Práctica 1
Bases de Datos Distribuidas

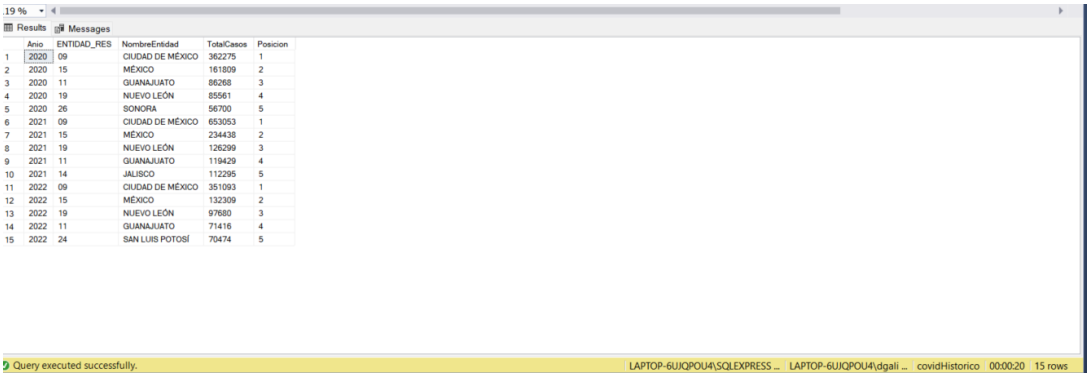
Equipo 7
Pérez López Leonardo
Galicia Cobaxin Daniel

6 de marzo de 2025

1. Consulta N° 1

Listar el top 5 de las entidades con más casos confirmados por cada uno de los años.

Información
Número de consulta: *#1 Listar el top 5 de las entidades con más casos confirmados por cada uno de los años.*
Requisitos: ▪ *Se utilizará la tabla* `datoscovid` *para el manejo de las cifras.* ▪ *Se utilizará la tabla* `cat_entidades` *para hacer coincidencia y saber exactamente a qué entidad en su atributo nombre estamos rankeando.* ▪ `FECHA_INGRESO`*: Para conocer el año de ingreso por enfermedad.* ▪ `ENTIDAD_RES`*: Para saber la entidad de residencia del paciente.* ▪ `CLASIFICACION_FINAL`*: Para recopilar solo los casos confirmados de covid.*
Significado de catálogos: ▪ `CLASIFICACION_FINAL`*: 1, 2, 3 refiere a casos positivos confirmados por diferentes medios.*
Responsable: *Daniel Galicia Cobaxin*
Comentarios: ▪ *Se utilizó* `WITH` *para crear una tabla temporal nombrada como* `Ranking`*, esto con ayuda del comando* `AS`*.* ▪ *Se nombró a la tabla* `datoscovid` *como* `d` *y a la tabla* `cat_entidades` *como* `c`*.* ▪ *Se recopiló la información de solamente el año del paciente siendo nombrado* `Anio`*, la entidad residencia y la suma total de los casos de dicha entidad nombrandolo como* `TotalCasos` *y siendo contados por el comando* `COUNT`*.* ▪ *Se ocupó el comando* `RANK` *para poder asignar un número de* `ranking` *a cada fila dentro de un grupo que en este caso está dado por* `GROUP BY` *la fecha de ingreso y la entidad residencia.* ▪ *Se dividen los datos según el año de la fecha ingreso, siendo ordenados de manera descendente y nombrando a esta columna como* `Posicion`*.* ▪ *Finalmente se recogen solamente las primeras 5 filas de cada año, dando así el top 5 del 2020, 2021 y 2022.* ▪ `RANK()`*: Asigna un número de ranking a cada fila dentro de un grupo de filas.* ▪ `PARTITION BY()`*: Divide los datos en grupos (o particiones) basados en una o más columnas.*



	Anio	ENTIDAD_RES	NombreEntidad	TotalCasos	Posicion
1	2020	09	CIUDAD DE MÉXICO	362275	1
2	2020	15	MÉXICO	161809	2
3	2020	11	GUANAJUATO	86268	3
4	2020	19	NUEVO LEÓN	85561	4
5	2020	26	SONORA	56700	5
6	2021	09	CIUDAD DE MÉXICO	653053	1
7	2021	15	MÉXICO	234438	2
8	2021	19	NUEVO LEÓN	126299	3
9	2021	11	GUANAJUATO	119429	4
10	2021	14	JALISCO	112295	5
11	2022	09	CIUDAD DE MÉXICO	351093	1
12	2022	15	MÉXICO	132309	2
13	2022	19	NUEVO LEÓN	97680	3
14	2022	11	GUANAJUATO	71416	4
15	2022	24	SAN LUIS POTOSÍ	70474	5

Figura 1: Resultado de la consulta 1

2. Consulta N° 2

Listar el municipio con más casos confirmados recuperados por estado y por año.

