BIG DATA ANALYTICS

MINI PROJECT

Experiment No.: 11

To implement a weather forecast dataset using MongoDB

Name	Roll No	Batch
Shreya Raju Katole	18ET1054	B/B2
Vidhi Hemant Kandalkar	18ET1093	B/B1

Under the Guidance: Mr.Dayanand Sir

Experiment No.: 11

Aim: To implement a weather forecast dataset using MongoDB.

• What will you learn by performing this experiment?
How different MongoDB operations are used in mini-project.

• Software/Languages Required: MongoDB.

• Theory:
What is MongoDB?
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 (SSPL).

MongoDB Community Server
As of October 2018, MongoDB is released under the Server Side Public License (SSPL), a license developed by the project. It replaces the GNU Affero General Public License, and is nearly identical to the GNU General Public License version 3, but requires that those making the software publicly available as part of a "service" must make the service's entire source code available under this license. The SSPL was submitted for certification to the Open Source Initiative but later withdrawn. The language drivers are available under an Apache License. In addition, MongoDB Inc. offers proprietary licenses for MongoDB. The last versions licensed as AGPL version 3 are 4.0.3 (stable) and 4.1.4.

MongoDB has been removed from the Debian, Fedora, and Red Hat Enterprise Linux distributions due to the licensing change. Fedora determined that the SSPL version 1 is not a free software license because it is "intentionally crafted to be aggressively discriminatory" towards commercial users.

What is Kaggle?
Kaggle allows users to find and publish data sets, explore and build models in a web-based data-science environment, work with other data scientists and machine learning engineers, and enter competitions to solve data science challenges.

- **Procedure/ Program:**

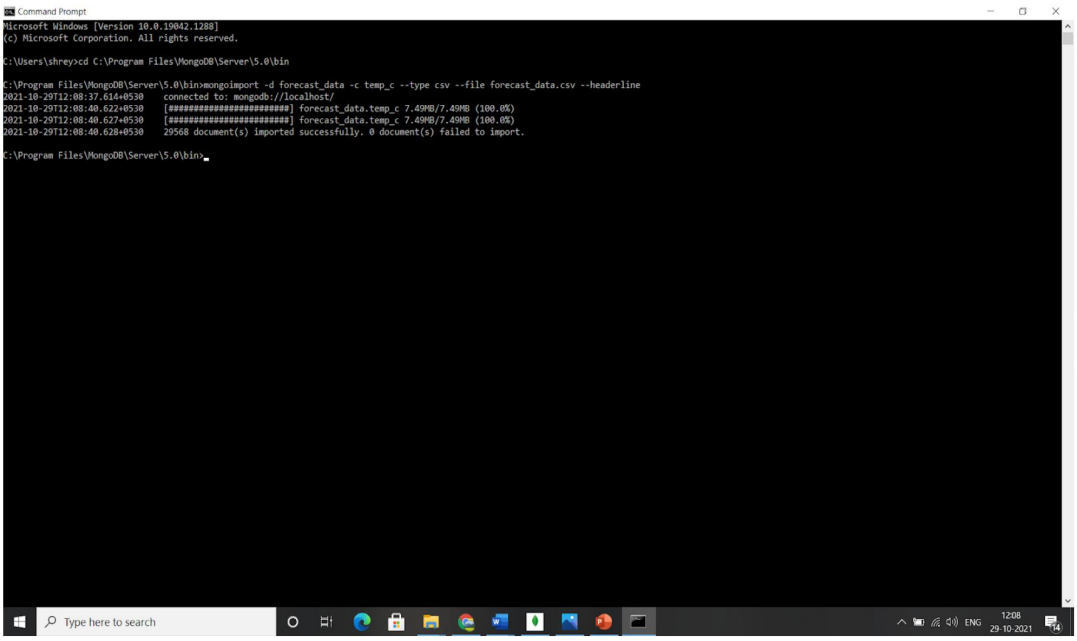
1. Get CSV dataset from Kaggle.com
2. Install MongoDB Community Server
3. Install MongoDB toolkit
4. In command prompt change directory where MongoDB bin folder is located.
5. Import and Convert CSV file to MongoDB code.
6. Open mongo.exe and execute the following commands.

