

Practical :4

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

Requirement:

- a. Python Package: PyMongo
- b. Mongo Database

Step A: Install Mongo database

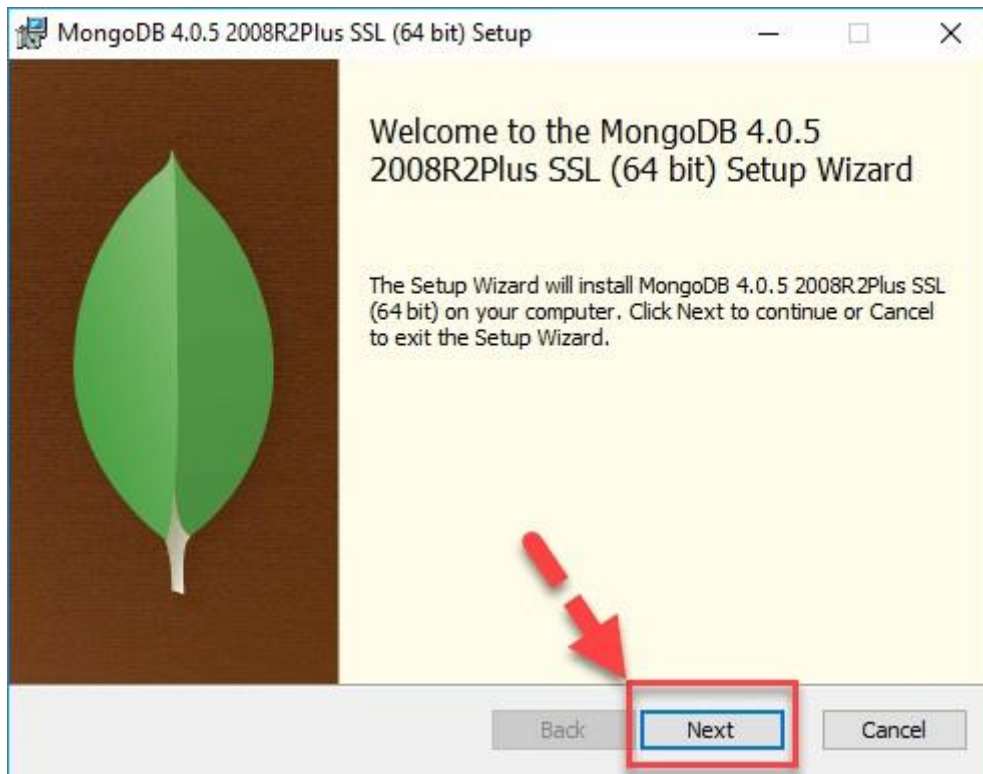
Step 1) Go to (<https://www.mongodb.com/download-center/community>) and Download MongoDB Community Server. We will install the 64-bit version for Windows.

Select the server you would like to run:



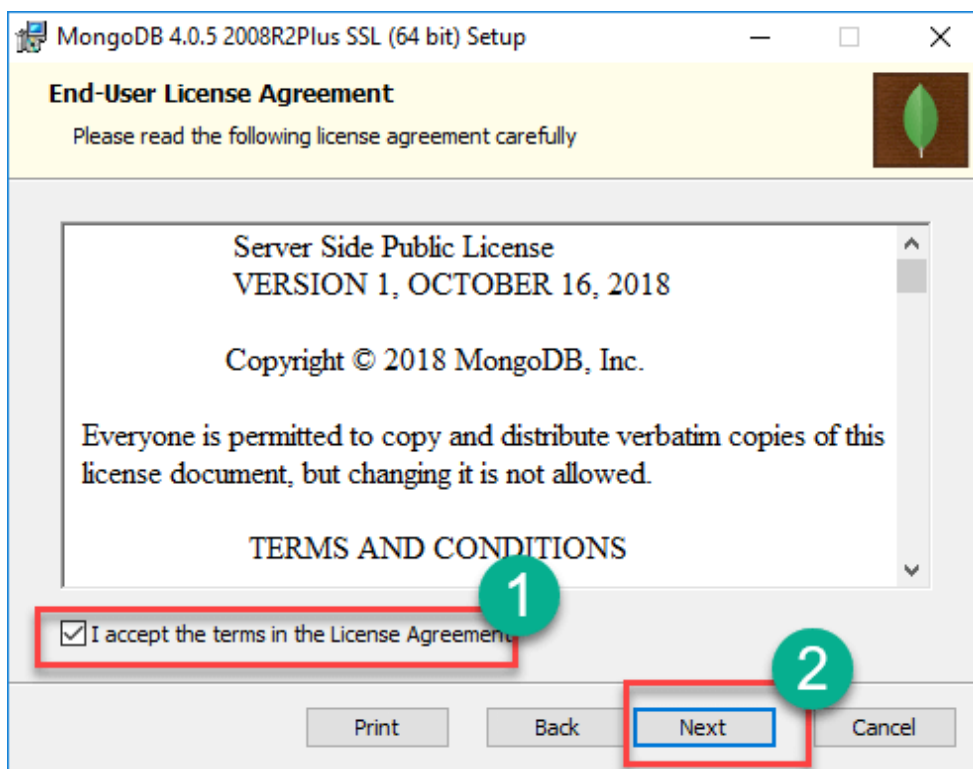
The screenshot shows the MongoDB download page. At the top, there are two buttons: "MongoDB Community Server" (highlighted in green) and "MongoDB Atlas". Below these, there are three dropdown menus: "Version" (set to "4.0.5 (current release)"), "OS" (set to "Windows 64-bit x64"), and "Package" (set to "MSI"). A red dashed arrow points from the "Version" dropdown to the "Download" button. The "Download" button is a green rectangle with the text "Download" in white. Below the button, the URL https://fastdl.mongodb.org/win32/mongodb-win32-x86_64-2008plus-ssl-4.0.5-signed.msi is displayed.

Step 2) Once download is complete open the msi file. Click Next in the start up screen

