

## Introduction to Graph Databases

### Activity 3 - Cypher

You will be querying two Neo4j databases, provided to you. These databases are: (1) A graph representation of the Northwind operational database, denoted **northwindhg.db**; (2) A graph representation of the Northwind data warehouse database, called **northwindDW.db**.

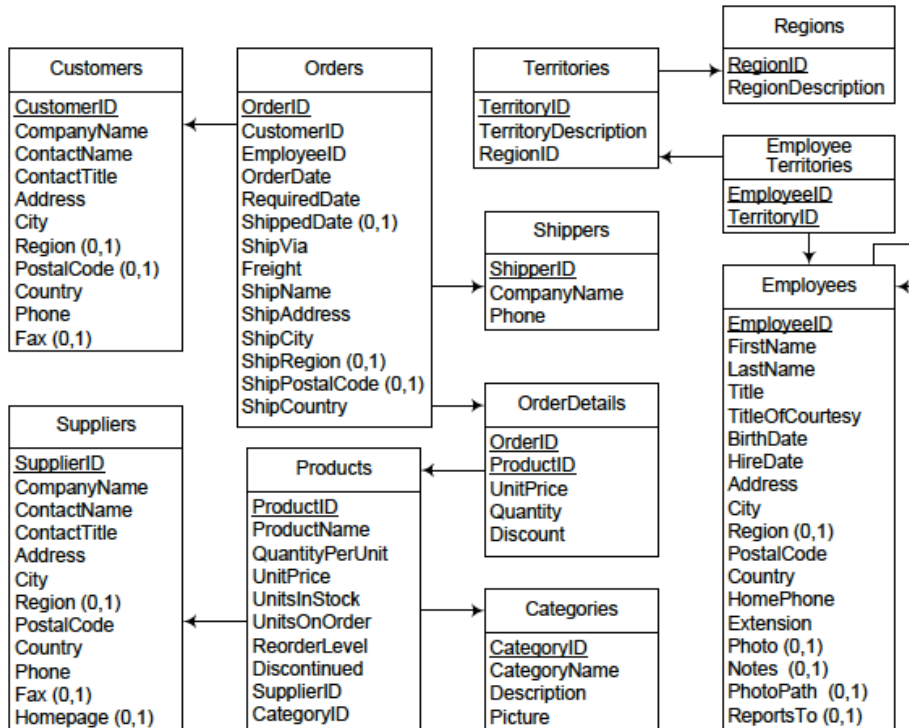
Before starting the Neo4j server, you need to choose the database you will work with. For this, you go to the **conf** folder, and edit the **neo4j.conf** file. You will find something like this:

```
#dbms.default_database=minigraphweb
#dbms.default_database=musicbrainz
#dbms.default_database=northwinddw
dbms.default_database=northwindhg
#dbms.default_database=neo4j
#dbms.default_database=webgraph3
#dbms.default_database=webdb
#dbms.default_database=telco
```

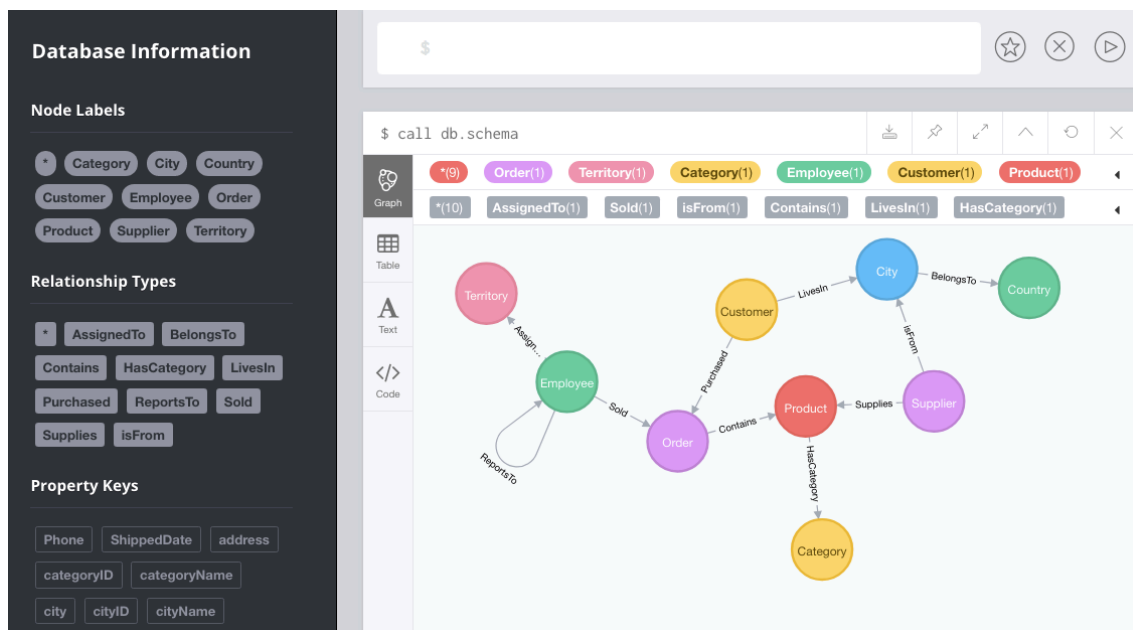
Since `dbms.default_database = northwindhg` is unmarked, to change the database to `northwinddw`, you mark `#dbms.default_database=rivers`, and unmark `dbms.default_database = northwindhg`. Save the changes, and quit the file. Then you run: `.../bin/neo4j console` to start the Server. Then, open a browser, and type the following url: **localhost:7474**. Now you can start writing Cypher queries.

