

# Trabajo Práctico Especial

## Parte 2



FACULTAD DE  
**CIENCIAS EXACTAS**  
UNICEN

Entrega N° 25

Programación 3

TUDAI

Ciencias Exactas

UNICEN

Fecha de entrega: 01/07/2022

### Alumnos

Guido Zoppo

zoppoguido4@gmail.com

Agustin Montes

montesagustin99@gmail.com

# Introducción

En el presente trabajo se desarrollará un análisis sobre la implementación de diferentes servicios partiendo desde un historial de búsqueda de géneros resultante de las búsquedas realizadas por los usuarios.

Los servicios a implementar son los siguientes:

1. Obtener los N géneros más buscados luego de buscar por el género A.
2. A partir de un género A encontrar, en tiempo polinomial, la secuencia de géneros que más alto valor de búsqueda posee. Vamos a definir el valor de búsqueda de una secuencia como la suma de los arcos que la componen.
3. Obtener el grafo únicamente con los géneros afines a un género A; es decir que, partiendo del género A, se consiguió una vinculación cerrada entre uno o más géneros que permitió volver al género de inicio.

# Desarrollo

Para implementar el grafo, se tuvo en cuenta que un vértice debe representar a un género y que éste puede llegar a tener más de un vértice adyacente (representando un género buscado inmediatamente después que éste), se procedió a utilizar la estructura HashMap donde las claves serán del tipo String, representando los vértices, y el valor será un ArrayList de arcos, donde cada arco tendrá un género origen, otro género destino y una etiqueta que indicará la cantidad de veces que se buscó el género destino inmediatamente después de haber buscado género origen.

A la hora de implementar el servicio 1, se realiza un solo llamado a la estructura elegida para almacenar los vértices y arcos del grafo, dicho llamado es para realizar la búsqueda de los arcos salientes del género que el usuario ingresa, con este último como clave. Analizando lo anterior, se puede afirmar que la estructura elegida para la implementación tiene un impacto positivo ya que este tipo de acciones, en promedio, tienen un costo computacional bajo.

Tiempo genero leyendas	
Archivo	Milisegundos para N=3
Dataset 1	0,95
Dataset 2	0,79
Dataset 3	1,4
Dataset 4	1,01

