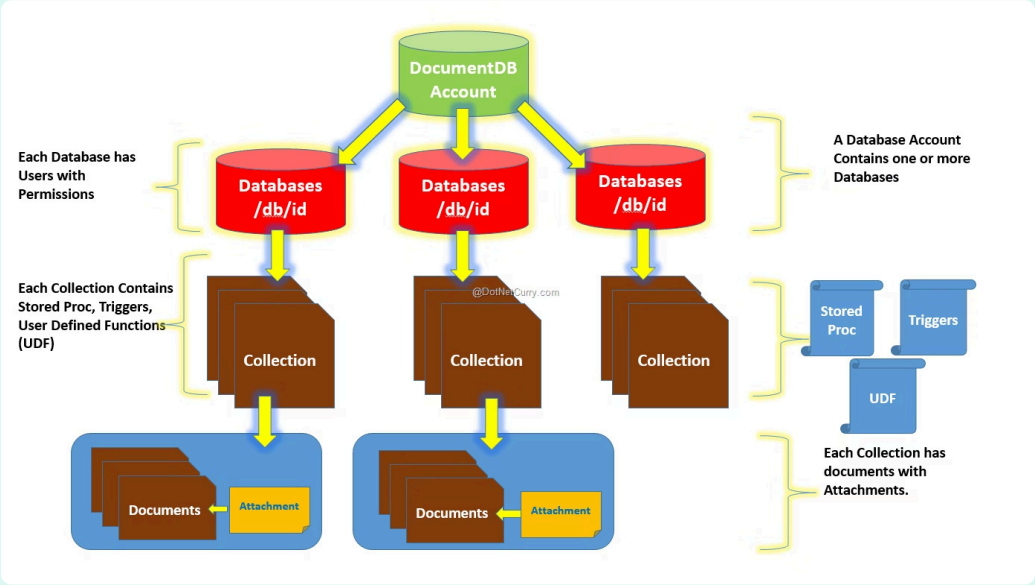
The Document Database

Ngoc My Nguyen

May 2025

Agenda



1. Documents and Document Databases

Document Databases

Non-relational databases that store data in the form of documents, typically JSON or BSON formats.

Flexible Data Models

Designed to handle semi-structured or unstructured data with dynamic schemas.

Scalability and Performance

Support rich querying capabilities, indexing, and horizontal scaling for modern applications.

2. Company Overview

2007

Founded as 10gen in New York City by former DoubleClick executives (Dwight Merriman, Eliot Horowitz, and Kevin Ryan)

2009

First production release of MongoDB database

2013

Company rebranded from 10gen to MongoDB

2014

Dev Ittycheria joined as CEO, leading the cloud transformation

2017

Launch of Atlas, MongoDB's cloud database service

Successful Initial public offering (IPO)

3. Licensing Evolution

Open Source Beginnings

MongoDB initially adopted the Affero General Public License (AGPL), ensuring that anyone modifying the code would need to share those modifications if distributed publicly.

SSPL Introduction

In 2018, MongoDB shifted to the Server Side Public License (SSPL), designed specifically to prevent cloud providers from offering MongoDB as a service without contributing back to the community.

Current Model

MongoDB offers both a free community edition as well as an enterprise version with additional features, enhanced security, monitoring, and support for on-premises deployments

