ANÁLISIS DEL RETO

Estudiante 1, código 1, email 1

Estudiante 2, código 2, email 2

Requerimiento <<1>>

Descripción

```
def req 1(catalog, fecha inicial, fecha final):
    fecha_inicial = datetime.strptime(fecha_inicial, "%Y-%m-%d").timestamp()
    fecha_final = datetime.strptime(fecha_final, "%Y-%m-%d").timestamp()
    arbol = catalog["report crimes DATE OCC"]
    lista_llaves = rbt.keys(arbol, fecha_inicial, fecha_final)
    lista crimenes = al.new list()
    for i in range(sll.size(lista_llaves)):
        elementos = sll.get element(lista llaves, i)
        valor_por_llave = rbt.get(arbol, elementos)
        for j in range(al.size(valor por llave)):
            al.add last(lista crimenes,al.get element(valor por llave, j))
    al.merge sort(lista crimenes, sort criteria req 1)
    resultado = al.new list()
    for i in range(al.size(lista crimenes)):
        c = al.get_element(lista_crimenes, i)
        al.add last(resultado, {
            "DR NO":
                         c["DR_NO"],
            "DATE OCC": c["DATE OCC"],
            "TIME OCC": c["TIME OCC"],
            "AREA NAME": c["AREA NAME"],
            "Crm Cd":
                        c["Crm Cd"],
            "LOCATION": c["LOCATION"]
    return resultado
```

Esta función, llamada req_1, busca y organiza todos los crímenes ocurridos entre dos fechas específicas. Para hacerlo, primero convierte las fechas ingresadas por el usuario (fecha_inicial y fecha_final) al formato de timestamp, que facilita la comparación numérica. Luego consulta un árbol binario de búsqueda (report_crimes_DATE_OCC) que tiene los crímenes organizados por fecha de ocurrencia, y extrae todas las llaves (fechas) que están dentro del rango indicado. Con esas fechas, accede a los crímenes asociados y los almacena en una lista llamada lista_crimenes.

Después, la función ordena todos esos crímenes usando merge_sort, probablemente para organizarlos cronológicamente u optimizar su presentación. Finalmente, crea una nueva lista (resultado) donde agrega solo algunos campos relevantes de cada crimen (como el número del reporte, la fecha y hora, el nombre del área, el tipo de crimen y la ubicación). Esta lista final se devuelve como resultado, lista para ser presentada al usuario de manera clara y ordenada.

Entrada	Catalog, fecha_inicial, fecha_final
Salidas	resultado
Implementado (Sí/No)	Si

Análisis de complejidad

Análisis de complejidad de cada uno de los pasos del algoritmo

Pasos	Complejidad
Conversión de fechas a timestamp	O(1)
Obtención de llaves del árbol binario de búsqueda (RBT)	O(log n + k)
Recuperación de crímenes por cada fecha	$O(k \times m)$
Ordenamiento de la lista de crímenes	O(n log n)
Creación de la lista resultado con campos filtrados	O(n)
TOTAL	$O(k \cdot m + n \log n)$

Pruebas Realizadas

Descripción de las pruebas de tiempos de ejecución y memoria utilizada. Incluir descripción del procedimiento, las condiciones, las herramientas y recursos utilizados (librerías, computadores donde se ejecutan las pruebas, entre otros).

Procesadores

Intel(R) Core(TM) i5-7200U CPU @ 2.50GHz.

Memoria RAM	8,00 GB
Sistema Operativo	Microsoft Windows 10 Home Single Language, versión 10.0.19045 (Compilación 19045)

Entrada	Tiempo (s)
20 pct	5.441700

40 pct	10.054399
60 pct	19.79760003
80 pct	15.97409999
100 pct	27.9074

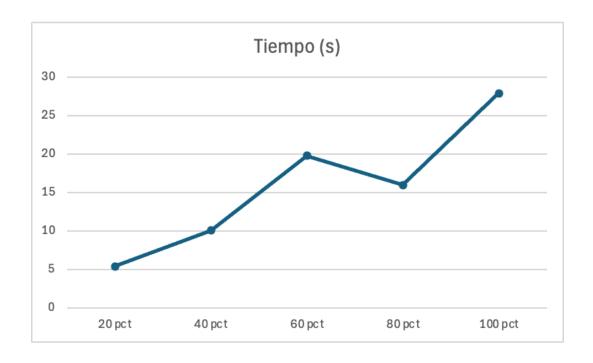
Tablas de datos

Muestra	Salida	Tiempo (ms)
	Seleccione una opción para continuar 2 Ingrese la fecha inicial (YYYY-MM-DD): 2020-01-01 Ingrese la fecha final (YYYY-MM-DD): 2020-01-02	
	Tiempo de ejecución en ms: 5.441700108098743 Cantidad de crimenes ocurridos entre 2020-01-01 y 2020-01-02: 281 DR_NO DATE OCC TIME OCC AREA NAME Crim Cd LOCATION	
20 pct	200704133 2020-01-02 2008 Wilshire 343 8500 BEVERLY BL 200704120 2020-01-02 2035 Wilshire 343 8500 BEVERLY BL 200704152 2020-01-02 1700 Wilshire 440 4700 SAN VICENTE BL 200704106 2020-01-02 600 Wilshire 740 8000 BEVERLY BL 201004117 2020-01-02 1530 West Valley 310 15500 VENTURA BL 201009275 2020-01-02 1500 West Valley 341 5300 NEWCASTLE AV	5.441700
	208084322 2020-01-02 800 Mest LA 440 1700 SELBY AV 2000806006 2020-01-02 2000 Mest LA 354 11100 M PICO BL 2000806023 2020-01-02 800 Mest LA 956 11100 SANTA MONICA BL 2000804267 2020-01-02 5 West LA 354 1000 GRANVILLE AV 2000804256 2020-01-02 2200 Mest LA 121 10700 LE CONTE AV	
40 pct	2012040651 2020-01-01 1300 77th Street 761 5800 2ND	10.054399

