/;<https://neo4j.com/graphacademy/online-training/introduction-to-neo4j/>

**Introduction to Neo4j Online Course and Tutorial**

**About the Course**

This online course is a collection of lessons and tutorials which help you build a good starting knowledge of graph databases in less than a day. It also teaches the core functionality of the Neo4j graph database. You will learn the foundation knowledge required to start building applications with Neo4j, including how to read and write Cypher queries.

**Course Outline**

* **Introduction to Graph Databases ✅**
* **Introduction to Neo4j ✅**
* **Setting up your Development Environment Tutorial ✅**
* **Introduction to Cypher ✅**
* Getting More out of Queries
* Creating Nodes and Relationships
* Getting More out of Neo4j

### [Introduction to Graph Databases](https://neo4j.com/graphacademy/online-training/introduction-to-neo4j/part-1/)

* Describe what a graph database is.
* Describe how to model relational data in a property graph model.

Estimated time: 30 minutes

### [Introduction to Neo4j](https://neo4j.com/graphacademy/online-training/introduction-to-neo4j/part-2/)

* Describe the Neo4j Graph Platform and its components.
* Describe the features and benefits of Neo4j.

Estimated time: 30 minutes

### [Setting Up Your Development Environment](https://neo4j.com/graphacademy/online-training/introduction-to-neo4j/part-3/)

* Set up your development environment for performing the hands-on exercises of this course which is one of:
  + Neo4j Sandbox
  + Neo4j Desktop

Estimated time: 30 minutes

### [Introduction to Cypher](https://neo4j.com/graphacademy/online-training/introduction-to-neo4j/part-4/)

* Use MATCH to retrieve nodes from the graph.
* Use MATCH to retrieve relationships from the graph.
* Use MATCH to retrieve properties from the graph.

Estimated time: 60 minutes

### [Getting More Out of Queries](https://neo4j.com/graphacademy/online-training/introduction-to-neo4j/part-5/)

* Use the WHERE clause for queries.
* Control query processing.
* Control how results are returned.
* Work with Cypher dates and lists.

Estimated time: 90 minutes

### [Creating Nodes and Relationships](https://neo4j.com/graphacademy/online-training/introduction-to-neo4j/part-6/)

* Create, update, and delete nodes and properties of nodes.
* Create, update, and delete relationships and properties of relationships.
* Merge data in the graph.

Estimated time: 90 minutes

### [Getting More Out of Neo4j](https://neo4j.com/graphacademy/online-training/introduction-to-neo4j/part-7/)

* Use parameters.
* Define constraints in the graph.
* Profile and monitor query execution.
* Define indexes in the graph.
* Import relational data into the graph.

Estimated time: 90 minutes

### [Summary](https://neo4j.com/graphacademy/online-training/introduction-to-neo4j/part-8/)

* Quiz results.
* Resources to learn more.
* Course feedback.

MODULE 1

Today’s business and user requirements demand applications that connect more and more of the world’s data, yet still expect high levels of performance and data reliability. Many applications of the future will be built using graph databases like Neo4j.

In this video, you will learn how the need for graph databases has evolved.

[ANA]

NEO4J is a transactional NOSQL database which still has the ACID properties

INTUITIVENESS --- SPEED --- AGILITY

**Cypher** is a declarative QUERY LANGUAGE (similar to SQL) that is optimized for GRAPHS.

### What Is a graph database?

A graph database is an online database management system with Create, Read, Update and Delete (CRUD) operations working on a graph data model. Graph databases are generally built for use with online transaction processing (OLTP) systems. Accordingly, they are normally optimized for transactional performance, and engineered with transactional integrity and operational availability in mind.

Unlike other databases, relationships take first priority in graph databases. This means your application doesn’t have to infer data connections using foreign keys or out-of-band processing, such as MapReduce.

By assembling the simple abstractions of nodes and relationships into connected structures, graph databases enable us to build sophisticated models that map closely to our problem domain.

### The case for graph databases

The biggest value that graphs bring to the development stack is their ability to store relationships and connections as first-class entities.

For instance, the early adopters of graph technology reimagined their businesses around the value of data relationships. These companies have now become industry leaders: LinkedIn, Google, Facebook and PayPal.

As pioneers in graph technology, each of these enterprises had to build their own graph database from scratch. Fortunately for today’s developers, that’s no longer the case, as graph database technology is now available off the shelf.

In this video, you will learn how graph databases help you to model real-world data that needs to be connected as well as how Neo4j is used to solve real problems facing enterprises today.

### What is a graph?

A graph is composed of two elements: **nodes** and **relationships**.

Each node represents an entity (a person, place, thing, category or other piece of data). With Neo4j, nodes can have **labels** that are used to define types for nodes. For example, a Location node is a node with the label Location. That same node can also have a label, Residence. Another Location node can also have a label, Business. A label can be used to group nodes of the same type. For example, you may want to retrieve all of the Business nodes.

Each relationship represents how two nodes are connected. For example, the two nodes Person and Location, might have the relationship LIVES\_AT pointing from a Person node to Location node. A relationship represents the verb or action between two entities. The MARRIED relationship is defined from one Person node to another Person node. Although the relationship is defined as directional, it can be queried in a non-directional manner. That is, you can query if two Person nodes have a MARRIED relationship, regardless of the direction of the relationship. For some data models, the direction of the relationship is significant. For example, in Facebook, using the KNOWS relationship is used to indicate which Person invited the other Person to be a friend.

This enables you to closely align data and connections in the graph to your real-world application. For example, a *Person* node might have a property, *name* and a *Location* node might have a property, *address*. In addition, a relationship, *MARRIED* , might have a property, *since*.

In this video, you will learn how to model property graphs containing nodes and relationships and how Cypher is used to access a graph database.

### Modeling relational to graph

Many applications’ data is modeled as relational data. There are some similarities between a relational model and a graph model:

| **Relational** | **Graph** |
| --- | --- |
| Rows | Nodes |
| Joins | Relationships |
| Table names | Labels |
| Columns | Properties |

But, there are some ways in which the relational model differs from the graph model:

| **Relational** | **Graph** |
| --- | --- |
| Each column must have a field value. | Nodes with the same label aren’t required to have the same set of properties. |
| Joins are calculated at query time. | Relationships are stored on disk when they are created. |
| A row can belong to one table. | A node can have many labels. |

#### **Run-time behavior: RDBMS vs graph**

How data is retrieved is very different between an RDBMS and a graph database:

#### **How we model: RDBMS vs graph**

How you model data from relational vs graph differs:

| **Relational** | **Graph** |
| --- | --- |
| Try and get the schema defined and then make minimal changes to it after that. | It’s common for the schema to evolve with the application. |
| More abstract focus when modeling i.e. focus on classes rather than objects. | Common to use actual data items when modeling. |

If we were modeling a football transfers graph in relational and graph databases these diagrams show what common approaches might look like.

