Introduction to Anaconda

Emilie Skoog

Summer 2021

Itinerary

- What is conda?
- What is a conda environment?
- Benefits of conda environments
- Installing conda on your local computer (tutorial)
- Creating conda environments (tutorial)
- Installing programs within a conda environment (tutorial)

What is conda?

“Anaconda is a distribution of the Python and R programming languages for scientific computing, that aims to simplify package management and deployment.” “It is an open source **package** and **environment** management system.”

Types of conda:

- Miniconda:
  Lightweight distribution of conda; only contains the necessary python packages.

- Anaconda:
  A **data science platform** distribution of conda; comes with a lot of scientific python packages.

Info from: https://geohackweek.github.io/datasharing/01-conda-tutorial/

What is a conda environment?

- A tool that helps to keep dependencies required by different projects separate by creating isolated spaces for them that contain per-project dependencies for them.

- You can think of it as a virtual space that you have created to perform certain tasks.

Analogy #1: Tents

You can think of each conda environment as a tent. You can enter into that tent and bring whatever belongings (data) you need to work on into that tent with whatever tools are in that tent (installed programs), and when you are done, you can exit that tent and enter another bringing whatever needed belongings (data) with you for further work (analysis).

Each tent is a different conda environment.

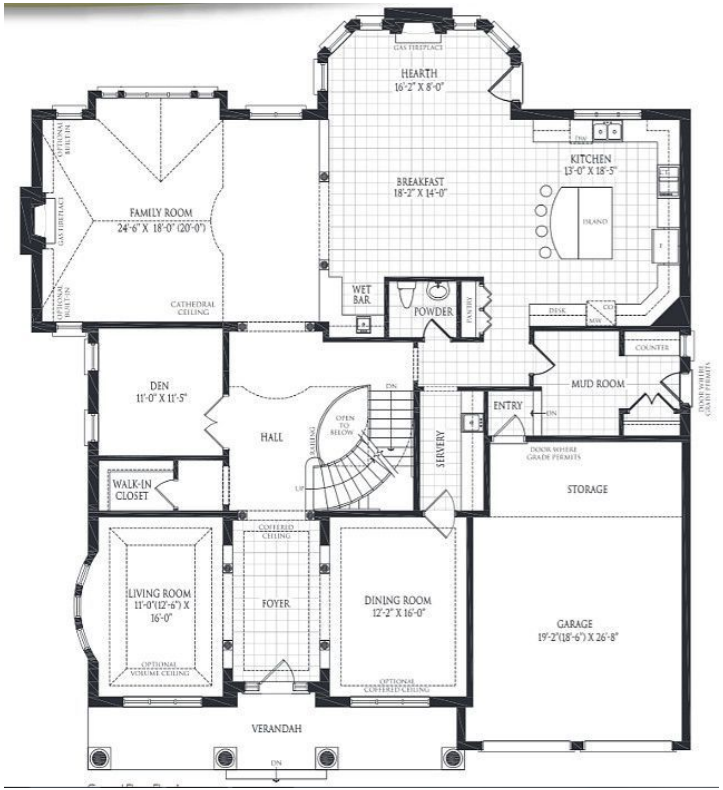
Analogy #2: House

You have many rooms in your house.

Each room has certain furniture/appliances/ tools that you use to do a specific task in that “environment.” For example, you cook in the kitchen, not in the bathroom (hopefully) and you use the bathroom in the bathroom rather than the bedroom (also hopefully).

Each environment has dedicated resources within it and you can enter or exit each of these environments to perform whatever tasks you want to do within that environment.

Having rooms or environments within your house creates organization and makes doing specific tasks a lot easier. (If you had all of your furniture piled on top of itself in the same room, then it would make it really hard to cook without catching your sofa on fire (programs can clash), for example).

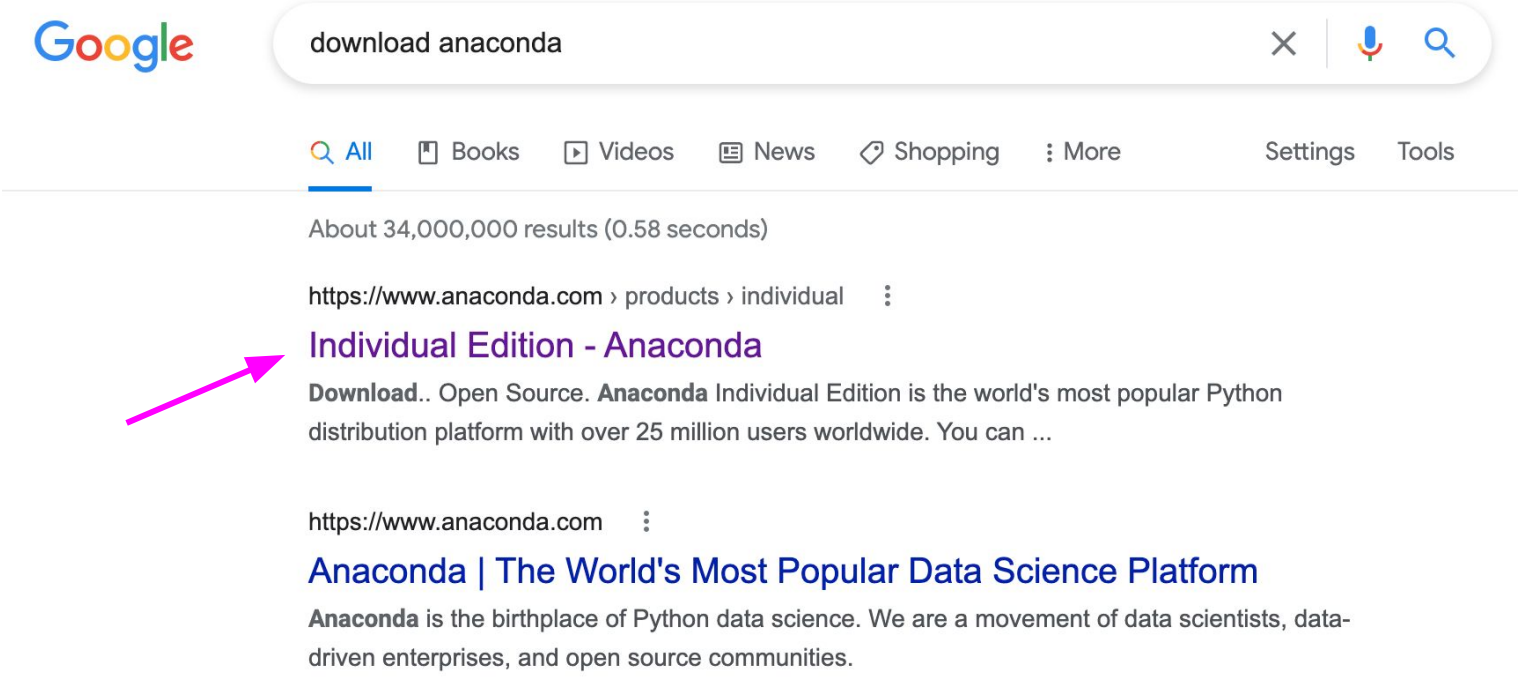


Benefits of creating and having conda environments

There are many more, but here are some reasons most relevant to us:

1. Keeps programs with different dependencies separate (no clashing dependencies)
2. Control over program versions
3. Organization
4. Reproducibility

Installing conda on your local computer

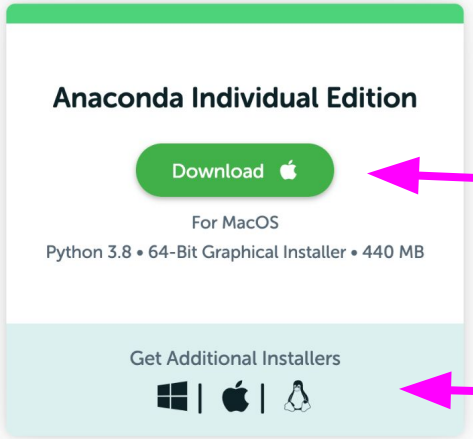


Installing conda on your local computer

Individual Edition

Your data science toolkit

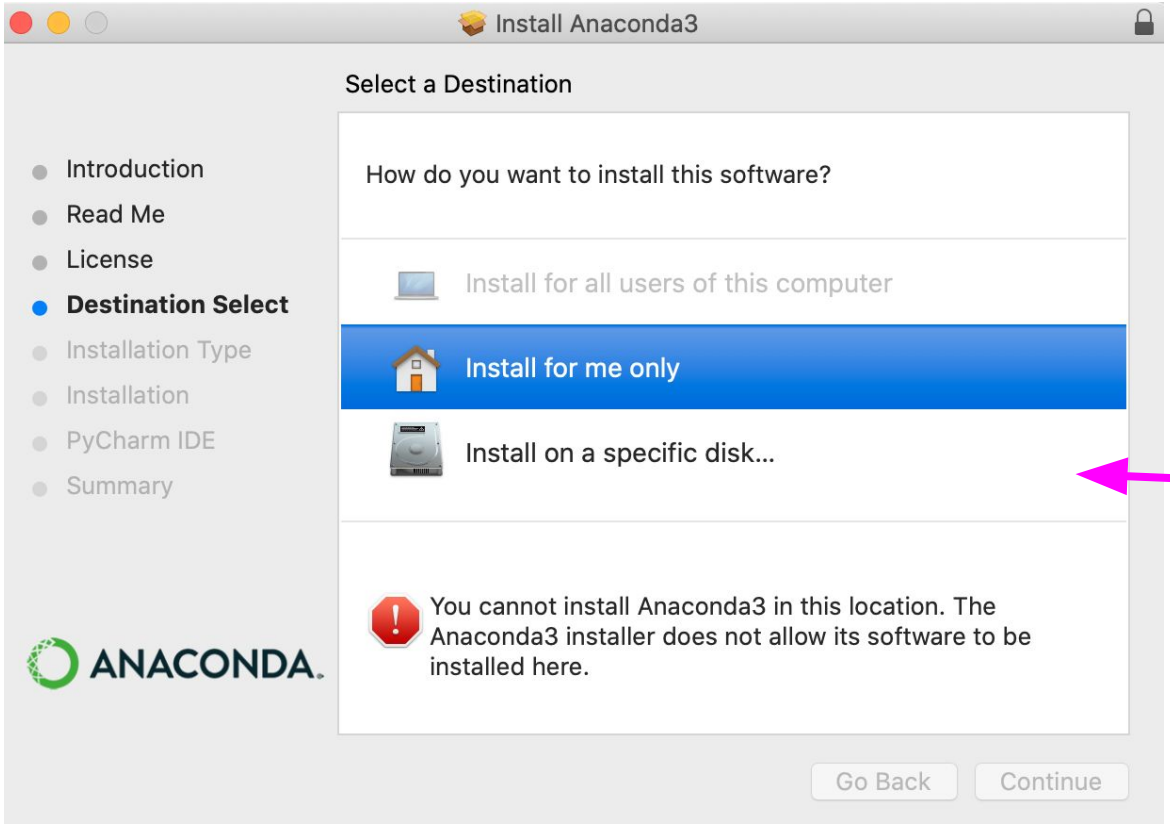
With over 25 million users worldwide, the open-source Individual Edition (Distribution) is the easiest way to perform Python/R data science and machine learning on a single machine. Developed for solo practitioners, it is the toolkit that equips you to work with thousands of open-source packages and libraries.



MacOS

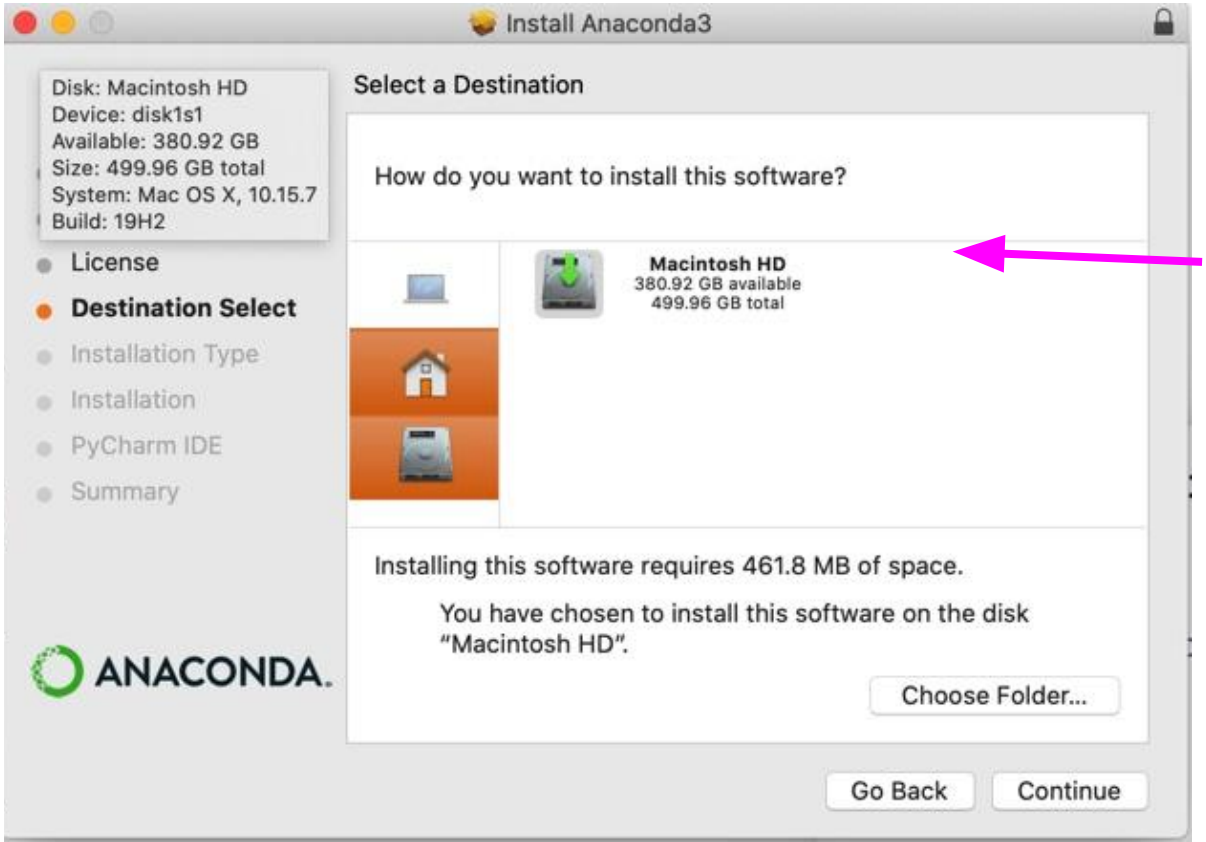
(For others)