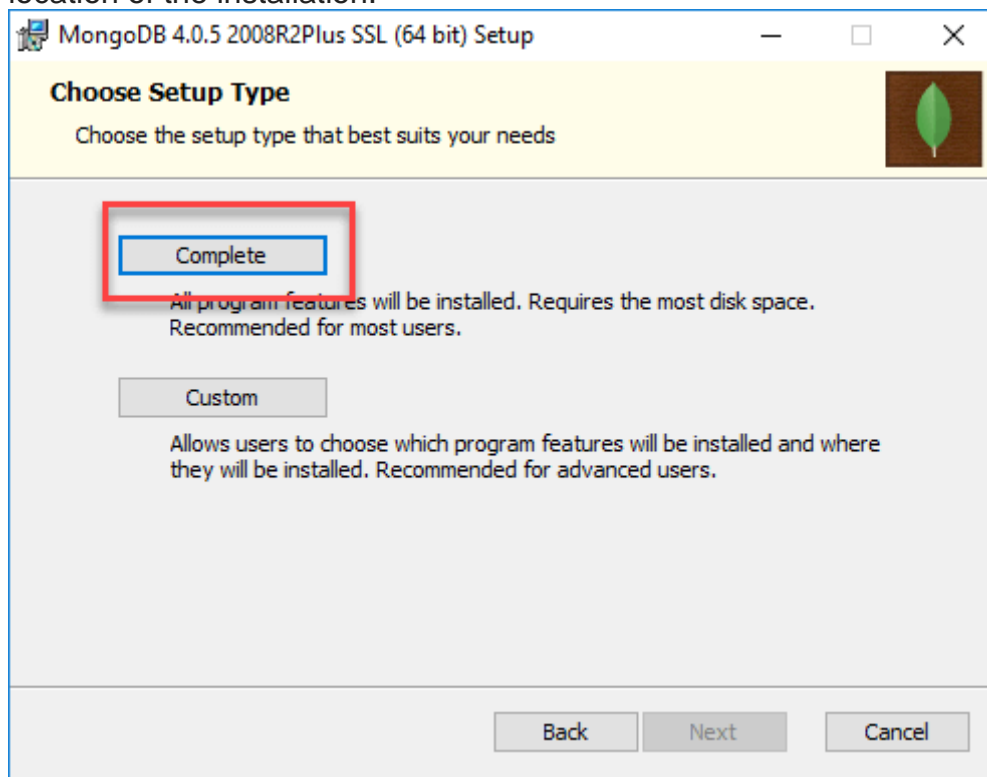


Step 3)

1. Accept the End-User License Agreement
2. Click Next



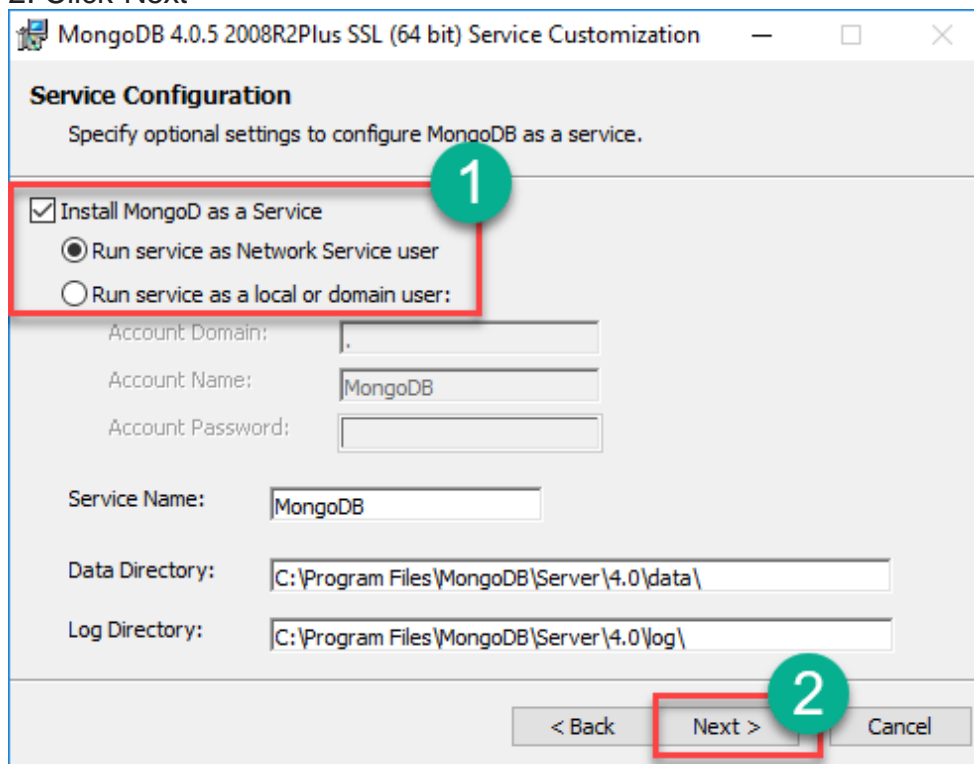
Step 4) Click on the "complete" button to install all of the components. The custom option can be used to install selective components or if you want to change the location of the installation.



