

# Evaluación Almacenamiento y Captura de datos

Claudio Aracena

Entrega: 5 de enero 2019

En grupos de máximo tres personas deberán resolver los dos problemas descritos abajo. Deben entregar el notebook con el código para resolver cada problema explicado el detalle.

1. Haga un proceso de web scraping para el sitio web <https://www.pcfactory.cl/>. Para ello realice las siguientes tareas:
  - a. Extraer todos los nombres, marcas y precios de los productos bajo la categoría de notebooks. (3 puntos)
  - b. Crear un dataframe con la información extraída. (1 punto)
  - c. Crea una base de datos con la información del dataframe. (1 punto)
  - d. Realizar la consulta SQL que entregue el precio promedio agrupado por marca. (1 punto)
  
2. Utilizando la herramienta Selenium, haga un proceso de web scraping para el sitio web <https://es.wikipedia.org>. Para ello realice las siguientes tareas:
  - a. Utilizar selenium para emular un navegador y cargar la página deseada. (1 punto)
  - b. Ingresar en el input de búsqueda el string "países por superficie" y presionar enter. (0,5 puntos)
  - c. Esperar por la página de resultados. (0,5 puntos)
  - d. Extraer la lista de países y territorios con su superficie desde la página de resultados (no incluir Planeta Tierra, Océanos y Tierras emergidas). (2 puntos)
  - e. Cerrar el navegador. (0,5 puntos)
  - f. Convertir la lista de países y territorios con su superficie en un dataframe. (0,5 puntos)
  - g. Calcular la suma, el conteo y el promedio de superficies utilizando pandas. (1 punto)

Cada pregunta tiene una escala de 1 a 7. El promedio de ambas preguntas resulta en la nota final de la tarea.