Análisis de los planes de ejecución de la base de datos COVID 19“Listado de consultas a programar para analizar planes de ejecución”

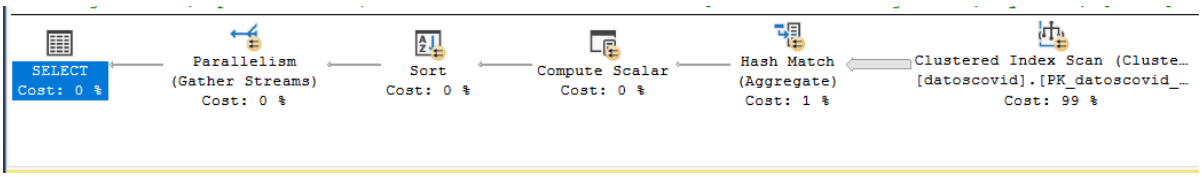
1.Listar los casos positivos por entidad de residencia

Consulta propia:

```
select ENTIDAD_RES,count(*) total_confimado from dbo.datoscovid
where CLASIFICACION_FINAL between 1 and 3
group by ENTIDAD_RES
order by ENTIDAD_RES
```

Consulta de otro equipo:

```
select ENTIDAD_RES, count(*) total_confirmado
from dbo.datoscovid
where CLASIFICACION_FINAL between 1 and 3
group by ENTIDAD_RES
order by ENTIDAD_RES
```



Clustered Index Scan (Clustered)

Scanning a clustered index, entirely or only a range.

Physical Operation	Clustered Index Scan
Logical Operation	Clustered Index Scan
Estimated Execution Mode	Batch
Storage	RowStore
Estimated Operator Cost	406.657 (99%)
Estimated I/O Cost	403.858
Estimated Subtree Cost	406.657
Estimated CPU Cost	2.79872
Estimated Number of Executions	1
Estimated Number of Rows to be Read	15265600
Estimated Number of Rows for All Executions	5613880
Estimated Number of Rows Per Execution	5613880
Estimated Row Size	19 B
Ordered	False
Node ID	4

Predicate
[covidHistorico].[dbo].[datoscovid].[CLASIFICACION_FINAL]>=(1) AND [covidHistorico].[dbo].[datoscovid].[CLASIFICACION_FINAL]<=(3)

Object
[covidHistorico].[dbo].[datoscovid].[PK_datoscovid_ID_Registro]

Output List
[covidHistorico].[dbo].[datoscovid].ENTIDAD_RES

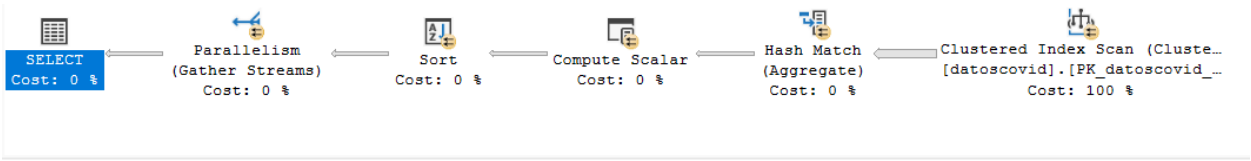
2. Listar los casos sospechosos por entidad

Consulta propia:

```
select ENTIDAD_UM, ENTIDAD_RES,count(*) total_sospechosos from dbo.datoscovid
where CLASIFICACION_FINAL = 6
group by ENTIDAD_RES, ENTIDAD_UM
order by ENTIDAD_UM
```

Consulta de otro equipo:

```
Select ENTIDAD_UM, ENTIDAD_RES, count(*) total_sospechosos
from dbo.datoscovid
where CLASIFICACION_FINAL = 6
group by ENTIDAD_UM, ENTIDAD_RES
order by ENTIDAD_UM
```



Clustered Index Scan (Clustered)

Scanning a clustered index, entirely or only a range.