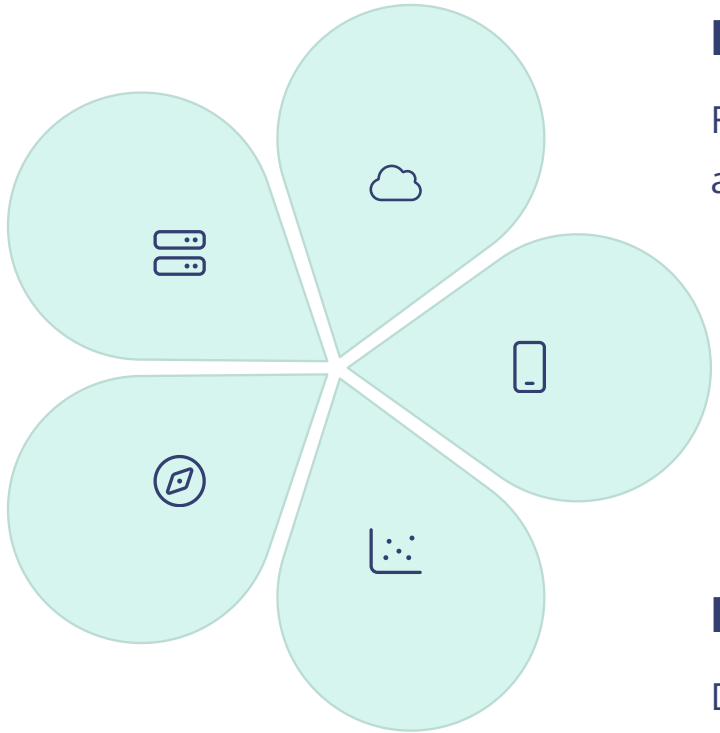
4. Product Portfolio

Core Database

Community & Enterprise Editions

MongoDB Compass

GUI for exploring and managing data



MongoDB Atlas

Fully-managed cloud database service across major cloud providers

MongoDB Realm

Mobile database and synchronization platform

MongoDB Charts

Data visualization tools

5. IT Architecture

Application Layer

Client applications using MongoDB drivers

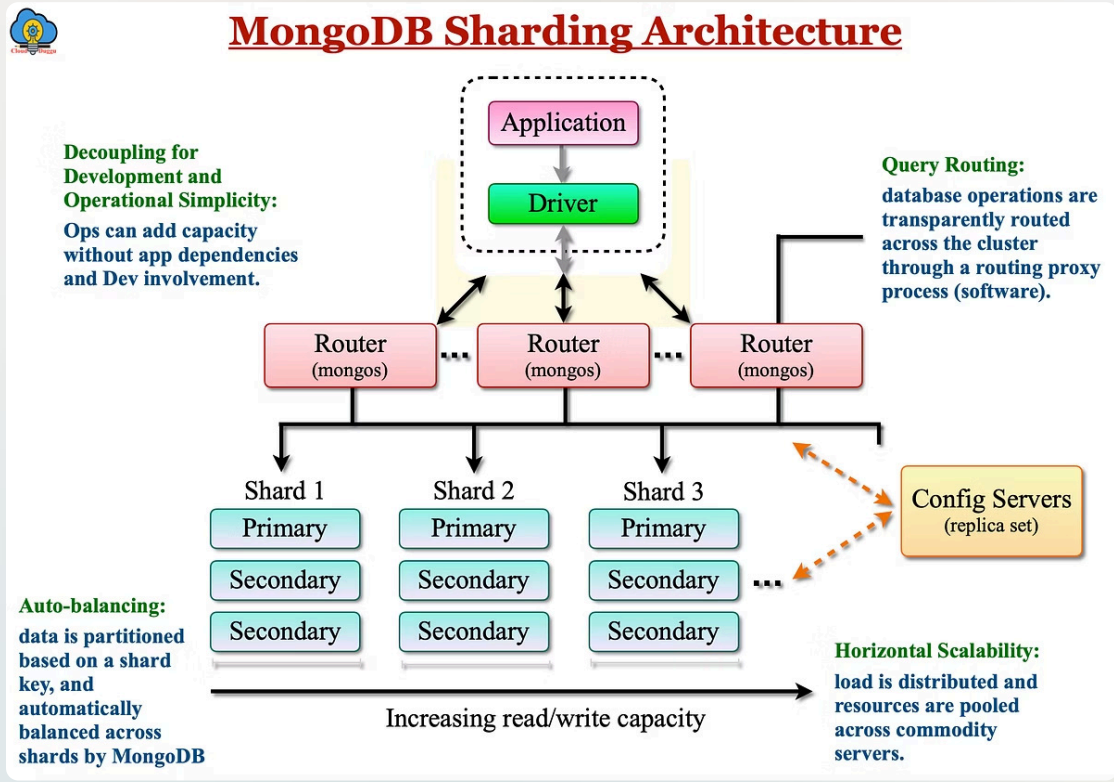
Database Layer

Distributed architecture with automatic sharding

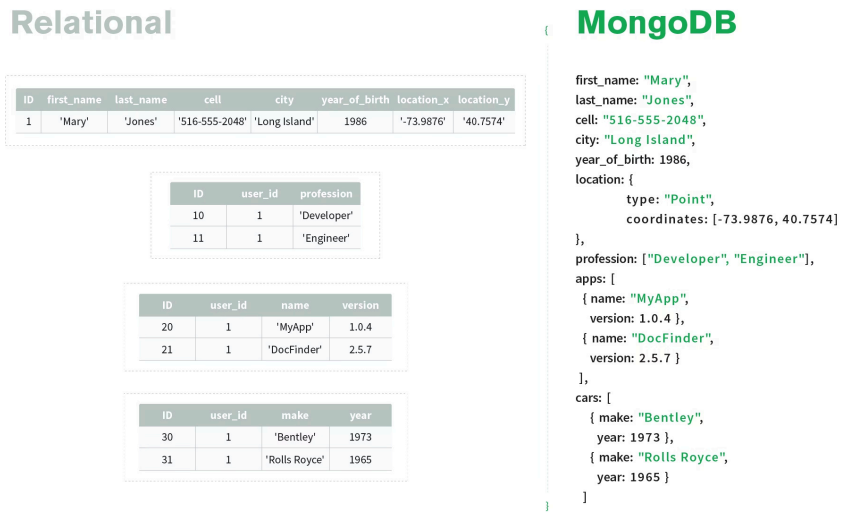
Storage Layer

WiredTiger storage engine (replaced memory-mapped storage)

MongoDB uses replica sets for high availability. Primary-secondary replication ensures data redundancy.



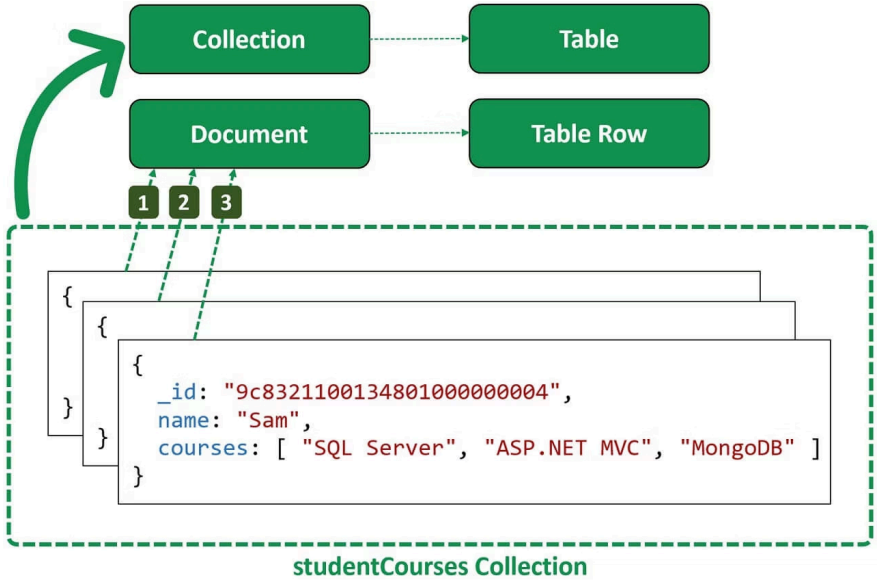
6. The Data Models



JSON-Like Documents

MongoDB stores data in flexible BSON format (binary JSON), which supports a rich variety of data types and nested structures.

- No rigid schema requirements
- Natural mapping to object-oriented code
- Intuitive data representation



Schema Flexibility

Documents in the same collection can have different fields, allowing applications to evolve naturally over time.

- Easy schema evolution
- No migration downtime
- Support for polymorphic data

7. Client Libraries and Programming

Python

Popular for data science and web applications with frameworks like Django and Flask.

Node.js

Excellent for JavaScript-based full-stack development with Express.js and MEAN/MERN stacks.

Java

Enterprise-grade driver with strong Spring Boot integration for robust backend systems.

C#/.NET

Deep integration with Microsoft ecosystem for building enterprise Windows applications.

Go

Lightweight, high-performance driver ideal for cloud-native and microservice architectures.

Ruby

Popular among web developers utilizing Rails, with a straightforward driver for MongoDB integration.

PHP

Commonly used in web development, with robust support for MongoDB in popular frameworks like Laravel.

and more ...

