

Descripción corta del challenge

Realización de consultas a una base de datos orientada a grafos que contiene información sobre si una persona está contagiada o no de COVID-19, y los lugares que ha visitado.

La base de datos

Adjunto capturas de pantalla de los primeros registros de los csv utilizados para creación de la BD, y los scripts para crear nodos y relaciones en Neo4j

Nodos_Persona.csv

id_persona	nombre_persona	estado_salud	hora_test_resultado	latitud_domicilio	longitud_domicilio
1	Lara Fernandez	Sano	2020-09-05T00:43:18	41.65659	-4.74017
2	Fidel Figueroa	Contagiado	2020-09-04T11:18:34	41.64418	-4.71989
3	Marcelino Rodriguez	Contagiado	2020-09-03T13:39:54	41.6451	-4.72687
4	Fatima Camino	Sano	2020-09-03T07:58:25	41.64108	-4.71349
5	Berta Romero	Contagiado	2020-09-02T05:36:03	41.66009	-4.72864
6	Celia Cuenca	Sano	2020-09-05T12:47:29	41.66049	-4.73338
7	Alvaro Fernandez	Sano	2020-09-01T07:37:28	41.65885	-4.7499
8	Diego Alonso	Contagiado	2020-09-04T12:58:16	41.65291	-4.72707
9	Sofía Vitores	Contagiado	2020-09-03T06:34:39	41.64347	-4.71747
10	Maria Soto	Sano	2020-09-02T17:48:23	41.64679	-4.73083

Nodos_Establecimientos.csv

id_ubicacion	nombre_establecimiento	tipo_establecimiento
1	Bar Comandante	Bar
2	Teatro Calderon	Teatro
3	Colegio San José	Colegio
4	Hospital Rio Horteiga	Hospital
5	Rio Shopping	Centro Comercial
6	Teatro Juan del Enzina	Teatro
7	Hospital Santisima Trinidad	Hospital
8	Bar La Chica de Ayer	Bar
9	Vialia	Centro Comercial
10	Colegio Amor de Dios	Colegio

Nodos_Visita.csv

id_visita	id_persona	id_ubicacion	inicio_visita	fin_visita
1	9	4	2020-09-01T05:31:39	01/09/2020 10:16
2	5	3	2020-08-31T09:36:03	31/08/2020 14:33
3	5	2	2020-09-01T18:00:03	01/09/2020 19:32
4	17	2	2020-09-02T17:59:42	02/09/2020 19:59
5	8	5	2020-09-02T17:18:16	02/09/2020 23:41
6	9	5	2020-09-02T18:17:39	02/09/2020 22:57
7	19	3	2020-09-02T09:29:02	02/09/2020 14:21
8	11	1	2020-09-01T16:12:30	01/09/2020 19:42
9	11	2	2020-09-02T17:57:35	02/09/2020 20:13
10	7	3	2020-08-31T09:31:25	31/08/2020 14:38