- **Results**

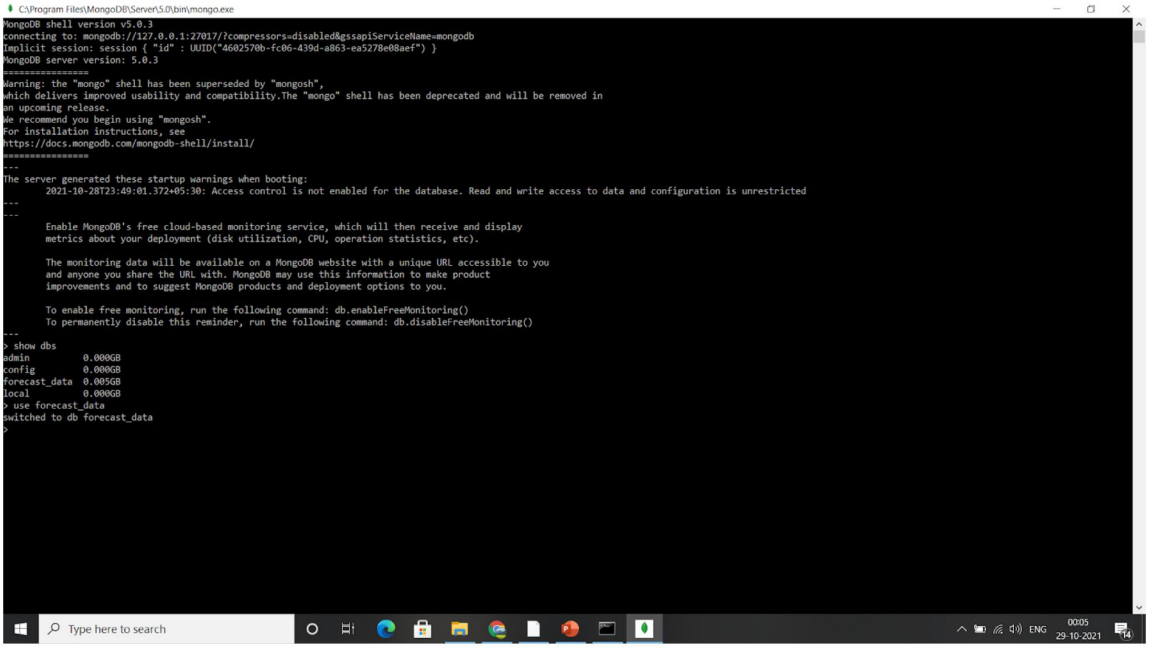
Step 1:

Import mongodb in cmd

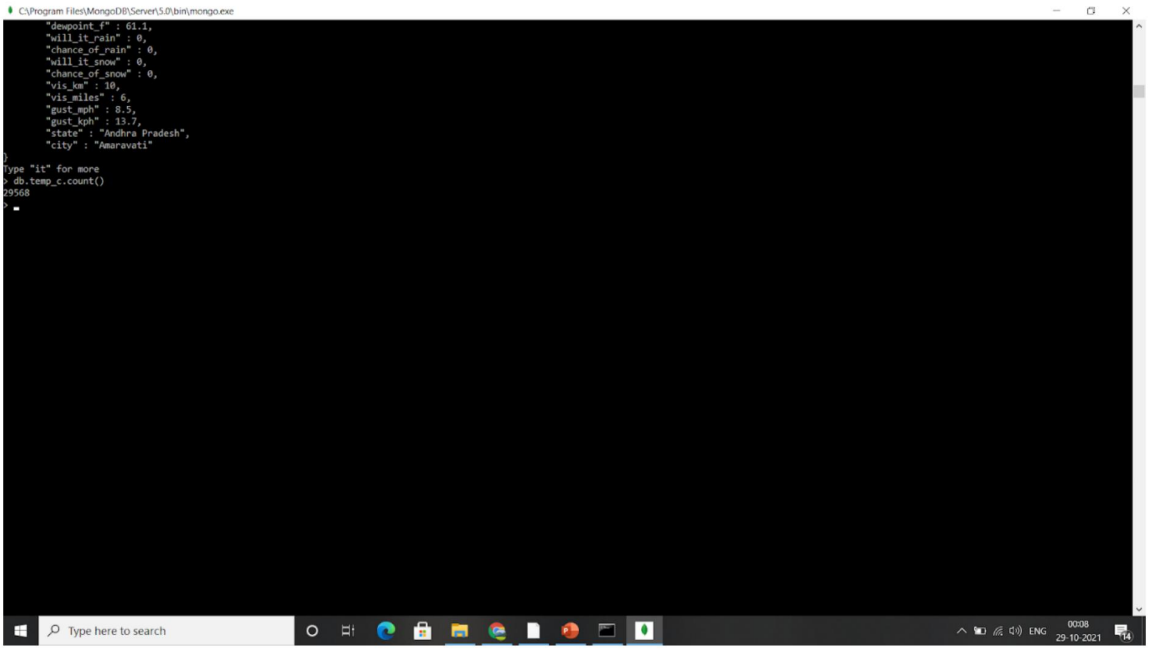
mongoimport -d forecast_data -c temp_c --type csv --file forecast_data.csv --headerline



