Cèlia Martínez

Marta Matute

Goretti Pena

Judit Ugas

Projecte Neo4j :  
Padrons

Bases de Dades no Relacionals

# Introducció

# Exercici 1: importació de dades

La importació de les dades per a crear el graf que farem servir per a les quèries es pot trobar al [GitHub](https://github.com/Marta-Matute/Projecte-Neo4j-Grup2), dins la carpeta “Code”. Hem optat per escriure tota la informació en anglès (noms de les propietats, relacions, etc). En total, la importació de dades està formada per 8 comandes executades una darrera l’altra:

Una comanda per a assegurar-nos que la base de dades està buida:

match (n) detach delete n;

Dues comandes per a carregar tots els nodes:

// LOAD PEOPLE DATA

LOAD CSV WITH HEADERS FROM 'https://docs.google.com/spreadsheets/d/e/2PACX-1vTfU6oJBZhmhzzkV\_0-avABPzHTdXy8851ySDbn2gq32WwaNmYxfiBtCGJGOZsMgCWjzlEGX4Zh1wqe/pub?output=csv' AS row

CREATE (p:PERSON {Id: toInteger(row.Id), Year: toInteger(row.Year), Name: row.name, Last\_Name: row.surname, Second\_Last\_Name: row.second\_surname});

// LOAD HOUSE DATA

LOAD CSV WITH HEADERS FROM 'https://docs.google.com/spreadsheets/d/e/2PACX-1vT0ZhR6BSO\_M72JEmxXKs6GLuOwxm\_Oy-0UruLJeX8\_R04KAcICuvrwn2OENQhtuvddU5RSJSclHRJf/pub?output=csv' AS row

CREATE (:HOUSE {Municipality: row.Municipi, Home\_Id: toInteger(row.Id\_Llar), Registry\_Year: toInteger(row.Any\_Padro), Street: row.Carrer, Number: toInteger(row.Numero)});

Dues comandes per a crear índexs en algunes de les propietats (això ens servirà per accelerar el temps de creació de les relacions i les quèries), una per a esborrar qualsevol índex anterior, i l’altre per a crear els que ens calen:

// DELTE ANY EXISTING INDEXES

CALL apoc.schema.assert({},{},true) YIELD label, key

RETURN \*;

// CREATE INDEXES (FOR FASTER DATA CREATION)

create index for (p:PERSON) on (p.Id);

create index for (p:PERSON) on (p.Year);

create index for (h:HOUSE) on (h.Home\_Id);

create index for (h:HOUSE) on (h.Municipality);

create index for (h:HOUSE) on (h.Registry\_Year);

I per últim tres comandes que creen les tres relacions entre nodes que necessitem:

// CREATION OF "LIVES" RELATIONSHIP BETWEEN A PERSON AND A HOUSE

LOAD CSV WITH HEADERS FROM 'https://docs.google.com/spreadsheets/d/e/2PACX-1vRM4DPeqFmv7w6kLH5msNk6\_Hdh1wuExRirgysZKO\_Q70L21MKBkDISIyjvdm8shVixl5Tcw\_5zCfdg/pub?output=csv' AS row

MATCH (p:PERSON {Id: toInteger(row.IND), Year: toInteger(row.Year)})

MATCH (h:HOUSE {Municipality: row.Location, Home\_Id: toInteger(row.HOUSE\_ID), Registry\_Year: toInteger(row.Year)})

MERGE (p)-[:LIVES {Year: toInteger(row.Year)}]->(h);

// CREATION OF "SAME\_AS" BETWEEN TWO PEOPLE

LOAD CSV WITH HEADERS FROM 'https://docs.google.com/spreadsheets/d/e/2PACX-1vTgC8TBmdXhjUOPKJxyiZSpetPYjaRC34gmxHj6H2AWvXTGbg7MLKVdJnwuh5bIeer7WLUi0OigI6wc/pub?output=csv' AS row

MATCH (p:PERSON {Id: toInteger(row.Id\_A)})

MATCH (q:PERSON {Id: toInteger(row.Id\_B)})

MERGE (p)-[:SAME\_AS]->(q);

// CREATION OF "FAMILY" RELATIONSHIP BETWEEN TWO PEOPLE

LOAD CSV WITH HEADERS FROM 'https://docs.google.com/spreadsheets/d/e/2PACX-1vRVOoMAMoxHiGboTjCIHo2yT30CCWgVHgocGnVJxiCTgyurtmqCfAFahHajobVzwXFLwhqajz1fqA8d/pub?output=csv' AS row

MATCH (p:PERSON {Id: toInteger(row.ID\_1)})

MATCH (q:PERSON {Id: toInteger(row.ID\_2)})

MERGE (p)-[rel:FAMILY {Relation: row.Relacio, Harmonized\_Relation: row.Relacio\_Harmonitzada}]->(q);

# Exercici 2: Queries

Un cop s’han carregat les dades, l’exercici 2 consistia en resoldre 10 queries:

1. Dels padró de 1866 de Castellví de Rosanes (CR), retorna el número d'habitants i la llista de noms. Elimina duplicats i nan.