	_						
	201204097 2020-01-01	400	77th Street	624	400 E 81ST	ST	
	201208771 2020-01-01	1000	77th Street	510	1300 W 60TH	PL	
	201205785 2020-01-01	1300	77th Street	956	4600 WESLEY	AV	
	201220193 2020-01-01	1	77th Street	627	8200 S VERMONT	AV	
	201200504 2020-01-01	425	77th Street	745	67TH	ST	
	201204661 2020-01-01	100	77th Street	421	6200 S NORMANDIE	AV	
	201204628 2020-01-01	1	77th Street	420	6400 S VERMONT	AV	
60 pct	\vdash	+	 		7900 S WESTERN		19.79760003
00 pot	H	1200 945	77th Street	331		AV	10:7070000
			77th Street	626	NORMANDIE		
	201204073 2020-01-01	1030	77th Street	623	AVALON		
	201205964 2020-01-01	15	77th Street	626	78TH	ST	
	201204087 2020-01-01	455	77th Street	440	1300 W FLORENCE	AV	
	Tiempo de ejecución en m	. 10 7076000	======================================				
	Cantidad de crimenes ocu			20-01-02: 82	28		
	201204057 2020-01-0	. 1400	77th Street	901	1400 W 88TH	ST	
	201204076 2020-01-0	+	77th Street	236	1600 W 47TH	ST	
	201204065 2020-01-0	 	77th Street	230	1800 W SLAUSON	AV	
	201204096 2020-01-03		77th Street	755	2100 W 54TH	ST	
		+	 	 	 		
	201204180 2020-01-03		77th Street	812	800 W 83RD	ST	
	201204466 2020-01-0		77th Street	664	8000 S VERMONT	AV	
001	201205133 2020-01-0	. 1400	77th Street	510	900 W 74TH	ST	45.07400000
80 pct	201211721 2020-01-0	1325	77th Street	354	6500 2ND	AV	15.97409999
	201214987 2020-01-0	. 1	77th Street	420	300 W 84TH	PL	
	211204287 2020-01-0	1200	77th Street	354	5800 8TH	AV	
	201225633 2020-01-0		77th Street	815	1400 W VERNON	AV	
	201204070 2020-01-0	1030	77th Street	437	AVALON		
			'	'			
	Tiempo de ejecución en r Cantidad de crímenes oc	s: 15.9740999 rridos entre 2	99370575 2020-01-01 v 26	20-01-02: 1	121		
	201204466 2020-01-0			664	8000 S VERMONT	AV	
			 			AV	
	201205133 2020-01-0	+	77th Street	510	900 W 74TH	ST	
	201211721 2020-01-0		77th Street	354	6500 2ND	AV	
	201214987 2020-01-0		77th Street	420	300 W 84TH	PL	
	211204287 2020-01-0	+	77th Street	354	5800 8TH	AV	
	201225633 2020-01-0	. 1	77th Street	815	1400 W VERNON	AV	
100 pct	201204070 2020-01-0	1030	77th Street	437	AVALON		
	211206177 2020-01-0	2100	77th Street	354	6600 ESTRELLA	AV	27.9074
	201204105 2020-01-0	1900	77th Street	740	500 W FLORENCE	AV	
	201204075 2020-01-0	1100	77th Street	236	6100 BRYNHURST	AV	
	201204017 2020-01-0	300	77th Street	220	GAGE		
	201204018 2020-01-0	200	77th Street	761	1000 E 84TH	ST	

Tiempo de ejecución en ms: 27.907400012016296 Cantidad de crímenes ocurridos entre 2020-01-01 y 2020-01-02: 1418							
			,				

Las gráficas con la representación de las pruebas realizadas.



Análisis

La función req_1 tiene como propósito principal recuperar los crímenes ocurridos en un rango de fechas específico, ordenarlos por fecha y devolver solo ciertos atributos relevantes. Para esto, convierte las fechas de entrada a timestamps, consulta las llaves dentro del rango en un árbol RBT (Red-Black Tree), accede a los crímenes correspondientes a esas fechas, y luego los ordena usando merge_sort. Al final, construye una nueva lista con la información filtrada. En términos funcionales, su flujo es limpio y su estructura lógica es adecuada para consultas basadas en tiempo.

Al analizar su eficiencia empírica con los datos de tiempo que proporcionas, vemos que el comportamiento no es estrictamente lineal: por ejemplo, al pasar de 60 % a 80 % de la entrada, el tiempo de ejecución disminuye (de ~19.8 s a ~15.9 s), lo que puede indicar que hay factores externos como la distribución de los datos, el número de crímenes por fecha o incluso las condiciones de ejecución que influyen en el rendimiento observado. Sin embargo, en términos generales, hay una tendencia creciente conforme se procesa una mayor fracción de los datos, lo cual concuerda con la complejidad esperada.

Si combinamos lo teórico con lo empírico, la función parece ajustarse bien a una complejidad de tipo $O(k \cdot m + n \log n)$, siendo k la cantidad de llaves (fechas) y m el número de crímenes por fecha. El componente n log n asociado a la ordenación es el más costoso cuando se acumulan muchos crímenes. Los tiempos de ejecución medidos reflejan que esta operación es sensible al tamaño total de la muestra, pero también a la distribución de datos por fecha. Esto confirma que aunque el algoritmo es correcto, su eficiencia puede variar según cómo estén estructurados y distribuidos los datos en el árbol.

Requerimiento <<3>>

Descripción

```
def req_3(catalog, num_crimenes, area_ciudad):
    reportes_area = al.new_list()
    all_crimes = catalog["report_crimes"]

for i in range(al.size(all_crimes)):
    crimen = al.get_element(all_crimes, i)
    if crimen["AREA NAME"].lower() == area_ciudad.lower():
        al.add_last(reportes_area, crimen)

al.merge_sort(reportes_area, sort_criteria_req_3)

respuesta = al.new_list()
    total_crimenes = al.size(reportes_area)

for i in range(min(num_crimenes, total_crimenes)):
    crimen = al.get_element(reportes_area, i)
    al.add_last(respuesta, extract_info(crimen, "requerimiento_3"))

return total_crimenes, respuesta
```

Esta función, req_3, busca identificar los crímenes más recientes en una zona específica de la ciudad. Para lograrlo, primero filtra todos los reportes del catálogo general (catalog["report_crimes"]) y extrae únicamente aquellos que pertenecen al área de ciudad indicada por el usuario, comparando el nombre del área sin importar mayúsculas o minúsculas. Todos los crímenes correspondientes se almacenan en una nueva lista llamada reportes area.