Here is the relational model:

### How does Neo4j support the property graph model?

* Neo4j is a **Database** – use it to reliably **store information** and **find it later**.
* Neo4j’s data model is a **Graph**, in particular a **Property Graph**.
* **Cypher** is Neo4j’s graph query language (**SQL for graphs!**).
* Cypher is a declarative query language: it describes **what** you are interested in, not **how** it is acquired.
* Cypher is meant to be very **readable** and **expressive**.

### Summary

You should now be able to:

* Describe what a graph database is.
* Describe some common use cases for using a graph database.
* Describe how real-world scenarios are modeled as a graph.

MODULE 2

# **Introduction to Neo4j**

### About this module

The Neo4j Graph Platform enables developers to create applications that are best architected as graph-powered systems that are built upon the rich connectedness of data.

At the end of this module, you should be able to:

* Describe the components and benefits of the Neo4j Graph Platform.

### Neo4j Graph Platform

The Neo4j Graph Platform includes components that enable you to develop your graph-enabled application. To better understand the Neo4j Graph Platform, you will learn about these components and the benefits they provide.

#### **Neo4j Database**

The heart of the Neo4j Graph Platform is the Neo4j Database. The Neo4j Graph Platform includes out-of-the-box tooling that enables you to access graphs in Neo4j Databases. In addition, Neo4j provides APIs and drivers that enable you to create applications and custom tooling for accessing and visualizing graphs.

#### **Neo4j Database: Index-free adjacency**

With index free adjacency, when a node or relationship is written to the database, it is stored in the database as connected and any subsequent access to the data is done using pointer navigation which is very fast. Since Neo4j is a native graph database (i.e. it has a graph as its core data model), it supports very large graphs where connected data can be traversed in constant time without the need for an index.

***Neo4j Database: ACID (Atomic, Consistent, Isolated, Durable)***

Transactionality is very important for robust applications that require an ACID (atomicity, consistency, isolation, and durability) guarantees for their data. If a relationship between nodes is created, not only is the relationship created, but the nodes are updated as connected. All of these updates to the database must all succeed or fail.

***Clusters***

Neo4j supports clusters that provide high availability, scalability for read access to the data, and failover which is important to many enterprises.

#### **Graph engine**

The Neo4j graph engine is used to interpret Cypher statements and also executes kernel-level code to store and retrieve data, whether it is on disk, or cached in memory. The graph engine has been improved with every release of Neo4j to provide the most efficient access to an application’s graph data. There are many ways that you can tune the performance of the engine to suit your particular application needs.

#### **Language and driver support**

Because Neo4j is open source, you can delve into the details of how the Neo4j Database is accessed, but most developers simply use Neo4j without needing a deeper understanding of the underlying code. Neo4j provides a full stack that implements all levels of access to the database and clustering layer where you can use our published APIs. The language used for querying the Neo4j database is Cypher, an open source language.

In addition, Neo4j supports Java, JavaScript, Python, C#, and Go drivers out of the box that use Neo4j’s bolt protocol for binary access to the database layer. Bolt is an efficient binary protocol that compresses data sent over the wire as well as encrypting the data. For example, you can write a Java application that uses the Bolt driver to access the Neo4j database, and the application may use other packages that allow data integration between Neo4j and other data stores or uses as common framework such as spring.

It is also possible for you to develop your own server-side extensions in Java that access the data in the database directly without using Cypher. The Neo4j community has developed drivers for a number of languages including Ruby, PHP, and R. You can also extend the functionality of Neo4j by creating user defined functions and procedures that are callable from Cypher.

#### **Libraries**

Neo4j has a published, open source Cypher library, Awesome Procedures on Cypher (APOC) that contain many useful procedures you can call from Cypher. Another Cypher library is the Graph Algorithms library, shown here, that can help you to analyze data in your graphs. Graph analytics are important because with Neo4j, the technology can expose questions about the data that you never thought to ask. And finally, you can use the GraphQL library (tree-based subset of a graph) to access a Neo4j Database. These libraries are available as plug-ins to your Neo4j development environment, but there are many other libraries that have been written by users for accessing Neo4j.

Tools

In a development environment, you will use the Neo4j Browser or a Web browser to access data and test your Cypher statements, most of which will be used as part of your application code. Neo4j Browser is an application that uses the JavaScript Bolt driver to access the graph engine of the Neo4j database server. Neo4j also has a new tool called **Bloom** that enables you to visualize a graph without knowing much about Cypher. In addition, there are many tools for importing and exporting data between flat files and a Neo4j Database, as well as an ETL tool.

In this video, you can see how Neo4j Bloom can be used to examine and modify a Graph, even when you know very little about Cypher:

#### **Whiteboard modeling**

With a property graph model, it is very easy to collaborate with colleagues to come up with a whiteboard model of your data that is easy to understand and easy modify. You then use the model to create the nodes, relationships, labels, and properties you will use for your Neo4j data. Even after the graph has been defined and populated with data, it is easy to modify the graph as your application needs change.

Neo4j Graph Platform architecture

Here is the big picture of the Neo4j Graph Platform. The Neo4j Database provides support for graph transactions and analytics. Developers use the Neo4j Desktop, along with Neo4j Browser to develop graphs and test them, as well as implement their applications in a number of languages using supported drivers, tools and APIs. Administrators use tools to manage and monitor Neo4j Databases and clusters. Business users use out-of-the box graph visualization tools or they use custom tools. Data analysts and scientists use the analytics capabilities in the Graph Algorithm libraries or use custom libraries to understand and report findings to the enterprise. Applications can also integrate with existing databases (SQL or NoSQL), layering Neo4j on top of them to provide rich, graph-enabled access to the data.

MODULE 3

# **Setting Up Your Development Environment**

### About this module

As a developer, you will create Neo4j Databases, add and update data in them, and query the data. When you learn to use Neo4j as a developer, you have three options⎼ Neo4j Desktop, Neo4j Aura, or Neo4j Sandbox. In this module you will learn how to use each of these development environments and select the option that is best for your needs while you are learning about Neo4j.

Many graph-enabled applications have been developed and deployed using Neo4j’s Community Edition (free). If your enterprise requires production features such as failover, clustering, monitoring, advanced access control, secure routing, etc. you will use Neo4j’s Enterprise Edition, or you can use Neo4j Aura which is the cloud instance of a Neo4j Enterprise Edition Database.

At the end of this module, you should be able to:

* Determine the development environment that is best for you:
  + Install and start using the Neo4j Desktop.
  + Create a Neo4j Database instance in the cloud with Neo4j Aura.
  + Create a Neo4j Sandbox for learning Neo4j.
* Start using Neo4j Browser.

Neo4j Desktop

The Neo4j Desktop includes the Neo4j Database server which has a graph engine and kernel so that Cypher statements can be executed to access a database on your system. It includes an application called Neo4j Browser. Neo4j Browser enables you to access a Neo4j database using Cypher. You can also call built-in procedures that communicate with the database server. There are a number of additional libraries and drivers for accessing the Neo4j database from Cypher or from another programming language that you can install in your development environment. If you are looking to use your system for application development and you want to be able to create multiple Neo4j databases on your machine, you should consider downloading the Neo4j Desktop (free download). The Neo4j Desktop runs on OS X, Linux, and Windows.