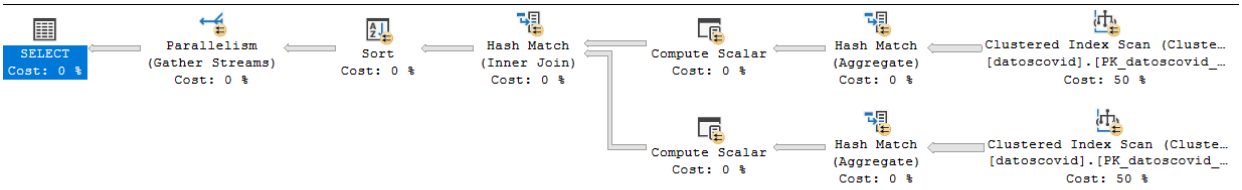
Physical Operation	Clustered Index Scan
Logical Operation	Clustered Index Scan
Estimated Execution Mode	Batch
Storage	RowStore
Estimated Operator Cost	406.657 (100%)
Estimated I/O Cost	403.858
Estimated Subtree Cost	406.657
Estimated CPU Cost	2.79872
Estimated Number of Executions	1
Estimated Number of Rows to be Read	15265600
Estimated Number of Rows for All Executions	572785
Estimated Number of Rows Per Execution	572785
Estimated Row Size	25 B
Ordered	False
Node ID	4

Predicate
[covidHistorico].[dbo].[datoscovid].[CLASIFICACION_FINAL]=(6)
Object
[covidHistorico].[dbo].[datoscovid].[PK_datoscovid_ID_Registro]
Output List
[covidHistorico].[dbo].[datoscovid].ENTIDAD_UM, [covidHistorico].[dbo].[datoscovid].ENTIDAD_RES

3. Listar el top 5 de municipios por entidad con el mayor número de casos reportados, indicando casos sospechosos y casos confirmados.

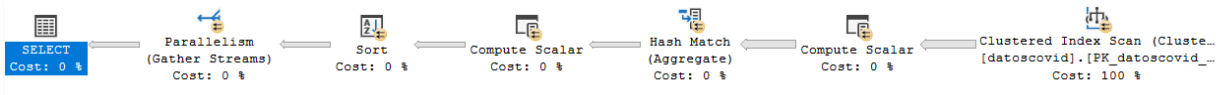
Consulta propia:

```
select cc.ENTIDAD_RES, cc.MUNICIPIO_RES, cc.confirmado, cs.sospechoso
from (select ENTIDAD_RES, MUNICIPIO_RES, count(*) as sospechoso
      from dbo.datoscovid where CLASIFICACION_FINAL = 6
      group by ENTIDAD_RES, MUNICIPIO_RES
      ) cs
inner join
(select ENTIDAD_RES, MUNICIPIO_RES, count (*) as confirmado
 from dbo.datoscovid where CLASIFICACION_FINAL between 1 and 3
 group by ENTIDAD_RES, MUNICIPIO_RES) cc
on cc.ENTIDAD_RES =  cs.ENTIDAD_RES and cs.MUNICIPIO_RES = cc.MUNICIPIO_RES
order by cc.ENTIDAD_RES
```



Consulta de otro equipo:

```
select ENTIDAD_RES,MUNICIPIO_RES,count(*) as reportados, count(case CLASIFICACION_FINAL when 1 then CLASIFICACION_FINAL
                                        when 2 then CLASIFICACION_FINAL
                                        when 3 then CLASIFICACION_FINAL
                                end) as confirmado,
            count(case CLASIFICACION_FINAL when 6 then CLASIFICACION_FINAL end) as sospechoso
from dbo.datoscovid
group by ENTIDAD_RES, MUNICIPIO_RES
order by ENTIDAD_RES, reportados desc
```



Clustered Index Scan (Clustered)

Scanning a clustered index, entirely or only a range.

Physical Operation	Clustered Index Scan
Logical Operation	Clustered Index Scan
Estimated Execution Mode	Batch
Storage	RowStore
Estimated Operator Cost	406.657 (100%)
Estimated I/O Cost	403.858
Estimated Subtree Cost	406.657
Estimated CPU Cost	2.79872
Estimated Number of Executions	1
Estimated Number of Rows to be Read	15265600
Estimated Number of Rows for All Executions	15265600
Estimated Number of Rows Per Execution	15265600
Estimated Row Size	27 B
Ordered	False
Node ID	5

Object

[covidHistorico].[dbo].[datoscovid].[PK_datoscovid_ID_Registro]

