

01MBID Fundamentos de la tecnología Big Data



viu

Universidad
Internacional
de Valencia

Sesión 3

Tema 2 + Actividad

De:



Planeta Formación y Universidades

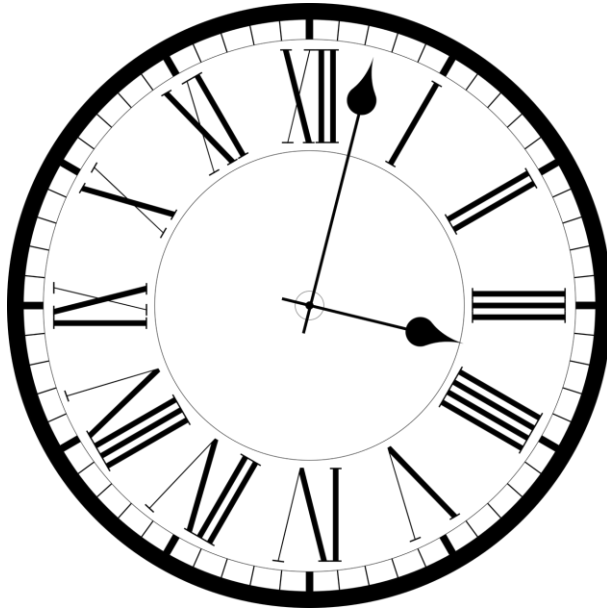
> Agenda

- **Dudas**
- **Tema 2 2de2: Fuentes de datos en entornos Big data**
- **Actividad**

> Agenda

- **Dudas**
- **Tema 2 2de2: Fuentes de datos en entornos Big data**
- **Actividad**





> Agenda

- Dudas
- **Tema 2 2de2: Fuentes de datos en entornos Big data**
- **Actividad**

> Fuentes de datos en entornos Big data

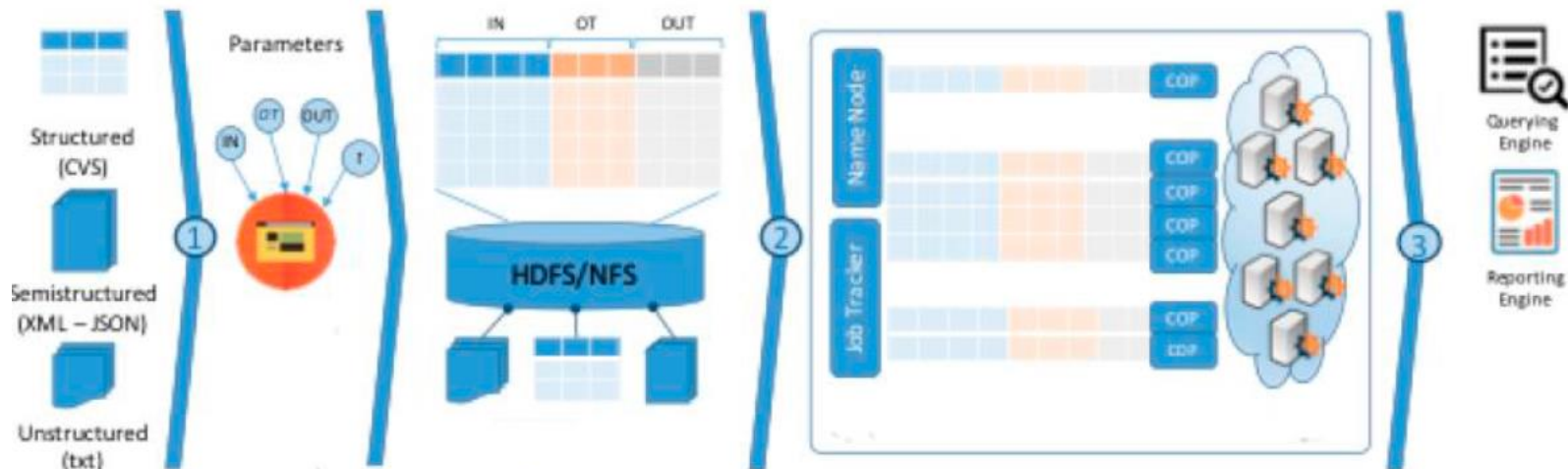
- 1) ¿Qué es una fuentes de datos?
- 2) Diferencias respecto a las tecnologías de datos tradicionales
- 3) Tipos de datos y flujo de datos**
- 4) Calidad de Datos**
- 5) Las V's del Big Data**

> Tipos de datos y flujos de datos

- Estructurados
 - formato + finitos
 - ejemplo, expediente académico
- No Estructurados
 - sin formato + indeterminado + herramientas específicas
 - ejemplo, video
- Semiestructurados
 - marcadores o organización interna
 - ejemplo, XML

> Tipos de datos y flujos de datos

Ejemplo:



> Fuentes de datos en entornos Big data

- 1) ¿Qué es una fuentes de datos?
- 2) Diferencias respecto a las tecnologías de datos tradicionales
- 3) Tipos de datos y flujo de datos
- 4) Calidad de Datos**
- 5) Las V's del Big Data**



Calidad de datos

ISO 25012:2008

Grado en que las características de los datos satisfacen las necesidades explícitas e implícitas cuando se usan bajo condiciones específicas.

“Calidad adecuación al uso”



Calidad de datos

Importancia en Big Data

- Objetivo: definir para qué y por qué se quieren los datos.
¿Qué **valor** me aporta?
- **Dimensiones de Calidad:** disponibilidad, usabilidad, confiabilidad, pertinencia y presentación.



Calidad de datos

Importancia en Big Data

Condición previa para el análisis y el uso de Big Data y para garantizar el valor de esos datos.



Calidad de datos

- Disponibilidad
- Usabilidad
- Confiabilidad
- Pertinencia
- Presentación



Disponibilidad

Accesibilidad:

- Interfaz de acceso a datos
- Los datos públicos o fáciles de adquirir

Oportunidad:

- Rango de tiempo, para que sean útiles
- Actualizados “regularmente”
- Intervalo de tiempo entre la recopilación, procesamiento y uso cumple los requisitos



Usabilidad

Credibilidad:

- Fuente de los datos: organizaciones especializadas de un país, campo o industria
- Expertos o especialistas auditan regularmente y comprueban la exactitud de los datos
- Los datos existen en el rango de valores conocidos o “aceptables”



Confiabilidad

Exactitud:

- Los datos proporcionados son precisos
- La representación de datos (o valor) refleja bien el estado real de la información de origen
- La representación de información (datos) no causara ambigüedad

Consistencia:

- Después de procesar los datos, sus conceptos, dominios de valor y formatos coinciden con los originales
- **Durante un cierto tiempo**, los datos permanecen consistentes
- **Durante un cierto tiempo**, los datos permanecen verificables



Confiabilidad

Integridad:

- El formato de los datos es claro y cumple los criterios
- Los datos son consistentes con la integridad estructural
- Los datos son consistentes con la integridad del contenido

Complejidad:

- Si una deficiencia de **un componente** afectara el uso de los datos para datos con componentes múltiples
- Si una deficiencia de **un componente** afectara la precisión y la integridad de los datos



Pertinencia

Conveniencia:

- Los datos recogidos no coinciden completamente con el tema, pero exponen un aspecto parcial del mismo.
- La mayoría de los conjuntos de datos recuperados están dentro de la temática de recuperación que los usuarios necesitan
- La temática de la información proporciona coincidencias con la temática de recuperación de los usuarios



Presentación

Legibilidad:

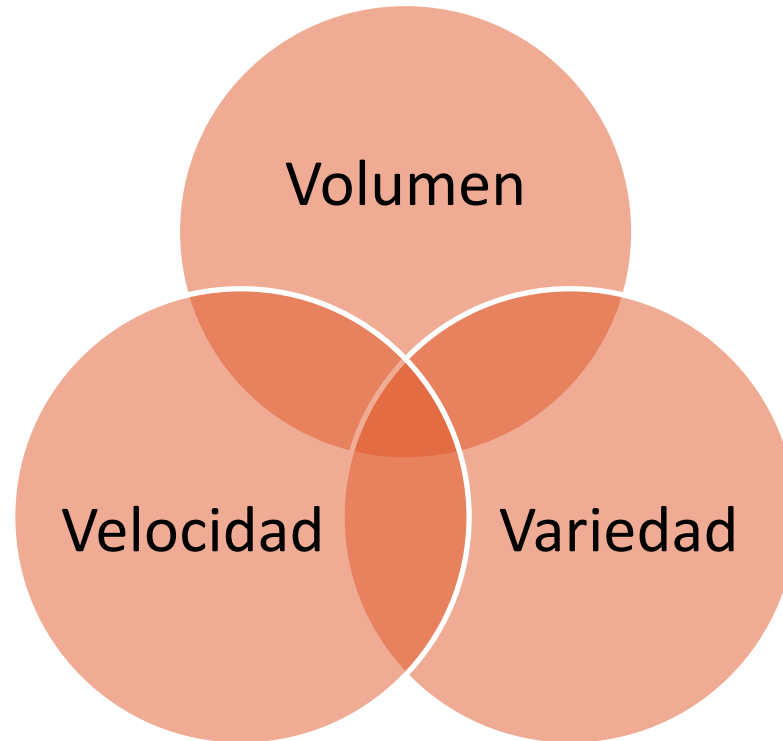
- Los datos (contenido, formato, etc.) son claros y comprensibles
- Es fácil juzgar que los datos facilitados satisfacen las necesidades
- La descripción de los datos, la clasificación y el contenido de codificación satisfacen la especificación y son fáciles de entender

