



Semantix

MongoDB - Básico

Aula 5

Quem sou eu?

Eu sou Rodrigo Augusto Rebouças.

Engenheiro de dados da Semantix
Instrutor do Semantix Academy

Você pode me encontrar em:
rodrigo.augusto@semantix.com.br





Agregação



Agregações MongoDB

- Maneiras de realizar agregações no MongoDB
 - Agregação de propósito único
 - Mais simples, porém limitadas
 - Função map-reduce
 - Forma tradicional de realizar agregações em Big Data
 - Faz uso de funções de JavaScript
 - Mais opções de manipulação de dados
 - Difícil a implementação
 - Pipeline de agregação
 - Melhor desempenho que map-reduce
 - MongoDB adiciona operadores novos a cada versão
 - MongoDB 4.4 - \$accumulator
 - \$function

Agregação Propósito único

- Agregam documentos de uma única coleção
 - Operações de agregações simples
 - Menos flexível
- Métodos
 - Count
 - Distinct
 - EstimatedDocumentCount
 - Não usa um filtro de consulta
 - Usa metadados para retornar a contagem de uma coleção

Agregação Propósito único

○ Count

- Sintaxe:

- `db.<nomeCollection>.count()`
- `db.<nomeCollection>.count(<filtro>) = db.<nomeCollection>.find({ <filtro> }).count()`

- Ex.

- `db.cliente.count()`
- `db.cliente.count({cidade: "São Paulo"})`
- `db.cliente.find({cidade: "São Paulo"}).count()`

○ Distinct

- Sintaxe: `db.<nomeCollection>.distinct("<atributo>")`

- Ex.: `db.cliente.count({nome})`

○ EstimatedDocumentCount

- `db.<nomeCollection>.estimatedDocumentCount()`

Agregação Pipeline

- Agrupar valores de vários documentos e executar operações nos dados agrupados para retornar um único resultado
- Sintaxe:
 - ```
db.<nomeCollection>.aggregate([
 {$<estagio>:<parâmetros>},
 {$<estagio>::<operadoresExpressão>},
 {$<estagio>:<parâmetro>, <operadoresExpressão>}
])
```
- Estágios
  - <https://docs.mongodb.com/manual/reference/operator/aggregation-pipeline/>
- Operadores de expressão
  - <https://docs.mongodb.com/manual/reference/operator/aggregation/#aggregation-expression-operators>

# Agregação Pipeline – Estágios

- \$addFields
- \$bucket
- \$bucketAuto
- \$collStats
- \$count
- \$currentOp
- \$facet
- \$geoNear
- \$graphLookup
- \$group
- \$indexStats
- \$limit
- \$listLocalSessions
- \$listSessions
- \$lookup
- \$match
- \$merge
- \$out
- \$planCacheStats
- \$project
- \$redact
- \$replaceRoot
- \$replaceWith
- \$sample
- \$set
- \$skip
- \$sort
- \$sortByCount
- \$unionWith
- \$unset
- \$unwind



# Agregação Pipeline – Operadores de Expressão

- Operadores de expressão aritmética
- Operadores de expressão de matriz
- Operadores de expressão booleana
- Operadores de expressão de comparação
- Operadores de expressão condicional
- Operadores de expressão de agregação personalizados
- Operadores de tamanho de dados
- Operadores de expressão de data
- Operador de expressão literal
- Operadores de expressão de objeto
- Definir operadores de expressão
- Operadores de expressão de string
- Operador de Expressão de Texto
- Operadores de expressão de trigonometria
- Operadores de expressão de tipo
- Acumuladores ( \$group)
- Acumuladores (em outras fases)
- Operadores de expressão variável

# Agregação Match

- Match

- Sintaxe:

- `db.<nomeCollection>.aggregate([  
 { $match: { <query> }  
}]`

- Exemplo:

```
db.funcionarios.aggregate([
 { $match: { status: "Ativo" } }
])
```

# Agregação Group

## ○ Group

- Sintaxe:

- db.<nomeCollection>.aggregate(  
 {\$group:{  
 \_id: "\$<atributo>",  
 <atributoNovo>: { <accumulator1> : "\$<atributoAgg>"},  
 ...  
 }  
})