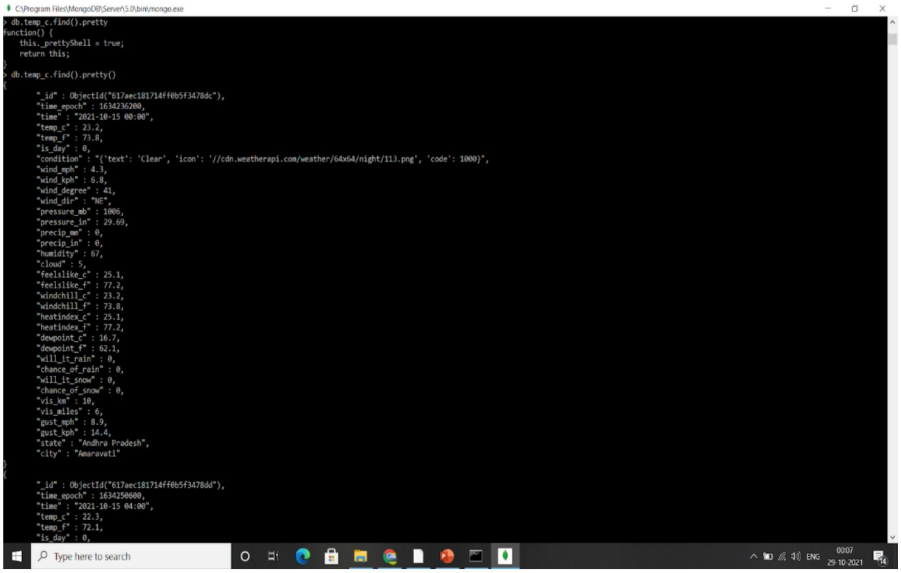
Step 2:
Show dbs

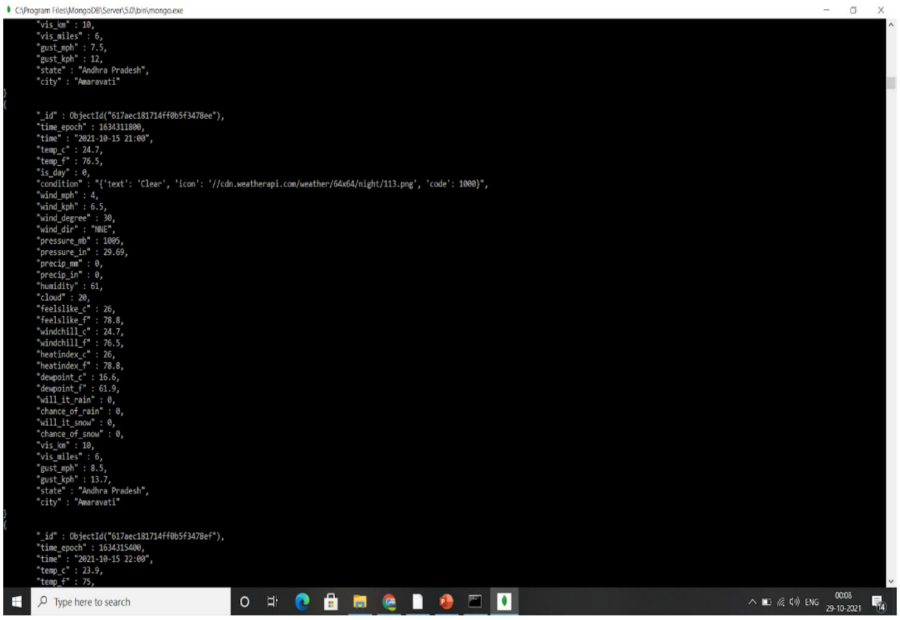


Step 3:
db.temp_c.count()

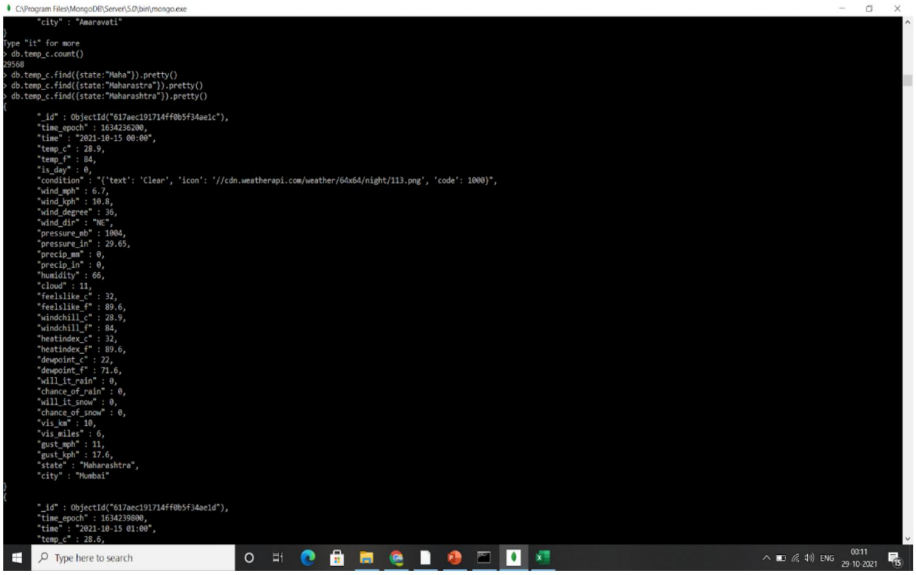


Step 4:
db.temp_c.find().pretty()