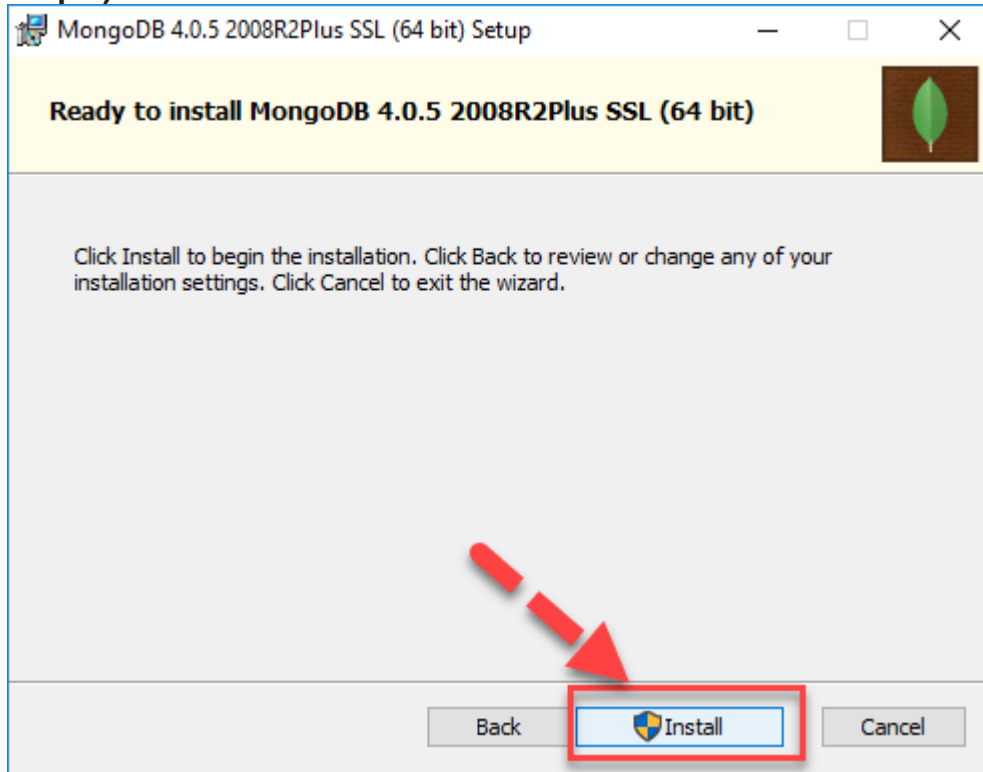
Step 5)