Installing conda on your local computer



Will likely need to select this option

Installing conda on your local computer



If so, select this

Check installation

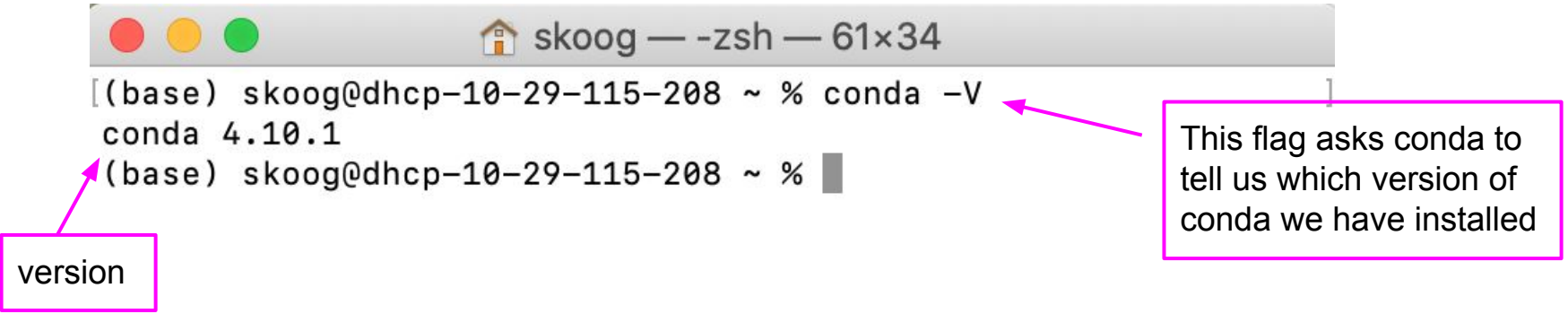
1. Open your terminal (IMPORTANT: if it was already open, close it and reopen it)

Check installation

1. Open your terminal (IMPORTANT: if it was already open, close it and reopen it)
2. Type conda -V (make sure the ‘V’ is capitalized)

Check installation

1. Open your terminal (IMPORTANT: if it was already open, close it and reopen it)
2. Type conda -V (make sure the ‘V’ is capitalized)



How do we create and use a conda environment?

Step 1: Create a conda environment

Step 2: Enter your conda environment

Step 3: Install programs inside of conda environment

Step 4: Use conda environment

Step 0: Checking out our conda environments

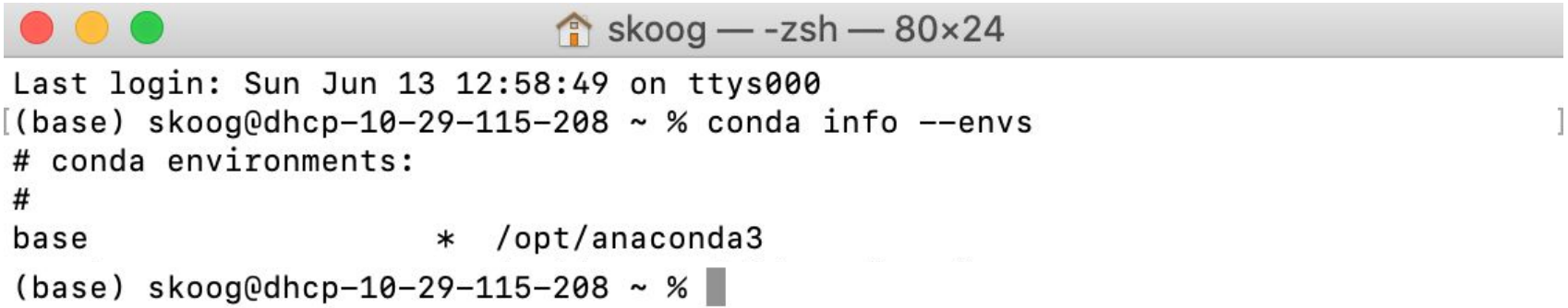
Conda command that enables us to see what environments we have:

```
conda info --envs
```

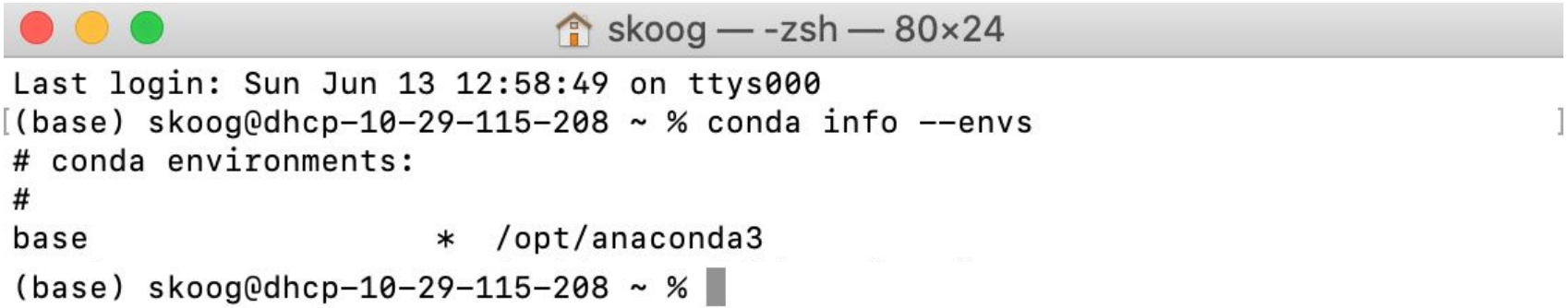
Can visit this site (also linked under resources on GitHub) to see conda cheat sheet for all commands for creating, checking, removing, etc. conda environments:

https://conda.io/projects/conda/en/latest/user-guide/tasks/manage-environments.html

