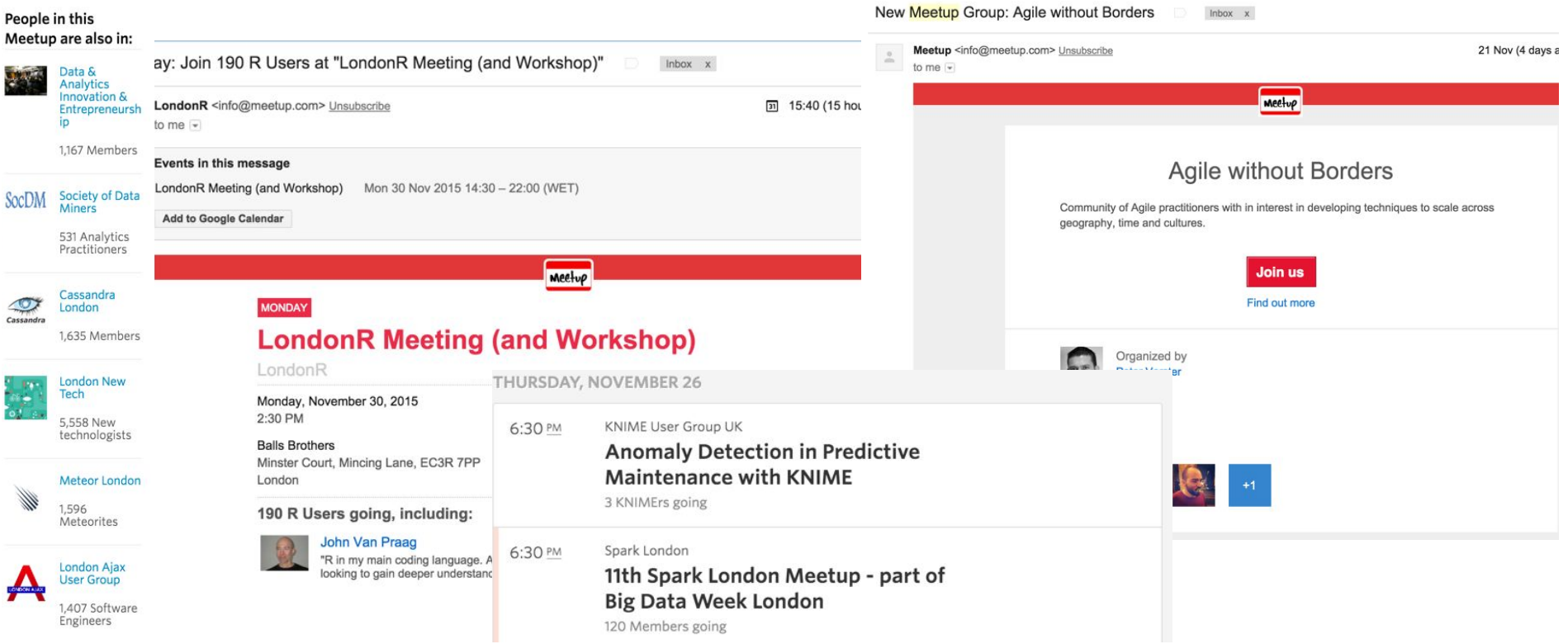
The Meetup Dataset

What are we going to do?

- Introduce the dataset
- Modeling workflow
- Intro to the property graph model
- Load CSV data

The Meetup dataset

Meetup makes recommendations



What recommendations can we make?

