

ANÁLISIS DEL RETO

(Estudiante 1)= Daniel Pérez, Código:202224568 de.perez@uniandes.edu.co

(Estudiante 2 y 3)=Nicolas Barreto B., código:202320496, n.barretob@uniandes.edu.co

Requerimiento <<5>>

Descripción

Este requerimiento se encarga de retornar:

- . El total de ofertas publicadas en la ciudad en el periodo de consulta.
- El total de empresas que publicaron por lo menos una oferta en la ciudad de consulta.
- Empresa con mayor número de ofertas y su conteo
- Empresa con menor número de ofertas (al menos una) y su conteo
- El listado de ofertas publicadas ordenadas cronológicamente por fecha y nombre de la empresa (v.gr. Para dos ofertas con la misma fecha, el orden lo decide la empresa de forma alfabética). Cada uno

de los elementos debe presentar la siguiente información:

- Fecha de publicación de la oferta
- Título de la oferta
- Nombre de la empresa de la oferta
- Tipo de lugar de trabajo de la oferta
- Tamaño de la empresa de la oferta
- Tipo de lugar de trabajo de la oferta

dado el Nombre de la ciudad, La fecha inicial del periodo a consultar (con formato "%Y-%m-%d"), La fecha final del periodo a consultar (con formato "%Y-%m-%d"), el tipo de lista en que es guardada la información, y la información en sí.

Pasos implementación:

1. # primero organiza la lista bajo el criterio necesitado
2. # luego la filtra y le saca el tamaño para obtener el total de ofertas publicadas en la ciudad en el periodo de consulta, un dato que se retorna al final de la función
3. después crea una lista fantasma en la que se guardan todas las empresas que publicaron, lo que luego es retornado retorna
4. # más tarde se hace una función iterativa para comparar el menor y el mayor de toda la lista, y así guardarlos para que luego sean retornados
5. # Y por último se retorna la lista ordenada inicial para sacar todos sus valores filtrados en el view

Entrada	<code>data_structs, city, fecha_inicial, fecha_final, tipo</code>
Salidas	<code>total_ofertas, dato_empresas, conteo_empresa, max_empresa, conteo_min, min_empresa, lista_ordenada</code>

Implementado (Sí/No)	Sí, lo implemento Nicolas Barreto.
-----------------------------	------------------------------------

Análisis de complejidad

Análisis de complejidad de cada uno de los pasos del algoritmo

Pasos	Complejidad
Paso 1 (mergesort)	$O(n \log(n))$
Paso 1 (mergesort)	$O(n \log(n))$
Paso 2,3 (Partes 1 y 2 del requerimiento)	$O(n)$
Paso 4,5,6 (partes 3,4, y 5 del requerimiento)	$O(n)$

Memoria utilizada extra = $O(n)$

Pruebas Realizadas

Las pruebas realizadas fueron realizadas en una maquina con las siguientes especificaciones. Los datos de entrada fueron el data_structs(carga datos), city (por usuario), fecha_inicial(por usuario), fecha_final(por usuario), tipo (por usuario).