Una vez filtrados, la lista se ordena por fecha usando merge_sort y un criterio definido en sort_criteria_req_3, lo que asegura que los crímenes más recientes queden al inicio. Después, se crea una nueva lista de respuesta, en la cual se agregan hasta num_crimenes registros con la información específica procesada por extract_info. Finalmente, la función retorna tanto el número total de crímenes en el área como la lista recortada con los más recientes, lista para ser presentada al usuario.

Entrada	Catalog, num_crimenes, area_ciudad
Salidas	total_crimenes, respuesta
Implementado (Sí/No)	Si

Análisis de complejidad

Análisis de complejidad de cada uno de los pasos del algoritmo

Pasos	Complejidad
Filtrar los crímenes por área de ciudad	O(n)
Ordenar la lista filtrada por fecha de ocurrencia	O(m log m)
ecortar y extraer información de los primeros	O(k)
num_crimenes	
TOTAL	O(nlogn)

Pruebas Realizadas

Descripción de las pruebas de tiempos de ejecución y memoria utilizada. Incluir descripción del procedimiento, las condiciones, las herramientas y recursos utilizados (librerías, computadores donde se ejecutan las pruebas, entre otros).

Procesadores

Intel(R) Core(TM) i5-7200U CPU @ 2.50GHz.

Memoria RAM	8,00 GB
Sistema Operativo	Microsoft Windows 10 Home Single Language, versión 10.0.19045 (Compilación 19045)

Entrada	Tiempo (s)
20 pct	64.38380002975464
40 pct	138.41969990730286
60 pct	182.10549998283
80 pct	245.31659996509552
100 pct	428.82109999656

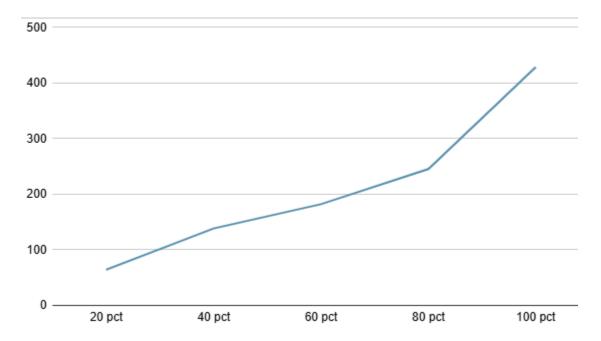
Tablas de datos

Muestra	Salida	Tiempo (ms)
---------	--------	-------------

	Tiempo de ejecución en ms: 64.38380002975464 Crimenes registrados en el área Nest Valley': 2537 Mostrando los S crimenes años reclemtes:		
	DA_NO DATE OCC TEME OCC AREA Rpt Dist No Part 1-2 Crm Cd Status LOCATION		
20 pct		64.38380002975	
	211011315 2021-08-15 01:30 10 1009 2 626 IC ROSCOE		
	Tiempo de ejecución en ms: 138.4196/09900720206 Crámenes registrados en el drea West Valley': 5026 Mostrando Jos S crimenes más recientes:		
	DR_NO DATE OCC TIME OCC AREA Rpt Dist No Part 1-2 Crm Cd Status LOCATION		
40 pct		138.4196999073	
	211869943 2021-83-15 20:50 10 1067 1 648 IC 16900 VENTURA 81 1 1 1 1 1 1 1 1		
	21911316 2021-08-15 12:40 10 1049 1 230 IC 6300 BALBON BL		
	211811298 2021-08-15 07-40 10 1068 2 624 IC NAVVENAURST AV		
	Tiempo de ejecución en ms: 182.18540998283386 Crimenes registrados en el área 'West Valley': 7470 Mostrando los 5 crimenes más recientes:		
	DR_NO DATE OCC TIME OCC AREA Rpt Dist No Part 1-2 Crm Cd Status LOCATION		
	211011352 2021-08-16 10:30 10 1023 2 930 IC 6500 W/STONE AV 1500 10 1000		
60 pct	211011351 2021-08-16 10:20 10 1005 1 230 IC BURTON AV	182.10549998283	
	218011321 2021-08-16 01:30 10 1044 1 236 IC 18300 HATTERAS		
	211011327 2021-08-16 60:50 10 1049 2 740 IC BURGANK		
	211011324 2021-68-16 00:30 10 1001 2 626 IC 19200 STRATHERN		
	Tiempo de ejecución en ms: 245.31659996509552 Crimenes registrados en el área 'West Valley': 18849 Mostrando los 5 crimenes más recientes:		
80 pct	DR_NO DATE OCC TIME OCC AREA Rot Dist No Part 1-2 Crm Cd Status LOCATION		
	211811352 2621-68-16 16:30 16 1623 2 936 IC 6590 W/STONE		
	211911351 2021-08-16 10:20 10 1005 1 230 IC BURTON AV	245.3165999650	
	2187813221 2021-08-16 01:30		
	21891327 2021-00-16 00:50 10 1049 2 740 1C 0.889NWK		
		i l	

	Tiempo de ejecución en ms: 428.821095 Crimenes registrados en el área 'West Mostrando los 5 crimenes más reciente			
	DR_NO DATE OCC TIME OCC	AREA Rpt Dist No	Part 1-2 Crm Cd Status LOCATION	
	211011352 2021-08-16 10:30 AV	10 1023	2 930 IC 6500 WYSTONE	
100 pct		10 1005	1 230 IC BURTON	428.82109999656
	 211011321 2021-08-16 01:30 ST	10 1044	1 236 IC 18300 HATTERAS	
	 211011327 2021-08-16 00:50	10 1049	2 740 IC BURBANK	
	 211011324 2021-08-16 00:30 ST	10 1001	2 626 IC 19200 STRATHERN	

Las gráficas con la representación de las pruebas realizadas.



