

# UNIDAD PROFESIONAL INTERDISCIPLINARIA EN INGENIERÍA Y TECNOLOGÍAS AVANZADAS UPIITA

PROFESOR: Carlos De La Cruz Sosa

**ASIGNATURA:** Bases de Datos Distribuidas

**GRUPO:** 3TM3

## Práctica I

# **EQUIPO 5:**

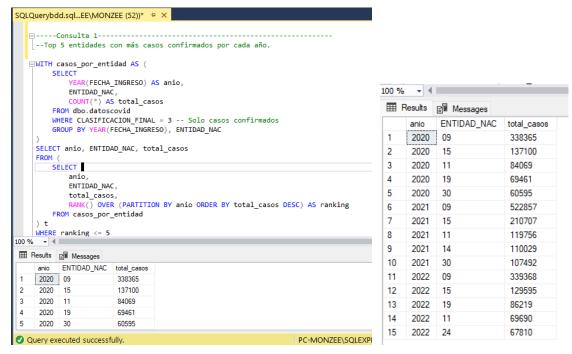
- Bernal Aguilar Yuvia Abigail
- Contreras Jimenez Mariana Montserrat
  - Medina Gómez Jimena Zarahí

**FECHA:** 06/03/2025

**Descripción:** En esta práctica aplicará los conceptos de programación de consultas en SQL revisados en las sesiones de clase del curso.

## Resultados:

### Consulta 1:



No. Consulta	1	
Descripción	Top 5 de las entidades con más casos confirmados por cada uno de los años registrados en la BDD.	
Significado de los valores de los catálogos	<ul> <li>ENTIDAD_NAC - Estado de nacimiento de una persona, su valor es un código asociado a una entidad federativa.</li> <li>FECHA_INGRESO - Fecha en el que el paciente fue registrado en la BDD.</li> <li>CLASIFICACIÓN_FINAL - El estado del caso = 3, significa "caso confirmado por prueba de laboratorio".</li> </ul>	
Responsable	Contreras Jiménez Mariana Montserrat	
Comentarios	<ul> <li>TOP 5 y ORDER BY - Se utilizó ORDER BY para ordenar los casos de mayor a menor y TOP para obtener solo las cinco entidades con más casos confirmados.</li> <li>YEAR(FECHA_INGRESO) - Se usó la función YEAR() para agrupar los casos por año y permitir un análisis anaual.</li> </ul>	

## Consulta 2:

```
SQLQuerybdd.sql...EE\MONZEE (52))* + ×
   -----Consulta 2-----
     --Municipio con más casos confirmados recuperados por estado y por año.
   SELECT
YEAR(FECHA_INGRESO) AS anio,
           ENTIDAD NAC
           MUNICIPIO_RES,
     COUNT(*) AS total_confirmados
FROM dbo.datoscovid
WHERE CLASIFICACION_FINAL = 3 -- Solo casos confirmados
GROUP BY YEAR(FECHA_INGRESO), ENTIDAD_NAC, MUNICIPIO_RES
HAVING COUNT(*) = (
SELECT MAX(casos)
FROM (
           FROM (
SELECT COUNT(*) AS casos
                                                                                                          100 % + 4
                FROM dbo.datoscovid
WHERE CLASIFICACION_FINAL = 3
GROUP BY YEAR(FECHA_INGRESO), ENTIDAD_NAC, MUNICIPIO_RES
                                                                                                           Results Messages
                                                                                                                 anio ENTIDAD_NAC MUNICIPIO_RES total_confirmados
          ) AS max_casos
                                                                                                                 2021 09
                                                                                                                                                007
                                                                                                                                                                      76158
     ORDER BY anio, ENTIDAD_NAC;
```