Neo4j Aura

Neo4j Aura enables you to create a Neo4j Database instance in the Cloud using a monthly subscription model. The amount per month depends on the amount of memory required for the database. This frees you from needing to install Neo4j on your system. Once you create a Neo4j Database at the [Neo4j Aura site](https://console.neo4j.io/), it will be managed by Neo4j. Backups are done automatically for you and the database is available 24X7. In addition, the Neo4j will ensure that the database instance is always up-to-date with the latest version of Neo4j.

Neo4j Sandbox

The Neo4j Sandbox is another way that you can begin development with Neo4j. It is a temporary, cloud-based instance of a Neo4j Server with its associated graph that you can access from any Web browser. The database in a Sandbox may be blank or it may be pre-populated. It is started automatically for you when you create the Sandbox.

By default, the Neo4j Sandbox is available for three days, but you can extend it for up to 10 days. If you do not want to install Neo4j Desktop on your system, consider creating a Neo4j Sandbox. You must make sure that you extend your lease of the Sandbox, otherwise you will lose your graph and any saved Cypher scripts you have created in the Sandbox. However, you can use Neo4j Browser Sync to save Cypher scripts from your Sandbox. We recommend you use the Neo4j Desktop or Neo4j Aura for a real development project. The Sandbox is intended as a temporary environment or for learning about the features of Neo4j as well as specific graph use-cases.

### Steps for setting up your development environment for this training

**If you are using Neo4j Desktop**:

1. Install Neo4j Desktop.
2. In a project, create a local graph (database).
3. Start the database.
4. Click the Neo4j Browser application.

**If you are using Neo4j Aura**:

1. Create a 1G Neo4j Database in your Neo4j Aura account that is already started for you.
2. Click the link to access Neo4j Browser.

**If you are using Neo4j Sandbox**:

1. Create a Neo4j Sandbox (blank Sandbox). The Sandbox has a db already started for you.
2. Click the link to access Neo4j Browser.

### ****Guided Exercise: Getting Started with Neo4j Desktop****

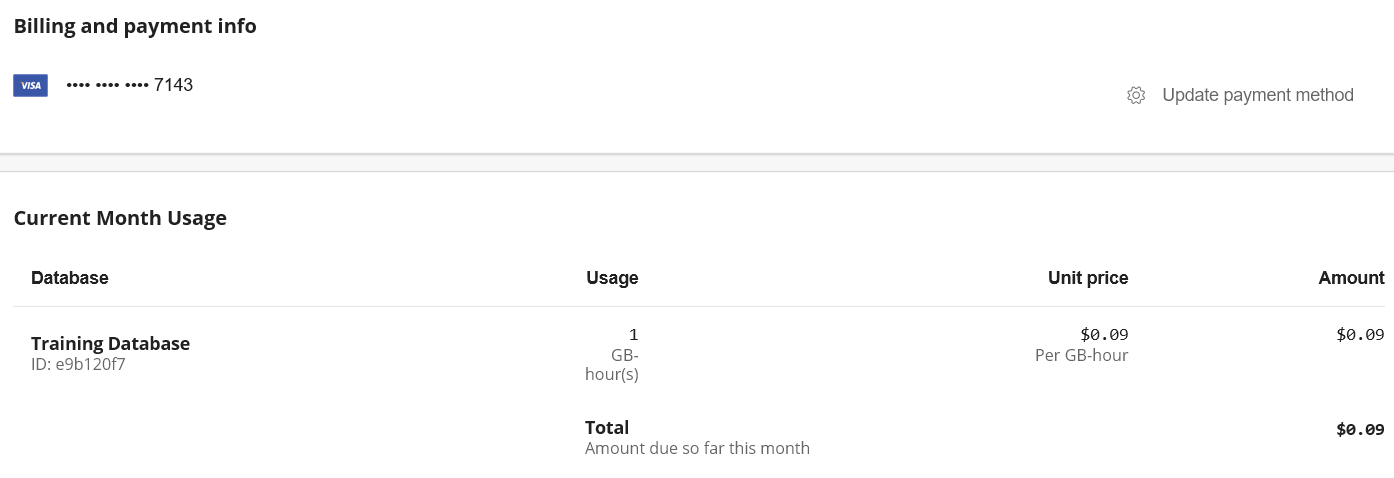
In order to perform the hands-on exercises for this training, you must install Neo4j Desktop, create a Database in Neo4j Aura, or create a Neo4j Sandbox.

If you want to download and install Neo4j Desktop on your system, follow along with one of these videos to download, install and get started using Neo4j Desktop. If you will be using Neo4j Desktop in your development environment, you can follow the steps in the video to create a TestMovies project with its corresponding Movies database.

[ANA] console.neo4j.io

wP5EmqOLSUiMGqF4NS6gCFLXEyajqbt8BfeOaq3pwjE

**DELETED!!!!!**



If you do not plan on using your database we recommend [deleting the database to avoid such charges](https://aura.support.neo4j.com/hc/en-us/articles/360045591454-Deleting-your-Neo4j-Aura-database).

No, you will only be billed for each hour, or fraction thereof, during which your database(s) existed. We display the monthly price, based on 24-hour days and 30-day months, merely as a guide to help you budget your use of the service.

Running continuously for a 30 day month will cost you $518.40

**Guided Exercise: Creating a Neo4j Sandbox**

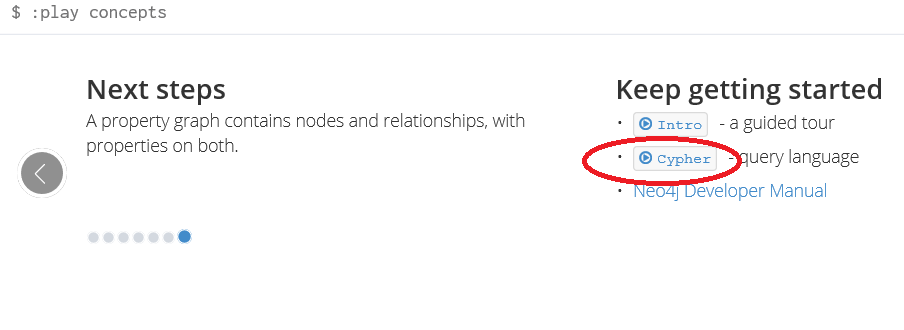
In order to perform the hands-on exercises for this training, you must install Neo4j Desktop, create a Database in Neo4j Aura, or create a Neo4j Sandbox.

If you want to run the Neo4j Database server in the cloud using a Neo4j Sandbox, follow along with this video to create a Neo4j Sandbox for training purposes. For this course, you should select the Blank Sandbox.

[ANA] sandbox.neo4j.com

IN the browser, click in START LEARNING, and advance through the guide!!

