

Práctica 1A: Cálculo de frecuencias a nivel de un carácter

Guillermo Ruiz

Crea un notebook para preprocesar el texto y convertirlo a minúsculas, quita acentos y los siguientes caracteres:

```
" ; : , . \ \ - \ " ' / ( ) [ ] _ { } ~ < > | « » -- ' "
```

Reemplaza las tabulaciones y saltos de línea con un espacio.

2. Calcula las frecuencias a nivel de un carácter, de los documentos text_1.txt al text_5.txt y genera los histogramas para cada archivo. Para generar las gráficas, ordena los caracteres de acuerdo a su frecuencia, de mayor a menor frecuencia en el histograma.

Los archivos son:

- text_1.txt
- text_2.txt
- text_3.txt
- text_4.txt
- text_5.txt

3. Interpreta cada histograma en referencia a qué caracteres destacan y dar indicios de por qué ocurren esas observaciones.