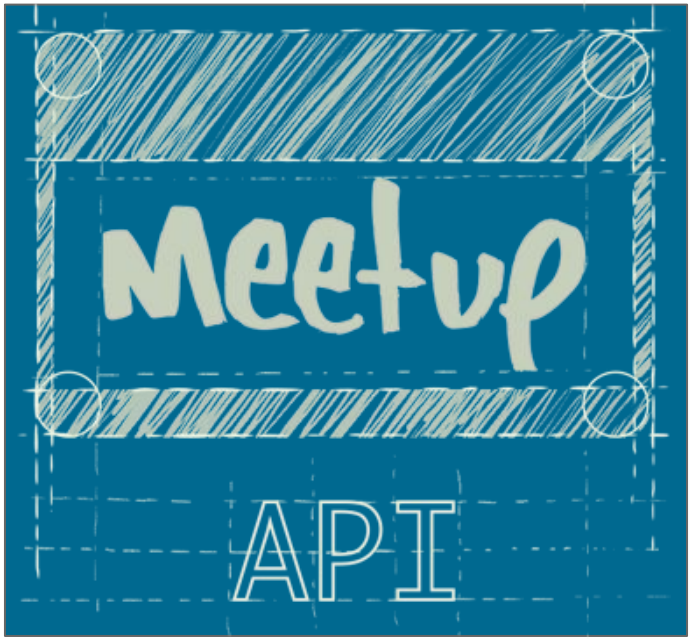
What recommendations can we make?

- Several different types
  - groups to join
  - topics to follow
  - events to attend

As a user of meetup.com trying to

find groups to join and events to attend

The data



meetup.com/meetup_api/

Find a Meetup Group Start a Meetup Group Invite Friends to Join

Messages Notifications

neo4j London User Group

Home Members Sponsors Photos Pages Discussions More Group tools My profile

neo4j

London, United Kingdom
Founded Jun 1, 2011

About us...

Invite friends

People	2,823
Group reviews	39
Upcoming Meetups	7
Past Meetups	154
Our calendar	

edit

Organizers:
Neo4j, Mark Needham

Contact

We're about:
Data Mining · New Technology · Web Development · Data Visualization · Data Analytics · Open Source

Graphs are everywhere. Use them.

+ SCHEDULE A NEW MEETUP

Upcoming 7 Past Drafts 3 Calendar

FEATURED MEETUP

Beyond SQL: Managing events & relationships in Social Care with Neo4j

Skillsmatter (CodeNode)
CodeNode, 10 South Place [illegible] EC2M 2RB, GB , London (map)

Wed Apr 13
6:30 PM
I'M GOING
41 going
1 comment

This month we'll have Andre[illegible]yson from Capita looking at how Ne[illegible]d other NoSQL technologies are [illegible]g used to build a new Social Care syste[illegible]ased upon a modern cloud-ready software architecture. The talk will include a l[illegible]at the graph model used to enable the mapping of person-centric events and relationships, as well as the... LEARN MORE

Hosted by: Mark Needham (Co-Organizer)

Cypher for SQL Developers

Neo Technology
5 - 11 Lavington Street, London, SE10NZ, London (map)

Tue Apr 5
6:30 PM
I'M GOING
3 days left
48 going
0 comments

This session is an introduction to graphs for SQL developers. We'll start with an overview of Neo4j and its query language Cypher before moving on to a worked example... LEARN MORE

Hosted by: Mark Needham (Co-Organizer), and Rik

What's new



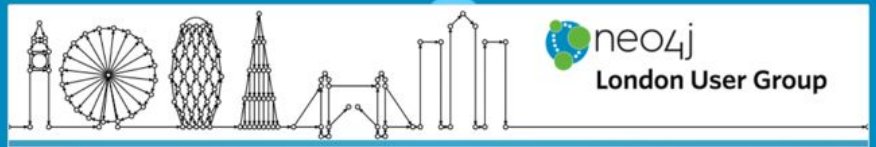
MORE

NEW MEMBER Mads Jensen joined
2 days ago

NEW MEMBER Laurie Clark-Michalek joined
2 days ago

NEW COMMENT Mark Needham commented on Real-Time Fraud Detection with Graphs - Dr Jim Webber, Chief Scientist at Neo4j
2 days ago

NEW MEMBER Gareth Chillingworth



neo4j London User Group

neo4j

London, United Kingdom
Founded Jun 1, 2011

About us...

Invite friends

People 2,823
Group reviews 39
Upcoming Meetups 7
Past Meetups 154
Our calendar

edit

Organizers:
Neo4j, Mark Needham

Contact

We're about:
Data Mining · New Technology · Web Development · Data Visualization · Data Analytics · Open Source

Graphs are everywhere. Use them.

