MongoDB

dr Sara Migliorini

NoSQL Systems – MongoDB

- MongoDB è un sistema per la gestione di basi di dati NoSQL di tipo document-based.
 - Una base di dati è formata da collezioni di documenti.
 - Ciascuna collezione corrisponde al concetto di tabella (o entità).
 - Ciascuna collezione è formata da insiemi di documenti (tuple).
- Altre tipologie di database NoSQL
 - Graph database (e.g. Neo4j)
 - Column-oriented database (e.g. Apache Cassandra)
 - Key-value database (e.g. Amazon DynamoDB)
 - Document-oriented database (e.g. MongoDB)

MongoDB: Caratteristiche

- Struttura a documenti che ricorda l'organizzazione ad oggetti dei linguaggi di programmazione moderni
- Flessibilità e assenza di schema (schema-less)
 - Dati la cui struttura evolve nel tempo
 - Dati non strutturati o semi-strutturati
- Scalabilità e performance
 - Fondamentalmente MongoDB è un sistema distribuito
 - La parte server è stata sviluppata pensando che la configurazione di default prevedesse un cluster (replica)
 - Ridondanza e alta disponibilità
- Linguaggio di interrogazione e indici per semplicare e velocizzare operazioni comuni, come findOne() e updateOne()

MongoDB: Tools

Administrative tools: per la gestione dei database

- MongoDB shell (mongosh): da riga di comando
- MongoDB Compass: interfaccia grafica
 - https://www.mongodb.com/try/download/compass

Backup e restore: per la creazione o ripristino di backup

- mongodbdump: per creare un backup binario
- mongodbrestore: per il ripristino di un backup

MongoDB: Tools

Performance and monitoring tools: per verificare le prestazioni del cluster e monitorare lo stato di salute del sistema

- mongostat: statistiche sulle prestazioni
- mongotop: monitorare il tempo richiesta dalle instanze per eseguire operazioni di lettura e scrittura

Developer and integration tools

MongoDB per Visual Studio Code

• . . .

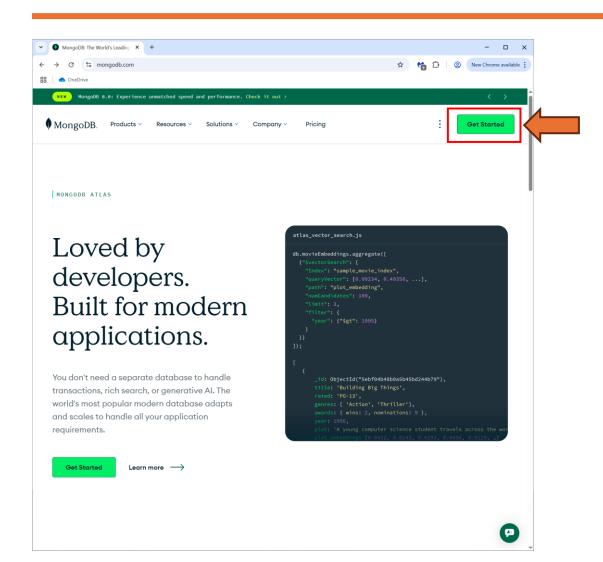
MongoDB: Installazione

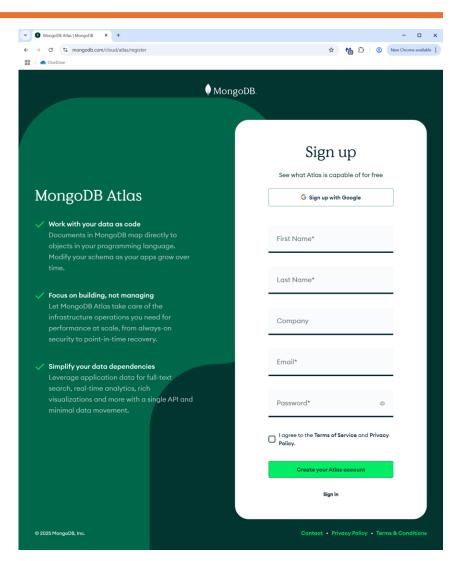
- MongoDB cluster:
 - Scaricare ed installare MongoDB Community Edition
 - https://www.mongodb.com/try/download/community



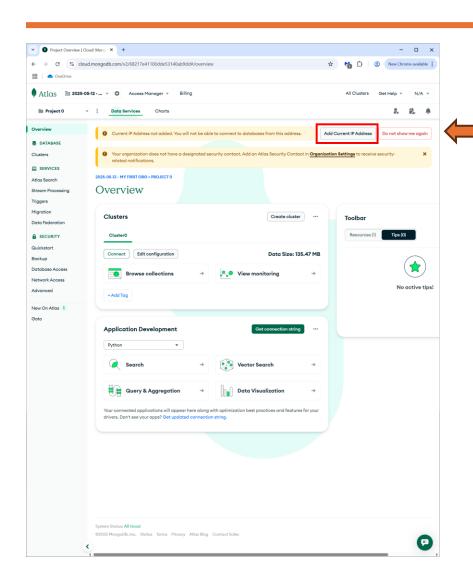
- Creare un account su MongoDB Atlas ed utilizzare il cluster gratuitio
 - https://www.mongodb.com/try

MongoDB: Cluster su Atlas



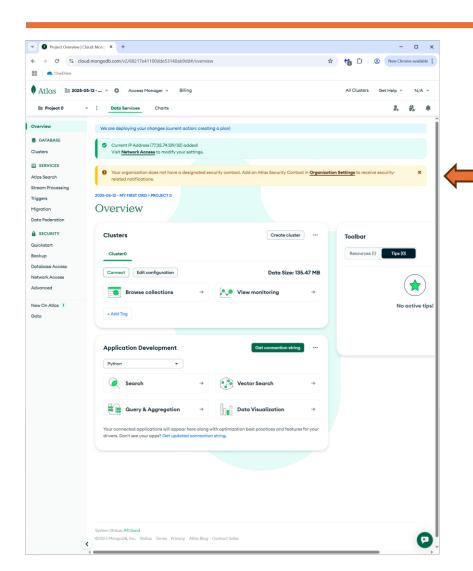


MongoDB: Creazione Cluster su Atlas



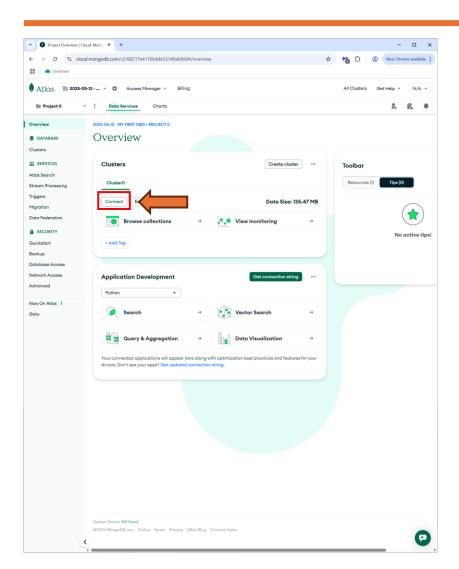
1. Aggiungere l'IP alla lista di quelli autorizzati alla connessione