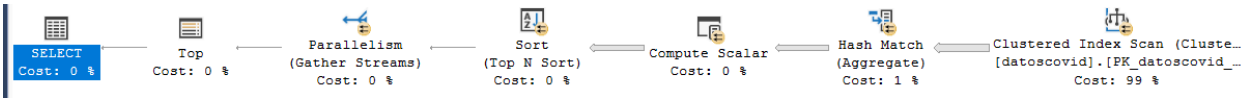
Output List

[covidHistorico].[dbo].[datoscovid].ENTIDAD_RES, [covidHistorico].[dbo].[datoscovid].MUNICIPIO_RES, [covidHistorico].[dbo].[datoscovid].CLASIFICACION_FINAL

4. Determinar el municipio con el mayor número de defunciones en casos confirmados.

Consulta propia:

```
select top 1 ENTIDAD_RES, MUNICIPIO_RES, count(*) as Difuntos from dbo.datoscovic
where FECHA_DEF !='9999-99-99'  and CLASIFICACION_FINAL between 1 and 3
group by ENTIDAD_RES, MUNICIPIO_RES
order by Difuntos desc
```



Clustered Index Scan (Clustered)

Scanning a clustered index, entirely or only a range.

Physical Operation	Clustered Index Scan
Logical Operation	Clustered Index Scan
Estimated Execution Mode	Batch
Storage	RowStore
Estimated Operator Cost	406.657 (99%)
Estimated I/O Cost	403.858
Estimated Subtree Cost	406.657
Estimated CPU Cost	2.79872
Estimated Number of Executions	1
Estimated Number of Rows to be Read	15265600
Estimated Number of Rows for All Executions	249476
Estimated Number of Rows Per Execution	249476
Estimated Row Size	49 B
Ordered	False
Node ID	5

Predicate

[covidHistorico].[dbo].[datoscovid].[CLASIFICACION_FINAL]>=(1) AND [covidHistorico].[dbo].[datoscovid].[CLASIFICACION_FINAL]<=(3) AND [covidHistorico].[dbo].[datoscovid].[FECHA_DEF]<>N'9999-99-99'

Object

[covidHistorico].[dbo].[datoscovid].[PK_datoscovid_ID_Registro]

Output List

[covidHistorico].[dbo].[datoscovid].ENTIDAD_RES, [covidHistorico].[dbo].[datoscovid].MUNICIPIO_RES

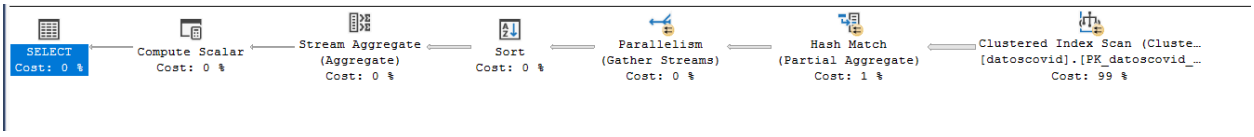
Consulta de otro equipo:

```
Select TOP 1 ENTIDAD_RES,MUNICIPIO_RES, count(case CLASIFICACION_FINAL when 1 then CLASIFICACION_FINAL
                                                 when 2 then CLASIFICACION_FINAL
                                                 when 3 then CLASIFICACION_FINAL
                                    end) as Confirmados_fallecidos
from dbo.datoscovid
where FECHA_DEF != '9999-99-99'
GROUP BY ENTIDAD_RES,MUNICIPIO_RES
ORDER BY Confirmados_fallecidos DESC;
```

5. Determinar por entidad, si de casos sospechosos hay defunciones reportadas asociadas a neumonía.

Consulta propia:

```
select ENTIDAD_RES,  count(*) from dbo.datoscovid
 where neumonia=1 and CLASIFICACION_FINAL=6 and FECHA_DEF !='9999-99-99'
 group by ENTIDAD_RES
```



Clustered Index Scan (Clustered)

Scanning a clustered index, entirely or only a range.