Crear nodos "Persona", "Ubicación":

```
LOAD CSV WITH HEADERS FROM 'file:///home/imfbigdata/neo4j/Nodos_Persona.csv' AS csv

CREATE (p:Persona{id:toInt(csv.id_persona), nombre: csv.nombre_persona, estado:
csv.estado_salud, hora_result_test: datetime(csv.hora_test_resultado),
ubicacion_domicilio: point({x: toFloat(csv.latitud_domicilio), y:
toFloat(csv.longitud_domicilio), crs:'wgs-84'})})

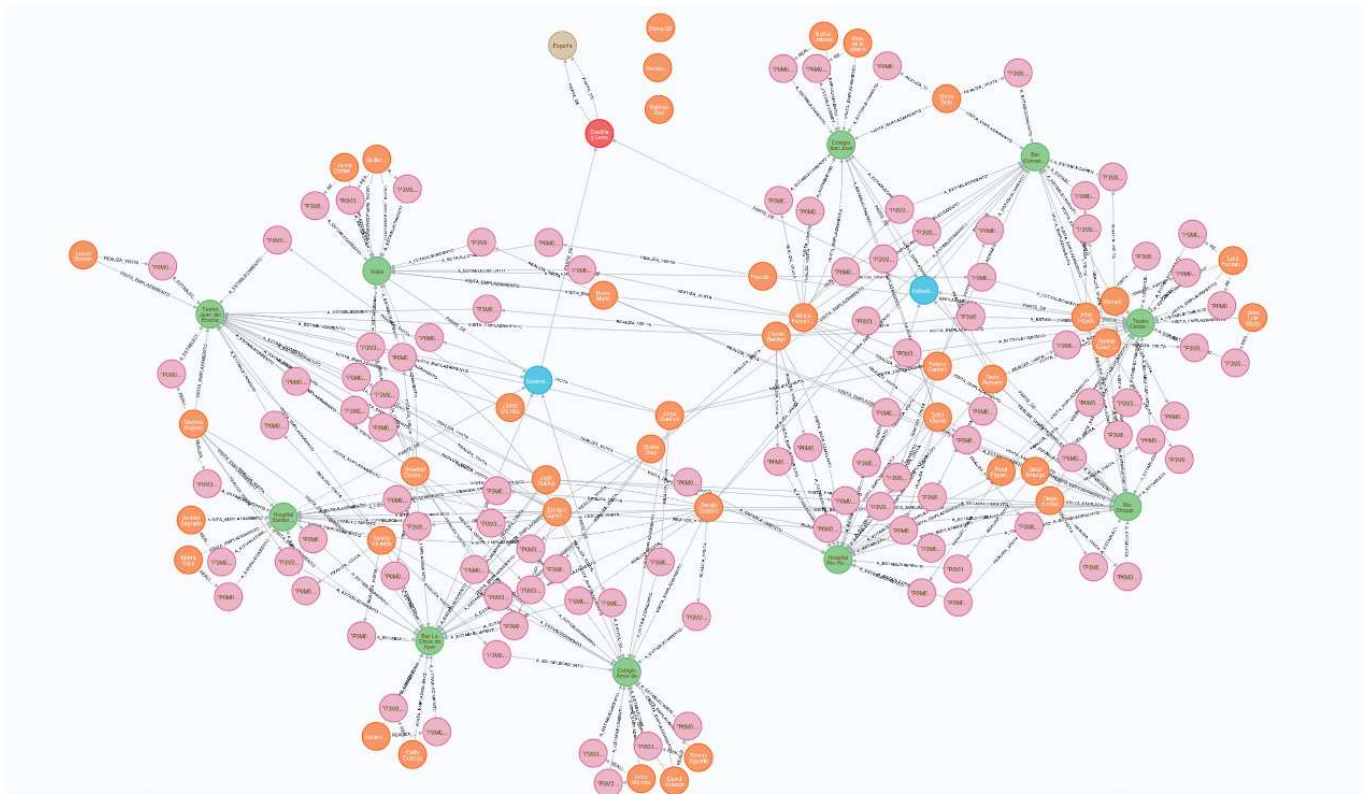
LOAD CSV WITH HEADERS FROM 'file:///home/imfbigdata/neo4j/Nodos_Establecimiento.csv'
AS csv

CREATE (u:Ubicacion{id:toInt(csv.id_ubicacion), nombre: csv.nombre_establecimiento,
tipo: csv.tipo_establecimiento})
```

Crear relaciones "Visita_Emplazamiento":

```
CREATE (p)-[vi:VISITA_EMPLAZAMIENTO{id: toInt(csv.id_visita),
inicio_visita:datetime(csv.inicio_visita), fin_visita:datetime(csv.fin_visita)}]->(u)
set vi.duration=duration.inSeconds(v.inicio_visita,v.fin_visita)
```

Vista del grafo:



Consulta 1

Enunciado

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:

Respuesta

```
1 MATCH (p:Persona)
2 WITH
3     CASE WHEN (p.estado = "Sano") THEN 1 ELSE 0 END AS Sanos,
4     CASE WHEN (p.estado = "Contagiado") THEN 1 ELSE 0 END AS Contagiados
5 RETURN SUM(Sanos) AS Total_Sanos, SUM(Contagiados) AS Total_Contagiados
```

```
$ MATCH (p:Persona) WITH CASE WHEN (p.estado = "Sano") THEN 1 ELSE 0 END AS Sanos, CASE WHEN
```

	Table	
	Total_Sanos	Total_Contagiados
Text	30	10
Code		

Consulta 2

Enunciado

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*.