MongoDB: Creazione Cluster su Atlas



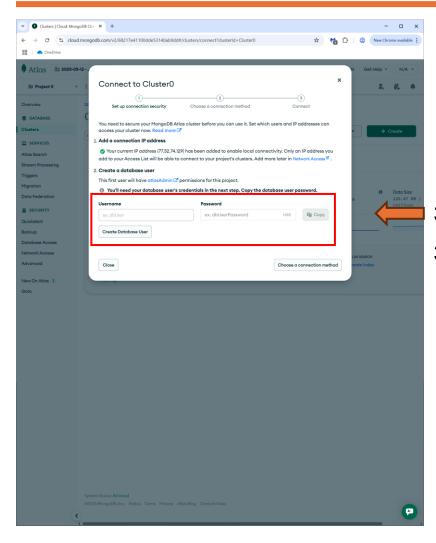
- 1. Aggiungere l'IP alla lista di quelli autorizzati alla connessione
- 2. Completare le configurazioni di sicurezza aggiungendo un'email di contatto

MongoDB: Creazione Cluster su Atlas

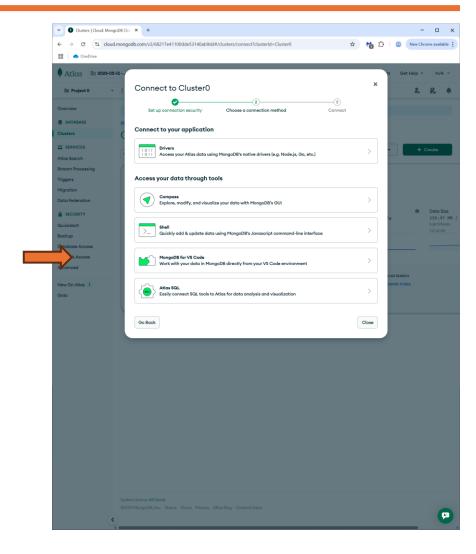


- 1. Aggiungere l'IP alla lista di quelli autorizzati alla connessione
- 2. Completare le configurazioni di sicurezza aggiungendo un'email di contatto
- 3. Connessione

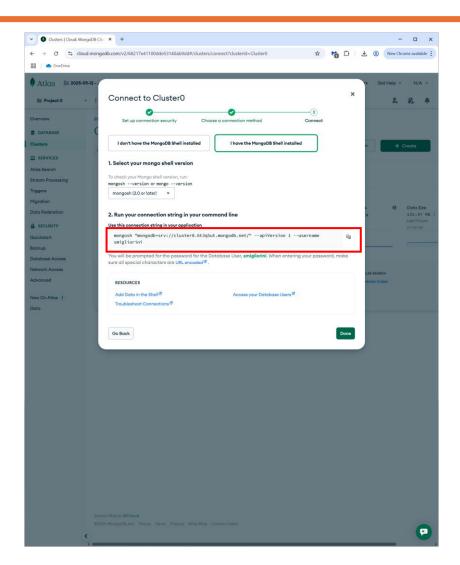
MongoDB: Creazione DB e Connessione



3a. Creazione utente3b. Scelta del client



Connessione a MongoDB tramite mongosh



Connessione a MongoDB tramite mongosh

- La shell mongosh può essere installata in vario modo:
 - Download al link: https://www.mongodb.com/try/download/shell
 - Usare NPM (Node Package Manager) se Node.js è installato nel Sistema
 - Aprire il terminale
 - npm install -g mongosh
- Connessione ad un server MongoDB
 - Comando generico:
 - mongosh "mongodb://username:password@hostname:port/test"
 - Cluster Altas (copia ed incolla dall'interfaccia web):
 - mongosh "mongodb+srv://cluster0.6t3q5ut.mongodb.net/" --apiVersion 1 -username smigliorini

Connessione a MongoDB tramite mongosh

```
mongosh mongodb+srv://<credentials>@cluster0.6t3q5ut.mongodb.net/
Microsoft Windows [Version 10.0.19045.5737]
(c) Microsoft Corporation. All rights reserved.
C:\Users\yoshi>mongosh "mongodb+srv://cluster0.6t3q5ut.mongodb.net/" --apiVersion 1 --username smigliorini
Enter password: *********
Current Mongosh Log ID: 6821bfb362f2d74ffe6c4bcf
Connecting to:
                        mongodb+srv://<credentials>@cluster0.6t3q5ut.mongodb.net/?appName=mongosh+2.5.1
Using MongoDB:
                        8.0.8 (API Version 1)
Using Mongosh:
                        2.5.1
For mongosh info see: https://www.mongodb.com/docs/mongodb-shell/
To help improve our products, anonymous usage data is collected and sent to MongoDB periodically (https://www.mongodb.com/legal/privacy-
You can opt-out by running the disableTelemetry() command.
Atlas atlas-swiyxj-shard-0 [primary] test>
```

Mostrare tutti i database

• test> show dbs

```
mongosh mongodb+srv://<credentials>@cluster0.6t3q5ut.mongodb.net/
                                                                                                                              - □ ×
Microsoft Windows [Version 10.0.19045.5737]
(c) Microsoft Corporation. All rights reserved.
C:\Users\yoshi>mongosh "mongodb+srv://cluster0.6t3q5ut.mongodb.net/" --apiVersion 1 --username smigliorini
Inter password: *********
Current Mongosh Log ID: 6821bfb362f2d74ffe6c4bcf
Connecting to:
                       8.0.8 (API Version 1)
Jsing MongoDB:
Jsing Mongosh:
For mongosh info see: https://www.mongodb.com/docs/mongodb-shell/
To help improve our products, anonymous usage data is collected and sent to MongoDB periodically (https://www.mongodb.com/legal/privacy-
You can opt-out by running the disableTelemetry() command.
Atlas atlas-swiyxj-shard-0 [primary] test> show dbs
sample mflix 121.84 MiB
admin
             348.00 KiB
               3.02 GiB
Atlas atlas-swiyxj-shard-0 [primary] test> 🕳
```