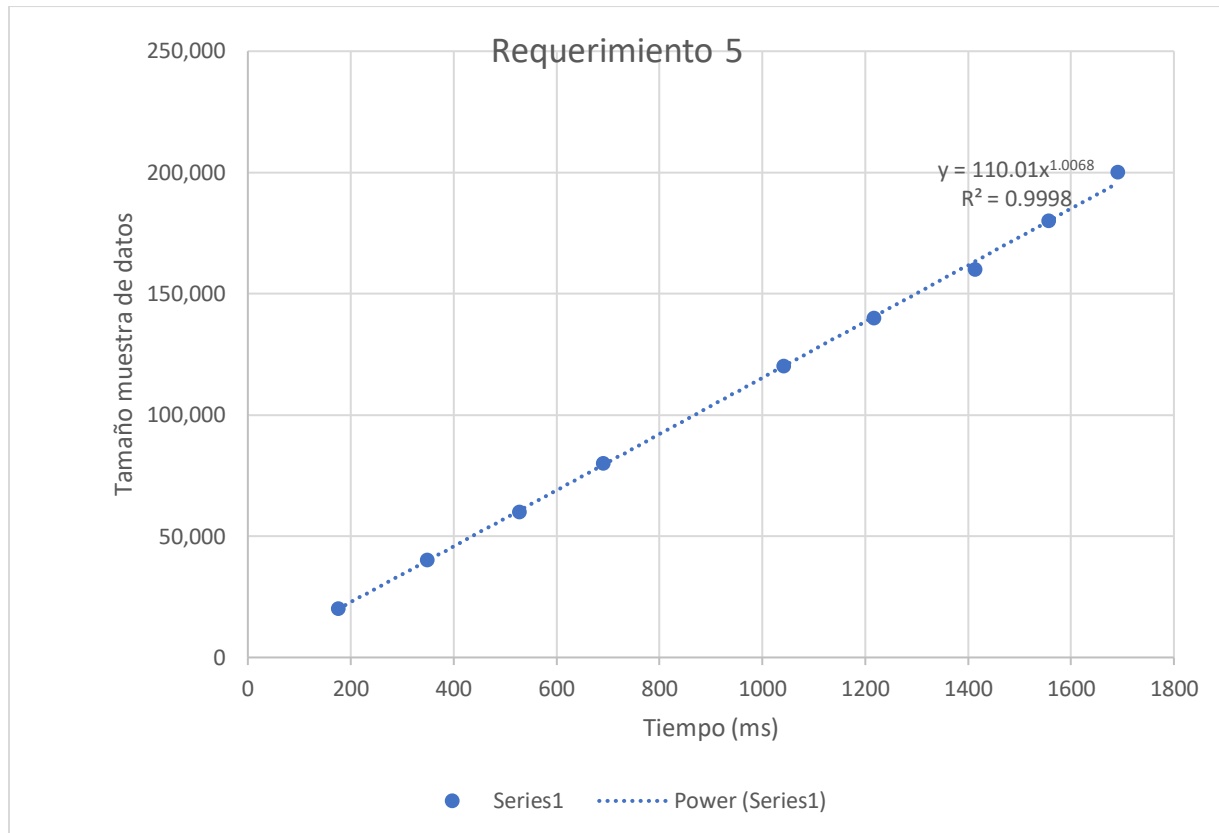
Tablas de datos

Procesadores	Chip M2 Apple
Memoria RAM	8 GB
Sistema Operativo	MacOS

Entrada	Tiempo (ms)
Data_structs(10% del archivo con una muestra de 20,000 elementos), city(Warszawa), fecha inicial(2022-04-14T14:57:00.000Z), fecha final(2023-08-30T09:00:00.000Z), tipo(0 = Array_list)	175.24 ms
Data_structs(20% del archivo con una muestra de 40,000 elementos), city(Warszawa), fecha inicial(2022-04-14T14:57:00.000Z), fecha final(2023-08-30T09:00:00.000Z), tipo(0 = Array_list)	348.29 ms
Data_structs(20% del archivo con una muestra de 60,000 elementos), city(Warszawa), fecha inicial(2022-04-14T14:57:00.000Z), fecha final(2023-08-30T09:00:00.000Z), tipo(0 = Array_list)	527.32 ms

Data_structs(30% del archivo con una muestra de 80,000 elementos), city(Warszawa), fecha inicial(2022-04-14T14:57:00.000Z), fecha final(2023-08-30T09:00:00.000Z), tipo(0 = Array_list)	690.69 ms
Data_structs(40% del archivo con una muestra de 120,000 elementos), city(Warszawa), fecha inicial(2022-04-14T14:57:00.000Z), fecha final(2023-08-30T09:00:00.000Z), tipo(0 = Array_list)	1042.12 ms
Data_structs(50% del archivo con una muestra de 140,000 elementos), city(Warszawa), fecha inicial(2022-04-14T14:57:00.000Z), fecha final(2023-08-30T09:00:00.000Z), tipo(0 = Array_list)	1216.62 ms
Data_structs(60% del archivo con una muestra de 160,000 elementos), city(Warszawa), fecha inicial(2022-04-14T14:57:00.000Z), fecha final(2023-08-30T09:00:00.000Z), tipo(0 = Array_list)	1413.12 ms
Data_structs(70% del archivo con una muestra de 180,000 elementos), city(Warszawa), fecha inicial(2022-04-14T14:57:00.000Z), fecha final(2023-08-30T09:00:00.000Z), tipo(0 = Array_list)	1557.15 ms
Data_structs(100% del archivo con una muestra de 200,000 elementos), city(Warszawa), fecha inicial(2022-04-14T14:57:00.000Z), fecha final(2023-08-30T09:00:00.000Z), tipo(0 = Array_list)	1691.31 ms

