

Enric Comuñas del Peral	1633263
Jordi Auñón Ansio	1636172
David Martí Felip	1633953
Joel Jumilla Fernández	1632465

## **Informe del projecte de MongoDB**

### **Exercici 1:**

#### **- Editorial:**

```
{_id: "Nom",  
  responsable: "Responsable",  
  adreça: "Adreça",  
  país: "País"}
```

Hem cregut adequat que no tingués cap atribut que fes referència a qualsevol altre col·lecció, és a dir, no hem aplicat cap patró de disseny sobre aquesta col·lecció. Això degut a què hem cregut més oportú fer el tractament de la relació entre les editorials i les col·leccions des de la col·lecció Col·lecció, ja que més d'una d'aquestes pot pertànyer a més d'una editorial. I hem cregut que, precisament, aquesta informació és més important que saber totes les col·leccions que té una editorial, és a dir, la referència complementària.

#### **- Col·lecció:**

```
{_id: "Nom",  
  total_exemplars: "Total d'exemplars",  
  generes: ["Gènere 1", ...],  
  idioma: "Idioma",  
  any_inici: "Any inici",  
  any_fi: "Any fi",  
  tancada: "Tancada?",  
  editorials: ["Nom editorial 1", ...]}
```

Hem decidit aplicar un patró de disseny de referència (a l'atribut "editorials") a l'hora de decidir com està estructurada, ja que així podrem relacionar documents d'aquesta amb els de la col·lecció Editorial pel nom ("\_id") de les editorials. Hem cregut oportú fer-ho amb aquest patró, i no de l'estesa, perquè no creiem que hi hagi cap altre atribut a la col·lecció referenciada que també serveixi per identificar mínimament un document concret seu.

#### **- Publicació:**

```
{_id: "ISBN",  
  titol: "Títol",  
  stock: "Stock",
```

autor: "Autor",  
preu: "Preu",  
num\_pagines: "Número de pàgines",  
guionistes: ["Nom artístic 1", ...],  
dibuixants: ["Nom artístic 1", ...],  
\_id\_coleccio: "Nom col·lecció",  
editorial: "Nom editorial"}

Hem decidit aplicar un patró de disseny de referència de tres col·leccions (Col·lecció, Artista i Editorial) sobre aquesta col·lecció Publicació, per estructurar-la d'aquesta manera junt amb els atributs requerits. Primer, hem decidit relacionar-la així amb Col·lecció, ja que a un document d'aquesta pot tenir varies publicacions, però cadascuna d'aquestes només pertany a una col·lecció. Això per evitar fer un "embedding" de totes les publicacions a les col·leccions corresponents, ja que llavors hi hauria molta informació i documents en una mateixa col·lecció. Segon, hem decidit estructurar així els atributs de "guionistes" i "dibuixants" perquè qualsevol artista pot prendre qualsevol paper, o ambdós, en qualsevol de les publicacions existents. I tercer, assumint la redundància i dependències que això genera, hem decidit afegir un atribut "editorial" que faci referència directament a Editorial i que contingui l'editorial a on pertany la col·lecció de la qual forma part.

#### - Artista:

{\_id: "Nom artístic",  
nom: "Nom",  
cognoms: "Cognoms",  
data\_naixement: "Data de naixement",  
pais: "País"}

En resulta el mateix cas que la col·lecció Editorial, és a dir, hem cregut adient estructurar-la només amb els atributs propis i sense aplicar cap patró de disseny, perquè pensem que aquesta és la col·lecció que ha de ser referenciada per Publicació, i no al revés. Per tant, aquesta només conté els seus atributs demandats als requisits.

#### - Personatge:

{\_id: "Nom",  
tipus: "Tipus",  
publicacions: ["ISBN publicació 1", ...]}

Sobre la col·lecció Personatge, hem decidit aplicar un patró de disseny de referència de Publicació, perquè creiem que pot generar menys redundància que no pas al revés. Això degut a què, encara que una publicació pot tenir diversos personatges amb certs papers, cadascun d'aquests personatges amb un cert paper pot formar part de diferents publicacions.

## **Exercici 2:**

L'script de Python utilitza la base de dades "projecte", amb 5 col·leccions, que són les presentades a l'exercici 1, amb la mateixa estructura. Aquestes col·leccions només es creen si encara no existeixen. Això es controla amb un try/except, per evitar errors.

Se li ha de passar 2 paràmetres:

- O bé el paràmetre "-f", que ha d'anar acompanyat del nombre del fitxer del que es vol fer la lectura, que en aquest cas serà "Dades.xlsx". Llavors, faríem "-f Dades.xlsx" si volem omplir la base de dades.
- O bé, si volem eliminar les dades de la base, haurem d'utilitzar el paràmetre "-delete\_all", acompanyat del nombre de la base de dades, en aquest cas "projecte". Llavors, faríem "-delete\_all projecte" per eliminar les dades de les diverses col·leccions de la base de dades.

Si hem seleccionat la primera opció, i si l'arxiu és correcte (es comprova amb un try/except), llavors s'anirà llegint cada pàgina de l'excel i es treurà d'allà la informació pertinent a cada col·lecció. La informació de cada pàgina es treu en format llista de diccionaris, on cada diccionari és una línia del document de dades, on les claus són les capçaleres i el valor el contingut.

S'itera sobre la llista, i s'afegeixen les dades que encara no s'havien obtingut a diccionaris que representen la informació de cadascuna de les col·leccions. Un cop es tenen totes les dades, s'afegeixen les dades a la base de dades mitjançant la comanda insert\_many.