Physical Operation	Clustered Index Scan
Logical Operation	Clustered Index Scan
Estimated Execution Mode	Row
Storage	RowStore
Estimated Operator Cost	406.657 (99%)
Estimated I/O Cost	403.858
Estimated Subtree Cost	406.657
Estimated CPU Cost	2.79872
Estimated Number of Executions	1
Estimated Number of Rows to be Read	15265600
Estimated Number of Rows for All Executions	37332.3
Estimated Number of Rows Per Execution	37332.3
Estimated Row Size	45 B
Ordered	False
Node ID	6

Predicate

[covidHistorico].[dbo].[datoscovid].[NEUMONIA]=(1) AND
[covidHistorico].[dbo].[datoscovid].[CLASIFICACION_FINAL]=(6) AND
[covidHistorico].[dbo].[datoscovid].[FECHA_DEF]<>N'9999-99-99'

Object

[covidHistorico].[dbo].[datoscovid].[PK_datoscovid_ID_Registro]

Output List

[covidHistorico].[dbo].[datoscovid].ENTIDAD_RES

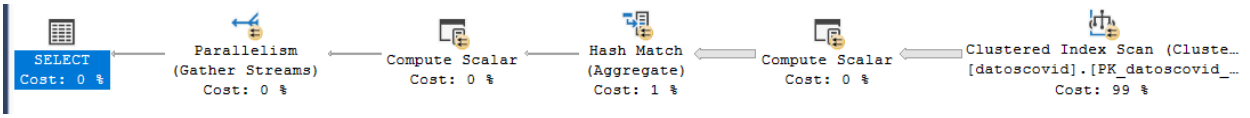
Consulta de otro equipo:

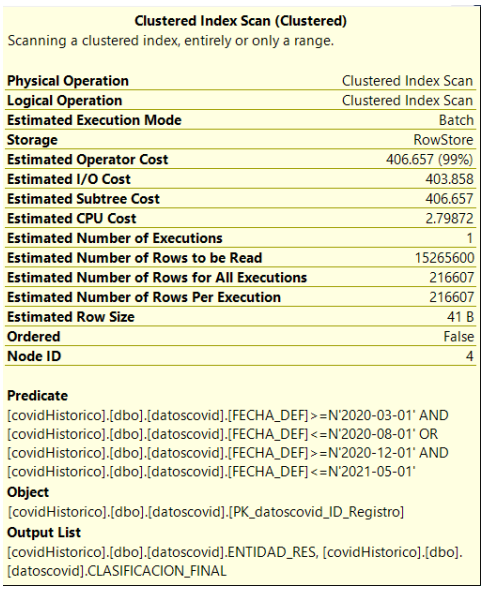
```
SELECT DISTINCT ENTIDAD_RES from dbo.datoscovid
where EXISTS(Select *from dbo.datoscovid Where NEUMONIA = 1 and FECHA_DEF != '9999-99-99')
GROUP BY ENTIDAD_RES
```

6. Listar por entidad el total de casos sospechosos, casos confirmados, total de defunciones en los meses de marzo a agosto 2020 y de diciembre 2020 a mayo 2021.

Consulta propia:

```
select ENTIDAD_RES, count(*) as difuntos, count(case CLASIFICACION_FINAL when 1 then CLASIFICACION_FINAL
                                                when 2 then CLASIFICACION_FINAL
                                                when 3 then CLASIFICACION_FINAL end)
as confirmado, count(case CLASIFICACION_FINAL when 6 then CLASIFICACION_FINAL end) as sospechoso
from dbo.datoscovid
where (FECHA_DEF between '2020-03-01' and '2020-08-01') or (FECHA_DEF between '2020-12-01' and '2021-05-01')
GROUP BY ENTIDAD_RES
```





Clustered Index Scan (Clustered)

Scanning a clustered index, entirely or only a range.

Physical Operation	Clustered Index Scan
Logical Operation	Clustered Index Scan
Estimated Execution Mode	Batch
Storage	RowStore
Estimated Operator Cost	406.657 (99%)
Estimated I/O Cost	403.858
Estimated Subtree Cost	406.657
Estimated CPU Cost	2.79872
Estimated Number of Executions	1
Estimated Number of Rows to be Read	15265600
Estimated Number of Rows for All Executions	216607
Estimated Number of Rows Per Execution	216607
Estimated Row Size	41 B
Ordered	False
Node ID	4