Respuesta

```
1 MATCH (psana:Persona{estado:"Sano"})-[visana:VISITA_EMPLAZAMIENTO]→(:Ubicacion)←[vicont:VISITA_EMPLAZAMIENTO]-(pcont:Persona{estado:"Contagiado"})
2 WHERE viconc.inicio_visita < visana.fin_visita
3 RETURN DISTINCT psana.nombre AS Nombre_persona_contagiada
```

```
$ MATCH (psana:Persona{estado:"Sano"})-[visana:VISITA_EMPLAZAMIENTO]→(:Ubicacion)←[vicont:VISITA_EMPLAZAMIENTO]-(pcont:...
```

Table	Nombre_persona_contagiada
Text	"Lara Fernandez"
Code	"Alba Hoyos"
	"Daniel Benitez"
	"Ramon Cuadrado"
	"Francisco Pardo"
	"Jorge Cuenca"
	"Enrique Caminero"
	"Sergio Barbero"
	"Fatima Camino"

Started streaming 19 records after 1 ms and completed after 2 ms.

Consulta 3

Enunciado

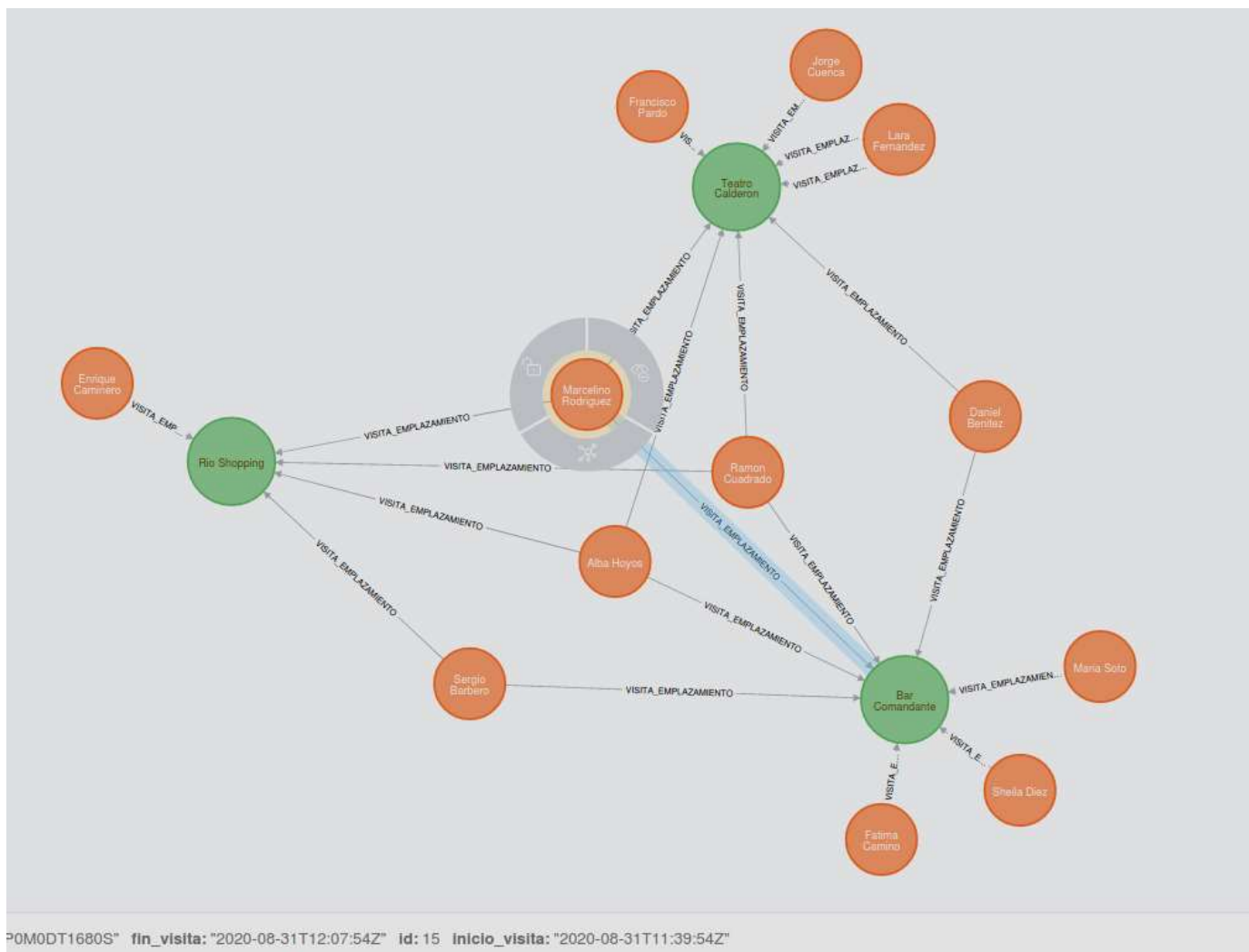
Mostrar grafo con las personas sanas que han coincidido con una persona contagiada, concretamente con "Marcelino Rodriguez". Hay que mostrar en un grafo con este nodo persona, junto con todos los lugares que ha visitado, y con los nodos de etiqueta persona sana que han visitado también ese lugar posteriormente.

