

# ANÁLISIS DEL RETO

Valentina España Cuellar 202414079, v.espana@uniandes.edu.co

Juan Sebastian Cortes Cortes 202411692, js.cortesc12@uniandes.edu.co

Tomas Alarcón Martínez Troncoso 202420126, t.alarcon1@uniandes.edu.co

## Carga de datos

### Descripción

```
1 def Load_data(catalog, filename):
2     """
3     Carga los datos del reto
4     """
5     # TODO: Realizar la carga de datos
6     nombre = data_dir + filename
7     archivo = csv.DictReader(open(nombre, encoding="utf-8"))
8     for registro in archivo:
9         fecha = datetime.strptime(registro["Load_time"], "%Y-%m-%d %H:%M:%S")
10        registro_data = {
11            "source": registro["source"],
12            "commodity": registro["commodity"],
13            "statical_category": registro["statical_category"],
14            "unit_measurement": registro["unit_measurement"],
15            "state_name": registro["state_name"],
16            "location": registro["location"],
17            "year_collection": registro["year_collection"],
18            "freq_collection": registro["freq_collection"],
19            "reference_period": registro["reference_period"],
20            "load_time": fecha,
21            "value": registro["value"]
22        }
23        al.add_last(catalog["registros"], registro_data)
24        update_bystate(catalog, registro_data)
25        update_bycategory(catalog, registro_data)
26        update_byyear(catalog, registro_data)
27        update_bycommodity(catalog, registro_data)
28
29    al.merge_sort(catalog["registros"], comp_fechas)
30
31    return catalog
```

Se crea un diccionario el cual es el catálogo, este contiene un array list con todos los registros y cuatro mapas divididos por ESTADO, CATEGORÍA, AÑO Y PRODUCTO para así facilitar el ingreso a los datos en los requerimientos.

Entrada	Catalogo, Nombre del archivo
Salidas	Catalogo
Implementado (Sí/No)	Implementado grupalmente

Análisis de complejidad

Pasos	Complejidad
Crear registros y añadirlos a los mapas	$O(n)$
Ordenar los registros en el array list con Merge Sort	$O(n \log n)$
<b>TOTAL</b>	<b><math>O(n \log n)</math></b>

Procesadores	12th Gen Intel(R) Core(TM) i7-1255U
Memoria RAM	32 GB
Sistema Operativo	Windows 11

Entrada	Tiempo (ms)
Agricultural 20%	2914.42
Agricultural 40%	6062.26
Agricultural 60%	9253.83
Agricultural 80%	13228.91
Agricultural 100%	16045.41

Tablas de datos

20% de la muestra:  
El tiempo de ejecución es: 2914.4183999970555  
Se cargaron 100000 registros  
El mayor año de recolección de registro fue: 2024

Primeros 5 registros:

Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medición	Tipo del producto	Valor de medición
SURVEY	2024	2024-11-25	WEEKLY	FLORIDA	PCT FAIR	CATTLE	33
SURVEY	2024	2024-11-25	WEEKLY	GEORGIA	PCT POOR	CATTLE	6
SURVEY	2024	2024-11-25	WEEKLY	NEW MEXICO	PCT ADEQUATE	FEED	21
SURVEY	2024	2024-11-25	WEEKLY	SOUTH CAROLINA	PCT FAIR	CATTLE	31
SURVEY	2024	2024-11-25	MONTHLY	ARKANSAS	HEAD	CHICKENS	192

Ultimos 5 registros:

Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medición	Tipo del producto	Valor de medición
SURVEY	1952	2012-01-01	MONTHLY	WYOMING	LB / HEAD, LIVE BASIS	CATTLE	950
SURVEY	1993	2012-01-01	POINT IN TIME	WYOMING	HEAD	CATTLE	12,000
SURVEY	2011	2011-05-31	POINT IN TIME	NORTH CAROLINA	HEAD	CHICKENS	135,000
SURVEY	2011	2011-05-31	POINT IN TIME	OKLAHOMA	HEAD	CHICKENS	3,700
SURVEY	2011	2011-05-31	POINT IN TIME	OKLAHOMA	HEAD	CHICKENS	1,700

40% de la muestra:

El tiempo de ejecucion es: 6062.26439999789

Se cargaron 20000 registros

El mayor año de recoleccion de registro fue: 2024

Primeros 5 registros:

Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
SURVEY	2024	2024-11-25	WEEKLY	FLORIDA	PCT FAIR	CATTLE	33
SURVEY	2024	2024-11-25	WEEKLY	GEORGIA	PCT POOR	CATTLE	6
SURVEY	2024	2024-11-25	WEEKLY	NEW MEXICO	PCT ADEQUATE	FEED	21
SURVEY	2024	2024-11-25	WEEKLY	SOUTH CAROLINA	PCT FAIR	CATTLE	31
SURVEY	2024	2024-11-25	WEEKLY	WEST VIRGINIA	PCT GOOD	SHEEP	15

Ultimos 5 registros:

Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
SURVEY	2011	2011-09-21	MONTHLY	OTHER STATES	HEAD	CATTLE	59,000
SURVEY	2011	2011-05-31	POINT IN TIME	NORTH CAROLINA	HEAD	CHICKENS	135,000
SURVEY	2011	2011-05-31	POINT IN TIME	OKLAHOMA	HEAD	CHICKENS	3,700
SURVEY	2011	2011-05-31	POINT IN TIME	OKLAHOMA	HEAD	CHICKENS	1,700
SURVEY	2011	2011-05-31	POINT IN TIME	OKLAHOMA	HEAD	CHICKENS	2,600

60% de la muestra:

El tiempo de ejecucion es: 9253.82840000093

Se cargaron 30000 registros

El mayor año de recoleccion de registro fue: 2024

Primeros 5 registros:

Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
SURVEY	2024	2024-11-25	WEEKLY	FLORIDA	PCT FAIR	CATTLE	33
SURVEY	2024	2024-11-25	WEEKLY	GEORGIA	PCT POOR	CATTLE	6
SURVEY	2024	2024-11-25	WEEKLY	MISSISSIPPI	PCT VERY POOR	LIVESTOCK TOTALS	2
SURVEY	2024	2024-11-25	WEEKLY	NEW MEXICO	PCT ADEQUATE	FEED	21
SURVEY	2024	2024-11-25	WEEKLY	SOUTH CAROLINA	PCT FAIR	CATTLE	31

Ultimos 5 registros:

Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
SURVEY	2011	2011-09-21	MONTHLY	OTHER STATES	HEAD	CATTLE	59,000
SURVEY	2011	2011-05-31	POINT IN TIME	NORTH CAROLINA	HEAD	CHICKENS	135,000
SURVEY	2011	2011-05-31	POINT IN TIME	OKLAHOMA	HEAD	CHICKENS	3,700
SURVEY	2011	2011-05-31	POINT IN TIME	OKLAHOMA	HEAD	CHICKENS	1,700
SURVEY	2011	2011-05-31	POINT IN TIME	OKLAHOMA	HEAD	CHICKENS	2,600

80% de la muestra:

El tiempo de ejecucion es: 13228.90849999711

Se cargaron 40000 registros

El mayor año de recoleccion de registro fue: 2024

Primeros 5 registros:

Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
SURVEY	2024	2024-11-25	WEEKLY	ALABAMA	PCT POOR	CATTLE	6
SURVEY	2024	2024-11-25	WEEKLY	COLORADO	PCT NORMAL	CATTLE	68
SURVEY	2024	2024-11-25	WEEKLY	FLORIDA	PCT FAIR	CATTLE	33
SURVEY	2024	2024-11-25	WEEKLY	GEORGIA	PCT POOR	CATTLE	6
SURVEY	2024	2024-11-25	WEEKLY	MISSISSIPPI	PCT VERY POOR	LIVESTOCK TOTALS	2

Ultimos 5 registros:

Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
SURVEY	2011	2011-05-31	POINT IN TIME	OKLAHOMA	HEAD	CHICKENS	3,700
SURVEY	2011	2011-05-31	POINT IN TIME	OKLAHOMA	HEAD	CHICKENS	1,700
SURVEY	2011	2011-05-31	POINT IN TIME	OKLAHOMA	HEAD	CHICKENS	2,600
SURVEY	2011	2011-05-31	POINT IN TIME	OKLAHOMA	HEAD	CHICKENS	52,000
SURVEY	2011	2011-05-31	POINT IN TIME	SOUTH CAROLINA	HEAD	CHICKENS	2,200

100% de la muestra:

El tiempo de ejecucion es: 16045.407299999148  
Se cargaron 50000 registros  
El mayor año de recoleccion de registro fue: 2024

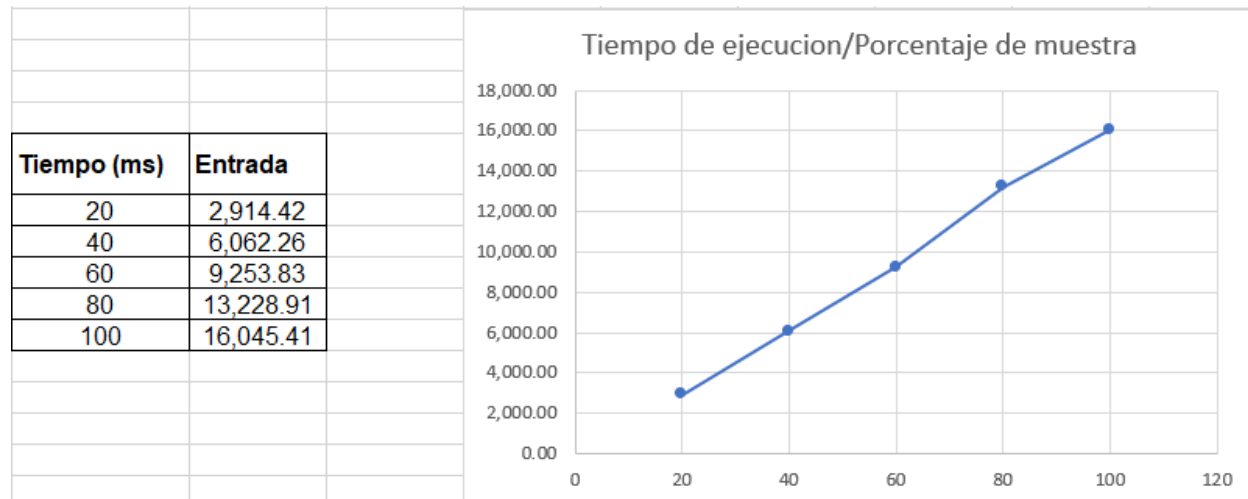
Primeros 5 registros:

Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
SURVEY	2024	2024-11-25	WEEKLY	ALABAMA	PCT POOR	CATTLE	6
SURVEY	2024	2024-11-25	WEEKLY	COLORADO	PCT NORMAL	CATTLE	68
SURVEY	2024	2024-11-25	WEEKLY	COLORADO	PCT HEAVY	SHEEP	0
SURVEY	2024	2024-11-25	WEEKLY	FLORIDA	PCT FAIR	CATTLE	33
SURVEY	2024	2024-11-25	WEEKLY	GEORGIA	PCT POOR	CATTLE	6

Ultimos 5 registros:

Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
SURVEY	2011	2011-05-31	POINT IN TIME	OKLAHOMA	HEAD	CHICKENS	1,700
SURVEY	2011	2011-05-31	POINT IN TIME	OKLAHOMA	HEAD	CHICKENS	2,600
SURVEY	2011	2011-05-31	POINT IN TIME	OKLAHOMA	HEAD	CHICKENS	52,000
SURVEY	2011	2011-05-31	POINT IN TIME	SOUTH CAROLINA	HEAD	CHICKENS	2,200
SURVEY	2011	2011-05-31	ANNUAL	SOUTH CAROLINA	HEAD	TURKEYS	2,700,000

## Gráfica



## Análisis

La carga de datos presenta una complejidad de  $O(n \log n)$ , resultado de insertar los registros en los mapas y ordenarlos con merge sort. Los resultados experimentales muestran un crecimiento coherente con la teoría, sin llegar a tiempos de ejecución tan grandes como se presentarían si no se usaran los mapas, que beneficio tanto el rendimiento de la carga de datos, como el resto de requerimientos.