Predicate

[covidHistorico].[dbo].[datoscovid].[FECHA_DEF]>=N'2020-03-01' AND [covidHistorico].[dbo].[datoscovid].[FECHA_DEF]<=N'2020-08-01' OR [covidHistorico].[dbo].[datoscovid].[FECHA_DEF]>=N'2020-12-01' AND [covidHistorico].[dbo].[datoscovid].[FECHA_DEF]<=N'2021-05-01'

Object

[covidHistorico].[dbo].[datoscovid].[PK_datoscovid_ID_Registro]

Output List

[covidHistorico].[dbo].[datoscovid].ENTIDAD_RES, [covidHistorico].[dbo].[datoscovid].CLASIFICACION_FINAL

Consulta de otro equipo:

```
select ENTIDAD_RES, count(*) as fallecidos, count(case CLASIFICACION_FINAL when 1 then CLASIFICACION_FINAL
                                                    when 2 then CLASIFICACION_FINAL
                                                    when 3 then CLASIFICACION_FINAL
                                end) as confirmado,
                                count(case CLASIFICACION_FINAL when 6 then CLASIFICACION_FINAL end) as sospechoso
from dbo.datoscovid
where (FECHA_DEF between '2020-03-01' and '2020-08-31') or (FECHA_DEF between '2020-12-01' and '2021-05-31')
GROUP BY ENTIDAD_RES
```

7. Listar los 5 municipios con el mayor número de casos confirmados en niños menos de 13 años con alguna comorbilidad reportada y cuantos de esos casos fallecieron.

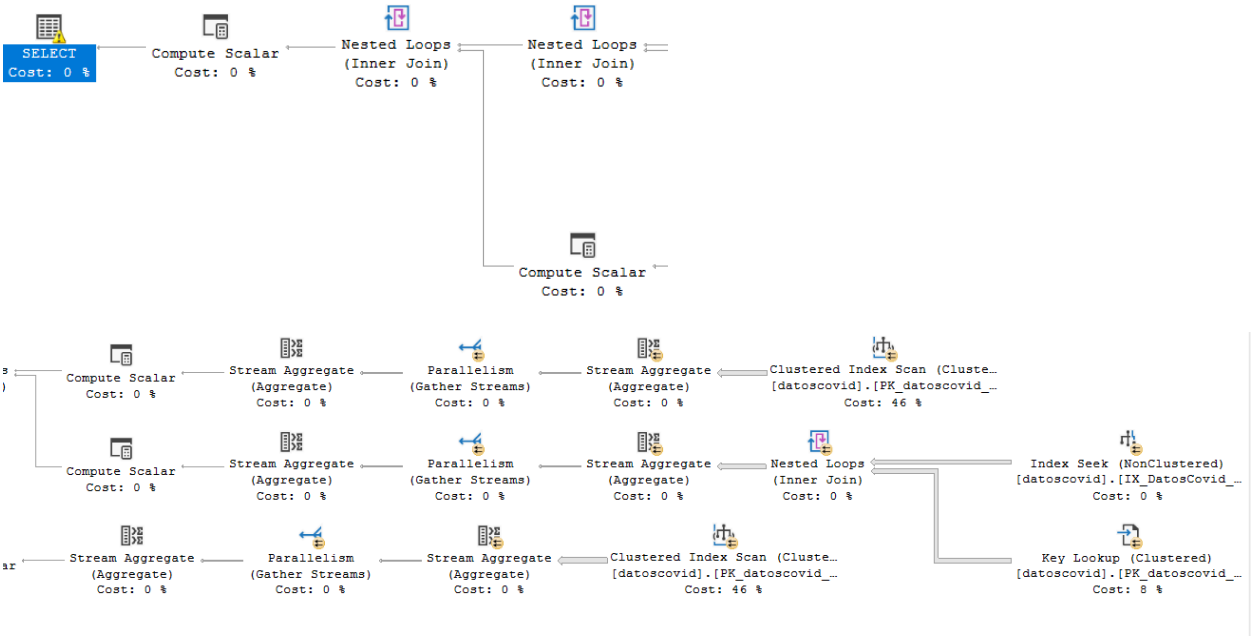
Consulta propia:

Consulta de otro equipo:

8. Determinar si en el año 2020 hay una mayor cantidad de defunciones menores de edad que en el año 2021 y 2022.

Consulta propia:

```
select
case  when sqrd20.DIFUNTOS2020>sqrd21.DIFUNTOS2021+sqrd22.DIFUNTOS2022 then 'fue mayor'
else 'no fue mayor'
end
from (select count(*) as DIFUNTOS2020 from dbo.datoscovid where edad<18 and FECHA_DEF  between '2020-01-01' and '2020-12-31') sqrd20,
(select count(*) as DIFUNTOS2021 from dbo.datoscovid where edad<18 and FECHA_DEF  between '2021-01-01' and '2021-12-31') sqrd21,
(select count(*) as DIFUNTOS2022 from dbo.datoscovid where edad<18 and FECHA_DEF  between '2022-01-01' and '2022-12-31') sqrd22
```



Key Lookup (Clustered)

Uses a supplied clustering key to lookup on a table that has a clustered index.

Physical Operation	Key Lookup
Logical Operation	Key Lookup
Estimated Execution Mode	Row
Storage	RowStore
Estimated I/O Cost	0.003125
Estimated Operator Cost	66.8977 (8%)
Estimated CPU Cost	0.0001581
Estimated Subtree Cost	66.8977
Estimated Number of Executions	20748.3
Estimated Number of Rows for All Executions	235789905.69
Estimated Number of Rows Per Execution	11364.3
Estimated Row Size	18 B
Ordered	True
Node ID	21

Predicate

CONVERT_IMPLICIT(int,[covidHistorico].[dbo].[datoscovid].[EDAD],0)<(18)

Object

[covidHistorico].[dbo].[datoscovid].[PK_datoscovid_ID_Registro]

Seek Predicates

Seek Keys[1]: Prefix: [covidHistorico].[dbo].[datoscovid].ID_REGISTRO = Scalar Operator([covidHistorico].[dbo].[datoscovid].[ID_REGISTRO])

Consulta de otro equipo:

```
select  Defunciones_2020.d_20,Defunciones_2021.d_21,Defunciones_2022.d_22,
        case  when Defunciones_2020.d_20>Defunciones_2021.d_21+Defunciones_2022.d_22 then 'Es mayor a 2021+2022'
        else 'Es menor a 2021+2022'
end as vs_SUMA_2122, case  when Defunciones_2020.d_20>Defunciones_2021.d_21 then 'Es mayor a 2021'
        else 'Es menor a 2021'
end as vs_2021, case  when Defunciones_2020.d_20>Defunciones_2022.d_22 then 'Es mayor a 2022'
        else 'Es menor a 2022'
end as vs_2022
    from(select count(*) as d_20 from dbo.datoscovid where edad<18 and FECHA_DEF  between '2020-01-01' and '2020-12-31') Defunciones_2020,
        (select count(*) as d_21 from dbo.datoscovid where edad<18 and FECHA_DEF  between '2021-01-01' and '2021-12-31') Defunciones_2021,
        (select count(*) as d_22 from dbo.datoscovid where edad<18 and FECHA_DEF  between '2022-01-01' and '2022-12-31') Defunciones_2022
```

9. Determinar si en el año 2021 hay un porcentaje mayor al 60 de casos reportados que son confirmados por estudios de laboratorio en comparación al año 2020.

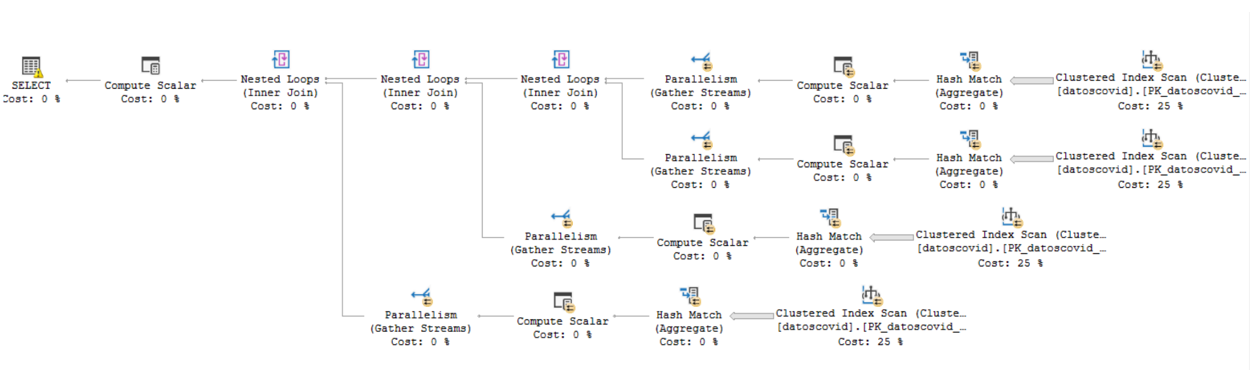
Consulta :