Try the CYPHER browser guide to learn simple statements:

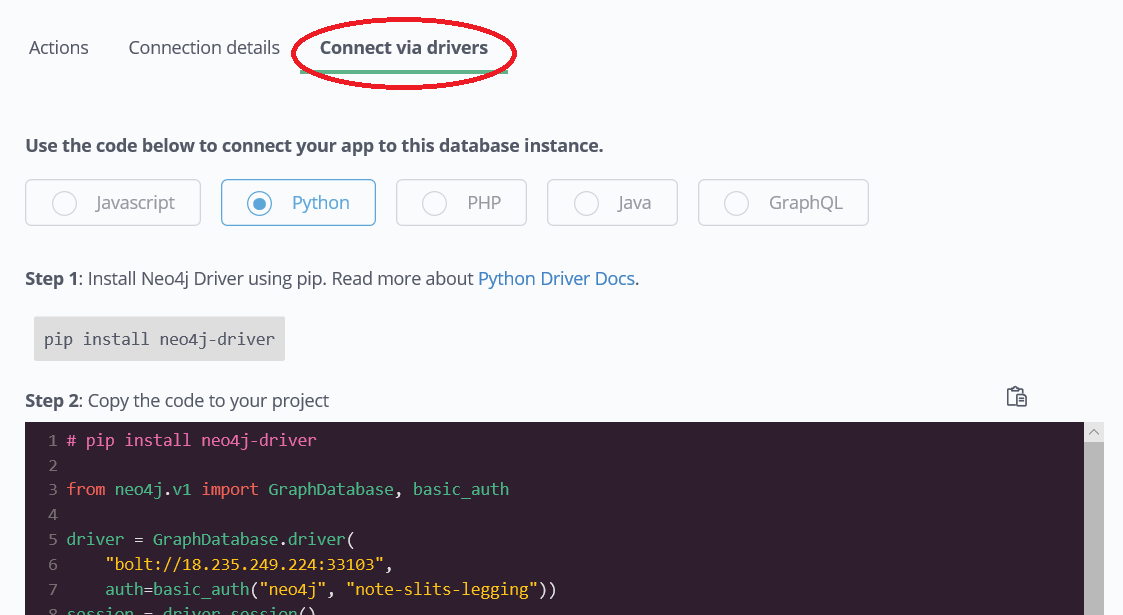


[ANA] click the example code (in the guide) and then execute!

CREATE (ee:Person { name: "Emil", from: "Sweden", klout: 99 })

CREATE (ee:Person { name: "Ana", from: "Brazil", klout: 90 })

To connect to the sandbox via PYTHON:



### Using Neo4j Browser

Neo4j Browser is a tool that enables you to access a Neo4j Database by executing Cypher statements to create or update data in the graph and to query the graph to return data. The data returned is typically visualized as nodes and relationships in a graph, but can also be displayed as tables. In addition to executing Cypher statements, you can execute a number of system calls that are related to the database being accessed by the Browser. For example, you can retrieve the list of queries that are currently running in the server.

There are two ways that you can use Neo4j Browser functionality:

* Open the Neo4j Browser application from Neo4j Desktop (database is local)
* Use the Neo4j Browser Web interface by specifying a URL in a Web browser using port 7474 (database is local, in Neo4j Aura, or in Neo4j Sandbox)

The user interface and behavior for the Neo4j Browser application and Neo4j Browser Web interface is identical, except that in the Web interface, you have the option of syncing your settings and scripts to the cloud.

### ****Guided Exercise: Getting Started with Neo4j Browser****

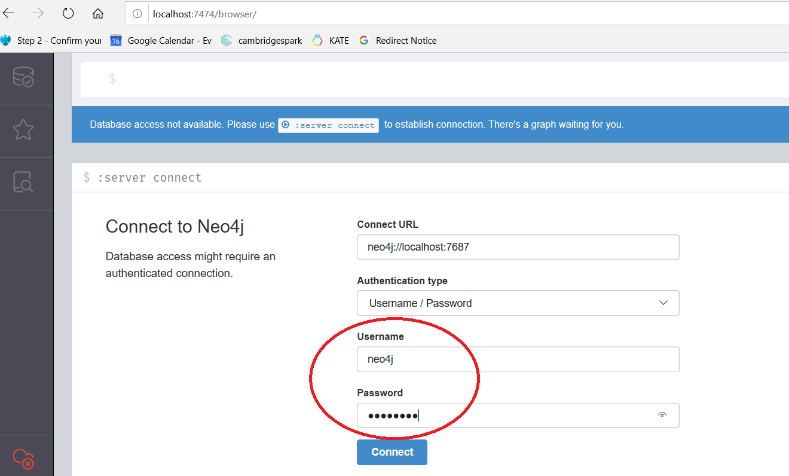
Follow along with this video to become familiar with common tasks in Neo4j Browser. You will use Neo4j Browser to populate the Movies database that is used for training as well as syncing your settings and scripts to the cloud.

Before you perform the tasks shown in this video, you must have either created and started the database in the Neo4j Desktop, created a Database in Neo4j Aura, or created a Neo4j Sandbox.

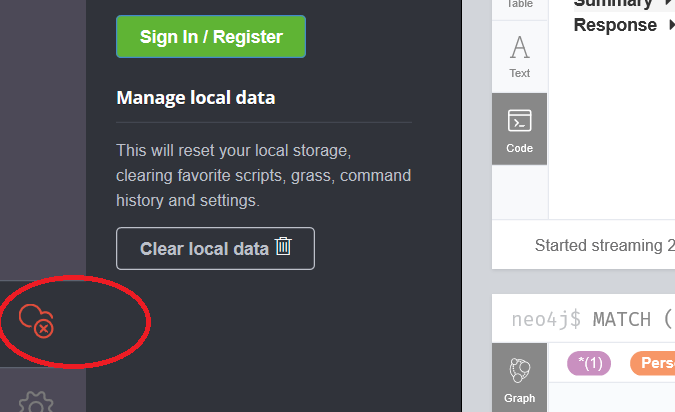
[ANA] To access the Neo4j Browser VIA WEB:

localhost:7474

username = neo4j and password = password (SET IN THE BROWSER!!)



[ANA] The only difference we see in the WEB browser is the cloud area:



[ANA] When we access the database for the first time the command below is executed:

:play start

A command begins with a colon => ‘:’

A Cypher statement does NOT begin with a colon!!!

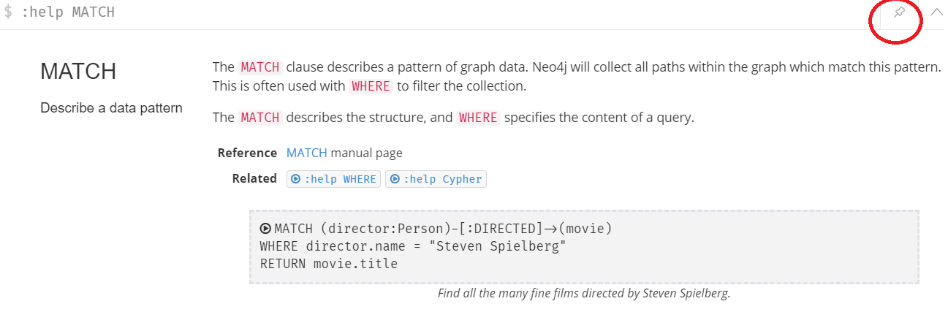
Try to type the commands below to get help for the command, keyboard shotcuts:

:help commands

:help keys

Quick help on Cypher keywords (MATCH = retrieves data from the database):

:help MATCH



The pane can be “pinned” it will always stay at this position. The result of the next command executed is placed below the pinned pane. SYSINFO shows info about the databases!

