Programming and frameworks for ML

NoSQL Additional Resources

About Me

Big Data Consultant at Indra / Big Data Lecturer

- More than 20 years of experience in different environments, technologies, customers, countries ...
- Passionate data and technology
- Enthusiastic Big Data world and NoSQL

Daniel Villanueva Jiménez

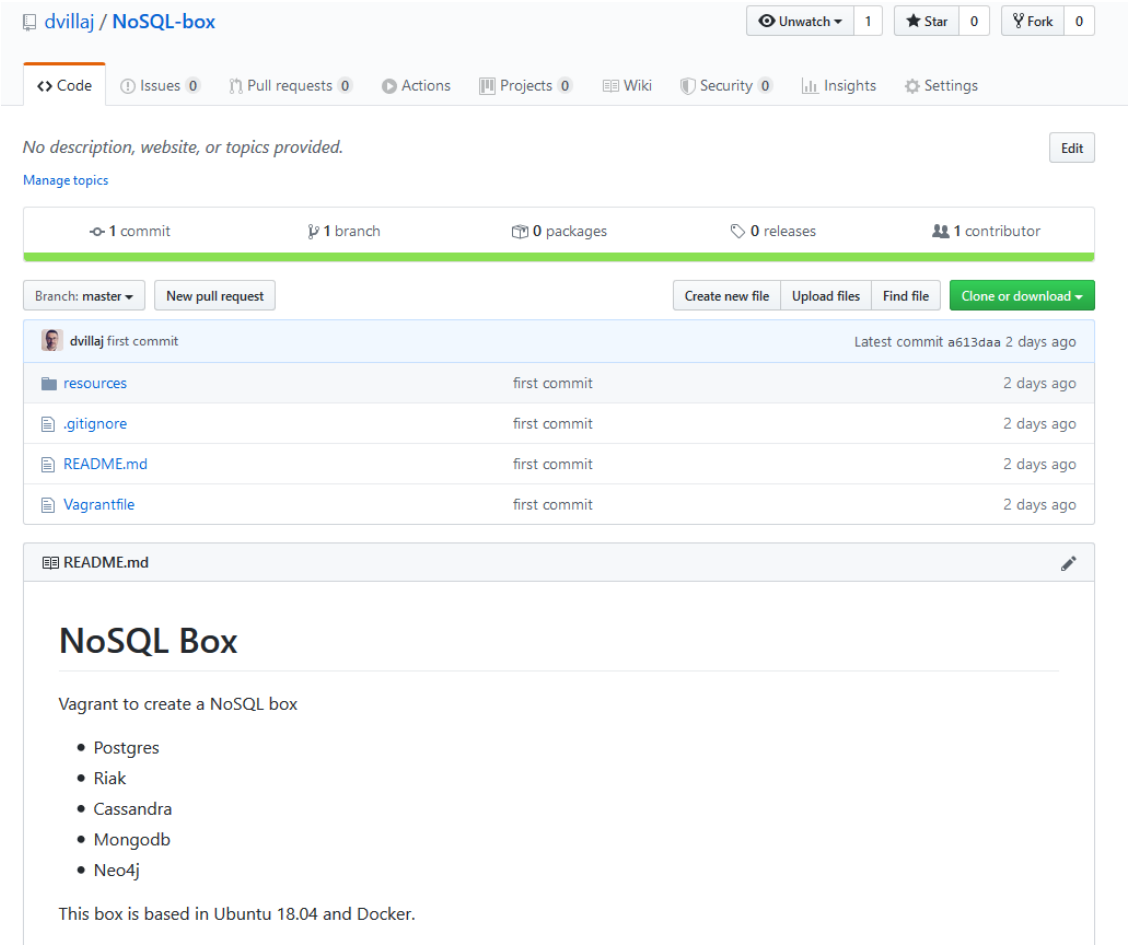
BigData Developer / Lecturer

INDRA • Universidad Pontificia de Salamanca

Agenda

- Infrastructure
- PostgreSQL
- Riak
- Cassandra
- MongoDB
- Neo4j
- Bibliography

NoSQL Virtual Machine



https://github.com/dvillaj/nosql-box

Docker

https://www.youtube.com/watch?v=1LRzlUoyZg4

Agenda

- Infrastructure
- PostgreSQL
- Riak
- Cassandra
- MongoDB
- Neo4j
- Bibliography

SQL Cheat Sheet

SQL Facts

- SQL stands for Structured Query Language
- SQL is pronounced "sequel"
- SQL is declarative language
- SQL is used to access & manipulate data in databases
- Top SQL DBs are MS SQL Server, Oracle, DB2, and MySQL

Database Definitions

- **RDBMS** (Relational Database Management System) – Software that stores and manipulates data arranged in relational database tables.
- **Table** – A set of data arranged in columns and rows. The columns represent characteristics of stored data and the rows represent actual data entries.

SQL Commands Categories

Data Query Language (DQL)

- SELECT - Retrieve data from table(s)

Data Manipulation Language (DML)

- INSERT - Insert data into db table
- UPDATE - Update data in db table
- DELETE - Delete data from table

Data Definition Language (DDL)

- CREATE - Create db object (table, view, etc.)
- ALTER - Modify db object (table, view, etc.)
- DROP - Delete db object (table, view, etc.)