No. Consulta	2	
Descripción	Municipio con más casos confirmados recuperados por estado y por año.	
Significado de los valores de los catálogos	<ul> <li>ENTIDAD_NAC - Estado de nacimiento de una persona, su valor es un código asociado a una entidad federativa.</li> <li>FECHA_INGRESO - Fecha en el que el paciente fue registrado en la BDD.</li> <li>CLASIFICACIÓN_FINAL - El estado del caso = 3, significa "caso confirmado por prueba de laboratorio".</li> <li>MUNICIPIO_RES - Municipio de residencia del paciente.</li> </ul>	
Responsable	Contreras Jiménez Mariana Montserrat	
Comentarios	<ul> <li>GROUP BY y MAX() - Utilizadas para encontrar el municipio con más casos en cada estado.</li> <li>ORDER BY y LIMIT 1 - Subconsulta utilizada para obtener el municipio con más casos por cada estado, asegurando que solo se devolviera uno</li> </ul>	

## Consulta 3:

```
Dispetes' AS morbilidad,

(COUNT(CASE WHEN DIABETES = 1 AND CLASIFICACION_FINAL = 3 THEN 1 END) * 100.0) / COUNT(CASE WHEN CLASIFICACION_FINAL = 3 THEN 1 END) AS porcenta FROM dbo.datoscovid

UNION ALL

SELECT

'Obesidad' AS morbilidad,

(COUNT(CASE WHEN DESIDAD = 1 AND CLASIFICACION_FINAL = 3 THEN 1 END) * 100.0) / COUNT(CASE WHEN CLASIFICACION_FINAL = 3 THEN 1 END) AS porcenta FROM dbo.datoscovid

UNION ALL

SELECT

'Hipertensión' AS morbilidad,

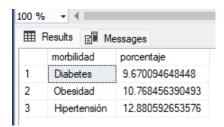
(COUNT(CASE WHEN DESIDAD = 1 AND CLASIFICACION_FINAL = 3 THEN 1 END) * 100.0) / COUNT(CASE WHEN CLASIFICACION_FINAL = 3 THEN 1 END) AS porcenta FROM dbo.datoscovid

UNION ALL

SELECT

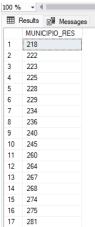
'Hipertensión' AS morbilidad,

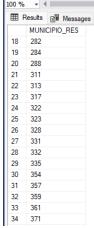
(COUNT(CASE WHEN HIPERTENSION = 1 AND CLASIFICACION_FINAL = 3 THEN 1 END) * 100.0) / COUNT(CASE WHEN CLASIFICACION_FINAL = 3 THEN 1 END) AS porcenta FROM dbo.datoscovid;
```

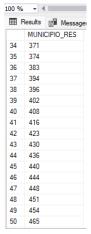


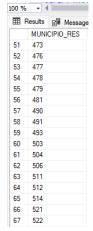
No. Consulta	3	
Descripción	Porcentaje de casos confirmados en diabetes, obesidad e hipertensión.	
Significado de los valores de los catálogos	<ul> <li>DIABETES - Indica si el paciente tiene diabetes. (1 = Sí, 2 = No).</li> <li>OBESIDAD - Indica si el paciente tiene obesidad (1 = Sí, 2 = No).</li> <li>CLASIFICACIÓN_FINAL - El estado del caso = 3, significa "caso confirmado por prueba de laboratorio".</li> </ul>	
Responsable	Contreras Jiménez Mariana Montserrat	
Comentarios	<ul> <li>COUNT(*) en comnbinación con SUM(CASE WHEN) para calcular la cantidad de pacientes con cada morbilidad y poder obtener el porcentaje.</li> </ul>	

#### Consulta 4:









Ⅲ	Results	Messages
	MUNI	CIPIO_RES
57	490	
58	491	
59	493	
60	503	
61	504	
62	506	
63	511	
64	512	
65	514	
66	521	
67	522	
68	527	
69	529	
70	541	
71	543	
72	564	
73	566	