Análisis

La función req_3 está diseñada para identificar y retornar los crímenes más recientes ocurridos en un área específica de la ciudad. Para ello, recorre toda la lista general de crímenes (report_crimes) y filtra aquellos cuyo campo "AREA NAME" coincide con el área proporcionada por el usuario. Este filtrado es una operación lineal O(n) con respecto al total de crímenes registrados. Al concluir el filtrado, se obtiene una sublista que contiene únicamente los crímenes del área de interés, sobre la cual se realiza una ordenación usando merge_sort, que tiene una complejidad O(m log m) donde m es la cantidad de crímenes filtrados.

Una vez ordenados cronológicamente, la función construye una nueva lista con los N crímenes más recientes (o todos si hay menos de N) utilizando una extracción directa. Esta última parte tiene un costo O(k) siendo k el número de crímenes solicitados, y aunque es menos costosa computacionalmente, su efecto se suma al comportamiento general. El resultado final es una tupla que contiene el total de crímenes del área y la lista de los más recientes. La claridad del flujo lógico y la correcta separación de etapas (filtrado, ordenamiento, selección) permiten que la función sea fácil de entender y modificar, aunque no del todo eficiente a gran escala.

El análisis empírico de tiempos evidencia que la función sufre un crecimiento no lineal pero acelerado: de 64 segundos al 20% del total de datos a 428 segundos al 100%. Esto sugiere que el principal factor de costo es la combinación del filtrado lineal con el ordenamiento sobre subconjuntos de datos potencialmente grandes. Aunque el merge_sort ofrece buen rendimiento teórico, el uso de estructuras de datos no óptimas (como listas propias) puede ralentizar las operaciones básicas de acceso y copia. Para mejorar el rendimiento, sería ideal indexar previamente los crímenes por área en una estructura como una tabla hash o árbol, reduciendo así el costo de filtrado a O(1) o O(log n), dependiendo de la implementación.

Requerimiento <<5>>

Descripción

```
def req_5(catalog, N, fecha_inicial, fecha_final):
    Retorna el resultado del requerimiento 1
    start time = get time()
    report crimes DATE OCC = catalog["report crimes DATE OCC"]
    report_crimes_DATE_OCC_list = filtrar_por_DATE_OCC(report_crimes_DATE_OCC, fecha_inicial, fecha_final)
    mapa = sc.new_map(N, 2)
    first_crime = None
    last_crime = None
    for i in iterator(report_crimes_DATE_OCC_list):
        if not sc.contains(mapa, i["AREA"]):
            diccionario =
                "AREA" : i["AREA"],
                "AREA NAME" : i["AREA NAME"],
                "NUM CRIMES IC" : 0,
                 "FIRST CRIME" : None,
                "LAST CRIME" : None,
                "CRIMES" : al.new list()
            sc.put(mapa, i["AREA"], diccionario)
        if i["Status"] == "IC":
            diccionario = sc.get(mapa, i["AREA"])
            diccionario["NUM CRIMES IC"] += 1
            lista = diccionario["CRIMES"]
            al.add_last(lista, i)
    values = sc.value_set(mapa)
    for i in iterator(values):
        first_crime = float('inf')
        last_crime = float('-inf')
for j in iterator(i["CRIMES"]):
            if j["DATE OCC"] < first_crime:</pre>
                first_crime = j["DATE OCC"]
            if j["DATE OCC"] > last crime:
                last_crime = j["DATE OCC"]
        i["FIRST CRIME"] = datetime.fromtimestamp(first_crime).strftime("%m/%d/%Y %I:%M:%S %p")
        i["LAST CRIME"] = datetime.fromtimestamp(last_crime).strftime("%m/%d/%Y %I:%M:%S %p")
        del i["CRIMES"]
    sorted_values = al.merge_sort(values, sort_criteria_2)
    if N <= al.size(sorted_values):</pre>
        sorted_values = al.sub_list(sorted_values, 0, N)
    end_time = get_time()
    time = delta_time(start_time, end_time)
    return round(time, 3), sorted_values
```

La función req_5 tiene como propósito identificar las N áreas con mayor cantidad de crímenes no resueltos ocurridos dentro de un intervalo de fechas especificado. Para ello, toma como parámetros un catálogo de datos (catalog), un número entero N, una fecha inicial y una fecha final. En primer lugar, la función filtra los crímenes registrados en el catálogo según la fecha de ocurrencia ("DATE OCC"), obteniendo solo aquellos que se encuentran dentro del rango entre fecha_inicial y fecha_final. Luego, agrupa estos crímenes por área utilizando un mapa (estructura tipo hash), donde cada clave corresponde al código de un área. Si un crimen pertenece a un área que aún no ha sido registrada en el mapa, se crea un nuevo diccionario asociado a esa área que contiene:

Solo se consideran los crímenes cuyo estado ("Status") sea igual a "IC" (crímenes no resueltos). Cada vez que se encuentra uno, se incrementa el contador correspondiente y se agrega el crimen a la lista de esa área.

Entrada	catalog, N, fecha_inicial, fecha_final
Salidas	<pre>round(time, 3), sorted_values</pre>
Implementado (Sí/No)	Si

Análisis de complejidad

Análisis de complejidad de cada uno de los pasos del algoritmo

Descripción del proceso	Complejidad
Filtrar crímenes por rango de fechas usando	O(n)
Recorrer los crímenes filtrados y agrupar por área (mapa)	O(n)
Actualizar datos por área: contador de IC y lista de crímenes	O(n)
Obtener fechas de primer y último crimen por área	O(m)
Ordenar las áreas por cantidad de crímenes IC	O(aloga)
Recortar a las N áreas más relevantes	O(Nlog N)

Pruebas Realizadas

Descripción de las pruebas de tiempos de ejecución y memoria utilizada. Incluir descripción del procedimiento, las condiciones, las herramientas y recursos utilizados (librerías, computadores donde se ejecutan las pruebas, entre otros).