# Requerimiento 1 (Grupal)

## Descripción

El Requerimiento No. 1 implementa una funcionalidad que permite identificar el último registro recopilado para un año de interés específico, basado en la fecha de carga del registro a la plataforma. La implementación utiliza una tabla de hash (catalog["map\_by\_year"]) para almacenar los registros indexados por año de recolección y una lista dinámica para manejar los registros asociados a cada año.

```
def req_1(catalog, year):
    """
    Retorna el resultado del requerimiento 1
    """
    # TODO: Modificar el requerimiento 1
    start_time = get_time()

    registros = lp.get(catalog["map_by_year"], year)

    if not registros or al.size(registros) == 0:
        return {"tiempo_ejecucion": delta_time(start_time, get_time()), "total_registros": 0, "ultimo_registro": None}

    registros = al.merge_sort(registros, comp_fechas)

    ultimo_registro = al.get_element(registros, 0)

    return {
        "tiempo_ejecucion": delta_time(start_time, get_time()),
        "total_registros": al.size(registros),
        "ultimo_registro": {
            "Año de recolección": ultimo_registro["year_collection"],
            "Fecha de carga": ultimo_registro["load_time"].strftime("%Y-%m-%d"),
            "Fuente": ultimo_registro["source"],
            "Frecuencia": ultimo_registro["freq_collection"],
            "Departamento": ultimo_registro["state_name"],
            "Producto": ultimo_registro["commodity"],
            "Unidad de medición": ultimo_registro["unit_measurement"],
            "Valor": ultimo_registro["value"]
        }
    }
```

Entrada	Catalog, year
Salidas	tiempo_ejecucion, total_registros, ultimo_registro, Año de recolección, Fecha de carga, Fuente, Frecuencia, Departamento, Producto, Unidad de medición, Valor
Implementado (Sí/No)	Implementado grupalmente

## Análisis de complejidad

Análisis de complejidad de cada uno de los pasos del algoritmo

Pasos	Complejidad
Consultar la tabla de hash para obtener la lista de registros del año.	O(1)
Verificar si la lista está vacía (al.size(registros))	O(1)
Ordenar los registros por fecha de carga con Merge Sort.	O(n log n)

Obtener el primer elemento de la lista ordenada (al.get_element).	$O(1)$
Construir y retornar el diccionario de salida.	$O(1)$
<b>TOTAL</b>	<b><math>O(n \log n)</math></b>

## Pruebas Realizadas

<b>Procesadores</b>	<b>Intel® Core Ultra 7</b>
<b>Memoria RAM</b>	32 GB
<b>Sistema Operativo</b>	Windows 11

<b>Entrada</b>	<b>Tiempo (s)</b>
Agricultural-20	338,83
Agricultural-40	436,01
Agricultural-60	1382,27
Agricultural-80	793,81
Agricultural-100	1095,94

## Tablas de datos

<b>Muestra</b>	<b>Salida</b>	<b>Tiempo (ms)</b>
Agricultural-20	Dato1	338,83
Agricultural-40	Dato2	436,01
Agricultural-60	Dato3	1382,27
Agricultural-80	Dato4	793,81
Agricultural-100	Dato5	1095,94

```

Tiempo de ejecución: 338.82570028305054 ms
Número total de registros encontrados: 12854

Último registro recopilado:
- Año de recolección: 2007
- Fecha de carga: 2020-06-09
- Fuente: SURVEY
- Frecuencia: MONTHLY
- Departamento: PUERTO RICO
- Producto: CATTLE
- Unidad de medición: LB / HEAD, LIVE BASIS
- Valor: 100

```

Imagen 1. Dato 1

```
Tiempo de ejecución: 436.0105001926422 ms
Número total de registros encontrados: 25715

Último registro recopilado:
- Año de recolección: 2007
- Fecha de carga: 2020-06-09
- Fuente: SURVEY
- Frecuencia: MONTHLY
- Departamento: PUERTO RICO
- Producto: CATTLE
- Unidad de medición: LB / HEAD, LIVE BASIS
- Valor: 100
```

*Imagen 2. Dato 2*

```
Ingrese el año de interés (formato YYYY): 2007

Tiempo de ejecución: 1382.2725999355316 ms
Número total de registros encontrados: 77470

Último registro recopilado:
- Año de recolección: 2007
- Fecha de carga: 2020-06-09
- Fuente: SURVEY
- Frecuencia: MONTHLY
- Departamento: PUERTO RICO
- Producto: CATTLE
- Unidad de medición: LB / HEAD, LIVE BASIS
- Valor: 100
```

*Imagen 3. Dato 3*

```
Ingrese el año de interés (formato YYYY): 2007

Tiempo de ejecución: 793.8137001991272 ms
Número total de registros encontrados: 51757

Último registro recopilado:
- Año de recolección: 2007
- Fecha de carga: 2020-06-09
- Fuente: SURVEY
- Frecuencia: MONTHLY
- Departamento: PUERTO RICO
- Producto: CATTLE
- Unidad de medición: LB / HEAD, LIVE BASIS
- Valor: 100
```

*Imagen 4. Dato 4*

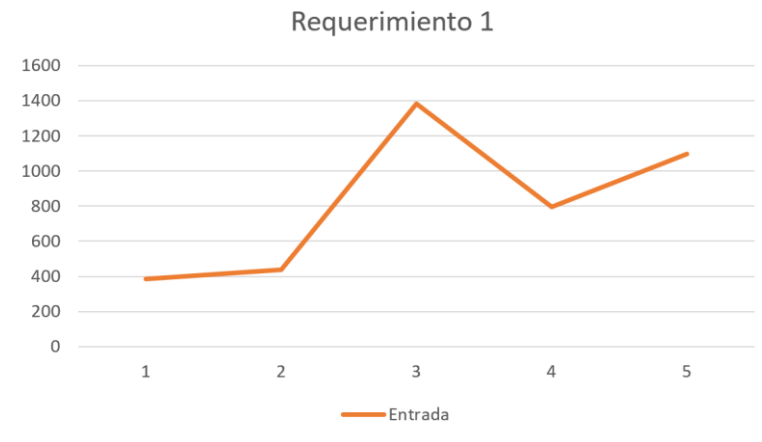
```
Ingrese el año de interés (formato YYYY): 2007

Tiempo de ejecución: 1095.9429998397827 ms
Número total de registros encontrados: 64698

Último registro recopilado:
- Año de recolección: 2007
- Fecha de carga: 2020-06-09
- Fuente: SURVEY
- Frecuencia: MONTHLY
- Departamento: PUERTO RICO
- Producto: CATTLE
- Unidad de medición: LB / HEAD, LIVE BASIS
- Valor: 100
```

*Imagen 5. Dato 5*

Graficas



## Análisis

El Requerimiento 1 se implementó exitosamente utilizando una tabla de hash para el acceso rápido a los registros por año y Merge Sort para ordenarlos por fecha de carga, con una complejidad de  $O(n \log n)$ . Las pruebas en un ordenador potente mostraron que el tiempo de ejecución aumenta con el tamaño de la muestra, pasando de 338,83 ms a 1095,94 ms, lo que confirma la dependencia del algoritmo con el volumen de datos. El rendimiento es adecuado para conjuntos pequeños y medianos.

## Requerimiento 2 (Grupal)

### Descripción



```

1 def req_2(catalog, state, n_registros):
2     """
3     Retorna el resultado del requerimiento 2
4     """
5     # TODO: Modificar el requerimiento 2
6     list_state = lp.get(catalog["map_by_state"], state)
7     lista = al.new_list()
8     list_state = al.merge_sort(list_state, comp_fechas)
9
10    if n_registros > al.size(list_state):
11        n_registros = al.size(list_state)
12    for i in range(0, n_registros):
13        al.add_last(lista, list_state["elements"][i])
14    return lista, al.size(list_state)

```

El requerimiento 2 consiste en listar los N últimos registros cargados dado un departamento de interés. La implementación utiliza una tabla de hash para acceder rápidamente a los registros del departamento solicitado, seguido de un ordenamiento con Merge Sort para organizarlos por fecha de carga.

<b>Entrada</b>	Catálogo, Departamento, Numero de registros
<b>Salidas</b>	Listado de los N registros cargados según su estado filtrado
<b>Implementado (Sí/No)</b>	Implementado grupalmente

## Análisis de complejidad

Análisis de complejidad de cada uno de los pasos del algoritmo

Pasos	Complejidad
Obtener la lista por estado	$O(n)$
Ordenar con Merge Sort	$O(n \log n)$
Comparación	$O(1)$
Agregar registros a la lista de resultados	$O(m)$ siendo m el número de registros
<b>TOTAL</b>	<b><math>O(n \log n)</math></b>

## Pruebas Realizadas

Las pruebas realizadas fueron realizadas en una maquina con las siguientes especificaciones. Los datos de entrada fueron el departamento/estado: US TOTAL y el numero de registros: 5

<b>Procesadores</b>	<b>AMD Ryzen 5 4500U</b>
<b>Memoria RAM</b>	8 GB
<b>Sistema Operativo</b>	Windows 11

<b>Entrada</b>	<b>Tiempo (ms)</b>
Agricultural-20	59,270

Agricultural-40	847,544
Agricultural-60	136,352
Agricultural-80	187,773
Agricultural-100	20214,616

Tablas de datos

Las tablas con la recopilación de datos de las pruebas.