No. Consulta	4	
Descripción	Municipios que no tienen casos confirmados en hipertensión, obesidad, diabetes, tabaquismo.	
Significado de los valores de los catálogos	<ul> <li>DIABETES, OBESIDAD, HIPERTENSION, TABAQUISMO         <ul> <li>Indican si el paciente tiene cada una de estas condiciones marcando 1 = Sí, 2 = No.</li> </ul> </li> <li>CLASIFICACIÓN_FINAL - Casos confirmados = 3.</li> </ul>	
Responsable	Contreras Jiménez Mariana Montserrat	
Comentarios	<ul> <li>NOT EXISTS se usa en subconsulta y sirve para excluir municipios donde haya al menos un caso con cuatro morbilidades.</li> <li>GROUP BY se utilizó para agrupar por municipio y filtrar aquellos que cumplen la condición.</li> </ul>	

## Consulta 5:

⊞ Results		
	estado	total_casos_con_neumonia
1	09	77254
2	15	56500
3	21	32046
4	30	27328
5	11	22587
6	14	19922
7	16	17528
8	12	16200
9	13	15239
10	25	15033
11	19	13824
12	24	13385
13	20	13268
14	26	12833
15	27	10854
16	08	10466
17	05	9901
18	02	8658
19	28	8386
20	22	8188

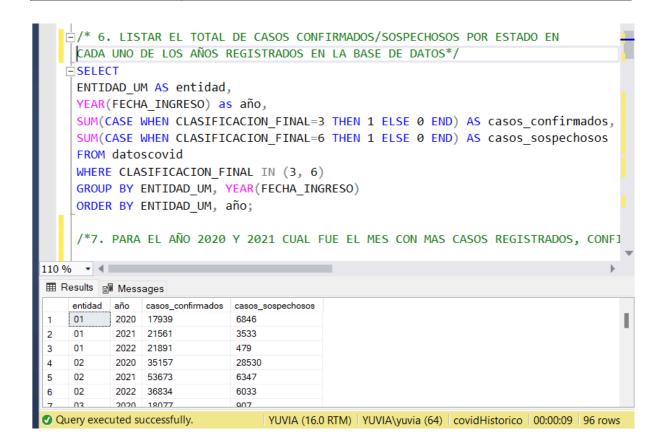
20	22	8188
21	31	8069
22	17	7161
23	32	6366
24	29	6070
25	10	5392
26	07	4619
27	18	4205
28	04	4031
29	01	3760
30	03	2628
31	06	2509
32	23	2275
33	99	1562

No. Consulta	5	
Descripción	Estados con más casos recuperados con neumonía.	
Significado de los valores de los catálogos	<ul> <li>ENTIDAD_NAC - Estado de nacimiento de una persona, su valor es un código asociado a una entidad federativa.</li> <li>NEUMONIA - Indica si el paciente tiene neumonía, en dado caso que sí = 1, 2 = No.</li> <li>CLASIFICACIÓN_FINAL - El estado del caso = 3, significa "caso confirmado por pruebade laboratorio".</li> <li>FECHA_DEF - Si es NULL, se asume que el paciente no falleció, es decir, posible recuperado.</li> </ul>	
Responsable	Contreras Jiménez Mariana Montserrat	
Comentarios	<ul> <li>FECHA_DEF IS NULL, se elimina este filtro ya que no se puede determinar exactamente quiénes eran recuperados, de esta forma se pudo obtener la cantidad total de casos con neumonía.</li> <li>COUNT(*) y GROUP BY se utilizó para contar los casos por estado y determinar cuáles tienen más casos.</li> <li>Estrategia: Ya que no hay una columna clara para poder identificar recuperados, contamos todos los casos confirmados con neumonía por estado sin filtrar los fallecidos, es decir, mostramos todos los casos de neumonía y luego revisamos las recuperaciones.</li> </ul>	