## ○ Exemplo:

```
db.funcionarios.aggregate(
 { $group: { _id: "$setor", total: { $sum: "$vendas" } } }
)
```

## ○ Atributos

- \_id

- Group By (SQL)

- "" - All

## ○ Acumulador

- \$sum

- \$push

- \$avg

- \$addToSet

- \$min

- \$stdDevPop

- \$max

- \$stdDevSamp

- \$first

- \$last

# Agregação Match e Group

- Pipeline de agregações

- Match e Group

```
db.funcionarios.aggregate([
 { $match: { status: "Ativo" } },
 { $group: {
 _id: "$setor",
 total: { $sum: "$vendas" } }
]})
```

# Agregação Match, Group, Sort e limit

- Pipeline de agregações

- Match e Group

```
db.funcionarios.aggregate([
 { $match: { status: "Ativo" } },
 { $group: {
 _id: "$setor",
 total: { $sum: "$vendas" },
 media: { $avg: "$vendas" },
 quantidade: { $count: 1 }
 },
 { $sort: { _id: 1 } },
 { $limit: 10 }
])
```

# Exercícios – Agregação I

1. Crie o banco de dados escola
2. Crie a collection alunos no banco de dados escola
3. Importe o arquivo “dataset\alunos.csv” para a collection alunos, com os seguintes atributos:
  - id\_discente: Number
  - nome: String
  - ano\_ingresso: Number
  - nivel: String
  - id\_curso: Number
4. Visualizar os valores únicos do “nivel” de cada “ano\_ingresso”
5. Calcular a quantidade de alunos matriculados por cada “id\_curso”
6. Calcular a quantidade de alunos matriculados por “ano\_ingresso” no "id\_curso": 1222
7. Visualizar todos os documentos do “nivel”: “M”
8. Visualizar o último ano que teve cada curso (id\_curso) dos níveis “M”
9. Visualizar o último ano que teve cada curso (id\_curso) dos níveis “M”, ordenados pelos anos mais novos de cada curso
10. Visualizar o último ano que teve os 5 últimos cursos (id\_curso) dos níveis “M”, ordenados pelos anos mais novos



# Agregação Lookup

- Lookup

- Left outer join

- Sintaxe:

- `db.<nomeCollection>.aggregate(  
 { $lookup: {  
 from: <collectionJoin>  
 localField: <atributoJoinLocal>  
 foreignField: <atributoJoinFrom>  
 as: <nomeArraySaida>  
 }  
}`

- <https://docs.mongodb.com/manual/reference/operator/aggregation/lookup/>

- Exemplo:

- `db.funcionario.aggregate([  
 { $lookup: {  
 from: "vendas"  
 localField: "cod_func"  
 foreignField: "cod_func"  
 as: "vendasFuncionario"  
 },  
 { $project: { "_id": 0, "cod_func": 1,  
 "vendasFuncionario.cod_cliente": 1 }  
 }  
])`

# Exercícios – Agregação II

1. Crie a collection cursos no banco de dados escola
2. Importe o arquivo “dataset\cursos.csv” para a collection cursos, com os seguintes atributos:
  - id\_curso: Number
  - id\_unidade: Number
  - nome: String
  - nivel: String
3. Realizar o left outer join da collection alunos e cursos, quando o id\_curso dos 2 forem o mesmo.
4. Realizar o left outer join da collection alunos e cursos, quando o id\_curso dos 2 forem o mesmo e visualizar apenas os seguintes campos
  - Alunos: id\_discente, nivel
  - Cursos: id\_curso, id\_unidade, nome