Información

Número de consulta: *#2 Listar el municipio con más casos confirmados recuperados por estado y por año.*

Requisitos:

- *Tabla* `datoscovid`.
- *Tabla* `cat_entidades`.
- `FECHA_INGRESO`: *Específicamente el año en el que ingresaron.*
- `ENTIDAD_RES`: *Entidades de residencia del paciente.*
- `entidad`: *Para visualizar el nombre de la entidad de residencia.*
- `MUNICIPIO_RES`: *Municipio de residencia del paciente.*
- `TIPO_PACIENTE`: *Para recopilar a los pacientes recuperados.*
- `FECHA_DEF`: *Para excluir a los fallecidos.*

Significado de catálogos:

- `TIPO_PACIENTE`: *1 para pacientes recuperados.*
- `FECHA_DEF`: *Formato '9999-99-99' para pacientes que no fallecieron.*

Responsable: *Daniel Galicia Cobaxin*

Comentarios:

- *Se crea una tabla temporal llamada* `Ranking`.
- *Se recopila solo el año de la* `FECHA_INGRESO` *y se nombra como* `Anio`.
- *Se nombra a la tabla* `datoscovid` *como* `d` *y a la tabla* `cat_entidades` *como* `c`.
- *Se recopila la entidad de residencia, entidad y el municipio de residencia.*
- *Se cuentan el total de casos con* `COUNT` *según la clasificación 1, 2, 3 en* `CLASIFICACION_FINAL`, *que el paciente sea de tipo 1 y la fecha de defunción coincida con el formato para los pacientes que no fallecieron.*
- *Se asigna un* `RANK` *por los grupos divididos según la fecha de ingreso y entidad, además de contarlos y ordenarlos de manera descendente, asignándole el nombre de* `Posicion`.
- *Con ayuda de un* `JOIN` *se hace coincidir la clave en* `datoscovid` *y* `cat_entidades` *para poder obtener el nombre exacto de la entidad de residencia.*
- *Se agrupan los datos por el año de la fecha de ingreso, entidad de residencia, entidad y municipio de residencia.*
- *Finalmente, solo se rescata el top 1 de cada estado, en cada año.*
- `RANK()`: *Asigna un número de ranking a cada fila dentro de un grupo de filas.*
- `PARTITION BY()`: *Divide los datos en grupos (o particiones) basados en una o más columnas.*

Results | Messages

	Anio	ENTIDAD_RES	NombreEntidad	MUNICIPIO_RES	TotalCasos
1	2020	01	AGUASCALIENTES	001	11402
2	2020	02	BAJA CALIFORNIA	002	11207
3	2020	03	BAJA CALIFORNIA SUR	003	8182
4	2020	04	CAMPECHE	002	2250
5	2020	05	COAHUILA DE ZARAGOZA	035	11525
6	2020	06	COLIMA	002	1706
7	2020	07	CHIAPAS	101	2357
8	2020	08	CHIHUAHUA	037	15909
9	2020	09	CIUDAD DE MÉXICO	007	43896
10	2020	10	DURANGO	005	12445
11	2020	11	GUANAJUATO	020	23751
12	2020	12	GUERRERO	001	8790
13	2020	13	HIDALGO	048	4174
14	2020	14	JALISCO	039	13383
15	2020	15	MÉXICO	033	13939
16	2020	16	MICHOACÁN DE OCAMPO	053	6622
17	2020	17	MORELOS	007	2315
18	2020	18	NAYARIT	017	3148
19	2020	19	NUEVO LEÓN	039	21369
20	2020	20	OAXACA	067	7671
21	2020	21	PUEBLA	114	24539
22	2020	22	QUERÉTARO	014	21753
23	2020	23	QUINTANA ROO	005	4961

Query executed successfully. LAPTOP-6UJQPOU4\SQLEXPRESS ... LAPTOP-6UJQPOU4\dgali ... covidHistorico 00:02:02 96 rows

Figura 2: Resultado de la consulta 2

3. Consulta N° 3

Listar el porcentaje de casos confirmados en cada una de las siguientes morbilidades a nivel nacional: diabetes, obesidad e hipertensión.

Información
Número de consulta: *#3 Listar el porcentaje de casos confirmados en cada una de las siguientes morbilidades a nivel nacional: diabetes, obesidad e hipertensión.*
Requisitos: ▪ *Tabla* `datoscovid`. ▪ `CLASIFICACION_FINAL`*: Para extraer solo casos confirmados.* ▪ *Columnas de morbilidades:* `Diabetes`, `Obesidad` *e* `Hipertension`.
Significado de catálogos: ▪ `CLASIFICACION_FINAL`*: Para filtrar solo los casos confirmados (*`IN (1, 2, 3)`*).* ▪ `DIABETES`*: Indica si el paciente tiene diabetes (1 = Sí).* ▪ `OBESIDAD`*: Indica si el paciente tiene obesidad (1 = Sí).* ▪ `HIPERTENSION`*: Indica si el paciente tiene hipertensión (1 = Sí).*
Responsable: *Daniel Galicia Cobaxin*
Comentarios: ▪ *Se utilizó la función* `SUM` *para sumar los casos en donde sí hay diabetes, obesidad o hipertensión (1).* ▪ *Cuando el caso es cierto, es decir, que hay alguna de las 3 morbilidades, se suma 1; de lo contrario, se suma 0.* ▪ *Al resultado de cada suma, se le multiplica por 100 y se divide entre el resultado de* `COUNT`*, es decir, de los casos confirmados.* ▪ *Finalmente, se obtienen solo los casos confirmados de* `CLASIFICACION_FINAL`.