## Consulta 6:

No. Consulta	6	
Descripcion	Listar el total de casos confirmados/sospechosos por estado en cada uno de los años registrados en la base de datos	
Significado de los valores de los catalogos	<ul> <li>3 = Casos confirmados de columna</li> <li>CLASIFICACION_FINAL</li> <li>6 = Casos sospechosos de columna</li> <li>CLASIFICACION_FINAL</li> </ul>	
Responsable de la consulta	Bernal Aguilar Yuvia Abigail	
Comentarios	SUM: Se uso para sumar el total de casos confirmados y sospechosos.  CASE WHEN: Determinaba que tipo de caso era dependiendo su numero correspondiente en CLASIFICACION_FINAL, si era 3 o 6, sumaba 1 si no, suma 0.  WHERE: Filtra los datos requeridos.  GROUP BY: Agrupa los datos por entidad y año.	

**ORDER BY:** Ordena el resultado por entidad seguido del año de forma ascendente.



## Consulta 7:

No. Consulta	7	
Descripcion	Para el año 2020 y 2021 cual fue el mes con más casos registrados, confirmados, sospechosos, por estado registrado en la base de datos.	
Significado de los valores de los catalogos	<ul> <li>3 = Casos confirmados de columna</li> <li>CLASIFICACION_FINAL</li> <li>6 = Casos sospechosos de columna</li> <li>CLASIFICACION_FINAL</li> </ul>	
Responsable de la consulta	Bernal Aguilar Yuvia Abigail	
Comentarios	WITH CASO AS: Estructura para crear una tabla temporal. Unicamente existe para esta consulta	

**SUM:** Se uso para sumar el total de casos confirmados y sospechosos.

**CASE WHEN:** Determinaba que tipo de caso era dependiendo su numero correspondiente en CLASIFICACION\_FINAL, si era 3 o 6, sumaba 1 si no, suma 0.

YEAR(FECHA\_INGRESO) IN (2020, 2021): Solo toma datos de esos años.

WHERE: Filtra los datos requeridos.

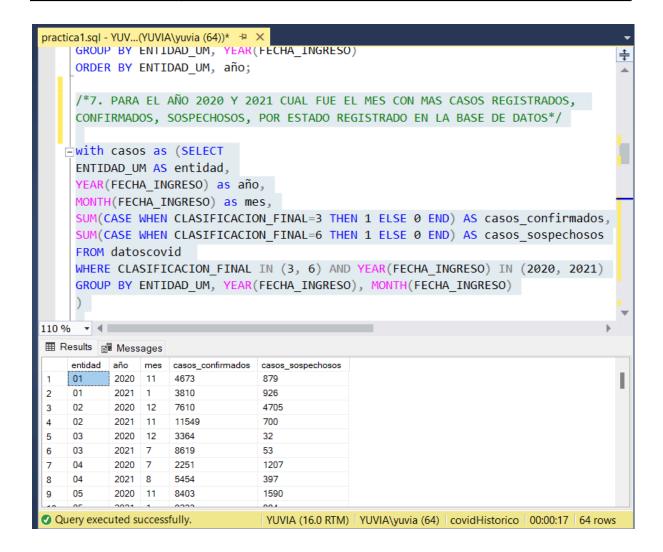
**GROUP BY:** Agrupa los datos por entidad, año y mes.

**MAX:** Se obtiene el mes con el mayor total de casos que se tiene al resultar la operacion que se establece entre parentesis.

JOIN: Combina la tabla temporal con una tabla.