## Exercise 1.

Consider the Northwind database, whose schema is:



This database has been exported to Neo4j, and you can find it at: [/...../data/databases/northwindhg](#). The graph schema is:



**Write in Cypher the following queries over the northwindhg.db database:**

**Query 1 - List products and their unit price.**

**Query 2 - List information about products 'Chocolade' & 'Pavlova'.**

**Query 3 - List information about products with names starting with a "C", whose unit price is greater than 50.**

**Query 4 - Same as 3, but considering the sales price, not the product's price.**

**Query 5 - Total amount purchased by customer and product.**

**Query 6 - Top ten employees, considering the number of orders sold.**

**Query 7 - For each employee, list the assigned territories.**

**Query 8 - For each city, list the companies settled in that city.**

**Query 9 - How many persons an employee reports to, either directly or transitively?**

**Query 10 - To whom do persons called "Robert" report to?**

**Query 11 - Who does not report to anybody?**

**Query 12 - Suppliers, number of categories they supply, and a list of such categories**

**Query 13 - Suppliers who supply beverages**

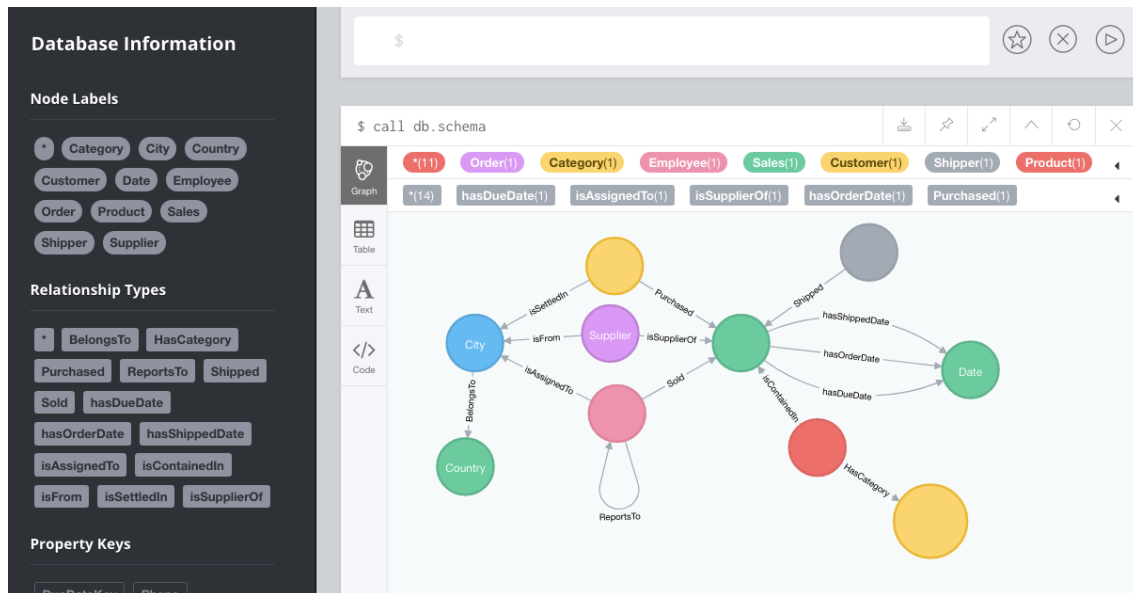
**Query 14 - Customer who purchases the largest amount of beverages**

**Query 15 - List the five most popular products (considering number of orders)**

**Query 16 - Products ordered by customers from the same country than their suppliers**

## Exercise 2.

Switch to the **northwinddw** database, doing the same steps as in **Assignment 2**. Now, the database is **northwinddw**. The schema is:



Write in Cypher the following queries over the northwindDW.db database:

Query 1. Total sales amount per customer, year, and product category

Query 2. Yearly sales amount for each pair of customer and supplier countries

Query 3. Three best-selling employees

Query 4. Best-selling employee per product and year

Query 5. Total sales and average monthly sales by employee and year

Query 6. Total sales amount and total discount amount per product and month

Query 7. Total sales amount, number of products, and sum of the quantities sold for each order

**Query 8. For each month, total number of orders, total sales amount, and average sales amount by order**

**Query 9. For each employee, total sales amount, number of cities, and number of states to which she is assigned**