Data Control Language (DCL)

- GRANT - Assign privilege
- REVOKE - remove privilege

How to select data from a table

SELECT <Column List>
FROM <Table Name>
WHERE <Search Condition>

Example:

```
SELECT FirstName, LastName, OrderDate
FROM Orders WHERE OrderDate > '10/10/2010'
```

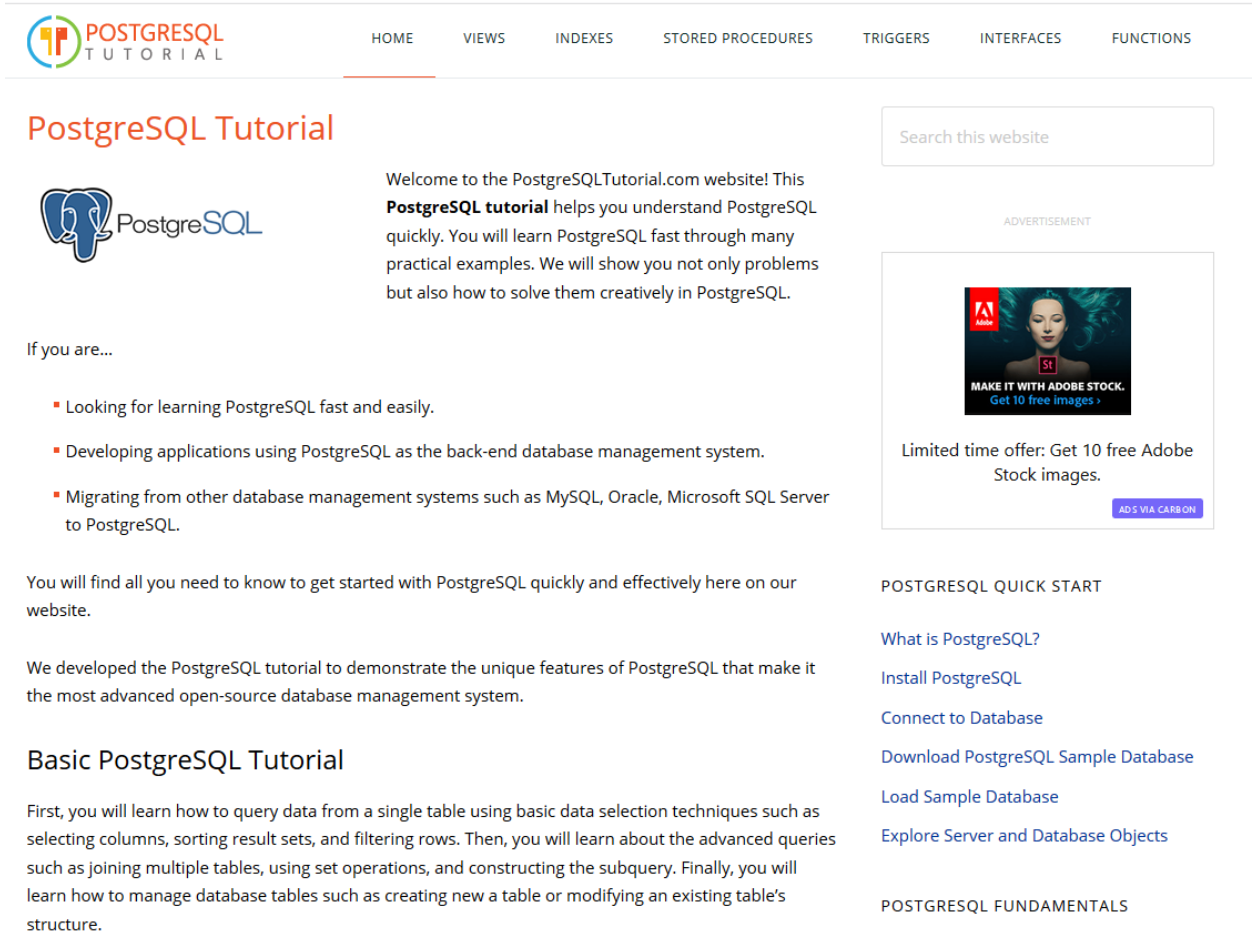
How to insert data in a table

INSERT INTO <Table Name>
(<Column List>) VALUES (<Values>)

Example:

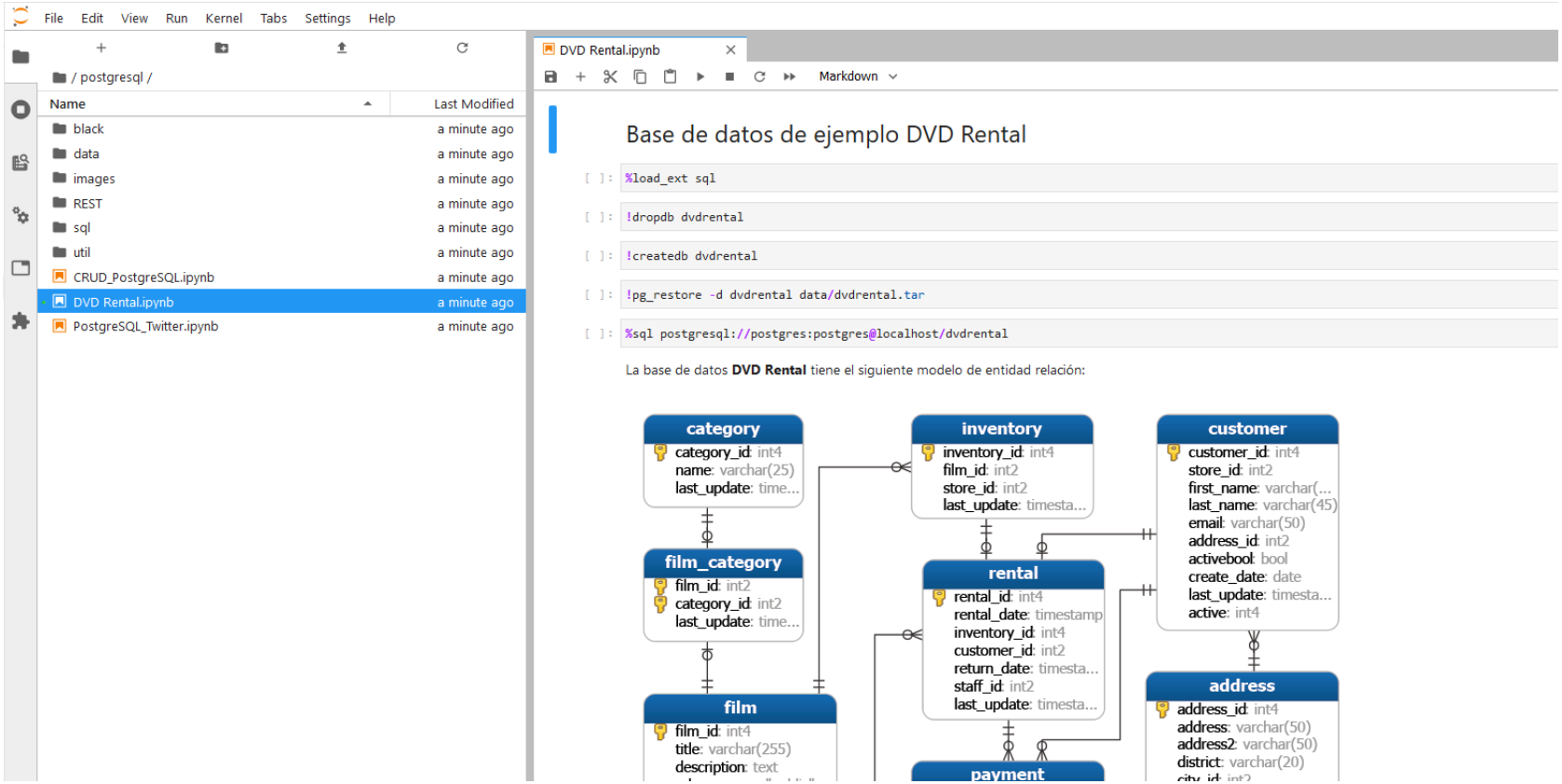
```
INSERT INTO Orders
(FirstName, LastName, OrderDate) VALUES
('John', 'Smith', '10/10/2010')
```

