MSDS 420 - MODULE 9 ASSIGNMENT: GRAPH DATA MODELS

Steve Desilets

May 28, 2023

Homework questions:

1. Use the provided Cypher script to create the graph database

a. You could use any names for your project and the graph database

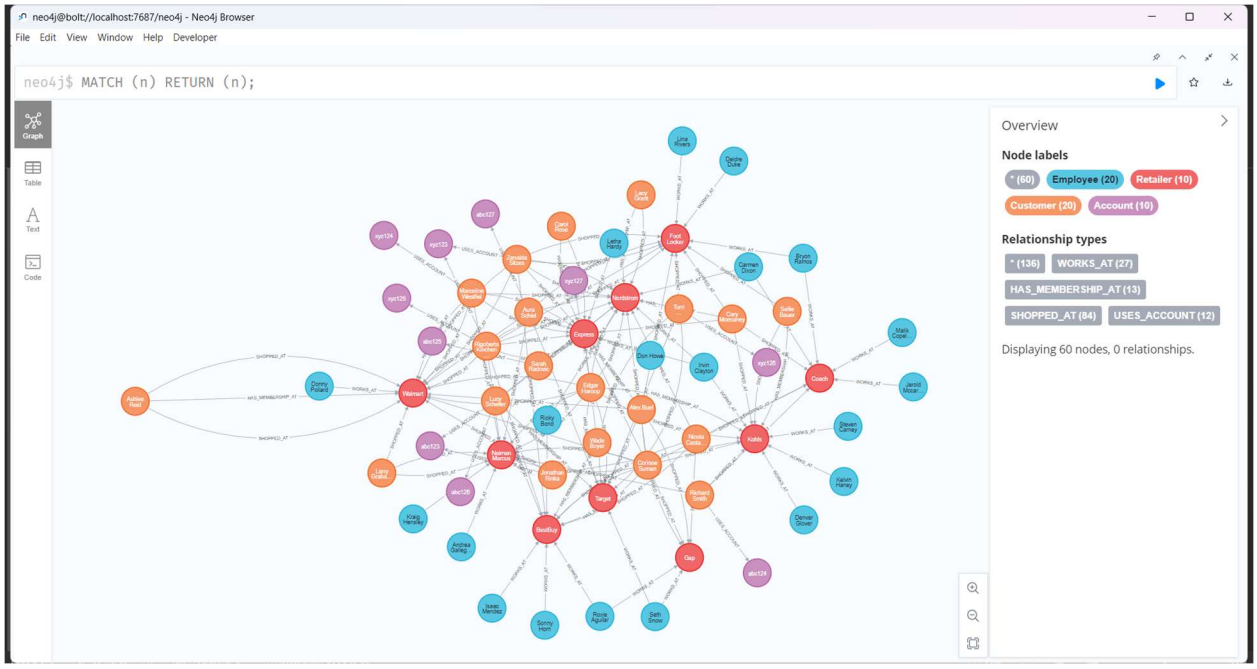
b. Copy the ENTIRE Cypher code in the script and paste it in ne4oj$ prompt and then click the blue play button on the right.

c. NOTE in step 15 above that your version may only allow one command at a time.

d. Run the command below. Find the Customer Ashlee Reid and pull the node to the far left of the screen. Include a screen capture of this view to show you were able to load the database (5 points)

MATCH (n) RETURN (n);