Muestra	Salida	Tiempo (ms)
Agricultural-20	Dato1	59,270
Agricultural-40	Dato2	847,544
Agricultural-60	Dato3	136,352
Agricultural-80	Dato4	187,773
Agricultural-100	Dato5	20214,616

Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
SURVEY	2024	2024-11-25	MONTHLY	US TOTAL	LB, LIVE BASIS	CHICKENS	70,400,000
SURVEY	2024	2024-11-25	MONTHLY	US TOTAL	LB, LIVE BASIS	TURKEYS	58,000
SURVEY	2024	2024-11-25	POINT IN TIME	US TOTAL	LB	PORK	25,063,000
SURVEY	2024	2024-11-25	MONTHLY	US TOTAL	HEAD	CHICKENS	10
SURVEY	2024	2024-11-25	MONTHLY	US TOTAL	HEAD	CHICKENS	702

Imagen 1. Dato 1

Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
SURVEY	2024	2024-11-25	MONTHLY	US TOTAL	LB, LIVE BASIS	CHICKENS	70,400,000
SURVEY	2024	2024-11-25	MONTHLY	US TOTAL	LB, LIVE BASIS	TURKEYS	58,000
SURVEY	2024	2024-11-25	POINT IN TIME	US TOTAL	LB	PORK	25,063,000
SURVEY	2024	2024-11-25	MONTHLY	US TOTAL	HEAD	CHICKENS	10
SURVEY	2024	2024-11-25	MONTHLY	US TOTAL	HEAD	CHICKENS	702

Imagen 2. Dato 2

Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
SURVEY	2024	2024-11-25	MONTHLY	US TOTAL	LB, LIVE BASIS	CHICKENS	70,400,000
SURVEY	2024	2024-11-25	MONTHLY	US TOTAL	LB, LIVE BASIS	TURKEYS	58,000
SURVEY	2024	2024-11-25	POINT IN TIME	US TOTAL	LB	PORK	25,063,000
SURVEY	2024	2024-11-25	MONTHLY	US TOTAL	HEAD	CHICKENS	10
SURVEY	2024	2024-11-25	MONTHLY	US TOTAL	HEAD	CHICKENS	702

Imagen 3. Dato 3

Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
SURVEY	2024	2024-11-25	MONTHLY	US TOTAL	LB, LIVE BASIS	CHICKENS	70,400,000
SURVEY	2024	2024-11-25	MONTHLY	US TOTAL	LB, LIVE BASIS	TURKEYS	58,000
SURVEY	2024	2024-11-25	POINT IN TIME	US TOTAL	LB	PORK	25,063,000
SURVEY	2024	2024-11-25	MONTHLY	US TOTAL	HEAD	CHICKENS	10
SURVEY	2024	2024-11-25	MONTHLY	US TOTAL	HEAD	CHICKENS	702

Imagen 4. Dato 4

Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recoleccion	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
SURVEY	2024	2024-11-25	MONTHLY	US TOTAL	LB, LIVE BASIS	CHICKENS	70,400,000
SURVEY	2024	2024-11-25	MONTHLY	US TOTAL	LB, LIVE BASIS	TURKEYS	58,000
SURVEY	2024	2024-11-25	POINT IN TIME	US TOTAL	LB	PORK	25,063,000
SURVEY	2024	2024-11-25	MONTHLY	US TOTAL	HEAD	CHICKENS	10
SURVEY	2024	2024-11-25	MONTHLY	US TOTAL	HEAD	CHICKENS	702

Imagen 5. Dato 5

### Graficas

Las gráficas con la representación de las pruebas realizadas.



### Análisis

Los tiempos de ejecución muestran un comportamiento no lineal, con un incremento significativo en el caso de Agrícola-100 (20214,616 ms). Esto sugiere que el algoritmo, aunque eficiente para conjuntos de datos pequeños y medianos, puede volverse costoso en términos de tiempo para volúmenes grandes debido al ordenamiento. La variabilidad en los tiempos también podría deberse a factores externos como la gestión de memoria durante las pruebas.

### Requerimiento 3 (Tomás Alarcón)

#### Descripción

Se implementa una funcionalidad para listar los registros recopilados según el nombre del departamento en un rango de años específico. La implementación utiliza un mapa indexado por estado/departamento ("map\_by\_state") para acceder rápidamente a los registros del departamento solicitado, filtra por el rango de años, contabiliza los registros por tipo de fuente (SURVEY/CENSUS) y ordena los resultados por fecha de carga de manera descendente.

```
def req_3(catalog, department, start_year, end_year):
    """
    Retorna el resultado del requerimiento 3
    """
    # TODO: Modificar el requerimiento 3
    list_deptos = lp.get(catalog["map_by_state"], department)
    lista = al.new_list()
    survey = 0
    census = 0
    list_deptos = al.merge_sort(list_deptos, comp_fechas)
    for i in range(al.size(list_deptos)):
        registro = al.get_element(list_deptos, i)
        if start_year <= int(registro["year_collection"]) and int(registro["year_collection"]) <= end_year:
            al.add_last(lista, registro)
            if registro["source"] == "SURVEY":
                survey += 1
            if registro["source"] == "CENSUS":
                census += 1
    return lista, al.size(lista), survey, census
```

<b>Entrada</b>	catalog, department, start_year, end_year
<b>Salidas</b>	tiempo_ejecucion, total_registros, total_survey, total_census, registros
<b>Implementado (Sí/No)</b>	Implementado por Tomás Alarcón

## Análisis de complejidad

Análisis de complejidad de cada uno de los pasos del algoritmo

Pasos	Complejidad
Consultar el mapa por departamento (lp.get)	$O(1)$
Verificar si la lista está vacía	$O(1)$
Filtrar registros por rango de años y contar por tipo de fuente	$O(n)$
Ordenar los registros filtrados con Merge Sort	$O(m \log m)$
Crear sublistas para los primeros y últimos 5 registros	$O(1)$
Construcción del diccionario de resultados	$O(1)$
<b>TOTAL</b>	<b><math>O(n \log n)</math></b>

## Pruebas Realizadas

<b>Procesadores</b>	Intel® Core Ultra 7
<b>Memoria RAM</b>	32 GB
<b>Sistema Operativo</b>	Windows 11

Entrada	Tiempo (ms)
Agricultural-20	7,15
Agricultural-40	28,17
Agricultural-60	46,16
Agricultural-80	28,15
Agricultural-100	68,36

## Tablas de datos

Muestra	Salida	Tiempo (ms)
Agricultural-20	Dato1	7,15
Agricultural-40	Dato2	28,17
Agricultural-60	Dato3	46,16
Agricultural-80	Dato4	28,15
Agricultural-100	Dato5	68,36

```

Ingrese el nombre del departamento a filtrar: FLORIDA
Ingrese el año inicial del periodo a consultar (YYYY): 2007
Ingrese el año final del periodo a consultar (YYYY): 2014

Tiempo de ejecución: 7.715299844741821 ms
Total de registros encontrados: 579
Total de registros con fuente 'SURVEY': 73
Total de registros con fuente 'CENSUS': 506

Primeros 5 registros:

```

Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia	Producto	Unidad de medición
SURVEY	2013	2019-06-25	MONTHLY	CHICKENS	HEAD
SURVEY	2014	2019-06-25	MONTHLY	CHICKENS	EGGS / 100 LAYER
SURVEY	2014	2019-06-25	POINT IN TIME	CHICKENS	PCT OF INVENTORY
SURVEY	2014	2019-05-09	POINT IN TIME	CATTLE	HEAD
SURVEY	2014	2016-04-29	MONTHLY	MILK	PCT

```

Últimos 5 registros:

```

Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia	Producto	Unidad de medición
CENSUS	2007	2012-01-01	ANNUAL	CATTLE	OPERATIONS
CENSUS	2007	2012-01-01	POINT IN TIME	HONEY	COLONIES
CENSUS	2007	2012-01-01	POINT IN TIME	EMUS	HEAD
CENSUS	2007	2012-01-01	POINT IN TIME	GOATS	OPERATIONS
CENSUS	2007	2012-01-01	POINT IN TIME	CATTLE	OPERATIONS

Imagen 1. Dato 1

```

Ingrese el nombre del departamento a filtrar: FLORIDA
Ingrese el año inicial del periodo a consultar (YYYY): 2007
Ingrese el año final del periodo a consultar (YYYY): 2014

Tiempo de ejecución: 28.16569995880127 ms
Total de registros encontrados: 1155
Total de registros con fuente 'SURVEY': 143
Total de registros con fuente 'CENSUS': 1012

Primeros 5 registros:

```

Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia	Producto	Unidad de medición
SURVEY	2013	2019-06-25	MONTHLY	CHICKENS	HEAD
SURVEY	2014	2019-06-25	MONTHLY	CHICKENS	EGGS / 100 LAYER
SURVEY	2014	2019-06-25	POINT IN TIME	CHICKENS	PCT OF INVENTORY
SURVEY	2014	2019-06-25	MONTHLY	CHICKENS	HEAD
SURVEY	2013	2019-06-25	MONTHLY	EGGS	EGGS

```

Últimos 5 registros:

```

Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia	Producto	Unidad de medición
CENSUS	2007	2012-01-01	ANNUAL	CATTLE	OPERATIONS
CENSUS	2007	2012-01-01	POINT IN TIME	EQUINE	OPERATIONS
CENSUS	2007	2012-01-01	POINT IN TIME	CATTLE	OPERATIONS
CENSUS	2007	2012-01-01	ANNUAL	HONEY	LB
CENSUS	2007	2012-01-01	ANNUAL	POULTRY, OTHER	OPERATIONS

Imagen 2. Dato 2

Ingrese el nombre del departamento a filtrar: FLORIDA  
 Ingrese el año inicial del periodo a consultar (YYYY): 2007  
 Ingrese el año final del periodo a consultar (YYYY): 2014

Tiempo de ejecución: 46.16380023956299 ms  
 Total de registros encontrados: 3554  
 Total de registros con fuente 'SURVEY': 446  
 Total de registros con fuente 'CENSUS': 3108

Primeros 5 registros:

Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia	Producto	Unidad de medición
SURVEY	2013	2019-06-25	MONTHLY	CHICKENS	HEAD
SURVEY	2014	2019-06-25	MONTHLY	CHICKENS	EGGS / 100 LAYER
SURVEY	2014	2019-06-25	POINT IN TIME	CHICKENS	PCT OF INVENTORY
SURVEY	2014	2019-06-25	MONTHLY	CHICKENS	HEAD
SURVEY	2013	2019-06-25	MONTHLY	EGGS	EGGS

Últimos 5 registros:

Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia	Producto	Unidad de medición
CENSUS	2007	2012-01-01	ANNUAL	CATTLE	OPERATIONS
CENSUS	2007	2012-01-01	ANNUAL	ANIMAL TOTALS	OPERATIONS
CENSUS	2007	2012-01-01	POINT IN TIME	CATTLE	HEAD
CENSUS	2007	2012-01-01	ANNUAL	EQUINE	OPERATIONS
CENSUS	2007	2012-01-01	ANNUAL	POULTRY TOTALS	OPERATIONS

Imagen 3. Dato 3

Ingrese el nombre del departamento a filtrar: FLORIDA  
 Ingrese el año inicial del periodo a consultar (YYYY): 2007  
 Ingrese el año final del periodo a consultar (YYYY): 2014