Official Drivers

Consistent API across all major programming languages

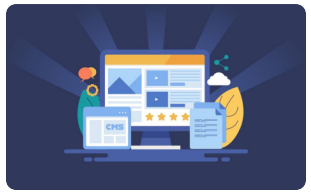
CRUD Operations

Simple, intuitive interface for database operations

Aggregation Framework

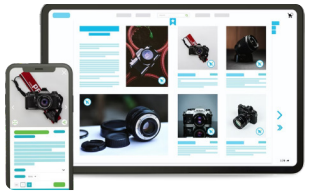
Powerful data processing and analysis capabilities

8. Real-World Applications



Content Management

MongoDB supports flexible content structures, making it ideal for dynamic content management systems that require rapid iteration and rich media support.



Catalogs

The document model efficiently handles diverse product catalogs with complex attributes and frequent updates.

Personalization

Document model excels at storing user profiles and preferences for customized experiences.

Real-Time Analytics

MongoDB's aggregation framework and time series collections enable powerful real-time dashboards and analytics applications.

Mobile Applications

The Realm SDK and synchronization capabilities make MongoDB ideal for offline-first mobile apps with seamless cloud sync.

Internet of Things

Flexible schema and time series collections handle diverse sensor data and high-volume telemetry ingestion.

9. Installation Options

Community Edition

- These documents offer guidance on installing MongoDB Community Edition, which is maintained by MongoDB.

Install on Linux

Install on macOS

Install on Windows

Install on Docker

- A community-managed Docker version of MongoDB is also available (**mongo - Official Image | Docker Hub**

Enterprise

MongoDB Enterprise is available for MongoDB Enterprise subscribers and includes several additional features including LDAP authentication, Kerberos authentication, and System Event Auditing.

Install on Linux

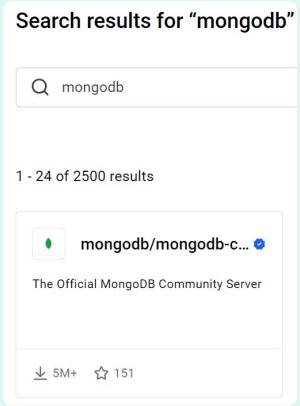
Install on macOS

Install on Windows

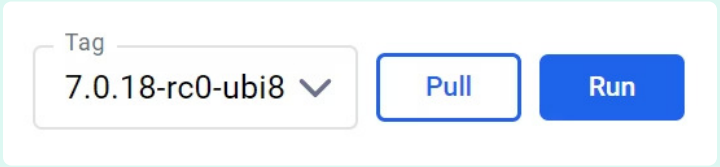
Install with Docker

9.1. Example: Easy installation and configuration with Docker Desktop (GUI)

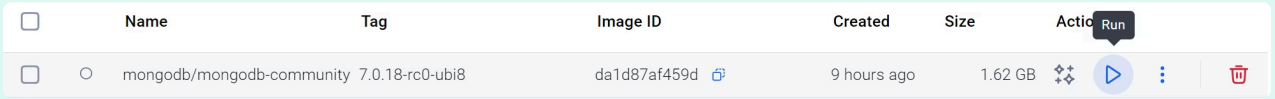
1. Search for "MongoDB Community Server" on Docker Hub and click on the result.



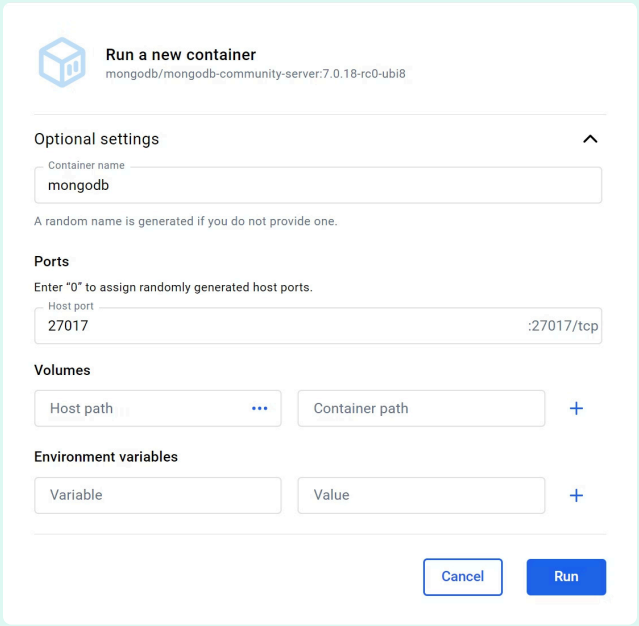
2. On the top right side, select and pull the desired version.