:sysinfo



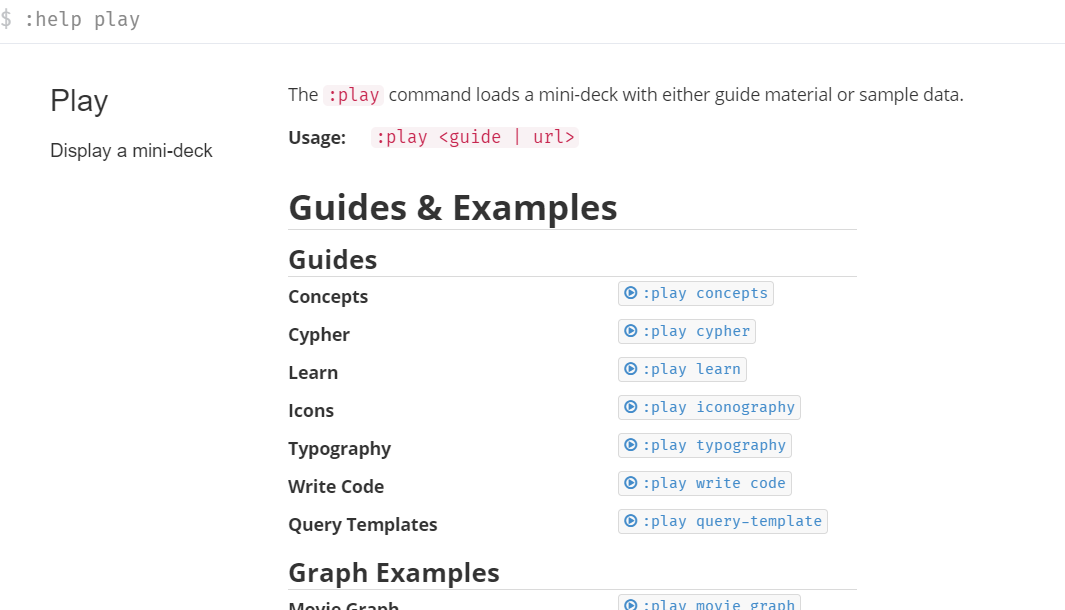
The command :history brings all commands, then just click on the one you want to run again:

:history

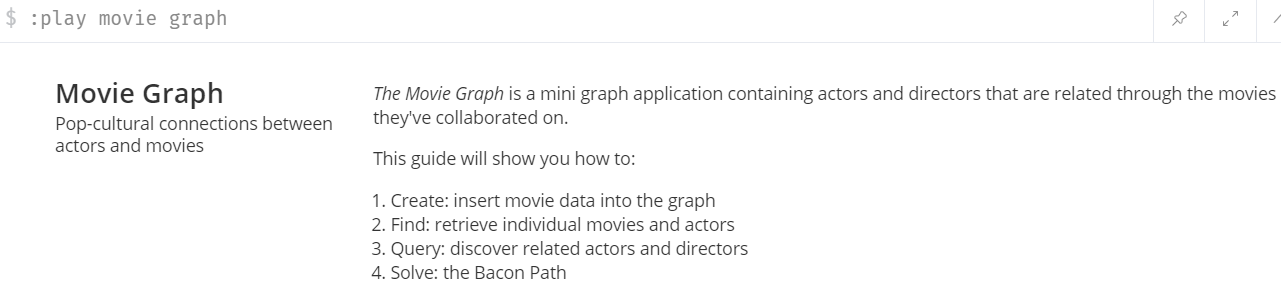
In SETTINGS we can change the # of result frames, and other limits! The command below also clears all the frames:

:clear

The :play command will execute a browser guide to show some help content



The command “:play movie” shows some things about CYPHER using the **movie** database



The command below does not return a graph, but a list of 5 “Persons”:

MATCH (p:Person) RETURN p.name LIMIT 5

To SAVE the commands I executed:

1. Click the Favorites icon on the left t
2. Click in the favorites TAB, create a folder to save my SCRIPT called Movies
3. RECALL the command that I used to create the movie DB on the right pane
4. To prevent another batch of the same records being re-created again, include the following line in the very beginning of the script:

MATCH (n) DETACH DELETE n;

1. (BEFORE that, go to SETTINGS and check the button “Enable multi statement query editor”)
2. Execute the script
3. Go back to the script and add the comment in the 1st line: “// reset database”
4. Then click the FAVORITE (star) icon
5. Then we can go to the favorites TAB and find the script there!!
6. Drag it and move it to the “Movies” folder

If you want to save your script in the CLOUD:

1. Click in the **cloud** button on the left pane
2. Sign in to the cloud
3. When you sign out all scripts will be saved in the cloud

MODULE 4

**Introduction to Cypher**

### About this module

Cypher is the query language you use to retrieve data from the Neo4j Database, as well as create and update the data.

At the end of this module, you should be able to write Cypher statements to:

* Retrieve nodes from the graph.
* Filter nodes retrieved using labels and property values of nodes.
* Retrieve property values from nodes in the graph.
* Filter nodes retrieved using relationships.

Throughout this training, you should refer to:

* [Neo4j Cypher Manual](https://neo4j.com/docs/cypher-manual/current/)
* [Cypher Reference card](http://neo4j.com/docs/cypher-refcard/current/)

### What is Cypher?

Cypher is a declarative query language that allows for expressive and efficient querying and updating of graph data. Cypher is a relatively simple and very powerful language. Complex database queries can easily be expressed through Cypher, allowing you to focus on your domain instead of getting lost in the syntax of database access.

Cypher is designed to be a human-friendly query language, suitable for both developers and other professionals. The guiding goal is to make the simple things easy, and the complex things possible.

#### **Cypher is ASCII art**

Optimized for being read by humans, Cypher’s construct uses English prose and iconography (called ASCII Art) to make queries more self-explanatory.

Being a declarative language, Cypher focuses on the clarity of expressing **what** to retrieve from a graph, not on **how** to retrieve it. You can think of Cypher as mapping English language sentence structure to patterns in a graph. For example, the nouns are nodes of the graph, the verbs are the relationships in the graph, and the adjectives and adverbs are the properties.

This is in contrast to imperative, programmatic APIs for database access. This approach makes query optimization an implementation detail instead of a burden on the developer, removing the requirement to update all traversals just because the physical database structure has changed.

Cypher is inspired by a number of different approaches and builds upon established practices for expressive querying. Many of the Cypher keywords like WHERE and ORDER BY are inspired by SQL. The pattern matching functionality of Cypher borrows concepts from SPARQL. And some of the collection semantics have been borrowed from languages such as Haskell and Python.