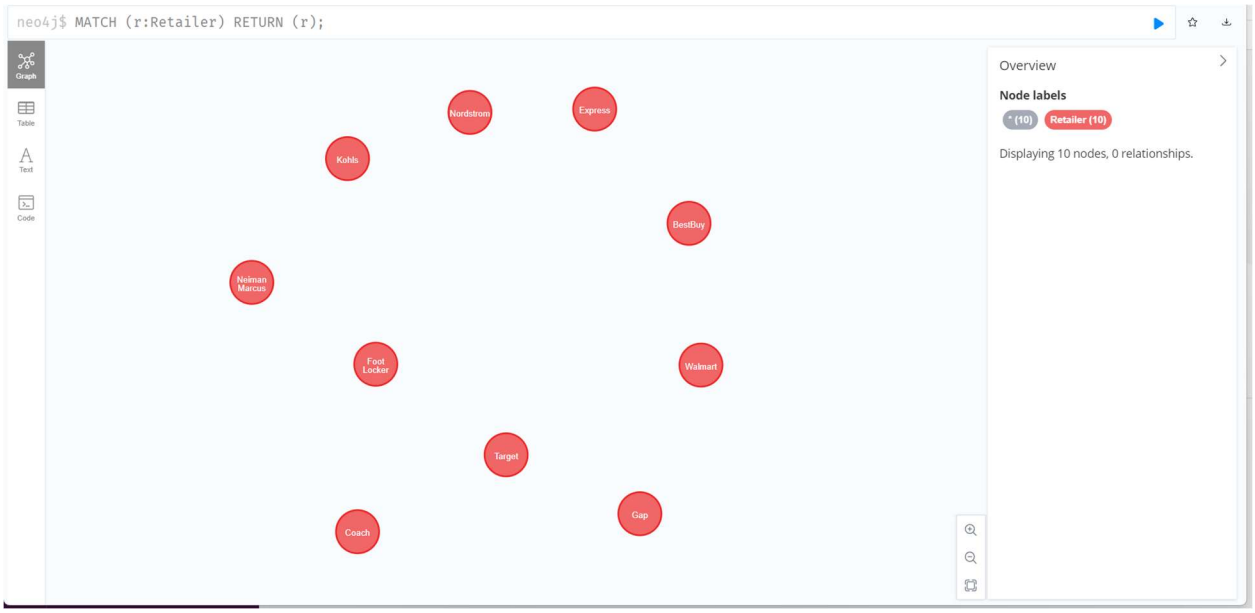
Please see the screenshot below for the output.



2. Execute the following Cypher code to get the list of retailers: (0 point)

MATCH (r:Retailer) RETURN (r);

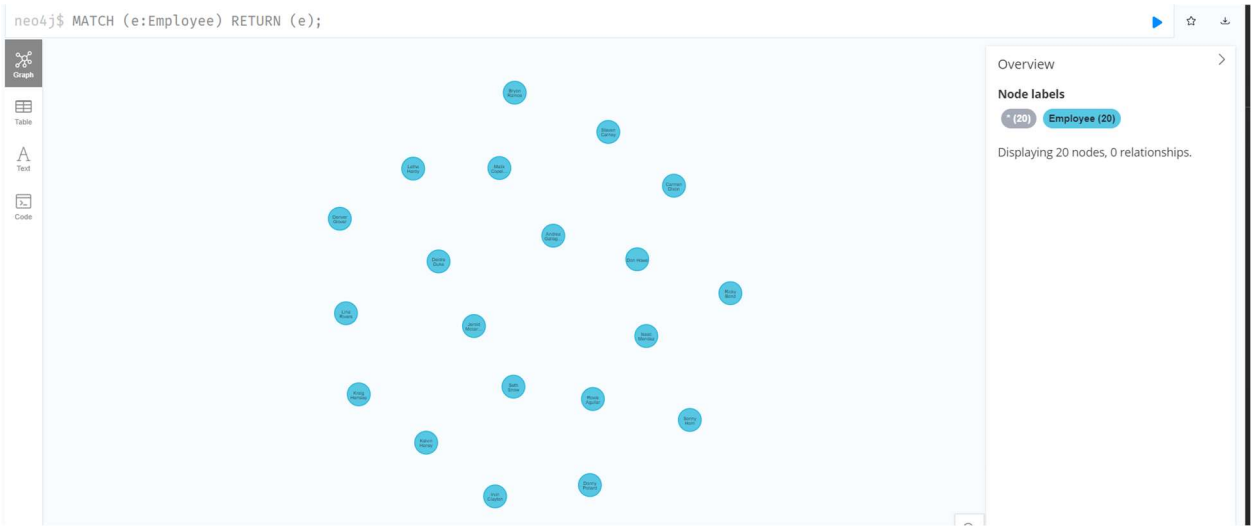
Please see the screenshot below for the output.