Results | Messages

	Porcentaje_Diabetes	Porcentaje_Obesidad	Porcentaje_Hipertension
1	9.617175382487	10.567170432354	12.787018025657

Figura 3: Resultado de la consulta 3

4. Consulta N° 4

4. Listar los municipios que no tengan casos confirmados en todas las morbilidades: hipertensión, obesidad, diabetes, tabaquismo.

Información
Número de consulta: *#4 Listar los municipios que no tengan casos confirmados en todas las morbilidades: hipertensión, obesidad, diabetes, tabaquismo.*
Requisitos: ▪ *Tabla* `datoscovid` *y* `cat_entidades`*.* ▪ `entidad`*: De la tabla* `cat_entidades` *para obtener los nombres de los estados.* ▪ `MUNICIPIO_RES`*: El municipio de residencia de cada paciente para las operaciones.* ▪ `CLASIFICACION_FINAL`*: Para excluir a los que sí están confirmados.* ▪ `HIPERTENSION`*,* `OBESIDAD`*,* `DIABETES` *y* `TABAQUISMO`*: Para verificar que cumplan con la condición de 1 (confirmación).*
Significado de catálogos: ▪ `CLASIFICACION_FINAL`*: Para filtrar solo los casos confirmados (*`IN (1, 2, 3)`*). Estos son de confirmación.* ▪ `HIPERTENSION`*,* `OBESIDAD`*,* `DIABETES`*,* `TABAQUISMO`*: Indican si el paciente tiene estas morbilidades (1 = Sí).*
Responsable: *Daniel Galicia Cobaxin*
Comentarios: ▪ *Se obtienen de manera única los datos de* `entidad` *de la tabla* `cat_entidades`*.* ▪ *Se obtiene el* `MUNICIPIO_RES` *de la tabla* `datoscovid`*.* ▪ *Se unen las tablas para obtener el nombre exacto del estado.* ▪ *Se utiliza* `WHERE NOT EXISTS` *para excluir a los municipios donde haya al menos 1 caso confirmado con todas las morbilidades.* ▪ *Se agrupan por estado finalmente.*

Results | Messages

	Estado	Municipio
1	OAXACA	218
2	OAXACA	222
3	OAXACA	223
4	OAXACA	228
5	OAXACA	229
6	OAXACA	234
7	OAXACA	236
8	OAXACA	240
9	OAXACA	264
10	OAXACA	267
11	OAXACA	268
12	OAXACA	274
13	OAXACA	275
14	OAXACA	281
15	OAXACA	282
16	OAXACA	284
17	OAXACA	288
18	OAXACA	311
19	OAXACA	313
20	OAXACA	317
21	OAXACA	322
22	OAXACA	328
23	OAXACA	331

Query executed successfully. | LAPTOP-6UJQPOU4\SQLEXPRESS ... | LAPTOP-6UJQPOU4\dgali ... | covidHistorico | 00:00:54 | 53 rows

Figura 4: Resultado de la consulta 4

5. Consulta N° 5

Listar los estados con más casos recuperados con neumonía.

Información
Número de consulta: *#5. Listar los estados con más casos recuperados con neumonía.*
Requisitos: ▪ *Tabla* `datoscovid` *y* `cat_entidades`*.* ▪ `entidad`*: Para obtener los nombres de los estados.* ▪ `ENTIDAD_RES`*: Para las operaciones de conteo.* ▪ `NEUMONIA`*: Condición para el conteo.*
Significado de catálogos: ▪ `CLASIFICACION_FINAL`*: Para filtrar solo los casos confirmados recuperados (=3).* ▪ `NEUMONIA`*: Para filtrar solo los casos con neumonía (=1).*
Responsable: *Daniel Galicia Cobaxin*
Comentarios: ▪ *Se obtiene* `entidad` *de la tabla* `cat_entidades`*, nombrando esta columna como* `Estado`*.* ▪ *Se hace el conteo con* `COUNT` *de los casos donde solo sean recuperados, es decir, el caso 3.* ▪ *Se hace el conteo agregando la condición obligatoria de que el elemento* `Neumonia` *sea igual a 1.* ▪ *A la columna del conteo se le nombró* `Total_Casos_Recuperados_Neumonia`*.* ▪ *Finalmente, se agrupan por entidad y se ordenan de manera descendente para tener hasta arriba los estados con más casos según la columna del conteo.*

Results | Messages

	Estado	Total_Casos_Recuperados_Neumonia
1	MÉXICO	70851
2	CIUDAD DE MÉXICO	68097
3	PUEBLA	30081
4	VERACRUZ DE IGNACIO DE LA LLAVE	21682
5	GUANAJUATO	21075
6	JALISCO	19226
7	BAJA CALIFORNIA	18454
8	NUEVO LEÓN	17991
9	HIDALGO	16214
10	MICHOACÁN DE OCAMPO	14290
11	SONORA	14092
12	SINALOA	13065
13	CHIHUAHUA	12201
14	GUERRERO	11969
15	SAN LUIS POTOSÍ	11577
16	QUERÉTARO	11450
17	TABASCO	10260
18	TAMAULIPAS	9324
19	MORELOS	8993
20	COAHUILA DE ZARAGOZA	8696
21	OAXACA	7928
22	QUINTANA ROO	7062
23	YUCATÁN	7048

Query executed successfully. LAPTOP-6UJQPOU4\SQLEXPRESS ... LAPTOP-6UJQPOU4\dgali ... covidHistorico 00:00:13 32 rows

Figura 5: Resultado de la consulta 5

6. Consulta N° 6

Listar el total de casos confirmados/sospechosos por estado en cada uno de los años registrados en la base de datos.