**ON:** Establece que condiciones se necesita para unir estas tablas.

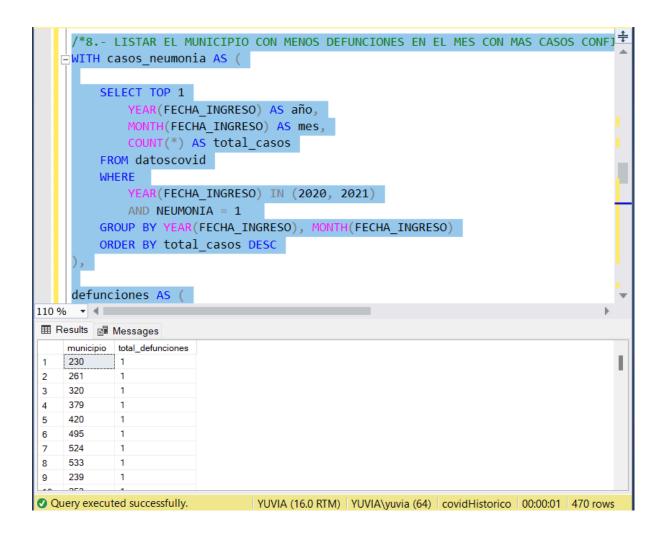
**ORDER BY:** Se ordena por entidad y año.



## Consulta 8:

No. Consulta	8

Descripción	Listar el municipio con menos defunciones en el mes con más caos confirmados con neumonia en los años 2020 y 2021
Significado de los valores de los catálogos	9999-99-99= Recuperacion del paciente de la columna FECHA_DEF
Responsable de la consulta	Bernal Aguilar Yuvia Abigail
Comentarios	WITH AS: Creacion de una tabla temporal. COUNT(*): Cuenta el total de registros donde hubo defuncion. WHERE: Filtra los datos requeridos. NEUMONIA=1: Solo de incluyen los casos que hubo neumonia. GROUP BY: Agrupa por año y mes. ORDER BY DESC: Ordena los resultados de mayor a menor. TOP 1: Con el orden descendente se toma el mes con mas casos de neumonia. YEAR(FECHA_INGRESO) IN (2020, 2021): Solo toma datos de esos años. FECHA_DEF <> '9999-99-99': Filtra los datos que sean diferentes al establecido entre comillas. SELECT *: Selecciona todas las columnas de la tabla. ORDER BY ASc: Ordena los resultados de menor a mayor.



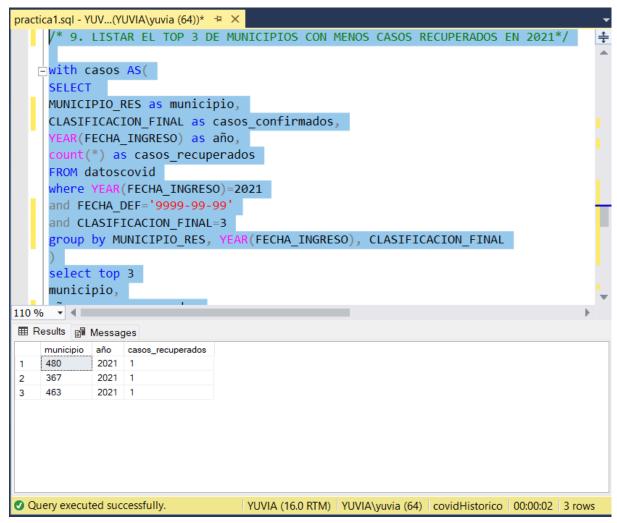
## Consulta 9:

No. Consulta	9
Descripción	Listar el top 3 de municipios con menos casos recuperados en el año 2021
Significado de los valores de los catálogos	<ul> <li>9999-99-99 = Recuperacion del paciente de la columna de FECHA_DEF</li> <li>3 = Casos confirmados de la columna CLASIFICACION_FINAL</li> </ul>
Responsable de la consulta	Bernal Aguilar Yuvia Abigail
Comentarios	WITH AS: Creacion de una tabla temporal. COUNT(*): Cuenta el total de registros donde hubo defuncion. WHERE: Filtra los datos requeridos. YEAR(FECHA_INGRESO): Especificamos el año que queremos que se extraigan los datos. FECHA_DEF = '9999-99-99': Indica que el paciente se recupero.

CLASIFICACION\_FINAL = 3: Filtra unicamente los casos confirmados.

SELECT TOP 3: Solo se seleccionar 3 filas con el menor numero de recuperados.

