

(Updated on 12/26/2021)

MongoDB Orientation: This document includes Orientation assignments, as well as other setup information.

Note: If you set up a separate Anaconda environment for this course and are not using (base), then please do this:

<https://medium.com/@nrk25693/how-to-add-your-conda-environment-to-your-jupyter-notebook-in-just-4-steps-abeab8b8d084>

Sign up for MDB Atlas free-tier cluster in the cloud (if you already have one, you will only have to add new collections as required)

- While you are following their instructions, *please, please, write down your Atlas username and password and your cluster/database-user username and password and save them somewhere easy to find.*
- Signup link: <https://www.mongodb.com/cloud/atlas?imp=docs>
- Below is a fairly recent youtube video walking you through the process of signing up for a free-tier cluster.
 - <https://www.youtube.com/watch?v=1duX6Nfevhc>
(third-party informative video)
- And below is my video, where I'm doing the same thing as the above video, with a more recent interface, updated on 12/05/2021

Video: [1-MDB-Atlas-Cluster](#)

Install MongoDB locally on Windows (so that we can use the Mongo shell in the cmd/terminal) (Mac instructions are below the Windows instructions, and screenshot submissions are below the Mac instructions)

- You can avoid all of this installation by simply using the VM!! You will still have to connect from Compass and from mongodb to your Atlas cluster, but the local installation has been done for you.

Important tip when using the VM: If you try to press CTRL-ALT-DEL when you are running the remote desktop on the VM, it will work on your local machine, not on the VM. In order to do a CTRL-ALT-DEL on the VM, you first have to press CTRL-SHIFT-ALT on your local computer. This will bring up a window on the left side of the VM which gives you an onscreen keyboard option. Select the onscreen keyboard option that allows a combination of keys. A large onscreen keyboard will then pop up. On that keyboard, you can then click CTRL, then ALT, then DEL, and it will bring up the control menu on the VM. You can select the task manager from there. You may need that if you exit from Compass (later, in the MongoDB part of the course) without clicking on connect-->exit. ***So it is very important to always click on Connect → Exit when using Compass, but essential when using the VM!!***

1. *Install a local instance of MongoDB:*
 - You can download the latest version for your OS [here](#). The current version of MongoDB as of this writing is 5.0.4. The Atlas Free-Tier is also providing you with a 4.4.10 instance. (If this changes, and you want to make sure that your local version is fully compatible, you can check [here](#). Sometimes, when running locally, you may get a warning that your local version doesn't match the Atlas installed version. So far, this has not posed any problems for our course.)
 - [Here is a clear, short video showing the installation of MongoDB on Windows](#), together with Compass, and adding it to the path. The video installs the community server.
 2. *Create a new folder (the video may have already instructed, but it bears repeating):*
 - a. `mkdir c:\data\db` (on Windows)
 3. *Install the Mongo Database Tools:*
 - a. With newer versions, you have to install the [database Tools](#) separately. Database tools include mongoimport, which we will be using in class. Download and installation instructions for your platform are [here](#).
 - b. [Here is another short video installing MongoDB, Mongo Database Tools](#) and Robo3T (an alternative to Compass that we are not using in the course, but you are free to try if you want), and adding them all to the path.
 4. *Optional: Install Mongo Shell (mongosh), a shell specifically designed for use with mongod. In my examples, I am using a regular terminal, but you can certainly use the new mongosh if you want.*
- I use the Enterprise edition in my installation. However the community edition will work fine for this course.
 - I tested everything with 4.x and with 5.x and it worked fine. Instructions have been updated, videos have been annotated. Of course, MongoDB may continue to make changes... We can't help that, but we can try to keep up!
 - If you prefer written instructions, they are below:
 - a. [Select your OS](#) from the drop-down on the right of the page and download.
 - b. For Windows, use the MSI package.
 - c. In the installation dialogue box, there is a checkbox to install Compass. Check that box!! If you don't, you'll have to install Compass separately later. (If you have chosen to install Robo3T instead, then don't install Compass. You don't need both.)
 - d. Find the folder `C:\Program Files\MongoDB\Server\5.x\bin`
 - i. Add this to the path
 1. Control Panel → System → Advanced System Settings
→ Environment Variables (a button in the settings dialogue box)

2. →under System Variables (the bottom half), click on:
 3. →Path→Edit→New
 4. Enter the path, and click MoveUp until it's at the top of the path
 5. Do the same for the Mongo Tools, which should be found in the same Server\5.x directory.
 6. Click OK.
- ii. Close and reopen terminal

Install MongoDB locally on MacOS: download center-->Enterprise server tab.