PostgreSQL tutorial



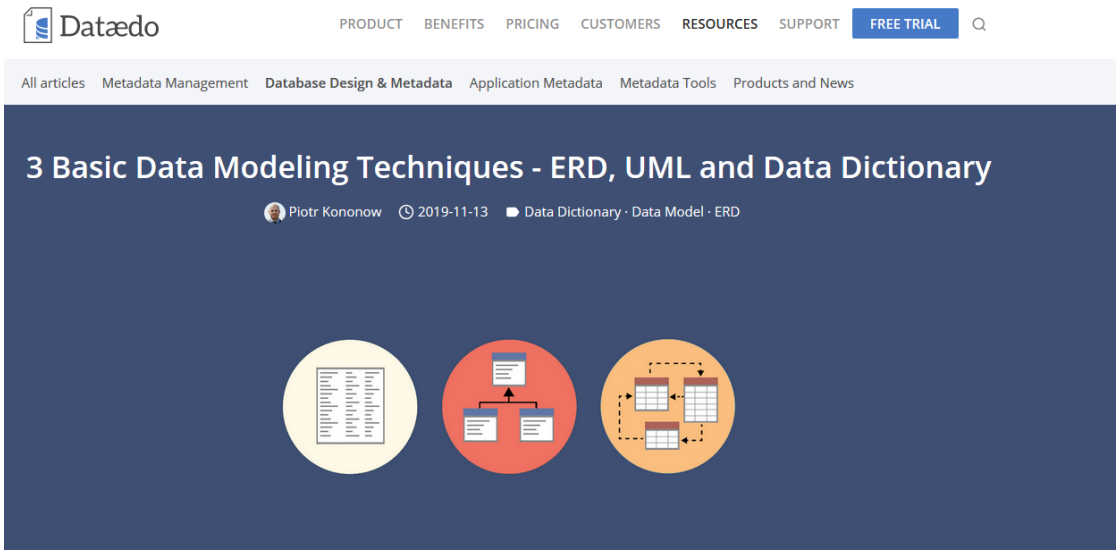
https://www.youtube.com/watch?v=1LRzlUoyZg4

PostgreSQL tutorial Sample Database



/postgresql/DVD Rental.ipynb

Data Modeling Techniques



In this article I will give you a brief overview of 3 basic data modeling techniques - ER Diagrams, UML Class Diagrams and a Data Dictionary.

Table of Contents:

1. Entity Relationship Diagrams
2. UML Class Diagrams
3. Data Dictionary

1. Entity Relationship Diagrams

Also referred to as **ER diagrams** or **ERDs**. Entity-Relationship modeling is a default technique for modeling and the design of relational (traditional) databases. In this notation architect identifies:

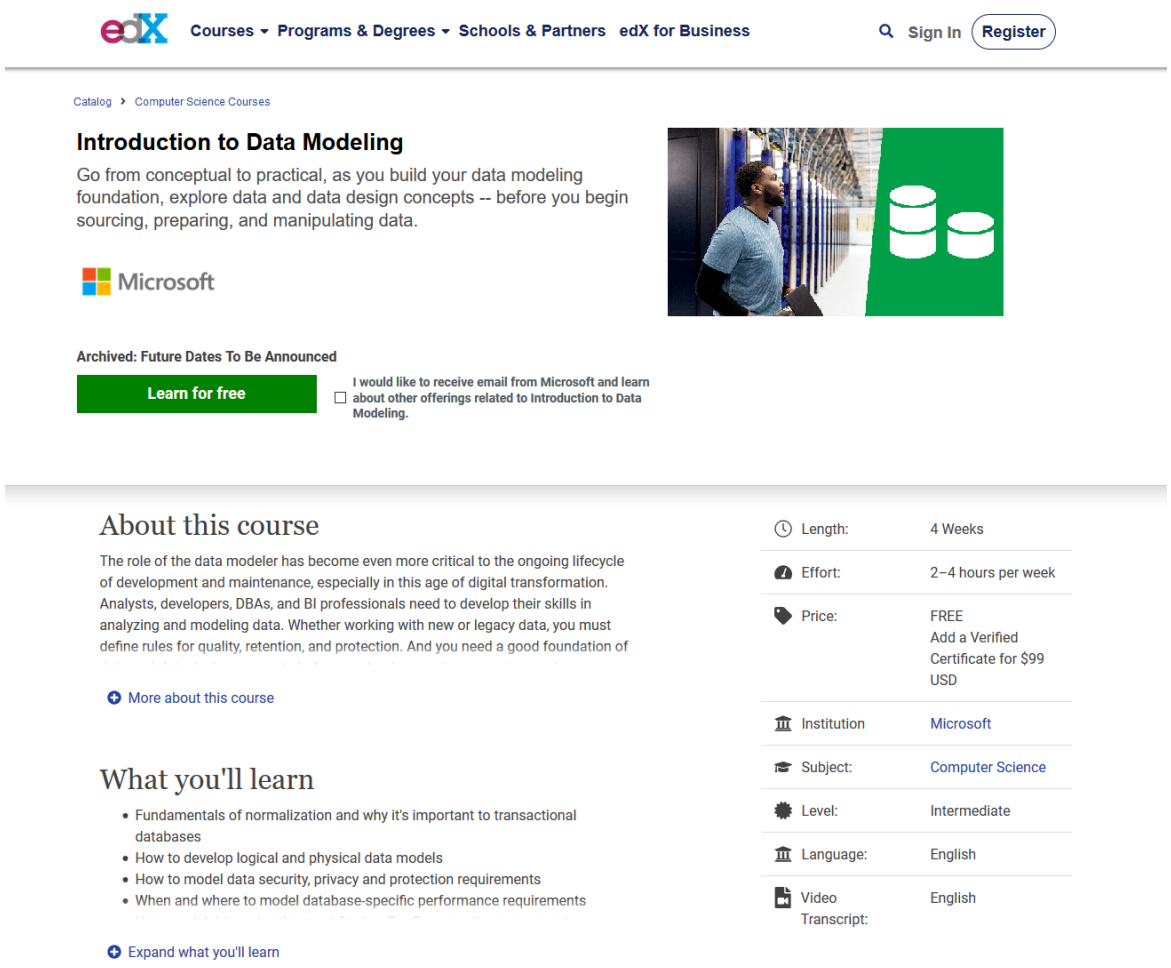
1. **Entities** representing objects (or tables in relational database),
2. **Attributes** of entities including **data type**,
3. **Relationships** between entities/objects (or foreign keys in a database).

ERDs work well if you want to design a relational (classic) database, Excel databases or CSV files. Basically, any kind of tabular data. They work well for visualization of database schemas and communication of top-level view of data.

Generate database docs
Generate convenient documentation of your databases in minutes and share it with your team.
See live sample