ORDER BY ... ASC: Ordena los resultados de menor a mayor.



## Consulta 10:

No. Consulta	10
Descripción	Listar el porcentaje de casos confirmados por genero en los años 2020 y 2021
Significado de los valores de los catálogos	<ul> <li>3 = Casos confirmados de la columna CLASIFICACION_FINAL</li> <li>1 = Paciente hombre de columna SEXO</li> <li>2 = Paciente mujer de columna SEXO</li> </ul>
Responsable de la consulta	Bernal Aguilar Yuvia Abigail

## **Comentarios**

WITH ... AS: Creacion de una tabla temporal.

**COUNT(\*) AS**: Se cuenta el numero total de registros.

WHERE: Filtramos

**CLASIFICACION\_FINAL = 3:** Solo se selecciona los casos

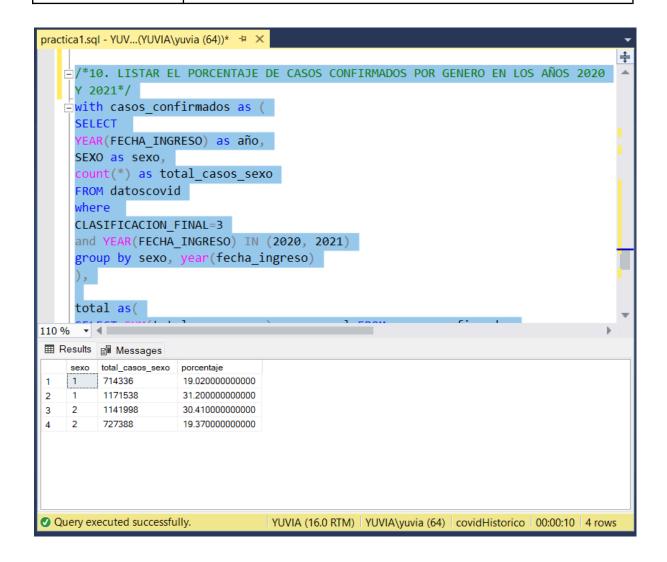
confirmados de COVID-19.

YEAR(FECHA\_INGRESO) IN (2020,2021): Unicamente se

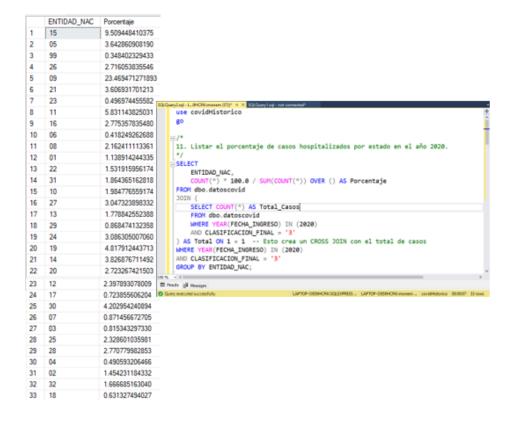
seleccionan los casos que tengan esos años.