IMPORTANT NOTE FOR POST-CATALINA MAC USERS:

The newer MacOS do not allow you to make changes in the root directory. However, MongoDB local install is still placing your database path where it is now read only. The newer mac instructions on the mongodb website are difficult to follow and I think a bit convoluted. There are instructions on the web using brew (homebrew) for the installation, but it's still a bit of a workaround. If you encounter any connection or installation issues, you should simply use the VM for the MongoDB labs.

That said, Brian Nguyen, a former DB student, has kindly outlined [instructions for successfully installing MongoDB on your post-Catalina Mac](#). You can try those instructions. Again, if they do not work for you, please just use the VM. Hopefully, by the time we get to the MongoDB module of the course, MongoDB will have updated their installation file, and these problems will have been fixed.

Official instructions from the mongodb website (very problematic):

- <https://docs.mongodb.com/manual/tutorial/install-mongodb-enterprise-on-os-x/>
- You may find the instructions at the above link sufficient with no further details needed. Note that the instructions have several warnings on changes that need to be made to conform with the new Mac OS architecture. Use at your own risk. Or use the VM.
 - Scroll down to step 1: "Download the tarball"
 - (Optionally continue down the page for full detailed instructions)
 - Click on "MongoDB Download Center"
 - Click on the green "Download" button
 - Navigate to your local folder where this was downloaded.
 - Double click on the .tgz file. This will unpack the file.
 - Good idea to put it into your home directory.
 - Also good idea to rename to something shorter. Ex: "mongodb-e"
 - Mongo is located in the bin folder

- Depending on the version of MacOS running, follow the next step accordingly. (To check the OS, go to the apple symbol in the top left and click on "About This Mac". There it will show macOS <name of OS>)
 - macOS Mojave (and older)
 - When in the home directory, update the path environment variable in a hidden file name ".bash_profile" (In your terminal, you can see hidden files using "ls -a" and you can see the ".bash_profile")
 - If you don't see one, create it, create a new file, save it as ".bash_profile" Be sure to save it in your home directory!!
 - To edit this file in the terminal (or create it for the first time) type in "nano .bash_profile".
 - You can add a line:
 - export PATH="/Users/<your username>/<your mongo path>/bin:\$PATH"
 - Click ^x (control + x) to exit the file, then click "y" to say yes to saving the file. If you did not provide a name when opening nano, it will prompt you for a name here.
 - Close and reopen terminal
 - Create the database folder:
 - sudo mkdir -p /data/db
 - macOS Catalina (and newer)
 - **Note: when you try to run any mongo commands in the terminal you may see a popup saying "<program name> cannot be opened because the developer cannot be verified".**
 - Click "cancel".
 - Open System Preferences and go to Security & Privacy.
 - Under the General tab and click the "Allow Anyway" button next to the statement with the same name of the program you are trying to run.
 - Run program again. Another popup will show, but this time click "open" and the program will run.
 - When in the home directory update the path environment variable in a hidden file name ".zshenv" (In your terminal, you can see hidden files using "ls -a" and you can see the ".zshenv")
 - If you don't see one, create it, create a new file, save it as ".zshenv" Be sure to save it in your home directory!!
 - To edit this file in the terminal (or create it for the first time) type in "nano .zshenv"
 - You can add a line:
 - export PATH="/Users/<your username>/<your mongo path>/bin:\$PATH"
 - Click ^x (control + x) to exit the file, then click "y" to say yes to saving the file. If you did not provide a name when opening nano, it will prompt you for a name here.
 - Close and reopen terminal
 - Create the database folder:
 - **Note: there is new security measures in place, so now the database folder cannot be at /data/db.**

- You can choose where the database folder goes as long as it is not in the root / directory. You may simply follow these directions:
 - Go into the mongo folder, previously named “mongodb-e”
 - `cd mongodb-e`
 - Inside that folder type in:
 - `mkdir -p ./data/db`
- **Note: when running mongo in the terminal, you must also specify the database path as an argument. If you followed the directions above you can run this when in the home directory:**
 - `mongod --dbpath=./mongodb-e/data/db`

FOR BOTH WINDOWS AND MAC USERS: TAKE SCREEN SHOTS AS INDICATED BELOW

- **Note: for Mac users on macOS Catalina and above, to run these commands below you *do not* need to specify the dbpath.**

5. Take a screen shot of the terminal after entering the following:

a. In the terminal/cmd window, type:

i. `mongod --version`

1. it should respond with the version

ii. `mongo --nodb`

1. It should respond with the version and also open the mongo shell, giving you a “>” cursor

iii. `quit()`

Course Environment:

- Assume that [Anaconda](#) is already installed
- Optionally, set up a new environment for the Mongo part of the course. (I am not for this course.)
 - [environment](#) (so as far as I'm concerned, this is optional)
- From inside Anaconda (or inside your specific environment within Anaconda): start jupyter notebook or, on windows:
 - Start-->Anaconda-->Jupyter
 - Runs on localhost:8888
- From inside Anaconda (or inside your specific environment within Anaconda):
 - `pip install pymongo dnspython`
 - (should work for the mac too)
 - FYI: [Support for mongodb+srv:// URIs requires dnspython:](#)
<http://api.mongodb.com/python/current/installation.html>
 - if it doesn't work for the mac, try looking at this, but I think things should run okay.
 - <https://support.opendns.com/hc/en-us/articles/227987667-Installing-and-Using-dnspython-clientsubnetoption>

- On your Desktop (or wherever you want to store your Jupyter notebooks, you may want to create a folder named intro-to-mongodb, only if you want your folder to match the way my videos look. It's not necessary to do this.

This is the end of orientation assignments for Mongo.

Below is additional to complete before lab 1

I have included the following in the same document as the Orientation; while not required for orientation, they should be completed before Lab 1.

Load data into MongoDB:

- the [Mongo folder](#) (also found on the syllabus, in the Course Schedule, just above the MongoDB section).
- Unzip that folder into your intro-to-mongodb folder that you created above, or wherever you want to store it. The movies_initial.csv file is in that folder. That's the file that you want to import to your mongodb Atlas.
- While you're at it, also [download the Jupyter notebooks](#) (also found in the Course Schedule of the syllabus, just above the MongoDB section). Unzip that into your intro-to-mongodb folder.
- import: using mongoimport command [Video: 2 MDB-Import-Movies Intial](#)
 - Go to [MongoAtlas, sign in.](#) Your mflix cluster should be displayed. If not, click on clusters.
 - Click on the "...", command-line tools, and copy the mongoimport command-line template statement. Substitute values for the green placeholders.

Data Import and Export Tools

Replace **PASSWORD** with the password for the admin user, **DATABASE** with the name of the database you wish to import/export to your cluster, and **COLLECTION** with the name of the collection you wish to import/export to your cluster. Replace **FILETYPE** with json or csv to specify the file type. Where applicable, replace **FILENAME** with the location and name of the output file (for export) or data source (for import).

NOTE: When exporting or importing CSV data, an additional --fields flag is often required. See documentation for the specific tool for additional details.

mongoimport | imports content from an Extended JSON, CSV, or TSV export

```
mongoimport --host atlas-zs3jq1-shard-0/mflix-shard-00-00.ep7ax.mongodb.net:27017,mflix-shard-00-01.ep7ax.mongodb.net:27017,mflix-shard-00-02.ep7ax.mongodb.net:27017 --ssl --username cfn --password <PASSWORD> --authenticationDatabase admin --db <DATABASE> --collection <COLLECTION> --type <FILETYPE> --file <FILENAME>
```

Copy

- You may want to copy this into Word/Sublime/Notepad or similar and substitute your values. You will be using it again, so you may want to save it.
- I used the name of the cluster (mflix) and the Database name (mflix).
- I added --headerline and --ssl. However, the latest version on Atlas includes --ssl
- **Note: if you see an error: "error connecting to host: could not connect to server: server selection error: server selection timeout", doublecheck on Atlas in the Network Security tab that the current IP address is on the IP whitelist, as shown on the video. If it is not, add it to the IP Whitelist and run the import statement again.**

- *run the mongoimport command from the regular terminal, NOT from the mongo shell!*
- *In your terminal, be sure to navigate to the folder that has your data file (in this case movies_initial.csv*

Using Compass:

- Like PHPMyAdmin, it's a locally-installed GUI tool that allows you to connect to a DB. Examine, update, delete and query.
- [Download Compass \(if it is not already installed from your MongoDB installation\)](#)
[Video: 3-Install-Compass](#)
- Install Compass
- Go to Atlas-->cluster-->Connect-->Connect using Compass. (You can use Connect through an application too.)
- Copy that string, as shown in the video
- Back in Compass, paste the string into the dialogue box.
 - **Note: if you see an error about a timeout, check on Atlas in the Network Security tab that the current IP address is on the IP whitelist. If it is not, add it to the IP Whitelist and run the import statement again.**
- Create database “mflix” and collection name “mflix” or whatever names you chose for your database and collection
- Click into mflix
- Click “Add Data” and choose the file movies_initial.csv
- Click import