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:

Respuesta

La consulta, y el grafo resultante, es el siguiente:

```
1 MATCH path = (psana:Persona{estado:"Sano"})-[visana:VISITA_EMPLAZAMIENTO]→(:Ubicacion)
  ←[vicont:VISITA_EMPLAZAMIENTO]-(pcont:Persona{nombre:"Marcelino Rodriguez"})
2 WHERE viconc.inicio_visita < visana.fin_visita
3 RETURN path
```



Explorando el grafo, vemos que hay 6 personas que sólo han tenido un contacto con Marcelino: todos los nodos que sólo hacen coincidencia en un establecimiento excepto Lara Fernández que realiza 2 visitas al teatro. Además de ser únicos, los contactos de estas 6 personas son INDIRECTOS, ninguno coincide en los establecimientos con Marcelino.

Consulta 4

Enunciado

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.

Respuesta

```
1 MATCH (psana:Persona{estado:"Sano"})-[visana:VISITA_EMPLAZAMIENTO]->(u:Ubicacion)<-[vicont:VISITA_EMPLAZAMIENTO]-
  (pcont:Persona{nombre:"Marcelino Rodriguez"})
2 WHERE viconc.inicio_visita < visana.fin_visita
3 RETURN pcont.nombre AS Esparcidor_virus, viconc.inicio_visita AS Comienzo_esparcimiento_virus, u.nombre AS
  Establecimiento, psana.nombre AS Persona_en_riesgo, visana.inicio_visita AS Inicio_visita_persona_riesgo
4 ORDER BY viconc.inicio_visita
```

\$ MATCH (psana:Persona{estado:"Sano"})-[visana:VISITA_EMPLAZAMIENTO]->(u:Ubicacion)<-[vicont:VISITA_EMPLAZAMIENTO]-(pcont...