**GROUP BY:** Agrupa los datos.

**SUM:** Suma el total de los casos requeridos.

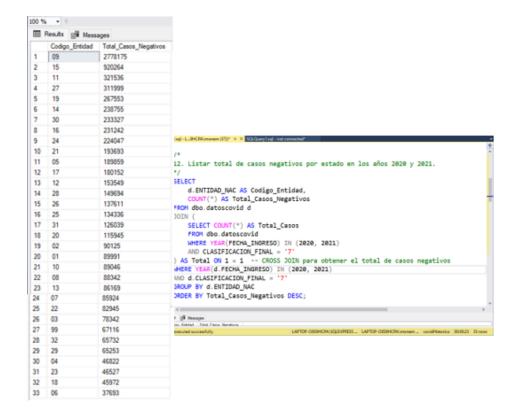


## Consulta 11:



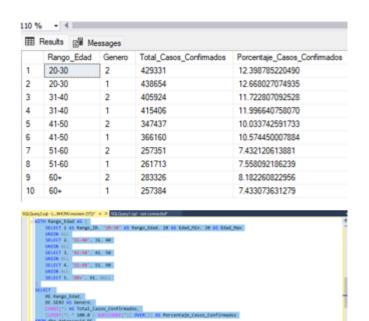
No. Consulta	11
Descripción	Listar el porcentaje de casos hospitalizados por estado en el año 2020.
Significado de los valores de los catálogos	<ul> <li>ENTIDAD_NAC: Estado de nacimiento según el código indicado.</li> <li>FECHA_INGRESO: La fecha en la que el paciente (usuario) se registra.</li> <li>CLASIFICACIÓN_FINAL: Estado del paciente, según el código es la condición. En esta consulta se utilizará el '3'.</li> </ul>
Responsable	Medina Gómez Jimena Zarahí
Comentarios	<ul> <li>COUNT(*): Cuenta el número de casos confirmados por entidad.</li> <li>SUM(COUNT(*)) OVER (): Suma total de casos confirmados en todas las entidades.</li> <li>COUNT(*) * 100.0 / SUM(COUNT(*)) OVER (): Calcula el porcentaje de casos por entidad respecto al total.</li> <li>Subconsulta: Se utiliza para localizar a los pacientes confirmados.</li> </ul>

## Consulta 12:



No. Consulta	12
Descripción	Listar total de casos negativos por estado en los años 2020 y 2021.
Significado de los valores de los catálogos	<ul> <li>ENTIDAD_NAC: Estado de nacimiento según el código indicado.</li> <li>FECHA_INGRESO: La fecha en la que el paciente (usuario) se registra.</li> <li>CLASIFICACIÓN_FINAL: Estado del paciente, según el código es la condición. En esta consulta se utilizará el '7'.</li> </ul>
Responsable	Medina Gómez Jimena Zarahí
Comentarios	<ul> <li>COUNT(*) AS Total_Casos_Negativos: Cuenta la cantidad de casos negativos.</li> <li>Subconsulta: Calcula el total de casos negativos (         CLASIFICACION_FINAL = '7') en los años 2020 y         2021.</li> <li>WHERE YEAR(d.FECHA_INGRESO) IN (2020, 2021):         Filtran los registros para considerar solo casos negativos.</li> </ul>

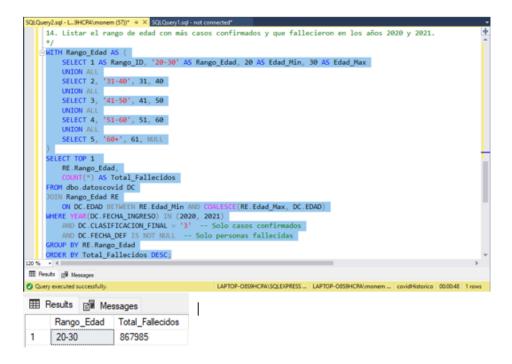
## Consulta 13:



RE Rango\_Edad, OC SERO OR BY RE Rango\_Edad, OC SERO DESC;

No. Consulta	13
Descripción	Listar porcentajes de casos confirmados por género en el rango de edades de 20 a 30 años, de 31 a 40 años, de 41 a 50 años, de 51 a 60 años y mayores a 60 años a nivel nacional.
Significado de los valores de los catálogos	<ul> <li>EDAD: Estado de nacimiento según el código indicado.</li> <li>SEXO: Genero del paciente. '1' -&gt; Hombre y '2' -&gt; Mujer</li> <li>FECHA_INGRESO: La fecha en la que el paciente (usuario) se registra.</li> <li>CLASIFICACIÓN_FINAL: Estado del paciente, según el código es la condición. En esta consulta se utilizará el '3'.</li> <li>WITH Rango_Edad AS: Rango de edades (20-30,31-40,41-50,51-60 y 60+):</li> </ul>
Responsable	Medina Gómez Jimena Zarahí
Comentarios	<ul> <li>WITH Rango_Edad AS (): Es una Expresión de Tabla Común (CTE, Common Table Expression).</li> <li>COUNT(*) AS Total_Casos_Confirmados: Cuenta la cantidad de casos confirmados por rango de edad y género.</li> <li>SUM(COUNT(*)) OVER(): Calcula el total de casos confirmados globalmente para luego obtener el porcentaje por grupo.</li> </ul>