3. Execute the following Cypher code to the get the list of employees: (0 point)

MATCH (e:Employee) RETURN (e);

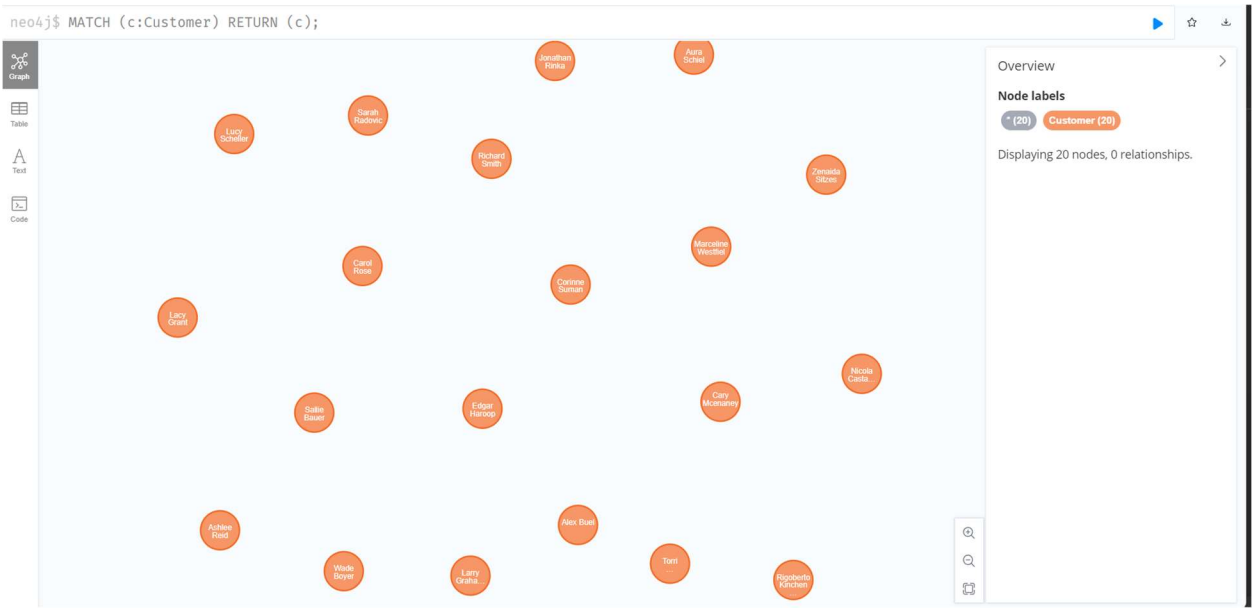
Please see the screenshot below for the output.



4. Execute the following Cypher code to the get the list of customers: (0 point)

MATCH (c:Customer) RETURN (c);

Please see the screenshot below for the output of the code.



5. Execute the following Cypher code to the get the list of all disputed transactions: (0 point)
MATCH (customer:Customer)-[transaction:SHOPPED_AT]->(retailer) WHERE transaction.status = "Disputed"
RETURN customer.name AS `Customer Name`, retailer.name AS `Retailer Name`, transaction.amount AS `Transaction Amount`, transaction.date AS `Transaction date`
ORDER BY `Transaction date` DESC
Please see the screenshot below for the output of the code.

```
MATCH (customer:Customer)-[transaction:SHOPPED_AT]→(retailer) WHERE transaction.status = "Disputed"
RETURN customer.name AS `Customer Name`, retailer.name AS `Retailer Name`, transaction.amount AS `Transaction Amount`,
transaction.date AS `Transaction date`
ORDER BY `Transaction date` DESC
```