Selezionare un particolare database su cui lavorare

- test> use sample_mflix
- switched to db sample_mflix

```
mongosh mongodb+srv://<credentials>@cluster0.6t3q5ut.mongodb.net/
 icrosoft Windows [Version 10.0.19045.5737]
(c) Microsoft Corporation. All rights reserved.
:\Users\yoshi>mongosh "mongodb+srv://cluster0.6t3q5ut.mongodb.net/" --apiVersion 1 --username smigliorini
Enter password: ********
Current Mongosh Log ID: 6821bfb362f2d74ffe6c4bcf
Connecting to:
Using MongoDB:
                        8.0.8 (API Version 1)
Using Mongosh:
For mongosh info see: https://www.mongodb.com/docs/mongodb-shell/
To help improve our products, anonymous usage data is collected and sent to MongoDB periodically (https://www.mongodb.com/legal/privacy-
You can opt-out by running the disableTelemetry() command.
Atlas atlas-swiyxj-shard-0 [primary] test> show dbs
sample mflix 121.84 MiB
               3.02 GiB
local
Atlas atlas-swiyxj-shard-0 [primary] test> use sample_mflix
switched to db sample_mflix
Atlas atlas-swiyxj-shard-0 [primary] sample_mflix> show collections
comments
embedded movies
novies
sessions
theaters
Atlas atlas-swiyxj-shard-0 [primary] sample mflix> _
```

Mostrare le collezioni di documenti presenti in un database

sample_mflix> show collections

```
mongosh mongodb+srv://<credentials>@cluster0.6t3q5ut.mongodb.net/
                                                                                                                                   Microsoft Windows [Version 10.0.19045.5737]
(c) Microsoft Corporation. All rights reserved.
C:\Users\yoshi>mongosh "mongodb+srv://cluster0.6t3q5ut.mongodb.net/" --apiVersion 1 --username smigliorini
Enter password: *********
Current Mongosh Log ID: 6821bfb362f2d74ffe6c4bcf
Using MongoDB:
                       8.0.8 (API Version 1)
Using Mongosh:
For mongosh info see: https://www.mongodb.com/docs/mongodb-shell/
To help improve our products, anonymous usage data is collected and sent to MongoDB periodically (https://www.mongodb.com/legal/privacy
You can opt-out by running the disableTelemetry() command.
Atlas atlas-swiyxj-shard-0 [primary] test> show dbs
sample mflix 121.84 MiB
               3.02 GiB
Atlas atlas-swiyxj-shard-0 [primary] test> use sample_mflix
switched to db sample_mflix
Atlas atlas-swiyxj-shard-0 [primary] sample_mflix> show collections
comments
embedded movies
movies
sessions
theaters
Atlas atlas-swiyxj-shard-0 [primary] sample mflix> _
```

Visualizare tutt i comandi disponibili

test> help

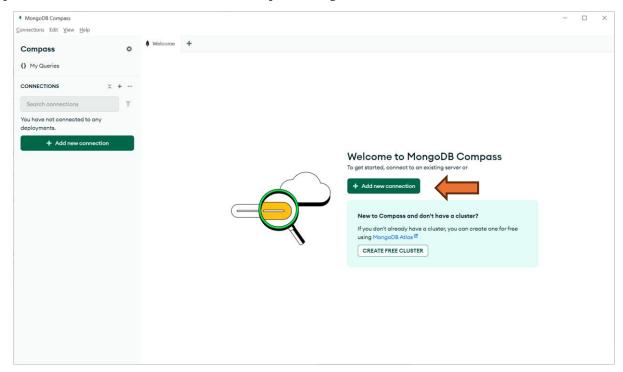
Uscire dalla shell

• test> exit

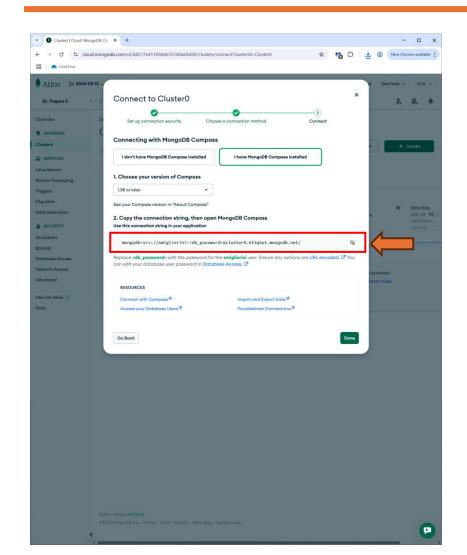
```
mongosh mongodb+srv://<credentials>@cluster0.6t3q5ut.mongodb.net/
embedded movies
movies
sessions
theaters
Atlas atlas-swiyxj-shard-0 [primary] sample mflix> help
  Shell Help:
   log
                                               'log.info(<msg>)': Write a custom info/warn/error/fatal/debug message to the log file
                                               'log.getPath()': Gets a path to the current log file
                                               Set current database
   use
   show
                                               'show databases'/'show dbs': Print a list of all available databases
                                               'show collections'/'show tables': Print a list of all collections for current database
                                               'show profile': Prints system.profile information
                                               'show users': Print a list of all users for current database
                                               'show roles': Print a list of all roles for current database
                                               'show log <type>': log for current connection, if type is not set uses 'global'
                                               'show logs': Print all logs
   exit
                                               Quit the MongoDB shell with exit/exit()/.exit
                                               Quit the MongoDB shell with quit/quit()
   quit
                                               Create a new connection and return the Mongo object. Usage: new Mongo(URI, options [optional])
   Mongo
                                               Create a new connection and return the Database object. Usage: connect(URI, username [optional], password [optional])
   connect
                                               result of the last line evaluated; use to further iterate
   version
                                               Shell version
                                               Loads and runs a JavaScript file into the current shell environment
   load
   enableTelemetry
                                               Enables collection of anonymous usage data to improve the mongosh CLI
   disableTelemetry
                                               Disables collection of anonymous usage data to improve the mongosh CLI
   passwordPrompt
                                               Prompts the user for a password
   sleep
                                               Sleep for the specified number of milliseconds
   print
                                               Prints the contents of an object to the output
   printjson
                                               Alias for print()
                                               Returns the hashed value for the input using the same hashing function as a hashed index. cls
   convertShardKeyToHashed
       Clears the screen like console.clear()
                                               Returns whether the shell will enter or has entered interactive mode
  For more information on usage: https://docs.mongodb.com/manual/reference/method
Atlas atlas-swiyxj-shard-0 [primary] sample_mflix> _
```

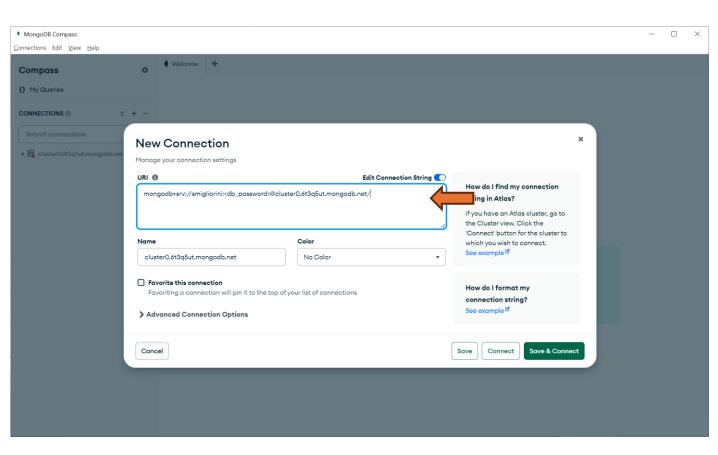
MongoDB Compass: Connessione

- Interfaccia grafica alternativa all'uso della shell mongosh
- Rappresentazione visuale dei database e delle collezioni
- Visualizza le performance di query ed indici

