

INSTITUTO POLITÉCNICO NACIONAL

UNIDAD PROFESIONAL INTERDISCIPLINARIA  
EN INGENIERIA Y TECNOLOGÍAS AVANZADAS

BASES DE DATOS DISTRIBUIDAS

GRUPO: 3TM3

*PRÁCTICA 3*

PROFESOR: DE LA CRUZ SOSA CARLOS

INTEGRANTES EQUIPO 2:

OCAMPO CABRERA ASael  
ROBERT ROA KARLA GUADALUPE

2022640070  
2022640710

### Práctica 3. Servidores vinculados y particionamiento

#### Distribución de los 3 Nodos:

NODO SQL KARLA	NODO SQL ASael	NODO MYSQL ASA
Región Occidente	Región Oriente	Región Sureste
Región Noroeste	Región Centro Sur	Región Suroeste
Región Noreste	Región Centro Norte	-----

#### Distribución de las Regiones:

Región	#	Estados
Noroeste	(02, 03, 26, 25, 08,10)	Baja California, Baja California Sur, Sonora, Sinaloa, Chihuahua, Durango
Noreste	(05, 19, 28)	Coahuila, Nuevo León, Tamaulipas
Centro Norte	(32, 01, 11, 22, 24)	Zacatecas, Aguascalientes, Guanajuato, Querétaro, San Luis Potosí
Centro Sur	(09, 15, 17)	CDMX, Estado de México, Morelos
Occidente	(18, 14, 06, 16)	Nayarit, Jalisco, Colima, Michoacán
Oriente	(30, 21, 29,13)	Veracruz, Puebla, Tlaxcala, Hidalgo
Suroeste	(12, 20,07)	Guerrero, Oaxaca, Chiapas
Sureste	(04, 23, 31, 27)	Campeche, Quintana Roo, Yucatán, Tabasco

Para realizar la fragmentación de la base de datos covidHistorico en 8 regiones, tomamos como referencia la clasificación de regiones de México publicada en Wikipedia, donde se agrupan los estados de manera político-económica. Esta división nos permitió organizar los fragmentos de forma lógica y balanceada.

La información fue extraída de:

Colaboradores de Wikipedia. (2025, 12 marzo). Regiones de México. Wikipedia, la enciclopedia libre. [https://es.wikipedia.org/wiki/Regiones\\_de\\_M%C3%A9xico](https://es.wikipedia.org/wiki/Regiones_de_M%C3%A9xico)

Gracias a esta organización, pudimos repartir la carga de trabajo entre los integrantes del equipo, de modo que cada uno se encargara de al menos tres consultas distribuidas en los distintos nodos, lo que facilitó la colaboración, redujo errores al momento de probar las consultas, y nos ayudó a entender mejor el funcionamiento de los sistemas distribuidos en la práctica.

#### Listado de Linked Server:

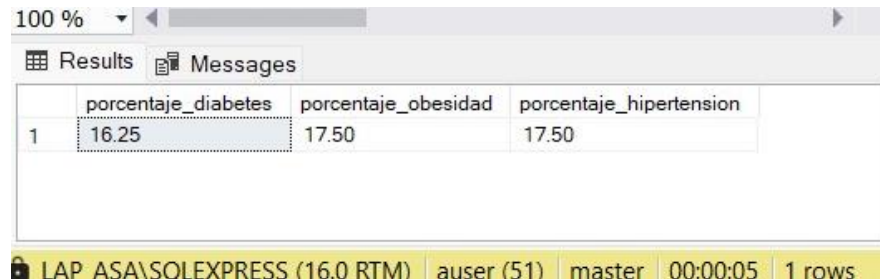
Laptop Karla	Laptop Asael
LS_KARLA_VPN	LS_SQLKARLA_VPN
LS_ASA_VPN	MYLS_VPN
MYSQLASAP3	MYSQL_ASA_P3

Cada integrante configuró tres linked servers desde su propia laptop usando VPN: uno apuntando a su propio servidor SQL local, otro hacia la laptop del compañero y un tercero hacia el nodo MySQL. Esta configuración nos permitió ejecutar las consultas distribuidas desde cualquier equipo, accediendo a los fragmentos de la base de datos sin importar la ubicación.