+ SCHEDULE A NEW MEETUP

Upcoming 7 | Past | Drafts 3 | Calendar

FEATURED MEETUP

Beyond SQL: Managing events & relationships in Social Care with Neo4j

Skillsmatter (CodeNode)
CodeNode, 10 South Place, London, EC2M 2RB, GB , London (map)

Wed Apr 13
6:30 PM
I'M GOING
41 going
1 comment

This month we'll have Andrew Bryson from Capita looking at how Neo4j and other NoSQL technologies are being used to build a new Social Care system based upon a modern cloud-ready software architecture. The talk will include a look at the graph model used to enable the mapping of person-centric events and relationships, as well as the... LEARN MORE

Hosted by: Mark Needham (Co-Organizer)

Cypher for SQL Developers

Neo Technology
5 - 11 Lavington Street, London, SE10NZ, London (map)

Tue Apr 5
6:30 PM
I'M GOING
3 days left
48 going
0 comments

This session is an introduction to graphs for SQL developers. We'll start with an overview of Neo4j and its query language Cypher before moving on to a worked example...
LEARN MORE

Hosted by: Mark Needham (Co-Organizer), and Rik

What's new

MORE

NEW MEMBER Mads Jensen joined
2 days ago

NEW MEMBER Laurie Clark-Michalek joined
2 days ago

NEW COMMENT Mark Needham commented on Real-Time Fraud Detection with Graphs - Dr Jim Webber, Chief Scientist at Neo4j
2 days ago

NEW MEMBER Gareth Chillingworth

What data do we have?

- Groups
- Members
- Events
- Topics
- Time & Date
- Location

First Recommendation: Find similar groups

As a member of <my-favourite-group>

I want to find other similar meetup groups

So that I can join those groups

What makes groups similar?

neo4j

We're about:
Data Mining · New Technology · Web Development · Data Visualization · Data Analytics · Open Source · Cloud Computing · Graph Databases · Big Data · NoSQL · Neo4j · Database Development · Java · Computer programming

We're about:
Open Source · Technology · Web Development · Computer programming · Agile Project Management · Java · Software Development

Deep Learning London

We're about:
Big Data Analytics · Artificial Intelligence · Computer programming · Big Data · Computer Science · Natural Language Processing · Machine Learning · Data Analytics · Data Visualization · Data Mining · Data Science · Algorithms · Deep Learning · neural networks

Data Science London

We're about:
Data Science · Machine Learning · Predictive Analytics · Data Mining · Big Data · Artificial Intelligence · Statistical Computing · Applied Statistics · Data Analytics · Open Source · Web Analytics · Text Analytics · Natural Language Processing · Hadoop · NoSQL

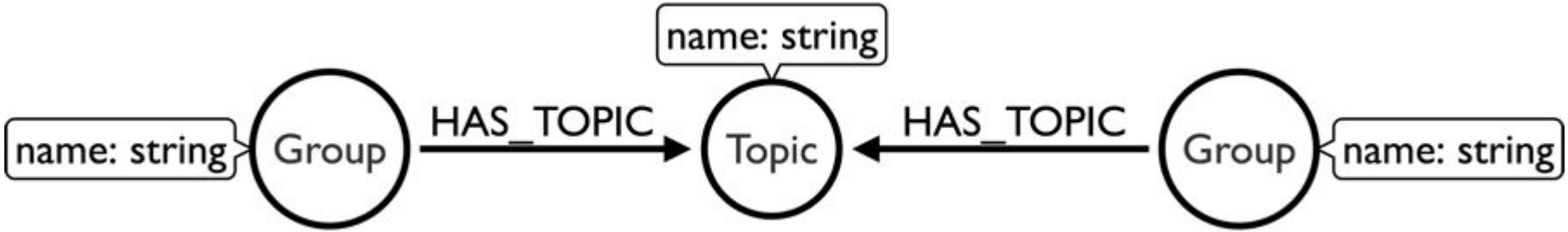
BIG DATA DEBATE

We're about:
BigData · Intellectual Discussion · Big Data · Debate · New Technology · Data Analytics · Data Visualization · Online Marketing · Database Development · Information Architecture · Information Science · Freedom · Cloud Computing · Business Strategy · Internet Professionals