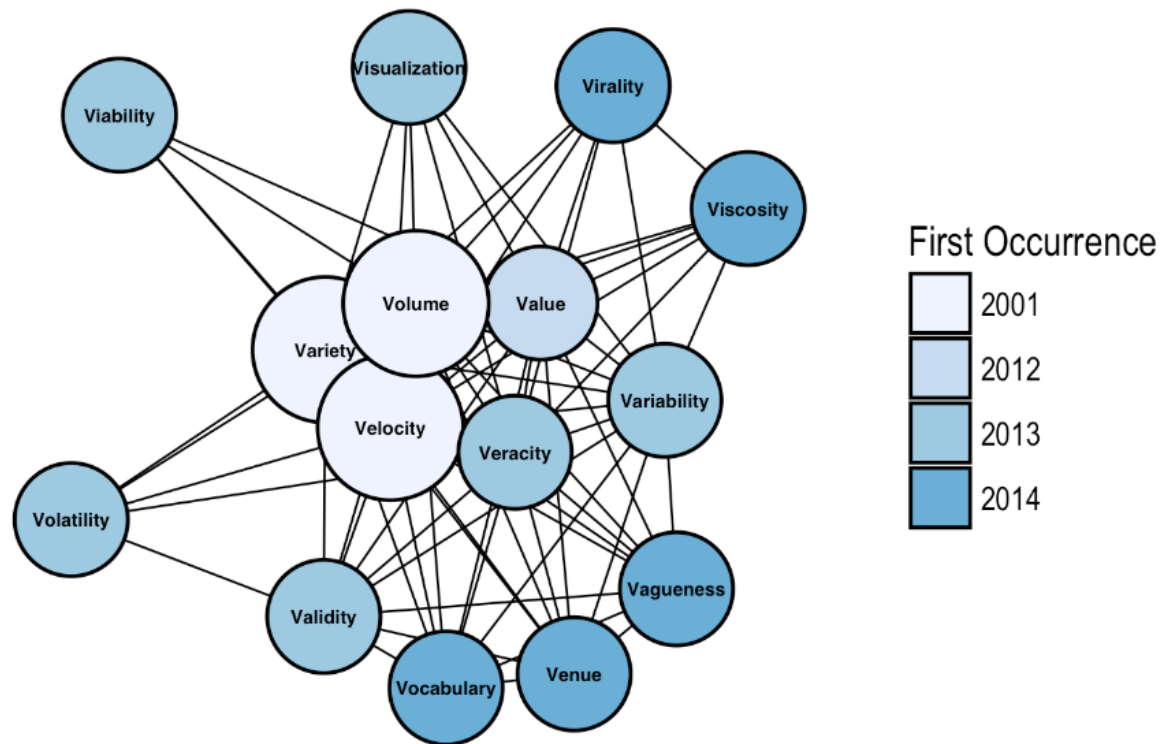
> Fuentes de datos en entornos Big data

- 1) ¿Qué es una fuentes de datos?
- 2) Diferencias respecto a las tecnologías de datos tradicionales
- 3) Tipos de datos y flujo de datos
- 4) Calidad de Datos
- 5) Las V's del Big Data**

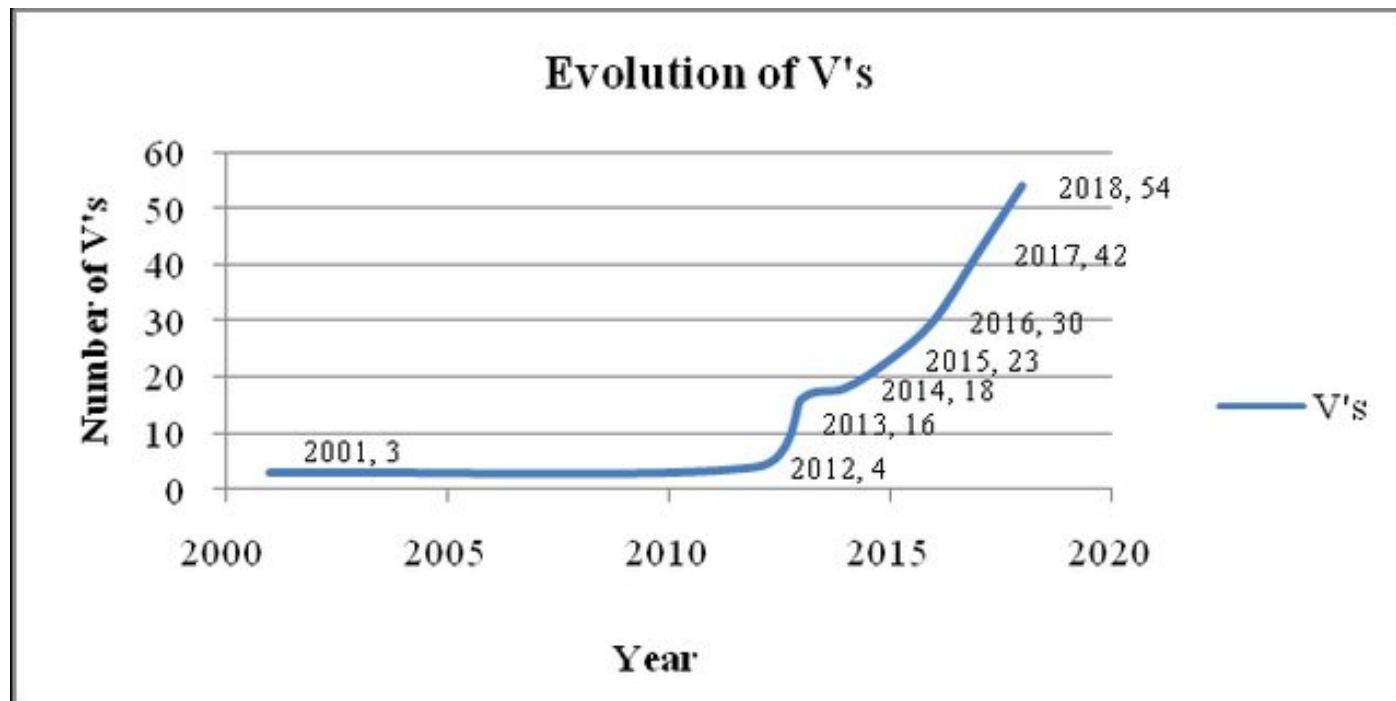


Las V's del Big Data





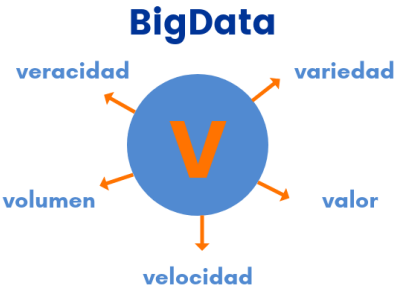
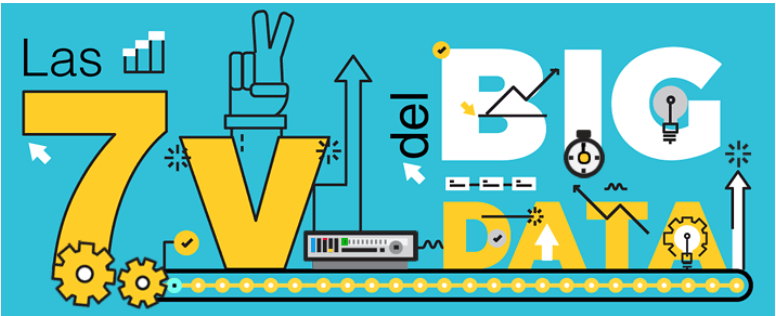
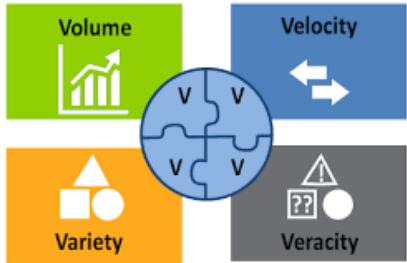
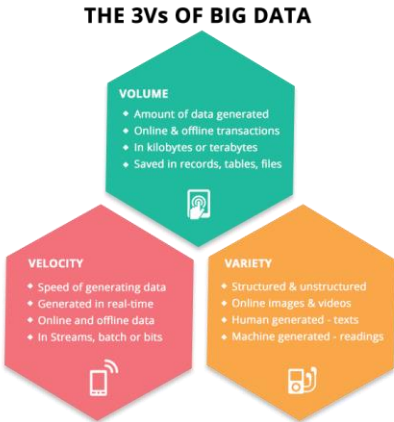
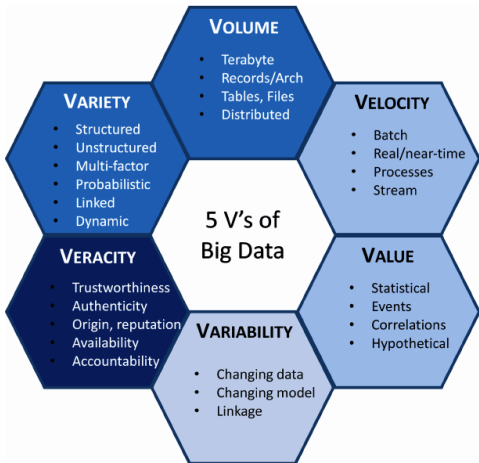
* <https://www.elderresearch.com/blog/42-v-of-big-data>



