

INDEX

Prac. No.	Practical	Date	Sign
1	Install, configure and run Hadoop and HDFS and explore HDFS		
2	Implement Decision tree classification techniques		
3	Implement SVM classification techniques.		
4	Implement of REGRESSION MODLE.		
5	Implement of Simple Linear Regression.		
6	Implement of Multiple Linear Regression.		
7	Implement of Logistic regression.		
8	Read a datafile grades_km_input.csv and apply k-means clustering		
9	Perform Apriori algorithm using Groceries dataset from the R rules package		
10	Implement an application that stores big data in Hbase / MongoDB and manipulate it using R / Python		

PRACTICAL 10

Aim: Implement an application that stores big data in Hbase / MongoDB and manipulate it using R / Python

Description:

MongoDB is a source-available cross-platform document-oriented database program. Classified as a NoSQL database program, MongoDB uses JSON-like documents with optional schemas.

MongoDB is developed by MongoDB Inc. and licensed under the Server Side Public License












Name your organization and project

Organization
Your organization can be a business, team, or an individual

Project Name
Use projects to isolate different environments (development/testing/production)

What is your preferred language?

We'll use this to customize code samples and content we share with you. You can always change this later.

 JavaScript	 C++	 C# / .NET	 Go
 Java	 C	 Perl	 PHP
 Python	 Ruby	 Scala	Other

[Skip](#) [Continue](#)

Step 1 : Sign up and create a cluster.




[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](#).


Cloud Provider & Region


AWS, Mumbai (ap-south-1) ▾




★ Recommended region ⓘ


NORTH AMERICA

 N. Virginia (us-east-1) ★


 Oregon (us-west-2) ★


ASIA

 Singapore (ap-southeast-1) ★


 Mumbai (ap-south-1)

EUROPE

 Frankfurt (eu-central-1) ★

 Ireland (eu-west-1) ★

AUSTRALIA

 Sydney (ap-southeast-2) ★

This is the home page of mongoDB Atlas.

The screenshot displays the MongoDB Atlas interface for a cluster named 'Cluster0' (Version 4.4.6) under the 'Sandbox' tier. The cluster is currently in the 'M0 Sandbox (General)' tier. A 'Connect to Atlas' checklist is visible on the left, showing progress at 20% with the first step 'Build your first cluster' completed. The main panel shows the cluster's configuration, including a warning that it is a Shared Tier Cluster and an 'Upgrade' button. Performance metrics for the last 6 hours are shown on the right, including Operations (R: 0, W: 0), Logical Size (0.0 B), and Connections (0).

Connect to Atlas

Follow this checklist to get started.

20%

- ☒ Build your first cluster
- ☐ Create your first database user
- ☐ Add IP Address to your Access List
- ☐ Load Sample Data (Optional)
- ☐ Connect to your cluster

[No thanks](#)

Cluster0
Version 4.4.6

[CONNECT](#) [METRICS](#) [COLLECTIONS](#) [...](#)

CLUSTER TIER
M0 Sandbox (General)

This is a Shared Tier Cluster
If you need a database that's better for high-performance production applications, upgrade to a dedicated cluster.

[Upgrade](#)

Logical Size: 0.0 B (512.0 MB max)

Operations: R: 0 W: 0 (100.0/s)

Connections: 0 (500 max)

System Status: All Good

©2021 MongoDB, Inc. [Status](#) [Terms](#) [Privacy](#) [Atlas Blog](#) [Contact Sales](#)

[Get Started](#) [Feature Requests](#)



Explore Your Data

- **Find:** run queries and interact with documents
- **Indexes:** build and manage indexes
- **Aggregation:** test aggregation pipelines
- **Search:** build search indexes

[Load a Sample Dataset](#)[Add My Own Data](#)[Learn more in Docs and Tutorials](#)

Step 2 : Click on collections to create and view existing databases.

Step 3 : Click on 'Add My Own Data' to create a database.

Create Database

DATABASE NAME ?

govind_db

COLLECTION NAME ?

