



IIC2115 – Programación como Herramienta para la Ingeniería (I/2024)

Laboratorio 6

Aspectos generales

- **Objetivo:** evaluar individualmente el aprendizaje sobre bases de datos relacionales, a través de la resolución de ejercicios asociados a la modelación de entidades y la posterior actualización de estas.
- **Lugar de entrega:** Parte 1 lunes 10/06 a las 17:30, Parte 2 domingo 16/06 a las 23:59, ambas en el repositorio privado.
- **Formato de entrega:** archivos Python Notebook L6_1.ipynb y L6_2.ipynb con las soluciones de las partes 1 y 2 del laboratorio. Los archivos deben estar ubicados en la carpeta L6. Utilice múltiples celdas de texto y código para facilitar el trabajo del cuerpo docente. Entregas que no cumplan el formato tendrán un descuento de 0,5 pts.
- **Entregas atrasadas:** el descuento por atraso para la Parte 1 es de 1 punto cada 10 minutos o fracción. El descuento por atraso para la Parte 2 es de 1 punto por cada hora o fracción.
- **Issues:** Las discusiones en las *issues* del Syllabus que sean relevantes para el desarrollo de la evaluación, serán destacadas y se considerarán como parte de este enunciado. Así mismo, el uso de librerías externas que solucionen aspectos fundamentales del problema no podrán ser utilizadas. Solo se podrán utilizar las que han sido aprobadas en las *issues*, previa consulta de los estudiantes.
- **Entregas con errores de sintaxis y/o que generen excepciones en todas las ejecuciones** serán calificados con nota 1.0.

Introducción

Con el fin practicar los contenidos de bases de datos relacionales, en este ejercicio deberá crear una base de datos con tablas que modelen las entidades presentes en un conjunto de datos que recopila información de aplicaciones disponibles para el sistema operativo Android y sus respectivos *reviews*.

Descripción de los datos

Considere la base de datos *Google Play Store Apps*, que recopila información de miles de aplicaciones disponibles para el sistema operativo Android, junto con múltiples *reviews* para ellas. La información se encuentra disponible 2 archivos *.csv*. El primero se llama *googleplaystore.csv*, el cual contiene la información de cada aplicación:

- **App:** nombre de la aplicación
- **Category:** categoría de la aplicación. Entre estas se encuentran ART_AND_DESIGN, FAMILY, GAME, etc.
- **RatingOverall:** *rating* de la aplicación
- **Reviews:** número de comentarios que tiene la aplicación. Esta columna debe ser ignorada. En otro archivo estan todos los comentarios y esa cantidad será la utilizada.
- **Size:** tamaño de la aplicación
- **Installs:** indicador que representa cantidad de descarga de la aplicación.
- **Type:** indica si es pagada o gratuita.
- **Price:** precio de la aplicación
- **Content Rating:** grupo de edad a la que va dirigida la aplicación.
- **Genres:** género de la aplicación, se puede considerar como una sub-categoría y cada aplicación puede tener más de un género. En ese caso, cada género está dividido por un punto y coma (“;”).
- **Last Updated:** fecha de la última vez que fue actualizada la aplicación
- **Current Ver:** versión de la aplicación.
- **Android Ver:** versión de *Android* requerida por la aplicación.

El segundo archivo lleva por nombre `googleplaystore_user_reviews.csv` y contiene diferentes comentarios de los usuarios para cada aplicación:

- **App:** nombre de la aplicación.
- **Translated_Review:** Comentario de la aplicación. Todas traducidas al inglés.
- **Sentiment:** Indica si el comentario es positivo, negativo o neutral.
- **Sentiment_Polarityws:** puntaje asociado a la polaridad del comentario, valor entre -1 y 1. Esta columna debe ser ignorada.
- **Sentiment_Subjectivity:** puntaje asociado al sentimiento entregado en el comentario. Esta columna debe ser ignorada.

Parte 1

- a) En base a la estructura de los datos almacenados en el archivo, deberá utilizar conceptos de bases de datos relacionales para crear tablas que modelen al menos tres entidades y las relaciones entre ellas, incluyendo las restricciones de integridad correspondientes. El modelo debe capturar la misma información que la disponible en los archivos.
- b) Una vez cumplida la misión anterior, deberá extraer los datos del archivo y cargarlos en las tablas correspondientes, creando y respetando las relaciones existentes en los datos.

Parte 2

En esta parte del laboratorio deberá extender lo realizado anteriormente, incorporando funciones para modificar la base de datos, además de introducir nueva información en esta.

- a) **Ediciones a la base de datos:** desarrolle 4 funciones para agregar, editar y eliminar información de su base de datos. Para cada una, incluya al menos 2 ejemplos de ejecución. Las funciones son:
 - **add_app(app_data):** esta función recibe un diccionario con la información de la aplicación. El formato del diccionario es:

Llave	Valor
name	nombre de la aplicación
category	categoría de la aplicación
rating	rating de la aplicación
size	tamaño de la aplicación
price	precio de la aplicación
version	versión de la aplicación
android	versión de <i>Android</i> de la aplicación
genres	géneros de la aplicación

Cuando se agrega la aplicación, debe suponer que esta tiene 0 comentarios, 0 instalaciones y que la última actualización fue cuando se agregó la aplicación a la base de datos. Para la lista de géneros, si algún género no existe, este debe ser agregado a la base de datos. En caso que un valor del diccionario no sea correcta, la función debe imprimir en pantalla “No es posible agregar la aplicación”.