"Esparcidor_virus"	"Comienzo_esparcimiento_virus"	"Establecimiento"	"Persona_en_riesgo"	"Inicio_visita_persona_riesgo"
"Marcelino Rodriguez"	"2020-08-30T18:39:54Z"	"Teatro Calderon"	"Ramon Cuadrado"	"2020-09-02T19:31:35Z"
"Marcelino Rodriguez"	"2020-08-30T18:39:54Z"	"Teatro Calderon"	"Daniel Benitez"	"2020-09-03T17:53:52Z"
"Marcelino Rodriguez"	"2020-08-30T18:39:54Z"	"Teatro Calderon"	"Francisco Pardo"	"2020-09-02T17:59:17Z"
"Marcelino Rodriguez"	"2020-08-30T18:39:54Z"	"Teatro Calderon"	"Lara Fernandez"	"2020-09-03T09:24:18Z"
"Marcelino Rodriguez"	"2020-08-30T18:39:54Z"	"Teatro Calderon"	"Lara Fernandez"	"2020-09-04T20:23:18Z"
"Marcelino Rodriguez"	"2020-08-30T18:39:54Z"	"Teatro Calderon"	"Jorge Cuenca"	"2020-09-02T17:56:22Z"
"Marcelino Rodriguez"	"2020-08-30T18:39:54Z"	"Teatro Calderon"	"Alba Hoyos"	"2020-09-02T17:57:35Z"
"Marcelino Rodriguez"	"2020-08-31T11:39:54Z"	"Bar Comandante"	"Sergio Barbero"	"2020-09-01T17:21:13Z"
"Marcelino Rodriguez"	"2020-08-31T11:39:54Z"	"Bar Comandante"	"Fatima Camino"	"2020-09-02T08:30:25Z"
"Marcelino Rodriguez"	"2020-08-31T11:39:54Z"	"Bar Comandante"	"Sheila Diez"	"2020-09-02T18:11:21Z"
"Marcelino Rodriguez"	"2020-08-31T11:39:54Z"	"Bar Comandante"	"Daniel Benitez"	"2020-08-31T16:52:36Z"

MAX COLUMN WIDTH

Consulta 5-1

Enunciado

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.

Respuesta

```
MATCH (psana:Persona{estado:"Sano"})-[visana:VISITA_EMPLAZAMIENTO]->(u:Ubicacion)<-[vicont:VISITA_EMPLAZAMIENTO]-(pcont:Persona{estado:"Contagiado"})
```

```
WITH psana.nombre AS en_riesgo, pcont.nombre AS contagiador, u.nombre AS local,
apoc.coll.max([visana.inicio_visita, viconc.inicio_visita]) AS maxStart,
apoc.coll.min([visana.fin_visita, viconc.fin_visita]) AS minEnd
```

```
WHERE maxStart <= minEnd
```

```
RETURN contagiador AS Persona_contagiada, COLLECT({Persona_en_contacto:en_riesgo,
Establecimiento:local, Fecha_comienzo_solapamiento:maxStart,
Fecha_fin_solapamiento:minEnd}) AS Contactos
```

```

"Marcelino Rodriguez" [
  {
    "Fecha_fin_solapamiento":
    "2020-09-01T21:18:37Z",
    "Establecimiento": "Rio
Shopping",
    "Fecha_comienzo_solapamiento":
    "2020-09-01T19:41:39Z",
    "Persona_en_contacto": "Sergio
Barbero"
  }
]

```

```

"Berta Romero" [
  {
    "Fecha_fin_solapamiento":
    "2020-08-31T22:16:03Z",
    "Establecimiento": "Rio
Shopping",
    "Fecha_comienzo_solapamiento":
    "2020-08-31T19:13:03Z",
    "Persona_en_contacto": "Alvaro
Fernandez"
  }
]

```

Consulta 5-2

Enunciado

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:

Respuesta

```

MATCH (psana:Persona{estado:"Sano"})-[visana:VISITA_EMPLAZAMIENTO]->(u:Ubicacion)<-
[vicont:VISITA_EMPLAZAMIENTO]-(pcont:Persona{estado:"Contagiado"})

```

```

WITH psana.nombre AS en_riesgo, pcont.nombre AS contagiador,
apoc.coll.max([visana.inicio_visita, vicont.inicio_visita]) AS maxStart,
apoc.coll.min([visana.fin_visita, vicont.fin_visita]) AS minEnd