Tiempo de ejecución: 28.14539988883667 ms  
 Total de registros encontrados: 2324  
 Total de registros con fuente 'SURVEY': 288  
 Total de registros con fuente 'CENSUS': 2036

Primeros 5 registros:

Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia	Producto	Unidad de medición
SURVEY	2012	2019-06-25	MONTHLY	EGGS	EGGS
SURVEY	2013	2019-06-25	MONTHLY	CHICKENS	HEAD
SURVEY	2014	2019-06-25	MONTHLY	CHICKENS	EGGS / 100 LAYER
SURVEY	2014	2019-06-25	POINT IN TIME	CHICKENS	PCT OF INVENTORY
SURVEY	2014	2019-06-25	MONTHLY	CHICKENS	HEAD

Últimos 5 registros:

Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia	Producto	Unidad de medición
CENSUS	2007	2012-01-01	ANNUAL	HOGS	OPERATIONS
CENSUS	2007	2012-01-01	POINT IN TIME	CATTLE	OPERATIONS
CENSUS	2007	2012-01-01	POINT IN TIME	SHEEP	OPERATIONS
CENSUS	2007	2012-01-01	POINT IN TIME	CATTLE	HEAD
CENSUS	2007	2012-01-01	ANNUAL	ORNAMENTAL FISH	OPERATIONS

Imagen 4. Dato 4

Tiempo de ejecución: 68.36180019378662 ms  
 Total de registros encontrados: 2891  
 Total de registros con fuente 'SURVEY': 359  
 Total de registros con fuente 'CENSUS': 2532

Primeros 5 registros:

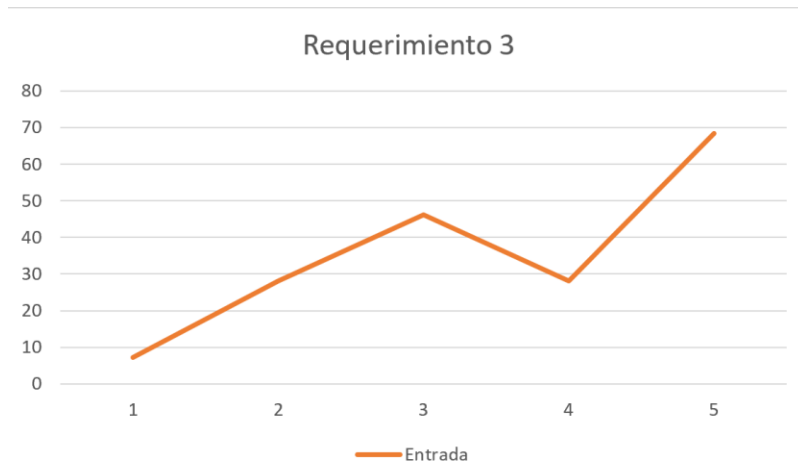
Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia	Producto	Unidad de medición
SURVEY	2012	2019-06-25	MONTHLY	EGGS	EGGS
SURVEY	2013	2019-06-25	MONTHLY	CHICKENS	HEAD
SURVEY	2014	2019-06-25	MONTHLY	CHICKENS	EGGS / 100 LAYER
SURVEY	2014	2019-06-25	POINT IN TIME	CHICKENS	PCT OF INVENTORY
SURVEY	2014	2019-06-25	MONTHLY	CHICKENS	HEAD

Últimos 5 registros:

Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia	Producto	Unidad de medición
SURVEY	2009	2012-01-01	WEEKLY	CHICKENS	HEAD
CENSUS	2007	2012-01-01	ANNUAL	ANIMAL TOTALS	OPERATIONS
CENSUS	2007	2012-01-01	POINT IN TIME	PAWS	HEAD
CENSUS	2007	2012-01-01	POINT IN TIME	CATTLE	OPERATIONS
CENSUS	2007	2012-01-01	ANNUAL	EQUINE	OPERATIONS

Imagen 5. Dato 5

## Graficas



## Análisis

El Requerimiento 3 se implementó con éxito utilizando un mapa indexado por departamento para acceder rápidamente a los registros y un proceso de filtrado por rango de años, con una complejidad total de  $O(n \log n)$  debido al uso de Merge Sort. Las pruebas en un ordenador avanzado mostraron variaciones en los tiempos de ejecución, oscilando entre 7,15 ms y 68,36 ms, con un comportamiento no completamente lineal, mostrando que el rendimiento es eficiente para volúmenes moderados de datos.

## Requerimiento 4 (Juan Sebastian Cortes)

### Descripción

```

def req_4(catalog, producto, anio_inicio, anio_fin):
    """
    Retorna el resultado del requerimiento 4
    """
    # TODO: Modificar el requerimiento 4

    start = get_time()
    total_census = 0
    total_survey = 0
    total_registros = 0

    registros_producto = lp.get(catalog["map_by_commodity"], producto)
    filtrados = al.new_list()

    for i in range(al.size(registros_producto)):
        registro = al.get_element(registros_producto, i)
        anio = int(registro["year_collection"])

        if anio_inicio <= anio <= anio_fin:
            al.add_last(filtrados, registro)
            total_registros += 1
            if registro["source"] == "CENSUS":
                total_census += 1
            elif registro["source"] == "SURVEY":
                total_survey += 1

    al.merge_sort(filtrados, comp_fechas)

    if total_registros > 20:
        primeros = al.sub_list(filtrados, 0, 5)
        ultimos = al.sub_list(filtrados, al.size(filtrados) - 5, 5)
        resultado = al.new_list()
        for i in range(al.size(primeros)):
            al.add_last(resultado, al.get_element(primeros, i))
        for i in range(al.size(ultimos)):
            al.add_last(resultado, al.get_element(ultimos, i))
    else:
        resultado = filtrados

    tiempo = delta_time(start, get_time())
    return tiempo, total_registros, total_census, total_survey, resultado

```

Permite consultar los registros recopilados para un tipo de producto específico en un rango de años determinado, ordenados por fecha de carga descendente. Retorna estadísticas por fuente de origen y muestra los primeros 5 y últimos 5 registros si hay más de 20.

<b>Entrada</b>	producto, año inicial y año final
<b>Salidas</b>	tiempo, numero de registros que cumplieron con el filtro, número de registros survey y census, y la lista de los registros
<b>Implementado (Sí/No)</b>	Sí, implementado por Sebastián Cortés

## Análisis de complejidad

Análisis de complejidad de cada uno de los pasos del algoritmo

Pasos	Complejidad
Buscar en el mapa	$O(1)$
Filtrado de registros	$O(n)$



Ordenar los registros válidos	$O(n\log(n))$
Ordenamiento visual (si son mas de 20 registros)	$O(1)$
<b>TOTAL</b>	<b><math>O(n\log(n))</math></b>

## Pruebas Realizadas

Las pruebas realizadas fueron realizadas en una máquina con las siguientes especificaciones. Los datos de entrada fueron el producto SHEEP, año inicial 2007 y año final 2010.

<b>Procesadores</b>	<b>12th Gen Intel(R) Core(TM) i7-1255U</b>
<b>Memoria RAM</b>	32 GB
<b>Sistema Operativo</b>	Windows 11

Entrada	Tiempo (ms)
Agricultural 20%	8.48
Agricultural 40%	21.27
Agricultural 60%	36.29
Agricultural 80%	40.46
Agricultural 100%	51.95

## Tablas de datos

20% de la muestra:

```
Tiempo de ejecución: 8.478499997407198 ms
Total de registros encontrados: 831
Total con fuente 'CENSUS': 559
Total con fuente 'SURVEY': 272
```

Listado de registros:

Fuente	Año	Fecha de carga	Frecuencia	Departamento	Unidad
SURVEY	2007	2020-06-09	MONTHLY	PUERTO RICO	LB, LIVE BASIS
SURVEY	2010	2015-02-03	ANNUAL	HAWAII	HEAD
CENSUS	2007	2012-01-01	POINT IN TIME	WYOMING	OPERATIONS
CENSUS	2007	2012-01-01	POINT IN TIME	WYOMING	HEAD
SURVEY	2010	2012-01-01	MONTHLY	WYOMING	LB / HEAD, LIVE BASIS
CENSUS	2007	2012-01-01	POINT IN TIME	WYOMING	HEAD

40% de la muestra:

Tiempo de ejecución: 21.265099998563528 ms  
Total de registros encontrados: 1680  
Total con fuente 'CENSUS': 1129  
Total con fuente 'SURVEY': 551

Listado de registros:

Fuente	Año	Fecha de carga	Frecuencia	Departamento	Unidad
SURVEY	2007	2020-06-09	MONTHLY	PUERTO RICO	LB, LIVE BASIS
SURVEY	2008	2020-06-09	MONTHLY	PUERTO RICO	HEAD
SURVEY	2007	2016-10-04	POINT IN TIME	INDIANA	\$ / HEAD
SURVEY	2007	2016-10-04	POINT IN TIME	NEBRASKA	\$
SURVEY	2009	2015-02-03	ANNUAL	COLORADO	HEAD
CENSUS	2007	2012-01-01	POINT IN TIME	WYOMING	OPERATIONS
SURVEY	2007	2012-01-01	MONTHLY	WYOMING	LB, LIVE BASIS
SURVEY	2010	2012-01-01	POINT IN TIME	WYOMING	HEAD
CENSUS	2007	2012-01-01	POINT IN TIME	WYOMING	OPERATIONS
CENSUS	2007	2012-01-01	ANNUAL	WYOMING	OPERATIONS

60% de la muestra:

Tiempo de ejecución: 36.290899999439716 ms  
Total de registros encontrados: 2508  
Total con fuente 'CENSUS': 1695  
Total con fuente 'SURVEY': 813

Listado de registros:

Fuente	Año	Fecha de carga	Frecuencia	Departamento	Unidad
SURVEY	2007	2020-06-09	MONTHLY	PUERTO RICO	LB, LIVE BASIS
SURVEY	2008	2020-06-09	MONTHLY	PUERTO RICO	HEAD
SURVEY	2007	2020-06-09	MONTHLY	PUERTO RICO	HEAD
SURVEY	2007	2016-10-04	POINT IN TIME	INDIANA	\$ / HEAD
SURVEY	2007	2016-10-04	POINT IN TIME	IOWA	\$
CENSUS	2007	2012-01-01	POINT IN TIME	WYOMING	OPERATIONS
CENSUS	2007	2012-01-01	POINT IN TIME	WYOMING	OPERATIONS
CENSUS	2007	2012-01-01	POINT IN TIME	WYOMING	OPERATIONS
CENSUS	2007	2012-01-01	POINT IN TIME	WYOMING	HEAD
CENSUS	2007	2012-01-01	POINT IN TIME	WYOMING	OPERATIONS

80% de la muestra:

Tiempo de ejecución: 40.461500000208616 ms  
Total de registros encontrados: 3351  
Total con fuente 'CENSUS': 2271  
Total con fuente 'SURVEY': 1080

Listado de registros:

Fuente	Año	Fecha de carga	Frecuencia	Departamento	Unidad
SURVEY	2007	2020-06-09	MONTHLY	PUERTO RICO	LB, LIVE BASIS
SURVEY	2008	2020-06-09	MONTHLY	PUERTO RICO	HEAD
SURVEY	2007	2020-06-09	MONTHLY	PUERTO RICO	HEAD
SURVEY	2007	2016-10-04	POINT IN TIME	INDIANA	\$ / HEAD
SURVEY	2007	2016-10-04	POINT IN TIME	IOWA	\$
SURVEY	2008	2012-01-01	MONTHLY	WYOMING	LB / HEAD, LIVE BASIS
CENSUS	2007	2012-01-01	ANNUAL	WYOMING	HEAD
CENSUS	2007	2012-01-01	POINT IN TIME	WYOMING	OPERATIONS
CENSUS	2007	2012-01-01	POINT IN TIME	WYOMING	OPERATIONS
SURVEY	2010	2012-01-01	MONTHLY	WYOMING	\$ / CWT

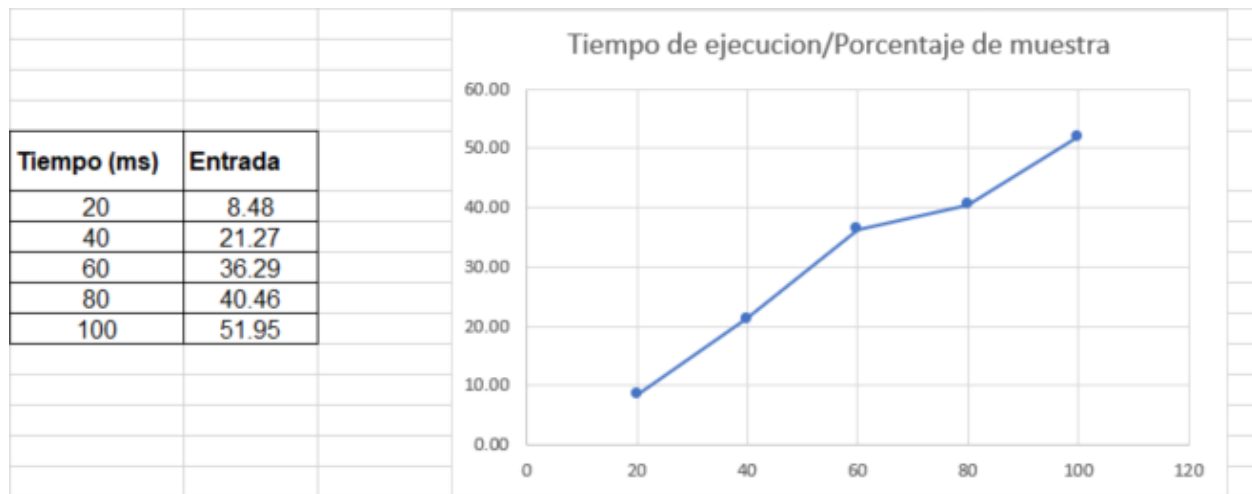
100% de la muestra:

Tiempo de ejecución: 51.95319999754429 ms  
Total de registros encontrados: 4162  
Total con fuente 'CENSUS': 2841  
Total con fuente 'SURVEY': 1321

Listado de registros:

Fuente	Año	Fecha de carga	Frecuencia	Departamento	Unidad
SURVEY	2007	2020-06-09	MONTHLY	PUERTO RICO	LB, LIVE BASIS
SURVEY	2008	2020-06-09	MONTHLY	PUERTO RICO	HEAD
SURVEY	2007	2020-06-09	MONTHLY	PUERTO RICO	HEAD
SURVEY	2008	2020-06-09	MONTHLY	PUERTO RICO	HEAD
SURVEY	2007	2016-10-04	POINT IN TIME	INDIANA	\$ / HEAD
CENSUS	2007	2012-01-01	POINT IN TIME	WYOMING	OPERATIONS
SURVEY	2007	2012-01-01	MONTHLY	WYOMING	\$ / CWT
SURVEY	2010	2012-01-01	MONTHLY	WYOMING	\$ / CWT
SURVEY	2008	2012-01-01	MONTHLY	WYOMING	HEAD
SURVEY	2009	2012-01-01	MONTHLY	WYOMING	\$ / CWT

## Gráfica



## Análisis

El Requerimiento 4 tiene una complejidad de  $O(n\log(n))$ . Los tiempos medidos muestran un crecimiento casi lineal al aumentar el tamaño de los datos, confirmando la eficiencia del algoritmo. El uso de mapas optimiza el filtrado, reduciendo en gran manera el tiempo de procesamiento, filtrado y por ende, de su ejecución total, incluso con volúmenes altos de datos.

## Requerimiento 5 (Valentina España)

### Descripción

```
1 def req_5(catalog, category, year_ini, year_fin):
2     """
3     Retorna el resultado del requerimiento 5
4     """
5     # TODO: Modificar el requerimiento 5
6     list_category = lp.get(catalog["map_by_category"], category)
7     lista = al.new_list()
8     list_category = al.merge_sort(list_category, comp_fechas)
9     for i in range(al.size(list_category)):
10        registro = al.get_element(list_category, i)
11        if year_ini <= int(registro["year_collection"]) and int(registro["year_collection"]) <= year_fin:
12            al.add_last(lista, registro)
13
14    return lista, al.size(lista)
```

El requerimiento 5 es listar los registros cargados según su categoría estadística para un rango de tiempo de interés. Utiliza un mapa por categoría para acceder a los registros, los ordena con Merge Sort y los filtra por el rango de años.

<b>Entrada</b>	Catálogo, Categoría, Año inicial, Año final
<b>Salidas</b>	Listado de los registros cargados según su categoría y el rango de años filtrado
<b>Implementado (Sí/No)</b>	Implementado por Valentina España Cuellar

## Análisis de complejidad

Análisis de complejidad de cada uno de los pasos del algoritmo

Pasos	Complejidad
Obtener la lista por categoria	$O(n)$
Ordenar con Merge Sort	$O(n \log n)$
Recorrido de la lista por categoria para agregar los registros a la lista de resultados	$O(n)$
<b>TOTAL</b>	<b><math>O(n \log n)</math></b>

## Pruebas Realizadas

Las pruebas realizadas fueron realizadas en una maquina con las siguientes especificaciones. Los datos de entrada fueron la categoría: INVENTORY, el año inicial: 2012 y el año final: 2015.

<b>Procesadores</b>	<b>AMD Ryzen 5 4500U</b>
<b>Memoria RAM</b>	8 GB
<b>Sistema Operativo</b>	Windows 11

Entrada	Tiempo (ms)
Agricultural-20	751,386
Agricultural-40	1359,810
Agricultural-60	1685,175
Agricultural-80	2327,077
Agricultural-100	4017,694

## Tablas de datos

Las tablas con la recopilación de datos de las pruebas.

Muestra	Salida	Tiempo (ms)
Agricultural-20	Dato1	751,386
Agricultural-40	Dato2	1359,810
Agricultural-60	Dato3	1685,175
Agricultural-80	Dato4	2327,077
Agricultural-100	Dato5	4017,694

Primeros 5 registros:							
Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
SURVEY	2013	2019-08-19	POINT IN TIME	NORTH CAROLINA	HEAD	HOGS	105,000
SURVEY	2015	2019-08-19	POINT IN TIME	NORTH CAROLINA	HEAD	HOGS	(0)
SURVEY	2015	2019-08-19	POINT IN TIME	NORTH CAROLINA	HEAD	HOGS	(0)
SURVEY	2015	2019-08-19	POINT IN TIME	NORTH CAROLINA	HEAD	HOGS	(0)
SURVEY	2015	2019-08-19	POINT IN TIME	NORTH DAKOTA	HEAD	HOGS	500
Ultimos 5 registros:							
Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	OPERATIONS	CATTLE	73
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	OPERATIONS	GOATS	34
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	OPERATIONS	CATTLE	232
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	OPERATIONS	CATTLE	16
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	OPERATIONS	CATTLE	14

Imagen 1. Dato 1

Primeros 5 registros:							
Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
SURVEY	2015	2019-08-19	POINT IN TIME	MONTANA	HEAD	HOGS	200
SURVEY	2013	2019-08-19	POINT IN TIME	NORTH CAROLINA	HEAD	HOGS	105,000
SURVEY	2015	2019-08-19	POINT IN TIME	NORTH CAROLINA	HEAD	HOGS	(0)
SURVEY	2015	2019-08-19	POINT IN TIME	NORTH CAROLINA	HEAD	HOGS	(0)
SURVEY	2015	2019-08-19	POINT IN TIME	NORTH CAROLINA	HEAD	HOGS	(0)
Ultimos 5 registros:							
Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	OPERATIONS	CHICKENS	239
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	HEAD	CATTLE	543
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	OPERATIONS	CATTLE	262
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	HEAD	CATTLE	8,005
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	OPERATIONS	CHICKENS	13

Imagen 2. Dato 2

Primeros 5 registros:							
Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
SURVEY	2015	2019-08-19	POINT IN TIME	MONTANA	HEAD	HOGS	200
SURVEY	2014	2019-08-19	POINT IN TIME	MONTANA	HEAD	HOGS	300
SURVEY	2015	2019-08-19	POINT IN TIME	MONTANA	HEAD	HOGS	20,500
SURVEY	2013	2019-08-19	POINT IN TIME	NORTH CAROLINA	HEAD	HOGS	105,000
SURVEY	2015	2019-08-19	POINT IN TIME	NORTH CAROLINA	HEAD	HOGS	(0)
Ultimos 5 registros:							
Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	HEAD	TURKEYS	157
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	OPERATIONS	GOATS	44
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	OPERATIONS	CATTLE	12
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	HEAD	CATTLE	365
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	OPERATIONS	EQUINE	465

Imagen 3. Dato 3

Primeros 5 registros:							
Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
SURVEY	2015	2019-08-19	POINT IN TIME	MONTANA	HEAD	HOGS	200
SURVEY	2014	2019-08-19	POINT IN TIME	MONTANA	HEAD	HOGS	300
SURVEY	2015	2019-08-19	POINT IN TIME	MONTANA	HEAD	HOGS	20,500
SURVEY	2013	2019-08-19	POINT IN TIME	NORTH CAROLINA	HEAD	HOGS	105,000
SURVEY	2015	2019-08-19	POINT IN TIME	NORTH CAROLINA	HEAD	HOGS	(0)
Ultimos 5 registros:							
Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	OPERATIONS	CATTLE	85
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	OPERATIONS	GOATS	70
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	OPERATIONS	CATTLE	3
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	OPERATIONS	LLAMAS	4
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	OPERATIONS	GOATS	13