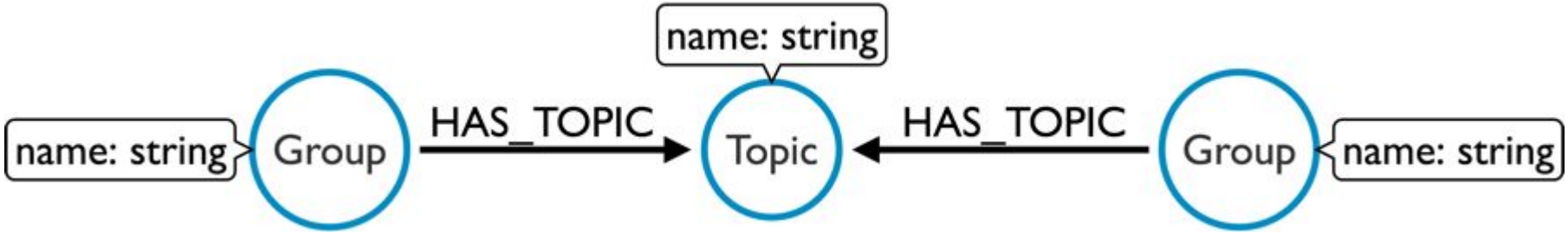
extract

We're about:
Marketing · Data · Software Development · New Technology · Web Technology · Business Intelligence · Cloud Computing · Business Strategy · Big Data · Machine Learning · Data Analytics · Data Visualization · Data Mining · Data Science · Big Data Analytics