* https://www.researchgate.net/publication/335619326_Unlock_Different_V's_of_Big_Data_for_Analytics



Las V's del Big Data



> Información sobre las V's

Infografía Big Data: las 7 V

<http://www.iic.uam.es/innovacion/big-data-infografia-7-v/>

<https://www.youtube.com/watch?v=S81frJYbjcE>

Big Data Technology with 8 V's

<https://www.m-brain.com/technology/>

The 10 Vs of Big Data

<https://tdwi.org/articles/2017/02/08/10-vs-of-big-data.aspx>

> Información sobre las V's

Las siete 'V' del Big Data

<https://bbvaopen4u.com/es/actualidad/las-siete-v-del-big-data>

The 42 V's of Big Data and Data Science

* <https://www.elderresearch.com/blog/42-v-of-big-data>

<https://www.kdnuggets.com/2017/04/42-vs-big-data-data-science.html>

Las 10 V's del Big Data

<https://www.grupo-novatech.com/las-10-vs-del-big-data/>

How to Ensure the Validity, Veracity, and Volatility of Big Data

<https://www.dummies.com/programming/big-data/engineering/how-to-ensure-the-validity-veracity-and-volatility-of-big-data/>

> Información sobre las V's

The Eight V's of Supercomputing and Big Data

<https://www.nimbix.net/eight-vs-supercomputing-big-data>

BIG DATA AND FIVE V'S CHARACTERISTICS

https://www.researchgate.net/publication/332230305_BIG_DATA_AND_FIVE_V'S_CHARACTERISTICS

Big Data: el futuro a través de los datos

http://www.upm.es/sfs/Rectorado/Gabinete%20del%20Rector/Revista%20UPM/NUMERO_31_VER.pdf

Los Big Data: Conceptos relacionados y algunas aplicaciones en pediatría

https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0370-41062019000400376#B13

> Información sobre las V's

Grandes datos, grandes desafíos para las ciencias sociales

http://www.scielo.org.mx/scielo.php?pid=S0188-25032018000200415&script=sci_arttext

The 23Vs of Big Data

<https://hackernoon.com/the-23vs-of-big-data-c9146716e0cb>

The five V's of big data

<https://www.bbva.com/en/five-vs-big-data/>

¿Cuáles son las 5 V's del Big Data?

<https://www.iebschool.com/blog/5-vs-del-big-data/>

> Información sobre las V's

WHAT ARE THE 7 V'S OF BIG DATA?

<https://www.yourtechdiet.com/blogs/7vs-big-data/>

LAS 10 V'S DEL BIG DATA

<https://www.datahack.es/blog/big-data/10-vs-del-big-data/>

...

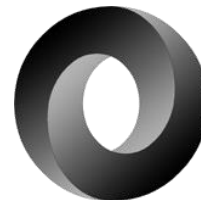
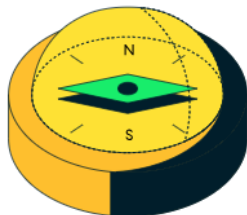


> Agenda

- Dudas
- Tema 2 2de2: Fuentes de datos en entornos Big data
- **Actividad**

> Actividad

MongoDB + Python + JSON



Conceptos clave

-MongoDB

-JSON

**MongoDB**

Muy grande

**mongoDB (Humongous)**

- SGBD Orientada a documentos
- Maneja "*colecciones*" similares a tablas

```
{
  Nombre: "Pedro",
  Apellidos: "Martínez Campo",
  Edad: 22,
  Aficiones: ["fútbol", "tenis", "ciclismo"],
  Amigos: [
    {
      Nombre: "Maria",
      Edad: 22
    },
    {
      Nombre: "Luis",
      Edad: 28
    }
  ]
}
```

```
{
  Nombre: "Luis",
  Estudios: "Administración y Dirección de Empresas",
  Amigos: 12
}
```


```
db.Clientes.find({Nombre:"Pedro"});
```



MongoDB

Documento

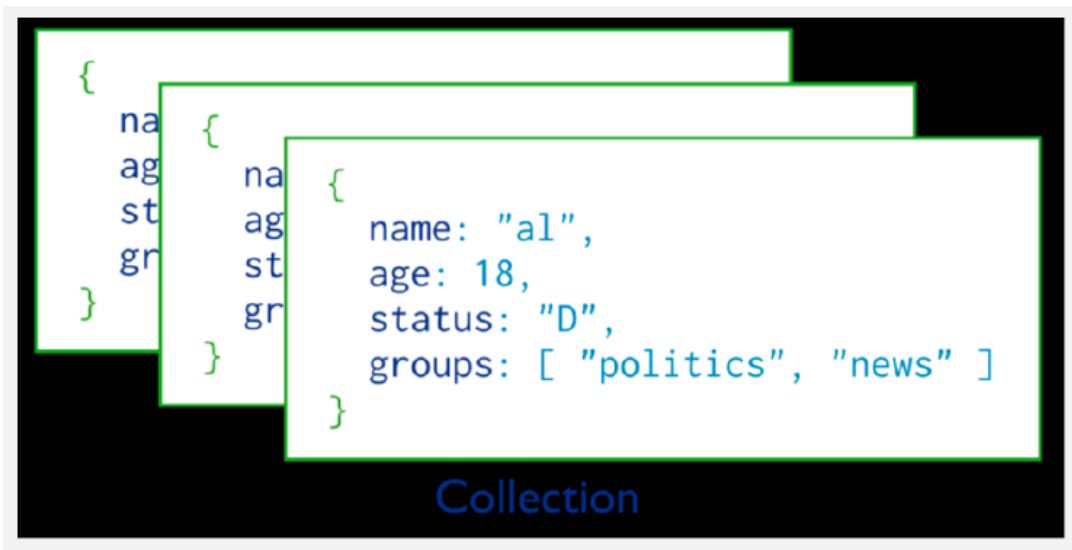
```
{  
  name: "sue",  
  age: 26,  
  status: "A",  
  groups: [ "news", "sports" ]  
}
```





MongoDB

Colección



Conjunto de documentos no necesariamente con la misma estructura

> MongoDB

Modelo Relacional	MongoDB
Tabla	Colección
registro	documento
columna	campo

<https://www.mongodb.com/docs/manual/reference/sql-comparison/#sql-to-mongodb-mapping-chart>



MongoDB

- ❑ MongoDB almacena internamente los datos en BSON, abreviatura de Binary-JSON.
- ❑ JSON es *Java Script Object Notation*, una notación estándar para el intercambio de datos.
- ❑ JSON especifica una gramática mediante expresiones regulares.