Graficas



Análisis

Como nos damos cuenta en la grafica, los datos recopilados tienen una implementación casi perfecta de $O(n \log n)$ de manera que tienen la forma esperada y crecen de forma casi lineal, ya que es una línea multiplicada por un logaritmo. Dado que excel no tiene la funcionalidad de hacer una línea de regresión lineal, se usó una regresión potencial al poder de 1,0068 que modela mejor esta característica. En cuanto a los valores de la tabla, se ve que crecen casi constantemente, ya que se le añade un poquito más de cambio cada vez, lo que demuestra nuestro pensamiento teórico de que la función tiene una complejidad $O(n \log n)$.

Requerimiento <<7>>

Descripción

Este requerimiento se encarga de retornar:

El total de ofertas de empleo

- Número de ciudades donde se ofertó en los países resultantes de la consulta.

- Nombre del país con mayor cantidad de ofertas y su conteo
- Nombre de la ciudad con mayor cantidad de ofertas y su conteo
- Para el conjunto de las ofertas de trabajo en los países resultantes de la consulta, por cada uno de los tres niveles de experticia (junior, mid y senior) calcule y presente la siguiente información:
 - Conteo de habilidades diferentes solicitadas en ofertas de trabajo
 - Nombre de la habilidad más solicitada y su conteo en ofertas de trabajo
 - Nombre de la habilidad menos solicitada y su conteo en ofertas de trabajo
 - Nivel mínimo promedio de las habilidades
 - Conteo de empresas que publicaron una oferta con este nivel
 - Nombre de la empresa con mayor número de ofertas y su conteo
 - Nombre de la empresa con menor número de ofertas (al menos una) y su conteo
 - Número de empresas que publicaron una oferta en este nivel que tienen una o más sedes

Dado el Numero de paises a consultar , La fecha inicial del periodo a consultar (con formato "%Y-%m-%d"), La fecha final del periodo a consultar (con formato "%Y-%m-%d"), el tipo de lista en que es guardada la informacion, y la informacion en si.

Entrada	data_structs, num_paises, fecha_inicial, fecha_final, tipo
Salidas	total_ofertas, total_ciudades, max_pais, cont_pais, max_ciudad, cont_ciudad, habilidades, max_habilidad, max_conteo, min_habilidad, min_conteo, niv_min_promedio, empresas_niv, max_empresas, max_conteoo,min_empresas, min_conteoo, empresas_multilocations
Implementado (Sí/No)	Sí, lo implemento Nicolas Barreto.

Análisis de complejidad

Análisis de complejidad de cada uno de los pasos del algoritmo

Pasos	Complejidad
Paso 1 (mergesort)	$O(n \log(n))$
Paso 2,3,4,5,6,7,8,9, (Partes 1, 2,3,4,5,6,7 y 8 del requerimiento)	$O(n^2)$
Paso 10,11,12,13,14 (partes 9,10,11,12,13 del requerimiento)	$O(n)$
TOTAL	$O(n^2)$

Memoria utilizada extra = $O(n)$

Pruebas Realizadas

Las pruebas realizadas fueron realizadas en una maquina con las siguientes especificaciones. Los datos de entrada fueron el data_structs(carga datos), num (por usuario), fecha_inicial(por usuario), fecha_final(por usuario), tipo (por usuario).

Tablas de datos

Procesadores	Chip M2 Apple
Memoria RAM	8 GB
Sistema Operativo	MacOS

Entrada	Tiempo (ms)
Data_structs(10% del archivo con una muestra de 20,000 elementos), num_paises (5), fecha inicial(2022-04-14T14:57:00.000Z), fecha final(2023-08-30T09:00:00.000Z), tipo(0 = Array_list)	

Graficas

Análisis

#Nota: Este laboratorio lo hizo Nicolas Barreto casi que completamente solo... por favor considerarlo al calificar.