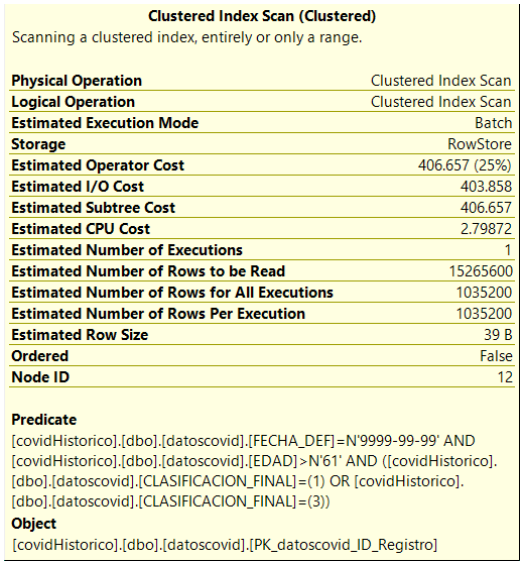
```sql
Select t1.Confirmados_2020,t2.Confirmados_2021,(((t2.Confirmados_2021-t1.Confirmados_2020)*100)/Confirmados_2020)as Porcentaje
from(select count(case CLASIFICACION_FINAL when 1 then CLASIFICACION_FINAL
                                           when 2 then CLASIFICACION_FINAL
                                           when 3 then CLASIFICACION_FINAL
                              end) as Confirmados_2020
                              from dbo.datoscovid where FECHA_SINTOMAS between '2020-01-01' and '2020-12-31') as t1,
                              (select count(case CLASIFICACION_FINAL when 1 then CLASIFICACION_FINAL
                                           when 2 then CLASIFICACION_FINAL
                                           when 3 then CLASIFICACION_FINAL
                              end) as Confirmados_2021
                              from dbo.datoscovid where FECHA_SINTOMAS between '2021-01-01' and '2021-12-31') as t2
```

10. Determinar en qué rango de edad: menor de edad, 19 a 40, 40 a 60 o mayor de 60 hay más casos reportados que se hayan recuperado.

Consulta propia:

```sql
select
case when ran1.Recuperado>ran2.Recuperado and ran1.Recuperado>ran3.Recuperado and ran1.Recuperado>ran4.Recuperado
    then 'El range de menor de edad tiene mas recuperados'
when  ran2.Recuperado>ran1.Recuperado and ran2.Recuperado>ran3.Recuperado and ran2.Recuperado>ran4.Recuperado
    then 'El rango de 18 a 40 tiene mas recuperados'
when  ran3.Recuperado>ran1.Recuperado and ran3.Recuperado>ran2.Recuperado and ran3.Recuperado>ran4.Recuperado
    then 'El rango de 18 a 40 tiene mas recuperados'
when  ran4.Recuperado>ran1.Recuperado and ran4.Recuperado>ran2.Recuperado and ran4.Recuperado>ran3.Recuperado
    then 'El rango de 18 a 40 tiene mas recuperados'
else 'error'
end
from
(select count(*) as Recuperado from dbo.datoscovid where FECHA_DEF ='9999-99-99' and (CLASIFICACION_FINAL= '1'
    or CLASIFICACION_FINAL = '3') and EDAD<18) ran1,
(select count(*) as Recuperado from dbo.datoscovid where FECHA_DEF ='9999-99-99' and (CLASIFICACION_FINAL= '1'
    or CLASIFICACION_FINAL = '3') and EDAD between '18' and '40') ran2,
(select count(*) as Recuperado from dbo.datoscovid where FECHA_DEF ='9999-99-99' and (CLASIFICACION_FINAL= '1'
    or CLASIFICACION_FINAL = '3') and EDAD between '41' and '60'   ) ran3,
(select count(*) as Recuperado from dbo.datoscovid where FECHA_DEF ='9999-99-99' and (CLASIFICACION_FINAL= '1'
    or CLASIFICACION_FINAL = '3') and EDAD >'61'  ) ran4
```





Clustered Index Scan (Clustered)

Scanning a clustered index, entirely or only a range.