Información
Número de consulta: *#6 Listar el total de casos confirmados/sospechosos por estado en cada uno de los años registrados en la base de datos.*
Requisitos: ▪ *Tabla* `cat_entidades` *y* `datoscovid`*.* ▪ `FECHA_INGRESO`*: Para extraer el año de la fecha de ingreso del paciente.* ▪ `entidad`*: Para extraer el nombre de la entidad residente dentro de la tabla* `cat_entidades`*.* ▪ `CLASIFICACION_FINAL`*: Para solo extraer los casos confirmados y sospechosos.*
Significado de catálogos: ▪ `CLASIFICACION_FINAL`*: Para contar solo casos confirmados y sospechosos (1, 2, 3, 6).*
Responsable: *Daniel Galicia Cobaxin*
Comentarios: ▪ *Se obtiene el año de la fecha de ingreso, la cual se nombra como* `Anio`*, y la entidad dentro de la tabla* `cat_entidades`*, la cual se nombra* `Estado`*.* ▪ *Se hace el conteo por estado en donde* `CLASIFICACION_FINAL` *haga coincidencia con 1, 2, 3 y 6, esto en 2020, 2021 y 2022.* ▪ *Se hace un* `JOIN` *entre las tablas y las claves de entidades para obtener el nombre de la entidad de residencia.* ▪ *Se agrupa y ordena por año, entidad y total de casos de manera descendente.*

Results | Messages

	Anio	Estado	Total_Casos
1	2020	CIUDAD DE MÉXICO	392919
2	2020	MÉXICO	200220
3	2020	NUEVO LEÓN	96075
4	2020	GUANAJUATO	95596
5	2020	JALISCO	67560
6	2020	BAJA CALIFORNIA	65984
7	2020	SONORA	61461
8	2020	COAHUILA DE ZARAGOZA	60815
9	2020	PUEBLA	58389
10	2020	VERACRUZ DE IGNACIO DE LA LLAVE	51784
11	2020	TABASCO	50534
12	2020	SAN LUIS POTOSÍ	44670
13	2020	TAMAULIPAS	43724
14	2020	CHIHUAHUA	42647
15	2020	CHIAPAS	37681
16	2020	QUERÉTARO	37013
17	2020	MICHOACÁN DE OCAMPO	35664
18	2020	SINALOA	33518
19	2020	YUCATÁN	33081
20	2020	OAXACA	32016
21	2020	HIDALGO	29656
22	2020	DURANGO	28531
23	2020	GUERRERO	28380

Query executed successfully. LAPTOP-6UJQPOU4\SQLEXPRESS ... LAPTOP-6UJQPOU4\dgali ... covidHistorico 00:00:22 96 rows

Figura 6: Resultado de la consulta 6

7. Consulta N° 7

Para el año 2020 y 2021 cuál fue el mes con más casos registrados, confirmados, sospechosos, por estado registrado en la base de datos.

Información
Número de consulta: *#7 Para el año 2020 y 2021 cuál fue el mes con más casos registrados, confirmados, sospechosos, por estado registrado en la base de datos.*
Requisitos: ▪ *Tabla* `cat_entidades` *y* `datoscovid`*.* ▪ `FECHA_INGRESO`*: Tanto* `Anio` *como* `Mes` *para las operaciones.* ▪ `CLASIFICACION_FINAL`*: Para el filtrado de confirmados y sospechosos.* ▪ `entidad`*: Para encontrar el nombre del estado en* `cat_entidades` *según las que coincida con* `datoscovid`*.*
Significado de catálogos: ▪ `CLASIFICACION_FINAL`*: Para distinguir entre casos confirmados (1, 2, 3) y sospechosos (6).*
Responsable: *Daniel Galicia Cobaxin*
Descripción de la consulta: ▪ *Se obtiene el mes y año, así como la entidad para la columna que se nombró como* `Estado`*.* ▪ *Se inicia el conteo, que se nombrará como* `Total_Casos`*.* ▪ *Se suman los casos cuando son confirmados y sospechosos, nombrándolo* `Total_Casos`*.* ▪ *Se hace la unión de las tablas para encontrar el nombre exacto de la entidad residente.* ▪ *Se filtra para solo los años 2020 y 2021.* ▪ *Se agrupa según el año, mes y por la entidad.* ▪ *Hasta este punto se tiene el conteo total por mes, estado y año (2020 y 2021).* ▪ *En la siguiente ventana, se tiene el* `Ranking mensual`*, el cual se ocupa por medio de* `Rank` *para asignar una 'posición' a cada grupo dividido con ayuda del comando* `PARTITION`*, sobre el año y estado, los cuales están ordenados por total de casos de manera descendente, llamado* `ranking`*.* ▪ *Finalmente, se extrae el ranking 1 de cada grupo.*
Comentarios: ▪ `RANK()`*: Asigna un número de ranking a cada fila dentro de un grupo de filas.* ▪ `PARTITION BY()`*: Divide los datos en grupos (o particiones) basados en una o más columnas.*