3. Locate the downloaded image and execute it to start running it.



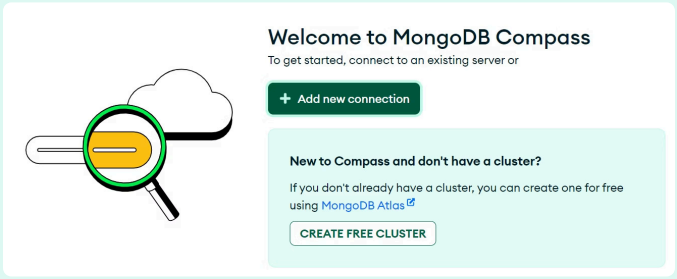
4. Customize the container settings by modifying the fields provided below.



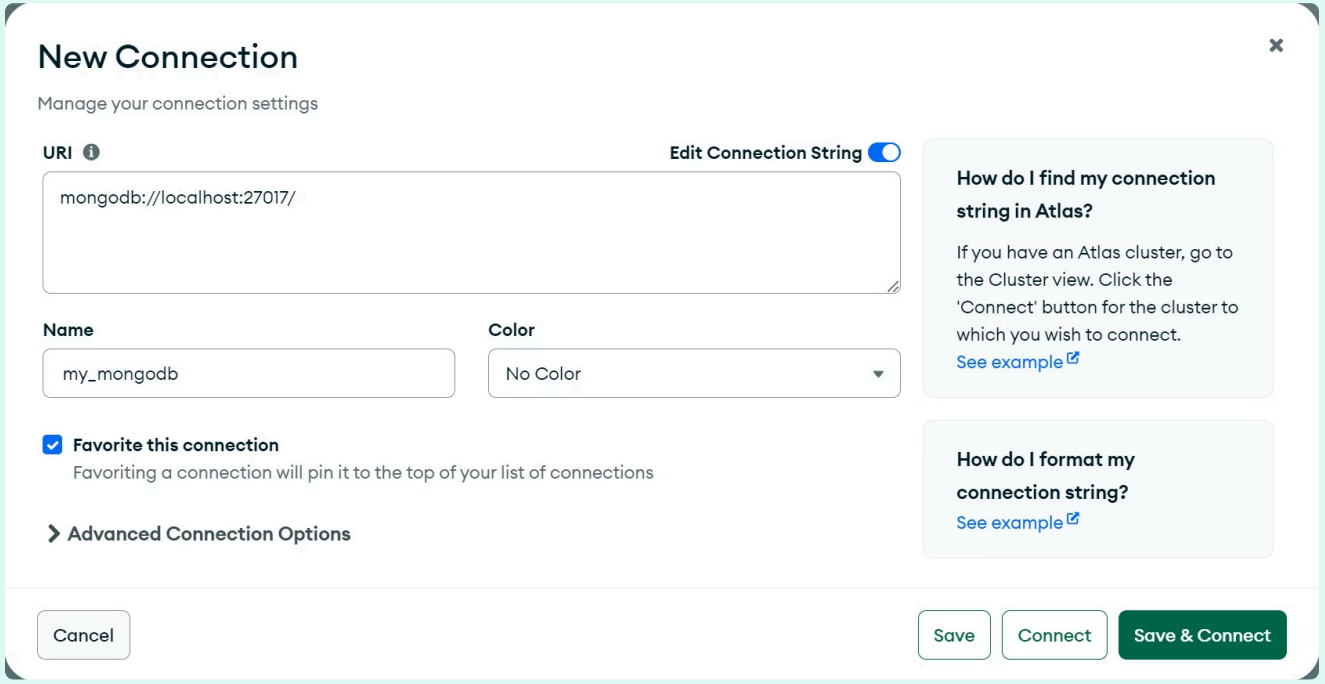
9.2. Example: Easily explore and manipulate database with MongoDB Compass (GUI)