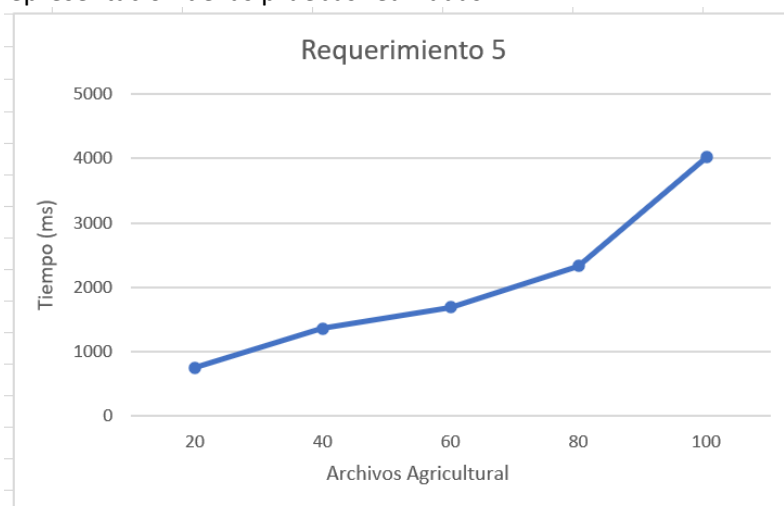
Imagen 4. Dato 4

Primeros 5 registros:							
Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
SURVEY	2015	2019-08-19	POINT IN TIME	MONTANA	HEAD	HOGS	200
SURVEY	2014	2019-08-19	POINT IN TIME	MONTANA	HEAD	HOGS	300
SURVEY	2015	2019-08-19	POINT IN TIME	MONTANA	HEAD	HOGS	20,500
SURVEY	2014	2019-08-19	POINT IN TIME	MONTANA	HEAD	HOGS	(0)
SURVEY	2013	2019-08-19	POINT IN TIME	NORTH CAROLINA	HEAD	HOGS	105,000
Ultimos 5 registros:							
Tipo de fuente	Año de recoleccion	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medicion	Tipo del producto	Valor de medicion
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	HEAD	CATTLE	9
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	OPERATIONS	CATTLE	246
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	HEAD	CATTLE	21
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	OPERATIONS	QUAIL	1
CENSUS	2012	2012-12-31	POINT IN TIME	WYOMING	OPERATIONS	LLAMAS	4

Imagen 5. Dato 5

## Graficas

Las gráficas con la representación de las pruebas realizadas.



## Análisis

Los tiempos de ejecución aumentan de manera proporcional al tamaño de la muestra, confirmando la complejidad teórica. El incremento es consistente, pasando de 751,386 ms para Agrícola-20 a 4017,694 ms para Agrícola-100. Esto indica que el algoritmo maneja eficientemente conjuntos de datos medianos, pero el tiempo de ejecución se vuelve significativo para conjuntos más grandes.

## Requerimiento 6 (Grupal)

### Descripción

Este requerimiento permite analizar los registros de un departamento específico dentro de un rango de fechas de carga determinado. Se accede a los registros mediante una tabla hash indexada por departamento, se filtran los registros que cumplen con el rango de fechas y se contabilizan los registros según su fuente de origen ("SURVEY" o "CENSUS"). Finalmente, los registros se ordenan de manera descendente por fecha de carga.

```
def req_6(catalog, department, start_date, end_date):  
    """  
    Retorna el resultado del requerimiento 6  
    """  
    # TODO: Modificar el requerimiento 6  
    start_time = get_time()  
  
    start_date = datetime.strptime(start_date, "%Y-%m-%d")  
    end_date = datetime.strptime(end_date, "%Y-%m-%d")  
  
    list_depto = lp.get(catalog["map_by_state"], department)  
    lista = al.new_list()  
    survey = 0  
    census = 0  
    list_depto = al.merge_sort(list_depto, comp_fecha)  
    for i in range(al.size(list_depto)):  
        registro = al.get_element(list_depto, i)  
        if start_date <= registro["Load_time"] and registro["Load_time"] <= end_date:  
            al.add_last(lista, registro)  
            if registro["source"] == "SURVEY":  
                survey += 1  
            if registro["source"] == "CENSUS":  
                census += 1  
  
    return lista, al.size(lista), census, survey
```

Entrada	Catalog, department, start_date, end_date
Salidas	Lista (registros filtrados dentro del rango de fechas y ordenados por fecha de carga), total_registros, total_census, total_survey
Implementado (Sí/No)	Implementado grupalmente

### Análisis de complejidad

Análisis de complejidad de cada uno de los pasos del algoritmo

Pasos	Complejidad
-------	-------------



Consultar el mapa por departamento	$O(1)$
Convertir las fechas de entrada a formato datetime	$O(1)$
Ordenar los registros por fecha de carga con Merge Sort	$O(n \log n)$
Filtrar los registros dentro del rango de fechas	$O(n)$
Contar los registros por tipo de fuente ("SURVEY" y "CENSUS")	$O(n)$
Construcción del resultado final	$O(1)$
<b>TOTAL</b>	<b><math>O(n \log n)</math></b>

### Pruebas Realizadas

Procesadores	Intel® Core Ultra 7
Memoria RAM	32 GB
Sistema Operativo	Windows 11

Entrada	Tiempo (ms)
Agricultural-20	38,56
Agricultural-40	75,75
Agricultural-60	214,40
Agricultural-80	136,86
Agricultural-100	162,65

### Tablas de datos

Muestra	Salida	Tiempo (ms)
Agricultural-20	Dato1	38,56
Agricultural-40	Dato2	75,75
Agricultural-60	Dato3	214,40
Agricultural-80	Dato4	136,86
Agricultural-100	Dato5	162,65

Ingrese el nombre del departamento a filtrar: ALABAMA  
 Ingrese la fecha inicial del periodo a consultar (YYYY-MM-DD): 2018-01-01  
 Ingrese la fecha final del periodo a consultar (YYYY-MM-DD): 2020-01-01  
 Tiempo de ejecución: 38.85629987716675 ms  
 Total de registros encontrados: 336  
 Total de registros con fuente 'SURVEY': 292  
 Total de registros con fuente 'CENSUS': 44  
 Primeros 5 registros:

Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medición	Tipo del producto	Valor de medición
SURVEY	2019	2019-11-04	WEEKLY	ALABAMA	PCT EXCELLENT	CATTLE	2
SURVEY	2018	2019-08-01	MONTHLY	ALABAMA	COLONIES	HONEY	1,700
SURVEY	2018	2019-08-01	MONTHLY	ALABAMA	COLONIES	HONEY	8,500
SURVEY	2019	2019-07-29	WEEKLY	ALABAMA	PCT FAIR	CATTLE	20
SURVEY	2017	2019-06-25	MONTHLY	ALABAMA	EGGS	EGGS	138,300,000

Ultimos 5 registros:

Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medición	Tipo del producto	Valor de medición
CENSUS	2017	2018-02-01	POINT IN TIME	ALABAMA	OPERATIONS	RABBITS	5
CENSUS	2017	2018-02-01	POINT IN TIME	ALABAMA	OPERATIONS	CATTLE	35
CENSUS	2017	2018-02-01	ANNUAL	ALABAMA	OPERATIONS	WOOL	1
CENSUS	2017	2018-02-01	POINT IN TIME	ALABAMA	HEAD	SHEEP	116
CENSUS	2017	2018-02-01	ANNUAL	ALABAMA	OPERATIONS	CATTLE	27

Imagen 1. Dato 1

Ingrese el nombre del departamento a filtrar: ALABAMA  
 Ingrese la fecha inicial del periodo a consultar (YYYY-MM-DD): 2018-01-01  
 Ingrese la fecha final del periodo a consultar (YYYY-MM-DD): 2020-01-01  
 Tiempo de ejecución: 75.75230002403259 ms  
 Total de registros encontrados: 631  
 Total de registros con fuente 'SURVEY': 553  
 Total de registros con fuente 'CENSUS': 78  
 Primeros 5 registros:

Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medición	Tipo del producto	Valor de medición
SURVEY	2019	2019-11-04	WEEKLY	ALABAMA	PCT EXCELLENT	CATTLE	2
SURVEY	2019	2019-08-12	WEEKLY	ALABAMA	PCT EXCELLENT	CATTLE	5
SURVEY	2019	2019-08-12	WEEKLY	ALABAMA	PCT FAIR	CATTLE	39
SURVEY	2019	2019-08-05	WEEKLY	ALABAMA	PCT EXCELLENT	CATTLE	10
SURVEY	2018	2019-08-01	MONTHLY	ALABAMA	COLONIES	HONEY	1,700

Ultimos 5 registros:

Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medición	Tipo del producto	Valor de medición
CENSUS	2017	2018-02-01	ANNUAL	ALABAMA	OPERATIONS	CATTLE	56
CENSUS	2017	2018-02-01	ANNUAL	ALABAMA	9	ANIMAL TOTALS	1,815,000
CENSUS	2017	2018-02-01	POINT IN TIME	ALABAMA	HEAD	HOOS	(0)
CENSUS	2017	2018-02-01	POINT IN TIME	ALABAMA	OPERATIONS	SHEEP	6
CENSUS	2017	2018-02-01	ANNUAL	ALABAMA	HEAD	GOATS	(0)

Imagen 2. Dato 2

Ingrese el nombre del departamento a filtrar: ALABAMA  
 Ingrese la fecha inicial del periodo a consultar (YYYY-MM-DD): 2018-01-01  
 Ingrese la fecha final del periodo a consultar (YYYY-MM-DD): 2020-01-01  
 Tiempo de ejecución: 214.3961000442505 ms  
 Total de registros encontrados: 1900  
 Total de registros con fuente 'SURVEY': 1668  
 Total de registros con fuente 'CENSUS': 232  
 Primeros 5 registros:

Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medición	Tipo del producto	Valor de medición
SURVEY	2019	2019-11-04	WEEKLY	ALABAMA	PCT EXCELLENT	CATTLE	2
SURVEY	2019	2019-11-04	WEEKLY	ALABAMA	PCT EXCELLENT	CATTLE	2
SURVEY	2019	2019-09-09	WEEKLY	ALABAMA	PCT GOOD	CATTLE	69
SURVEY	2019	2019-09-09	WEEKLY	ALABAMA	PCT GOOD	CATTLE	69
SURVEY	2019	2019-08-12	WEEKLY	ALABAMA	PCT EXCELLENT	CATTLE	5

Ultimos 5 registros:

Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medición	Tipo del producto	Valor de medición
CENSUS	2017	2018-02-01	POINT IN TIME	ALABAMA	HEAD	CHICKENS	86
CENSUS	2017	2018-02-01	POINT IN TIME	ALABAMA	HEAD	CATTLE	(0)
CENSUS	2017	2018-02-01	POINT IN TIME	ALABAMA	HEAD	CATTLE	22,099
CENSUS	2017	2018-02-01	POINT IN TIME	ALABAMA	OPERATIONS	CATTLE	73
CENSUS	2017	2018-02-01	POINT IN TIME	ALABAMA	OPERATIONS	GOATS	44