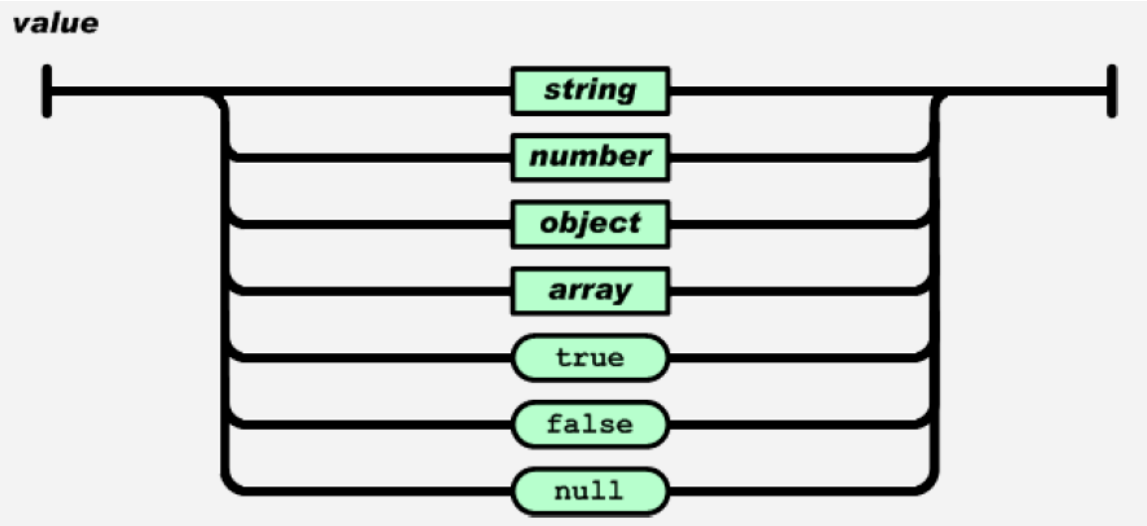
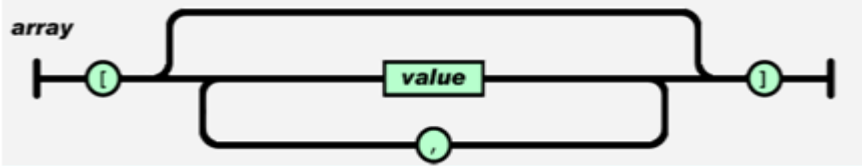
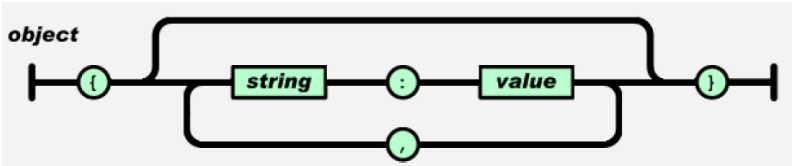


MongoDB JSON

```
object ::= '{ ' members' }' *  
members ::=  
    pair | pair ',' members  
pair ::= string : value  
array ::= '[' element ']' *  
element ::=  
    value | value ',' element
```

```
value ::=  
    String |  
    number |  
    object |  
    array |  
    true |  
    false |  
    null
```


JSON



Diccionario Python = JSON



```
{
  "Space X":
  [
    {"VehiculosLanzamiento":
      [
        {"Nombre": "Falcon 1", "CargaUtil": 670},
        {"Nombre": "Falcon 9", "CargaUtil": 22800},
        {"Nombre": "Falcon Heavy", "CargaUtil": 63800}
      ]
    },
    {"Naves":
      [
        {"Nombre": "Dragon 1", "Descripcion": "Cargo Dragon"},
        {"Nombre": "Dragon 2", "Descripcion": "Crew Dragon"}
      ]
    }
  ]
}
```

```
1 {
2   "Space X": [
3     {
4       "VehiculosLanzamiento": [
5         {
6           "Nombre": "Falcon 1",
7           "CargaUtil": 670
8         },
9         {
10          "Nombre": "Falcon 9",
11          "CargaUtil": 22800
12        },
13        {
14          "Nombre": "Falcon Heavy",
15          "CargaUtil": 63800
16        }
17      ]
18    },
19    {
20      "Naves": [
21        {
22          "Nombre": "Dragon 1",
23          "Descripcion": "Cargo Dragon"
24        },
25        {
26          "Nombre": "Dragon 2",
27          "Descripcion": "Crew Dragon"
28        }
29      ]
30    }
31  ]
32 }
```


* <https://jsonformatter.org/>



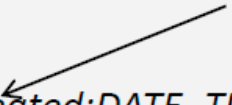
MongoDB - Twitter

```
{_id:POST_ID,  
title:TITLE_OF_POST,  
description:POST_DESCRIPTION,  
by:POST_BY,  
url:URL_OF_POST  
tags:[TAG1,TAG2,TAG3],  
likes:TOTAL_LIKES,  
comments:  
  [  
    {user:'COMMENT_BY',message:TEXT,dateCreated:DATE_TIME,like:LIKES},  
    {user:'COMMENT_BY',message:TEXT,dateCreated:DATE_TIME,like:LIKES}  
  ]  
}
```

Los tags se integran como array.



Los comentarios se integran como array de subdocumentos



MongoDB Consultas + Actualizaciones

Equivalencias de SQL en MongoDB

<https://www.mongodb.com/docs/manual/reference/sql-comparison/#sql-to-mongodb-mapping-chart>

SQL Terms, Functions, and Concepts	MongoDB Aggregation Operators
WHERE	\$match
GROUP BY	\$group
HAVING	\$match
SELECT	\$project
ORDER BY	\$sort
LIMIT	\$limit
SUM()	\$sum
COUNT()	\$sum
join	No existe, se puede hacer algo similar con \$unwind con los campos embebidos en un subdocumento.

MongoDB CRUD Operations

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

find

count

sort

group

sum

unwind

limit

MongoDB Pipeline

<https://docs.mongodb.com/manual/core/aggregation-pipeline/>

```
db.orders.aggregate( [  
    { $match: { status: "A" } },  
    { $group: { _id: "$cust_id",total:  {$sum: "$amount" } } }  
])
```

MongoDB Pipeline

Collection

↓
`db.orders.aggregate([`
 `{ $match: { status: "A" } },`
 `{ $group: { _id: "$cust_id", total: { $sum: "$amount" } } }`
`]`)

<pre>{ cust_id: "A123", amount: 500, status: "A" }</pre>
<pre>{ cust_id: "A123", amount: 250, status: "A" }</pre>
<pre>{ cust_id: "B212", amount: 200, status: "A" }</pre>

MongoDB Pipeline

Collection

↓
`db.orders.aggregate([`
 `$match stage → { $match: { status: "A" } },`
 `{ $group: { _id: "$cust_id", total: { $sum: "$amount" } } }`
 `]`
`)`

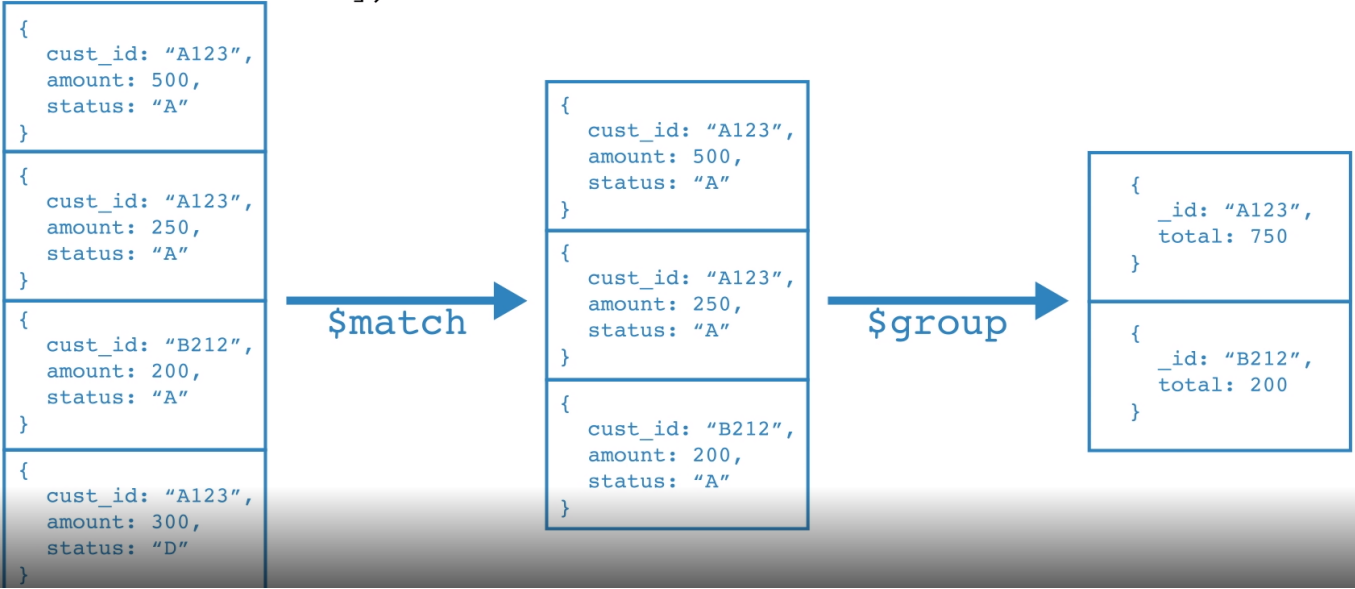
{ cust_id: "A123", amount: 500, status: "A" }
{ cust_id: "A123", amount: 250, status: "A" }
{ cust_id: "B212", amount: 200, status: "A" }
{ cust_id: "A123", amount: 300, status: "D" }

→
`$match`

{ cust_id: "A123", amount: 500, status: "A" }
{ cust_id: "A123", amount: 250, status: "A" }
{ cust_id: "B212", amount: 200, status: "A" }