1. Select "Run service as Network Service user". make a note of the data directory, we'll need this later.

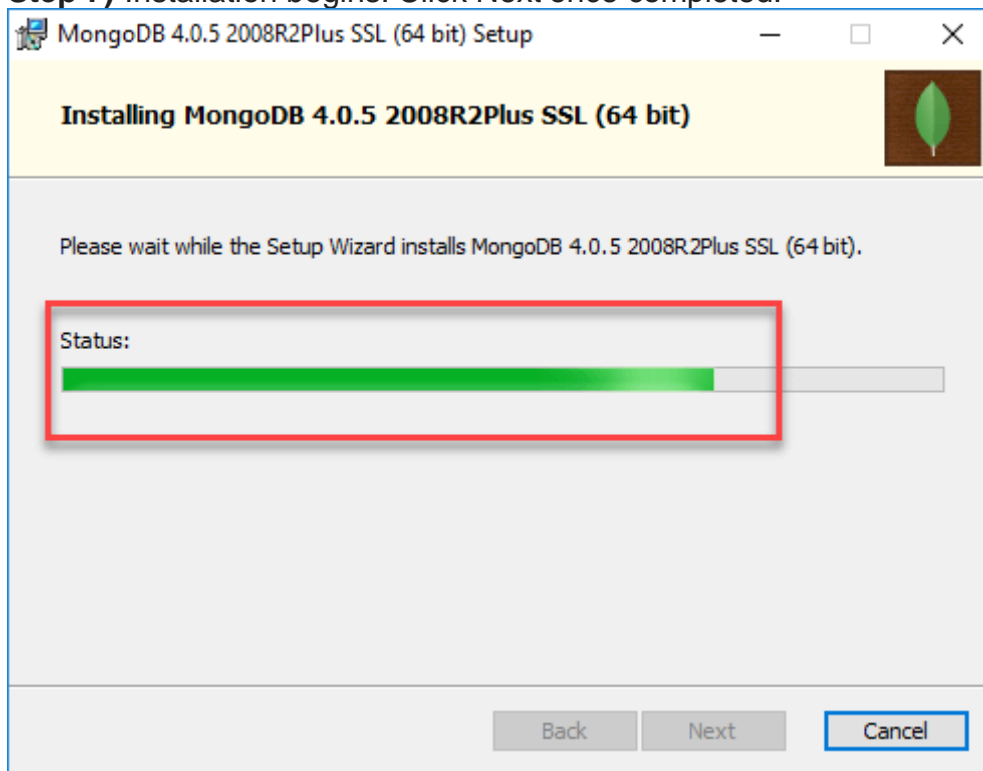
2. Click Next



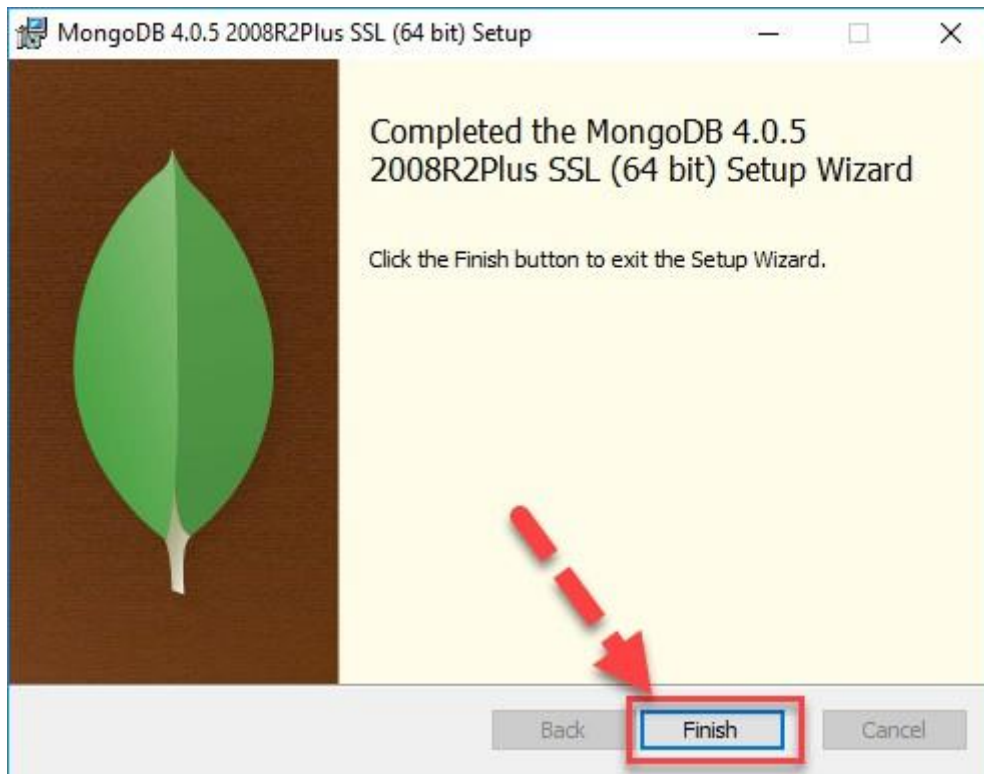
Step 6) Click on the Install button to start the installation.



Step 7) Installation begins. Click Next once completed.