govind

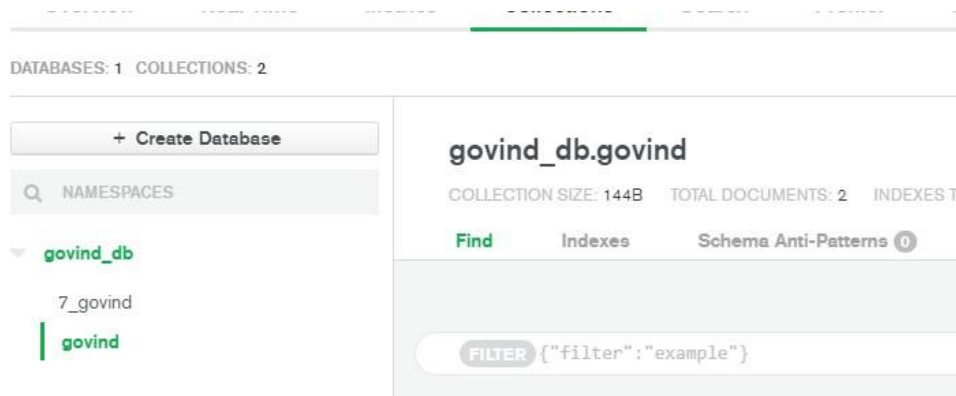
☐ Capped Collection

Before MongoDB can save your new database, a collection name must be specified at the time of creation.

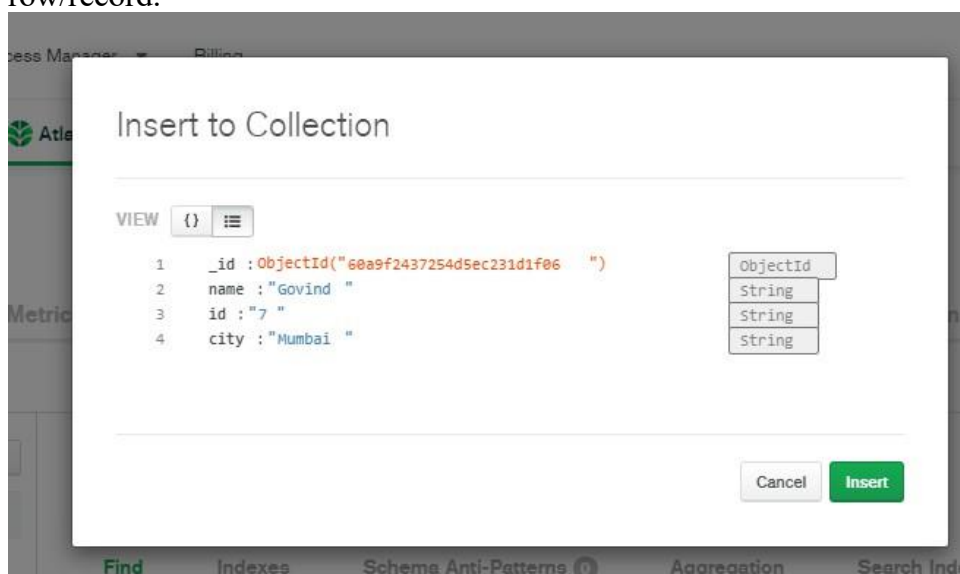
Cancel

Create

Step 4 : Click on insert document to add records.



Since MongoDB is a No-SQL database, so you can add 'n' number of columns for any row/record.