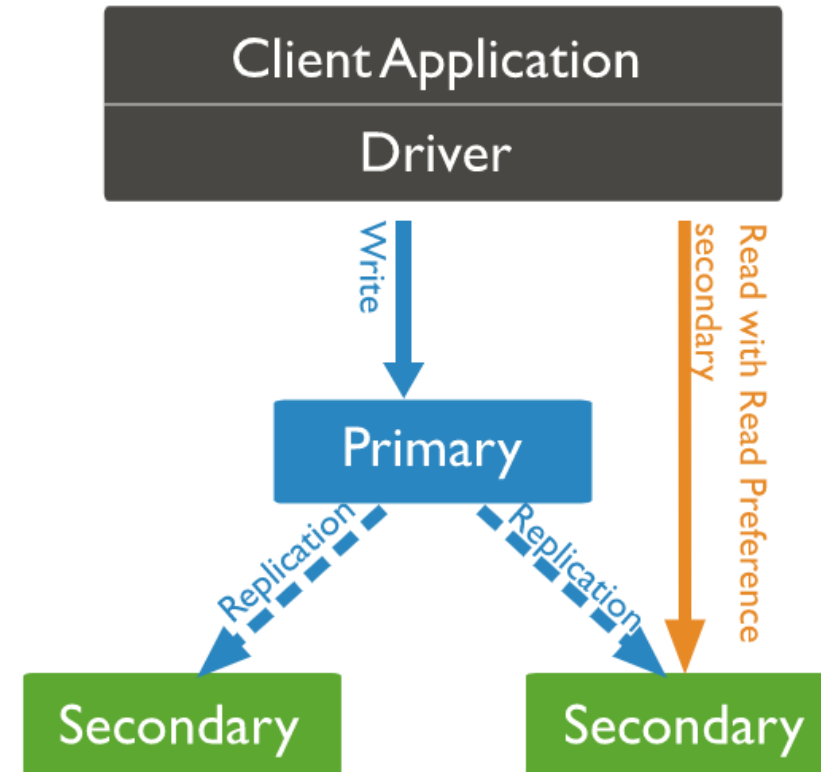
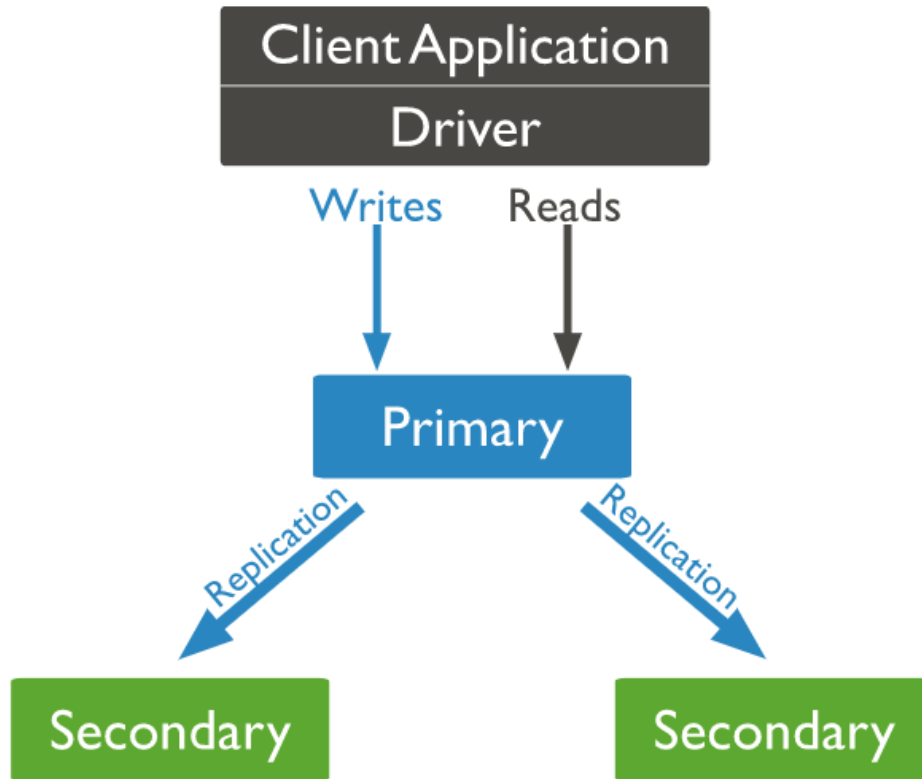
## Replicação e Shard

# Replicação MongoDB

- Ambientes de produção
  - Redundância e alta disponibilidade
  - Tolerância a Falhas
- Denominado como Replica Set
  - Grupo de instâncias mongod (nós) que mantém o mesmo conjunto de dados
  - Nós
    - 1 Primário
      - Recebe as solicitações de escrita e leitura
      - Registrar todas as alterações em seus datasets em seu log de operação (oplog)
    - N Secundários
      - Réplica do oplog do nó primário
      - Ganhar desempenho na leitura de dados
    - Distribuir solicitações de leitura nos secundários