MongoDB Pipeline

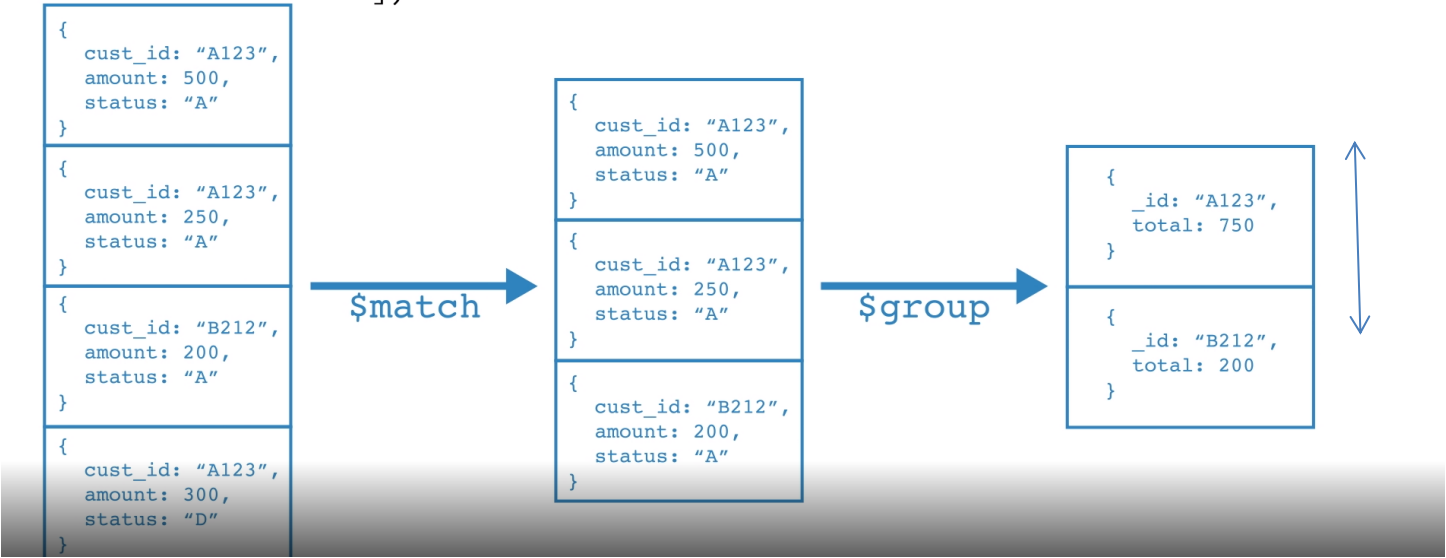
Collection
↓
db.orders.aggregate([
 \$match stage → { \$match: { status: "A" } },
 \$group stage → { \$group: { _id: "\$cust_id",total: { \$sum: "\$amount" } } }
])



MongoDB Pipeline

```
Collection
↓
db.orders.aggregate( [
  $match stage → { $match: { status: "A" } },
  $group stage → { $group: { _id: "$cust_id", total: { $sum: "$amount" } } }
])
```

\$sort
, { \$sort : {total : 1}}



MongoDB Pipeline

```
db.orders.aggregate([{$match: {status: "A"}}, {$group: {_id: "$cust_id", total: {$sum: "$amount"}}}, {$sort: {total: 1}}])
```

```
/* 1 */
{
  "_id" : "B123",
  "total" : 200
}

/* 2 */
{
  "_id" : "A123",
  "total" : 750
}
```

MongoDB Pipeline

\$unwind desagregar arrays, por cada elemento del array crea un nuevo documento duplicando los demás datos.

```
/* 1 */
{
  "_id" : ObjectId("5eaa6b2f804aacfaf16311ce"),
  "nombre" : "roger",
  "deportes" : [ "rally", "basquet" ]
}
```

```
/* 2 */
{
  "_id" : ObjectId("5eaa6b5d804aacfaf1631bce"),
  "nombre" : "jordi",
  "deportes" : [ "egames" ]
}
```

```
/* 3 */
{
  "_id" : ObjectId("5eaa6b95804aacfaf163286d"),
  "nombre" : "monica",
  "deportes" : [ "basquet", "zumba" ]
}
```

```
/* 4 */
{
  "_id" : ObjectId("5eaa6bb7804aacfaf1632ffd"),
  "nombre" : "jony",
  "deportes" : [ "futbol", "egames" ]
}
```

MongoDB Pipeline

```
db.getCollection('deportesPreferidos').aggregate([{$unwind : "$deportes"}])
```

deportesPreferidos 0.123 sec.			
	_id	nombre	deportes
1	<input type="checkbox"/> ObjectId("5...	<input type="checkbox"/> roger	<input type="checkbox"/> rally
2	<input type="checkbox"/> ObjectId("5...	<input type="checkbox"/> roger	<input type="checkbox"/> basquet
3	<input type="checkbox"/> ObjectId("5...	<input type="checkbox"/> jordi	<input type="checkbox"/> egames
4	<input type="checkbox"/> ObjectId("5...	<input type="checkbox"/> monica	<input type="checkbox"/> basquet
5	<input type="checkbox"/> ObjectId("5...	<input type="checkbox"/> monica	<input type="checkbox"/> zumba
6	<input type="checkbox"/> ObjectId("5...	<input type="checkbox"/> jony	<input type="checkbox"/> futbol
7	<input type="checkbox"/> ObjectId("5...	<input type="checkbox"/> jony	<input type="checkbox"/> egames

MongoDB Pipeline

```
db.getCollection('deportesPreferidos').aggregate([{$unwind : "$deportes"}, {$group:
{_id:"$deportes", total: {$sum:1}}}]])
```

deportesPreferidos		0.119 sec.
	_id	total
1	rally	1.0
2	basquet	2.0
3	egames	2.0
4	futbol	1.0
5	zumba	1.0

Actualizar: Query Selectors

Name	Description
<code>\$eq</code>	Matches values that are equal to a specified value.
<code>\$gt</code>	Matches values that are greater than a specified value.
<code>\$gte</code>	Matches values that are greater than or equal to a specified value.
<code>\$in</code>	Matches any of the values specified in an array.
<code>\$lt</code>	Matches values that are less than a specified value.
<code>\$lte</code>	Matches values that are less than or equal to a specified value.
<code>\$ne</code>	Matches all values that are not equal to a specified value.
<code>\$nin</code>	Matches none of the values specified in an array.

Actualizar

❑ *\$set cambia valor*

```
db.Partidos.update({siglas:{$eq:"PSOE"}},{ $set:{programa:"Blabla"}})
```

❑ *\$push: añade un valor especificado a un array.*

```
db.Partidos.update( { _id:1}, { $push:{votos : 89} } )
```

<https://docs.mongodb.com/manual/reference/method/db.collection.update/>
<https://docs.mongodb.com/manual/reference/operator/query/#query-selectors>

Buscar

```
db.collection.find(<criteria>, <projection>)
```

```
db.Partidos.find({siglas:{$eq:"PSOE"}}, {nombre:1})
```

Projection = nombre atributo y booleano (1 mostrar y 0 no mostrar, lista blanca o lista negra)

<https://docs.mongodb.com/manual/reference/method/db.collection.find/>

<https://docs.mongodb.com/manual/reference/method/db.collection.find/#find-projection>

Buscar – Operadores Lógicos

- \$all: Devuelve cierto si el array almacenado como valor del atributo tiene los mismos elementos que el proporcionado en la condición.
- \$exists: Para comprobar que el atributo existe en el documento.
- \$mod: Para comprobar el resto de una división del valor del atributo por un número.
- \$ne: Devuelve cierto si no es igual (non equal)
- \$in: Devuelve cierto si el valor está entre alguno de los proporcionados.
- \$nin: Contrario de \$in.
- \$or: Para indicar que se debe cumplir al menos una condición de entre un grupo de condiciones.
- \$nor: Contrario de \$or.
- \$size: Devuelve true si coincide con el número de elementos de un array.
- \$type: Para comprobar el tipo del valor almacenado (número, cadena...)

Consulta agregados

db.collection.aggregate(pipeline)

- \$first, \$last, \$min, \$max, \$avg, \$sum,
- \$push o \$addToSet para generar un array con los valores que devuelve, etc.
- \$group: indica el atributo por el que se quiere agregar
- \$match: criterio a cumplir

Consulta agregados

db.collection.aggregate(pipeline)

db.Partidos.aggregate([{\$group:{_id:"\$tendencia", "num":{"\$sum":1}}}]])



The screenshot shows a MongoDB query interface with the following components:

- Query Editor:** Contains the aggregation query: `db.Partidos.aggregate([{$group:{_id:"$tendencia", "num":{"$sum":1}}}]])`
- Execution Status:** Shows the query was executed on the 'Partidos' collection in 0.122 seconds.
- Results Table:** Displays the aggregated data with three columns: Key, Value, and Type.

Key	Value	Type
▼ (1) liberal	{ 2 fields }	Object
_id	liberal	String
num	2.0	Double
▼ (2) derecha	{ 2 fields }	Object
_id	derecha	String
num	1.0	Double
▼ (3) izquierda	{ 2 fields }	Object
_id	izquierda	String
num	1.0	Double

> Dudas



CONFIGURACIÓN

MongoDB Atlas <https://www.mongodb.com/>

MONGODB ATLAS

MongoDB Atlas.
MongoDB totalmente
gestionada en la nube.