http://www.sql-tutorial.net/sql-cheat-sheet.pdf

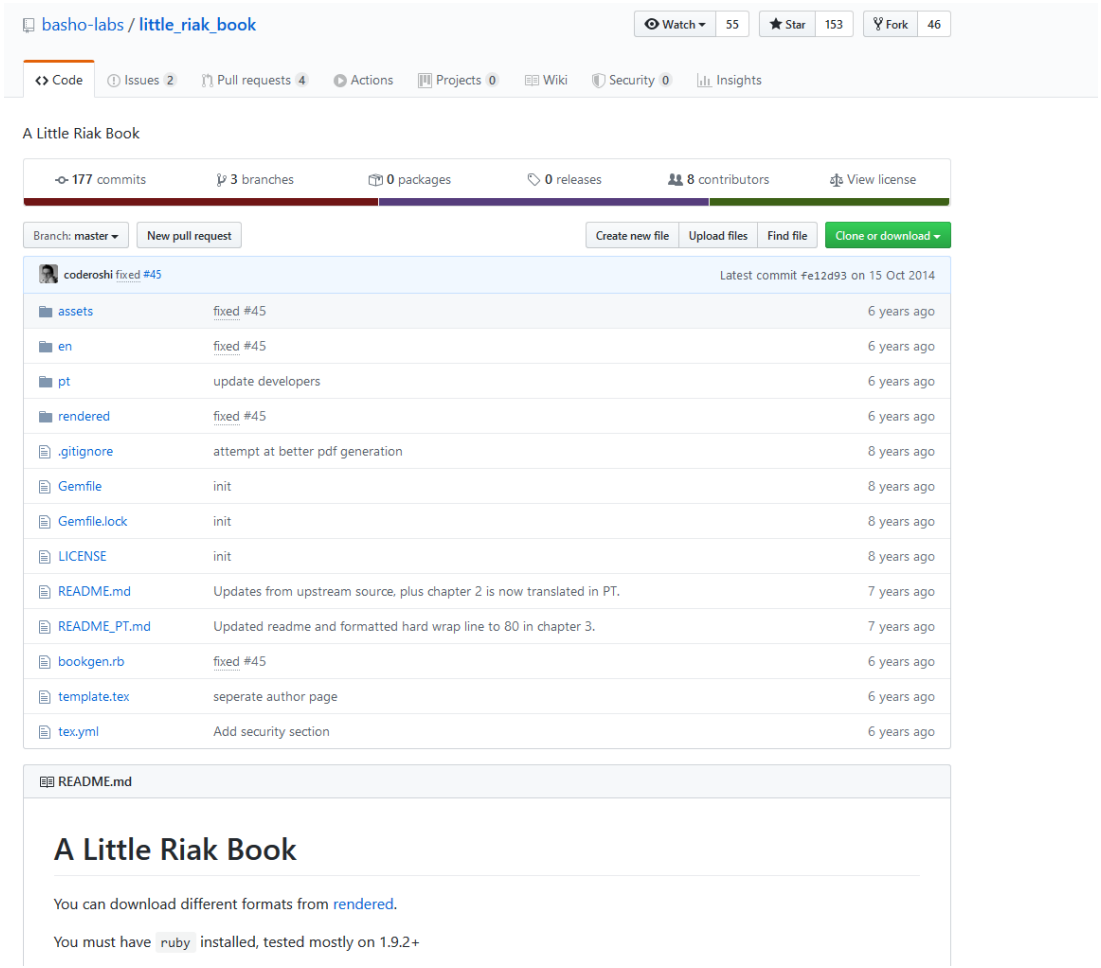
Data Modeling



Agenda

- Infrastructure
- PostgreSQL
- Riak
- Cassandra
- MongoDB
- Neo4j
- Bibliography

Little Riak a Book



- https://github.com/basho-labs/little_riak_book

Agenda

- Infrastructure
- PostgreSQL
- Riak
- Cassandra
- MongoDB
- Neo4j
- Bibliography

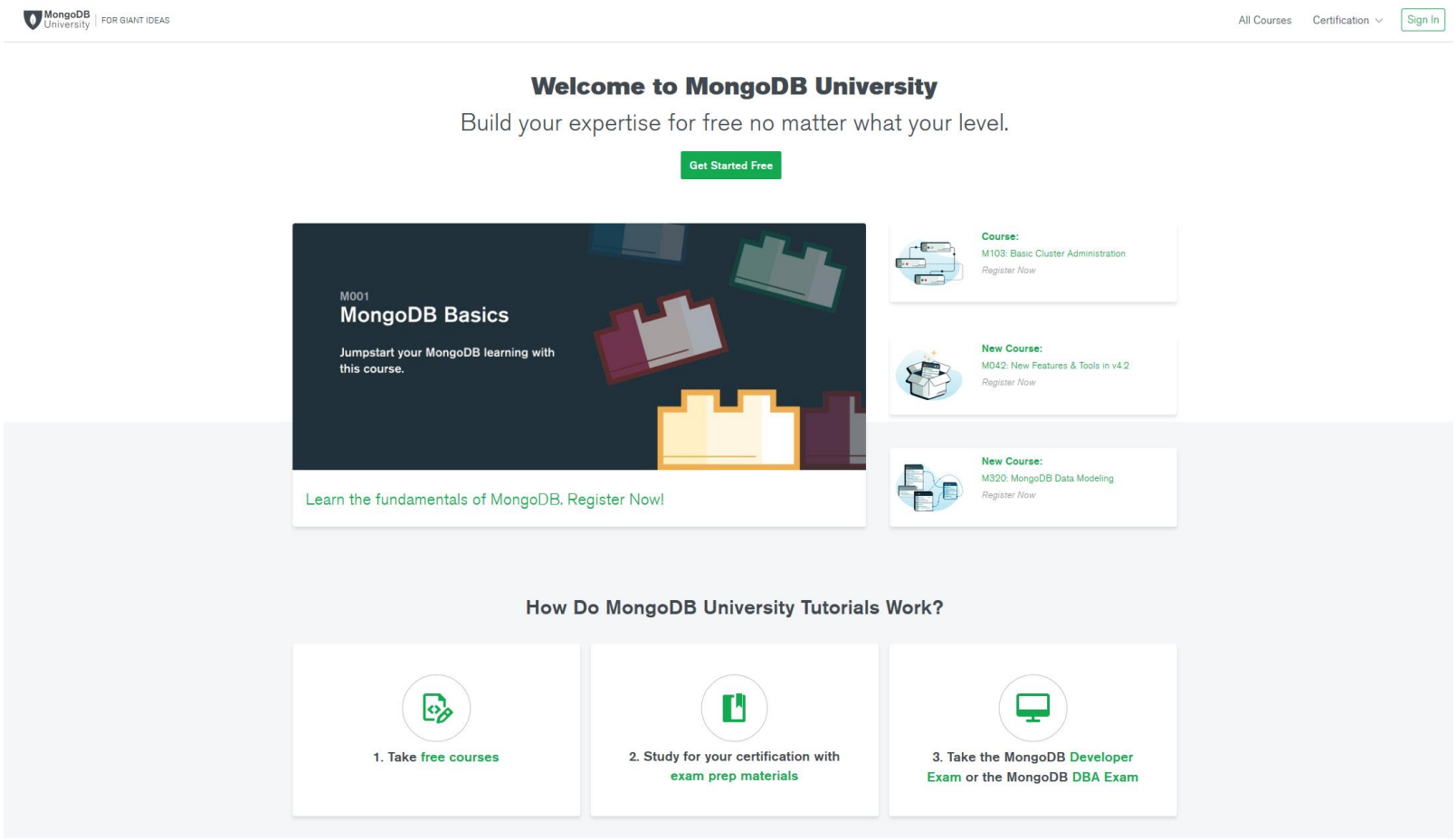
Cassandra Academy

- https://academy.datastax.com/

Agenda

- Infrastructure
- PostgreSQL
- Riak
- Cassandra
- MongoDB
- Neo4j
- Bibliography

