



Laboratorio # 3 - Conjunto de Problemas # 3

Laboratorio y Conjunto de Problemas: Proyecto (Repositorio, Propuesta, Fuente de Datos)

Instructor del Laboratorio: Ing. Gonzalo Armando Maradiaga Solano

Profesor: Dr. Servio Palacios Valor: 5 % del total de la clase (1 % el laboratorio, 4 % los problemas).

Resumen

Este laboratorio y conjunto de problemas pretende ayudar a los grupos/equipos para crear el repositorio del Proyecto en Github y revisar/explorar/limpiar la fuente de datos del Proyecto.^{1 2}

1. Repositorio del Proyecto (Valor: 2 puntos)

Para este laboratorio, creamos un documento de plantilla que pueden llenar. La plantilla es llamada **"Laboratorio 03 - Libro de Trabajo"**.

Para esta sección del laboratorio y problemas, los estudiantes deben crear una organización y al menos un repositorio para el Proyecto.

Por favor ver el video relacionado a la creación del repositorio del Proyecto y cómo se clonan repositorios (**Proyecto - Github** en Blackboard).

Para esta sección se debe hacer lo siguiente:

1. Seguir el video **"Proyecto - Github"** que está en Blackboard para crear una organización para el equipo del proyecto, y al menos un repositorio con la solución propuesta. **Este paso solo es necesario hacerlo una vez, lo pueden hacer como grupo o designen una persona del grupo que lo haga** (revisar video). En este paso, como se muestra en el video, la persona encargada debe agregar a sus compañeras/compañeros de grupo. Esta Organización es nueva y exclusiva para su Proyecto. Todo el código fuente y documentación deben subirse en este organización/repositorio.
2. Al crear el repositorio, como se muestra en el video, deben crear un archivo **README.md** con la información del proyecto. Para este paso, pueden utilizar la información de la propuesta del proyecto (se explica el archivo readme en el video). Por favor analicen el enlace del video:
<https://github.com/CoolSoftwareTeam/CoolSoftware>
3. Cada miembro del grupo debe subir/editar el README.md (commit, push, como se explica en el video) con su información desde su Github (agregar su nombre e información relevante en la sección **Integrantes**). Debe verse claramente el commit/push de su información a este archivo **README.md**. Ver ejemplo en: <https://github.com/CoolSoftwareTeam/CoolSoftware>, los nombres en la lista deben ser editados por cada uno de los miembros del grupo.
4. Compartir el nombre del organización/repositorio en el documento.

¹Última edición (versión 2.0): 14 de Septiembre del 2022, 19:16 ET

²Debido a algunos caracteres de escape y formato, tildes y otros caracteres pueden ser omitidos para claridad en el código fuente.

2. Fuente de Datos (Valor: 2 puntos)

En esta sección, los estudiantes analizarán la fuente de datos seleccionada. Algunos de los objetivos de esta sección incluyen:

- Definir el **formato** de los datos (CSV, JSON).
- Analizar la fuente de datos (data source o dataset) para realizar el proceso de limpieza de los datos [1].
- Definir las propiedades importantes de la fuente de datos.
- Estandarizar la fuente de datos y nombrar los archivos de datos. Utilizar directorios y nombres de archivos consistentes.
- Discutir procesos estadísticos o de análisis (algoritmos) que se pueden ejecutar sobre la fuente de datos.

Después de analizar la fuente de datos, hacer lo siguiente:

1. Utilizar la información del análisis de la fuente de datos y subirla al repositorio. Esto puede ser parte del README.md (una sección llamada **Fuente de Datos**). Por favor revisar el ejemplo del Video, también pueden encontrar un ejemplo acá: <https://github.com/CoolSoftwareTeam/CoolSoftware>
2. Subir la fuente de datos al repositorio (subirlo a un subdirectorio **data**).

Es responsabilidad de cada grupo tener el repositorio, fuente de datos, y cualquier documentación necesaria antes de la reunión de definición de alcance del proyecto.

3. Laboratorio (1 Punto)

Un ejemplo de como debe lucir la organización/repositorio en Github se muestra en:

<https://github.com/CoolSoftwareTeam/CoolSoftware>

Recuerden que cada integrante del grupo debe hacer un commit/push (ver video) en el repositorio.

Para este laboratorio haga lo siguiente:

1. Conteste las preguntas del documento **Laboratorio 03 - Libro de Trabajo**.
2. Crear el repositorio de la Sección 1. Sigán las instrucciones en la Sección 1.
3. Incluir a los integrantes del grupo en el repositorio del Proyecto. Sigán las instrucciones en la Sección 1.
4. Subir únicamente el documento **Laboratorio 03 - Libro de Trabajo** a Blackboard (se revisará la organización, repositorio, y los commits de cada estudiante). Este documento debe contener el enlace al repositorio hacia el cuál se efectuó el trabajo del grupo. Toda la documentación, código fuente, y fuente de datos debe estar en el repositorio.

Para la creación de esta Organización y Repositorio, utilizar mejores prácticas.

- Utilizar nombres consistentes y breves.
- Utilizar directorios en la medida de lo posible. Un directorio para los datos (data), un directorio para el código (src), etc.

- Si el grupo ya empezó con el desarrollo del Software/Programa, por favor debuggear el código fuente utilizando los métodos vistos en clase. De forma similar, se explicó cómo buscar ayuda en línea (StackOverflow [2], W3Schools [3], y cualquier otra fuente que consideren necesaria).

Referencias

- [1] J. Takagi, “Cleaning up data and turning a csv file into json using python.” <https://medium.com/analytics-vidhya/cleaning-up-data-and-turning-a-csv-file-into-json-using-python-72e67c2ee76e>, 2022. [Online; accessed Septiembre 10, 2022].
- [2] Stackoverflow, “Stack Overflow.” <https://stackoverflow.com/>, 2022. [Online; accessed Septiembre 14, 2022].
- [3] w3Schools, “Python tutorial.” <https://www.w3schools.com/python/default.asp>, 2022. [Online; accessed Septiembre 14, 2022].