The Cypher language has been made available to anyone to implement and use via openCypher (opencypher.org), allowing any database vendor, researcher or other interested party to reap the benefits of our years of effort and experience in developing a first class graph query language.

#### **Nodes**

Cypher uses a pair of parentheses like (), (n) to represent a node, much like a circle on a whiteboard. Recall that a node typically represents an entity in your domain. An anonymous node, (), represents one or more nodes during a query processing where there are no restrictions of the type of node or the properties of the node. When you specify (n) for a node, you are telling the query processor that for this query, use the variable n to represent nodes that will be processed later in the query for further query processing or for returning values from the query.

#### **Labels**

Nodes in a graph are typically labeled. Labels are used to group nodes and filter queries against the graph. That is, labels can be used to optimize queries. In the Movie database you will be working with, the nodes in this graph are labeled Movie or Person to represent two types of nodes.

For example, you can see the labels in the database by simply clicking the Database icon in Neo4j Browser:

You can filter the types of nodes that you are querying, by specifying a **label** for a node. A node can have zero or more labels.

Here are simplified syntax examples for specifying a node:

()

(variable)

(:Label)

(variable:Label)

(:Label1:Label2)

(variable:Label1:Label2)

Notice that a node must have the parentheses. The labels and the variable for a node are optional.

Here are examples of specifying nodes in Cypher:

() // anonymous node not be referenced later in the query

(p) // variable p, a reference to a node used later

(:Person) // anonymous node of type Person

(p:Person) // p, a reference to a node of type Person

(p:Actor:Director) // p, a reference to a node of types Actor and Director

A node can have multiple labels. For example a node can be created with a label of Person and that same node can be modified to also have the label of Actor and/or Director.

#### **Comments in Cypher**

In Cypher, you can place a comment (starts with //) anywhere in your Cypher to specify that the rest of the line is interpreted as a comment.

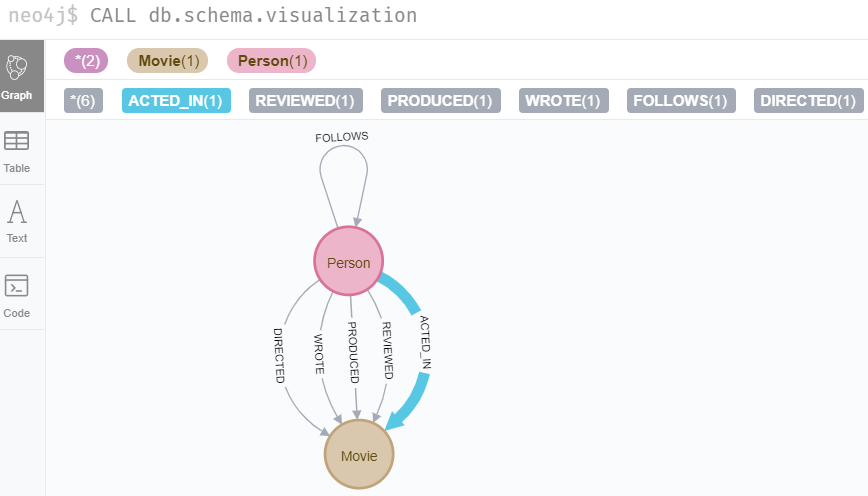
### Examining the data model

When you are first learning about the data (nodes, labels, etc.) in a graph, it is helpful to examine the data model of the graph. You do so by executing CALL db.schema, which calls the Neo4j procedure that returns information about the nodes, labels, and relationships in the graph.

For example, when we run this procedure in our training environment, we see the following in the result pane. Here we see that the graph has 2 labels defined for nodes, Person and Movie. Each type of nodes is displayed in a different color. The relationships between nodes are also displayed, which you will learn about later in this module.

[ANA] to get information about the DATA MODEL (not specific nodes):

CALL db.schema.visualization



### Using MATCH to retrieve nodes

In this video, you will be introduced to using the MATCH statement to retrieve nodes from the graph in Neo4j Browser.

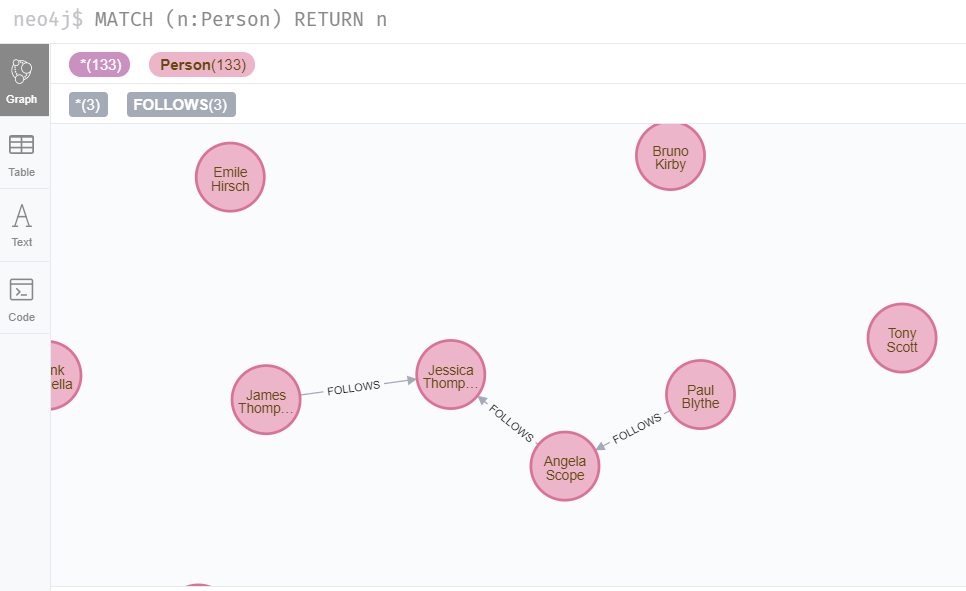
[ANA] When I click in the label “Person”, the following statement is automatically executed:

MATCH (n:Person) RETURN n LIMIT 25

It will also show relationships between the PERSON nodes (SETTING => “Connect result nodes” checkbox is checked!!!), there are THREE FOLLOWS relationships!

MATCH (n:Person) RETURN n

ZOOM out to find the relationships!!!



The most widely used Cypher clause is MATCH. The MATCH clause performs a pattern match against the data in the graph. During the query processing, the graph engine traverses the graph to find all nodes that match the graph pattern. As part of query, you can return nodes or data from the nodes using the RETURN clause. The RETURN clause must be the last clause of a query to the graph. Later in this training, you will learn how to use MATCH to select nodes and data for updating the graph. First, you will learn how to simply return nodes.

Here are simplified syntax examples for a query:

MATCH (variable)

RETURN variable

MATCH (variable:Label)

RETURN variable

Notice that the Cypher keywords MATCH and RETURN are upper-case. This coding convention is described in the Cypher Style Guide and will be used in this training. This MATCH clause returns all nodes in the graph, where the optional Label is used to return a subgraph if the graph contains nodes of different types. The variable must be specified here, otherwise the query will have nothing to return.