MongoDB University

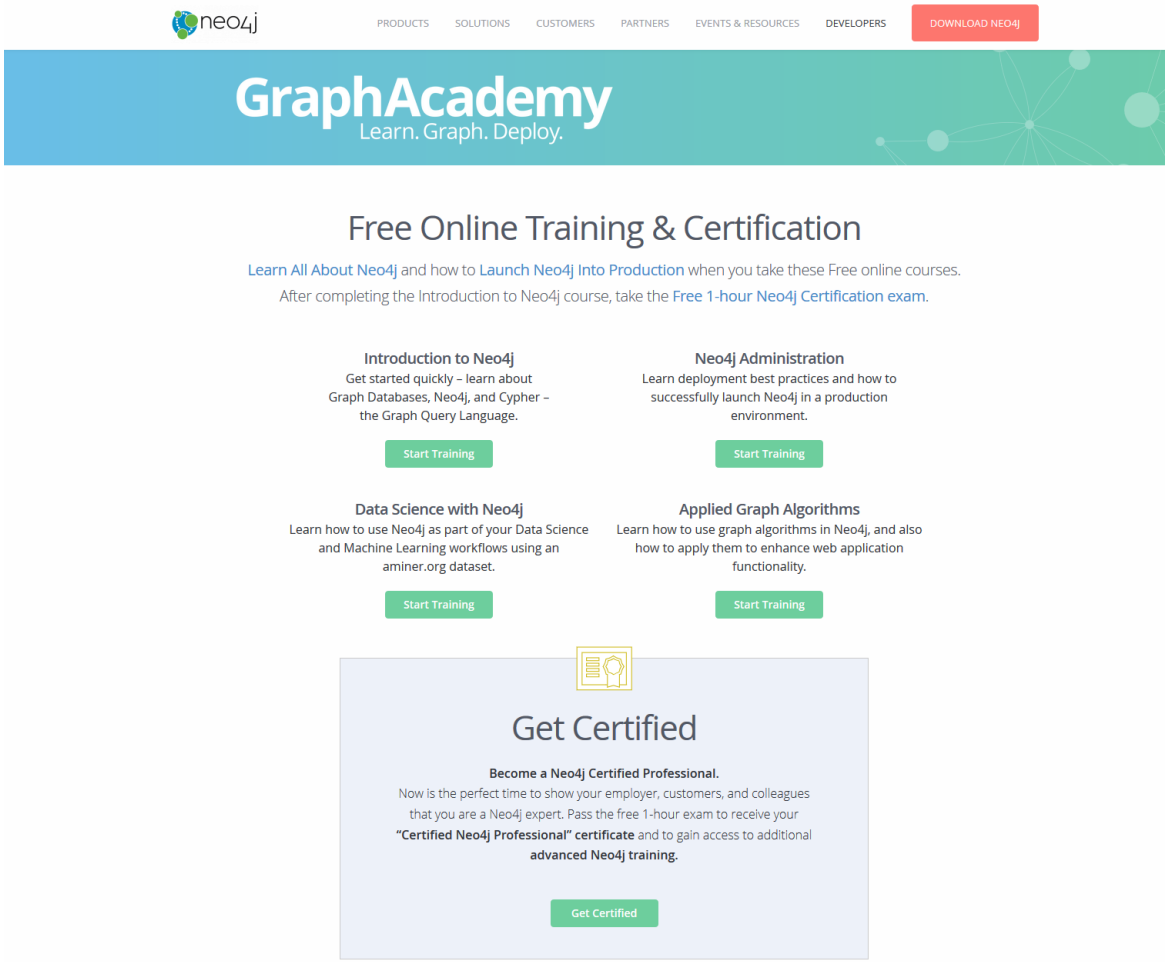


- https://university.mongodb.com/

Agenda

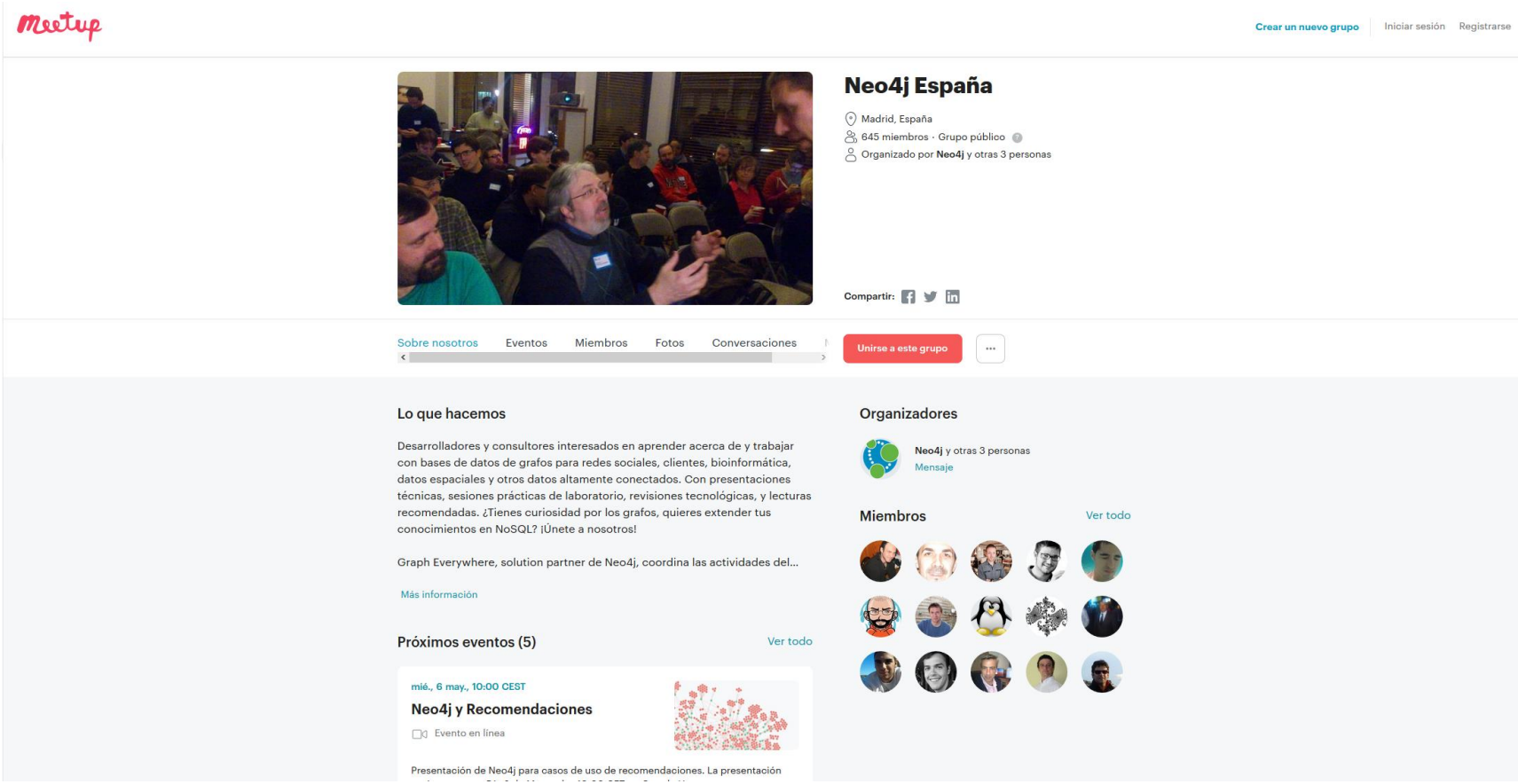
- Infrastructure
- PostgreSQL
- Riak
- Cassandra
- MongoDB
- Neo4j
- Bibliography