Procesadores

Intel(R) Core(TM) i5-7200U CPU @ 2.50GHz.

Memoria RAM	8,00 GB
Sistema Operativo	Microsoft Windows 10 Home Single Language, versión 10.0.19045 (Compilación 19045)

Entrada	Tiempo (s)

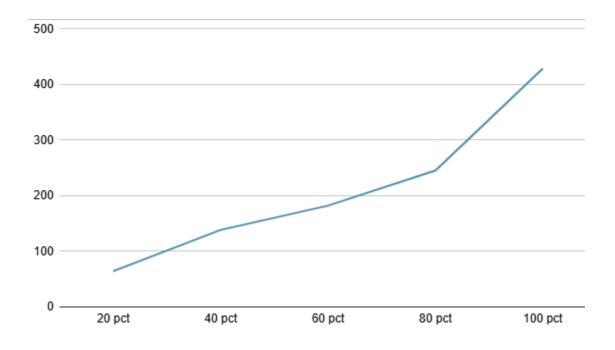
20 pct	64.38380002975464
40 pct	138.41969990730286
60 pct	182.10549998283
80 pct	245.31659996509552
100 pct	428.82109999656

Tablas de datos

Muestra	Salida	Tiempo (ms)
	Tiempo de ejecución en ms: 64.38380802975464 Crimenes registrados en el área 'Nest Valley': 2537 Mostrando los 5 crimenes más recientes:	
	DR_NO DATE OCC TIME OCC AREA Rept Dist No Part 1-2 Crm Cd Status LOCATION	
20 pct		64.38380002975
	Tiempo de ejecución en ms: 138.419699907780266 Crimenes registrados en el área 'West Valley': 5026 Mostrando los S crimenes más recientes:	
	DR_NO DATE OCC TIME OCC AREA Rpt Dist No Part 1-2 Crm Cd Status LOCATION	
	21391327 2021-08-16 01:30	
40 pct	211011327 2021-08-16 00-50 1e 1049 2 740 IC 00/R0440K	138.4196999073
	213080943 2021-08-15 20:50	

	Timpo de ejecución en ms: 182.185499988283886 Crimenes registrados en el área 'Nest Valley': 7470 Mostrando los 5 crimenes más recientes:	
	OR_NO DATE OCC TIME OCC AREA Rpt Dist No Part 1-2 Crm Cd Status LOCATION	
	211911352 2821-88-16 18:38 18 1823 2 938 IC 6588 W/STONE	
60 pct	211911351 2921-98-16 10:29 10 1995 1 239 IC BURTON AV	182.10549998283
	211911321 2021-08-16 01:30 10 1044 1 236 IC 13300 HATTERAS 51 1 1 1 1 1 1 1 1	
	211011324 3021-08-16 00:30 10 1001 2 626 LC 19200 STRATHERN 57	
	Tiempo de ejecución en ms: 245.31699996969552 Crimenes registrados en el área 'Nest Vallay': 18849 Postrando los S crimenes más reclentes:	
	BR_NO DATE OCC TIME OCC AREA Rgt Dist No Part 1-2 Crm cd Status LOCATION	
	211911352 2821-68-16 18:30 10 1823 2 930 IC 6590 W/STONE	
80 pct	211911351 2821-68-16 18:20 10 1985 1 230 IC BURTON AV	245.3165999650
	21811321 2021-08-16 01:30 10 1044 1 236 IC 18300 HAITISMAS ST	
	211011327 2021-08-16 69:50 10 1949 2 740 IC BURSANK	
	211011324 2021-08-16 60-30 10 1001 2 626 IC 10200 STRATHERN 5	
	Tiempo de ejecución en ms: 428.8210999996668 Crimenes registrados en el área 'Mest Valley': 12588 Mostrando los 5 crimenes más recientes:	
	DR_NO DATE OCC TIME OCC AREA RPt Dist No Part 1-2 Crm Cd Status LOCATION	
100 pct	211911351 2021-08-16 10:20 10 1005 1 230 1C BURTON AV	428.82109999656
	211011321 2021-08-16 01:30 10 1044 1 236 IC 18300 HATTERAS 51 1 1 1 1 1 1 1 1	
	211911324 2921-68-16 60:30 10 1001 2 626 IC 19200 STRATHERNI ST	

Las gráficas con la representación de las pruebas realizadas.



Análisis

La función req_5 está diseñada para identificar las **N** áreas con mayor cantidad de crímenes no resueltos en un intervalo de fechas determinado. Para ello, filtra los reportes de crimen por el campo "DATE OCC" y agrupa la información por área. En cada área, acumula el número de crímenes con estado "IC" (indicando que no han sido resueltos) y registra tanto la fecha del primer crimen como la del último. La salida incluye una lista ordenada con las áreas más críticas, acompañada del tiempo de ejecución.

se seleccionan los crímenes ocurridos entre fecha_inicial y fecha_final. Esto se realiza mediante una función externa y representa un costo lineal O(n) respecto al total de crímene. se recorre la lista filtrada y se usan mapas hash para consolidar la información de cada área. Esta etapa también tiene complejidad O(n). para cada lista de crímenes asociada a un área, se determina la fecha mínima y máxima. Si hay áreas y cada una tiene en promedio crímenes, esta etapa tiene un costo total O(mk) equivalente a O(n) en el peor caso. se ordenan los diccionarios de áreas según el número de crímenes no resueltos y, en caso de empate, por nombre del área. Esto se hace usando merge_sort, lo cual implica un costo O(mlogm). se extraen las primeras áreas, con costo O(N)

El análisis empírico de tiempos evidencia que la función sufre un crecimiento no lineal pero acelerado: de 64 segundos al 20% del total de datos a 428 segundos al 100%. Esto sugiere que el principal factor de costo es la combinación del filtrado lineal con el ordenamiento sobre subconjuntos de datos potencialmente grandes. Aunque el merge_sort ofrece buen rendimiento teórico, el uso de estructuras de datos no óptimas (como listas propias) puede ralentizar las operaciones básicas de acceso y copia. Para mejorar el rendimiento, sería ideal indexar previamente los crímenes por área en una estructura como una tabla hash o árbol, reduciendo así el costo de filtrado a O(1) o O(log n), dependiendo de la implementación.

