

## Ejercicio 2:

Se necesita realizar el análisis de un texto y para ello es necesario realizar un preprocesamiento que incluye una serie de transformaciones y normalizaciones. Este preprocesamiento incluye las siguientes etapas:

1. Eliminación de stop words
2. Convertir a mayúsculas

Los stop words serán reconocidos a través de un archivo de texto que se recibirá como parámetro y se eliminarán del texto a analizar. ([https://en.wikipedia.org/wiki/Stop\\_words](https://en.wikipedia.org/wiki/Stop_words))

Una vez finalizado el preprocesamiento se deberá contabilizar la frecuencia de aparición de cada palabra en el texto y generar un reporte ordenando los resultados de mayor a menor frecuencia en un archivo de salida frecuencias\_{Nombre archivo entrada}\_{yyyy-mm-dd\_hh:mm:ss}.out con formato CSV y por pantalla sólo se mostrarán las 5 palabras con mayor frecuencia.

Parámetros (el orden de los parámetros no debe ser fijo):

- -StopWords “archivo stopwords”: path absoluto o relativo del archivo con los stopwords.
- -Resultado “directorio resultado”: path absoluto o relativo del directorio donde se generará el archivo de salida. Opcional. Si no se informa se generará en el directorio de ejecución.
- -Entrada “archivo a analizar”: path absoluto o relativo del archivo de texto a analizar.

Resultado esperado:

1. Archivo de salida de las frecuencias del archivo analizado
2. Por consola se mostrarán las 5 palabras con mayor frecuencia absoluta

Por ejemplo:

```
COVID,20
CASOS,15
ARGENTINA,10
BRASIL,9
URUGUAY,5
```