Neo4j Graph Academ¡``p0oiu



- https://neo4j.com/graphacademy/

Neo4j Meetup

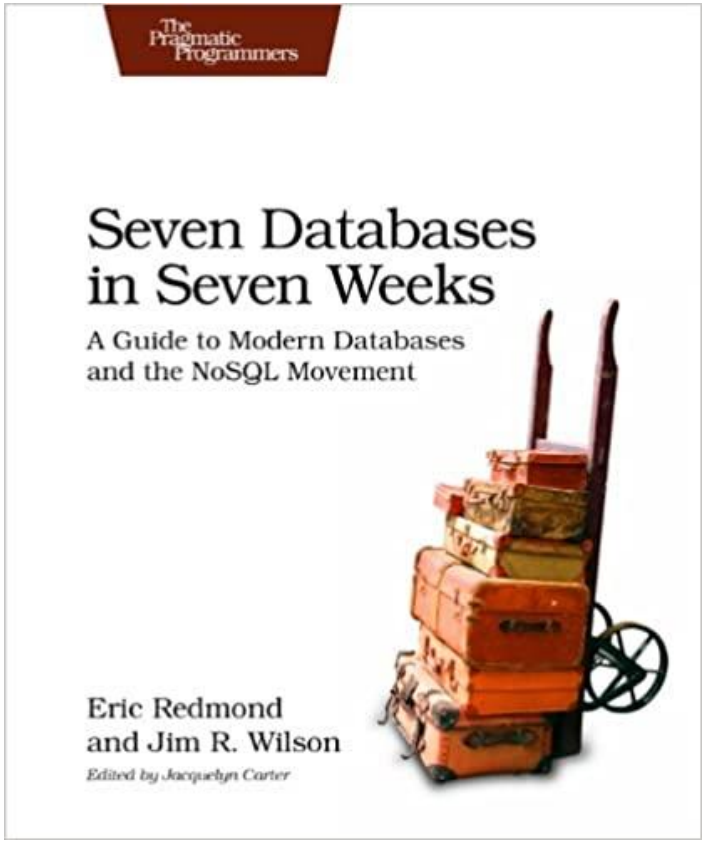


- https://www.meetup.com/es-ES/graphdb-spain/

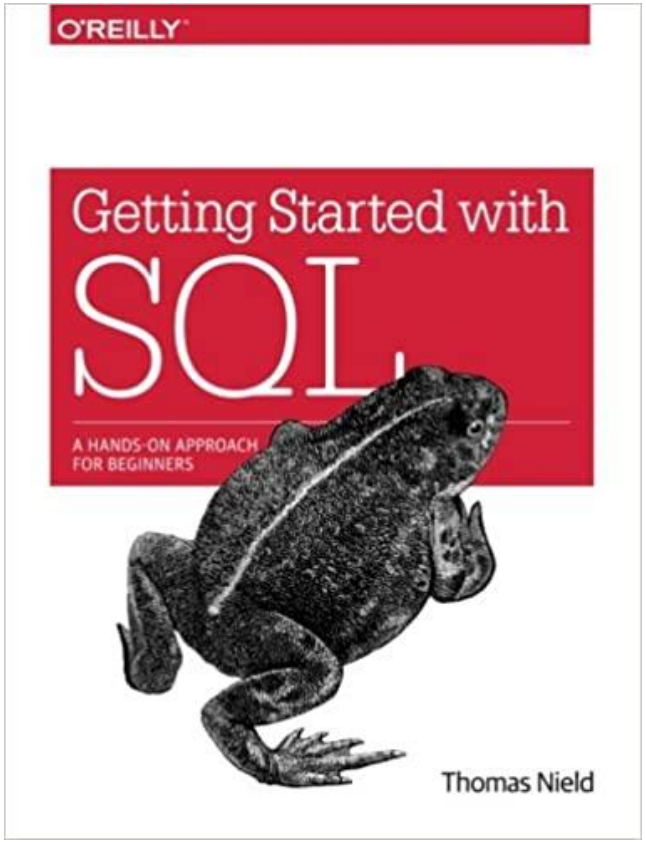
Agenda

- Infrastructure
- PostgreSQL
- Riak
- Cassandra
- MongoDB
- Neo4j
- Bibliography