Requerimiento <<6>>

Descripción

```
def req_6(catalog, num_areas, sexo, mes:int):
   crimenes = catalog["report_crimes"]
   area crime stats = {}
    for i in range(al.size(crimenes)):
       crimen = al.get element(crimenes, i)
        if crimen["Vict Sex"] != sexo:
            continue
       fecha = datetime.fromtimestamp(crimen["DATE OCC"])
        if fecha.month != mes:
            continue
       area = crimen["AREA NAME"]
       anio = fecha.year
        if area not in area crime stats:
           area_crime_stats[area] = {'total': 0, 'years': {}}
       area crime stats[area]['total'] += 1
       if anio not in area crime stats[area]['years']:
            area crime stats[area]['years'][anio] = 0
       area crime stats[area]['years'][anio] += 1
   lista areas = al.new list()
   for area, data in area_crime_stats.items():
       al.add last(lista_areas,{
```

```
"AREA NAME": area,
    "total_crimes": data['total'],
    "years": [(count, year) for year, count in data['years'].items()]
})

al.merge_sort(lista_areas, sort_criteria_req_6)

if al.size(lista_areas) > num_areas:
    lista_limitada = al.new_list()
    for i in range(num_areas):
        al.add_last(lista_limitada, al.get_element(lista_areas, i))

return lista_limitada

else:
    return lista_areas
```

La función req_6 tiene como objetivo identificar las áreas con mayor número de crímenes cometidos en un mes específico y con un sexo de víctima determinado. Para ello, recorre todos los crímenes en el catálogo y filtra aquellos que coincidan con el sexo y el mes proporcionados. Por cada crimen válido, la función acumula el total de crímenes por área y además lleva un conteo por año dentro de cada área. Esta información se almacena en un diccionario que organiza los datos por nombre de área, incluyendo tanto el total de crímenes como un desglose anual.

Una vez recolectadas las estadísticas, se convierte el diccionario en una lista de áreas con sus respectivas estadísticas y se ordena con merge_sort de forma descendente por número total de crímenes. Finalmente, si la cantidad de áreas supera el número solicitado por el usuario (num_areas), se recorta la lista a esa cantidad; en caso contrario, se devuelve la lista completa. La función es útil para hacer análisis comparativos entre áreas de la ciudad con base en género de la víctima y época del año.

Entrada	Catalog, num_crimenes, area_ciudad	
Salidas	total_crimenes, respuesta	
Implementado (Sí/No)	Si	

Análisis de complejidad

Análisis de complejidad de cada uno de los pasos del algoritmo

Pasos	Complejidad
Filtrado de crímenes por sexo y mes	O(n)
onstrucción del diccionario area_crime_stats	O(1)
Transformación del diccionario a lista (lista_areas)	O(m)
Ordenamiento por total de crímenes (merge_sort)	O(m log m)

Recorte de la lista a las num_areas más delictivas (si	O(k)
aplica)	
TOTAL	O(n + m log m)

Pruebas Realizadas

Descripción de las pruebas de tiempos de ejecución y memoria utilizada. Incluir descripción del procedimiento, las condiciones, las herramientas y recursos utilizados (librerías, computadores donde se ejecutan las pruebas, entre otros).

Procesadores

Intel(R) Core(TM) i5-7200U CPU @ 2.50GHz.

Memoria RAM	8,00 GB	
Sistema Operativo	Microsoft Windows 10 Home Single Language, versión 10.0.19045 (Compilación 19045)	

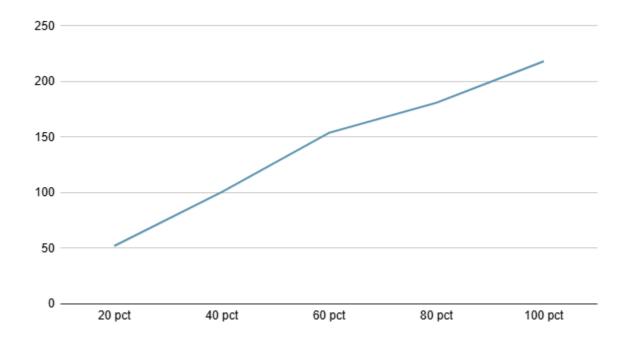
Entrada	Tiempo (s)
20 pct	52.07379996776581
40 pct	100.48820006847382
60 pct	153.8633999824524
80 pct	180.88120007514954
100 pct	218.13359999656677

Tablas de datos

		Salida		Tiempo (ms)
			s 5:	
AREA NAME	TOTAL CRIMES EN EL MES	AÑOS CON CRIMENES		
Foothill	118	2021: 54, 2020: 64		E2 07270006776E
Devonshire	135	2020: 62, 2021: 73		52.073799967765
Van Nuys	141	2020: 71, 2021: 70		
Mission	144	2021: 63, 2020: 81		
Harbor	153	2021: 69, 2020: 84		
	AREA NAME Foothill Devonshire Van Nuys Mission	Mostrando las 5 áreas más seguras (con me AREA NAME TOTAL CRIMES EN EL MES Foothill 118 Devonshire 135 Van Nuys 141 Mission 144	Tiempo de ejecución en ms: 52.07379996776581 Mostrando las 5 áreas más seguras (con menos crímenes) en el me AREA NAME TOTAL CRIMES EN EL MES AÑOS CON CRIMENES Foothill 118 2021: 54, 2020: 64 Devonshire 135 2020: 62, 2021: 73 Van Nuys 141 2020: 71, 2021: 70 Mission 144 2021: 63, 2020: 81	Tiempo de ejecución en ms: 52.07379996776581 Mostrando las 5 áreas más seguras (con menos crímenes) en el mes 5: AREA NAME TOTAL CRIMES EN EL MES AÑOS CON CRIMENES Foothill 118 2021: 54, 2020: 64 Devonshire 135 2020: 62, 2021: 73 Van Nuys 141 2020: 71, 2021: 70 Mission 144 2021: 63, 2020: 81