Results | Messages

	Año	Estado	Mes	Total_Casos
1	2020	AGUASCALIENTES	11	5594
2	2020	BAJA CALIFORNIA	12	12772
3	2020	BAJA CALIFORNIA SUR	12	3417
4	2020	CAMPECHE	7	3571
5	2020	CHIAPAS	6	13189
6	2020	CHIHUAHUA	10	14556
7	2020	CIUDAD DE MÉXICO	12	135939
8	2020	COAHUILA DE ZARAGOZA	11	10248
9	2020	COLIMA	8	2007
10	2020	DURANGO	11	7511
11	2020	GUANAJUATO	12	21713
12	2020	GUERRERO	7	5361
13	2020	HIDALGO	12	6315
14	2020	JALISCO	12	11875
15	2020	MÉXICO	12	43140
16	2020	MICHOACÁN DE OCAMPO	8	5767
17	2020	MORELOS	12	2882
18	2020	NAYARIT	7	2522
19	2020	NUEVO LEÓN	12	15867
20	2020	OAXACA	7	5504
21	2020	PUEBLA	7	11510
22	2020	QUERÉTARO	12	11143
23	2020	QUINTANA ROO	7	4833

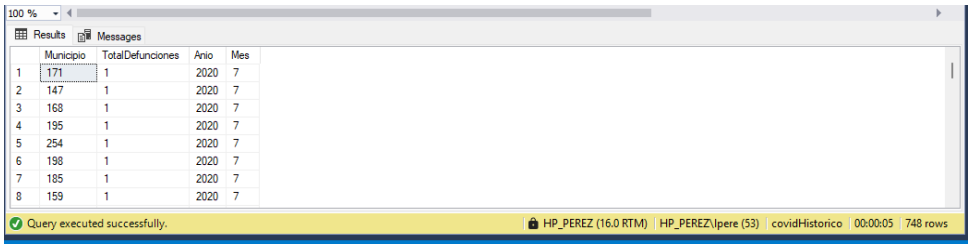
Query executed successfully. LAPTOP-6UJQPOU4\SQLEXPRESS ... LAPTOP-6UJQPOU4\dgali ... covidHistorico 00:01:57 64 rows

Figura 7: Resultado de la consulta 7

8. Consulta N° 8

Listar el municipio con menos defunciones en el mes con más casos confirmados con neumonía en los años 2020 y 2021

Información
Número de consulta: 8
Requisitos: N/A
Significado de catálogos: Neumonia = 1, hace referencia al catalogo Çátalogo SI_NO", donde 1 significa "Si". Clasificacion_Final (1,2,3) hace referencia al catalogo Çat CLASIFICACION_FINAL_COVID"donde 1,2,3 son aquellos casos confirmados
Responsable: Pérez López Leonardo
Comentarios: Uso ÇOUNT(*)"para contar casos de COVID con nuemonía, después, ORDER BY COUNT(*) DESC selecciona el mes y año con más casos y uso subconsultas para extraer el año y el mes en los que hubo más casos. Bsuco solo los casos donde Fecha_Def IS NOT NULL (fallecidos). En la subconulta uso YEAR(TRY_CONVERT(DATE, NULLIF(Fecha_Def, ''))) = (subconsulta) para el año con más neumonía y MONTH(TRY_CONVERT(DATE, NULLIF(Fecha_Def, ''))) = (subconsulta) para el mes con más neumonía. Agrupo por ENTIDAD_RES y ordeno por menor cantidad de defunciones y uso: COUNT(*) AS TotalDefunciones cuenta los fallecidos en cada entidad, ORDER BY TotalDefunciones ASC para obtener el municipio con menos defunciones y finalmente, SELECT TOP 1 devuelve solo el estado con menos fallecidos en ese mes y año.



100 %

Results | Messages

	Municipio	TotalDefunciones	Anio	Mes
1	171	1	2020	7
2	147	1	2020	7
3	168	1	2020	7
4	195	1	2020	7
5	254	1	2020	7
6	198	1	2020	7
7	185	1	2020	7
8	159	1	2020	7

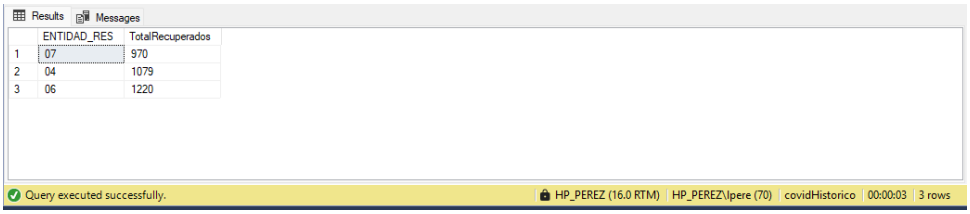
Query executed successfully. | HP_PEREZ (16.0 RTM) | HP_PEREZ\lpere (53) | covidHistorico | 00:00:05 | 748 rows