Physical Operation	Clustered Index Scan
Logical Operation	Clustered Index Scan
Estimated Execution Mode	Batch
Storage	RowStore
Estimated Operator Cost	406.657 (25%)
Estimated I/O Cost	403.858
Estimated Subtree Cost	406.657
Estimated CPU Cost	2.79872
Estimated Number of Executions	1
Estimated Number of Rows to be Read	15265600
Estimated Number of Rows for All Executions	1035200
Estimated Number of Rows Per Execution	1035200
Estimated Row Size	39 B
Ordered	False
Node ID	12

Predicate

[covidHistorico].[dbo].[datoscovid].[FECHA_DEF]=N'9999-99-99' AND [covidHistorico].[dbo].[datoscovid].[EDAD]>N'61' AND ([covidHistorico].[dbo].[datoscovid].[CLASIFICACION_FINAL]=(1) OR [covidHistorico].[dbo].[datoscovid].[CLASIFICACION_FINAL]=(3))

Object

[covidHistorico].[dbo].[datoscovid].[PK_datoscovid_ID_Registro]

Consulta de otro equipo:

```
select
    case when R1.P_recuperadas>R2.P_recuperadas and R1.P_recuperadas>R3.P_recuperadas and R1.P_recuperadas>R4.P_recuperadas
           then 'Menores de edad se recuperan mas'
         when R2.P_recuperadas>R1.P_recuperadas and R2.P_recuperadas>R3.P_recuperadas and R2.P_recuperadas>R4.P_recuperadas
           then '19 a 40 años se recuperan mas'
         when R3.P_recuperadas>R1.P_recuperadas and R3.P_recuperadas>R2.P_recuperadas and R3.P_recuperadas>R4.P_recuperadas
           then '41 a 60 años se recuperan mas'
         when R4.P_recuperadas>R1.P_recuperadas and R4.P_recuperadas>R2.P_recuperadas and R3.P_recuperadas>R4.P_recuperadas
         then 'Mayores de 60 se recuperan mas'
end as Supervivencia
from
        (select count(*) as P_recuperadas from dbo.datoscovid where FECHA_DEF ='9999-99-99' and (CLASIFICACION_FINAL= '1'
            or CLASIFICACION_FINAL = '3') and (EDAD<=18) AND TIPO_PACIENTE!=2) R1,
        (select count(*) as P_recuperadas from dbo.datoscovid where FECHA_DEF ='9999-99-99' and (CLASIFICACION_FINAL= '1'
        or CLASIFICACION_FINAL = '3') and (EDAD between '19' and '40') AND TIPO_PACIENTE!=2) R2,
        (select count(*) as P_recuperadas from dbo.datoscovid where FECHA_DEF ='9999-99-99' and (CLASIFICACION_FINAL= '1'
        or CLASIFICACION_FINAL = '3') and (EDAD between '41' and '60') AND TIPO_PACIENTE!=2 ) R3,
        (select count(*) as P_recuperadas from dbo.datoscovid where FECHA_DEF ='9999-99-99' and (CLASIFICACION_FINAL= '1'
        or CLASIFICACION_FINAL = '3') and (EDAD >'60') AND TIPO_PACIENTE!=2) R4
```

Conclusiones

En este ejercicio fue un poco más complejo la decisión de establecer índices que ayuden a logar resultados con mayor optimización para las soluciones que se llegaron de las consultas propuestas, debido al diccionario de datos y la forma en la que los mismos están estructurados resulto una tarea con mayor dificultad en comparación con las antes vistas en clase, esto principalmente a que la información se manejaba en distintas tablas, las cuales brindaban mayor posibilidad de obtener fácilmente índices en base a las llaves primarias, esto queda ejemplificado con el hecho de que se decide tomar como índice agrupado/llave primaria