1. Download MongoDB Compass **here**

2. Click "Add New Connection" to connect to the downloaded MongoDB Community Edition instance

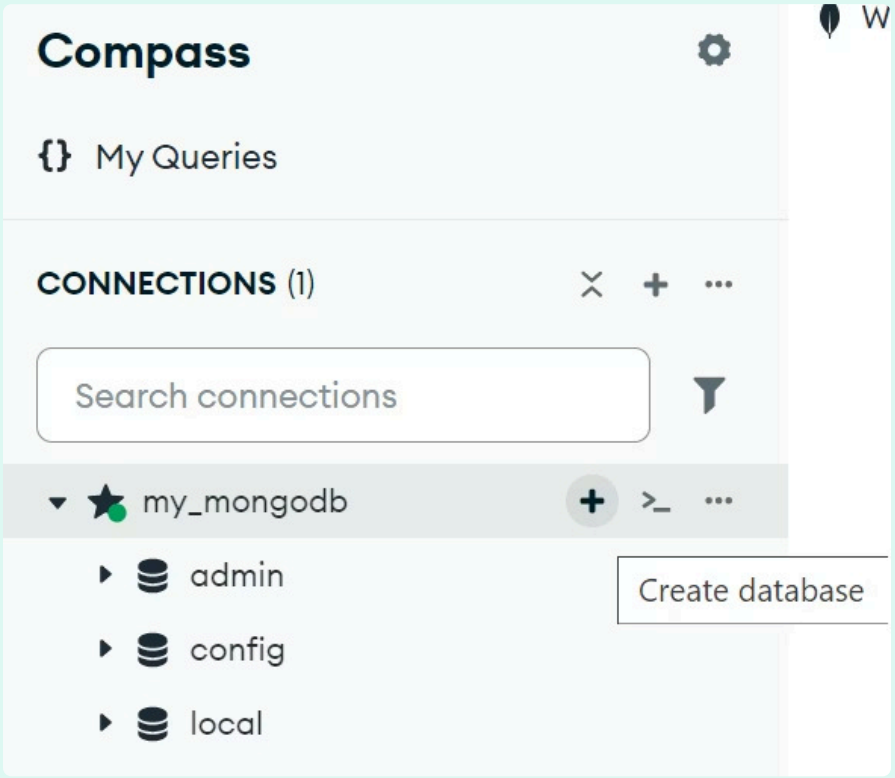


3. Modify the fields provided to complete the configuration, then click "Save & Connect."

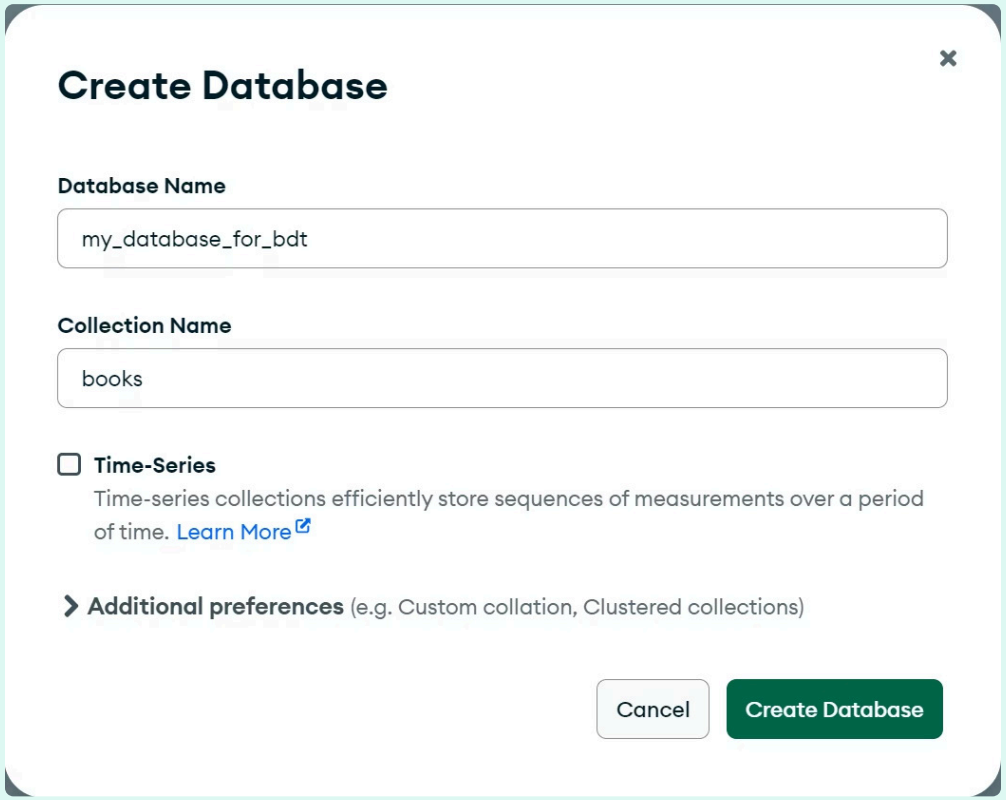


9.3. Example: Create a database and a collection with MongoDB Compass (GUI)

4. The connection will appear under "Connections." Click the plus symbol next to your connection to create a new database.