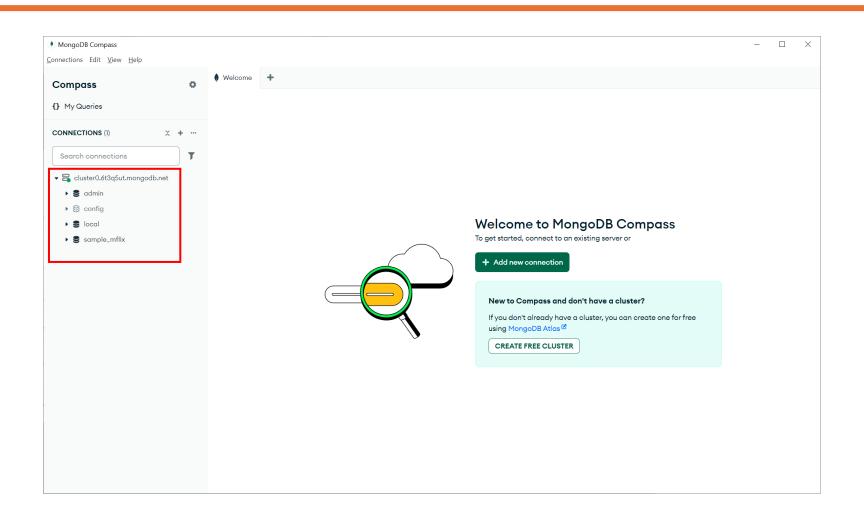


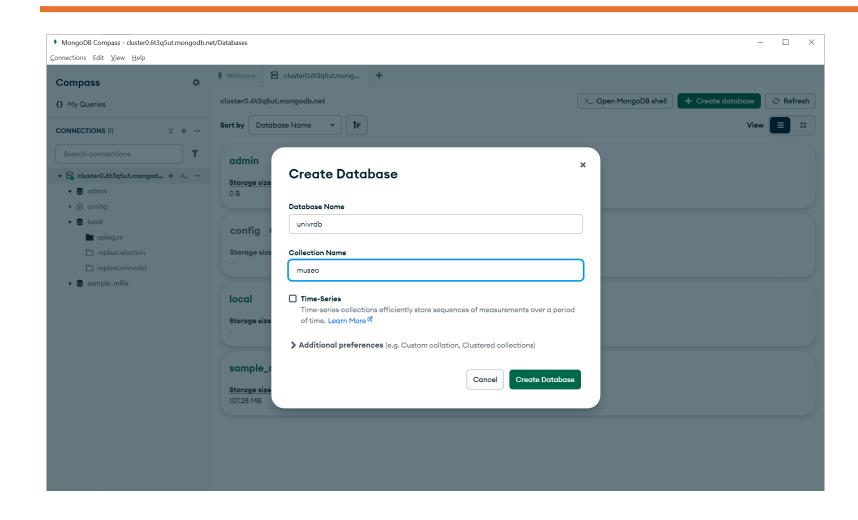
MongoDB Compass: Connessione



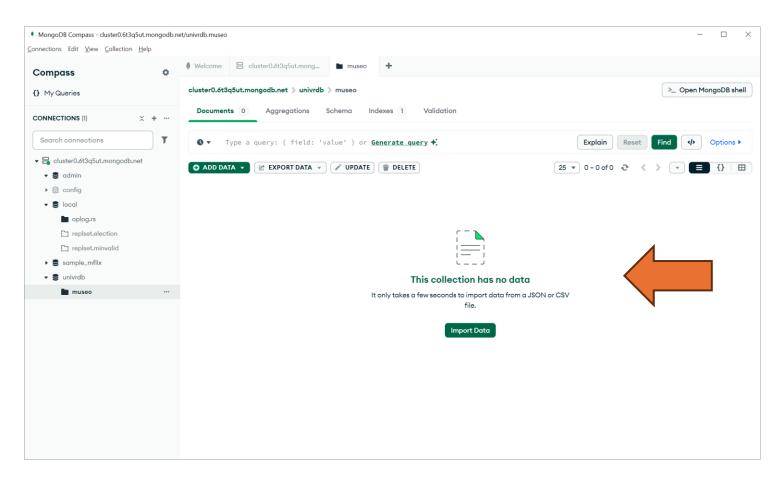


MongoDB Compass: Consultazione DB

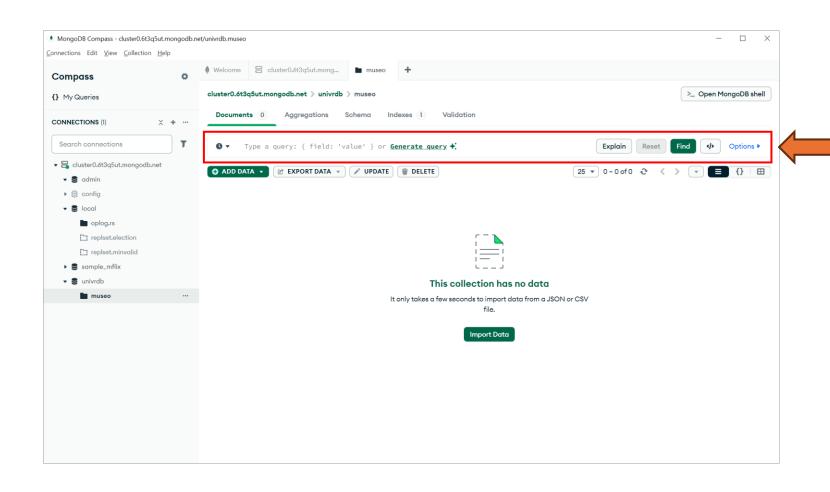




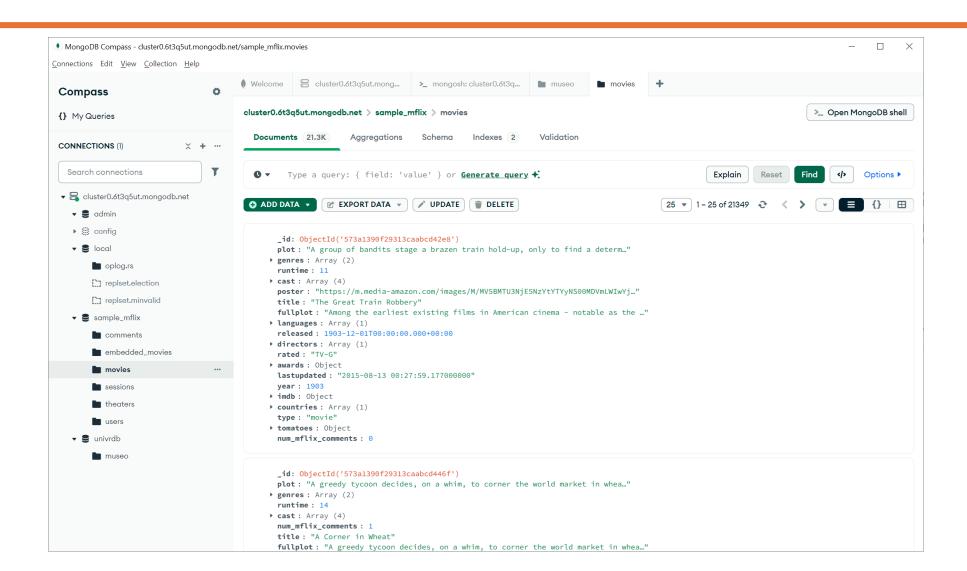
Un database contiene almeno una collezione di documenti.

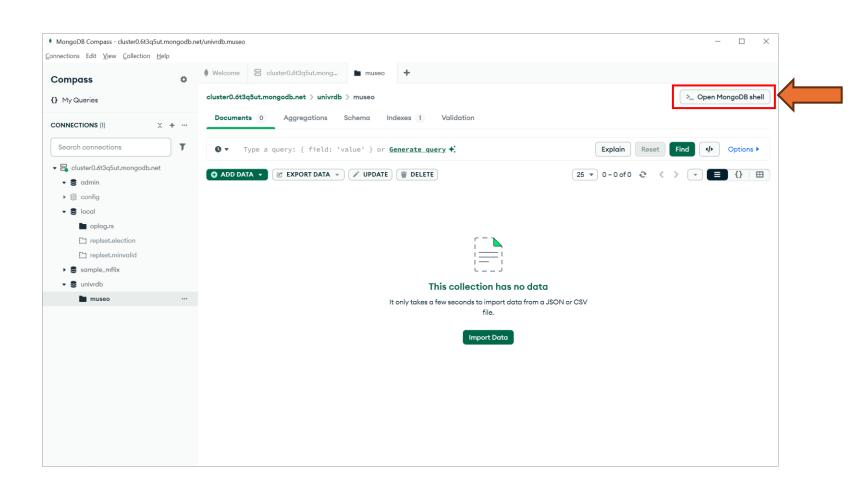