Aproveche el poder de sus datos creando y administrando
sus datos en la nube.

Empiece gratis

Consulte los precios →

Registrarse

Vea de forma gratuita lo que Atlas es capaz de hacer

 Regístrese con Google

Nombre*

Roger

Apellido*

Clotet

Empresa

Correo Electrónico*

Contraseña*

.....

☒ Acepto las **Terms of Service** y la **Privacy Policy**.

Cree su cuenta de Atlas

CONFIGURACIÓN



Great, now verify your email



Check your inbox at **rclozetm@unemi.edu.ec** and click the verification link inside to complete your registration. This link will expire shortly, so verify soon!

Don't see an email? Check your spam folder.

Link expired? [Resend verification email](#)



Email successfully verified!



Continue

CONFIGURACIÓN



Enable your multi-factor authentication methods

Secure your account across MongoDB

Secure your MongoDB account across MongoDB Atlas, University, Community, and Support with Multi-Factor-Authentication (MFA). Setting up MFA will help reduce risks from outside threats to your account. [Learn More](#)

Setup Backup MFA Methods

In addition, you will also be setting up an additional method in this setup. This is a precaution in the case you lose access to your first MFA enabled device and to avoid being locked out.

Set up now

Remind me later



Multi-factor authentication setup is now complete

✓ Your MFA methods have now been enabled. You can always add or remove more methods under your Account Application inside of the Security page.

Multi-factor authentication (MFA) methods

In the case that you lose access to your first enabled MFA method, set up another one to avoid being locked out. [Learn more](#)



Email



Okta Verify Mobile App



Security Key/Biometric

SETUP



Authenticator App

SETUP



SMS

SETUP

Why is this important for me?

MFA will enhance your account's security by requiring you to identify yourself by more than your email and password.

[Learn how this protects your account](#)

Continue

CONFIGURACIÓN

Overview



Create a cluster

Choose your cloud provider, region, and specs.

+ Create

CONFIGURACIÓN

Deploy your cluster

Use a template below or set up advanced configuration options. You can also edit these configuration options once the cluster is created.

M10

\$0.09/hour

For production applications with sophisticated workload requirements.

STORAGE

10 GB

RAM

2 GB

vCPU

2 vCPUs

Serverless

For application development and testing, or workloads with variable traffic.

STORAGE

Up to 1 TB

RAM

Auto-scale

vCPU

Auto-scale

M0

Free

For learning and exploring MongoDB in a cloud environment.

STORAGE

512 MB

RAM

Shared

vCPU

Shared

✔ **Free forever!** Your M0 cluster is ideal for experimenting in a limited sandbox. You can upgrade to a production cluster anytime.

Name
You cannot change the name once the cluster is created.

Cluster0

- ☒ Automate security setup ⓘ
- ☒ Preload sample dataset ⓘ

Provider

aws

Google Cloud

Azure

Go to Advanced Configuration

Create Deployment

CONFIGURACIÓN

CLUSTERS > CREATE A SHARED CLUSTER

Create a Shared Cluster

Welcome to MongoDB Atlas! We've recommended some of our most popular options, but feel free to customize your cluster to your needs. For more information, check our [documentation](#).

Serverless

Dedicated

FREE Shared

For learning and exploring MongoDB in a sandbox environment. Basic configuration controls. No credit card required to start. Upgrade to dedicated clusters for full functionality. Explore with sample datasets. Limit of one free cluster per project.

Cloud Provider & Region

GCP, Sao Paulo (southamerica-east1)

aws

Google Cloud

Azure

★ Recommended region ⓘ ⓘ Dedicated tier region ⓘ ⓘ Low carbon emissions region ⓘ

NORTH AMERICA / SOUTH AMERICA

🇺🇸 Iowa (us-central) ★ ⓘ

🇧🇪 Belgium (europe-west) ★ ⓘ

🇦🇺 Sydney (australia-southeast) ★ ⓘ

🇻🇪 Sao Paulo (southamerica-east1) ★ ⓘ

🇵🇱 Warsaw (europe-central2) ★ ⓘ

🇦🇺 Melbourne (australia-southeast2) ★ ⓘ

🇺🇸 South Carolina (us-east) ★ ⓘ

🇫🇮 Finland (europe-north) ★ ⓘ

🇦🇺 Tokyo (asia-northeast) ★

🇺🇸 N. Virginia (us-east-4) ★ ⓘ

🇬🇧 London (europe-west2) ★ ⓘ

🇮🇳 Mumbai (asia-south) ★

🇺🇸 Ohio (us-east) ★ ⓘ

🇩🇪 Frankfurt (europe-west3) ★ ⓘ

🇮🇳 Singapore (asia-southeast) ★

🇺🇸 Los Angeles (us-west2) ★ ⓘ

🇳🇱 Netherlands (europe-west4) ★ ⓘ

🇮🇩 Jakarta (asia-southeast2) ★

🇺🇸 Salt Lake City (us-west3) ★ ⓘ

🇨🇭 Zurich (europe-west5) ★ ⓘ

🇰🇷 Seoul (asia-northeast3) ★

🇺🇸 Las Vegas (us-west4) ★ ⓘ

🇮🇹 Milan (europe-west5) ★ ⓘ

🇰🇷 Seoul (asia-northeast3) ★

Atlas

Access Manager

Billing

Project 0

Data Services

App Services

Charts

Overview

DEPLOYMENT

Database

Data Lake

SERVICES

Device & Edge Sync

Triggers

Data API

Data Federation

Atlas Search

Stream Processing

Migration

SECURITY

Quickstart

Backup

Database Access

Network Access

Advanced

New On Atlas 4

Goto

Security Quickstart

To access data stored in Atlas, you'll need to create users and set up network security controls. [Learn more about security se](#)

1

How would you like to authenticate your connection?

Your first user will have permission to read and write any data in your project.

Username and Password

Certificate

ⓘ

We autogenerated a username and password for your first database user in this project using your MongoDB Cloud registration information.

×

Create a database user using a username and password. Users will be given the *read and write to any database privilege* by default. You can update these permissions and/or create additional users later. Ensure these credentials are different to your MongoDB Cloud username and password.

Username

rclorem

Password ⓘ

Autogenerate Secure Password

Copy

Create User

MO Cluster Provisioning...

This process will take 3-5 minutes.

×


to connect from?

CONFIGURACIÓN

 Your cluster has finished provisioning. 

✔ Where would you like to connect from?

Enable access for any network(s) that need to read and write data to your cluster.



**My Local Environment**

Use this to add network IP addresses to the IP Access List. This can be modified at any time.

Cloud Environment

Use this to configure network access between Atlas and your cloud or on-premise environment. Specifically, set up IP Access Lists, Network Peering, and Private Endpoints.

ADVANCED

 We added your current IP address. You can connect to your cluster locally from this device. 

Add entries to your IP Access List

Only an IP address you add to your Access List will be able to connect to your project's clusters. You can manage existing IP entries via the [Network Access Page](#).

IP Address	Description
<input type="text" value="Enter IP Address"/>	<input type="text" value="Enter description"/>
<input type="button" value="Add My Current IP Address"/>	
<input type="button" value="Add Entry"/>	

IP Access List	Description
0.0.0.0/0	All

Finish and Close

CONFIGURACIÓN

Overview

Clusters

Create cluster



Cluster0

Connect

Edit configuration



Add data



Migrate data



Load sample data

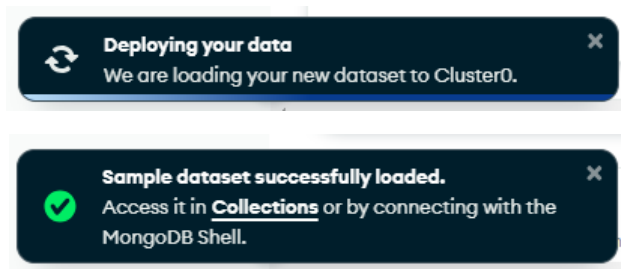


Data modeling templates



+ Add Tag

CONFIGURACIÓN



CONFIGURACIÓN



TOOLS

MongoDB Command Line Database Tools Download

The MongoDB Database Tools are a collection of command-line utilities for working with a MongoDB deployment. These tools release independently from the MongoDB Server schedule enabling you to receive more frequent updates and leverage new features as soon as they are available. See the [MongoDB Database Tools](#) documentation for more information.