Imagen 3. Dato 3

Ingrese el nombre del departamento a filtrar: ALABAMA  
 Ingrese la fecha inicial del periodo a consultar (YYYY-MM-DD): 2018-01-01  
 Ingrese la fecha final del periodo a consultar (YYYY-MM-DD): 2020-01-01  
 Tiempo de ejecución: 136.85839986081147 ms  
 Total de registros encontrados: 1243  
 Total de registros con fuente 'SURVEY': 1093  
 Total de registros con fuente 'CENSUS': 150  
 Primeros 5 registros:

Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medición	Tipo del producto	Valor de medición
SURVEY	2019	2019-11-04	WEEKLY	ALABAMA	PCT EXCELLENT	CATTLE	2
SURVEY	2019	2019-10-28	WEEKLY	ALABAMA	PCT POOR	CATTLE	5
SURVEY	2019	2019-09-09	WEEKLY	ALABAMA	PCT GOOD	CATTLE	69
SURVEY	2019	2019-08-12	WEEKLY	ALABAMA	PCT EXCELLENT	CATTLE	5
SURVEY	2019	2019-08-12	WEEKLY	ALABAMA	PCT FAIR	CATTLE	19

Ultimos 5 registros:

Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medición	Tipo del producto	Valor de medición
CENSUS	2017	2018-02-01	POINT IN TIME	ALABAMA	OPERATIONS	CATTLE	20,004
CENSUS	2017	2018-02-01	POINT IN TIME	ALABAMA	HEAD	CATTLE	155
CENSUS	2017	2018-02-01	POINT IN TIME	ALABAMA	OPERATIONS	CATTLE	106
CENSUS	2017	2018-02-01	POINT IN TIME	ALABAMA	OPERATIONS	ELK	1
CENSUS	2017	2018-02-01	POINT IN TIME	ALABAMA	HEAD	HOOS	2,110

Imagen 4. Dato 4

```

Ingresar el nombre del departamento a filtrar: ALABAMA
Ingresar la fecha inicial del periodo a consultar (YYYY-MM-DD): 2018-01-01
Ingresar la fecha final del periodo a consultar (YYYY-MM-DD): 2020-01-01
Tiempo de ejecución: 162.64750027056555 ms
Total de registros encontrados: 1560
Total de registros con fuente 'SURVEY': 1370
Total de registros con fuente 'CENSUS': 195
Primeros 5 registros:

```

Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medición	Tipo del producto	Valor de medición
SURVEY	2019	2019-11-18	WEEKLY	ALABAMA	PCT FAIR	CATTLE	30
SURVEY	2019	2019-11-04	WEEKLY	ALABAMA	PCT EXCELLENT	CATTLE	2
SURVEY	2019	2019-10-26	WEEKLY	ALABAMA	PCT POOR	CATTLE	5
SURVEY	2019	2019-09-16	WEEKLY	ALABAMA	PCT VERY POOR	CATTLE	1
SURVEY	2019	2019-09-09	WEEKLY	ALABAMA	PCT GOOD	CATTLE	69

```

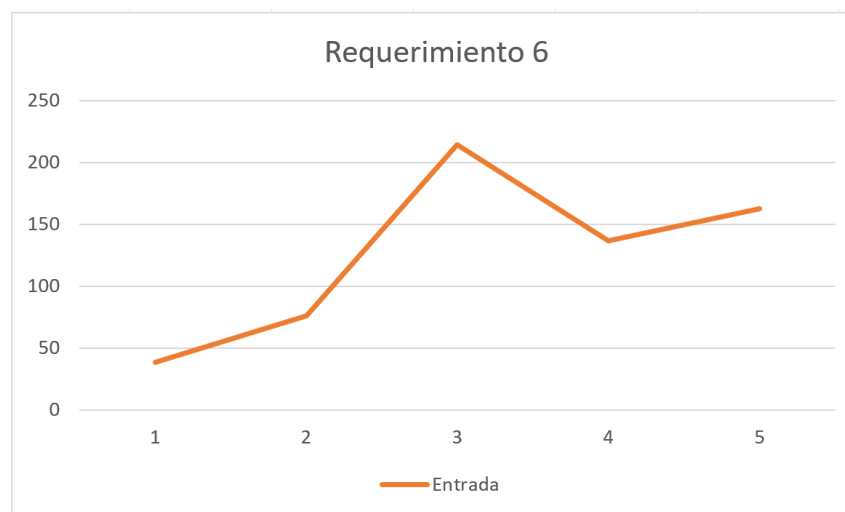
Últimos 5 registros:

```

Tipo de fuente	Año de recolección	Fecha de carga	Frecuencia de recolección	Departamento	Unidad de medición	Tipo del producto	Valor de medición
CENSUS	2017	2018-02-01	POINT IN TIME	ALABAMA	HEAD	EQUINE	727
CENSUS	2017	2018-02-01	ANNUAL	ALABAMA	\$	ANIMAL TOTALS	18,212,000
CENSUS	2017	2018-02-01	POINT IN TIME	ALABAMA	OPERATIONS	CATTLE	1,765
CENSUS	2017	2018-02-01	POINT IN TIME	ALABAMA	OPERATIONS	POULTRY, OTHER	8
CENSUS	2017	2018-02-01	POINT IN TIME	ALABAMA	OPERATIONS	POULTRY, OTHER	58

Imagen 5. Dato 5

## Graficas



## Análisis

El requerimiento 6 permite analizar los registros de un departamento en un rango de fechas determinado, proporcionando estadísticas sobre la cantidad total de registros y su distribución por fuente de origen (CENSUS o SURVEY). Al emplear el mapa se accede rápidamente a los datos del departamento, se filtra los registros por fecha de carga y los ordena en orden descendente utilizando Merge Sort. La complejidad total del algoritmo es  $O(n \log n)$  debido al proceso de ordenamiento, siendo eficiente para grandes volúmenes de datos. Sin embargo, los tiempos de ejecución muestran que el rendimiento puede verse afectado por el tamaño de la muestra, indicando que el tiempo de filtrado y conteo puede volverse costoso a medida que los datos crecen.

## Requerimiento 7 (Grupal)

## Descripción

```
def req_7(catalog, estado, anio_inicio, anio_final, orden):  
    """  
    Retorna el resultado del requerimiento 7  
    """  
    # TODO: Modificar el requerimiento 7  
  
    start = get_time()  
    total_filtrados = 0  
    resumen = al.new_list()  
  
    registros_estado = lp.get(catalog["map_by_state"], estado)  
    ingresos_por_anio = lp.new_map(100, 0.5)  
    invalidos = ["(D)", "(Z)", "(Y)", "(X)", "(NA)"]  
    for i in range(al.size(registros_estado)):  
        registro = al.get_element(registros_estado, i)  
        anio = int(registro["year_collection"])  
        if anio_inicio <= anio <= anio_final and "$" in registro["unit_measurement"]:  
            valor = registro["value"]  
            fuente = registro["source"]  
            total_filtrados += 1  
  
            if not lp.contains(ingresos_por_anio, anio):  
                lp.put(ingresos_por_anio, anio, {  
                    "anio": anio,  
                    "ingresos": 0,  
                    "registros": 0,  
                    "invalidos": 0,  
                    "survey": 0,  
                    "census": 0  
                })  
  
            datos = lp.get(ingresos_por_anio, anio)  
  
            if valor in invalidos:  
                datos["invalidos"] += 1  
            else:  
                datos["ingresos"] += float(valor.replace(",", ""))  
  
            datos["registros"] += 1  
            if fuente == "SURVEY":  
                datos["survey"] += 1  
            elif fuente == "CENSUS":  
                datos["census"] += 1  
  
    resumen = lp.value_set(ingresos_por_anio)  
  
    if orden.upper() == "ASCENDENTE":  
        al.merge_sort(resumen, comparar_ingresos_asc)  
    else:  
        al.merge_sort(resumen, comparar_ingresos_desc)  
  
    if al.size(resumen) > 0:  
        al.get_element(resumen, 0)["indicador"] = "MAYOR" if orden.upper() == "DESCENDENTE" else "MENOR"  
        al.get_element(resumen, al.size(resumen) - 1)["indicador"] = "MENOR" if orden.upper() == "DESCENDENTE" else "MAYOR"  
  
    if al.size(resumen) > 20:  
        top_5 = al.sub_list(resumen, 0, 5)  
        bottom_5 = al.sub_list(resumen, al.size(resumen) - 5, 5)  
        nuevo_resumen = al.new_list()  
        for i in range(al.size(top_5)):  
            al.add_last(nuevo_resumen, al.get_element(top_5, i))  
        for i in range(al.size(bottom_5)):  
            al.add_last(nuevo_resumen, al.get_element(bottom_5, i))  
        resumen = nuevo_resumen  
  
    tiempo = delta_time(start, get_time())  
    return tiempo, total_filtrados, resumen
```

El Requerimiento 7 calcula y ordena los ingresos anuales de un departamento en un rango de años dado, considerando solo los registros en dólares. Para cada año se reportan los ingresos totales,

cantidad de registros válidos e inválidos, y número de registros por fuente. El resultado se ordena según el criterio del usuario.

<b>Entrada</b>	estado, tipo de ordenamiento, año inicial y año final
Salidas	tiempo, número de registros que cumplieron con el filtro y la lista de los años ordenada.
<b>Implementado (Sí/No)</b>	Si, implementado grupalmente

## Análisis de complejidad

Análisis de complejidad de cada uno de los pasos del algoritmo

Pasos	Complejidad
Buscar en el mapa	$O(1)$
Filtrado de registros	$O(n)$
Ordenar los registros válidos	$O(n\log(n))$
Ordenamiento visual (si son mas de 20 registros)	$O(1)$
<b>TOTAL</b>	<b><math>O(n\log(n))</math></b>

## Pruebas Realizadas

Las pruebas realizadas fueron realizadas en una máquina con las siguientes especificaciones. Los datos de entrada fueron el estado CALIFORNIA, orden ASCENDENTE, año inicial 2007 y año final 2010.