Customer Name	Retailer Name	Transaction Amount	Transaction date
" Nicola Castanon "	"Coach"	"721"	"7/17/2020"
" Zenaida Sitzes "	"Express"	"1884"	"5/7/2020"
" Marcelne Westfield "	"Express"	"533"	"5/6/2020"
"Edgar Haroop"	"Neiman Marcus"	"1732"	"5/26/2020"
"Edgar Haroop"	"Kohls"	"1021"	"5/23/2020"
"Lucy Scheller"	"BestBuy"	"424"	"5/20/2020"
"Larry Grahamn"	"Walmart"	"425"	"5/18/2020"
"Larry Grahamn"	"Neiman Marcus"	"475"	"5/19/2020"
"Richard Smith"	"Kohls"	"875"	"5/13/2020"
" Rigoberto Kinchen "	"BestBuy"	"424"	"5/10/2020"
"Jonathan Rinka"	"Neiman Marcus"	"375"	"4/19/2020"
" Torri Pettway "	"Foot Locker"	"62"	"4/17/2020"
"Carol Rose"	"Express"	"721"	"4/13/2020"
"Edgar Haroop"	"Nordstrom"	"1415"	"4/1/2020"
" Rigoberto Kinchen "	"Express"	"721"	"4/1/2020"
"Edgar Haroop"	"Walmart"	"654"	"3/20/2020"
" Rigoberto Kinchen "	"Walmart"	"914"	"3/18/2020"
" Zenaida Sitzes "	"Walmart"	"1149"	"3/18/2020"
"Richard Smith"	"Coach"	"1145"	"3/18/2020"
"Ashlee Reid"	"Walmart"	"1149"	"3/18/2020"
"Sarah Radovic"	"Nordstrom"	"516"	"3/15/2020"
" Aura Schiel "	"Neiman Marcus"	"830"	"3/13/2020"
" Cary Mcenaney "	"Kohls"	"468"	"2/29/2020"
"Edgar Haroop"	"Walmart"	"1848"	"2/20/2020"
" Rigoberto Kinchen "	"Nordstrom"	"1003"	"2/20/2020"
" Corinne Suman "	"Nordstrom"	"816"	"2/20/2020"
"Lacy Grant"	"Nordstrom"	"1003"	"2/20/2020"
"Jonathan Rinka"	"Kohls"	"1345"	"2/18/2020"
" Zenaida Sitzes "	"BestBuy"	"378"	"2/10/2020"
"Sallie Bauer"	"Foot Locker"	"378"	"2/10/2020"
" Torri Pettway "	"Target"	"605"	"1/27/2020"
"Jonathan Rinka"	"Walmart"	"945"	"1/27/2020"
" Zenaida Sitzes "	"Nordstrom"	"1790"	"1/20/2020"

6. Write the Cypher code to get the number of disputed transactions for every retailer. The output should show the Retailer name and the number of disputes. Sort with highest number of disputes on top. (10 points)

Please see the screenshot below for the code and output.

```
1 MATCH (customer:Customer)-[transaction:SHOPPED_AT]→(r:Retailer)
2 WHERE transaction.status = "Disputed"
3 RETURN r.name AS Retailer_Name, count(transaction) AS Disputed_Transaction_Count
4 ORDER BY `Disputed_Transaction_Count` DESC;
5
```

Table | Text | Code

Retailer_Name	Disputed_Transaction_Count
"Walmart"	7
"Nordstrom"	6
"Express"	4
"Kohls"	4
"Neiman Marcus"	4
"BestBuy"	3
"Foot Locker"	2
"Coach"	2
"Target"	1

7. Write the Cypher code to get the number of disputed transactions and the list of customer names for these disputed transactions for every retailer. The output should show the Retailer and the customer name(s). You can consider using a collect() container, but it is not required. (10 points)

Please see the screenshot below for the code and output.

```
1 MATCH (c:Customer) - [transaction:SHOPPED_AT] → (r:Retailer)
2 WHERE transaction.status = "Disputed"
3 RETURN r.name AS Retailer_Name, collect(c.name) AS Customer_Names, count(transaction) AS Disputed_Transaction_Count
4 ORDER BY Disputed_Transaction_Count DESC;
```

Table | Text | Code

	Retailer_Name	Customer_Names	Disputed_Transaction_Count
1	"Walmart"	["Ashlee Reid", "Edgar Haroop", "Jonathan Rinka", "Edgar Haroop", " Zenaida Sitzes ", "Larry Grahamr", " Rigoberto Kinchen "]	7
2	"Nordstrom"	[" Zenaida Sitzes ", " Rigoberto Kinchen ", " Corinne Suman ", "Edgar Haroop", "Sarah Radovic", "Lacy Grant"]	6
3	"Express"	[" Marceline Westfield ", " Rigoberto Kinchen ", " Zenaida Sitzes ", "Carol Rose"]	4
4	"Kohls"	["Jonathan Rinka", "Richard Smith", " Cary Mcenaney ", "Edgar Haroop"]	4
5	"Neiman Marcus"	["Jonathan Rinka", "Edgar Haroop", " Aura Schiel ", "Larry Grahamr"]	4
6	"BestBuy"	["Lucy Scheller", " Rigoberto Kinchen ", " Zenaida Sitzes "]	3
7	"Foot Locker"	[" Torri Pettway ", "Sallie Bauer"]	2
8	"Coach"	[" Nicola Castanon ", "Richard Smith"]	2
9	"Target"	[" Torri Pettway "]	1

8. Write the Cypher code to get the number of disputed transactions for every customer that has more than one disputed transaction (10 points)