# Replicação MongoDB

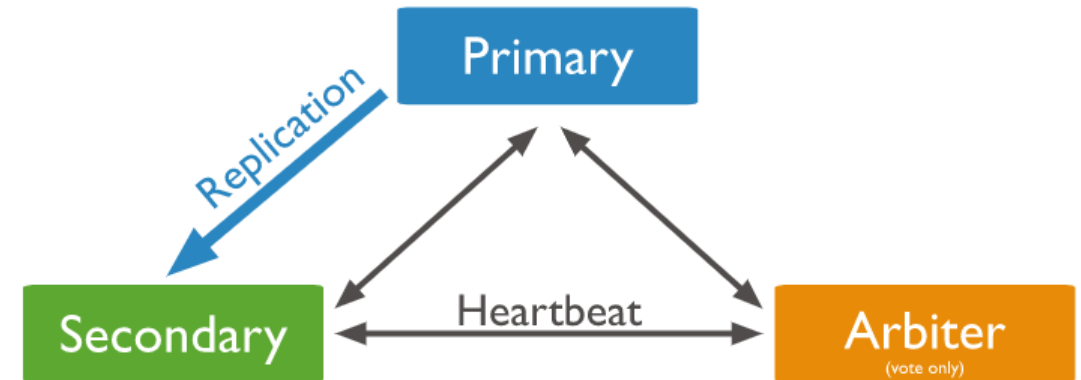
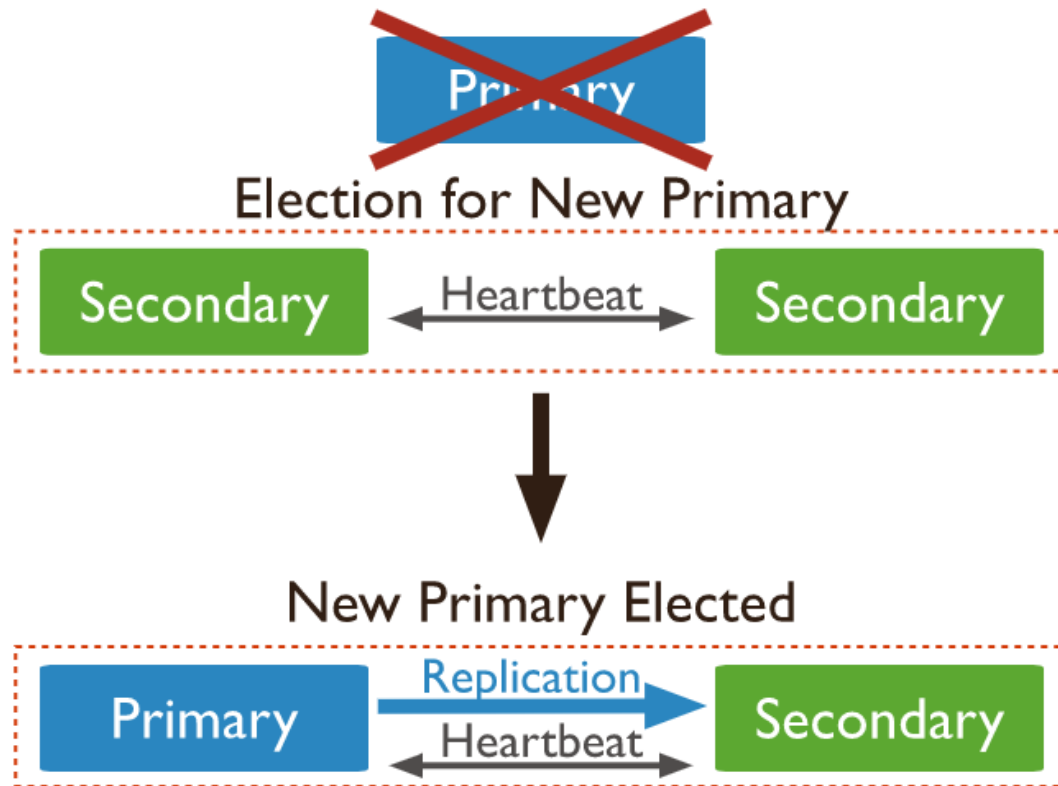
- Replica Set