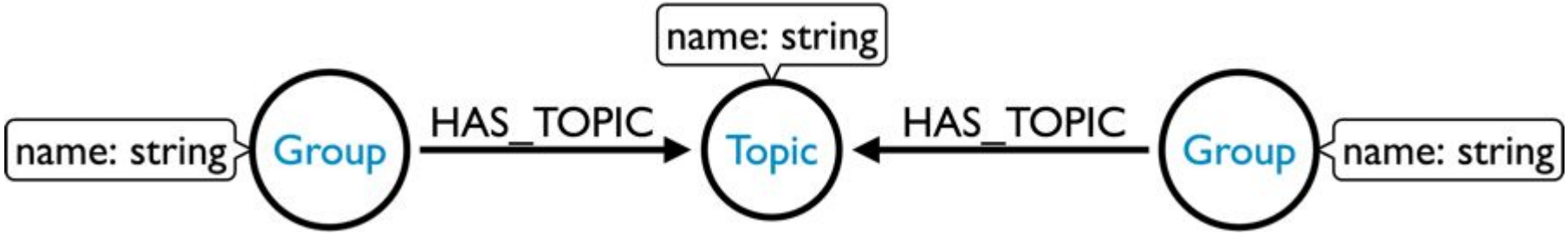
Labeled property graph



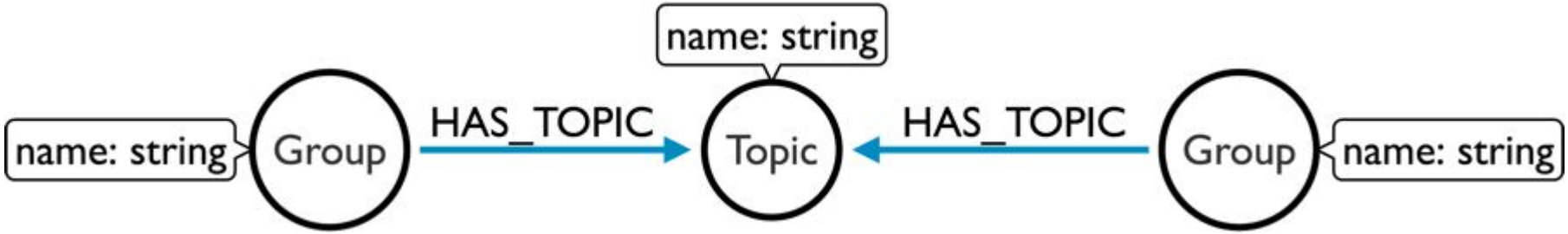
Nodes


name: string
name: string
Group
HAS_TOPIC
Topic
HAS_TOPIC
Group
name: string

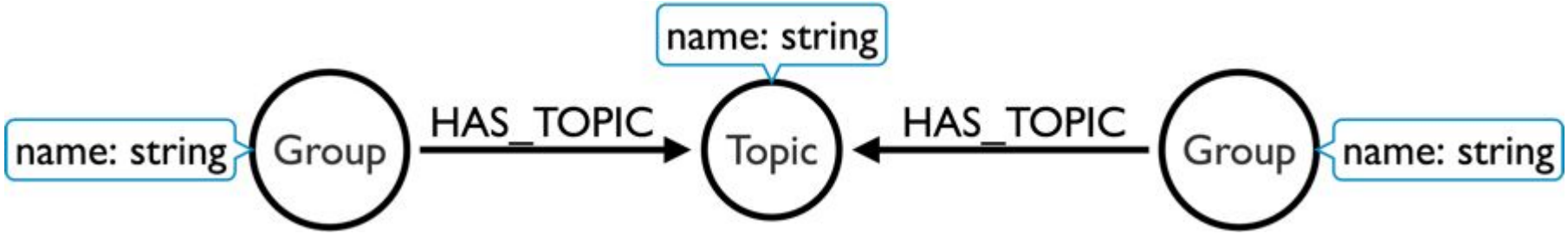
Labels


name: string
name: string
Group
HAS_TOPIC
Topic
HAS_TOPIC
Group
name: string

Relationships



Properties


name: string
name: string
Group
HAS_TOPIC
Topic
HAS_TOPIC
Group
name: string

Prepare your Neo4j graph.db directory

Copy folders

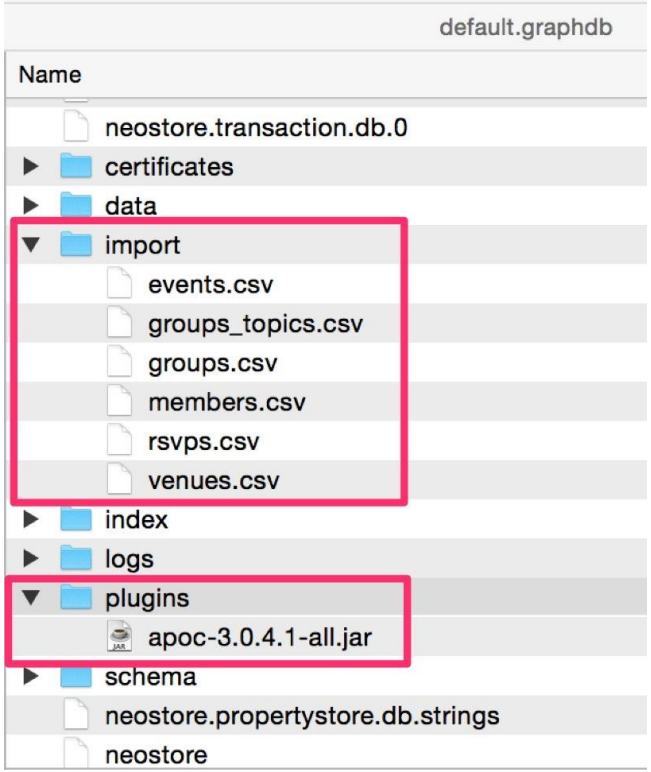
* `import`
* `plugins`

from **USB Stick**

to the `default.graphdb`

folder

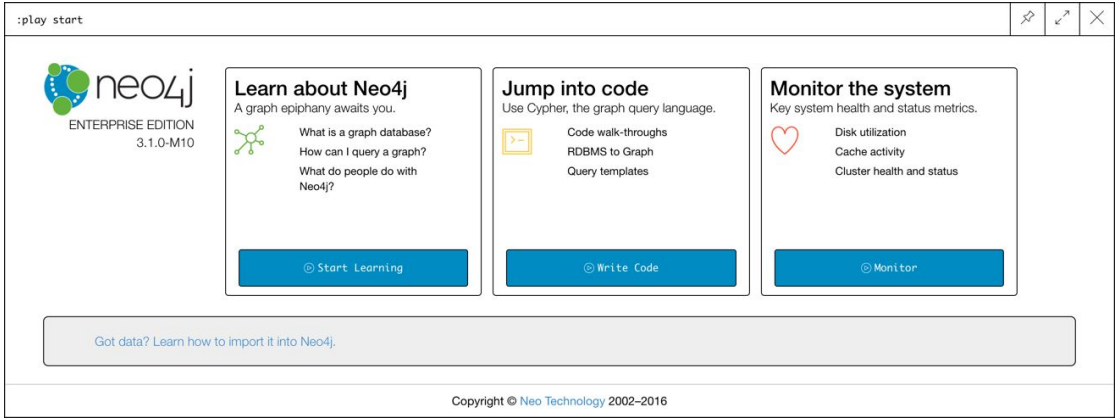
(or `$NEO4J_HOME`)