Step 0: Checking out our conda environments



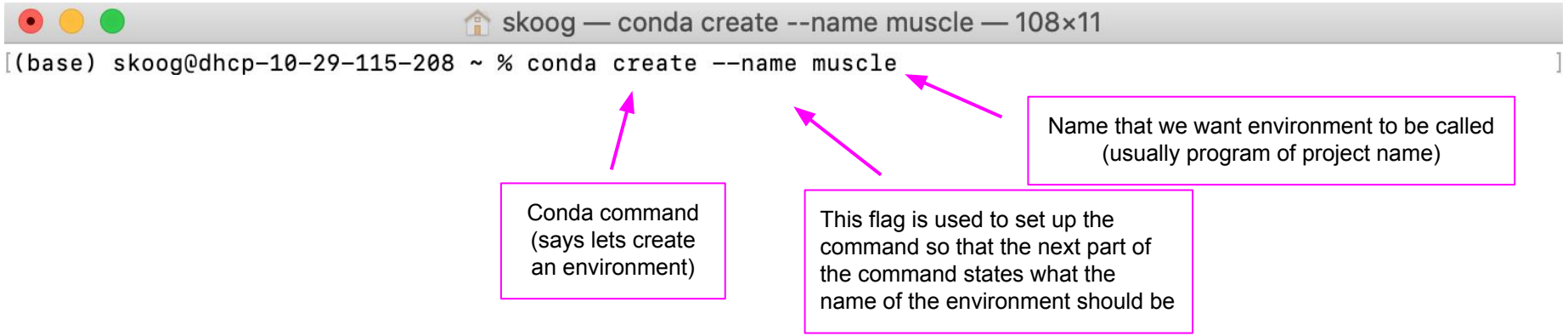
Step 0: Checking out our conda environments



‘base’ conda environment is our simplest conda ‘environment’ which basically just lets us execute commands to install programs and build other environments.

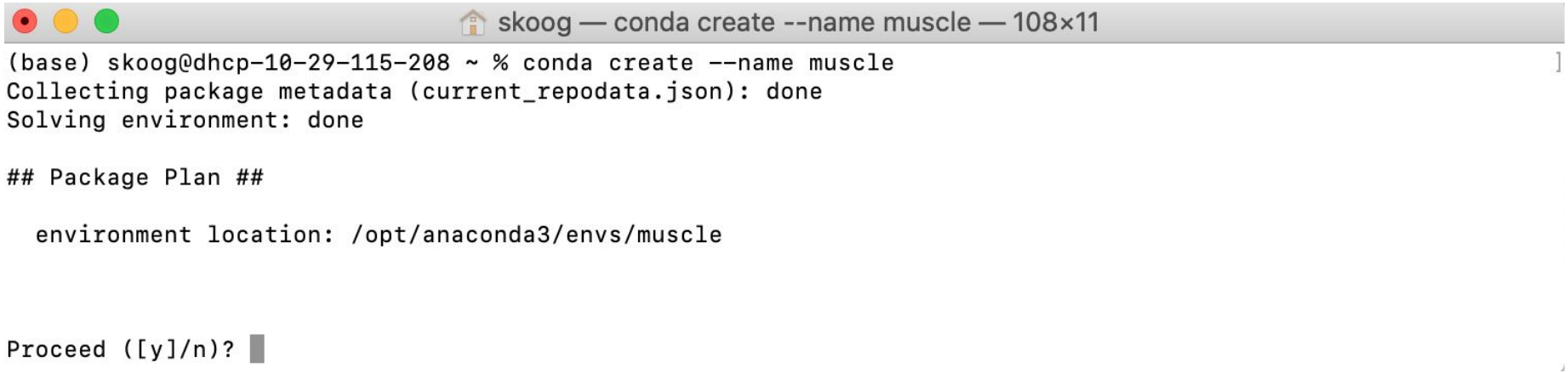
If you installed conda properly, you will have a ‘base’ conda environment.

Step 1: Create conda environment

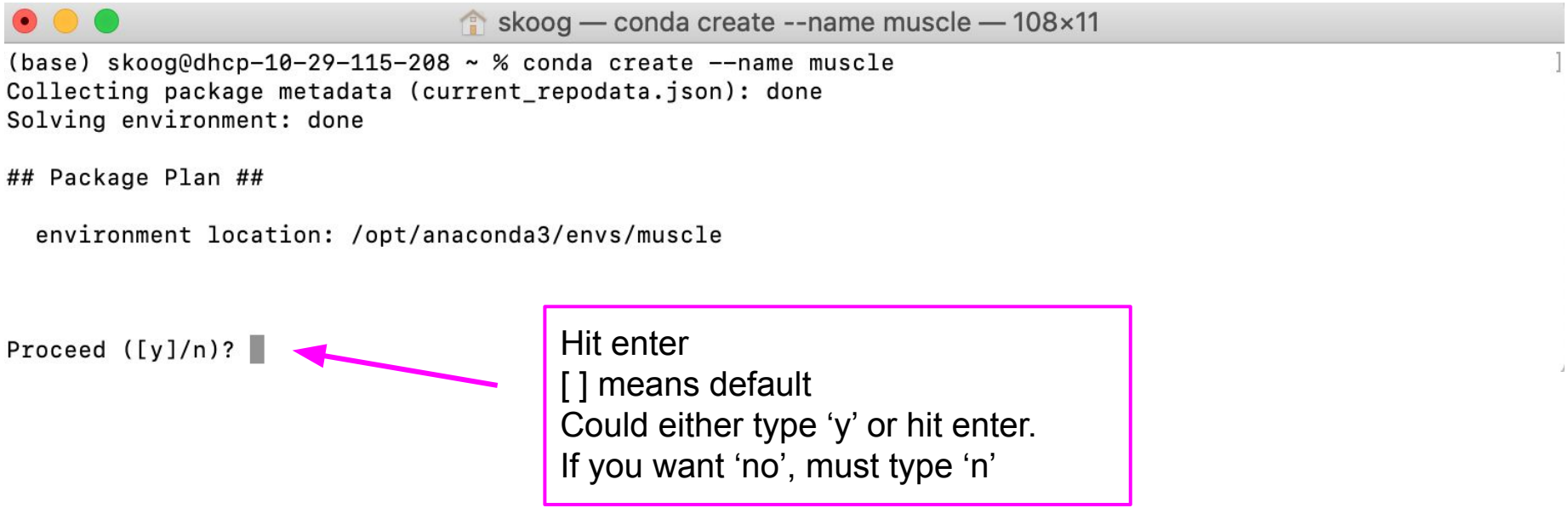


(Use the online conda cheat sheet to see what command to use to create a new environment)