<https://docs.mongodb.com/manual/replication/>

# Replicação MongoDB

- Falha no Nó primário
  - Não pode processar operações de gravação



<https://docs.mongodb.com/manual/replication/>



# Replicação MongoDB

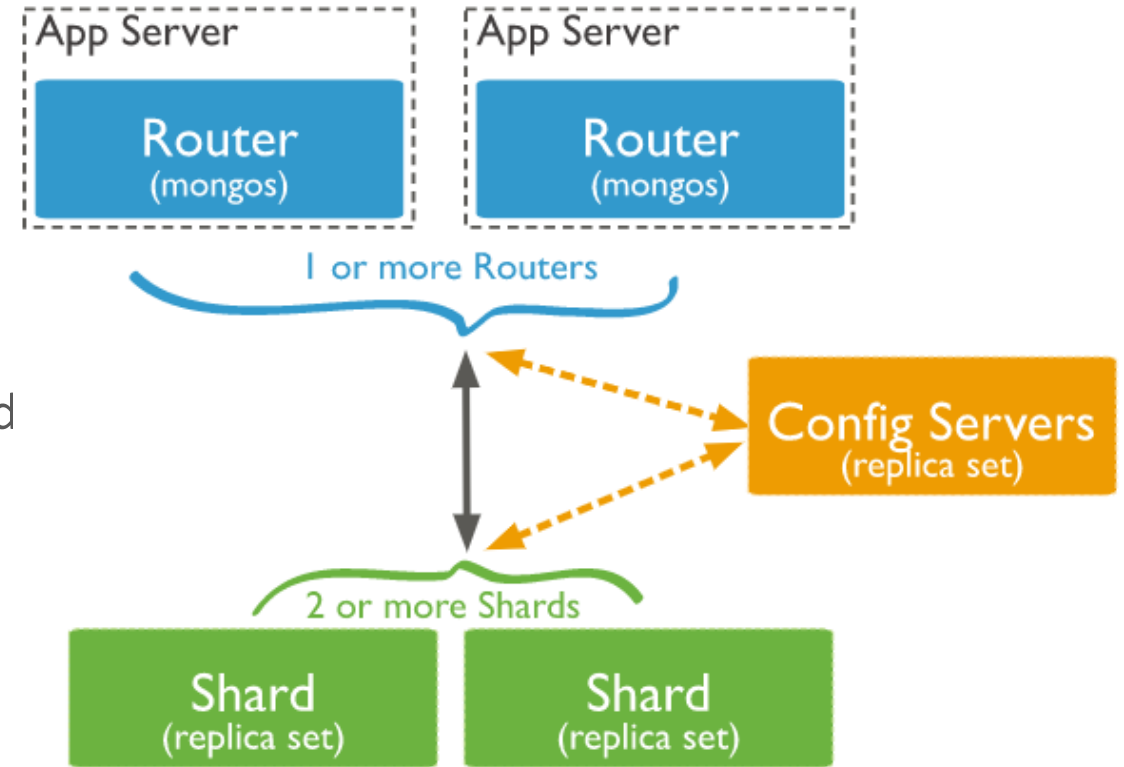
- Visualizar configuração do Replica Set
  - `rs.conf()`
- Verificar o status do Replica Set
  - `rs.status()`

```
> rs.status()
{
 "ok" : 0,
 "errmsg" : "not running with --replSet",
 "code" : 76,
 "codeName" : "NoReplicationEnabled"
}
```

- Configuração
  - <https://docs.mongodb.com/manual/tutorial/deploy-replica-set/>

# Shard MongoDB

- MongoDB fragmenta dados no nível de collection
  - Distribui os dados entre os shards no cluster
- Arquitetura
  - Shard
    - Contém um subconjunto dos dados fragmentados
    - Pode ser implementado como um replica set
    - Escalabilidade Horizontal
  - Config Server
    - Armazena os metadados
    - Definições de configuração do cluster
  - Mongos
    - Interface entre o aplicação cliente e o sharded cluster





# MongoDB Atlas



# MongoDB Atlas

- Banco de dados em nuvem global
- DBaaS
  - Database as a Service

## Cloud Provider & Region














Choose your preferred cloud provider and the region nearest to clients

Select a cloud provider to see its region availability.