Una volta creato un database ed una collezione è possibile importare i dati (documenti) che fanno parte della conessione tramite file CSV oppure JSON.

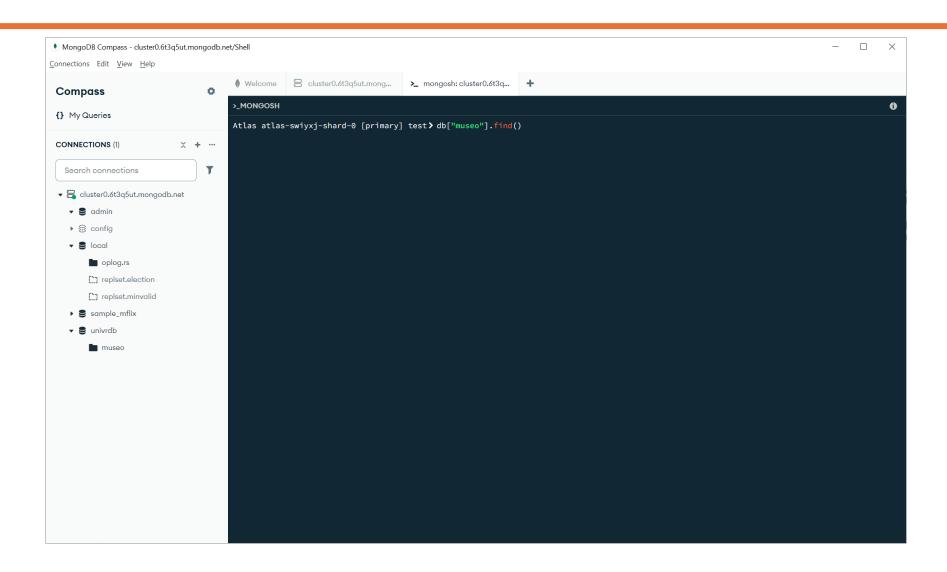


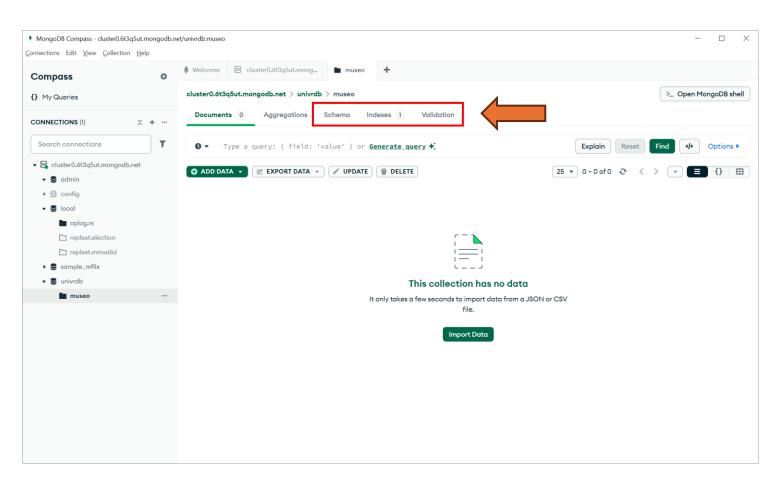
Selezionata una collezione è possibile vederne un campione oppure interrogarla tramite Query textbox.



