

# ÍNDICE

1. Caso de negocio.....	2
1.1. Instrucciones para resolver el caso de evaluación .....	4
1.2. Cuestiones de evaluación.....	6

# 1. CASO DE NEGOCIO

Una vez se ha construido y preparado completamente la base de datos “Rastreo\_COVID” durante el ejercicio de repaso, se procede a resolver las cuestiones que se plantean en el caso de evaluación.

En primer lugar, se recuerdan los nodos y relaciones que forman la BBDD:

## ❖ **DATASET**

Está dividido en tres CSV ubicados en el directorio siguiente de la máquina virtual:  
**/home/imfbigdata/neo4j**

### 1. Nodo persona. Propiedades:

- Id de la persona (id\_persona).
- Nombre de la persona (nombre\_persona).
- Estado de la persona (estado\_salud: contagiado/sano).
- *Timestamp* que representa el momento en que se notifica a una persona que se había realizado el test PCR, si está o no contagiada (hora\_test\_resultado).
- Latitud: coordenada de latitud del domicilio donde reside la persona (latitud\_domicilio).
- Longitud: coordenada de longitud del domicilio donde reside la persona (longitud\_domicilio). Corresponden a coordenadas de Valladolid y Salamanca.

### 2. Nodo Ubicación. Propiedades:

- Id del lugar (id\_ubicacion).
- Nombre del establecimiento (nombre\_establecimiento).
- Tipo de establecimiento (tipo\_establecimiento).

### 3. Nodo visita (de personas a distintos lugares (hospital, colegio, bar, restaurante, etc.) de dos ciudades: Valladolid y Salamanca). Propiedades:

- Identificador de la visita (id\_visita).
- Identificador de la persona (id\_ubicacion).
- Hora de comienzo de la visita (inicio\_visita).
- Hora de fin de la visita (fin\_visita).

Hay tres tipos de relaciones entre los nodos:

- ❖ (Persona) [VISITA\_EMPLAZAMIENTO] (Ubicacion)
- ❖ (Persona) [REALIZA\_VISITA] (Visita) [A\_ESTABLECIMIENTO] (Ubicacion)

## 1.1. INSTRUCCIONES PARA RESOLVER EL CASO DE EVALUACIÓN

Para cada cuestión, será necesario copiar el comando que se ha ejecutado en el editor de Neo4J para responder lo que se pide en cada cuestión. Además, se deberá adjuntar una captura de la tabla (formato *table* o *text*) o grafo resultante con las columnas que se piden.

Habrà alguna cuestión en la que además de adjuntar las capturas, se pide responder a alguna pregunta concreta analizando los datos obtenidos como resultado de la consulta.

Un ejemplo podría ser el siguiente:

```
MATCH (p:Persona {estado:"Contagiado"}) -  
[v:VISITA_EMPLAZAMIENTO] -> (u:Ubicacion) - [:PARTE_DE] -  
> (c:Ciudad)  
  
RETURN p.nombre AS Persona_contagiada, u.nombre AS  
Establecimiento, c.nombre as Ciudad
```

- Tabla con *query* y tabla resultante en formato *table*:

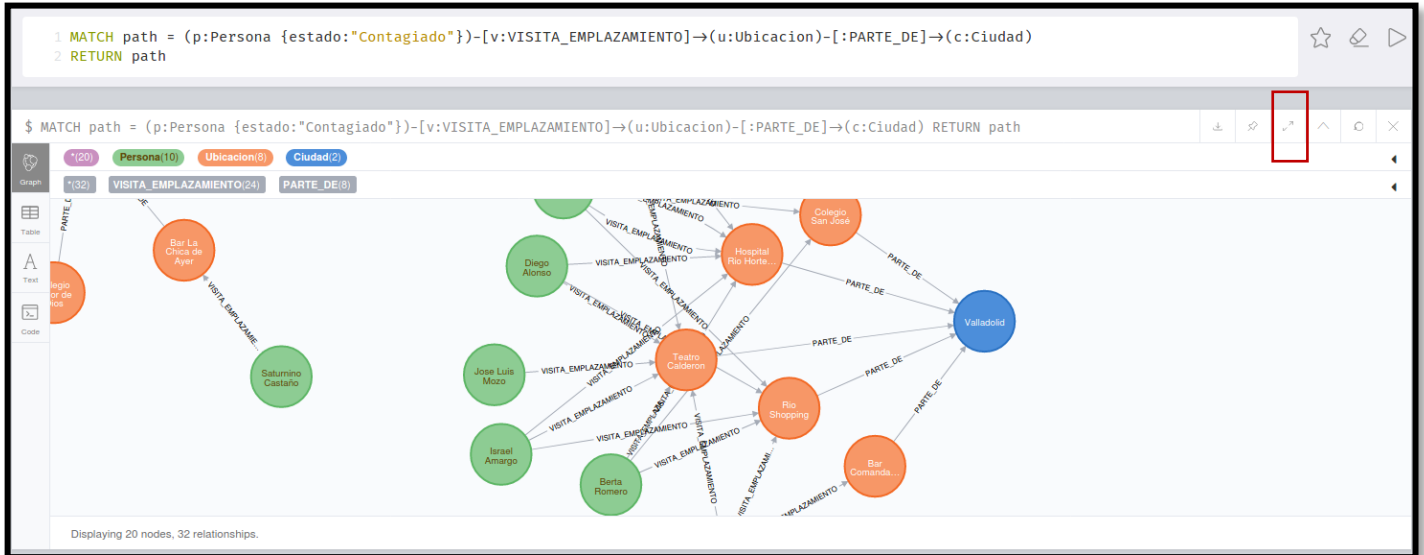
<pre>1 MATCH (p:Persona {estado:"Contagiado"})-[v:VISITA_EMPLAZAMIENTO]→(u:Ubicacion)-[:PARTE_DE]→(c:Ciudad) 2 RETURN p.nombre AS Persona_contagiada, u.nombre AS Establecimiento, c.nombre as Ciudad</pre>			
\$ MATCH (p:Persona {estado:"Contagiado"})-[v:VISITA_EMPLAZAMIENTO]→(u:Ubicacion)-[:PARTE_DE]→(c:Ciudad) RETURN p.nombre AS Persona_contagiada,			
<div>Table</div>	Persona_contagiada	Establecimiento	Ciudad
<div>Text</div>	"Marcelino Rodriguez"	"Bar Comandante"	"Valladolid"
<div>Code</div>	"Diego Alonso"	"Teatro Calderon"	"Valladolid"
	"Marcelino Rodriguez"	"Teatro Calderon"	"Valladolid"
	"Fidel Figueroa"	"Teatro Calderon"	"Valladolid"
	"Jose Luis Mozo"	"Teatro Calderon"	"Valladolid"
	"Israel Amargo"	"Teatro Calderon"	"Valladolid"
	"Berta Romero"	"Teatro Calderon"	"Valladolid"
	"Israel Amargo"	"Hospital Rio Hortega"	"Valladolid"
	"Berta Romero"	"Hospital Rio Hortega"	"Valladolid"
Started streaming 24 records after 1 ms and completed after 9 ms.			

- Si de esta misma consulta se quisiera obtener un grafo:

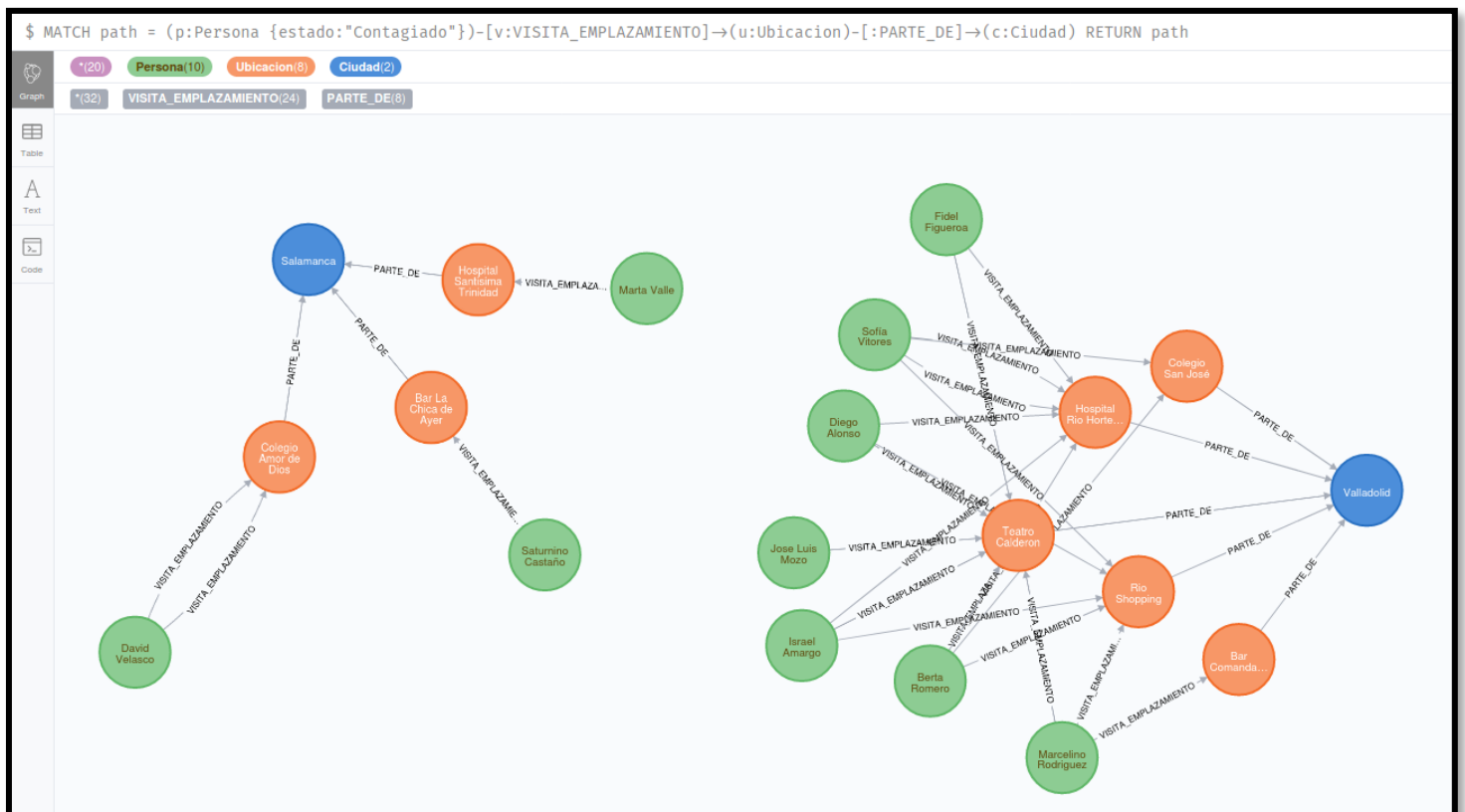
```

MATCH path = (p:Persona {estado:"Contagiado"})-
[v:VISITA_EMPLAZAMIENTO]->(u:Ubicacion)-[:PARTE_DE]-
>(c:Ciudad)
RETURN path