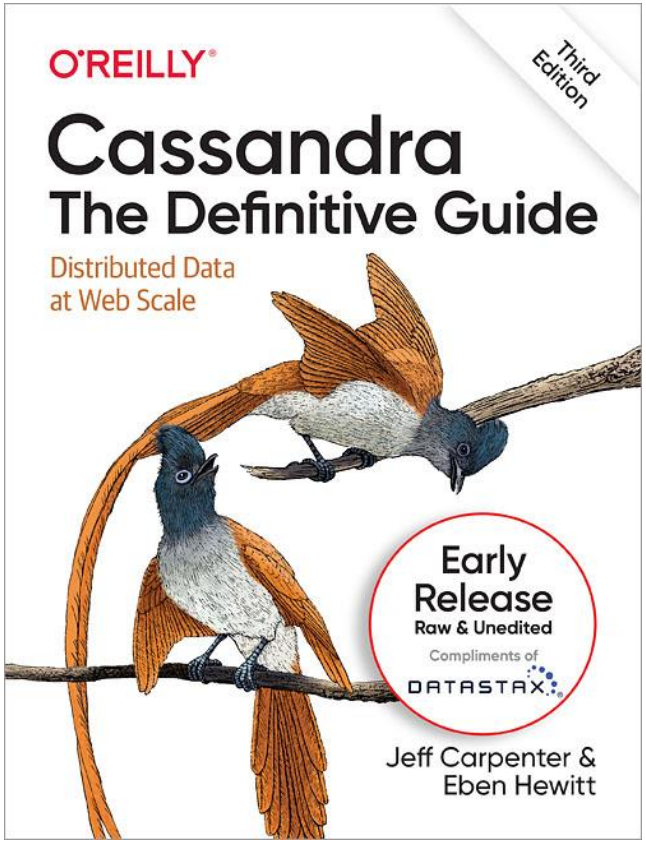
Bibliography



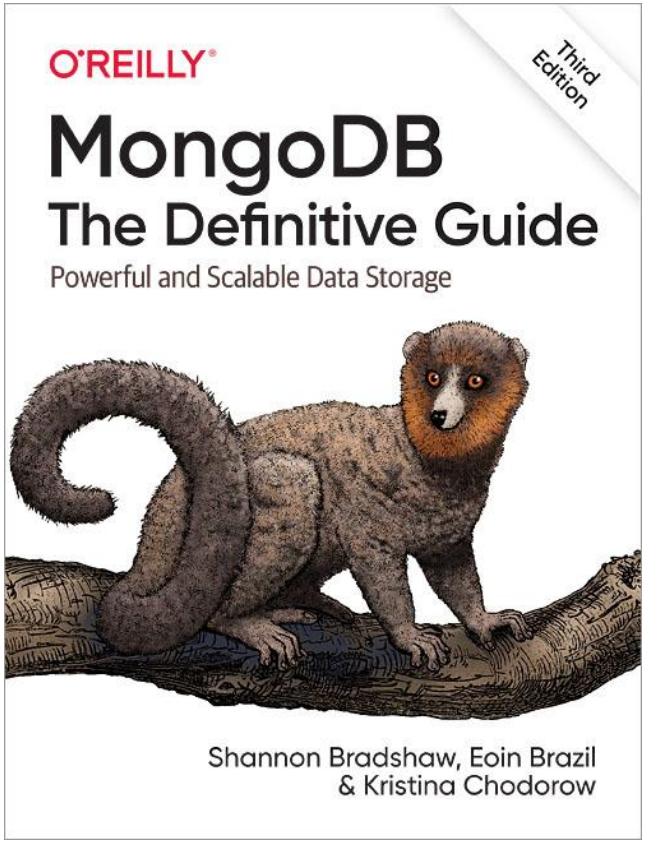
Bibliography



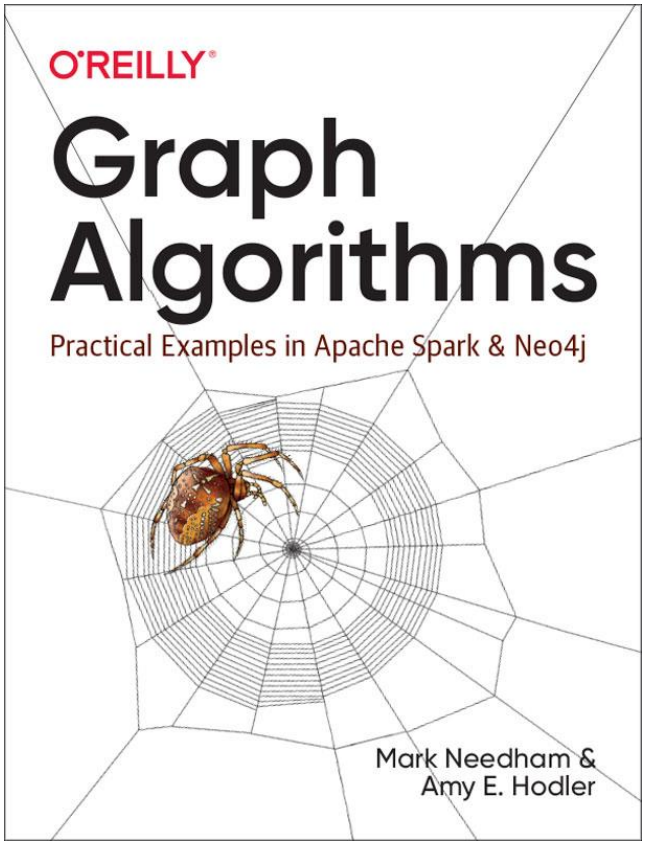
Bibliography



Bibliography



Bibliography



- https://neo4j.com/graph-algorithms-book/

THANKS FOR YOUR ATTENTION

Daniel Villanueva Jiménez
daniel.villanueva@immune.institute
@dvillaj