- **add_comment(app_name, text, sentiment)**: agregar un nuevo comentario de una aplicación. Debe recibir el nombre de la aplicación, un texto y un sentimiento (positivo, neutro o negativo). No puede suponer que siempre va a existir el nombre de la aplicación.
- **download_app(app_name)**: descarga una aplicación. Esto incrementa en 1 el número del atributo descargas. No puede suponer que siempre va a existir el nombre de la aplicación.
- **delete_app(app_name)**: elimina toda la información asociada a la aplicación con el nombre ingresado. No puede suponer que siempre va a existir el nombre de la aplicación.

- b) Cree y llene tablas de información agregada (o de resumen) para la columna **Rating**, **Installs** y **Content Rating** (una tabla por columna). Qué aspectos deben agregarse o resumirse por tabla debe ser decidido por usted.
- c) Incorpore nueva información en la base de datos, utilizando *webscraping* para descargar contenido de internet. Para esto, puede basarse en el material disponible en el curso o utilizar otras herramientas. No existe un número mínimo de filas a incorporar a la base de datos, ni es necesario que la nueva información cubra exactamente las mismas columnas existentes, solo es necesario que la nueva información se descargue con *webscraping* y que lógicamente tenga relación con la base de datos.

Corrección

Es importante que deje todas las celdas de su trabajo ejecutadas antes de subir el archivo, de lo contrario se le aplicará un descuento de 0,5 puntos al puntaje total.

Para la corrección se revisarán los procedimientos desarrollados para responder los diferentes ejercicios y cómo utiliza las librerías en ellos. Se espera que las interacciones con la base de datos que no sean triviales, sean claramente justificadas.

Política de Integridad Académica

Los/as estudiantes de la Escuela de Ingeniería de la Pontificia Universidad Católica de Chile deben mantener un comportamiento acorde a la Declaración de Principios de la Universidad. En particular, se espera que mantengan altos estándares de honestidad académica. Cualquier acto deshonesto o fraude académico está prohibido; los/as estudiantes que incurran en este tipo de acciones se exponen a un Procedimiento Sumario. Es responsabilidad de cada estudiante conocer y respetar el documento sobre Integridad Académica publicado por la Dirección de Docencia de la Escuela de Ingeniería.

Específicamente, para los cursos del Departamento de Ciencia de la Computación, rige obligatoriamente la siguiente política de integridad académica. Todo trabajo presentado por un/a estudiante para los efectos de la evaluación de un curso debe ser hecho **individualmente** por el/la estudiante, **sin apoyo en material de terceros**. Por “trabajo” se entiende en general las interrogaciones escritas, las tareas de programación u otras, los trabajos de laboratorio, los proyectos, el examen, entre otros.

En particular, si un/a estudiante copia un trabajo, o si a un/a estudiante se le prueba que compró o intentó comprar un trabajo, **obtendrá nota final 1.1 en el curso** y se solicitará a la Dirección de Pregrado de la Escuela de Ingeniería que no le permita retirar el curso de la carga académica semestral.

Por “copia” se entiende incluir en el trabajo presentado como propio, partes hechas por otra persona. En caso que corresponda a “copia” a otros estudiantes, la sanción anterior se aplicará a todos los involucrados. En todos los casos, se informará a la Dirección de Pregrado de la Escuela de Ingeniería para que tome sanciones adicionales si lo estima conveniente.

También se entiende por copia extraer contenido sin modificarlo sustancialmente desde fuentes digitales como Wikipedia o mediante el uso de asistentes inteligentes como ChatGPT, Gemini o Copilot. Se entiende que una modificación sustancial involucra el análisis crítico de la información extraída y en consecuencia todas las modificaciones y mejoras que de este análisis se desprendan. Cualquiera sea el caso, el uso de

fuentes bibliográficas, digitales o asistentes debe declararse de forma explícita, y debe indicarse cómo el/la estudiante mejoró la información extraída para cumplir con los objetivos de la actividad evaluativa.

Obviamente, está permitido usar material disponible públicamente, por ejemplo, libros o contenidos tomados de Internet, **siempre y cuando se incluya la referencia correspondiente**.

Lo anterior se entiende como complemento al Reglamento del Estudiante de la Pontificia Universidad Católica de Chile (<https://registrosacademicos.uc.cl/reglamentos/estudiantiles/>). Por ello, es posible pedir a la Universidad la aplicación de sanciones adicionales especificadas en dicho reglamento.

Compromiso del Código de Honor

Este curso suscribe el Código de Honor establecido por la Universidad, el que es vinculante. Todo trabajo evaluado en este curso debe ser propio. En caso que exista colaboración permitida con otros/as estudiantes, el trabajo deberá referenciar y atribuir correctamente dicha contribución a quien corresponda. Como estudiante es un deber conocer el Código de Honor (<https://www.uc.cl/codigo-de-honor/>).