Here are example queries to the Movie database:

MATCH (n) // returns all nodes in the graph

RETURN n

MATCH (p:Person) // returns all Person nodes in the graph

RETURN p

When we execute the Cypher statement, MATCH (p:Person) RETURN p, the graph engine returns all nodes with the label Person. The default view of the returned nodes are the nodes that were referenced by the variable p.

When you specify a pattern for a MATCH clause, you should always specify a node label if possible. In doing so, the graph engine uses an index to retrieve the nodes which will perform better than not using a label for the MATCH.

#### **Viewing nodes as table data**

We can also view the nodes as table data where the nodes and their associated property values are shown in a JSON-style format:

When nodes are displayed as table values, the node labels and ids are not shown, only the property values for the nodes. Node ids are unique identifiers and are set by the graph engine when a node is created.

### ****Exercise 1: Retrieving nodes****

In the query edit pane of Neo4j Browser, execute the browser command: :play intro-neo4j-exercises and follow the instructions for Exercise 1.

### Properties

In Neo4j, a node (and a relationship, which you will learn about later) can have properties that are used to further define a node. A property is identified by its property key. Recall that nodes are used to represent the entities of your business model. A property is defined for a node and not for a type of node. All nodes of the same type need not have the same properties.

For example, in the Movie graph, all Movie nodes have both title and released properties. However, it is not a requirement that every Movie node has a property, tagline.

Properties can be used to filter queries so that a subset of the graph is retrieved. In addition, with the RETURN clause, you can return property values from the retrieved nodes, rather than the nodes.

#### **Examining property keys**

As you prepare to create Cypher queries that use property values to filter a query, you can view the values for property keys of a graph by simply clicking the Database icon in Neo4j Browser. Alternatively, you can execute CALL db.propertyKeys, which calls the Neo4j library method that returns the property keys for the graph.

Here is what you will see in the result pane when you call the method to return the property keys in the Movie graph. This result stream contains all property keys in the graph. It does not display which nodes utilize these property keys.

#### **Retrieving nodes filtered by a property value**

You have learned previously that you can filter node retrieval by specifying a label. Another way you can filter a retrieval is to specify a value for a property. Any node that matches the value will be retrieved.

Here are simplified syntax examples for a query where we specify one or more values for properties that will be used to filter the query results and return a subset of the graph:

MATCH (variable {propertyKey: propertyValue})

RETURN variable

MATCH (variable:Label {propertyKey: propertyValue})

RETURN variable

MATCH (variable {propertyKey1: propertyValue1, propertyKey2: propertyValue2})

RETURN variable

MATCH (variable:Label {propertyKey: propertyValue, propertyKey2: propertyValue2})

RETURN variable

Here is an example where we filter the query results using a property value. We only retrieve Person nodes that have a born property value of 1970.

MATCH (p:Person {born: 1970})

RETURN p

Here is an example where we specify two property values for the query.

MATCH (m:Movie {released: 2003, tagline: 'Free your mind'})

RETURN m

As it turns out, there is only one movie with the tagline, Free your mind in the Movie database, but if we had specified a different year, the query would not have returned a value because when you specify properties, both properties must match.

#### **Returning property values**

In this video, you will see how to return property values to the output stream when you retrieve nodes from the graph in Neo4j Browser.

[ANA] retrieving specific properties (using ALIAS):

MATCH (m:Movie {released: 2006}) RETURN m.title, m.released AS `Year Released`

Thus far, you have seen how to retrieve nodes and return nodes (entire graph or a subset of the graph). You can use the RETURN clause to return property values of nodes retrieved.

Here are simplified syntax examples for returning property values, rather than nodes:

MATCH (variable {prop1: value})

RETURN variable.prop2

MATCH (variable:Label {prop1: value})

RETURN variable.prop2

MATCH (variable:Label {prop1: value, prop2: value})

RETURN variable.prop3

MATCH (variable {prop1:value})

RETURN variable.prop2, variable.prop3

In this example, we use the born property to refine the query, but rather than returning the nodes, we return the name and born values for every node that satisfies the query.

MATCH (p:Person {born: 1965})

RETURN p.name, p.born

#### **Specifying aliases**

If you want to customize the headings for a table containing property values, you can specify **aliases** for column headers.

Here is the simplified syntax for specifying an alias for a property value:

MATCH (variable:Label {propertyKey1: propertyValue1})

RETURN variable.propertyKey2 AS alias2

If you want a heading to contain a space between strings, you must specify the alias with the back tick ` character, rather than a single or double quote character. In fact, you can specify any variable, label, relationship type, or property key with a space also by using the back tick ` character.

Here we specify aliases for the returned property values:

MATCH (p:Person {born: 1965})

RETURN p.name AS name, p.born AS `birth year`

### ****Exercise 2: Filtering queries using property values****

In the query edit pane of Neo4j Browser, execute the browser command: :play intro-neo4j-exercises and follow the instructions for Exercise 2.

### Relationships

Relationships are what make Neo4j graphs such a powerful tool for connecting complex and deep data. A relationship is a **directed** connection between two nodes that has a **relationship type** (name). In addition, a relationship can have properties, just like nodes. In a graph where you want to retrieve nodes, you can use relationships between nodes to filter a query.

#### **ASCII art**

Thus far, you have learned how to specify a node in a MATCH clause. You can specify nodes and their relationships to traverse the graph and quickly find the data of interest.

Here is how Cypher uses ASCII art to specify path used for a query:

() // a node

()--() // 2 nodes have some type of relationship

()-->() // the first node has a relationship to the second node

()<--() // the second node has a relationship to the first node

#### **Querying using relationships**

In your MATCH clause, you specify how you want a relationship to be used to perform the query. The relationship can be specified with or without direction.

Here are simplified syntax examples for retrieving a set of nodes that satisfy one or more directed and typed relationships:

MATCH (node1)-[:REL\_TYPE]->(node2)

RETURN node1, node2

MATCH (node1)-[:REL\_TYPEA | :REL\_TYPEB]->(node2)

RETURN node1, node2

where:

|  |  |
| --- | --- |
| *node1* | is a specification of a node where you may include node labels and property values for filtering. |
| *:REL\_TYPE* | is the type (name) for the relationship. For this syntax the relationship is from *node1* to *node2*. |
| *:REL\_TYPEA* , *:REL\_TYPEB* | are the relationships from *node1* to *node2.* The nodes are returned if at least one of the relationships exists. |
| *node2* | is a specification of a node where you may include node labels and property values for filtering. |

#### **Examining relationships**

You can run CALL db.schema to view the relationship types in the graph. In the Movie graph, we see these relationships between the nodes:

Here we see that this graph has a total of 6 relationship types between the nodes. Some Person nodes are connected to other Person nodes using the FOLLOWS relationship type. All of the other relationships in this graph are from Person nodes to Movie nodes.

#### **Using a relationship in a query**

Here is an example where we retrieve the Person nodes that have the ACTED\_IN relationship to the Movie, The Matrix. In other words, show me the actors that acted in The Matrix.