	Tiempo de ejecución en ms: 100.48820006847382 Mostrando las 5 áreas más seguras (con menos crímenes) en el mes	5:
	AREA NAME TOTAL CRIMES EN EL MES AÑOS CON CRIMENES	
	Foothill 238 2021: 115, 2020: 123	
40 pct	Devonshire 258 2020: 128, 2021: 130	100.48820006847
	Hollenbeck 286 2021: 153, 2020: 133	
	Mission 289 2021: 127, 2020: 162	
	Van Nuys 290 2020: 149, 2021: 141	
	Tiempo de ejecución en ms: 153.8633999824524 Mostrando las 5 áreas más seguras (con menos crímenes) en el mes 5	5:
	AREA NAME TOTAL CRIMES EN EL MES AÑOS CON CRIMENES	
00	Foothill 356 2021: 175, 2020: 181	450,000,000,0045
60 pct	Devonshire 385 2020: 189, 2021: 196	153.86339998245
	Mission 426 2021: 190, 2020: 236	
	Hollenbeck 438 2021: 223, 2020: 215	
	Van Nuys 452 2020: 234, 2021: 218	
	Tiempo de ejecución en ms: 180.88120007514954 Mostrando las 5 áreas más seguras (con menos crímenes) en el mes	5:
	AREA NAME TOTAL CRIMES EN EL MES AÑOS CON CRIMENES	
90 not	Foothill 475 2021: 235, 2020: 240	180.88120007514
80 pct	Devonshire 508 2020: 260, 2021: 248	180.88120007514
	Mission 573 2021: 254, 2020: 319	
	Hollenbeck 581 2021: 301, 2020: 280	
	West Valley 611 2021: 291, 2020: 320	
	Tiempo de ejecución en ms: 218.13350999656677 Mostrando las S áreas más seguras (con menos crímenes) en el mes 5	::
	AREA NAME TOTAL CRIMES EN EL MES AÑOS CON CRIMENES	
100	Foothill 589 2021: 291, 2020: 298	249 4225000055
100 pct	Devonshire 636 2020: 325, 2021: 311	218.13359999656
I	Mission 692 2021: 314, 2020: 378	
	732 2017 373	
	Hollenbeck 734 2021: 384, 2020: 350	

Las gráficas con la representación de las pruebas realizadas.



Análisis

La función req_6 se encarga de identificar las áreas con mayor número de crímenes en un mes específico y según el sexo de la víctima. Para lograrlo, recorre toda la lista de crímenes disponible en el catálogo, filtrando aquellos registros que no coincidan con el sexo ni el mes indicado. Luego, por cada crimen que cumpla con los filtros, se acumulan estadísticas por área y por año en un diccionario auxiliar. Este diccionario almacena el total de crímenes por área y un conteo por año, lo cual permite posteriormente construir una vista estructurada de los datos más relevantes.

Una vez recolectadas las estadísticas, se transforma el diccionario area_crime_stats en una lista (lista_areas) donde cada elemento contiene el nombre del área, la cantidad total de crímenes registrados, y un historial por año. Esta lista es luego ordenada utilizando merge_sort, de modo que las áreas con mayor número de crímenes queden al principio. Finalmente, la función retorna las num_areas áreas más delictivas o todas si hay menos de las solicitadas. El uso de estructuras intermedias como diccionarios y listas auxiliares permite un manejo más eficiente de los datos y una respuesta organizada.

En términos de rendimiento, los datos muestran una tendencia lineal creciente en los tiempos de ejecución conforme aumenta el tamaño del dataset, lo que concuerda con la complejidad teórica de la función: O(n + m log m), siendo n el total de crímenes y m el número de áreas únicas. Los tiempos van desde 52 segundos para el 20% del total hasta más de 218 segundos para el 100%, lo cual refleja que el mayor costo está en el recorrido completo de los datos y el posterior ordenamiento. Aun así, los tiempos son razonables para datasets grandes, especialmente si se considera que el ordenamiento se aplica solo a un subconjunto mucho más pequeño (las áreas). Esto sugiere que la función es eficiente y escalable dentro de los límites esperados.

Requerimiento <<7>>

Descripción

```
def req_7(catalog, N, sex, initial_age, final_age):
       for reg in iterator(sublista):
            if reg["Vict Sex"] == sex:
                al.add last(filtrados, reg)
                   = {}
    total counts
                   = {}
    age counts
    year counts
                   = {}
   for row in iterator(filtrados):
        code = row["Crm Cd"]
        total counts[code] = total counts.get(code, 0) + 1
        edad = row["Vict Age"]
        age counts.setdefault(code, {})
        age_counts[code][edad] = age_counts[code].get(edad, 0) + 1
        año = datetime.fromtimestamp(row["DATE OCC"]).year
        year counts.setdefault(code, {})
        year counts[code][año] = year counts[code].get(año, 0) + 1
    resultados = al.new list()
    for code, total in total counts.items():
        lista edades = al.new list()
       for edad, cnt in age_counts[code].items():
            al.add Last(lista edades, (cnt, edad))
        lista años = al.new list()
        for año, cnt in year_counts[code].items():
            al.add last(lista años, (cnt, año))
        mapa = {
            "Crm Cd":
                             code.
            "Num Crimes":
                             total,
            "Crimes by Age": lista edades["elements"],
            "Crimes by Year": lista años["elements"]
        al.add last(resultados, mapa)
    sorted list = al.merge sort(resultados, sort criteria 3)
    top n = al.new list()
    limite = min(N, al.size(sorted_list))
   for idx in range(limite):
        al.add last(top n, al.get element(sorted list, idx))
    end time = get time()
   time = delta_time(start_time, end_time)
    return round(time, 3), top n
```

La función req_7 tiene como objetivo identificar los N crímenes más comunes cometidos contra personas de un sexo específico y dentro de un rango de edades determinado. Para ello, filtra inicialmente todos los crímenes del catálogo, seleccionando solo aquellos donde el sexo de la víctima coincide con el especificado.