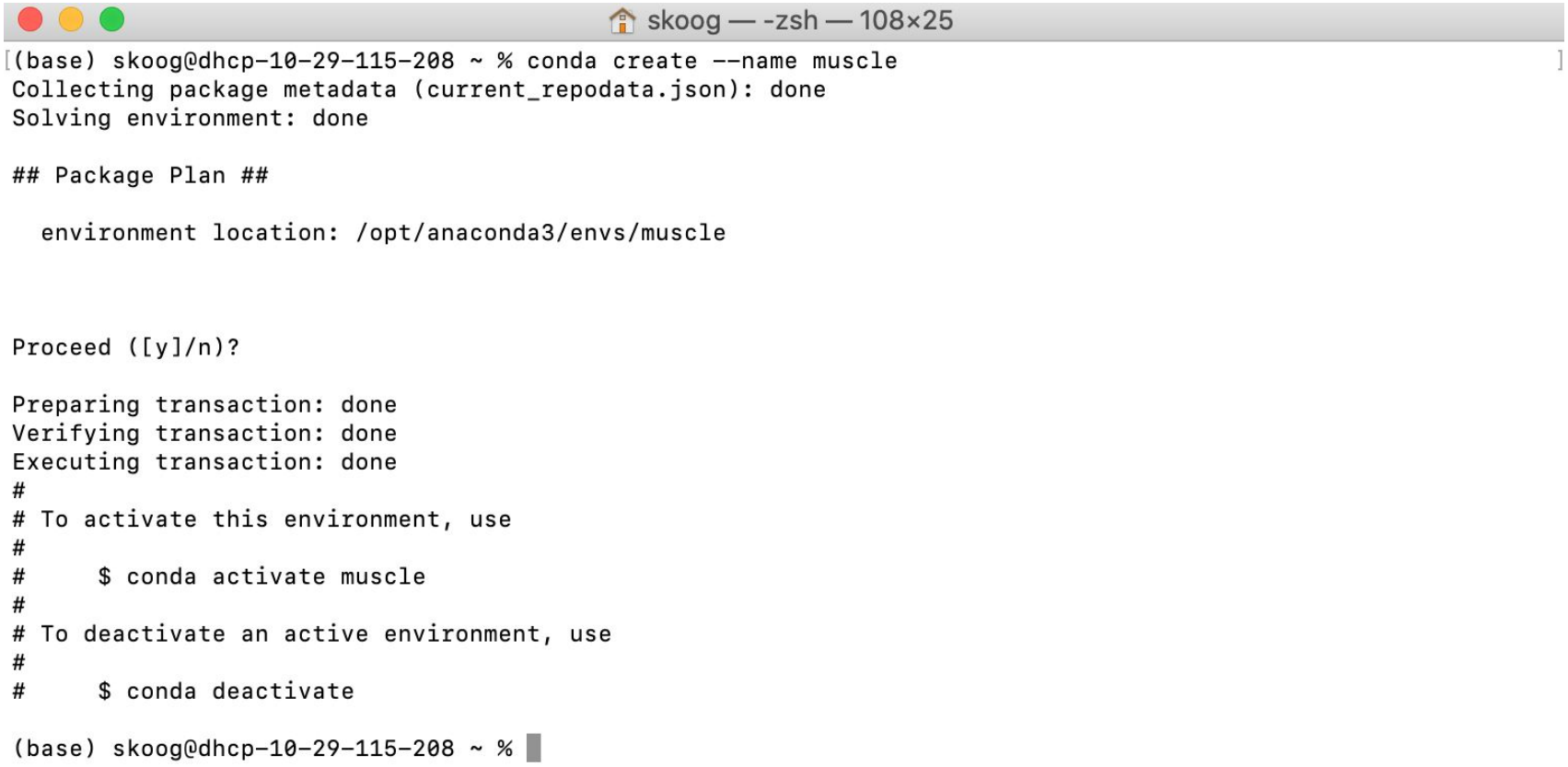
Step 1: Create conda environment



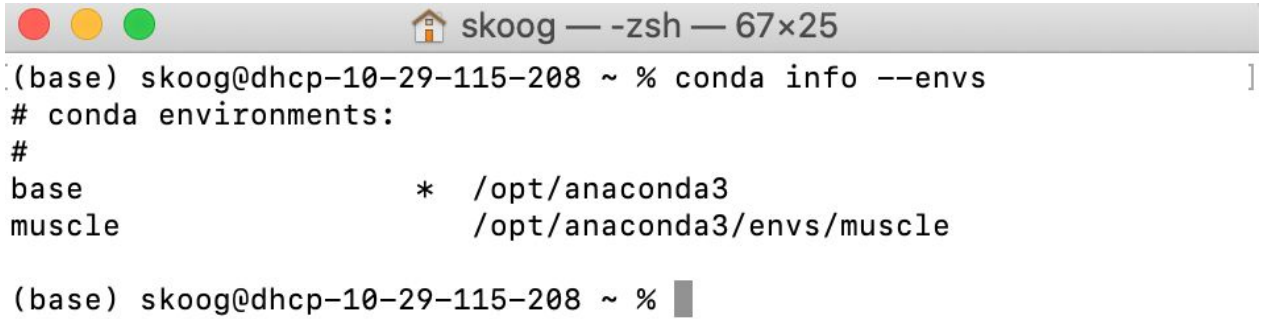
Step 1: Create conda environment



Step 1: Create conda environment



We can now check to see that we have a new environment



How do we create and use a conda environment?

