

## EVALUACIÓN TALLER 2

### colaborativo

semestre: 2024-1

**Curso: Lenguaje de Programación 2**

**Objetivo:**

Integrar los elementos y conceptos desarrollados en clase a través de su aplicación en el desarrollo de un proyecto de extracción de datos de la web desarrollado de manera colaborativa.

### ELABORACIÓN y DESARROLLO DE UN PROYECTO DE WEB SCRAPING

#### **Primera parte: INSCRIPCIÓN DEL PROYECTO**

- Inscribir su tema del proyecto, detallando las fuentes de datos que se va a extraer información(**webs o APIS a utilizar**), sus objetivos y producto final (idea)
- Indicar los nombres de los integrantes con sus usuarios en GitHub en el foro correspondientes.

#### **Segunda parte: PRESENTACIÓN - DESARROLLO DE SU PROYECTO Y PRODUCTO FINAL**

- Presentación final (informe) de su proyecto, con el código documentado, el diseño y planeamiento de la extracción y estructuración de los datos o información extraída y combinada. Además, de los datos extraídos e integrados (en una o más hojas del tipo .csv o base de datos).
- La presentación reporte consta de la url del repositorio GitHub de trabajo, el cual deben ser compartida en el aula virtual en la actividad correspondiente.
- La presentación oral consta de exposición del proyecto enfatizando en la relevancia y dificultades encontradas. Tiempo de exposición entre 20 y 25 minutos

#### **Indicaciones:**

- EL trabajo es grupal colaborativo (mínimo 2 y máximo 3 integrantes) utilizando Git y Git Hub. NO SE ACEPTARÁN NI CALIFICARÁN TRABAJOS INDIVIDUALES.
- Debe descargarse información de al menos dos fuentes de información: una utilizando scraping de páginas webs (webscraping 'duro'), otra usando APIs y/o descarga de .csv o similares.
- Deben crear un repositorio en GitHub en el que se desarrollará el proyecto de forma colaborativa. Es importante no solo colgar el .ipynb final, sino la planeación del diseño de la descarga y estructuración de los datos al combinarlos, además de la documentación que evidencie el trabajo del diseño y del desarrollo del proyecto, o cualquier otra información o documento que permita evidenciar el detalle del trabajo desarrollado en equipo.
- El código debe desarrollarse en Python (JupyterNotebook)

#### **ALGUNAS IDEAS DE PROYECTOS EN WEB-SCRAPING:**

<https://rayobyte.com/blog/web-scraping-ideas/>

<https://www.mygreatlearning.com/blog/top-web-scraping-projects/>  
<https://www.projectpro.io/article/web-scraping-projects-ideas/475>  
<https://www.zenrows.com/blog/web-scraping-project-ideas>