Step 8) Click on the Finish button to complete the installation



Program 1: Displaying the database name:

```
import pymongo
myclient = pymongo.MongoClient("mongodb://localhost:27017/")
mydb = myclient["mybigdata"]
print(myclient.list_database_names())
```

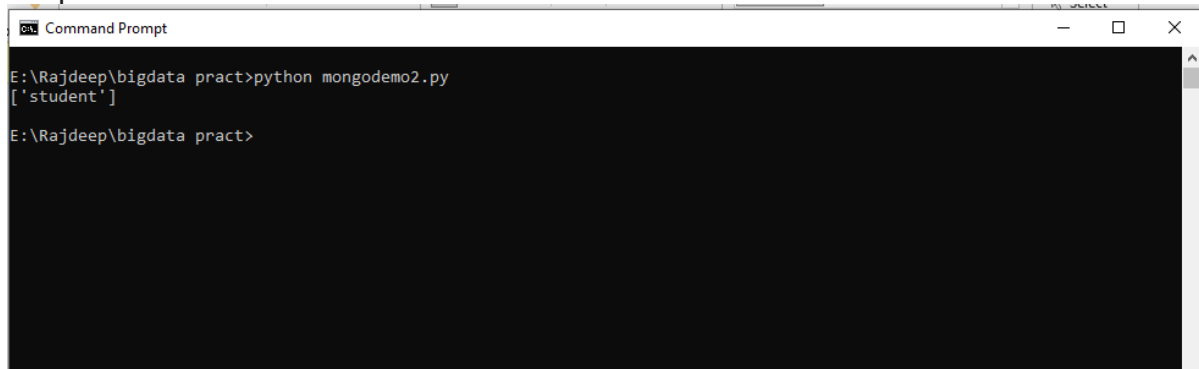
Output:

```
Command Prompt
E:\Rajdeep\bigdata pract>python mongodemo1.py
['admin', 'config', 'local', 'mybigdata']
E:\Rajdeep\bigdata pract>
```

Program 2: Creating collection:

```
import pymongo
myclient = pymongo.MongoClient("mongodb://localhost:27017/")
mydb = myclient["mybigdata"]
mycol=mydb["student"]
print(mydb.list_collection_names())
```

Output:

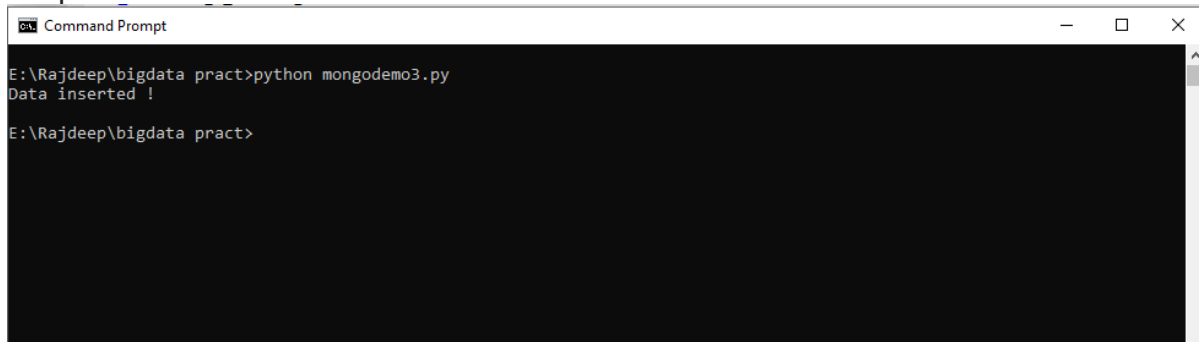


```
Command Prompt
E:\Rajdeep\bigdata pract>python mongodemo2.py
['student']
E:\Rajdeep\bigdata pract>
```

Program 3: Inserting Data

```
import pymongo
myclient = pymongo.MongoClient("mongodb://localhost:27017/")
mydb = myclient["mybigdata"]
mycol=mydb["student"]
mydict={"name":"vai", "address":"bhy"}
x=mycol.insert_one(mydict)
print("Data inserted !")
```

Output:

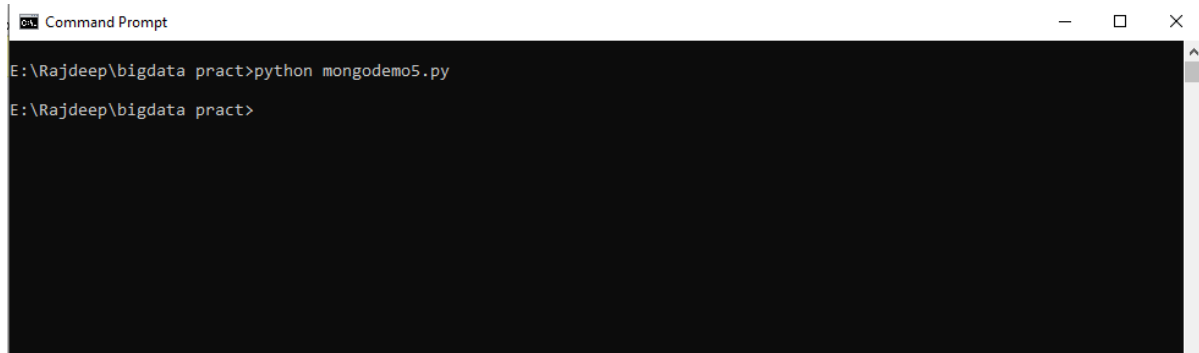


```
Command Prompt
E:\Rajdeep\bigdata pract>python mongodemo3.py
Data inserted !
E:\Rajdeep\bigdata pract>
```

Program 4: Insert Multiple data into Collection

```
import pymongo
myclient = pymongo.MongoClient("mongodb://localhost:27017/")
mydb = myclient["mybigdata"]
mycol=mydb["student"]
mylist=[{"name":"Ganesh", "address":"Mumbai"}, {"name":"Varun",
"address":"Mumbai"},
{"name":"Prasoon", "address":"Pune"}, {"name":"Satish", "address":"Pune"},]
x=mycol.insert_many(mylist)
print("Data inserted !")
```

Output:



```
Command Prompt
E:\Rajdeep\bigdata pract>python mongodemo5.py
E:\Rajdeep\bigdata pract>
```

Program 5: Displaying the collection data:

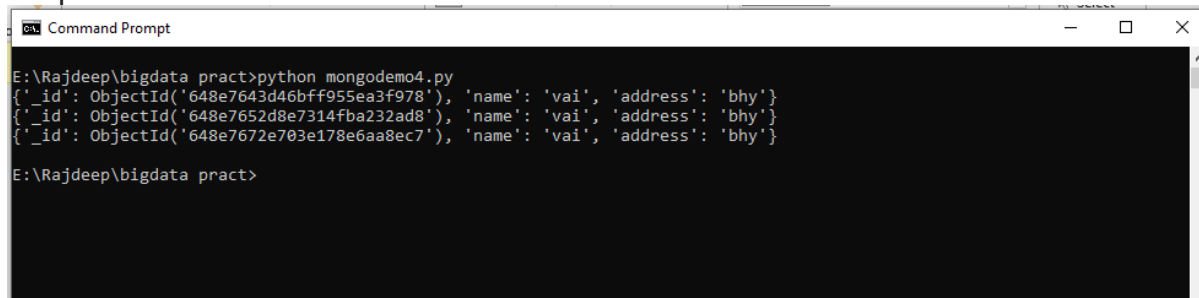
```
import pymongo
myclient = pymongo.MongoClient("mongodb://localhost:27017/")
mydb = myclient["mybigdata"]
mycol = mydb["student"]

myquery = { "name": "Vai" }

mydoc = mycol.find(myquery)

for x in mydoc:
    print(x)
```

Output:



```
Command Prompt
E:\Rajdeep\bigdata pract>python mongodemo4.py
{'_id': ObjectId('648e7643d46bff955ea3f978'), 'name': 'vai', 'address': 'bhy'}
{'_id': ObjectId('648e7652d8e7314fba232ad8'), 'name': 'vai', 'address': 'bhy'}
{'_id': ObjectId('648e7672e703e178e6aa8ec7'), 'name': 'vai', 'address': 'bhy'}
E:\Rajdeep\bigdata pract>
```