Step 1: Create a conda environment 

Step 2: Enter your conda environment

Step 3: Install programs inside of conda environment

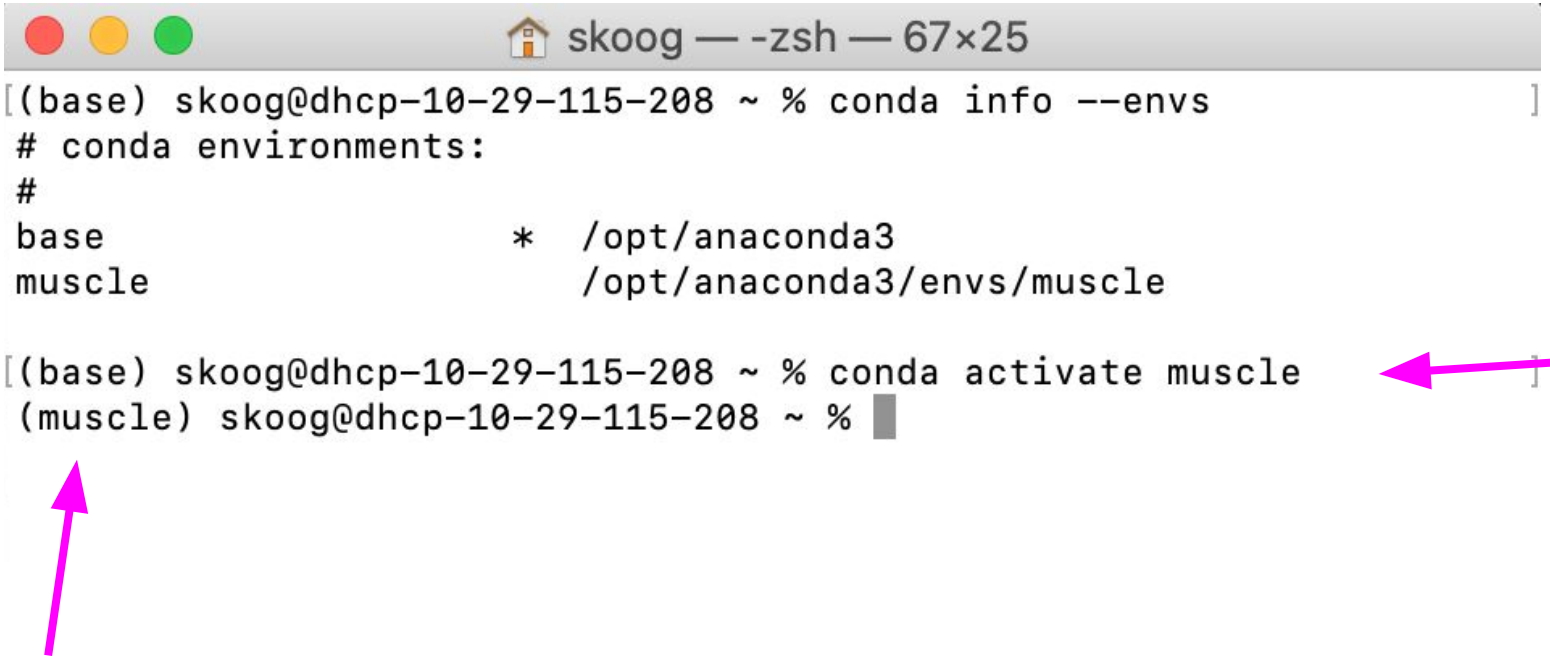
Step 4: Use conda environment

Step 2: Enter your environment

skoog — -zsh — 108×25

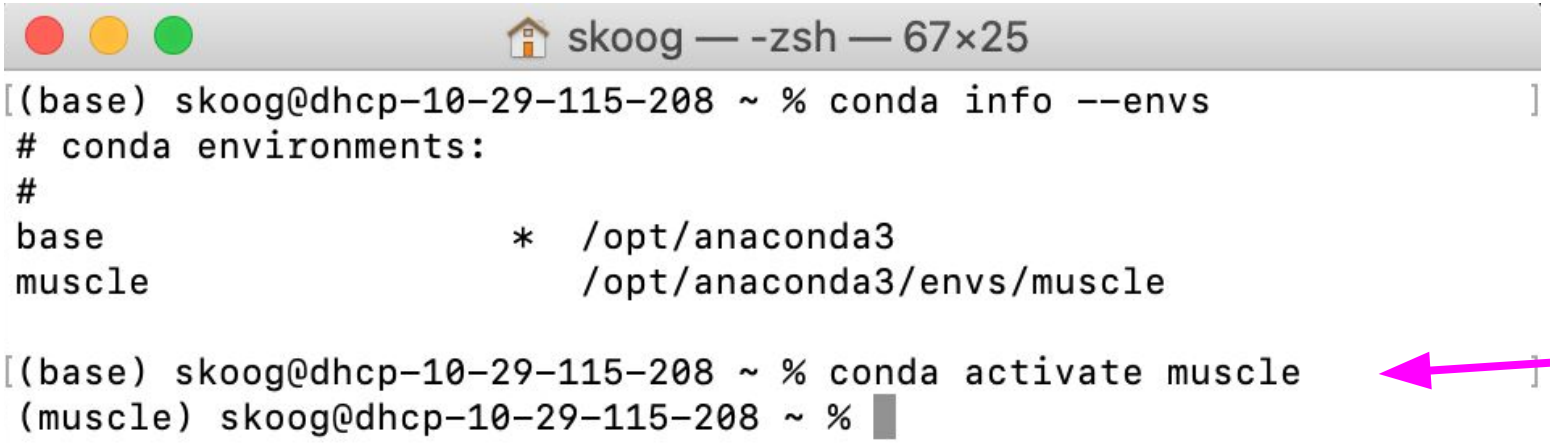
```
(base) skoog@dhcp-10-29-115-208 ~ % conda create --name muscle
Collecting package metadata (current_repodata.json): done
Solving environment: done

## Package Plan ##

  environment location: /opt/anaconda3/envs/muscle



Proceed ([y]/n)?

Preparing transaction: done
Verifying transaction: done
Executing transaction: done
#
# To activate this environment, use
#
#     $ conda activate muscle
#
# To deactivate an active environment, use
#
#     $ conda deactivate

(base) skoog@dhcp-10-29-115-208 ~ %
```

Step 2: Enter your environment



Notice change in conda environment

Step 2: Enter your environment



Notice change in conda environment

Note that so far, we have only created an environment with the name of the program that we want (could have also named it anything else). We still have to ADD (or install) that program inside the conda environment. (It's like having a kitchen without any appliances. We need to install the appliances before a room called the kitchen can actually function as our kitchen!)

How do we create and use a conda environment?