Imagen 1.1 - Análisis de tiempo del servicio 1  
para genero Leyendas

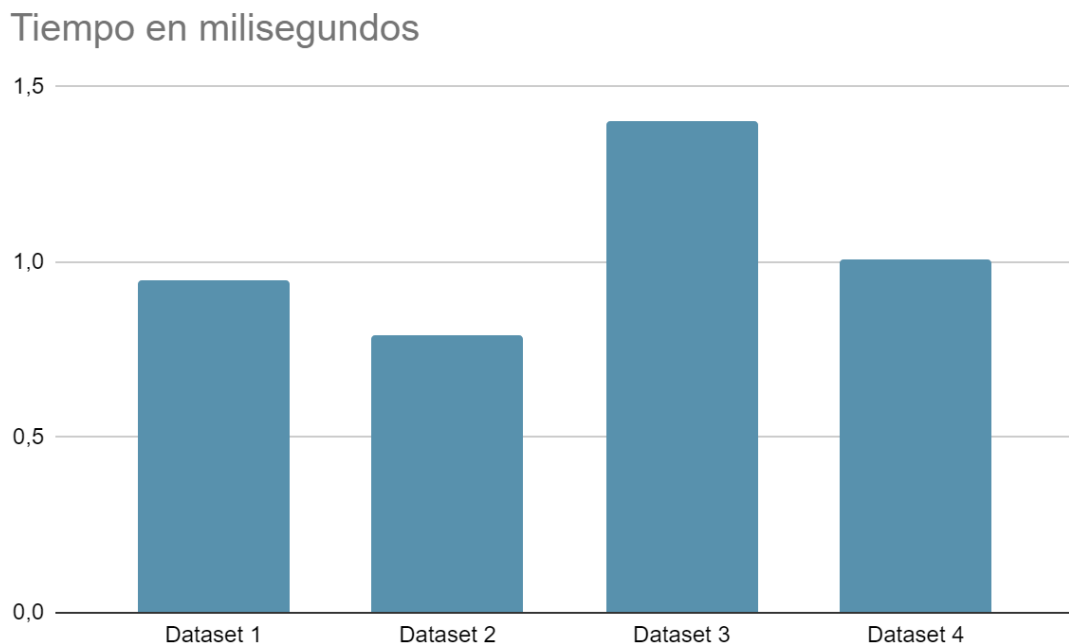


Imagen 1.2 - Gráfico que representa la imagen 1.1

Tiempo genero viajes	
<i>Archivo</i>	<i>Milisegundos para N=3</i>
Dataset 1	0,61
Dataset 2	0,75
Dataset 3	1,05
Dataset 4	1,11

Imagen 1.3 - Análisis de tiempo del servicio 1 para genero Viajes

### Tiempo en milisegundos

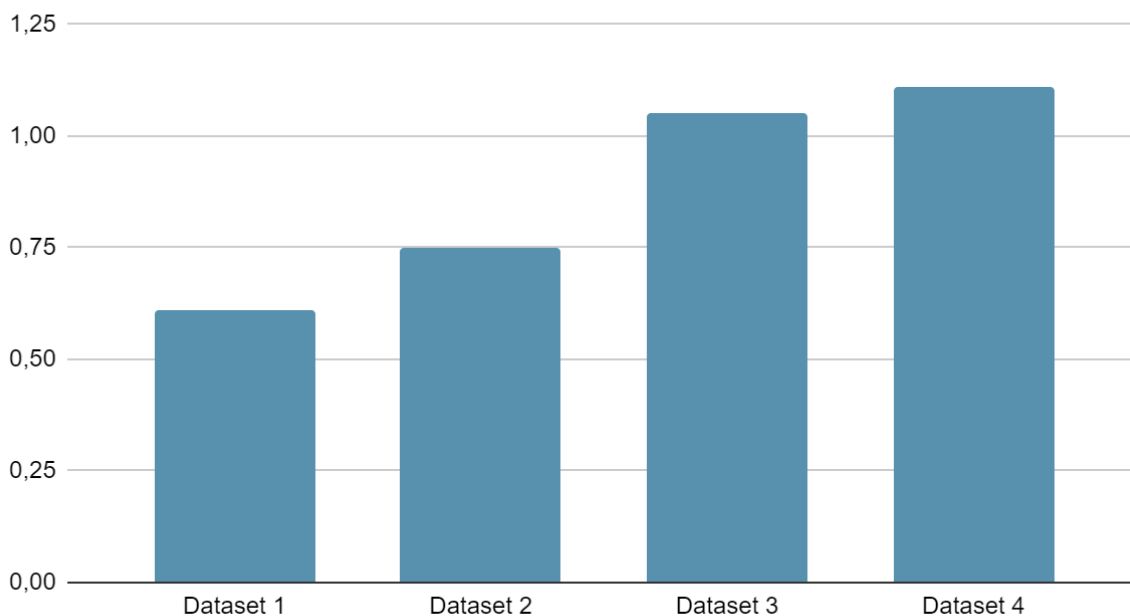


Imagen 1.4 - Gráfico que representa la imagen 1.3

En estos gráficos se ve reflejado el comportamiento en proporción al tiempo. Se puede observar como el tiempo no crece proporcionalmente a la cantidad de datos que contiene cada DataSet (cabe remarcar que el dataset1 es el que tiene menor volumen de datos y el dataset4 es el que más volumen de datos tiene).

Podemos observar en la imagen 1.2, que el Dataset3 cuenta con mayor tiempo que el Dataset4, aunque este último sea superior en tamaño. Esto resulta de esta manera ya que el Dataset3 cuenta con mayor volumen de datos cuando buscamos por el género “viajes”.

A la hora de implementar el servicio 2 realizamos un llamado a nuestra estructura a partir de una clave (género) para obtener todos los arcos que sean parte de este género. Una vez obtenidos los arcos, donde cada uno es un posible candidato a la solución, los ordenamos de mayor a menor (comparando por su etiqueta) para así poder seleccionar el género del arco que se repita mayor cantidad de veces.

Una vez seleccionado el género siguiente, repetiremos este comportamiento hasta quedarnos sin candidatos y así generar la secuencia de géneros que más alto valor de búsqueda posee.