Please see the screenshot below for the code and output.

```
1 MATCH (customer:Customer)-[transaction:SHOPPED_AT]→(r:Retailer)
2 WHERE transaction.status = "Disputed"
3 WITH customer.name AS Customer_Name, count(transaction) AS Disputed_Transaction_Count
4 WHERE Disputed_Transaction_Count > 1
5 RETURN Customer_Name, Disputed_Transaction_Count
6 ORDER BY Disputed_Transaction_Count DESC
```

Table | Text | Code

Customer_Name	Disputed_Transaction_Count
"Edgar Haroop"	5
" Zenaida Sitzes "	4
" Rigoberto Kinchen "	4
"Jonathan Rinka"	3
" Torri Pettway "	2
"Larry Grahamr"	2
"Richard Smith"	2

9. Write the Cypher code to get the list of stores on LaSalle street that have disputed transactions and the number of disputed transactions for every store; the store list must be sorted by store name in ascending order. (10 points)

Please see the screenshot below for the code and output.

```
1 MATCH (customer:Customer)-[transaction:SHOPPED_AT]→(r:Retailer)
2 WHERE transaction.status = "Disputed" AND r.street CONTAINS "LaSalle"
3 RETURN r.name  AS Retailer_Name, count(transaction) AS Disputed_Transaction_Count
4 ORDER BY `Retailer_Name` ASC;
```

Table | Text | Code

Retailer_Name	Disputed_Transaction_Count
"Neiman Marcus"	4
"Nordstrom"	6

10. Write the Cypher code to get the list of Employees who work in at least 2 stores where disputed transactions reported in these retailers (10 points)

Please see the screenshot below for the code and output.

```
1 MATCH (c1:Customer) - [s1:SHOPPED_AT] → (r1:Retailer) ← [w1:WORKS_AT] - (e:Employee) - [w2:WORKS_AT] → (r2:Retailer) ←
  [s2:SHOPPED_AT] - (c2:Customer)
2 WHERE s1.status = "Disputed" AND s2.status = "Disputed"
3 RETURN DISTINCT e.name as Employee_Name
4 ORDER BY Employee_Name
```

Table | Text | Code

	Employee_Name
1	"Bryon Ramos"
2	"Carmen Dixon"
3	"Irvin Clayton"
4	"Ricky Bond"

11. Write the Cypher code to show the total amount customers spent shopping at retailers. List the customer's name and the total amount spent. (10 points)

Please see the screenshot below for the code and output.

```
1 MATCH (c:Customer)-[transaction:SHOPPED_AT]→(r:Retailer)
2 RETURN c.name AS Customer_Name, sum(toInteger(transaction.amount)) AS Total_Amount_Spent
3 ORDER BY `Total_Amount_Spent` DESC;
```

Table | Text | Code

	Customer_Name	Total_Amount_Spent
1	"Edgar Haroop"	8371
2	" Zenaida Sitzes "	7172
3	" Rigoberto Kinchen "	4937
4	" Corinne Suman "	3722
5	" Cary Mcenaney "	3159
6	" Nicola Castanon "	2738
7	"Jonathan Rinka"	2665
8	" Marceline Westfield "	2651
9	" Alex Buel "	2551
10	"Richard Smith"	2285
11	" Aura Schiel "	2043
12	"Wade Boyer"	1884
13	"Ashlee Reid"	1762
14	"Lucy Scheller"	1272
15	"Larry Grahamr"	1224
16	"Lacy Grant"	1003
17	" Torri Pettway "	843
18	"Carol Rose"	721
19	"Sallie Bauer"	721
20	"Sarah Radovic"	516

12. Write the Cypher code to show the average amount spent at each Retailer. List the Retailer and the average amount spent. Sort with highest amount on top (10 points)

Please see the screenshot below for the code and output.

```
1 MATCH (c:Customer)-[transaction:SHOPPED_AT]→(r:Retailer)
2 RETURN r.name AS Retailer_Name, round(avg(toInteger(transaction.amount)),2) AS Average_Amount_Spent
3 ORDER BY `Average_Amount_Spent` DESC;
```

Table | Text | Code

	Retailer_Name	Average_Amount_Spent
1	"Nordstrom"	888.08
2	"Walmart"	801.43
3	"Kohls"	734.0
4	"Express"	713.5
5	"Coach"	603.5
6	"Neiman Marcus"	578.55
7	"BestBuy"	448.83
8	"Target"	343.67
9	"Gap"	308.5
10	"Foot Locker"	292.22