

Universitat Autònoma de Barcelona
Facultat de Ciències



PRÀCTICA NEO4J

Autors:

Manuel Arnau & Sofia Di Capua & Gerard Lahuerta & Ona Sánchez
1597487 — 1603685 — 1601350 — 1601181

18 de Juny de 2023

Contents

1	Introducció	3
2	Exercicis	4
2.1	Exercici 1	4
2.1.1	Importació i creació del node Habitatge	4
2.1.2	Importació i creació del node Individu	4
2.1.3	Importació i creació de la relació Família	5
2.1.4	Importació i creació de la relació SAME_AS	5
2.1.5	Importació i creació de la relació VIU	5
2.2	Exercici 2	7
2.3	Exercici 3	8
3	Treball en equip	9

1 Introducció

En aquest projecte es treballa el disseny, la implementació i la consulta a una base de dades en NEO4J. A partir dels requisits i les dades que s'han subministrat, s'ha implementat un script en Cypher que processa i insereix les dades en una base de dades de NEO4J. Seguidament, s'han implementat les consultes demanades per corroborar l'ús correcte de la base de dades.

L'objectiu d'aquest projecte és afermar els conceptes ensenyats a classe mitjançant una aplicació diversa i realista d'una base de dades.

Podeu veure tots els codis a la següent pàgina de github:

https://github.com/Gerard-Lahuerta/Projecte_NEO4J.git

2 Exercicis

2.1 Exercici 1

Per tal de poder treballar amb NEO4J s'ha hagut d'importar les dades que s'hi disposen en els fitxers (*FAMILIA*, *HABITATGES*, *INDIVIDUAL*, *SAME_AS* i *VIU*)¹ en format *CSV* al format Cypher.

Per tal de fer la correcta importació de les dades amb els tractaments de no duplicitat de dades, no importació de dades amb codis identificatius *null* i tipus de variables correctes, s'ha creat un document en Cypher encarregat de fer tots aquests processos: .

Mostrem a continuació el contingut del fitxer i els processos que fa.

2.1.1 Importació i creació del node Habitatge

```
LOAD CSV WITH HEADERS FROM
" https:// docs . google . com / spreadsheets / d / e / 2PACX-1vT0ZhR6BSO_
M72JEmxXKs6GLuOwxm_Oy-0UruLJeX8_R04KAeICuvrwn2OENQhtuvddU5RS
JSclHRJf/pub?output=csv " AS row

WITH row.Municipi AS Municipi ,
     toInteger(row.Id_Llar) AS LlarID ,
     toInteger(row.Any_Padro) AS YearPadro ,
     row.Carrer AS Carrer ,
     toInteger(row.Numero) AS Num

FOREACH ( ignoreMe in CASE WHEN LlarID is not null
  THEN [1]
  ELSE [] END
  | FOREACH (ignoreMe in CASE WHEN Municipi is not null
  THEN [1]
  ELSE [] END
  | MERGE (h:Habitatge {LlarID:LlarID})
    SET h.Municipi = Municipi ,
        h.YearPadro = YearPadro ,
        h.Carrer = Carrer ,
        h.Num = Num))

CREATE INDEX HabitatgeID FOR
(h:Habitatge) ON (h.LlarID , h.Municipi , h.YearPadro)

CREATE CONSTRAINT UniqueHabitatgeID FOR
(h:Habitatge) REQUIRE h.HabitatgeID IS UNIQUE
```

2.1.2 Importació i creació del node Individu

¹Aquets fitxer per a un desembolupament adient i escalable del treball en el nostre grup s'han decidit suministrar-los en el fitxer mitjançant un url a un directori drive on són descarгат quan es generen les dades al NEO4J.

```

LOAD CSV WITH HEADERS FROM
"https://docs.google.com/spreadsheets/d/e/
2PACX-1vTfU6oJBZhnhzzkV_0-avABPzHTdXy8851ySDbn2gq32WwaNmYxfiBtCGJGOZ
sMgCWjzIEGX4Zhlwqe/pub?output=csv" as row
WITH row
WHERE row.Id is not NULL
MERGE (i:Individu {IndividuID: row.Id})
SET i.Year = toInteger(row.Year), i.Name = row.name,
    i.Surname = row.surname, i.SecondSurname = row.second_surname

CREATE CONSTRAINT UniqueIndividuID FOR (i:Individu)
REQUIRE i.IndividuID IS UNIQUE;

```

2.1.3 Importació i creació de la relació Família

```

LOAD CSV WITH HEADERS FROM
"https://docs.google.com/spreadsheets/d/e/2PACX1vRVOoMAMoxHiGboT
jCIHo2yT30CCWgVHgocGnVJxiCTgyurtmqCfAFahHajobVzwXFLwhqajz1fqA8d/
pub?output=csv" AS row
WITH row.ID_1 AS Id_1, row.ID_2 AS Id_2, row.Relacio AS relacio ,
    row.Relacio_Harmonitzada AS relacio_harmonitzada
MATCH (p:Individu {IndividuID:Id_1})
MATCH (o:Individu {IndividuID:Id_2})
MERGE (o)-[rel:Relacio_Familiar {relacio: relacio ,
    relacio_harmonitzada:relacio_harmonitzada}]->(p)
RETURN count(rel);

```

2.1.4 Importació i creació de la relació SAME_AS

```

LOAD CSV WITH HEADERS FROM
"https://docs.google.com/spreadsheets/d/e/2PACX1vTgC8TBmdXhjUOPKJ
xyiZSpetPYjaRC34gmXHj6H2AWvXTGbg7MLKVdJnwh5bIeer7WLUi0OigI6wc/pub
?output=csv" AS row WITH row.Id_A AS Id_A, row.Id_B AS Id_B
MATCH (p:Individu {IndividuID:Id_A})
MATCH (o:Individu {IndividuID:Id_B})
MERGE (o)-[rel: SAME_AS]-(p)
RETURN count(rel);

```

2.1.5 Importació i creació de la relació VIU

```

LOAD CSV WITH HEADERS FROM
"https://docs.google.com/spreadsheets/d/e/2PACX1vRM4DPeqFmv7w6kLH5msNk6_Hdh
shVixl5Tcw_5zCfdg/pub?output=csv" AS vivencia
WITH vivencia
MATCH (p:Individu {IndividuID:vivencia.IND})
MATCH (h:Habitatge{LlarID:toInteger(vivencia.HOUSE_ID),
    Municipi:vivencia.Location,YearPadro:toInteger(vivencia.Year)})
MERGE (p) -[rel:VIU_A]- (h)

```

```
return count(rel)
```

2.2 Exercici 2

2.3 Exercici 3

3 Treball en equip

Per tal de treballar de forma més eficient s'ha decidit separar les diverses tasques entre el grup.

D'aquesta forma, cada individu s'ha encarregat d'una part.

Cada integrant del grup s'ha encarregat de fer els següents treballs:

- Ona Sánchez:
- Manuel Arnau:
- Sofia Di Capua:
- Gerard Lahuerta:

Cada integrant del grup es va encarregar de descriure en aquest informe la part que havia realitzat.

Comentar a més, que tot el treball ha sigut documentat i treballat mitjançant la plataforma *GitHub*; per la qual cosa es pot accedir a la tasca a través de [l'enllaç](#) on també es pot observar el desenvolupament cronològic de l'entrega.