Tiempo genero cine	
<i>Archivo</i>	<i>Milisegundos</i>
Dataset 1	0,89
Dataset 2	3,01
Dataset 3	4,37
Dataset 4	3,65

Imagen 2.1 - Análisis de tiempo del servicio 2  
para genero Cine

### Tiempo en milisegundos

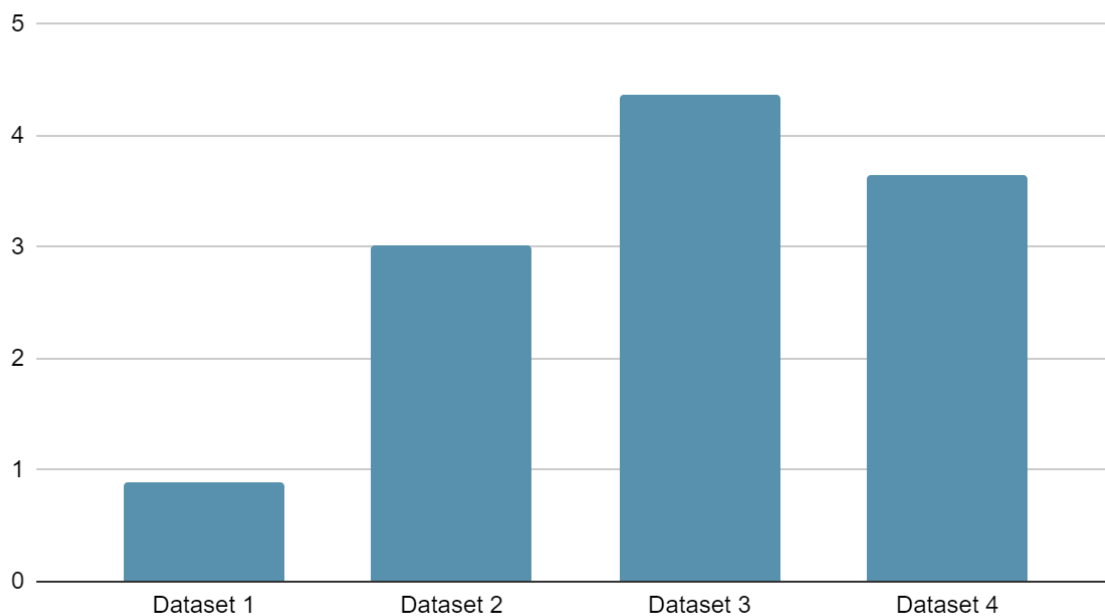


Imagen 2.2 - Gráfico que representa la imagen 2.1

Tiempo genero psicologia	
Archivo	Milisegundos
Dataset 1	0,82
Dataset 2	3,01
Dataset 3	4,19
Dataset 4	4,63