Version	100.10.0	▼
Platform	Windows x86_64	▼
Package	zip	▼

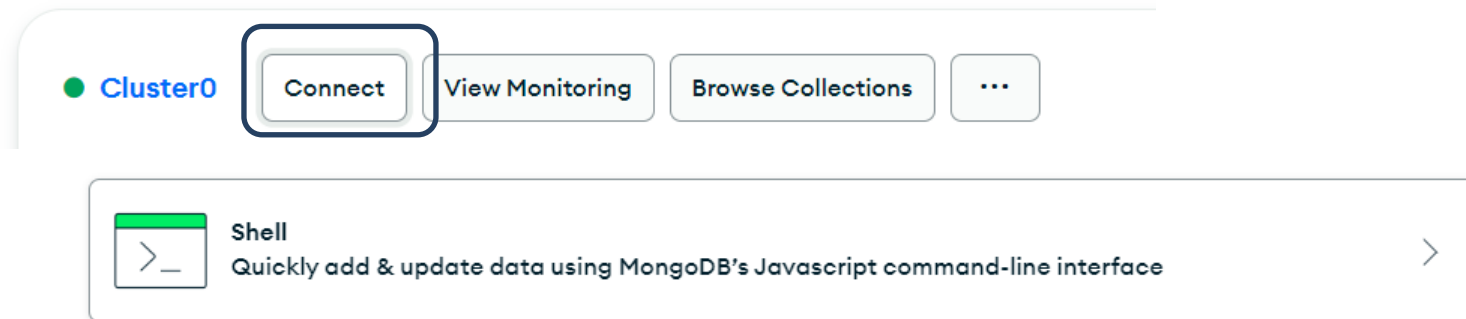
Download

Copy link

More Options ...

CONFIGURACIÓN

Mongoimport



2. Run your connection string in your command line

Use this connection string in your application

```
mongo "mongodb+srv://cluster0.sit6c.mongodb.net/" --username rclotetm
```



CONFIGURACIÓN

Mongoimport

```
mongoimport "mongodb+srv://cluster0.sit6c.mongodb.net/" --username rclotetm --password ????? --db VIU --collection tweets --file tweets_UNEMI_2.json --type json
```

```
mongoimport "mongodb+srv://cluster0.sit6c.mongodb.net/" --username rclotetm --password ????? --db VIU --collection cuentas --file cuentas_UNEMI_2.json --type json
```

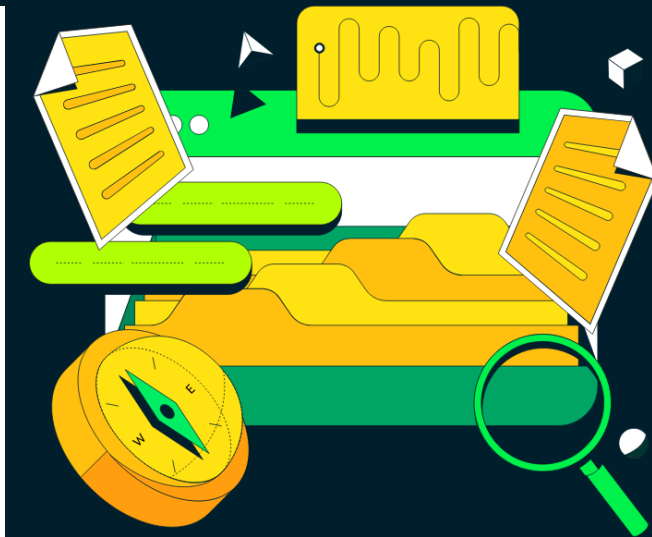
CONFIGURACIÓN

Mongoimport

```
mongoimport "mongodb+srv://cluster0.sit6c.mongodb.net/" --username rclotetm --password
2024-09-22T18:49:06.048-0500 connected to: mongodb+srv://cluster0.sit6c.mongodb.net/
2024-09-22T18:49:09.051-0500 [#####] UNEMI.tweets 3.06MB/42.4MB (7.2%)
2024-09-22T18:49:12.050-0500 [#####] UNEMI.tweets 6.75MB/42.4MB (15.9%)
2024-09-22T18:49:15.054-0500 [#####] UNEMI.tweets 6.75MB/42.4MB (15.9%)
2024-09-22T18:49:18.063-0500 [#####] UNEMI.tweets 13.2MB/42.4MB (31.2%)
2024-09-22T18:49:21.060-0500 [#####] UNEMI.tweets 16.6MB/42.4MB (39.1%)
2024-09-22T18:49:24.060-0500 [#####] UNEMI.tweets 16.6MB/42.4MB (39.1%)
2024-09-22T18:49:27.052-0500 [#####] UNEMI.tweets 20.2MB/42.4MB (47.5%)
2024-09-22T18:49:30.063-0500 [#####] UNEMI.tweets 23.8MB/42.4MB (56.0%)
2024-09-22T18:49:33.056-0500 [#####] UNEMI.tweets 23.8MB/42.4MB (56.0%)
2024-09-22T18:49:36.063-0500 [#####] UNEMI.tweets 26.6MB/42.4MB (62.9%)
2024-09-22T18:49:39.055-0500 [#####] UNEMI.tweets 30.3MB/42.4MB (71.5%)
2024-09-22T18:49:42.054-0500 [#####] UNEMI.tweets 30.3MB/42.4MB (71.5%)
2024-09-22T18:49:45.062-0500 [#####] UNEMI.tweets 34.2MB/42.4MB (80.5%)
2024-09-22T18:49:48.056-0500 [#####] UNEMI.tweets 34.2MB/42.4MB (80.5%)
2024-09-22T18:49:51.051-0500 [#####] UNEMI.tweets 38.9MB/42.4MB (91.6%)
2024-09-22T18:49:54.049-0500 [#####] UNEMI.tweets 42.4MB/42.4MB (100.0%)
2024-09-22T18:49:57.053-0500 [#####] UNEMI.tweets 42.4MB/42.4MB (100.0%)
2024-09-22T18:49:58.483-0500 [#####] UNEMI.tweets 42.4MB/42.4MB (100.0%)
2024-09-22T18:49:58.483-0500 11991 document(s) imported successfully. 0 document(s) failed to import.
```

CONFIGURACIÓN

Compass. The GUI for MongoDB.



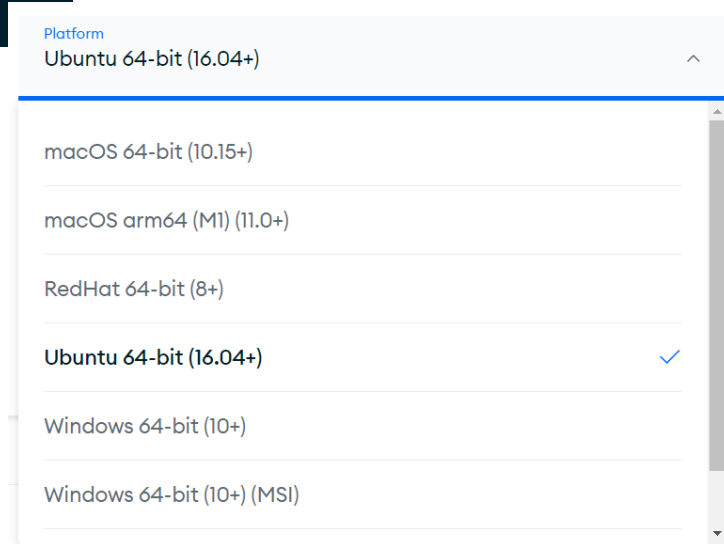
Trabaje fácilmente en con sus datos en Compass, la interfaz gráfica de usuario desarrollado por y para MongoDB. Compass proporciona de todo, desde el análisis de esquemas hasta la optimización de índices y pipelines de agregación en una única interfaz centralizada.

- Compass está disponible gratuitamente para descargar y usar
- Disponible en Linux, Mac o Windows

CONFIGURACIÓN

Compass. The GUI for MongoDB.