Graphical user interface, text, application, email

Description automatically generated

1. Dels padrons de Sant Feliu de Llobregat (SFLL) d'abans de l'any 1840 (no inclòs), retorna la població, l'any del padró i la llista d'identificadors dels habitatges de cada padró. Ordena els resultats per l'any de padró.

Graphical user interface, text, application, email

Description automatically generated

1. Retorna el nom de les persones que vivien al mateix habitatge que "rafel marti" (no té segon cognom) segons el padró de 1838 de Sant Feliu de Llobregat (SFLL). Retorna la informació en mode graf i mode llista.

En mode de graf:

Diagram

Description automatically generated

I en mode de llista:

Graphical user interface, text, application, email

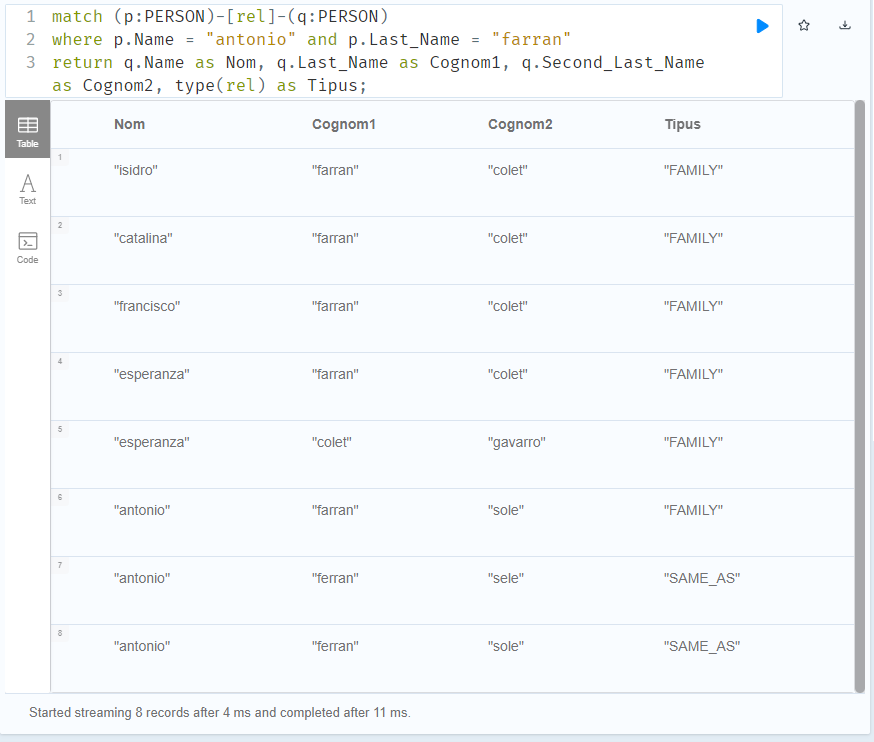
Description automatically generated

1. Retorna totes les aparicions de "Miguel ballester". Fes servir la relació SAME\_AS per poder retornar totes les instancies, independentment de si hi ha variacions lèxiques (ex. diferents formes d'escriure el seu nom/cognoms). Mostra la informació en forma de subgraf.

Diagram, schematic

Description automatically generated

1. Mostra totes les persones relacionades amb "antonio farran". Mostra la informació en forma de taula: el nom, cognom1, cognom2, i tipus de relació.



1. Graphical user interface, text, application, email

   Description automatically generatedLlisteu totes les relacions familiars que hi ha.
2. Identifiqueu els nodes que representen el mateix habitatge (carrer i numero) al llarg dels anys de Sant Feliu del Llobregat (SFLL). Mostreu el resultat dels habitatges que tingueu totes dues informacions (carrer i numero), el nombre total d’habitatges, el llistat d’anys dels padrons i el llistat de les Ids de les llars. Ordeneu de més a menys segons el total d’habitatges i mostreu-ne els 10 primers.

Graphical user interface, text

Description automatically generated

1. Mostreu les famílies de Castellví de Rosanes amb més de 3 fills. Mostreu el nom i cognoms del cap de família i el nombre de fills. Ordeneu-les pel nombre de fills fins a un límit de 20, de més a menys.

Graphical user interface

Description automatically generated with low confidence

1. Mitja de fills a Sant Feliu del Llobregat l’any 1881 per família. Mostreu el total de fills, el nombre d’habitatges i la mitja.

Graphical user interface, text, application, email

Description automatically generated

1. Per cada any que hi ha a la base de dades, quin és el carrer amb menys habitants de Sant Feliu de Llobregat?

Graphical user interface, text, application

Description automatically generated

# Exercici 3: Anàlisis del graf

* Explicació raonada de cada exercici
* Treball en equip