En canvi, si hem seleccionat la segona opció, s'agafarà una llista de totes les col·leccions que tingui la base de dades mitjançant la comanda list\_collection\_names(), i s'iterarà sobre cadascuna d'aquestes col·leccions, de les quals es farà un drop() per eliminar-les.

### **Exercici 3:**

**1. Les 5 publicacions amb major preu. Mostrar només el títol i preu.**

```
db.publicacions.find({}, {_id:0, titol:1, preu:1}).sort({preu: -1}).limit(5)
```

**2. Valor màxim, mínim i mitjà del preus de les publicacions de l'editorial Juniper Books.**

```
db.publicacions.aggregate([
  {
    $match: {
      editorial: "Juniper Books"
    }
  },
  {
    $group: {
      _id: "$editorial",
      precioMaximo: { $max: "$preu" },
      precioMinimo: { $min: "$preu" },
      precioMedio: { $avg: "$preu" }
    }
  }
])
```

**3. Artistes (nom artístic) que participen en més de 5 publicacions com a dibuixant.**

```
db.publicacions.aggregate([
  {$unwind: "$dibuixants"},
  {$group: {_id: "$dibuixants", count: {$sum: 1}}},
  {$match: {count: {$gt: 5}}},
  {$project: {_id: 1}}
])
```

**4. Número de col·leccions per gènere. Mostra gènere i número total.**

```
db.coleccions.aggregate([{$unwind: "$genere"},
  { $group: { _id: "$genere", total: { $sum: 1 } } },
  { $project: { total: 1 } }
])
```

**5. Per cada editorial, mostrar el recompte de col·leccions finalitzades i no finalitzades.**

```
db.coleccions.aggregate([
  {
    $unwind: "$editorials"
  },
  {
    $group: {
      _id: "$editorials",
      total: { $sum: 1 }
    }
  }
])
```

```
{
  $group:{
    _id: "$editorials",
    finalitzades : {$sum: { $cond:{ if :{ $eq: ["$tancada", true] }, then: 1, else: 0}}},
    no_finalitzades : {$sum: { $cond:{ if :{ $eq: ["$tancada", true] }, then: 0,else: 1}}}
  }
}
])
```

**6. Mostrar les 2 col·leccions ja finalitzades amb més publicacions. Mostrar editorial i nom col·lecció.**

```
db.coleccions.aggregate([
  {$match:{tancada:true}},
  {$sort:{total_exemplars:-1}},
  {
    $unwind:"$editorials"
  },
  {$limit:2},
  {$project:{_id:1,editorials:1}}
])
```

**7. Mostrar el país d'origen de l'artista o artistes que han fet més guions.**

```
db.artistes.aggregate([
  {
    $lookup:{
      from:"publicacions",
      localField:"_id",
      foreignField:"guionistes",
      as:"guions"
    }
  },
  {$addFields:{num_guions:{$size:"$guions"}}},
  {$sort:{num_guions:-1}},
  {$limit:1},
  {$project:{_id:0,pais:1}}
])
```

**8. Mostrar les publicacions amb tots els personatges de tipus "heroe".**

```
db.publicacions.aggregate([
```

```

{
  $lookup:{
    from: "personatges",
    localField: "_id",
    foreignField: "isbn",
    as: "personatges"
  }
},
{
  $match:{$and:[{"personatges.0":{"$exists:true"}},{personatges:{$not:{$elemMatch:{tipus:{$ne:"
  heroe"}}}}]}
},
  {$project: {_id:1}}
])

```

**9. Modificar el preu de les publicacions amb stock superior a 20 exemplars i incrementar-lo un 25%.**

```
db.publicacions.updateMany({"stock": {$gt: 20}}, {$mul: {"preu": 1.25}})
```

**10. Mostrar ISBN i títol de les publicacions conjuntament amb tota la seva informació dels personatges.**

```

db.publicacions.aggregate([
  {
    $lookup: {
      from: "personatges",
      localField: "_id",
      foreignField: "isbn",
      as: "personatges"
    }
  },
  {
    $project: {
      _id: 1,
      titol: 1,
      personatges: 1
    }
  }
])

```

### **Treball en equip:**

Cada membre ha estat involucrat en major o menor mesura en tots els exercicis.

Si parlem de la principal aportació de cada membre per cada exercici:

- El primer exercici ha estat pensat i fet principalment per l'Enric. Tot i això, com el disseny de les col·leccions té molt de pes en tots els altres exercicis, tothom ha revisat i suggerit canvis on era pertinent per agilitzar la feina, sobretot pel que fa a les consultes.
- El segon exercici ha estat pensat i fet principalment pel Joel, com es pot observar en els commitments del nostre repositori de GitHub. Per fer-ho, ha utilitzat l'esquema de l'exercici 1, i és l'script fet servir per a la realització del tercer exercici, amb les dades que aquest script ha carregat a la base de dades. Els encarregats del tercer exercici han informat sobre alguns errors que hi podien haver inicialment amb les dades carregades a l'hora de fer consultes, així que s'han anat fent modificacions a mesura que es necessitaven.
- El tercer exercici ha estat pensat i fet principalment pel Jordi, el David i l'Enric. Com ja s'ha comentat abans, han utilitzat les dades carregades amb l'script de Python, inspirat en el disseny de l'exercici 1.

A més, també ens hem hagut d'informar del funcionament de GitHub. D'això ha estat encarregat l'Enric, que va veure el vídeo del Campus Virtual i va ser qui va crear el repositori. Finalment, el principal redactor de l'informe ha estat el Joel, tot i que cada membre s'ha encarregat de la redacció del seu apartat d'exercicis i, per tant, l'única feina ha estat la redacció d'aquest apartat de cloenda.

Repositori de GitHub:

<https://github.com/1633263/bdnr2023uab.git>