Imagen 2.3 - Análisis de tiempo del servicio 2 para genero Psicología

### Tiempo en milisegundos

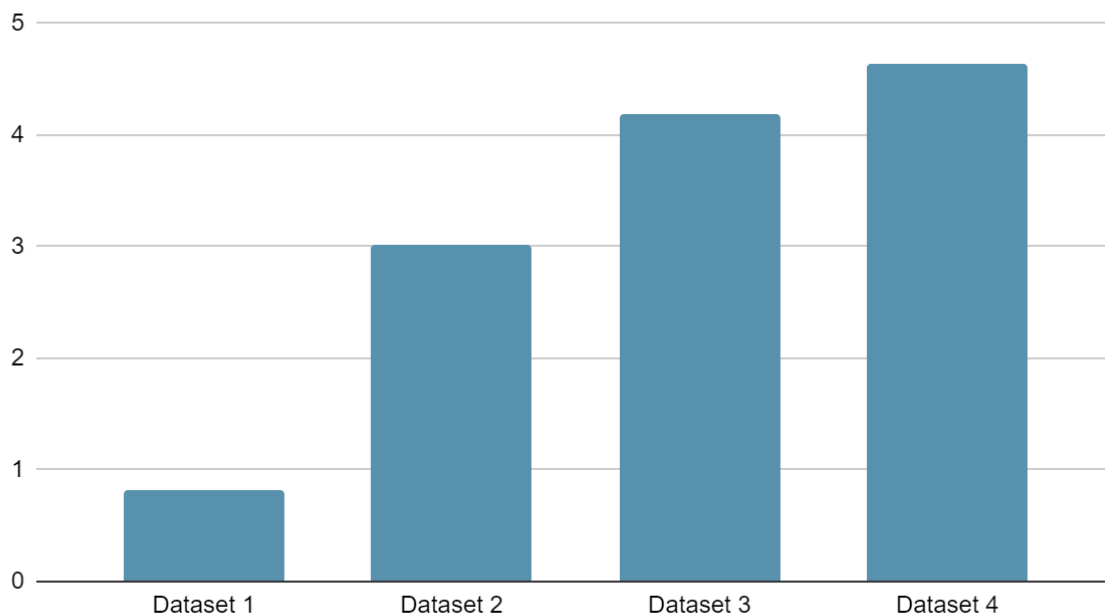


Imagen 2.4 - Gráfico que representa la imagen 2.3

En estos gráficos se puede observar que el comportamiento es similar al del servicio 1, es decir, que el tiempo no va a depender en su totalidad de la cantidad de datos que tenga cada DataSet sino que va a depender más de la cantidad de géneros que componen la secuencia final.

Por ejemplo, si se busca la secuencia de mayor valor de un género x, el género puede no estar en el dataset4 o puede tener una secuencia de dos géneros nada más y quizás en el dataset2, que tiene muchos menos datos que el dataset4, puede tener una secuencia más larga y eso hará que el tiempo de ejecución sea más largo para este último caso que para el primero.

A la hora de implementar el servicio 3, partimos de un género A y realizamos un recorrido del grafo en profundidad a partir de dicho género. A medida que vamos detectando los distintos ciclos, empezaremos a guardar los vértices que componen este ciclo en un arreglo, para así luego iterar sobre este arreglo e ir generando el grafo con sus vértices y arcos.

Tiempo genero psicologia	
Archivo	Milisegundos
Dataset 1	10,23
Dataset 2	13,3
Dataset 3	5,71
Dataset 4	6,67

