** 15% (Equivale al 15% de la nota final del curso)
- **Puntaje:** 60 puntos
- **Calificación:** Escala de 1 a 7, con una exigencia de 50%. La nota mínima para aprobar es 4.0.

Instrucciones

1. Antes de comenzar, debes haber revisado las siguientes clases y las lecturas del módulo 1 y 2.
2. Lee con atención el enunciado de la tarea que se adjunta a continuación
3. Debes entregar un archivo .ipynb con tus respuestas. Se adjunta un formulario de respuestas para que uses como template. Descargalo y desarrolla allí tu tarea.
4. Instrucciones:
 - Haz clic en el botón para agregar entrega. Se abrirá una nueva ventana que permite arrastrar el archivo y subirlo.
 - Comprueba que el archivo arrastrado es el correcto y presiona el botón para guardar cambios. El documento quedará guardado en la plataforma.

Enunciado

Introducción

En esta tarea continuaremos desarrollando el caso de uso visto en la Tarea 1. Para esto, te entregaremos una descripción aumentada de la base de datos, junto con su esquema, y varios datos que debes cargar a las tablas. Después de esto, te pediremos analizar a la base de datos ocupando SQL. En esta tarea recibirás: en enunciado, y un Jupyter notebook para probar tu solución (los datos se cargarán directamente a través del notebook).

Descripción del problema

Tu esfuerzo de armar una base de datos sobre talleres realizados en la municipalidad de Fantasilandia fue un tremendo éxito, y ahora un consorcio internacional contrató a tu empresa para armar una aplicación que permite guardar la información sobre talleres en varios países. Obviamente, la expansión de la base de clientes de tu aplicación también requiere ampliar un poco el caso de uso, y agregar ciertas cosas a tu esquema inicial.

En particular, la información que debes guardar ahora es la siguiente:

Personas. Esta tabla debe guardar la información de personas que participan en talleres. Para cada persona, debes guardar su nombre, rut, correo, y un número de teléfono. Lo mismo cómo en Tarea 1.

Talleres. Tu base de datos también debe tener la información sobre talleres que se pueden realizar. Para cada taller debes guardar su nombre, nombre del profesor que lo dicta, la fecha, la hora, el valor que tiene el taller, y su duración. Adicionalmente, tu aplicación debe permitir guardar la información sobre dos tipos distintos de talleres: talleres deportivos, y talleres instrumentales. Un taller deportivo tiene, en adición a los atributos de cualquier taller, el nombre de deporte, número de equipos, y el número de personas por equipo. Por otro lado, un taller instrumental también contiene el nombre de instrumento, y la cantidad de instrumentos disponibles.

Lugares. También debes guardar la información sobre lugares dónde se realizan los talleres. Un lugar tiene una dirección, compuesta por calle, número, código postal, una ciudad, y un país.

Participación. Tus datos deben modelar la participación de personas en talleres. Aquí, las personas, y los talleres deben ser identificado con las llaves de sus tablas respectivas. Para modelar esto te puedes inspirar con nuestro ejemplo de actores y películas.

Ubicación. Similarmente cómo con participación, debes tener una o más tablas que te permite decir en cual lugar se realiza cual taller. De nuevo, para identificar los talleres y los lugares, te debes referir a cada uno con su llave de su tabla respectiva.

Esta vez te entregaremos un esquema que modela estos requerimientos, y que se ve de la siguiente manera (nosotros no especificaremos los tipos de datos, sino solo los nombres de los atributos).

Persona(rut,nombre,correo,teléfono)

**TallerDeportivo(id,nombreTaller, nombreProfesor,
fecha,hora,valor,duración,deporte,nrEquipos, nrPersonas)**

**TallerInstrumental(id,nombreTaller, nombreProfesor,
fecha,hora,valor,duración,instrumento,nrInstrumentos)**

Lugar(id,calle,número,código,nombreCiudad,nombrePaís)

Participa(rut,idTaller)

EstáEn(idTaller,idLugar)

Pregunta 1

En la primera pregunta debes validar que los datos se cargaron correctamente, y que tu notebook pudo cargar la base de datos de manera correcta. Para esto, debes ejecutar la primera celda en el notebook que viene con esta tarea. Después, para validar que la carga fue exitosa, debes correr el comando:

SELECT *

FROM NombreTabla

LIMIT 5

para cada tabla de nuestro esquema. Por ejemplo, para validar que se cargaron correctamente las personas, debes correr:

SELECT *

FROM Persona

LIMIT 5.

Un comentario: nótese que la tabla se llama Persona, y no Personas.

Pregunta 2

Ahora que tenemos nuestra base de datos preparada, la podemos analizar en más detalles. En esta primera pregunta te pedimos encontrar la(s) persona(s) que participa(n) en máxima cantidad de talleres que se realizan en Honduras. Se requiere escribir una consulta SQL que recupera esta información sobre nuestro esquema. Esto quiere decir que la consulta debe devolver esta misma información sobre cualquier instancia de nuestro esquema, y no solo con la instancia que estamos revisando.

Pregunta 3

Ahora nos interesaría encontrar el nombre del deporte más popular. Para esto, escriba la consulta SQL que encuentra el nombre de deporte de taller deportivo con la mayor cantidad de participantes. De nuevo, tu consulta debe funcionar sobre cualquier instancia **de nuestro esquema**.

Pregunta 4

Ahora para cada taller instrumental cuenta el número de gente con el correo "@uc" participando. Debes escribir una consulta SQL que devuelve el id del taller, nombre del instrumento, y el número de personas con el correo "@uc" inscritas a este taller. Para filtrar la gente participando, ocupa el comando LIKE, que permite buscar sobre un string.

Pregunta 5

Para cada persona, computa la cantidad de talleres en cuales participa esta persona. Esto debe incluir tantos talleres deportivos, cómo talleres instrumentales. Una manera de

resolver esta consulta es ocupando consultas anidadas en la cláusula FROM, computando dicho número para cada tipo de taller.

Aspectos formales

Considera los aspectos formales que se describen a continuación:

- Letra Arial 12 normal, interlineado simple.
- Tamaño página: Letter (carta).
- Utilizar formato APA en citas al interior del texto y en la bibliografía.

¡Mucho éxito!

Importante: la fecha de entrega está indicada en el calendario del curso. Cuidar la redacción y la ortografía. Si tienes alguna duda sobre los contenidos o sobre cómo realizar esta actividad, puedes utilizar la herramienta "Mensajes" y enviar tu pregunta. Recibirás la respuesta de su tutor con las orientaciones correspondientes