```



- En el caso de los grafos, conviene ampliar el grafo (recuadro marcado en rojo en la imagen anterior) para tomar la captura de pantalla:



# 1.2. CUESTIONES DE EVALUACIÓN

❖ **CASO 1:** identificar el número de personas contagiadas y de personas sanas en la muestra de 40 personas. Devolver el resultado en formato *table* o *text*. Los campos a devolver se muestran a continuación:

<div><div>Table</div></div>	Total_Contagiados	Total_Sanos
-----------------------------	-------------------	-------------

- ❖ **CASO 2:** Encontrar las personas sanas que han estado en contacto con una persona que ha dado positivo.

Debido a que nos encontramos en las primeras fases de estudio de la enfermedad, se da por supuesto que la COVID puede infectar a personas sanas que hayan estado en un mismo lugar en el que ha estado una persona contagiada, aunque sea en diferentes días.

Identificar todas las personas sanas que han estado en el mismo lugar (sin importar fecha ni hora) en el que también ha estado una persona contagiada (lógicamente el comienzo de la visita de la persona sana tiene que ser posterior al de la persona enferma, ya que es imposible contagiarse si la persona sana ha ido antes que la contagiada).

Devolver el resultado en formato *table* o *text*.


- ❖ **CASO 3:** mostrar grafo con las personas sanas que han coincidido con una persona contagiada, concretamente con “Marcelino Rodriguez”. Hay que mostrar en un grafo con este nodo persona, junto con todos los lugares que ha visitado, y con los nodos de etiqueta persona sana que han visitado también ese lugar posteriormente.

Observando los resultados del grafo, comentar seis personas de las que han estado en alguna ubicación en la que ha estado Marcelino después de haber estado él, tienen menos riesgo que el resto de haberse contagiado que el resto:



❖ **CASO 4:** construir la misma consulta anterior, pero mostrando el resultado como una tabla y no como un grafo ( formato *table* o *text*) mostrando los campos:

- Esparcidor\_virus.
- Comienzo\_esparcimiento\_virus.
- Establecimiento.
- Persona\_en\_riesgo.
- Inicio\_visita\_persona\_en\_riesgo.

 Table	Esparcidor_virus	Comienzo_esparcimiento_virus	Ubicacion	Inicio_visita_persona_en_riesgo	Persona_en_riesgo
---	------------------	------------------------------	-----------	---------------------------------	-------------------

❖ **CASO 5.1:** construir una tabla (formato *text* y *table*) que identifique para cada persona contagiada (columna uno), las personas sanas con las que ha coincidido en un establecimiento en el mismo tiempo. Construir como segunda columna un *array* de elementos JSON llamado “Contactos” con claves:

- Persona\_en\_contacto.
- Establecimiento.
- Fecha\_comienzo\_solapamiento.
- Fecha\_fin\_solapamiento.