```

```

WHERE maxStart <= minEnd

```

```

WITH contagiador AS Persona_contagiada, COUNT(en_riesgo) AS Numero_contactos_sanos

```

```

RETURN Persona_contagiada, Numero_contactos_sanos

```

```

ORDER BY Numero_contactos_sanos DESC

```

"Persona_contagiada"	"Numero_contactos_sanos"
"Israel Amargo"	6
"Berta Romero"	5
"Sofía Vitores"	5
"Saturnino Castaño"	5
"Jose Luis Mozo"	3
"David Velasco"	3
"Marcelino Rodriguez"	1
"Diego Alonso"	1

Consulta 6

Enunciado

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:

Respuesta

```
1 MATCH (pcont:Persona{estado:"Contagiado"})-[vicont:VISITA_EMPLAZAMIENTO]→(u:Ubicacion)
2 WHERE pcont.hora_result_test < vicont.inicio_visita
3 RETURN pcont.nombre, vicont.inicio_visita AS Fecha_inicio_visita, pcont.hora_result_test AS Fecha_conf_contagio
```

\$ MATCH (pcont:Persona{estado:"Contagiado"})-[vicont:VISITA_EMPLAZAMIENTO]→(u:Ubicacion) WHERE pcont.hora_result_test < ...

Table	"pcont.nombre"	"Fecha_inicio_visita"	"Fecha_conf_contagio"
Text	"Marta Valle"	"2020-09-03T09:23:21Z"	"2020-09-01T19:07:22Z"
Code			

Consulta 7

Enunciado

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*.

Respuesta

```
MATCH (psana:Persona{estado:"Sano"})-[visana:VISITA_EMPLAZAMIENTO]->(u:Ubicacion)-[vicont:VISITA_EMPLAZAMIENTO]-(pcont:Persona{estado:"Contagiado"})

WITH psana.nombre AS en_riesgo, pcont.nombre AS contagiador,
apoc.coll.max([visana.inicio_visita, vicont.inicio_visita]) AS maxStart,
apoc.coll.min([visana.fin_visita, vicont.fin_visita]) AS minEnd

WHERE maxStart <= minEnd

RETURN contagiador AS Persona_contagiada, en_riesgo AS Persona_en_riesgo,
ROUND(toFloat(duration.inSeconds(maxStart,minEnd).seconds)/3600*10000)/10000 AS
Duracion_contacto

ORDER BY Persona_en_riesgo
```

"Persona_contagiada"	"Persona_en_riesgo"	"Duracion_contacto"
"Jose Luis Mozo"	"Alba Hoyos"	1.0369
"Israel Amargo"	"Alba Hoyos"	1.9942
"Berta Romero"	"Alvaro Fernandez"	3.05
"Berta Romero"	"Alvaro Fernandez"	4.95
"Sofía Vitores"	"Alvaro Fernandez"	4.9503
"Saturnino Castaño"	"Celia Cuenca"	1.6978
"Sofía Vitores"	"Daniel Benitez"	1.9731
"Saturnino Castaño"	"Enrique Caminero"	1.1383
"Berta Romero"	"Esther Liebana"	4.7175
"Sofía Vitores"	"Esther Liebana"	5.0575
"Jose Luis Mozo"	"Francisco Pardo"	1.0086

Consulta 8

Enunciado

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).

