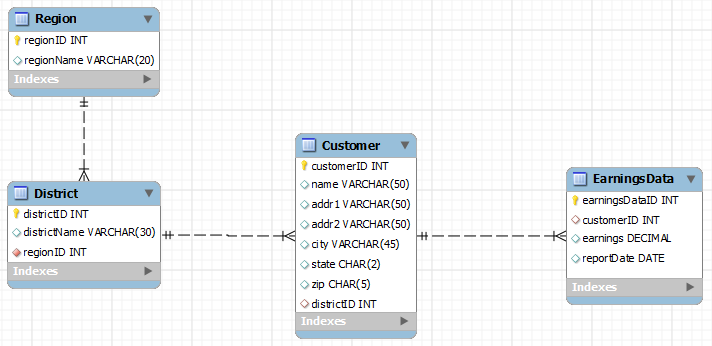
ISTE-438 Contemporary Databases

Practice Exercise – Neo4j Loading Data / Functions II

**Name: \_\_\_\_\_\_\_Archit Jain\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

Given the following ERD, we created the first two nodes, Earnings and Customer in class.



Part 1: Create the remaining nodes using the names provided:

Download the remaining two csv files from MyCourses (CSVFiles2.csv) and create the Region and District nodes. Connect the Region and District nodes with a relationship :CONTAINS\_DISTRICT and connect the District and Customer nodes with a relationship :CONTAINS\_CUSTOMER. Note that the headers with the property names may not match the ERD. Leave the files alone but make the adjustments in your create statements.

LOAD CSV WITH HEADERS FROM 'file:///CSVFiles/District.csv' as line

CREATE (:District { distID: line.distID, distName: line.distName, regionID: line.regionID})

LOAD CSV WITH HEADERS FROM 'file:///CSVFiles/ Region.csv' as line

CREATE (:Region { regionID: line.regionID, regionName: line. regionName })

MATCH (n1:District)

MATCH (n2:Region)

WHERE n1.regionID = n2.regionID

CREATE (n2)-[: CONTAINS\_DISTRICT]->(n1)

MATCH (n1:District)

MATCH (n2:Customer)

WHERE n1.distID = n2.districtID

CREATE (n2)-[: CONTAINS\_CUSTOMER]->(n1)

Part 2: Queries

1. How many customers are in the Eastern region? (you can use a different fieldname than the line.name)
2. What are the total earnings in the Eastern region? (use a longer relationship path that usual)
3. What are the average earnings for customers in the Eastern region?
4. Repeat the query above but show the totals for the Eastern and Western regions. Use the IN function. Show the region names and the totals.
5. List the region name, district name, customer name and the total earnings for each customer for all customers with names starting with the letter “S” and earnings in 2010.
6. Write update statements that will add separate properties to the Earnings nodes for the month, day and year. (Note: the report date is in the mm/dd/yyyy format) and add them individual fields and as YYYYMMDD. SPLIT on / will break it apart. This is a hard one that will require some thought. You will need SUBSTRING and SIZE to get leading zeros in MM and DD and concatenation (+) to put them together. You may want to do this in pieces with RETURNs before you do SET.

Paste statements used and the results into a MS Word document and submit to the MyCourses dropbox.

**Signature: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Have your instructor or TA sign here when PE is completed.**

MATCH (c:Region)-[r]->(e1:District)

where c.regionName ="Eastern"

with e1.distID as distID

Match (a:Customer)

where distID = a.districtID

return count(a)

╒══════════╕

│"count(a)"│

╞══════════╡

│13 │

└──────────┘

2.) MATCH (c:Region)-[r]->(e1:District)

where c.regionName ="Eastern"

with e1.distID as distID

Match (a:Customer)-[r]->(b:Earnings)

where distID = a.districtID

return Sum(b.earnings)

╒═════════════════╕

│"Sum(b.earnings)"│

╞═════════════════╡

│5321.0 │

└─────────────────┘

3.) MATCH (c:Region)-[r]->(e1:District)

where c.regionName ="Eastern"

with e1.distID as distID

Match (a:Customer)-[r]->(b:Earnings)

where distID = a.districtID

return Avg(b.earnings)

╒═════════════════╕

│"Avg(b.earnings)"│

╞═════════════════╡

│177.3666666666667│

└─────────────────┘

4.) MATCH (c:Region)-[r]->(e1:District)

where c.regionName IN ['Eastern','Western']

with e1.distID as distID,c.regionName as region

Match (a:Customer)-[r]->(b:Earnings)

where distID = a.districtID

return Distinct Sum(b.earnings),region

╒═════════════════╤═════════╕

│"Sum(b.earnings)"│"region" │

╞═════════════════╪═════════╡

│5321.0 │"Eastern"│

├─────────────────┼─────────┤

│5649.0 │"Western"│

└─────────────────┴─────────┘

5)

MATCH (c:Region)-[r]->(e1:District)

with e1.distID as distID,c.regionName as region,e1.distName as distname

Match (a:Customer)-[r]->(b:Earnings)

where distID = a.districtID

and a.name =~ '(?i)^s.\*'

return Distinct Sum(b.earnings),region,distname,a.name

╒═════════════════╤═════════╤═══════════════╤════════════════════════════════════╕

│"Sum(b.earnings)"│"region" │"distname" │"a.name" │

╞═════════════════╪═════════╪═══════════════╪════════════════════════════════════╡

│114.0 │"Eastern"│"Southeastern" │"Saint Andrews Golf Corp" │

├─────────────────┼─────────┼───────────────┼────────────────────────────────────┤

│329.0 │"Central"│"South Central"│"Scp Pool Corp" │

├─────────────────┼─────────┼───────────────┼────────────────────────────────────┤

│152.0 │"Central"│"Great Lakes" │"Synaptic Pharmaceutical Corp" │

├─────────────────┼─────────┼───────────────┼────────────────────────────────────┤

│920.0 │"Western"│"Northwestern" │"Sterling Financial Corp Washington"│

├─────────────────┼─────────┼───────────────┼────────────────────────────────────┤

│1665.0 │"Western"│"Southwestern" │"Shaw Brothers" │

└─────────────────┴─────────┴───────────────┴────────────────────────────────────┘

6) Match(a:Earnings)

Set a.month = Split(a.reportDate,"/")[0],

a.days= Split(a.reportDate,"/")[1],

a.year = Split(a.reportDate,"/")[2]

return a.month, a.days, a.year