Figura 8: Resultado de la consulta 8

9. Consulta N°9

Listar el top 3 de municipios con menos casos recuperados en el año 2021

Información
Número de consulta: 9
Requisitos: N/A
Significado de catálogos: Clasificacion_Final (1,2,3) hace referencia al catalogo Çat CLASIFICACION_FINAL_COVID"donde 1,2,3 son aquellos casos confirmados
Responsable: Pérez López Leonardo
Comentarios: Aqui use COUNT, que nos ayuda a contar el numero de filas que coinciden con una condición específica, en este caso lo utilizamos para el total de recuperados por Entidad. Como se explico en la consulta anterior "FECHA_DEF.^{es} de tipo nvarchar, por lo cual tenemos que hacer una conversión del dato y para eso nos apoyamos de la funcion TRY_CONVERT, posterior a esto, agrupamos por Entidad y ordenamos de manera ascendente



Results | Messages

	ENTIDAD_RES	TotalRecuperados
1	07	970
2	04	1079
3	06	1220

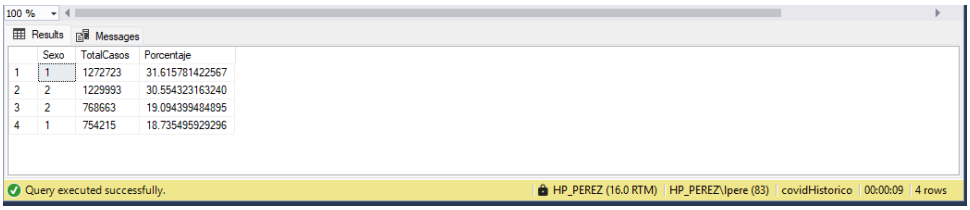
Query executed successfully. | HP_PEREZ (16.0 RTM) | HP_PEREZ\lpere (70) | covidHistorico | 00:00:03 | 3 rows

Figura 9: Resultado de la consulta 9

10. Consulta N° 10

Listar el porcentaje de casos confirmados por género en los años 2020 y 2021.

Información
Número de consulta: 10
Requisitos: N/A
Significado de catálogos: Clasificacion_Final (1,2,3) hace referencia al catálogo Çat CLASIFICACION_FINAL_COVID"donde 1,2,3 son casos confirmados. Sexo (1 y 2): 1. Mujer, 2. Hombre.
Responsable: Pérez López Leonardo
Comentarios: Solo consideré registros donde Clasificacion_Final_Covid indique casos confirmados (1, 2, 3), después, solo incluí los años 2020 y 2021, asegurando que Fecha_Def sea una fecha válida con TRY_CONVERT(), agrupé por la columna Sexo para contar los casos por cada género, calculé el porcentaje con COUNT(*) * 100.0 / SUM(COUNT(*)) OVER () que nos da el porcentaje de cada género respecto al total y finalmente, se ordena por Porcentaje DESC para mostrar primero el género con más casos.



100 % | Results | Messages

	Sexo	TotalCasos	Porcentaje
1	1	1272723	31.615781422567
2	2	1229993	30.554323163240
3	2	768663	19.094399484895
4	1	754215	18.735495929296

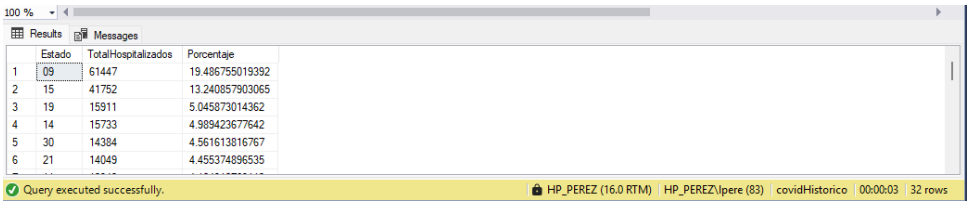
Query executed successfully. | HP_PEREZ (16.0 RTM) | HP_PEREZ\lpere (83) | covidHistorico | 00:00:09 | 4 rows

Figura 10: Resultado de la consulta 10

11. Consulta N° 11

Listar el porcentaje de casos hospitalizados por estado en el año 2020.

Información
Número de consulta: 11
Requisitos: N/A
Significado de catálogos: Clasificacion_Final (1,2,3) hace referencia al catálogo Çat CLASIFICACION_FINAL_COVID"donde 1,2,3 son casos confirmados. Del catálogo de ENTIDAD_UM obtenemos el número de la entidad.
Responsable: Pérez López Leonardo
Comentarios: En entidad_UM almacena la entidad federativa donde fue atendido el paciente, posteriormente, solo consideré registros donde Clasificacion_Final_Covid indique casos confirmados (1, 2, 3), hice un filtrado por año (2020) usando YEAR(TRY_CONVERT(DATE, NULLIF(Fecha_Def, ”))) = 2020, que convierte Fecha_Def a DATE, evitando errores con valores inválidos o vacíos, después, hice un filtrado por hospitalización con Tipo_Paciente = 2 (según el catálogo ”TIPO_PACIENTE”, 1 = Ambulatorio, 2 = Hospitalizado), luego, calculé el porcentaje de casos hospitalizados por estado respecto al total de hospitalizados con (COUNT(*) * 100.0) / SUM(COUNT(*)) OVER () y finalmente ordené por porcentaje descendente con Porcentaje DESC para mostrar los estados con más hospitalizaciones primero.