Configure a **free tier cluster** by first selecting a region labeled with **FREE TIER AVAILABLE** then choose the M0 option in the Cluster Tier below.

★ recommended region ⓘ

| NORTH AMERICA                                                                                                                                                            |                                                                                                                                                                              | EUROPE                                                                                                                                         | ASIA |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <div><div></div><div>N. Virginia (us-east-1) ★</div><div>FREE TIER AVAILABLE</div></div> | <div><div></div><div>Ireland (eu-west-1) ★</div></div>                                      | <div><div></div><div>Tokyo (ap-northeast-1)</div></div>       |      |
| <div><div></div><div>Ohio (us-west-1) ★</div></div>                                    | <div><div></div><div>London (eu-west-2)</div></div>                                       | <div><div></div><div>Seoul (ap-northeast-2)</div></div>     |      |
| <div><div></div><div>N. California (us-west-1)</div></div>                             | <div><div></div><div>Frankfurt (eu-central-1) ★</div><div>FREE TIER AVAILABLE</div></div> | <div><div></div><div>Singapore (ap-southeast-1)</div></div> |      |
| <div><div></div><div>Oregon (us-west-2) ★</div></div>                                  |                                                                                                                                                                              | <div><div></div><div>Mumbai (ap-south-1)</div></div>        |      |
| <div><div></div><div>Montreal (ca-central-1)</div></div>                               | SOUTH AMERICA                                                                                                                                                                |                                                                                                                                                |      |
|                                                                                                                                                                          | <div><div></div><div>São Paulo (sa-east-1)</div></div>                                    |                                                                                                                                                |      |

- Modos de uso do Cluster
  - Compartilhado
    - 512 MB de armazenamento (Grátis)
  - Dedicado
  - Dedicado com multiregiões
- Acesso:
  - <https://www.mongodb.com/cloud/atlas>





# Certificações



# Certificações MongoDB

- MongoDB University
  - <https://university.mongodb.com/certification>



C100DEV: MongoDB Certified Developer,  
Associate Level



C100DBA: MongoDB Certified DBA,  
Associate Level





## MongoDB for Developers Learning Path

6 courses

Start by learning the fundamentals of MongoDB. Then, build and deploy an application leveraging any MongoDB driver.

- >> Start with MongoDB fundamentals
- >> Set up your servers with basic administration
- >> Analyze your data with MongoDB Aggregation Framework
- >> Build your first app with a MongoDB Driver (in Python, Java, or JavaScript)
- >> Optimize the performance of your deployment
- >> Learn MongoDB's data modeling methodology

[Get Started](#) [Learn More](#)



## MongoDB for DBAs Learning Path

5 courses

Learn MongoDB administration fundamentals. Then securely manage your cluster and debug and optimize the performance of your deployment.

- >> Start with MongoDB fundamentals
- >> Set up your servers with basic administration
- >> Optimize the performance of your deployment
- >> Secure your cluster
- >> Debug your deployment

[Get Started](#) [Learn More](#)

<https://university.mongodb.com/courses/catalog>

## Practice Exams

Preparing for an exam can be daunting. Exam takers typically prepare by reviewing the MongoDB documentation, taking a free online course from MongoDB University, or reviewing other materials that we make available.

Our Practice Exams make this process simpler. Build confidence by familiarizing yourself with the subject areas and format of the certification exam. After completing your practice exam, identify knowledge gaps by taking a look at your incorrect answers. Each answer has detailed explanations with links to relevant resources.

The Practice Exams were written by the same MongoDB experts who build our Certification Exams.

Register Now

Developer Practice Exam

DBA Practice Exam

<https://university.mongodb.com/certification/exam-prep>



# Semantix

## Obrigado!

Alguma pergunta?



Você pode me encontrar em:  
[rodrigo.augusto@semantix.com.br](mailto:rodrigo.augusto@semantix.com.br)

**GET SMARTER**