

1 Task 0: setup

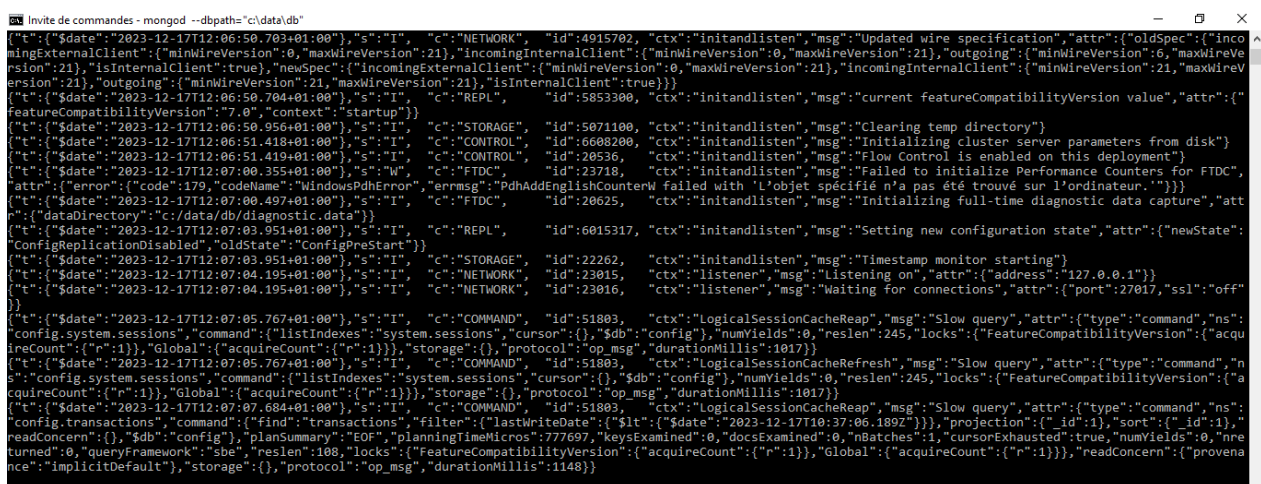
Create a data store directory for the MongoDB server. Launch the server (mongod) and write down the commands you used. On which port does it run?