Step 1: Create a conda environment 

Step 2: Enter your conda environment 

Step 3: Install programs inside of conda environment

Step 4: Use conda environment

Step 3: Install program inside of environment

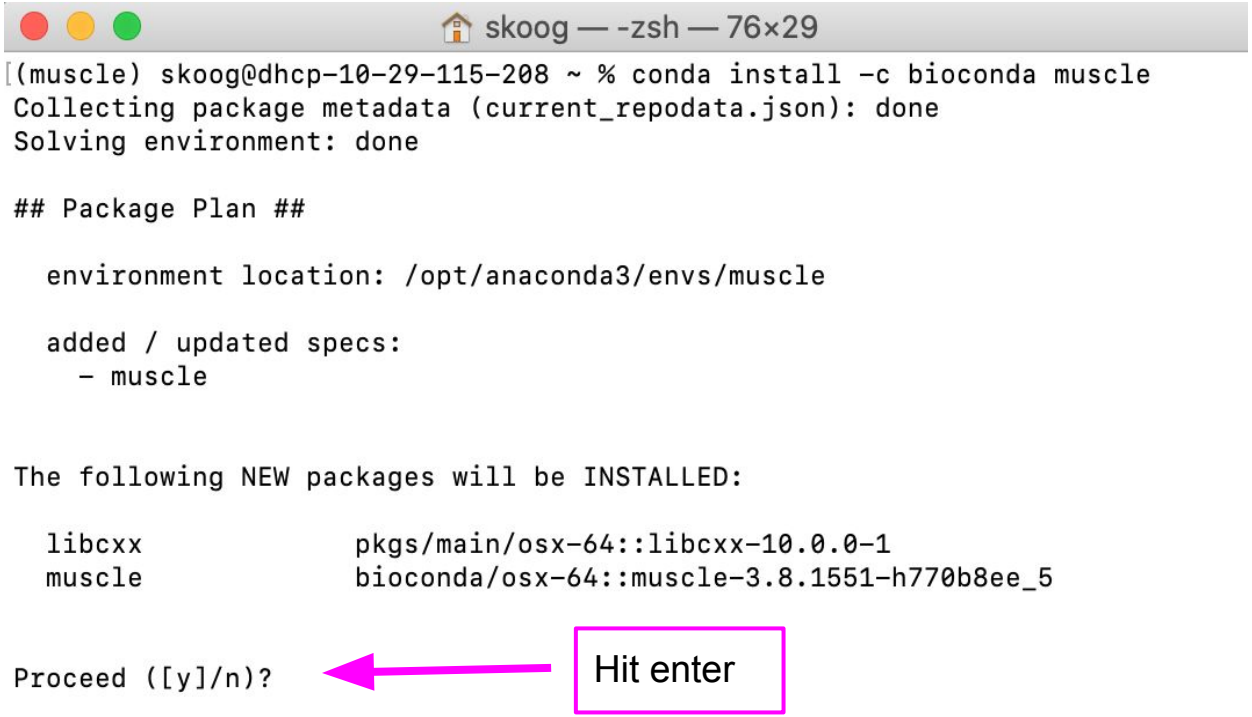
First, always check to make sure you are actually inside of your environment.

Second, install program.

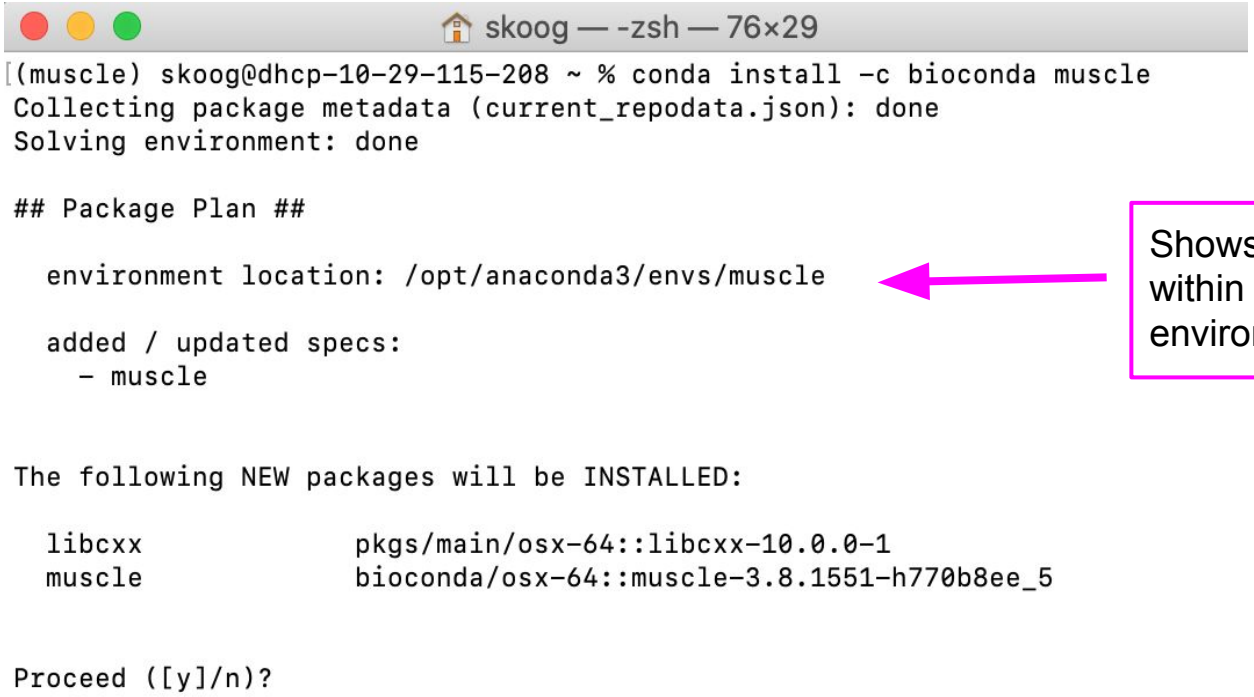
Note: To figure out how to install any program using anaconda, literally google “install [program] anaconda” (ex. install muscle anaconda)