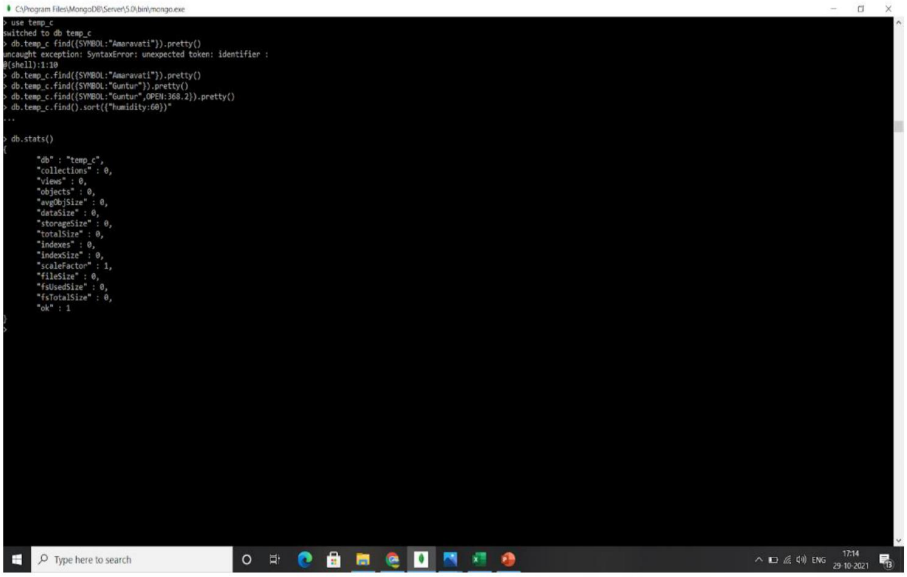




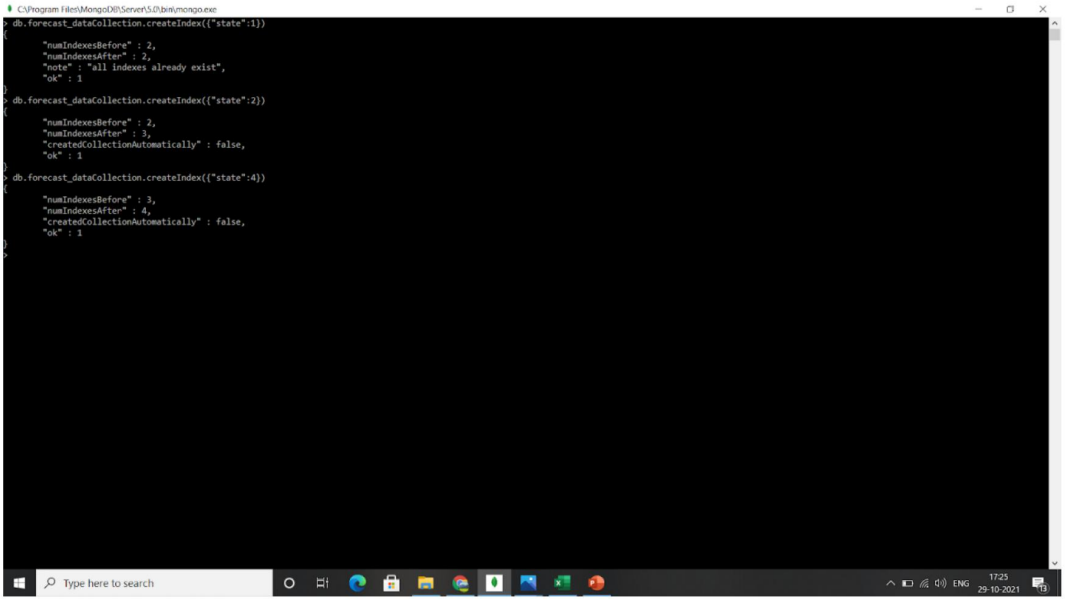
Step 5:
db.temp_c.find({state:"Maharashtra"}).pretty()