Commands to create the data directory:

```
cd C:\  
md "\\data\db"
```

To start the MongoDB database:

```
"C:\Program Files\MongoDB\Server\7.0\bin\mongod.exe" --dbpath="c:\data\db"
```



```
Invite de commandes - mongod --dbpath="c:\data\db"  
{ "t": { "$date": "2023-12-17T12:06:50.703+01:00" }, "s": "I", "c": "NETWORK", "id": 4015702, "ctx": "initandlisten", "msg": "Updated wire specification", "attr": { "oldSpec": { "incomingExternalClient": { "minWireVersion": 0, "maxWireVersion": 21 }, "incomingInternalClient": { "minWireVersion": 0, "maxWireVersion": 21 }, "outgoing": { "minWireVersion": 6, "maxWireVersion": 21 }, "isInternalClient": true }, "newSpec": { "incomingExternalClient": { "minWireVersion": 0, "maxWireVersion": 21 }, "incomingInternalClient": { "minWireVersion": 21, "maxWireVersion": 21 }, "outgoing": { "minWireVersion": 21, "maxWireVersion": 21 }, "isInternalClient": true } } }, "featureCompatibilityVersion": 7, "context": "startup" }, { "t": { "$date": "2023-12-17T12:06:50.956+01:00" }, "s": "I", "c": "STORAGE", "id": 5071100, "ctx": "initandlisten", "msg": "Clearing temp directory" }, { "t": { "$date": "2023-12-17T12:06:51.418+01:00" }, "s": "I", "c": "CONTROL", "id": 6608200, "ctx": "initandlisten", "msg": "Initializing cluster server parameters from disk" }, { "t": { "$date": "2023-12-17T12:06:51.419+01:00" }, "s": "I", "c": "CONTROL", "id": 20536, "ctx": "initandlisten", "msg": "Flow Control is enabled on this deployment" }, { "t": { "$date": "2023-12-17T12:07:00.355+01:00" }, "s": "W", "c": "FTDC", "id": 23718, "ctx": "initandlisten", "msg": "Failed to initialize Performance Counters for FTDC", "attr": { "error": { "code": 179, "codeName": "WindowsPdhError", "errmsg": "PdhAddEnglishCounterW failed with 'L'objet sp\u00e9cifi\u00e9 n'a pas \u00e9t\u00e9 trouv\u00e9 sur l'ordinateur.'" } } }, { "t": { "$date": "2023-12-17T12:07:00.497+01:00" }, "s": "I", "c": "FTDC", "id": 20625, "ctx": "initandlisten", "msg": "Initializing full-time diagnostic data capture", "attr": { "dataDirectory": "c:/data/db/diagnostic.data" } }, { "t": { "$date": "2023-12-17T12:07:03.951+01:00" }, "s": "I", "c": "REPL", "id": 6015317, "ctx": "initandlisten", "msg": "Setting new configuration state", "attr": { "newState": "configReplicationDisabled", "oldState": "ConfigPreStart" } }, { "t": { "$date": "2023-12-17T12:07:03.951+01:00" }, "s": "I", "c": "STORAGE", "id": 22262, "ctx": "initandlisten", "msg": "Timestamp monitor starting" }, { "t": { "$date": "2023-12-17T12:07:04.195+01:00" }, "s": "I", "c": "NETWORK", "id": 23015, "ctx": "listener", "msg": "Listening on", "attr": { "address": "127.0.0.1" } }, { "t": { "$date": "2023-12-17T12:07:04.195+01:00" }, "s": "I", "c": "NETWORK", "id": 23016, "ctx": "listener", "msg": "Waiting for connections", "attr": { "port": 27017, "ssl": "off" } }, { "t": { "$date": "2023-12-17T12:07:05.767+01:00" }, "s": "I", "c": "COMMAND", "id": 51803, "ctx": "LogicalSessionCacheReap", "msg": "Slow query", "attr": { "type": "command", "ns": "config.system.sessions", "command": { "listIndexes": "system.sessions", "cursor": {}, "$db": "config" }, "numYields": 0, "reslen": 245, "locks": { "FeatureCompatibilityVersion": { "acquireCount": { "r": 1 }, "Global": { "acquireCount": { "r": 1 } }, "storage": {}, "protocol": "op_msg", "durationMillis": 1017 } } }, { "t": { "$date": "2023-12-17T12:07:05.767+01:00" }, "s": "I", "c": "COMMAND", "id": 51803, "ctx": "LogicalSessionCacheRefresh", "msg": "Slow query", "attr": { "type": "command", "ns": "config.system.sessions", "command": { "listIndexes": "system.sessions", "cursor": {}, "$db": "config" }, "numYields": 0, "reslen": 245, "locks": { "FeatureCompatibilityVersion": { "acquireCount": { "r": 1 }, "Global": { "acquireCount": { "r": 1 } }, "storage": {}, "protocol": "op_msg", "durationMillis": 1017 } } }, { "t": { "$date": "2023-12-17T12:07:07.684+01:00" }, "s": "I", "c": "COMMAND", "id": 51803, "ctx": "LogicalSessionCacheReap", "msg": "Slow query", "attr": { "type": "command", "ns": "config.transactions", "command": { "find": "transactions", "filter": { "lastWriteDate": { "$lt": { "$date": "2023-12-17T10:37:06.180Z" } } }, "projection": { "_id": 1 }, "sort": { "_id": 1 }, "readConcern": { "level": "snapshot" }, "planSummary": "EOF", "planningTimeMicros": 777697, "keysExamined": 0, "docsExamined": 0, "batches": 1, "cursorExhausted": true, "numYields": 0, "returned": 0, "queryFramework": "sbe", "reslen": 108, "locks": { "FeatureCompatibilityVersion": { "acquireCount": { "r": 1 }, "Global": { "acquireCount": { "r": 1 } }, "readConcern": { "provenance": "implicitDefault" }, "storage": {}, "protocol": "op_msg", "durationMillis": 1148 } } }
```

The MongoDB start running in port 27017 as it displays in the message:

```
{ "t": { "$date": "2023-12-17T12:07:04.195+01:00" }, "s": "I", "c": "NETWORK", "id": 23016, "ctx": "listener", "msg": "Waiting for connections", "attr": { "port": 27017, "ssl": "off" } }
```

Import the file moviepeople-10.jsonl into the server?

Importing a JSON file into MongoDB while in the command line can be made by using mongoimport and defining the database and collection to upload to.

```
mongoimport --db Lab_2 --collection movie --file  
.\lab_2_datasets\moviepeople-10.jsonl
```

```
C:\Users\user1\Music>mongoimport --db Lab_2 --collection movie --file .\lab_2_datasets\moviepeople-10.jsonl  
2023-12-17T12:27:30.462+0100 connected to: mongod://localhost/  
2023-12-17T12:27:30.538+0100 10 document(s) imported successfully. 0 document(s) failed to import.
```

Launch a mongo shell client, and ask a query to retrieve all the data?