Se muestra a continuación un ejemplo del formato de cada documento JSON y de los dos campos a obtener:

```
{
  "Fecha_fin_solapamiento":
    "2020-09-02T18:44:54Z",
  "Establecimiento": "Rio
Shopping",
  "Fecha_comienzo_solapamiento":
    "2020-09-02T17:18:16Z",
  "Persona_en_contacto": "Ramon
Cuadrado"
}
```

 Table	Persona_Contagiada	Contactos
---	--------------------	-----------

- ❖ **CASO 5.2:** una vez obtenida la consulta anterior, conseguir, añadiendo tres sentencias a esta, una tabla que tenga el nombre de la persona contagiada y otra columna con el número de personas sanas con las que ha tenido contacto, obteniendo ese número a partir de los elementos del *array* de elementos JSON. Ordenar los resultados por “Numero\_de\_contactos\_sanos” descendente:

Persona_Contagiada	Numero_de_contactos_sanos
--------------------	---------------------------

- ❖ **CASO 6:** encontrar a aquellas personas (si es que hay alguna) que visitaron un establecimiento incluso después de saber que habían dado positivo en el test. Los campos de salida a mostrar son:

Contagiado	Fecha_inicio_visita	Fecha_Confirmacion_Contagio
------------	---------------------	-----------------------------

- ❖ **CASO 7:** ahora que se han obtenido todas las personas sanas que coincidieron en algún establecimiento con alguna contagiada, se quiere averiguar el tiempo exacto (duración) que coincidió cada persona sana con la persona contagiada.

Expresar la duración en horas y redondeada a cuatro decimales. Devolver el resultado en formato *table* o *text*.

- ❖ **CASO 8:** una persona ha estado en dos sitios diferentes con personas contagiadas, en uno estuvo una hora y media en contacto y en el otro dos. El total de exposición de esa persona habrá sido de tres horas y media.

La duración de cada contacto entre persona sana y contagiada será el resultado que se obtenga en la cuestión anterior. Por tanto, se puede utilizar la consulta anterior como base y será necesario añadirle algo más para conseguir el resultado esperado.

Si una persona sana coincidió con dos contagiadas el mismo día en el mismo establecimiento, también se sumará el tiempo que estuvo en contacto con cada contagiado, entendiendo que el haber estado rodeado de más contagiados supone que esa persona tenga un mayor riesgo de contraer la enfermedad.


Solamente se mostrarán en la tabla (formato *table* o *text*) las cinco personas sanas con más tiempo de exposición. A esas cinco personas se les realizará inmediatamente una llamada para que comiencen a guardar cuarentena. El tiempo total se mostrará en horas con redondeo a cuatro decimales (por ejemplo: 9,4972 horas, que serán nueve horas y 30 minutos).

 Table	Persona_Sana	Tiempo_total_exposicion
---	--------------	-------------------------

- ❖ **CASO 9:** se pretende tratar de reducir la afluencia e implementar aún más medidas de precaución en aquellos establecimientos en los que hayan estado más tiempo personas contagiadas.

Se pide devolver una tabla que contenga cada establecimiento, que ha sido visitado por al menos una persona contagiada, el total de visitas de contagiados en cada establecimiento, el total de visitas en cada establecimiento, el porcentaje de visitas de contagiados respecto al total de visitas de cada establecimiento, y la ciudad a la que pertenece el establecimiento. Expresar el porcentaje redondeado a dos decimales.

Comentar cuáles son los dos establecimientos con mayor y los dos con menor porcentaje de visitas de contagiados respecto del total de cada establecimiento.

 Tabla	Establecimiento	Num_visitas_Contagiados	Num_total_visitas	Porcentaje_visitas_de_personas_contagiadas	Ciudad
---	-----------------	-------------------------	-------------------	--	--------

- ❖ **CASO 10:** mostrar las distancias entre los domicilios de los contagiados de Valladolid que hayan ido a un mismo establecimiento aunque haya sido en diferentes fechas.

Hay que obtener el resultado mediante una sola consulta en el editor de Neo4J. Se quiere devolver solo los tres registros **Persona1 - Persona2 CONTAGIADAS** de Valladolid que vivan a más distancia.

**Importante:** evitar obtener dos registros con la misma distancia y nodos intercambiados como este:

Persona_Sana_1	Persona_Sana_2	Ciudad	Distancia_domicilios_km
"Berta Romero"	"Sofia Vitores"	"Valladolid"	
"Sofia Vitores"	"Berta Romero"	"Valladolid"	

**Nota:** no es necesario que los contagiados hayan coincidido en tiempo en el mismo momento en el mismo establecimiento.

Devolver el nombre de ambas personas, la ciudad a la que pertenecen (que tendrá que ser Valladolid) y la distancia entre los domicilios de ambas personas contagiadas como se muestra en la imagen anterior.

Tener en cuenta que aproximadamente Valladolid tiene una distancia de unos 7,5 km de punta a punta, por lo que si alguna de las distancias es mayor de este valor, será indicativo de que hay algún tipo de error.