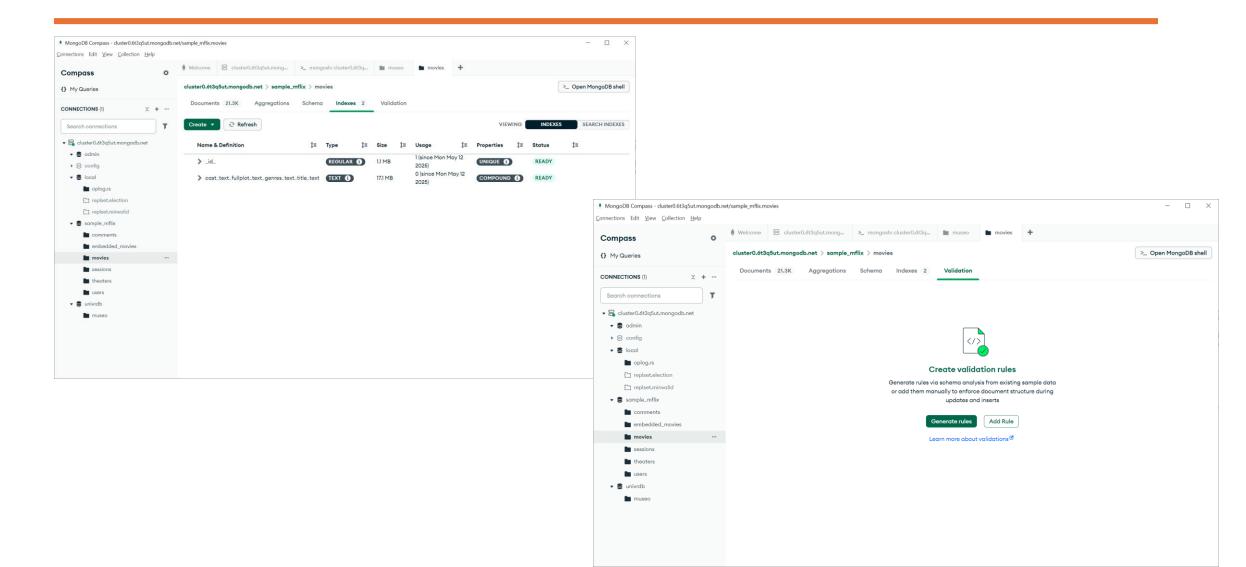


Usare la shell MongoDB integrata in Compass.





- Vedere lo schema dei documenti della collezione.
- Consultare gli indici della collezione.
- Validation rules: forzare delle regole di struttura dei documenti che dovno essere soddisfatte durante gli aggiornamenti e gli inserimenti.



Operazioni CRUD

Operazioni tipiche su una base di dati MongoDB

- Create: inserire un nuovo documento in un database MongoDB
- Read: interrogare uno o più documenti
- Update: modificare un documento esistente
- Delete: rimuovere un documento esistente

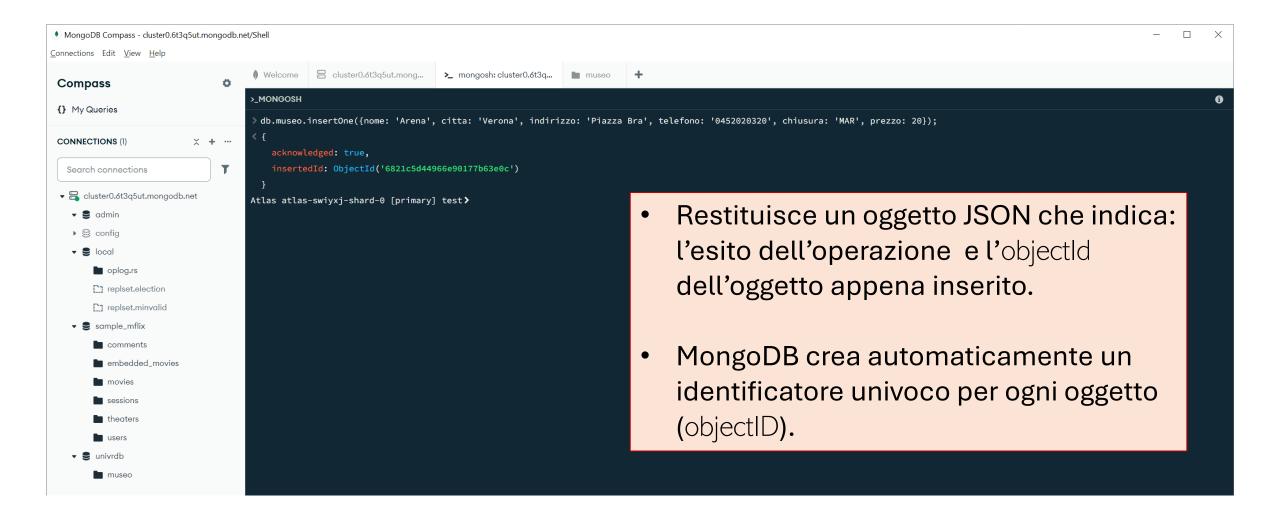
Punto di partenza

- show dbs
- use <dbname>

Inserimento di un documento in una collezione (insertOne)

```
db.museo.insertOne({ nome: 'Arena',
       citta: 'Verona',
       indirizzo: 'Piazza Bra',
       telefono: '04500303',
       chiusura: 'MAR',
       prezzo: 20,
 });
Nota:
   • db.museo = db["museo"] = db.collection('museo')
```





Inserimento di più documenti in una collezione (insertMany)

```
citta: 'Verona',
      . . .
     { nome: 'Castelvecchio',
      citta: 'Verona',
      . . .
```

Riassunto dei Metodi disponibili:

- db.collection.insertOne()
- db.collection.insertMany()

Read: Ricerca

MongoDB Compass - cluster0.6t3q5ut.mongodb.net/Shell

Ricerca di un documento in una collezione con un certo filtro

db.museo.find({citta: 'Verona'});

- Riceve in ingresso un oggetto JSON contente il filtro (opzionale).
- Restituisce gli oggetti JSON che soddisfano il filtro.

```
Connections Edit View Help

☐ cluster0.6t3a5ut.mona... 
➤ monaosh; cluster0.6t3a...
 Compass
 {} My Queries
                                           db.museo.find({citta: 'Verona'});
 CONNECTIONS (1)
                                              _id: ObjectId('6821c5d44966e90177b63e0c'),
  Search connections
                                              nome: 'Arena',
                                              citta: 'Verona',
 ▼ 

cluster0.6t3q5ut.mongodb.net
                                              indirizzo: 'Piazza Bra',
    ▼ 2 admin
                                              telefono: '0452020320',
                                              chiusura: 'MAR',
    ▶ 8 config
                                             prezzo: 20
    ▼ S local
        oplog.rs
                                         Atlas atlas-swiyxj-shard-0 [primary] test>
        replset.election
        replset.minvalid
```

Read: Ricerca

- db.museo.find({})
 - Recupera tutti i documenti della collezione museo
- db.museo.find({ <field1>: <value1>, ... })
 - Recupera tutti i documenti della collezione museo che soddisfano il filtro
 - Le condizioni sono in AND.
 - Operatori generici di confronto:
 - { <field1>: { <operator1>: <value1> }, ... }
 - \$in, \$nin
 - \$eq, \$ne
 - \$gt, \$gte, \$lt, \$lte
 - ecc...