Respuesta

```
MATCH (psana:Persona{estado:"Sano"})-[visana:VISITA_EMPLAZAMIENTO]->(u:Ubicacion)-[vicont:VISITA_EMPLAZAMIENTO]-(pcont:Persona{estado:"Contagiado"})

WITH psana.nombre AS en_riesgo, pcont.nombre AS contagiador,
apoc.coll.max([visana.inicio_visita, vicont.inicio_visita]) AS maxStart,
apoc.coll.min([visana.fin_visita, vicont.fin_visita]) AS minEnd

WHERE maxStart <= minEnd

WITH en_riesgo,
SUM(ROUND(toFloat(duration.inSeconds(maxStart,minEnd).seconds)/3600*10000)/10000) AS
Tiempo_total_exposición

RETURN en_riesgo AS Persona_en_riesgo, Tiempo_total_exposición

ORDER BY Tiempo_total_exposición DESC

LIMIT 5
```

"Persona_en_riesgo"	"Tiempo_total_exposición"
"Alvaro Fernandez"	12.9503
"Sergio Barbero"	9.7883
"Esther Liebana"	9.775
"Maria Soto"	9.4972
"Noelia Valentin"	7.2006

Consulta 9

Enunciado

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.

Respuesta

// Al explorar los datos, me di cuenta de que hay valores raros en las duraciones de las visitas.

```
MATCH (p:Persona)-[v:VISITA_EMPLAZAMIENTO]->(u:Ubicacion)
```

```
RETURN p.nombre,v.duration,v.id,u.nombre
```

```
ORDER BY v.duration
```

```
$ MATCH (p:Persona)-[v:VISITA_EMPLAZAMIENTO]→(u:Ubicacion) RETURN p.nombre,v.duration
```

"p.nombre"	"v.duration"	"v.id"	"u.nombre"
"Fatima Camino"	"P0M0DT-78660S"	25	"Hospital Rio Hortega"
"Lara Fernandez"	"P0M0DT-69480S"	14	"Teatro Calderon"
"Soledad Cadenas"	"P0M0DT0S"	82	"Teatro Juan del Enzina"
"Guillermo Carpintero"	"P0M0DT0S"	62	"Vialia"
"Marcelino Rodriguez"	"P0M0DT1680S"	15	"Bar Comandante"
"Fatima Camino"	"P0M0DT3240S"	20	"Bar Comandante"
"Marcelino Rodriguez"	"P0M0DT4500S"	28	"Teatro Calderon"
"Jose Luis Mozo"	"P0M0DT5280S"	37	"Teatro Calderon"
"Berta Romero"	"P0M0DT5520S"	3	"Teatro Calderon"
"Enrique Caminero"	"P0M0DT5549S"	86	"Teatro Juan del Enzina"
"Noelia Valentin"	"P0M0DT5572S"	68	"Teatro Juan del Enzina"

"Diego Alonso"	"P0M0DT22980S"	5	"Rio Shopping"
"Marta Valle"	"P0M0DT25943S"	76	"Hospital Santisima Trinidad"
"Bruno Marin"	"P0M0DT189464S"	49	"Vialia"

// Teniendo en cuenta estos valores anómalos, procedemos con el cálculo de número de visitas:

// Decidimos eliminar visitas con duraciones menor o igual que 0, o superior a 12h

```

// Vamos a necesitar filtrar las visitas teniendo en cuenta la premisa que hemos
indicado.

// Y posteriormente, tendremos que usar las visitas filtradas para calcular el número
de visitas

// En primer lugar, hacemos esta operación para las visitas de los contagiados,
guardando los resultados en un array de documentos con el comando COLLECT.

MATCH (pcont:Persona{estado:"Contagiado"})-[vicont:VISITA_EMPLAZAMIENTO]-
>(u:Ubicacion)-[]->(c:Ciudad)

WITH u.nombre AS Ubicacion, c.nombre AS Ciudad, vicont.duration.seconds AS
visitas_cont_secs

WHERE visitas_cont_secs > 0 AND visitas_cont_secs < 12*3600

WITH Ubicacion, Ciudad, COUNT(visitas_cont_secs) AS visitas_total_contagiados

WITH COLLECT ({Ubicacion:Ubicacion, Ciudad:Ciudad,
Visitas_total_contagiados:visitas_total_contagiados}) as data_contagiados


//Repetimos la operación, pero para las visitas de personas sanas.

MATCH (psano:Persona{estado:"Sano"})-[visano:VISITA_EMPLAZAMIENTO]->(u:Ubicacion)

WITH u.nombre AS Ubicacion, visano.duration.seconds AS visitas_sano_secs,
data_contagiados

WHERE visitas_sano_secs > 0 AND visitas_sano_secs< 12*3600

WITH Ubicacion, COUNT(visitas_sano_secs) AS visitas_total_sanos, data_contagiados

WITH COLLECT ({Ubicacion:Ubicacion, Visitas_total_sanos:visitas_total_sanos}) as
data_sanos, data_contagiados


// Extraemos la información con UNWIND de los arrays que hemos guardado

UNWIND data_contagiados AS c

UNWIND data_sanos AS s


// Juntamos toda la información de ambos arrays, incluimos la sentencia WHERE para
evitar multiplicación de registros

WITH c.Ubicacion AS Establecimiento, s.Ubicacion AS Establecimiento_s, c.Ciudad AS
Ciudad, c.Visitas_total_contagiados AS Num_visitas_contagiados,
c.Visitas_total_contagiados+s.Visitas_total_sanos AS Num_total_visitas

WHERE Establecimiento=Establecimiento_s