We start a mongo shell client by executing the command mongosh inside the terminal.

```

C:\Users\User1\Music\mongosh-2.1.1-win32-x64\bin>mongosh
Current Mongosh Log ID: 657edd149ed354db0ded866b
Connecting to:  mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+2.1.1
Using MongoDB:  7.0.4
Using Mongosh:   2.1.1

For mongosh info see: https://docs.mongodb.com/mongosh-shell/

-----
The server generated these startup warnings when booting
2023-12-17T12:18:16.080+01:00: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted
2023-12-17T12:18:16.081+01:00: This server is bound to localhost. Remote systems will be unable to connect to this server. Start the server with --bind_ip <address>
to specify which IP addresses it should serve responses from, or with --bind_ip_all to bind to all interfaces. If this behavior is desired, start the server with --bind_ip 127.0.0.1 to disable this warning
-----

test>

```

To query all the results in the client we use:

```

use Lab_2
db.movie.find( { } )

```

```

mongosh mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000

test> use Lab_2
switched to db Lab_2
Lab_2> db.movie.find( { } )
[
  {
    _id: ObjectId('657edaec3cd40322f05aa79f'),
    'person-name': 'Hughart, Ron',
    info: {
      birthnotes: [ 'Los Angeles, California, USA' ],
      birthdate: [ '18 June 1961' ],
      birthname: [ 'Hughart, Ronald P' ]
    }
  },
  {
    _id: ObjectId('657edaec3cd40322f05aa7a0'),
    'person-name': 'Cáceres, Luciano',
    info: {
      trivia: [
        "Grew up in Buenos Aires's neighborhood Bajo Flores.",
        "Was classmates of 'Sergio Sunraco' (qv), 'Claudio Tolcachir' (qv), 'Lautaro Delgado' (qv) and 'Melina Petriella' (qv) at the drama school.",
        "Stepfather of 'Ángela Torres (III)' (qv).",
        "Became a father for the 1st time at age 32 when his wife 'Gloria Carrá' (qv) gave birth to their daughter Amelia Cáceres on October 7, 2009.",
        "Met wife 'Gloria Carrá' (qv) in 2001, but they didn't become a couple until 2007.",
        "Is an accomplished stage actor/director.",
        "Was in a relationship with actress 'Laura Lopez Moyano' (qv) (2005-2007).",
        "His mother, political activist Haydeé Bello, died of cancer on June 15, 2002.",
        "Studied acting with 'Alejandra Boero' (qv).",
        "He started taking acting classes at the age of 9.",
        "Son of 'Ernesto Cáceres' (qv) and Haydeé Bello."
      ],
      birthnotes: [ 'Autonomous City of Buenos Aires, Argentina' ],
      birthdate: [ '24 January 1977' ],
      height: [ '187 cm' ],
      spouse: [
        "'Gloria Carrá' (qv) (27 September 2008 - 2015) (separated); 1 child"
      ],
      nicknames: [ 'Lu' ],
      trademark: [ 'Curly hair', 'Bright blue eyes' ],
      interviews: [
        "Clarín (Argentina), 9 January 2007, by: María Ana Rago, Luciano Cáceres: El teatro en la sangre"
      ]
    }
  }
]

```

2 Task 1: import and querying

Import the files moviepeople-3000.jsonl and cities.jsonl into the server (the one set up in Task 0)?

```

C:\Users\user1\Music\lab_2_datasets>mongoimport --db Lab_2 --collection movie --file .\moviepeople-3000.jsonl
2023-12-17T12:49:22.455+0100    connected to: mongodb://localhost/
2023-12-17T12:49:23.009+0100    3000 document(s) imported successfully. 0 document(s) failed to import.

```

Figure 1: Importing moviepeople-3000.jsonl

```

C:\Users\user1\Music\lab_2_datasets>mongoimport --db Lab_2 --collection movie --file .\cities.jsonl
2023-12-17T12:50:07.326+0100    connected to: mongodb://localhost/
2023-12-17T12:50:10.329+0100    [#####.....] Lab_2.movie  5.46MB/13.7MB (40.0%)
2023-12-17T12:50:13.336+0100    [#####.....] Lab_2.movie  12.7MB/13.7MB (93.0%)
2023-12-17T12:50:13.799+0100    [#####.....] Lab_2.movie  13.7MB/13.7MB (100.0%)
2023-12-17T12:50:13.799+0100    99838 document(s) imported successfully. 0 document(s) failed to import.

```

Figure 2: Importing cities.jsonl

In the mongo client shell, write queries to find?