## Desarrollo:

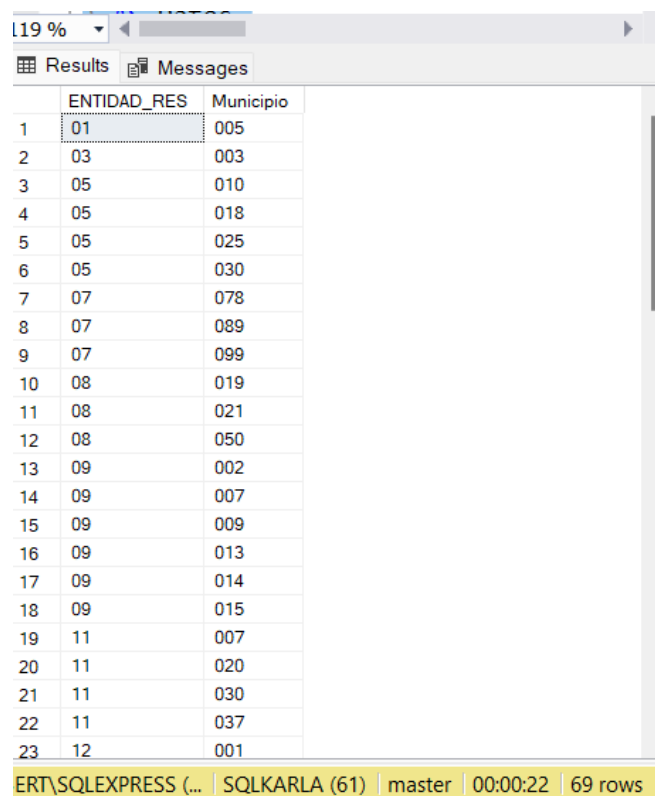
Proponer una fragmentación por regiones (recuperar una clasificación de regiones políticas y/o económicas del país) y alojar cada fragmento en una base de datos distinta. Interconectar los nodos y modificar las consultas 3,4,5 y 7 de la práctica 1 para que sean consultas distribuidas.

**Consulta número 3.** Listar el porcentaje de casos confirmados en cada una de las siguientes morbilidades a nivel nacional: diabetes, obesidad e hipertensión.



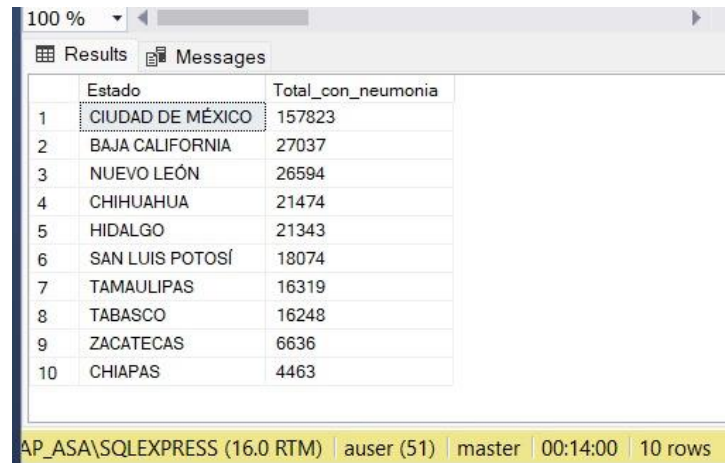
	porcentaje_diabetes	porcentaje_obesidad	porcentaje_hipertension
1	16.25	17.50	17.50

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



	ENTIDAD_RES	Municipio
1	01	005
2	03	003
3	05	010
4	05	018
5	05	025
6	05	030
7	07	078
8	07	089
9	07	099
10	08	019
11	08	021
12	08	050
13	09	002
14	09	007
15	09	009
16	09	013
17	09	014
18	09	015
19	11	007
20	11	020
21	11	030
22	11	037
23	12	001