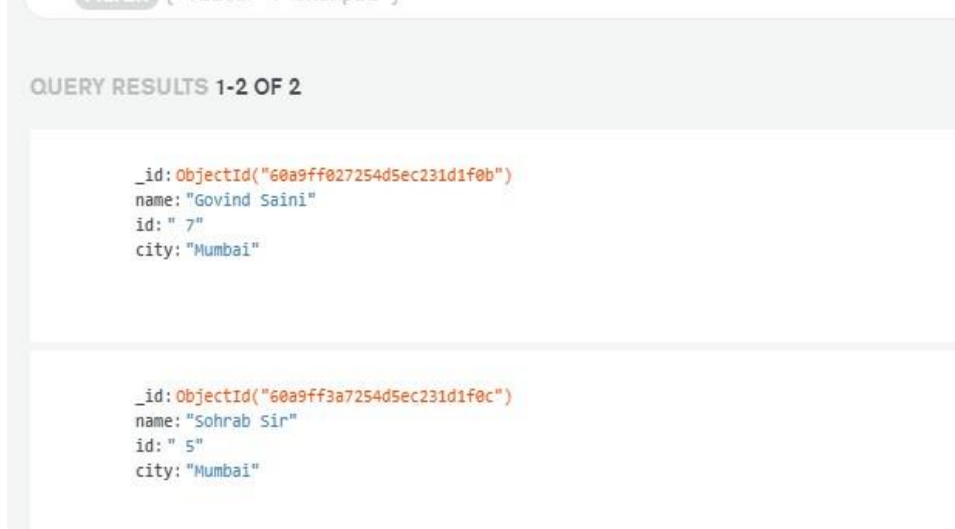
Perform updating data



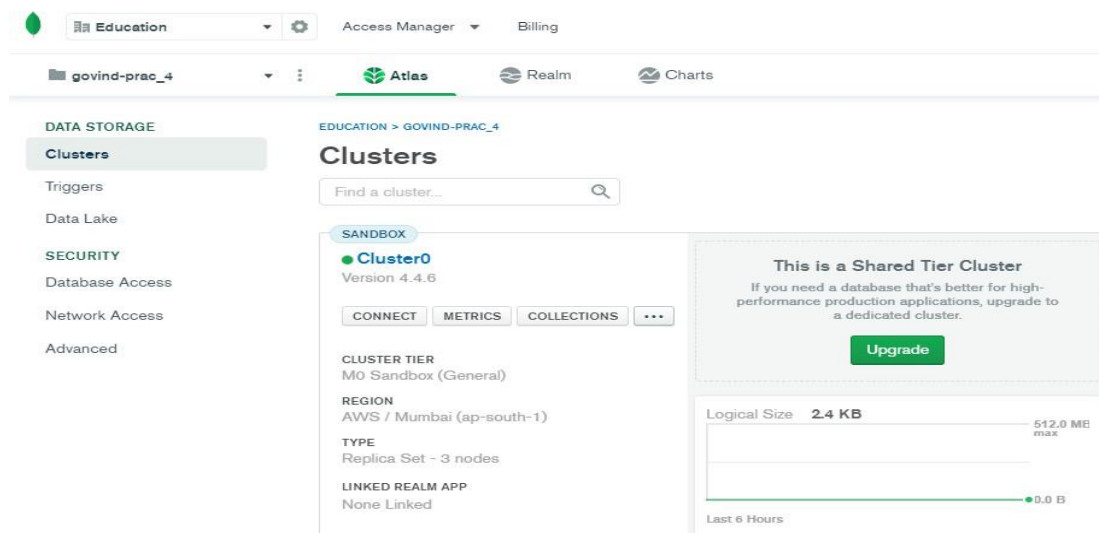
Performing deleting data



Performing Insert data



Step 5 : To start with the connection click on Overview, and then click on Connect.



Step 6 : Select on add your current IP and create a MongoDB user.

×

Connect to Cluster0

Setup connection security

Choose a connection method

Connect

You need to secure your MongoDB Atlas cluster before you can use it. Set which users and IP addresses can access your cluster now. [Read more](#)

You're ready to connect. Choose how you want to connect in the next step.

1

Add a connection IP address

✓ An IP address has been added to the IP Access List. *Add another address in the [IP Access List](#) tab.*

2

Create a Database User

✓ A MongoDB user has been added to this project. *Not yours? Create one in the [MongoDB Users](#) tab.*

You'll need your MongoDB user's credentials in the next step.

Close

Choose a connection method

Step 7 : Click on 'Connect your application'.

×

Connect to Cluster0


✓ Setup connection security

Choose a connection method

Connect


Choose a connection method [View documentation](#)

Get your pre-formatted connection string by selecting your tool below.




Connect with the mongo shell
Interact with your cluster using MongoDB's interactive Javascript interface

>



Connect your application
Connect your application to your cluster using MongoDB's native drivers

>



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

>

Go Back

Close

Step 8 : Select the driver as 'Python' and version as '3.6 or later'. (Select the version as 3.6 or later only if your Python's version is 3.6 or later.)

Connect to Cluster0

✓ Setup connection security

✓ Choose a connection method

Connect

1 Select your driver and version

DRIVER

Python

VERSION

3.6 or later

2 Add your connection string into your application code

☐ Include full driver code example

```
mongodb+srv://dbGovind:<password>@cluster0.m6shh.mongodb.net/myFirstDatabase?retryWrites=true&w=majority
```

Replace **<password>** with the password for the **dbGovind** user. Replace **myFirstDatabase** with the name of the database that connections will use by default. Ensure any option params are [URL encoded](#).

Having trouble connecting? [View our troubleshooting documentation](#)

Step 9 : Write the code given below in a Python file.

```

prc4.py - C:/Python27/prc4.py (2.7.17)
File Edit Format Run Options Window Help

import pymongo
from pymongo import MongoClient
client = pymongo.MongoClient("mongodb+srv://dbGovind:GmongoDB123@cluster0.m6shh.mongodb.net/myFirstDatabase?retryWrites=true&w=majority")
db = client.get_database('govind_db')
records = db.govind
db = client.test
print(records.count_documents({}))
print(list(records.find()))
  
```

Output :

```

===== RESTART: C:/Python27/prc4.py =====
2[{"_id":{"$_id":"60a9ff027254d5ec231dlf0b"},"name":"Govind Saini","id":" 7","city":"Mumbai"}{"_id":{"$_id":"60a9ff3a7254d5ec231dlf0c"},"name":"Sohrab Sir","id":" 5","city":"Mumbai"}]
>>> |
  
```