```

// Calculamos el resto de columnas que nos pedían en base a las que tenemos

WITH Establecimiento, Ciudad, Num_visitas_contagiados, Num_total_visitas,
ROUND(toFloat(Num_visitas_contagiados)/Num_total_visitas*10000)/100 AS
Porcentaje_visitas_personas_contagiadas

RETURN Establecimiento, Num_visitas_contagiados, Num_total_visitas,
Porcentaje_visitas_personas_contagiadas, Ciudad

ORDER BY Porcentaje_visitas_personas_contagiadas DESC

"Establecimiento"	"Num_visitas_contagiados"	"Num_total_visitas"	"Porcentaje_visitas_personas_contagiadas"	"Ciudad"
"Hospital Rio Hortega"	6	10	60.0	"Valladolid"
"Teatro Calderon"	6	12	50.0	"Valladolid"
"Rio Shopping"	5	13	38.46	"Valladolid"
"Colegio San José"	2	7	28.57	"Valladolid"
"Colegio Amor de Dios"	2	10	20.0	"Salamanca"
"Bar Comandante"	1	8	12.5	"Valladolid"
"Hospital Santisima Trinidad"	1	8	12.5	"Salamanca"
"Bar La Chica de Ayer"	1	11	9.09	"Salamanca"

// Los dos establecimientos con mayor porcentaje de visitas de contagiados son en Valladolid, mientras que los establecimientos que no aparecen en esta tabla por no tener visitas de contagiados, pertenecen a Salamanca. En general, vemos que los establecimientos de Valladolid tienen muchísima más afluencia de contagiados que los de Salamanca.

Consulta 10

Enunciado

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.

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.

Respuesta

```
MATCH (p1:Persona{estado:"Contagiado"})-[v1:VISITA_EMPLAZAMIENTO]->(u:Ubicacion)<-  
[v2:VISITA_EMPLAZAMIENTO]-(p2:Persona{estado:"Contagiado"})  
  
MATCH (u:Ubicacion)-[]->(c:Ciudad)  
  
WHERE (ID(p1) < ID(p2)) AND (c.nombre="Valladolid")  
  
RETURN DISTINCT p1.nombre AS Persona_contagiada_1, p2.nombre AS Persona_contagiada_2,  
c.nombre AS Ciudad,  
ROUND(distance(p1.ubicacion_domicilio,p2.ubicacion_domicilio))/1000 AS  
Distancia_domicilios_km  
  
ORDER BY Distancia_domicilios_km DESC  
  
LIMIT 3
```

"Persona_contagiada_1"	"Persona_contagiada_2"	"Ciudad"	"Distancia_domicilios_km"
"Berta Romero"	"Sofía Vitores"	"Valladolid"	2.224
"Fidel Figueroa"	"Berta Romero"	"Valladolid"	2.016
"Marcelino Rodriguez"	"Berta Romero"	"Valladolid"	1.675