<https://www.mongodb.com/try/download/compass>



CONFIGURACIÓN

Access your data through tools



Compass

Explore, modify, and visualize your data with MongoDB's GUI



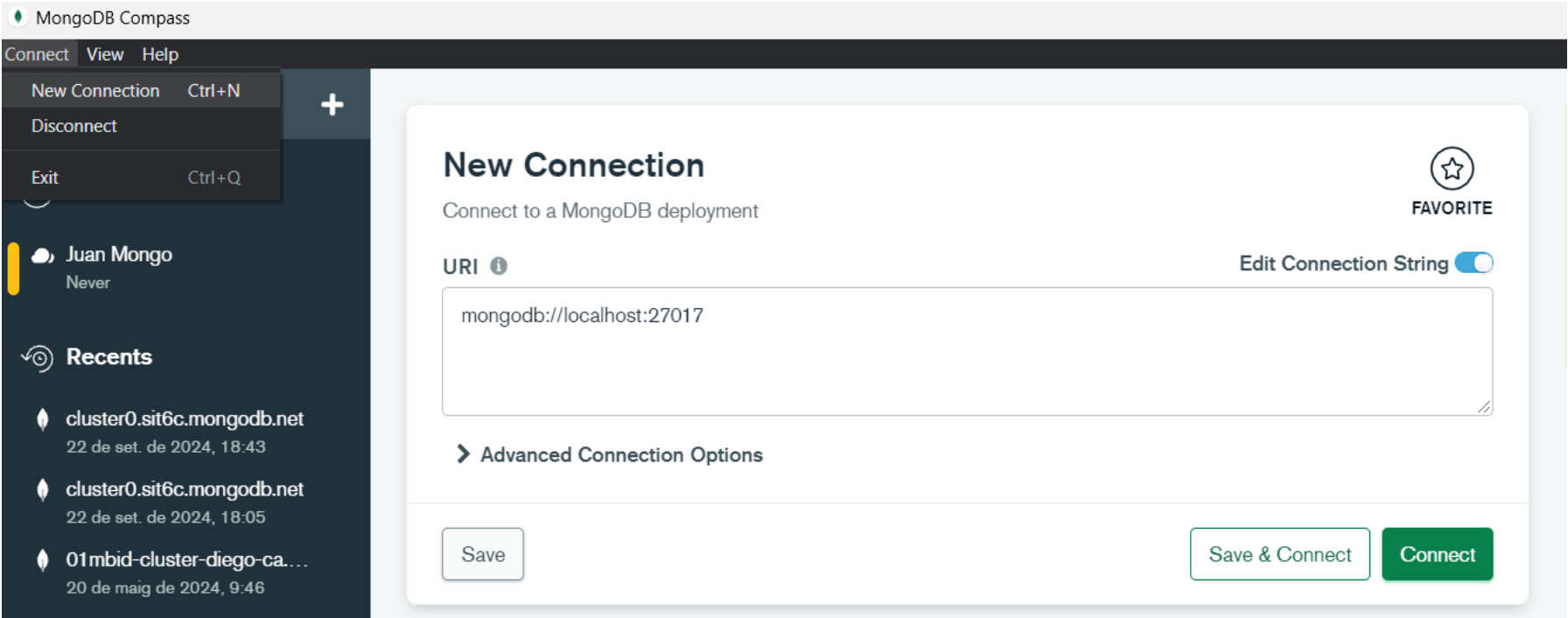
2. Copy the connection string, then open MongoDB Compass

Use this connection string in your application

```
mongodbsrv://rclotetm:<db_password>@cluster0.sit6c.mongodb.net/
```



CONFIGURACIÓN



CONFIGURACIÓN

MongoDB Compass - cluster0.sit6c.mongodb.net/UNEMI

Connect View Help

cluster0.sit6c.mongodb.net

12 DBS 24 COLLECTIONS

☆ FAVORITE

HOSTS

cluster0-shard-00-01.sit6c...

cluster0-shard-00-02.sit6c...

cluster0-shard-00-00.sit6c...

CLUSTER

Replica Set (atlas-zk6sjv-s...

3 Nodes

EDITION

MongoDB 7.0.12 Enterprise

{ } My Queries

☰ Databases

🔍 Filter your data

Collections

Create collection View

Sort by Collection Name

tweets

Storage size:	Documents:	Avg. document size:	Indexes:	Total index size:
8.53 MB	12 K	3.54 kB	1	552.96 kB

CONFIGURACIÓN

Tweet Object

<https://developer.x.com/en/docs/x-api/v1/data-dictionary/object-model/tweet>

Visor JSON de un documento (tweet)

<https://codebeautify.org/jsonviewer>

CONFIGURACIÓN




Project 0 Data Services App Services Charts



Quick, easy, real-time business insights

Atlas Charts makes it easy to create dashboards with multiple charts in just a few clicks.

 Loading instance

CONFIGURACIÓN

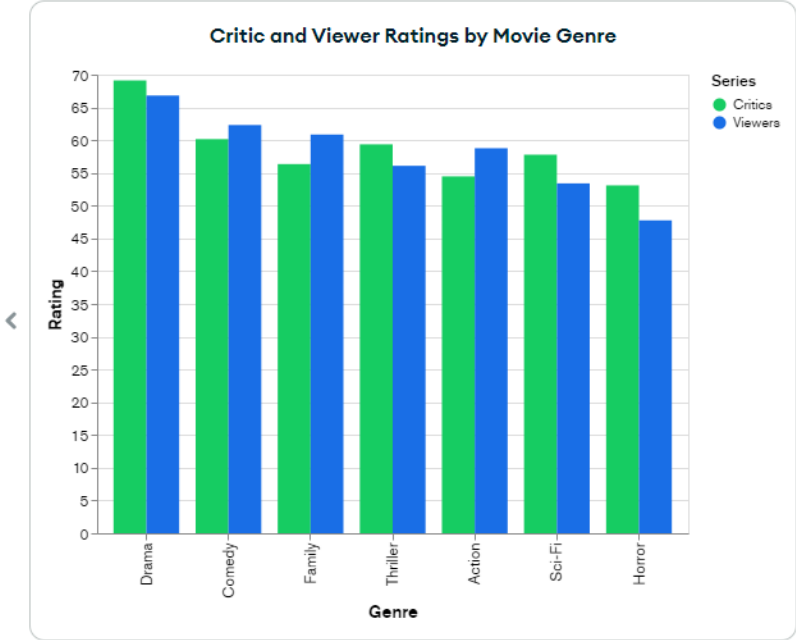


Hi Roger!
Welcome to Atlas Charts

Charts is the best way to create visualizations of your data.

- ✓ Zero setup – Build a chart from any collection in your project
- ✓ Create, share and collaborate on charts and dashboards
- ✓ Embed charts and dashboards into applications
- ✓ Completely free to get started and no user-based pricing

Start



CONFIGURACIÓN



Where would you like to start?



Chart builder

Quickly create a chart with data in your project or sample data.

Select



Sample dashboard

Explore the Atlas Charts features with the sample dashboard.

Select

Select Data Source

Suggested

Project


Sample


> Cluster0 (Cluster)

Cancel



CONFIGURACIÓN

 **Add Chart**

 Roger's Dashboard

Sample Data: Movies

Prompt

Describe a chart or insight you want from the data. Including fields in the prompt will help generate more relevant charts. [Learn more](#) about prompt writing.

Graph movie by country

Generate chart

Suggested Prompts

"What are the top 10 movies with the highest IMDB ratings and how do they compare to each other?"

SELECT

"How does the runtime of movies change over the years, binned by each decade?"

SELECT

"Are movies that have won more awards generally released

SELECT

Classic

Natural Language

Get help

Cancel

Save and close

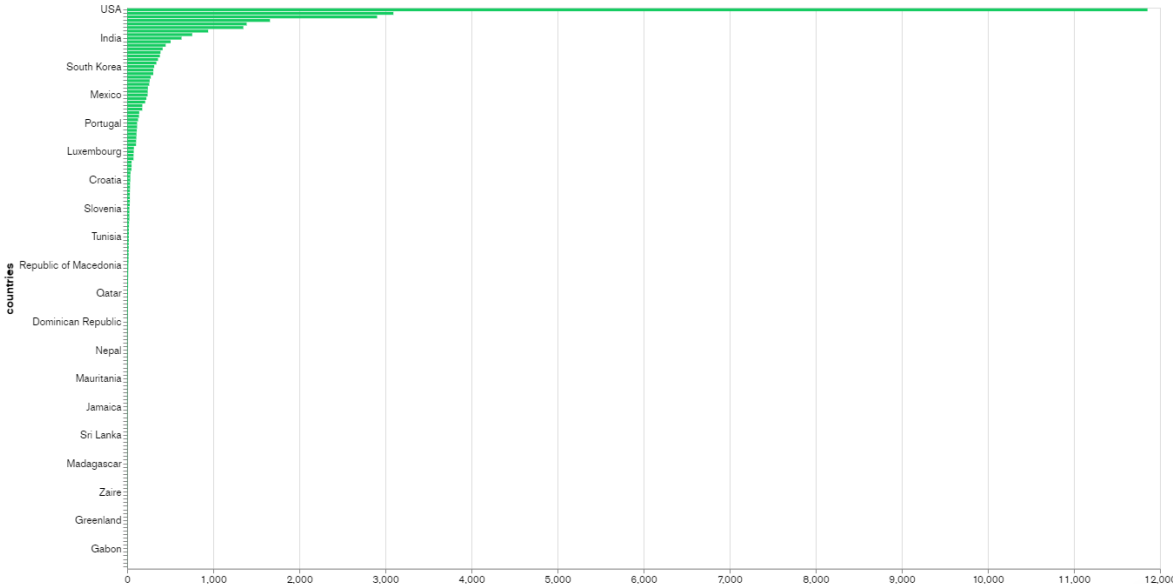
PREVIEW

COMPLETE

SUBSET



Cinematic Popularity across Nations: A Comparative Analysis

Enter a description



Country	Count (titles)
USA	12000
India	1000
South Korea	800
Mexico	600
Portugal	400
Luxembourg	300
Croatia	200
Slovenia	150
Tunisia	100
Republic of Macedonia	80
Qatar	60
Dominican Republic	50
Nepal	40
Mauritania	30
Jamaica	20
Sri Lanka	15
Madagascar	10
Zaire	5
Greenland	2
Gabon	1

Rate this generated chart

> Dudas



01MBID

Roger

roger.clotet@professor.universidadviu.com

Gracias



viu

Universidad
Internacional
de Valencia

universidadviu.com

De:



Planeta Formación y Universidades