MATCH (p:Person)-[rel:ACTED\_IN]->(m:Movie {title: 'The Matrix'})

RETURN p, rel, m

For this query, we are using the variable p to represent the Person nodes during the query, the variable m to represent the Movie node retrieved, and the variable rel to represent the relationship for the relationship type, ACTED\_IN. We return a graph with the Person nodes, the Movie node and their ACTED\_IN relationships.

**Important:** You specify node labels whenever possible in your queries as it optimizes the retrieval in the graph engine. That is, you should **not** specify this same query as:

MATCH (p)-[rel:ACTED\_IN]->(m {title:'The Matrix'})

RETURN p,m

#### **Querying by multiple relationships**

Here is another example where we want to know the movies that Tom Hanks acted in and directed:

MATCH (p:Person {name: 'Tom Hanks'})-[:ACTED\_IN|:DIRECTED]->(m:Movie)

RETURN p.name, m.title

Notice that there are multiple rows returned for the movie, That Thing You Do. This is because Tom Hanks acted in and directed that movie.

#### **Using anonymous nodes in a query**

Suppose you wanted to retrieve the actors that acted in The Matrix, but you do not need any information returned about the Movie node. You need not specify a variable for a node in a query if that node is not returned or used for later processing in the query. You can simply use the anonymous node in the query as follows:

MATCH (p:Person)-[:ACTED\_IN]->(:Movie {title: 'The Matrix'})

RETURN p.name

[ANA] ->(:Movie NO NODE VARIABLE SPECIfIED HERE!!!!

A best practice is to place named nodes (those with variables) before anonymous nodes in a MATCH clause.

#### **Using an anonymous relationship for a query**

Suppose you want to find all people who are in any way connected to the movie, The Matrix. You can specify an empty relationship type in the query so that all relationships are traversed and the appropriate results are returned. In this example, we want to retrieve all Person nodes that have any type of connection to the Movie node, with the title, The Matrix. This query returns more nodes with the relationships types, DIRECTED, ACTED\_IN, and PRODUCED.

MATCH (p:Person)-->(m:Movie {title: 'The Matrix'})

RETURN p, m

Here are other examples of using the anonymous relationship:

MATCH (p:Person)--(m:Movie {title: 'The Matrix'})

RETURN p, m

MATCH (m:Movie)<--(p:Person {name: 'Keanu Reeves'})

RETURN p, m

In this training, we will use -->, --, and <-- to represent anonymous relationships as it is a Cypher best practice.

#### **Retrieving the relationship types**

There is a built-in function, type() that returns the relationship type of a relationship.

Here is an example where we use the rel variable to hold the relationships retrieved. We then use this variable to return the relationship types.

MATCH (p:Person)-[rel]->(:Movie {title:'The Matrix'})

RETURN p.name, type(rel)

#### **Retrieving properties for relationships**

Recall that a node can have as set of properties, each identified by its property key. Relationships can also have properties. This enables your graph model to provide more data about the relationships between the nodes.

Here is an example from the Movie graph. The movie, The Da Vinci Code has two people that reviewed it, Jessica Thompson and James Thompson. Each of these Person nodes has the REVIEWED relationship to the Movie node for The Da Vinci Code. Each relationship has properties that further describe the relationship using the summary and rating properties.

Just as you can specify property values for filtering nodes for a query, you can specify property values for a relationship. This query returns the name of of the person who gave the movie a rating of 65.

MATCH (p:Person)-[:REVIEWED {rating: 65}]->(:Movie {title: 'The Da Vinci Code'})

RETURN p.name

### Using patterns for queries

Thus far, you have learned how to specify nodes, properties, and relationships in your Cypher queries. Since relationships are directional, it is important to understand how patterns are used in graph traversal during query execution. How a graph is traversed for a query depends on what directions are defined for relationships and how the pattern is specified in the MATCH clause.

Here is an example of where the FOLLOWS relationship is used in the Movie graph. Notice that this relationship is directional.

We can perform a query that returns all Person nodes who follow Angela Scope:

MATCH (p:Person)-[:FOLLOWS]->(:Person {name:'Angela Scope'})

RETURN p

If we reverse the direction in the pattern, the query returns different results:

MATCH (p:Person)<-[:FOLLOWS]-(:Person {name:'Angela Scope'})

RETURN p

#### **Traversing relationships**

Since we have a graph, we can traverse through nodes to obtain relationships further into the traversal.

For example, we can write a Cypher query to return all followers of the followers of Jessica Thompson.

MATCH (p:Person)-[:FOLLOWS]->(:Person)-[:FOLLOWS]->(:Person {name:'Jessica Thompson'})

RETURN p

This query could also be modified to return each person along the path by specifying variables for the nodes and returning them. In addition, you can assign a variable to the path and return the path as follows:

MATCH path = (:Person)-[:FOLLOWS]->(:Person)-[:FOLLOWS]->(:Person {name:'Jessica Thompson'})

RETURN path

#### **Using relationship direction to optimize a query**

When querying the relationships in a graph, you can take advantage of the direction of the relationship to traverse the graph. For example, suppose we wanted to get a result stream containing rows of actors and the movies they acted in, along with the director of the particular movie.

Here is the Cypher query to do this. Notice that the direction of the traversal is used to focus on a particular movie during the query:

MATCH (a:Person)-[:ACTED\_IN]->(m:Movie)<-[:DIRECTED]-(d:Person)

RETURN a.name, m.title, d.name

### Cypher style recommendations

Here are the **Neo4j-recommended** Cypher coding standards that we use in this training:

* Node labels are CamelCase and begin with an upper-case letter (examples: Person, NetworkAddress). Note that node labels are case-sensitive.
* Property keys, variables, parameters, aliases, and functions are camelCase and begin with a lower-case letter (examples: businessAddress, title). Note that these elements are case-sensitive.
* Relationship types are in upper-case and can use the underscore. (examples: ACTED\_IN, FOLLOWS). Note that relationship types are case-sensitive and that you cannot use the “-” character in a relationship type.
* Cypher keywords are upper-case (examples: MATCH, RETURN). Note that Cypher keywords are case-insensitive, but a best practice is to use upper-case.
* String constants are in single quotes, unless the string contains a quote or apostrophe (examples: ‘The Matrix’, “Something’s Gotta Give”). Note that you can also escape single or double quotes within strings that are quoted with the same using a backslash character.
* Specify variables only when needed for use later in the Cypher statement.
* Place named nodes and relationships (that use variables) before anonymous nodes and relationships in your MATCH clauses when possible.
* Specify anonymous relationships with -->, --, or <--

Here is an example showing some best coding practices:

MATCH (:Person {name: 'Diane Keaton'})-[movRel:ACTED\_IN]->

(:Movie {title:"Something's Gotta Give"})

RETURN movRel.roles

We recommend that you follow the [Cypher Style Guide](https://github.com/opencypher/openCypher/blob/master/docs/style-guide.adoc) when writing your Cypher statements.

### ****Exercise 3: Filtering queries using relationships****

In the query edit pane of Neo4j Browser, execute the browser command: :play intro-neo4j-exercises and follow the instructions for Exercise 3.