Imagen 3.1 - Análisis de tiempo del servicio 3  
para género Psicología

Tiempo en milisegundos

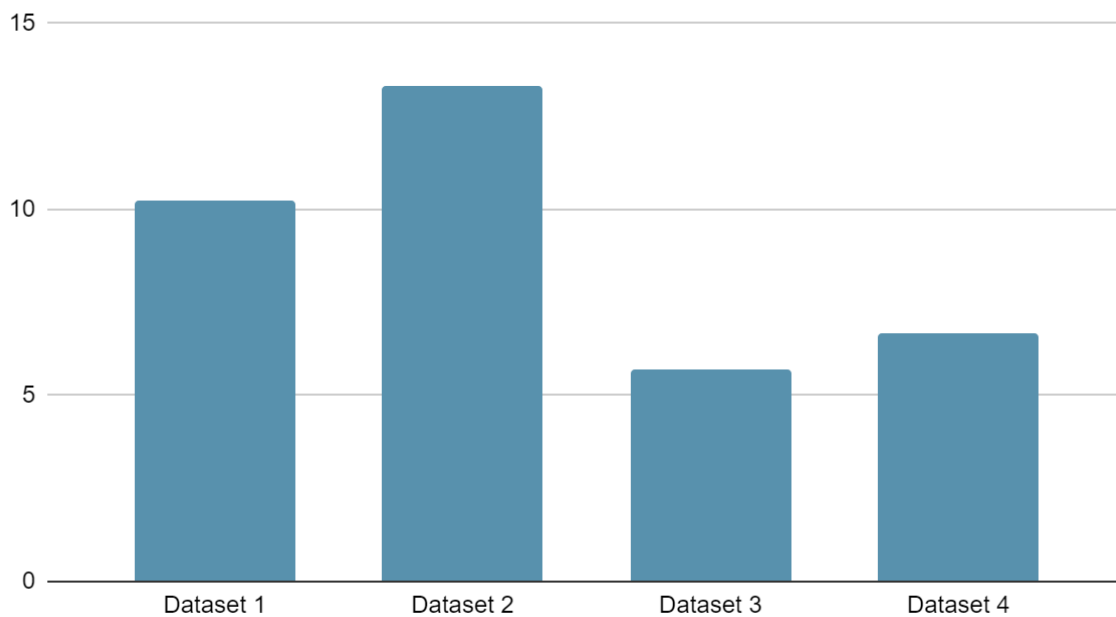


Imagen 3.2 - Gráfico que representa la imagen 3.1

Tiempo genero arte	
<i>Archivo</i>	<i>Milisegundos</i>
Dataset 1	17,74
Dataset 2	14,79
Dataset 3	3,51
Dataset 4	5,77

Imagen 3.3 - Análisis de tiempo del servicio 3 para género arte

### Tiempo en milisegundos

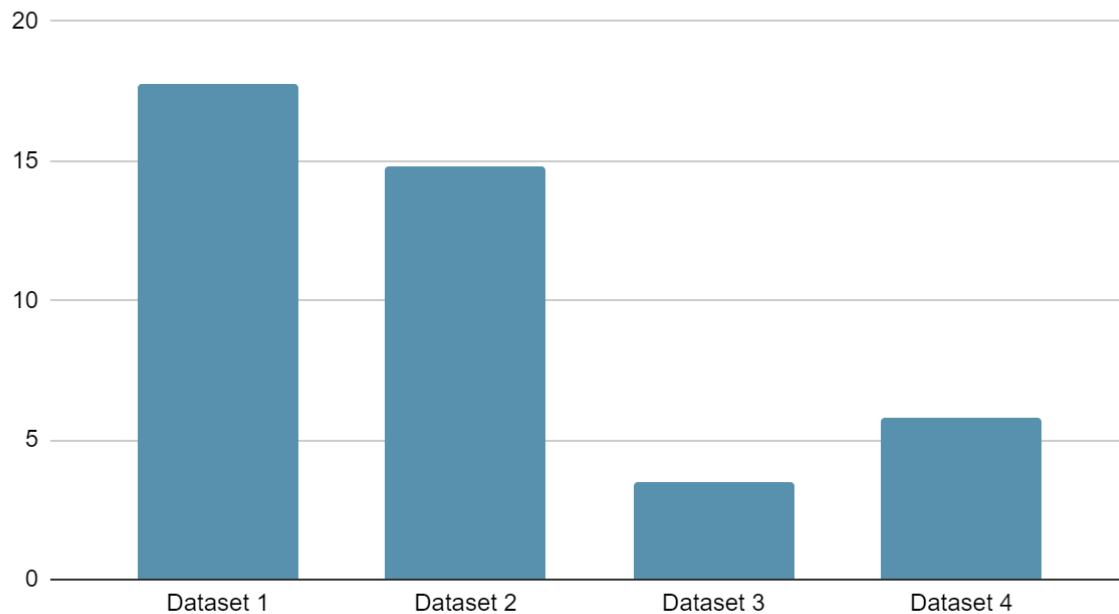


Imagen 3.4 - Gráfico que representa la imagen 3.3

En los anteriores gráficos podemos observar como el tiempo va a aumentar dependiendo de la cantidad de ciclos que contenga el grafo a partir de un género determinado. Va a ocurrir algo similar que en los servicio 1 y 2, donde el tiempo no siempre aumenta con relación al tamaño de los dataset sino que aumenta en relación al contenido buscado.



# Conclusiones

Para finalizar, podemos afirmar que para los servicios implementados los tiempos de búsqueda no van a depender solamente de la cantidad de datos de los archivos de entrada, sino que dependerá más de cómo estarán formados internamente y las veces que participa el género objetivo.

A la hora de hablar del costo de ejecución, al usar una estructura `HashMap<Género, Arcos>`, ésta va a generar un impacto positivo ya que a la hora de buscar los arcos de un género (acción que más veces realizamos) nos permitirá hacerlo a través de una clave y así recuperar los datos que la acompañen, en este caso, los arcos.

La mayor dificultad que obtuvimos fue a la hora de realizar el servicio 3. Sabemos que no es la manera más óptima de realizar el servicio ya que debemos crear el array con los vértices y luego volver a recorrer ese array para crear el grafo.