100 %

Results | Messages

	Estado	TotalHospitalizados	Porcentaje
1	09	61447	19.486755019392
2	15	41752	13.240857903065
3	19	15911	5.045873014362
4	14	15733	4.989423677642
5	30	14384	4.561613816767
6	21	14049	4.455374896535

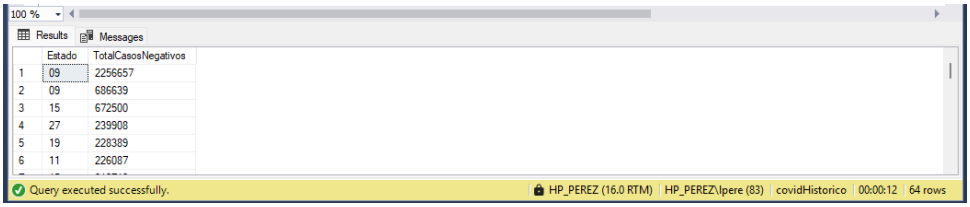
Query executed successfully. | HP_PEREZ (16.0 RTM) | HP_PEREZ\lpere (83) | covidHistorico | 00:00:03 | 32 rows

Figura 11: Resultado de la consulta 11

12. Consulta N° 12

Listar total de casos negativos por estado en los años 2020 y 2021.

Información
Número de consulta: 12
Requisitos: N/A
Significado de catálogos: Clasificacion_Final (1,2,3) hace referencia al catálogo Çat CLASIFICACION_FINAL_COVID"donde 1,2,3 son casos confirmados. Del catálogo de ENTIDAD_UM obtenemos el número de la entidad.
Responsable: Pérez López Leonardo
Comentarios: Entidad_UM AS Estado representa la entidad federativa donde se atendió el paciente, luego hice un filtrado de casos negativos con Clasificacion_Final_Covid NOT IN (1, 2, 3) donde se excluyen los casos confirmados (1, 2, 3), ya que los negativos tienen otros valores en la base de datos, luego, hice un filtrado por años (2020 y 2021) usando YEAR(TRY_CONVERT(DATE, NULLIF(Fecha_Def, ”))) IN (2020, 2021), posterior a esto, hice un agrupamiento por ENTIDAD_UM y finalmente ocupé ORDER BY TotalCasosNegativos DESC para ordenar el TotalCasosNegativos por orden descendente.



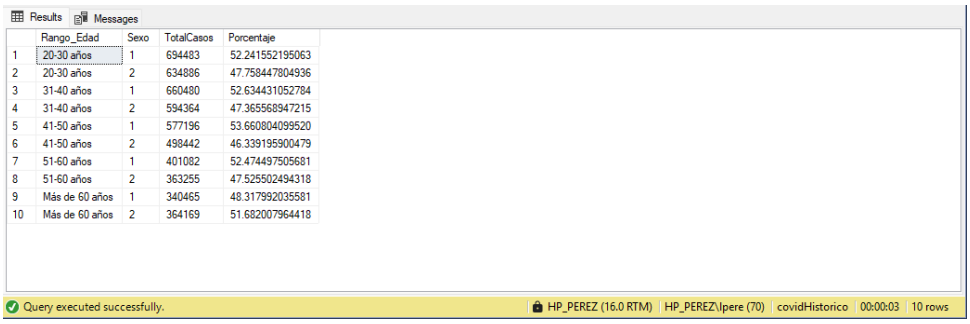
	Estado	TotalCasosNegativos
1	09	2256657
2	09	686639
3	15	672500
4	27	239908
5	19	228389
6	11	226087

Figura 12: Resultado de la consulta 12

13. Consulta N° 13

Listar el porcentaje 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.

Información
Número de consulta: 13
Requisitos: N/A
Significado de catálogos: Clasificacion_Final (1,2,3) hace referencia al catálogo Çat CLASIFICACION_FINAL_COVID"donde 1,2,3 son casos confirmados. En el caso del sexo 1 y 2 pertenecen a: 1. Mujeres, 2. Hombres.
Responsable: Pérez López Leonardo
Comentarios: En esta consulta me apoyé de varias instrucciones, la que más resalta es la de ÇASE-WHEN", que nos sirve para comparar una expresión con un conjunto de expresiones sencillas para determinar el resultado. Como se puede ver en la consulta, el ÇASE-WHEN"nos ayudó para poder comparar entre todos los rangos de edades solicitados en la problemática, agrupando entre sus rangos de edades y el sexo. Posterior a esto, hicimos uso de un ÇOUNT"para contar el número de filas con las que coinciden, después calculamos los promedios y usamos "GROUP BY"para agrupar y ordenamos por RANGO_EDADƨ "SEXO".



