

# PREPARE MONGODB

## Create file repo

```
sudo vim /etc/yum.repos.d/mongodb-org-6.0.repo

[mongodb-org-6.0]
name=MongoDB Repository
baseurl=https://repo.mongodb.org/yum/amazon/2/mongodb-org/
6.0/x86_64/
gpgcheck=1
enabled=1
gpgkey=https://www.mongodb.org/static/pgp/server-6.0.asc
```

## Install MongoDB

```
sudo yum install -y mongodb-org
```

## Start MongoDB and check status

```
sudo systemctl start mongod
sudo systemctl status mongod
```

## Load data from AWS S3 to MongoDB

Here we use database name `transaction_db` and the data in `card_transactions.csv` will be loaded into collection `card_transactions`.

```
aws s3 cp s3://history-transactions/card_transactions.csv - |  
mongoimport --db transaction_db --collection  
card_transactions --type csv --headerline
```

Check if data is correctly loaded.

```
[[hadoop@ip-172-31-66-21 ~]$ aws s3 cp s3://history-transactions/card_transactions.csv - | mongoimport --db transaction_  
db --collection card_transactions --type csv --headerline  
2023-05-13T18:13:42.314+0000    connected to: mongodb://localhost/  
2023-05-13T18:13:43.037+0000    53292 document(s) imported successfully. 0 document(s) failed to import.
```

```
[hadoop@ip-172-31-74-203 ~]$ mongosh  
Current Mongosh Log ID: 645d819a9891c8d64e68a09d  
Connecting to:      mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+1.8.2  
Using MongoDB:      6.0.5  
Using Mongosh:      1.8.2  
  
For mongosh info see: https://docs.mongodb.com/mongodb-shell/  
  
To help improve our products, anonymous usage data is collected and sent to MongoDB periodically (https://www.mongodb.com/legal/privacy-policy).  
You can opt-out by running the disableTelemetry() command.  
  
-----  
The server generated these startup warnings when booting  
2023-05-12T00:00:00.266+00:00: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted  
2023-05-12T00:00:00.266+00:00: vm.max_map_count is too low  
-----  
  
Enable MongoDB's free cloud-based monitoring service, which will then receive and display  
metrics about your deployment (disk utilization, CPU, operation statistics, etc).  
  
The monitoring data will be available on a MongoDB website with a unique URL accessible to you  
and anyone you share the URL with. MongoDB may use this information to make product  
improvements and to suggest MongoDB products and deployment options to you.  
  
To enable free monitoring, run the following command: db.enableFreeMonitoring()  
To permanently disable this reminder, run the following command: db.disableFreeMonitoring()  
-----  
  
test> use transaction_db  
switched to db transaction_db  
transaction_db> db.card_transactions.findOne()  
{  
  _id: ObjectId("645d8194389154a318c0c087"),  
  card_id: Long("348702330256514"),  
  member_id: Long("37495066290"),  
  amount: 4310362,  
  postcode: 33946,  
  pos_id: Long("614677375609919"),  
  transaction_dt: '11-02-2018 00:00:00',  
  status: 'GENUINE'  
}
```