**Número 5.** Listar los estados con más casos recuperados con neumonía.



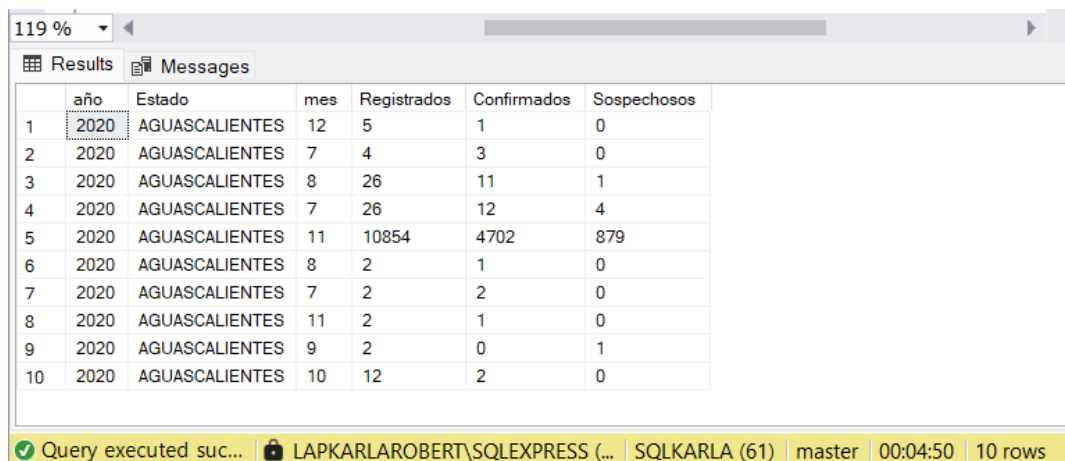
100 %

Results Messages

	Estado	Total_con_neumonia
1	CIUDAD DE MÉXICO	157823
2	BAJA CALIFORNIA	27037
3	NUEVO LEÓN	26594
4	CHIHUAHUA	21474
5	HIDALGO	21343
6	SAN LUIS POTOSÍ	18074
7	TAMAULIPAS	16319
8	TABASCO	16248
9	ZACATECAS	6636
10	CHIAPAS	4463

AP\_ASA\SQLEXPRESS (16.0 RTM) | auser (51) | master | 00:14:00 | 10 rows

**Número 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.



119 %

Results Messages

	año	Estado	mes	Registrados	Confirmados	Sospechosos
1	2020	AGUASCALIENTES	12	5	1	0
2	2020	AGUASCALIENTES	7	4	3	0
3	2020	AGUASCALIENTES	8	26	11	1
4	2020	AGUASCALIENTES	7	26	12	4
5	2020	AGUASCALIENTES	11	10854	4702	879
6	2020	AGUASCALIENTES	8	2	1	0
7	2020	AGUASCALIENTES	7	2	2	0
8	2020	AGUASCALIENTES	11	2	1	0
9	2020	AGUASCALIENTES	9	2	0	1
10	2020	AGUASCALIENTES	10	12	2	0

Query executed suc... | LAPKARLAROBERT\SQLEXPRESS (... | SQLKARLA (61) | master | 00:04:50 | 10 rows

## CONCLUSIONES

Esta práctica nos permitió aplicar de manera práctica los conceptos vistos en clase, tales como la fragmentación horizontal, la configuración de servidores vinculados (linked server) y la ejecución de consultas distribuidas.

Partimos de la base de datos covidHistorico, la cual fragmentamos en 8 regiones de México. Estas regiones se distribuyeron entre tres nodos, de la siguiente forma:

- **Nodo SQL Karla:** Región Occidente, Noroeste y Noreste
- **Nodo SQL Asael:** Región Oriente, Centro Sur y Centro Norte
- **Nodo MySQL ASA:** Región Sureste y Suroeste

La conexión entre los nodos se realizó mediante la creación de linked servers, permitiendo así que desde cualquier servidor SQL Server se pudiera consultar datos remotos. Para esto utilizamos instrucciones del tipo OPENQUERY.

Durante el desarrollo, enfrentamos diversos retos relacionados con la conectividad. En particular, la red escolar Wi-Fi del IPN presentó un rendimiento muy bajo, incluso al usar VPN, con tiempos de respuesta que superaban los 20 minutos en consultas con grandes volúmenes de datos. Esto hacía impráctica la ejecución y prueba de las consultas distribuidas. Como solución, optamos por utilizar una conexión de red doméstica con VPN, lo cual redujo significativamente los tiempos: las consultas más pesadas tardaban máximo 14 minutos, mientras que las de menor carga respondían en menos de 5 minutos, e incluso en segundos. Sin embargo, también pudimos notar que el hecho de ocupar una VPN como solución a las restricciones de conectividad de red publica en ciertos proveedores de internet, así como, la red publica de UPIITA, seguía siendo una solución no tan optima, ya que seguía un sistema de consulta en el que tenia que primero direccionar hacia la VPN para así poder pasar a los servidores vinculados, proceso que no era del todo practico. Por lo cual, consideramos que soluciones como utilizar un celular solo como Access Point podría solucionar el tema del rendimiento en cuestión al tiempo de ejecución de las consultas.

Las consultas 3, 4, 5 y 7 de la Práctica 1 fueron modificadas para ejecutarse en este entorno distribuido. Gracias a ello, logramos consultar datos segmentados por región de manera eficiente desde cualquier nodo SQL Server, demostrando la efectividad del modelo de fragmentación implementado.

En resumen, esta práctica nos permitió entender y poner en práctica cómo se estructura una infraestructura de bases de datos distribuidas en entornos reales, cómo se enlazan nodos (SQL y MySQL), y cómo herramientas como VPN pueden ser críticas para garantizar el rendimiento de las consultas en escenarios donde se maneja información distribuida a gran escala.