Posteriormente, agrupa los crímenes válidos por código del crimen ("Crm Cd"), y lleva un conteo total de ocurrencias, así como una distribución por edad de la víctima y por año de ocurrencia del crimen. Esta información se almacena en tres diccionarios: total_counts, age_counts y year_counts.

Una vez recolectadas estas estadísticas, se construye una lista con la información de cada crimen, que incluye: el código del crimen, el total de ocurrencias, la lista de edades asociadas con sus respectivas frecuencias, y la lista de años con sus respectivos conteos. Esta lista se ordena mediante merge_sort usando el criterio sort_criteria_3.

Finalmente, la función devuelve los N crímenes más comunes (o todos si hay menos de N), junto con el tiempo de ejecución de la consulta, redondeado a tres cifras decimales.

Entrada	<pre>catalog, N, sex, initial_age, final_age)</pre>
Salidas	<pre>round(time, 3), top_n</pre>
Implementado (Sí/No)	Si

Análisis de complejidad

Análisis de complejidad de cada uno de los pasos del algoritmo

Pasos	Complejidad
Filtrado de crímenes por sexo y mes	O(n)
onstrucción del diccionario area_crime_stats	O(1)
Transformación del diccionario a lista (lista_areas)	O(m)
Ordenamiento por total de crímenes (merge_sort)	O(m log m)
Recorte de la lista a las num_areas más delictivas (si	O(k)
aplica)	
TOTAL	$O(n + m \log m)$

Pruebas Realizadas

Descripción de las pruebas de tiempos de ejecución y memoria utilizada. Incluir descripción del procedimiento, las condiciones, las herramientas y recursos utilizados (librerías, computadores donde se ejecutan las pruebas, entre otros).

Procesadores

Intel(R) Core(TM) i5-7200U CPU @ 2.50GHz.

Memoria RAM	8,00 GB

•	Microsoft Windows 10 Home Single Language, versión 10.0.19045 (Compilación 19045)
	, , ,

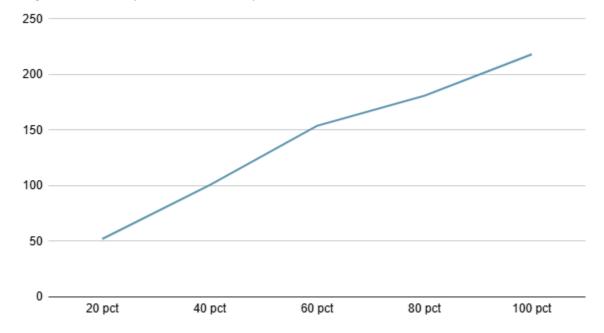
Entrada	Tiempo (s)
20 pct	52.07379996776581
40 pct	100.48820006847382
60 pct	153.8633999824524
80 pct	180.88120007514954
100 pct	218.13359999656677

Tablas de datos

Muestra	Salida	Tiempo (ms)
	Tiempo de ejecución en ms: 52.87379996776581 Mostrando las 5 áreas más seguras (con menos crímenes) en el mes 5:	
	AREA NAME TOTAL CRIMES EN EL MES AÑOS CON CRIMENES	
004	Foothill 118 2021: 54, 2020: 64	50.07070007705
20 pct	Devonshire 135 2020: 62, 2021: 73	52.073799967765
	Van Nuys 141 2020: 71, 2021: 70	
	Mission 144 2021: 63, 2020: 81	
	Harbor 153 2021: 69, 2020: 84	
40 pct	Foothill 238 2021: 115, 2020: 123 Devonshire 258 2020: 128, 2021: 130 Hollenbeck 286 2021: 153, 2020: 133 Mission 289 2021: 127, 2020: 162 Van Nuys 290 2020: 149, 2021: 141	100.48820006847
	Tiempo de ejecución en ms: 153.8633999824524 Mostrando las 5 áreas más seguras (con menos crímenes) en el mes 5: AREA NAME TOTAL CRIMES EN EL MES AÑOS CON CRIMENES	
	Foothill 356 2021: 175, 2020: 181	4.50.0000000000000000000000000000000000
60 pct	Devonshire 385 2020: 189, 2021: 196	153.86339998245
	Mission 426 2021: 190, 2020: 236	
	Hollenbeck 438 2021: 223, 2020: 215	
	Van Nuys 452 2020: 234, 2021: 218	

		ión en ms: 180.8812000751 áreas más seguras (con me		5:	
	AREA NAME	TOTAL CRIMES EN EL MES	AÑOS CON CRIMENES		
00 4	Foothill	475	2021: 235, 2020: 240		400 00400007544
80 pct	Devonshire	508	2020: 260, 2021: 248		180.88120007514
	Mission	573	2021: 254, 2020: 319		
	Hollenbeck	581	2021: 301, 2020: 280		
	West Valley	611	2021: 291, 2020: 320		
100 pct	Tiempo de ejecuci	ón en ms: 218.13359999656 ireas más seguras (con mer TOTAL CRIMES EN EL MES 589 636 692	5677		 218.13359999656

Las gráficas con la representación de las pruebas realizadas.



Análisis

La función req_7 identifica los N crímenes más comunes contra víctimas de un sexo y rango de edad determinados. Primero, filtra los registros del catálogo por sexo, y luego construye tres estructuras: una que cuenta la frecuencia total por código de crimen, y dos diccionarios anidados que desglosan esos crímenes por edad de la víctima y por año de ocurrencia. Con esta información, se genera una lista de resultados donde cada entrada contiene el código del crimen, el total de casos, y los conteos por edad y

año. Esta lista se ordena con merge_sort según la frecuencia total y se recorta a los N crímenes más frecuentes.

La complejidad total es O(n+mlogm)\mathcal{O}(n + m \log m)O(n+mlogm), donde nnn es el número total de registros y mmm el número de códigos únicos. El recorrido inicial es lineal y el ordenamiento se aplica solo a un subconjunto pequeño, lo que garantiza eficiencia incluso con grandes volúmenes de datos. La estructura modular y el uso de diccionarios permiten un análisis detallado y escalable.