Information about the person called Barbara Norton de Matos:

```
db.movie.find({"person-name": "Norton de Matos, Barbara"})
```

```
Lab_2> db.movie.find({"person-name": "Norton de Matos, Barbara"})
[
  {
    _id: ObjectId('657ee04201520e10058772ab'),
    'person-name': 'Norton de Matos, Bárbara',
    info: {
      trivia: [
        "Daughter of actor 'Norton de Matos' (qv).",
        "Engaged to Ricardo Areias and their wedding is scheduled for June 2017. They have a daughter together.",
        "Daughter of Luís Maria Cabral Norton de Matos and his former and first wife and distant cousin Maria Rita Cardoso Franco de Matos. Her former stepmother, Alexandra Maria Pires Nunes, is a fashion model with the name Xana Nunes. Stepmother: Joana Leitão Tito Martins. Has an older brother named Luís Maria Franco de Sousa Norton de Matos, a commercial pilot, and two younger half-sisters: Carolina Pires Nunes Norton de Matos (b. 17 Aug 1989) and Maria Tito Martins Norton de Matos (b. 10 Dec 1987). Has a daughter, Luz Norton de Matos Pina e Melo (b. 7 Apr 2006). Her father is Gonçalo Pina e Melo (b. 1977). Her surname Norton is of English origin. She also has distant English and Welsh ancestry.",
        "Gave birth to her 2nd child at age 36, a daughter named Flor Areias on May 11, 2016. Child's father is her fiance, Ricardo Areias.",
        "Her father is a first cousin of actor José Maria Norton de Matos.",
        "She and Gonçalo Pina e Melo (Luz's father) split in the beginning of 2007."
      ],
      birthnotes: [ 'Lisbon, Portugal' ],
      birthdate: [ '7 July 1979' ],
      birthname: [ 'de Matos, Bárbara Franco de Sousa Norton' ]
    }
  }
]
```

Figure 3: Finding Barbara

The birthplace of Steven Spielberg:

```
db.movie.find( { "person-name": "Spielberg, Steven" }, { "_id": 0, "info": { "birthnotes":
  ↳ 1 } } )
```

```
Lab_2> db.movie.find( { "person-name": "Spielberg, Steven" }, { "_id": 0, "info": { "birthnotes": 1 } } )
[
  { info: { birthnotes: [ 'Cincinnati, Ohio, USA' ] ] }
]
```

Figure 4: Finding birthplace

The number of people born in Lisbon

```
db.movie.aggregate([ { $match: { "info.birthnotes": { $in: ["Lisbon, Portugal"] } } }, {
  ↳ $group: { "_id": "$info.birthnotes", "count": { $count: { } } } } ] )
```

```
Lab_2> db.movie.aggregate([ { $match: { "info.birthnotes": { $in: ["Lisbon, Portugal"] } } }, { $group: { "_id": "$info.birthnotes", "count": { $count: { } } } } ] )
[ { _id: [ 'Lisbon, Portugal' ], count: 171 } ]
```

Figure 5: Counting number

The people taller than 170cm

```
db.movie.find({ $or: [ { $and: [ { "info.height": {$regex: ".*cm"} }, { "info.height": {$gt
  ↳ : "170 cm"} } ] }, { $and: [ { "info.height": {$regex: "(?!.*?cm)^.*$"} }, { "info.
  ↳ height": {$gte: "5' 7"} } ] } ] }, {"person-name": 1, "info.height": 1 } )
```

Figure 6: Finding people

```
db.movie.find( {"info.otherworks": {$regex: "Opera"}}, {"person-name": 1} )
```

Figure 7: Name of people

```
db.movie.aggregate( [ {$project: {"person-name": 1, "info.spouse": 1}}, {$unwind: {path: "
    ↳ $info.spouse"}}, { $group: { "_id": "$person-name", "lastSpouse": {$last: "$info.
    ↳ spouse"}, "countSpouse": {$count: {}} } }, {$match: {"countSpouse": {$gte: 1}} }, {
    ↳ $project: {"person-name": "$_id", "lastSpouse": 1, "_id": 0}} ] )
```

Figure 8: Last Spouse

```
db.movie.aggregate([ { $project: { "person-name": 1, "city": { $first: { $split: [ { $first: "$info.birthnotes" }, " ", " ], } }, } }, { $match: { "city": {$ne: null} } }, { $lookup: { from: "cities", localField: "city", foreignField: "name", as: "city" } }, { $match: { "city": {$ne: null, $not: {$size: 0}} } } ] )
```

3 Task 2: replication

Create data store directories for three MongoDB servers?

We create the data store directories as we did in task 0.

```
md c:\data\db1  
md c:\data\db2  
md c:\data\db3
```

Create a replica set for a collection called small-movie. Hint: to do that you will need to launch the three MongoDB servers (in different shells) and let them run?