### Consulta 14:



No. Consulta	14
Descripción	Listar el rango de edad con más casos confirmados y que fallecieron en los años 2020 y 2021.
Significado de los valores de los catálogos	<ul> <li>EDAD: Estado de nacimiento según el código indicado.</li> <li>FECHA_INGRESO: La fecha en la que el paciente (usuario) se registra.</li> <li>CLASIFICACIÓN_FINAL: Estado del paciente, según el código es la condición. En esta consulta se utilizará el '3'.</li> <li>FECHA_DEF: fecha en la que el paciente falleció.</li> <li>WITH Rango_Edad AS: Rango de edades (20-30,31-40,41-50,51-60 y 60+):</li> </ul>
Responsable	Medina Gómez Jimena Zarahí
Comentarios	<ul> <li>WITH Rango_Edad AS (): Es una Expresión de Tabla Común (CTE, Common Table Expression).</li> <li>COUNT(*) AS Total_Fallecidos: Cuenta la cantidad de casos confirmados de COVID-19 en los que el paciente falleció.</li> <li>SELECT TOP 1: Obtiene solo el rango de edad con más fallecimientos.</li> </ul>

## Conclusiones generales de la práctica:

En esta práctica, el enfoque principal fue desarrollar habilidades en el uso de SQL, aplicando técnicas de consulta y optimización de bases de datos distribuidas para obtener información relevante (consultas). Utilizamos diversas funciones en SQL como MAX(), MIN(), COUNT(), SUM(), y JOIN para realizar consultas complejas que permitieran extraer y organizar datos con criterios específicos. Aprendimos a aplicar filtros con WHERE para limitar los resultados a ciertos rangos de fechas o condiciones particulares, así como a utilizar GROUP BY para agrupar los resultados por categorías clave y calcular agregados como sumas y promedios.

El uso de funciones como <u>COALESCE</u> y el operador <u>JOIN</u> fue esencial para combinar datos de diferentes columnas en una sola tabla y obtener resultados más detallados y completos. Además, implementamos técnicas de ordenamiento con <u>ORDER BY</u> para organizar los resultados de manera que fueran fácilmente interpretables y la mayoría ordenados de mayor a menor (siempre que aplicara el caso). Las consultas también incluyeron la utilización de subconsultas y agregaciones para identificar análisis de ciertas clasificaciones (columnas) por edad y comorbilidades, lo que nos permite manejar grandes cantidades de información de forma estructurada y evitando datos "basura" o innecesarios para las consultas trabajadas.

Además, al trabajar con bases de datos distribuidas, tuvimos que asegurarnos de que nuestras consultas se enfocarán a las consultas correctas, algo que ayudó mucho es que la estructura de la base separó bastante claro cada columna con sus datos correctos así como aplicando filtros para garantizar que las consultas no afectarán el tiempo de respuesta. Esta práctica nos permitió comprender mejor la importancia de diseñar consultas optimizadas que no solo responden a la pregunta específica, sino que también fueran escalables y eficientes, especialmente cuando se manejan grandes volúmenes de datos.

Por último en cuanto a la organización del trabajo en equipo, utilizamos herramientas como GitHub para compartir el código y mantener un flujo de trabajo organizado. Además, la comunicación constante a través de WhatsApp permitió resolver dudas y la organización de quien realizaría que consultas así como apoyarnos con cualquier duda tanto de las consultas como el reporte de la práctica.