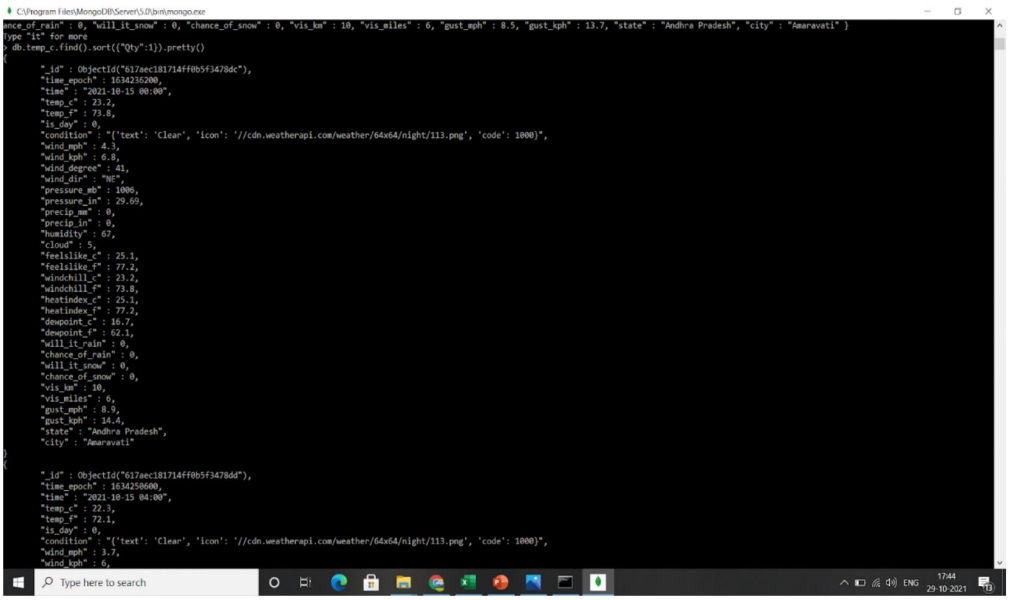
Step 6:
db.stats()



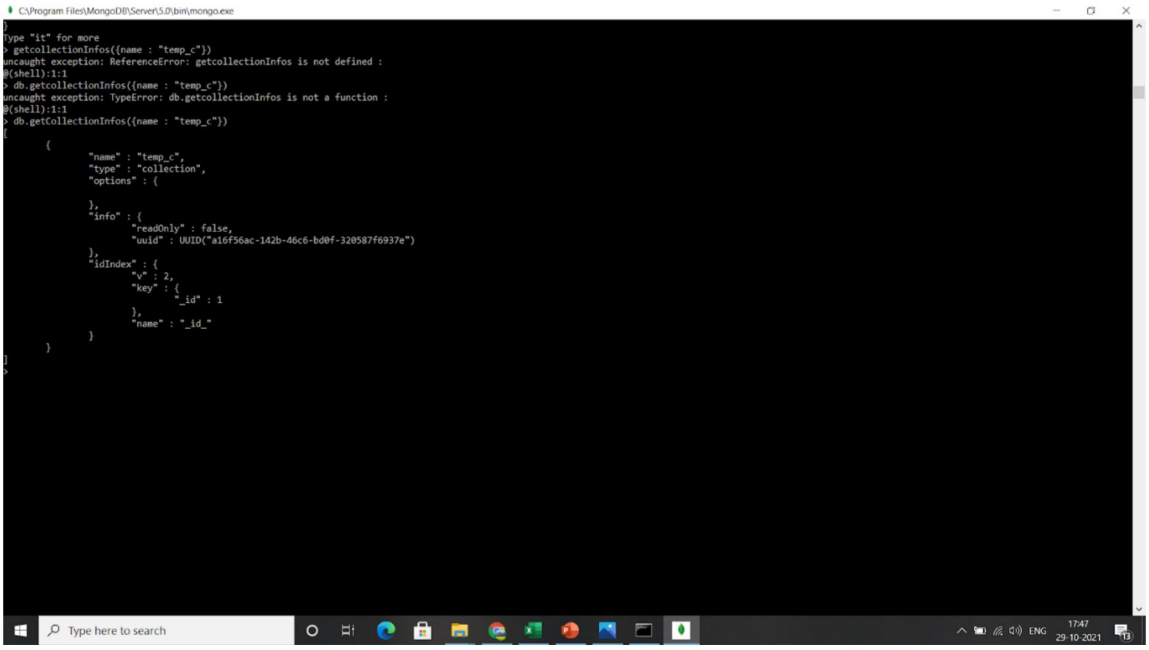
Step 7:
db.forecast_dataCollection.createIndex({“state”: 1})



Step 8:
db.temp_c..find().sort({“Qty”:1}).pretty()



Step 9:
db.getcollectionInfos({name : “temp_c”})



- **Conclusion and Discussion:**

Using MongoDB we are able to analyze the real-time data of Weather Forecast Easily and efficiently. By using different commands we were able to find open, find data related to respective cities or states.