Read: Ricerca

Ricerca con valori nulli

Ricerca con valori nulli

- db.museo.find({ item: null });
- db.museo.find({ item: { \$ne : null } });

Read: Ricerca con proiezione

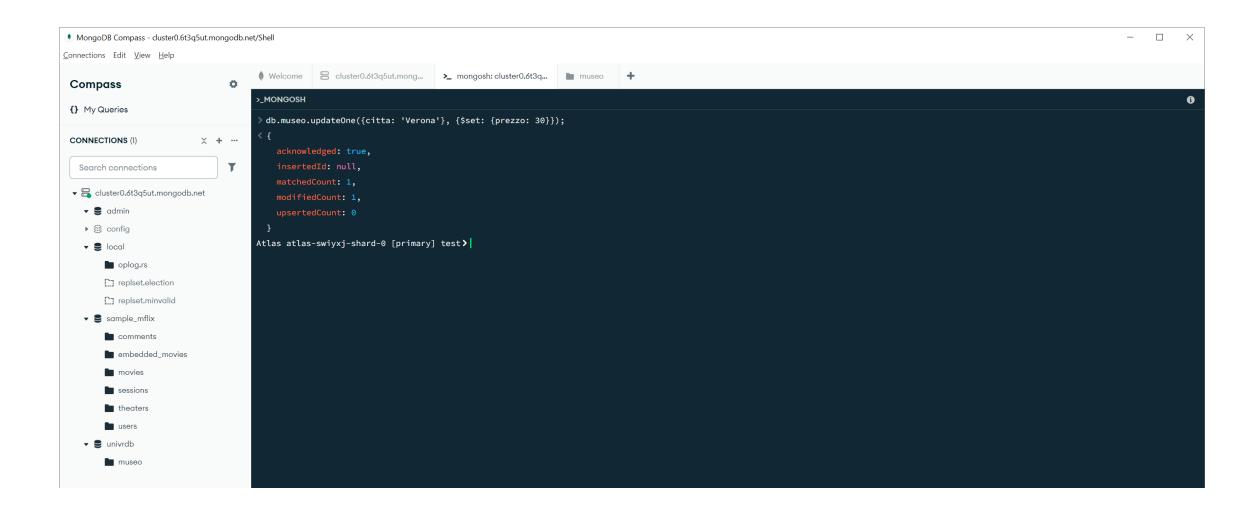
Ricerca con proiezione

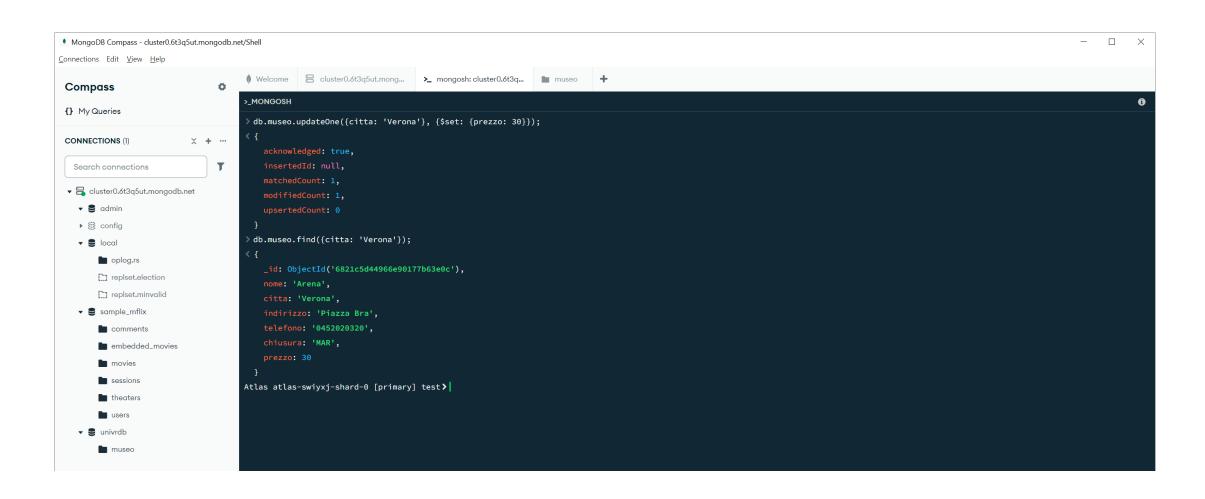
```
    db.museo.find({ status: 'A' })
    .project({ nome: 1, citta: 1, indirizzo: 0, chiusura: 0, prezzo: 0 });
```

- db.museo.find({ status: 'A' }).project({ indirizzo: 0, chiusura: 0, prezzo: 0 });
 - Esclude specifici attribute (valore = 0), default = 1

Aggiornare un documento: updateOne()

- db.museo.updateOne({citta: 'Verona'}, {\$set: {price: 30}});
 - L'oggetto JSON passato come primo parametro serve da filtro
 - \$set: permette di specifare un insieme di campi con il valore aggiornato
 - Ritorna un oggetto JSON che contiene dei metadati:
 - matchedCount: notifica il numero di documenti che soddisfano il filtro
 - modifiedCount: notifica il numero di documenti che sono stati modificati
 - Con updateOne() viene modificato sempre un solo documento → il primo che soddisfa il criterio





Riassunto dei metodi disponibili

- db.collection.updateOne()
- db.collection.updateMany()
- db.museo.replaceOne(<filter>, <replacement>, options)
 - Sostituisce un singolo documento mantenendo inalterato l'ObjectId

```
db.museo.replaceOne(
{ item: 'paper' },
    { item: 'paper',
        instock: [
            { warehouse: 'A', qty: 60 },
            { warehouse: 'B', qty: 40 }
            ]
        });
```

Cancellazione: Delete

Cancellazione di un documento da una collezione:

- db.museo.deleteOne({citta: 'Verona'});
- Parametro: oggetto JSON che permette di selezionare particolari documenti.
- deleteOne() cancella sempre il primo documento che rispetta la condizione di filtro.
- Ritorna un oggetto JSON con dei metadati, tra cui:
 - deletedCount: numero di documenti che sono stati cancellati
- deleteMany()

Schema in Documenti MongoDB

- Il modello relazionale predilige la proprietà di normalizzazione del dato: una stessa informazione è presente una sola volta all'interno della base di dati e le relazioni permettono di creare collegamenti da informazioni (evitare ridondanza!).
- In un database NoSQL è prevista la possibilità di avere ridondanza e strutture de-normalizzate per incrementare le prestazioni e garantire flessibilità (dato non strutturato o semi-strutturato).
- I documenti MongoDB consentono il concetto di incapsulamento: la proprietà di un documento può essere un altro documento.
 - L'incapsulamento riduce la necessità di JOIN e facilita il recupero dell'informazione.
 - La migliore struttura ed il miglior criterio di incapsulamento dipende dai requisiti dell'applicazione (tipico modo di interrogare l'informazione).

Incapsulamento

```
{ "_id": ObjectId("f34303859302"ù),
 "title": "Journey to the Center of the Earth",
 "published_year": 1864,
 "genres": ["Adventures", "Science Fiction"],
 "author": {
           "first_name": "Jules",
           "last_name": "Verne",
           "birth_year": 1828,
           "nationality": "French",
           "biography": "A pioneering author in the science fiction...."
 "rating": 4.5,
 "copies_sold": 1500000
```

Incapsulamento \rightarrow ridondanza Cosa succede se ho un altro libro dello stesso autore?