<b>Procesadores</b>	<b>12th Gen Intel(R) Core(TM) i7-1255U</b>
<b>Memoria RAM</b>	32 GB
<b>Sistema Operativo</b>	Windows 11

Entrada	Tiempo (ms)
Agricultural 20%	1.63
Agricultural 40%	3.45
Agricultural 60%	5.5
Agricultural 80%	6.06
Agricultural 100%	6.23

## Tablas de datos

20% de la muestra:

Tiempo de ejecución: 1.6294999979436398 ms

Total de registros válidos encontrados: 16

Año	Etiqueta	Ingresos (\$)	# Registros	# Inválidos	# Survey	# Census
2010	MENOR	14.85	2	0	2	0
2009		87.5	2	0	2	0
2008		5.74321e+07	4	0	4	0
2007	MAYOR	5.14538e+08	8	2	1	7

40% de la muestra:

Tiempo de ejecución: 3.247900001704693 ms

Total de registros válidos encontrados: 26

Año	Etiqueta	Ingresos (\$)	# Registros	# Inválidos	# Survey	# Census
2010	MENOR	14.85	2	0	2	0
2009		227.6	4	0	4	0
2008		5.74321e+07	6	1	5	1
2007	MAYOR	1.08822e+09	14	5	2	12

60% de la muestra:

Tiempo de ejecución: 5.501199997961521 ms

Total de registros válidos encontrados: 43

Año	Etiqueta	Ingresos (\$)	# Registros	# Inválidos	# Survey	# Census
2010	MENOR	93.75	3	0	3	0
2008		5.74321e+07	6	1	5	1
2009		1.02093e+08	8	0	8	0
2007	MAYOR	2.85014e+09	26	7	3	23

80% de la muestra:

Tiempo de ejecución: 6.061300002038479 ms  
Total de registros válidos encontrados: 58

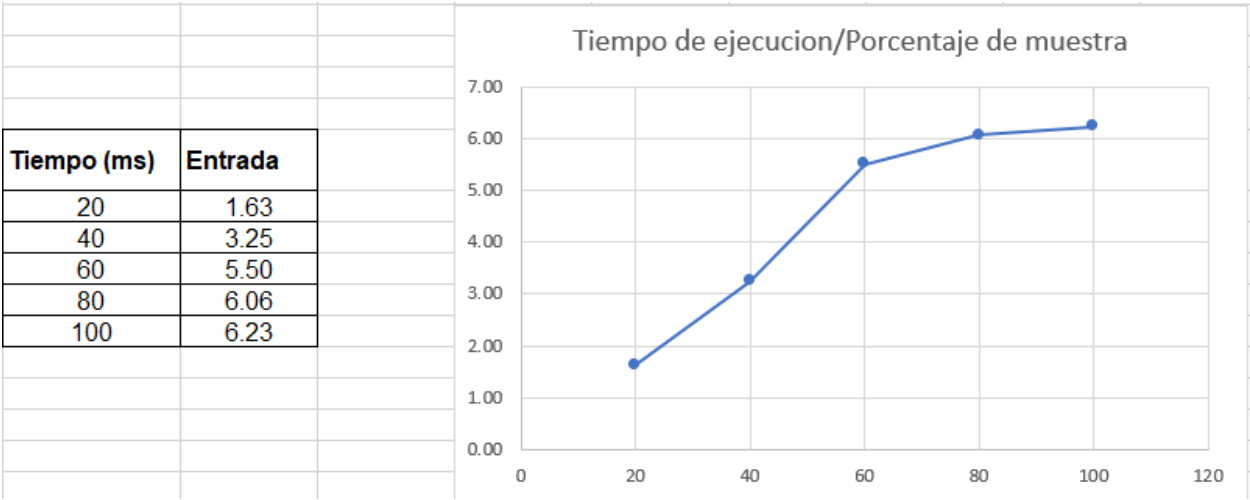
Año	Etiqueta	Ingresos (\$)	# Registros	# Inválidos	# Survey	# Census
2010	MENOR	269.95	5	0	5	0
2008		8.77741e+07	8	2	5	3
2009		1.02093e+08	8	0	8	0
2007	MAYOR	2.86588e+09	37	9	6	31

100% de la muestra:

Tiempo de ejecución: 6.230000000447035 ms  
Total de registros válidos encontrados: 71

Año	Etiqueta	Ingresos (\$)	# Registros	# Inválidos	# Survey	# Census
2010	MENOR	368.55	6	0	6	0
2008		8.77741e+07	9	2	6	3
2009		1.02094e+08	11	0	11	0
2007	MAYOR	2.87916e+09	45	13	7	38

Gráfica



## Análisis

El Requerimiento 7 tiene una complejidad de  $O(n)$ . Los tiempos medidos muestran un crecimiento casi lineal al aumentar el tamaño de los datos, confirmando la eficiencia del algoritmo. El uso de mapas optimiza el filtrado, reduciendo en gran manera el tiempo de procesamiento, filtrado y por ende, de su ejecución total, incluso con volúmenes altos de datos.



# Requerimiento 8 (Bono)

## Descripción

```
1 def req_8(catalog, num_estados, orden):
2     """
3     Retorna el resultado del requerimiento 8
4
5     # TODO: Modificar el requerimiento 8
6     """
7     start = get_time()
8     list_state_names = lp.key_set(catalog["map_by_state"])
9     array_list = al.new_list()
10
11     deptos = {}
12     for i in list_state_names["elements"]:
13
14         list_registros = lp.get(catalog["map_by_state"], i)
15
16         for registro in list_registros["elements"]:
17             if registro["unit_measurement"] == "(D)":
18                 continue
19             if registro["state_name"] not in deptos:
20                 deptos[registro["state_name"]] = {
21                     "state_name": registro["state_name"],
22                     "diferencia": 0,
23                     "menor_anio": int(registro["year_collection"]),
24                     "mayor_anio": int(registro["year_collection"]),
25                     "menor_tp": registro["load_time"].year - int(registro["year_collection"]),
26                     "mayor_tp": registro["load_time"].year - int(registro["year_collection"]),
27                     "survey": 0,
28                     "census": 0,
29                     "total_registros": 0,
30                     "promedio": 0
31                 }
32             depto_stats = deptos[registro["state_name"]]
33             diferencia = registro["load_time"].year - int(registro["year_collection"])
34
35             deptos[registro["state_name"]]["diferencia"] += diferencia
36             deptos[registro["state_name"]]["total_registros"] += 1
37
38             if registro["source"] == "SURVEY":
39                 deptos[registro["state_name"]]["survey"] += 1
40             elif registro["source"] == "CENSUS":
41                 deptos[registro["state_name"]]["census"] += 1
42
43             deptos[registro["state_name"]]["menor_tp"] = min(depto_stats["menor_tp"], diferencia)
44             deptos[registro["state_name"]]["mayor_tp"] = max(depto_stats["mayor_tp"], diferencia)
45             deptos[registro["state_name"]]["menor_anio"] = min(depto_stats["menor_anio"], int(registro["year_collection"]))
46             deptos[registro["state_name"]]["mayor_anio"] = max(depto_stats["mayor_anio"], int(registro["year_collection"]))
47
48     tiempo_promedio = 0
49
50     for depto in deptos:
51         deptos[depto]["promedio"] = round((deptos[depto]["diferencia"] / deptos[depto]["total_registros"]), 2)
52         tiempo_promedio += deptos[depto]["promedio"]
53
54     tiempo_promedio = tiempo_promedio / len(deptos)
55
56     for depto in deptos:
57         al.add_last(array_list, deptos[depto])
58
59     if orden == "desc":
60         array_list = al.merge_sort(array_list, comp_fechas_desc)
61     elif orden == "asc":
62         array_list = al.merge_sort(array_list, comp_fechas_asc)
63
64     if num_estados > al.size(list_state_names):
65         num_estados = al.size(list_state_names)
66     lista_n_deptos = al.sub_list(array_list, 0, num_estados)
67
68     end = get_time()
69     tiempo_total = delta_time(start, end)
70
71     return tiempo_total, len(deptos), tiempo_promedio, menor_año_recoleccion(catalog), mayor_año_recoleccion(catalog), lista_n_deptos
72
```

El requerimiento 8 consiste en listar los N departamentos con mayor o menor promedio de tiempo entre la recopilación de los registros y su carga a la plataforma de entre todos los registros. Utiliza un recorrido lineal para calcular estadísticas por departamento y un ordenamiento para clasificarlos.

<b>Entrada</b>	Catalogo, Numero de departamentos, Orden
<b>Salidas</b>	Información de los N departamentos con mayor o menor diferencia promedio de tiempo de recolección y publicación de registros
<b>Implementado (Sí/No)</b>	Implementado grupalmente

## Análisis de complejidad

Análisis de complejidad de cada uno de los pasos del algoritmo

Pasos	Complejidad
Recorrer el catálogo para obtener registros	$O(n)$
Filtrar registros por unidad de medida	$O(1)$
Actualizar estadísticas por departamento	$O(1)$
Calcular el promedio de diferencia por depto	$O(1)$
Ordenar departamentos por promedio (Merge Sort)	$O(m \log m)$ M= Numero de departamentos
Seleccionar los N primeros o últimos depts	$O(1)$
<b>TOTAL</b>	<b><math>O(n)</math></b>

## Pruebas Realizadas

Las pruebas realizadas fueron realizadas en una maquina con las siguientes especificaciones. Los datos de entrada fueron Numero de departamentos: 5 y orden: desc (descendente)

<b>Procesadores</b>	AMD Ryzen 5 4500U
<b>Memoria RAM</b>	8 GB
<b>Sistema Operativo</b>	Windows 11

Entrada	Tiempo (ms)
Agricultural-20	311,038
Agricultural-40	981,553
Agricultural-60	595,508
Agricultural-80	817,264
Agricultural-100	1444,783

## Tablas de datos

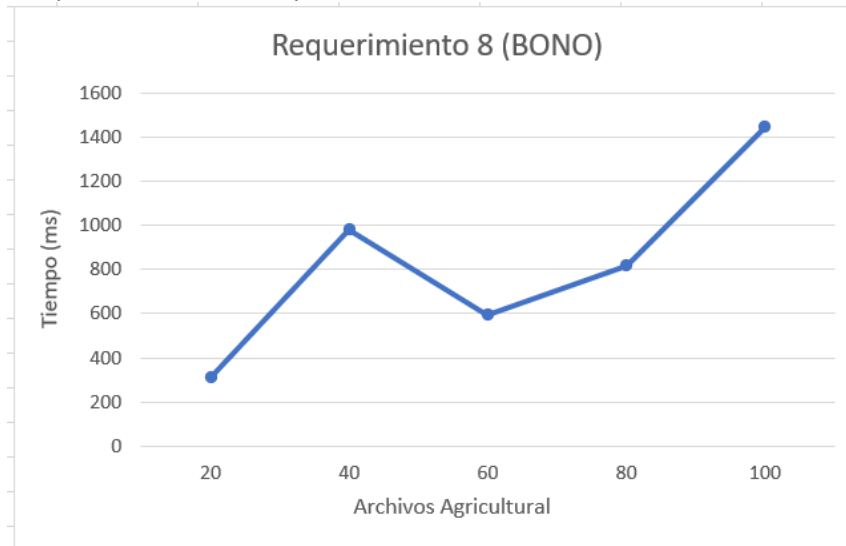
Las tablas con la recopilación de datos de las pruebas.

Muestra	Salida	Tiempo (ms)
Agricultural-20	Dato1	311,038
Agricultural-40	Dato2	981,553
Agricultural-60	Dato3	595,508

Agricultural-80	Dato4	817,264
Agricultural-100	Dato5	1444,783

## Graficas

Las gráficas con la representación de las pruebas realizadas.



## Análisis

Los tiempos de ejecución muestran un comportamiento lineal en general, aunque con algunas variaciones (por ejemplo, Agricultural-60 es más rápido que Agricultural-40). Esto confirma la eficiencia del algoritmo para conjuntos de datos grandes, ya que la complejidad lineal asegura un rendimiento escalable. Las fluctuaciones podrían deberse a la gestión de memoria o a la distribución de los datos en las pruebas.