	Rango_Edad	Sexo	TotalCasos	Porcentaje
1	20-30 años	1	694483	52.241552195063
2	20-30 años	2	634886	47.758447804936
3	31-40 años	1	660480	52.634431052784
4	31-40 años	2	594364	47.365568947215
5	41-50 años	1	577196	53.660804099520
6	41-50 años	2	498442	46.339195900479
7	51-60 años	1	401082	52.474497505681
8	51-60 años	2	363255	47.525502494318
9	Más de 60 años	1	340465	48.317992035581
10	Más de 60 años	2	364169	51.682007964418

Figura 13: Resultado de la consulta 13

14. Consulta N° 14

Listar el rango de edad con más casos confirmados y que fallecieron en los años 2020 y 2021.

Información
Número de consulta: 14
Requisitos: N/A
Significado de catálogos: Clasificacion_Final (1,2,3) hace referencia al catálogo Çat CLASIFICACION_FINAL_COVID"donde 1,2,3 son casos confirmados.
Responsable: Pérez López Leonardo
Comentarios: Dentro de la subconsulta, clasifiqué las edades en los rangos que necesitamos (así como en la consulta anterior), con FECHA_DEF IS NOT NULL se filtran solo casos de covid confirmados y fallecidos y me aseguré que la fecha sea de 2020 y 2021 convirtiendo el tipo de dato con "TRY_CONVERT". Agrupé los datos por el RANGO_EDAD, ordeno el TOTAL_FALLECIDOS en orden descendente y filtro el top 1.

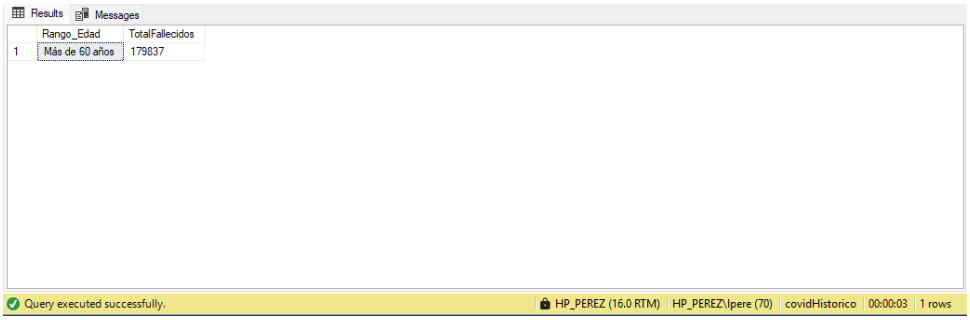


Figura 14: Resultado de la consulta 14

15. Conclusiones

Esta prática nos ayudó para reforzar o bien, recordar algunos conceptos clave que se utilizan en Bases de Datos. Si bien, algunas consultas fueron un poco "laboriosas"(Ejmeplo: Consulta 8) estas se lograron completar al 100Para poder conseguir cada resultado se siguió una estrategia, la cual no fue crear vistas (views), asociar o mas bien, crear tablas, o algo por el estilo. Para poder llegar a la solución de cada consulta necesitamos analizar como estaban estructurados los datos, ¿A qué nos referimos con esto?, como se sabe, esta base de datos de çovidHistoricoçuenta con sus propios catalogos y descriptores, lo cual nos ayudó bastante para poder abordar cada reto de esta prática. Principalmente nos enfocamos en hacer un "filtrado"de las columnas que ibamos a utilizar (CLASIFICACION_FINAL, FECHA_DEF, FECHA_INGRESO, etc), resaltando estas como las más importantes para obtener la información correcta. Nos enfrentamos a la problemática que, la base de datos no contaba con PK o FK, lo cual hacia el proceso mucho más lento. Para haber podido solucionar esa problemática, bien pudimos haber creado tablas y establecer PK y FK como se mencionó al principio pero, esto nos llevaría mas tiempo y probablemente no hubiésemos podido completar todos los retos propuestos, por lo cual, vuelvo a enfatizar que nosotros nos guiamos mucho del diccionario de datos proporcionado por el profesor.

Al principio de esta conclusión mencionamos que esta práctica nos ayudó a refozar conceptos sobre bases de datos pero, ¿Qué conceptos? o, ¿Qué fue lo nuevo que aprendimos?, bien, en cursos pasados de BD (bases de datos) utilizamos algunos conceptos como: COUNT(), GROUP BY, ORDER BY, HAVING, etc, pero como se dice, "si dejas de practicar las cosas, las llegas a olvidar", entonces, algunos de estos conceptos o funciones básicas y clave en las DB ya no estaban tan "frescas".

A parte de recordar, también aprendimos, y algunas de las funciones nuevas aprendidas fueron: YEAR(TRY_CONVERT(DATE, NULLIF(","))) que principalmente nos ayuda a la conversión de datos, OVER() que determina la creación de particiones y el orden de un conjunto de filas antes de aplicar la función de ventana asociada.

Nuestra estrategia no fue la mejor ni la peor, a nosotros nos funcionó y pudimos completar las consultas, quizá, para ejercicios posteriores tengamos que implementar algo diferente, pero la llección fue que, cada equipo tiene su estrategia y dentro de cada equipo, cada integrante tiene su forma de ver, analizar y resolver una problemática planteada. Este ejericicio nos ayudó a poder reforzar el trabajo en equipo y sobre todo la organización.