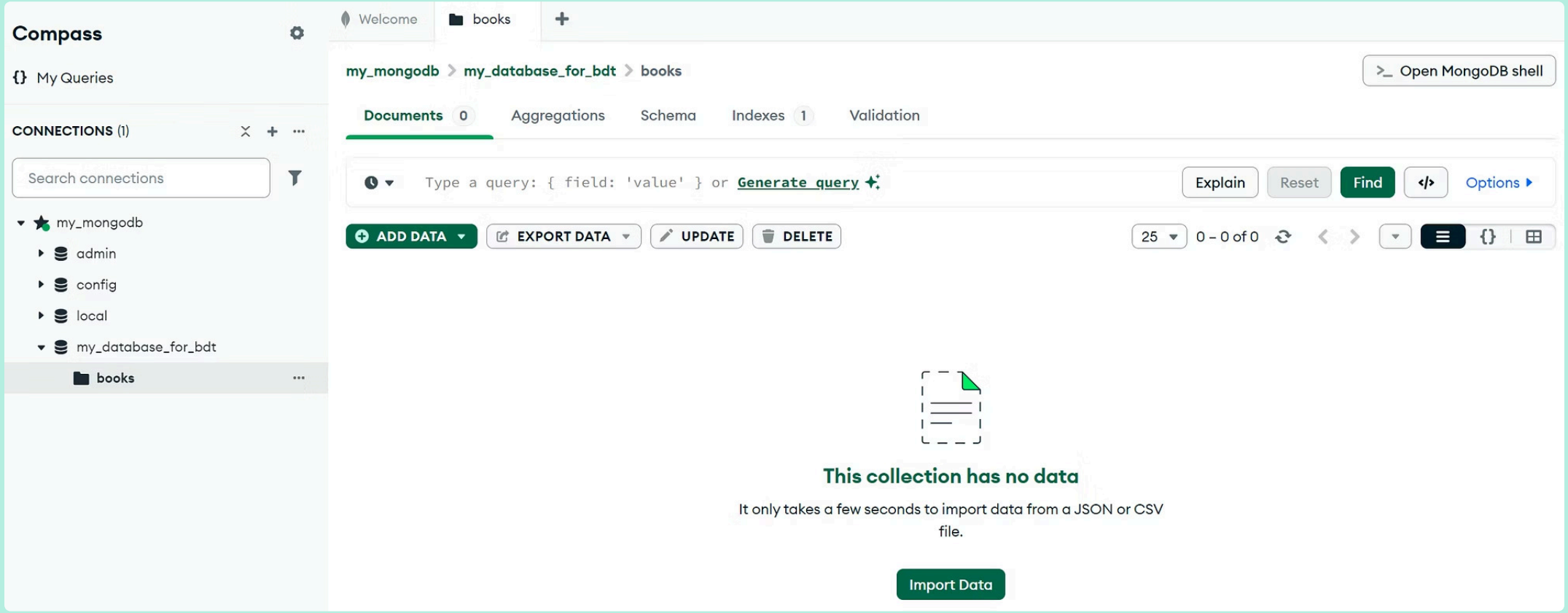


5. Enter the database and collection name. Then click "Create Database" to finalize the setup.



10. Manage the database

The database and collections can be managed directly within MongoDB Compass, and they can also be accessed directly in your programming environment.



For programming examples, please refer to the Jupyter Notebook that has been provided.