

PRÀCTICA 1 ATS



GRUP 16
Max González Andrés
Gerard Simon Estadella

ESQUEMA DE LA BASE DE DADES

- ONE-TO-MANY

```
{
  "$jsonSchema": {
    "bsonType": "object",
    "required": ["name", "address", "type_of_food", "rating"],
    "properties": {
      "name": { "bsonType": "string" },
      "address": { "bsonType": "string" },
      "type_of_food": { "bsonType": "string" },
      "rating": {
        "bsonType": "int",
        "minimum": 0,
        "maximum": 5
      },
      "outcode": { "bsonType": "string" },
      "postcode": { "bsonType": "string" }
    }
  }
}
```

```
{
  "$jsonSchema": {
    "bsonType": "object",
    "required": ["restaurant_id", "date", "result"],
    "properties": {
      "restaurant_id": { "bsonType": "string" },
      "date": { "bsonType": "string" },
      "result": { "bsonType": "string" },
      "certificate_number": { "bsonType": "string" },
      "business_name": { "bsonType": "string" },
      "sector": { "bsonType": "string" },
      "address": { "bsonType": "object" }
    }
  }
}
```

CONSULTES SIMPLES

```
db.inspections.find({ "result": { $ne: "No Violation Issued" } }).sort({ "date": 1 }) |
```

```
db.restaurants.find({  
  "type_of_food": "Chinese"  
})
```

```
db.restaurants.find({  
  "rating": { $gt: 4 } })
```

ÚS D'AGREGACIONS

```
db.restaurants.aggregate([
  {
    $group: {
      _id: "$type_of_food",
      avgRating: { $avg: "$rating" },
      count: { $sum: 1 }
    }
  },
  {
    $sort: { avgRating: -1 }
  }
])
```

```
db.inspections.aggregate([
  {
    $group: {
      _id: "$result",
      count: { $sum: 1 }
    }
  },
  {
    $group: {
      _id: null,
      total: { $sum: "$count" },
      results: { $push: { result: "$_id", count: "$count" } }
    }
  },
  { $unwind: "$results" },
  {
    $project: {
      _id: 0,
      result: "$results.result",
      count: "$results.count",
      percentage: {
        $multiply: [
          { $divide: ["$results.count", "$total" ] },
          100
        ]
      }
    }
  }
])
```

```
db.restaurants.aggregate([
  {
    $addFields: {
      id_string: { $toString: "$_id" }
    }
  },
  {
    $lookup: {
      from: "inspections",
      localField: "id_string",
      foreignField: "restaurant_id",
      as: "inspections"
    }
  },
  {
    $project: {
      id_string: 0
    }
  },
  { $limit: 5 }
])
```

OPTIMITZACIÓ RENDIMENT

Índexs Restaurant:

type_of_food

rating

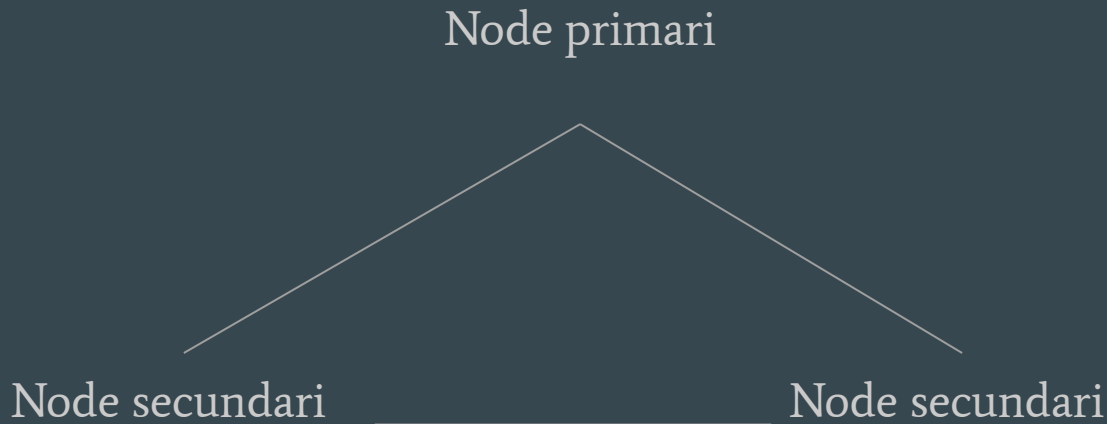
Índexs Inspeccions:

restaurant_id

result + data

ESTRATÈGIES D'ESCALABILITAT

Sharding hashejat sobre restaurant_id de inspections, ja que inspections és la colecció més gran i restaurant_id té molts valors diferents



Possibles colls d'ampolla i solucions

1. Consultes lentes per COLLSCAN → índexs adequats
2. \$lookup amb mal rendiment → índex sobre el camp restaurant_id a inspections.
3. Caiguda de nodes → Replica Set amb 3 nodes, amb recanvi automàtic de node
4. Ordenacions lentes amb sort → índexs compostos que inclouen els camps de les ordenacions

Gràcies