Make sure you've got Neo4j running

1. Start the server.
2. It should be running on: http://localhost:7474
3. Log-in with default credentials
   - user: **neo4j** password: **neo4j**
4. Choose a **new** password

We're good to go!



Recommend groups by topic

Open your browser to http://localhost:7474

and execute the following command:

```
:play http://guides.neo4j.com/reco/file
```

Make sure to follow the setup instructions for the CSV files and import directory.

Then start the guide for this module `Recommend Groups by Topic`

Follow the guides in your browser until you see...

Constraints and Indexes

MERGE

MERGE is used to uniquely create graph structures.

It tries to find the provided pattern in the graph and creates it if not found.

For **unique nodes** it uses constraints for quick lookup and locking.

```
MERGE (node:Label {key:value})
ON CREATE SET node.property = 'some value'
```

It can also be used for **unique relationships** when both nodes are known (bound).

Unique Constraints

We create unique constraints to:

- ensure uniqueness
- allow fast lookup of nodes which match label-property pairs.

Unique Constraints

We create unique constraints to:

- ensure uniqueness
- allow fast lookup of nodes which match label-property pairs.

```
CREATE CONSTRAINT ON (label:Label)
ASSERT label.property IS UNIQUE
```

Indexes

We create indexes to:

- allow fast lookup of nodes which match label-property pairs.

Indexes

We create indexes to:

- allow **fast lookup** of nodes which match label-property pairs.

```
CREATE INDEX ON :Label(property)
```

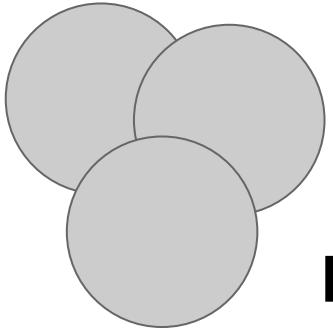
What are these fast lookups?

The following predicates use indexes:

- Equality
- `STARTS WITH`
- `CONTAINS`
- `ENDS WITH`
- Range searches
- (Non-) existence checks

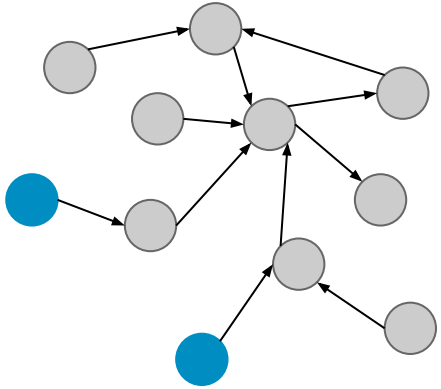
How are indexes used in neo4j?

Indexes are **only** used to find the **starting points** for queries.



Relational

Use index scans to look up rows in tables and join them with rows from other tables



Graph

Use indexes to find the starting points for a query.

Continue with the guide

Continue with the guide in your browser

Answers

Answers

Type the following command into the Neo4j browser to see the answers:

```
:play http://guides.neo4j.com/reco/answers/1.html
```

Continue with the guide

Continue with the guide in your browser

End of Module
The Meetup Dataset

Questions ?

LOAD CSV

```
[USING PERIODIC COMMIT] // optionally batch transactions
LOAD CSV      // load csv data
WITH HEADERS // optionally use first header row as keys in "row" map
FROM "url"    // file:// URL relative to $NEO4J_HOME/import or http://
AS row        // return each row of the CSV as list of strings or map
[FIELDTERMINATOR ";"] // optionally alt. delimiter
// ... rest of the Cypher statement ...
```