Riferimento

```
// Course Collection
 "_id": ObjectId("d04940004ncdd"),
 "course_name": "Introduction to Artificial Intelligence",
 "professor_id": ObjectId("4590f394ks"),
 "duration": "6 months",
 "credits": 6
// Professor Collection
 "_id": ObjectId("4590f394ks"),
 "first_name": "Alan",
 "last_name": "Turing",
 "department": "Computer science",
 "publications": 45
```

- Attraverso i riferimenti è possibile creare dei collegamenti tra i dati tramite «link» o «puntatori».
- Le applicazioni risolveranno questi riferimenti andando a recuperare i dati relativi.

Denormalizzazione

```
// Course Collection
 "_id": ObjectId("d04940004ncdd"),
 "course_name": "Introduction to Artificial Intelligence",
 "professor_id": ObjectId("4590f394ks"),
 "professor_name": "Alan Turin",
 "duration": "6 months",
 "credits": 6
// Professor Collection
 "_id": ObjectId("4590f394ks"),
 "first name": "Alan",
 "last_name": "Turing",
 "department": "Computer science",
 "publications": 45
```

 Approccio a metà tra incapsulamento e riferimento che consente di velocizzare l'accesso a certe informazioni.

MongoDB Formato e tipi di dato

- MongoDB memorizza le informazioni nel formato BSON (Binary JSON)
 che è del tutto simile a JSON ma offre una più ampia varietà di tipi base
 e delle ottimizzazioni per la memorizzazione e la ricerca.
- Tipi di dato supportati da BSON:
 - "_id": ObjectId("407f1f77bcf86cd799439012"),
 - "doubleField": 55.4,
 - "stringField": "Hello, world!",
 - "objectField": {"name": "John", "age": 34},
 - "arrayField": [1,2,3,4,5],
 - "binaryField": BinData(0, "binary-content"),
 - "booleanField": true,
 - "dataField": ISODate("2024-05-12T10:00:00Z"),

MongoDB Formato e tipi di dato

- "nullField": null,
- "regexField": /^abc/,
- "javascriptField": Code("function() { print('success!'); }"),
- "javascriptWithScopeField": Code("return this.x + this.y; {x: 3, y: 5}"),
- "intField": 42,
- "timestampField": Timestamp(1601410400, 1),
- "int64Field": NumberLong(9876543210),
- "decimal128Field": NumberDecimal("0.123456789"),
- "minKeyField": MinKey(),
- "maxKeyField": MaxKey(),

MongoDB: Aggregazioni

 Le aggregazioni sono uno strumento che consente di combinare più trasformazioni e interrogazioni, definendo una pipeline di processamento.

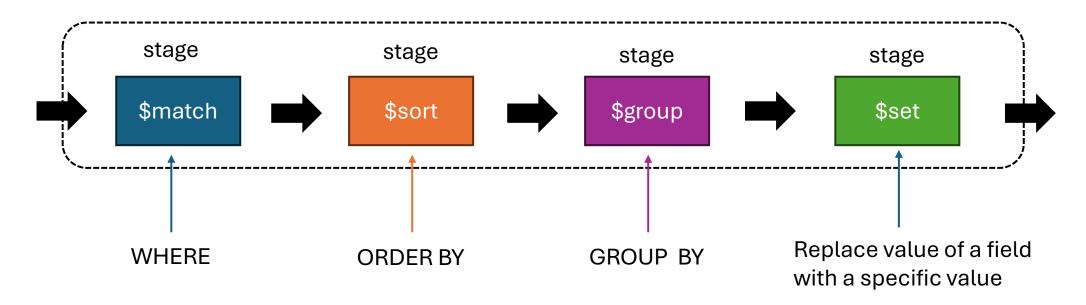
```
{
    "_id": ObjectId( "059440..."),
    "student_id": ObjectId( "4d034930..."),
    "course_id": ObjectId( "2f402002...",
    "grade": 76,
    "semester": "Fall 2023",
}

db.student_grades.aggregate([
    {$match: { semester: "Fall 2023" }},
    {$group: {_id: "student_id"}},
    {$project:{ $divide: [$totalGrade", $totalCourses]}}
])
```

Calcola la media per studente

MongoDB: Aggregazioni

• In una pipeline di trasformazione, ad ogni passo viene passato all'operazione corrente un insieme di documenti, mentre l'insieme ottenuto dalla trasformazione è passato all'operazione successiva.



MongoDB: Aggregazioni (operatori)

SQL clause, term, function	MongoDB aggregation operators
WHERE	\$match
GROUP BY	\$group
HAVING	\$match
SELECT	\$project
LIMIT	\$limit
OFFSET	\$skip
ORDER BY	\$ort
SUM()	\$sum
COUNT()	\$sum AND \$sortByCount
JOIN	\$lookup
SELECT INTO NEW_TABLE	\$out
MERGE INTO TABLE	\$merge
UNION ALL	\$unionWith

MongoDB: Aggregazioni (operatori)

SQL clause, term, function	MongoDB aggregation operators
WHERE	\$match
GROUP BY	\$group
HAVING	\$match
SELECT	\$project
LIMIT	\$limit
OFFSET	\$skin
Esistono altri operatori: consultare la documentazione!	
SUM()	\$sum
COUNT()	\$sum AND \$sortByCount
JOIN	\$lookup
SELECT INTO NEW_TABLE	\$out
MERGE INTO TABLE	\$merge
UNION ALL	\$unionWith

MongoDB: Join

• L'operatore di aggregazione \$lookup permette di eseguire un left outer join rispetto ad una collezione della stessa base di dati.

```
$lookup:
  localField: <field from the input documents>, ---- Attributo del documento corrente da usare per il (equi) JOIN
  foreignField: <field from the documents of the "from" collection Attributo del documento esterno da usare per il JOIN
  let: { <var_1>: <expression>, ..., <var_n>: <expression> },
  pipeline: [ <pipeline to run> ],
  as: <output array field> ————— Nome dell'attributo che conterrà l'elenco dei documenti collegati
```

MongoDB: Join

```
db.customers.aggregate([
   $lookup: {
    // The collection to join.
    from: "orders".
    // The field from the customers collection.
    localField: "_id",
    // The field from the orders collection.
    foreignField: "customerld",
    // The output array containing joined documents.
    as: "orders"
```

```
_id": "cust100",
"email": "jane@example.com", 
"name": "Jane Doe",
"orders":
     _id": "ord200",
    "customerld": "cust100",
    "itemIds": |
       "item300".
       "item301"
    " _id": "ord202",
    "customerld": "cust100",
    "itemIds": [
       "item301"
       "item303"
```