

# Tarea 1 - Programación con Python

---

## Taller de Python para Ciencia de Datos

### 1. Enunciado

Desarrollar un programa utilizando el lenguaje de programación Python que lea un documento de texto con codificación UTF-8 y que permita realizar las siguientes acciones:

- Entregar una estadística del documento con la siguiente información (los caracteres de control, como los saltos de líneas, no se deben considerar en el conteo de caracteres)(los saltos de líneas vacíos, no se deben considerar en el conteo de párrafos):
  - Número de palabras en total.
  - Número de palabras no repetidas.
  - Número de caracteres con espacio.
  - Número de caracteres sin espacio.
  - Número de párrafos (saltos de líneas).
- Buscar:
  - Buscar una palabra completa sin distinción de mayúsculas y minúsculas indicando cuantas veces aparece.
  - Buscar una palabra completa con distinción de mayúsculas y minúsculas indicando cuantas veces aparece.
  - Buscar una palabra o substring sin distinción de mayúsculas y minúsculas indicando cuantas veces aparece.

- Buscar una palabra o substring con distinción de mayúsculas y minúsculas indicando cuantas veces aparece.
- Reemplazar:
  - Reemplazar una palabra completa sin distinción de mayúsculas y minúsculas guardando los cambios en el archivo.
  - Reemplazar una palabra completa con distinción de mayúsculas y minúsculas guardando los cambios en el archivo.
  - Reemplazar una palabra o substring sin distinción de mayúsculas y minúsculas guardando los cambios en el archivo.
  - Reemplazar una palabra o substring con distinción de mayúsculas y minúsculas guardando los cambios en el archivo.

En la codificación del documento debe utilizar al menos dos de las cuatro estructuras especiales de Python (listas, tuplas, diccionarios y conjuntos). Sólo se debe utilizar los métodos nativos o integrados de Python y no ayudarse con alguna biblioteca externa que ya implemente los requerimientos solicitados.

## 2. Programa

El código se puede desarrollar usando una metodología de programación modular utilizando celdas de Python con alguna definición de función o bloque de código que implementen una funcional de manera integral.

Cada celda de código Python se debe explicar por alguna celda Markdown que describa su funcionalidad. También cada celda de Python debe tener algún comentarios que explique alguna de las instrucciones utilizadas.

## 3. Evaluación

El trabajo debe de ser original y no copiado desde Internet o de algún compañero (copiar y pegar), de lo contrario será evaluado con la nota NCR.

El programa que se entregue deberá ser revisado con el profesor para verificar su apropiada codificación y ejecución.

La pauta de evaluación es la siguiente:

- Avance (trabajo efectivo desarrollado en la semana) 20 %
- Interfaz (presentación, manejo de errores, manual de instrucciones, Readme) 10 %.
- Código (originalidad, uso de tipos de datos e instrucciones de Python) 30 %.
- Funcionalidad (nivel de implementación de los requerimiento solicitados) 40 %.

## **4. Entregas**

La tarea se debe subir en la plataforma EVA antes de la fecha definida, ya seas en en el Syllabus de la asignatura o por un acuerdo en clases.