Step 3: Install program inside of environment

Step 3: Install program inside of environment

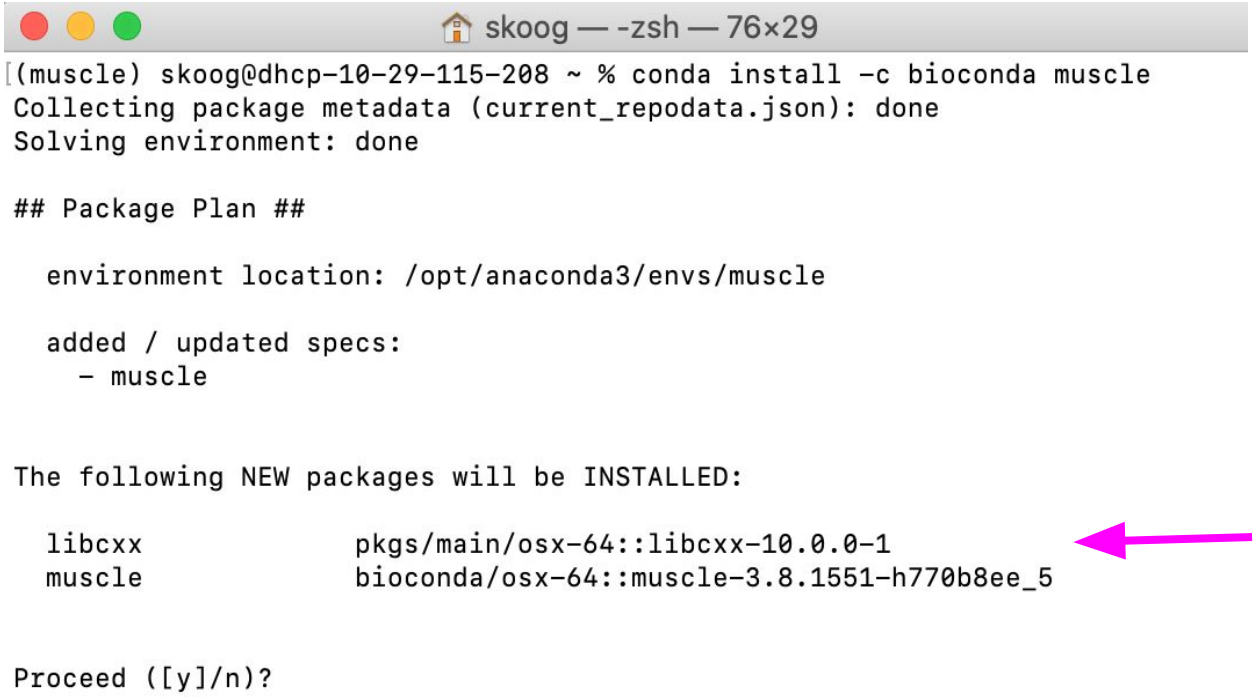


Step 3: Install program inside of environment



Shows us path (location) of program within conda. Here located in the conda environment (envs) directory

Step 3: Install program inside of environment



Shows us the packages that will be installed within the muscle program

Step 3: Install program inside of environment

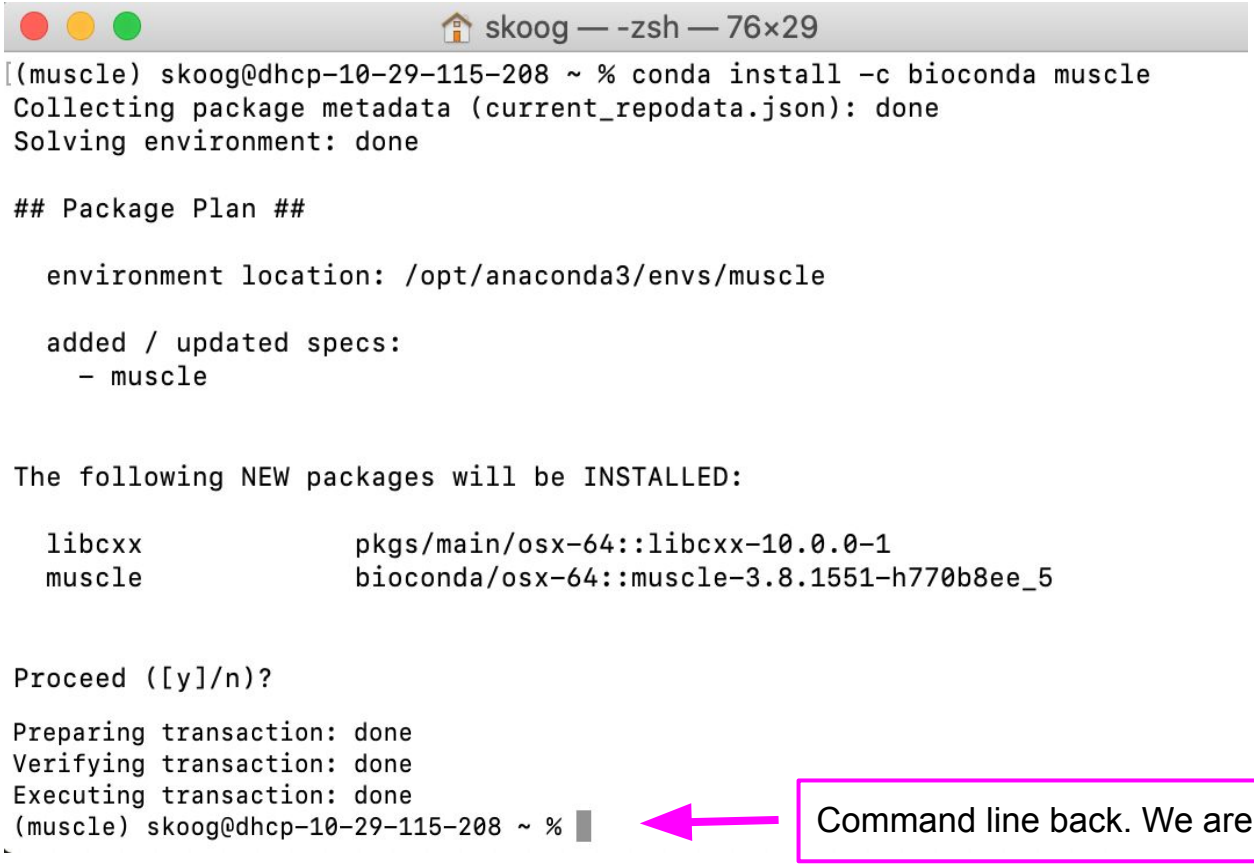
Step 3: Install program inside of environment

skoog — -zsh — 76×29

```
(muscle) skoog@dhcp-10-29-115-208 ~ % conda install -c bioconda muscle
Collecting package metadata (current_repodata.json): done
Solving environment: done

## Package Plan ##

  environment location: /opt/anaconda3/envs/muscle

  added / updated specs:
    - muscle


The following NEW packages will be INSTALLED:

  libcxx             pkgs/main/osx-64::libcxx-10.0.0-1
  muscle             bioconda/osx-64::muscle-3.8.1551-h770b8ee_5


Proceed ([y]/n)?

Preparing transaction: done
Verifying transaction: done
Executing transaction: done
(muscle) skoog@dhcp-10-29-115-208 ~ %
```

Successful

Step 3: Install program inside of environment



Step 3: Install program inside of environment

If, for example, you want to have a specific environment for one of your projects and you want several different programs inside this one environment (assuming they don't clash with one another and cause problems (clashing dependency versions, for example)), **you can install many programs within one environment**.

Maybe you want a single environment called "tree_building" where you have all the programs you need to align sequences and then use this output to make a tree, then you can install muscle (alignment program) and iqtree (a tree building program) within the same environment (... or you could also keep them separate).

How do we create and use a conda environment?

Step 1: Create a conda environment 

Step 2: Enter your conda environment 

Step 3: Install programs inside of conda environment 

Step 4: Use conda environment

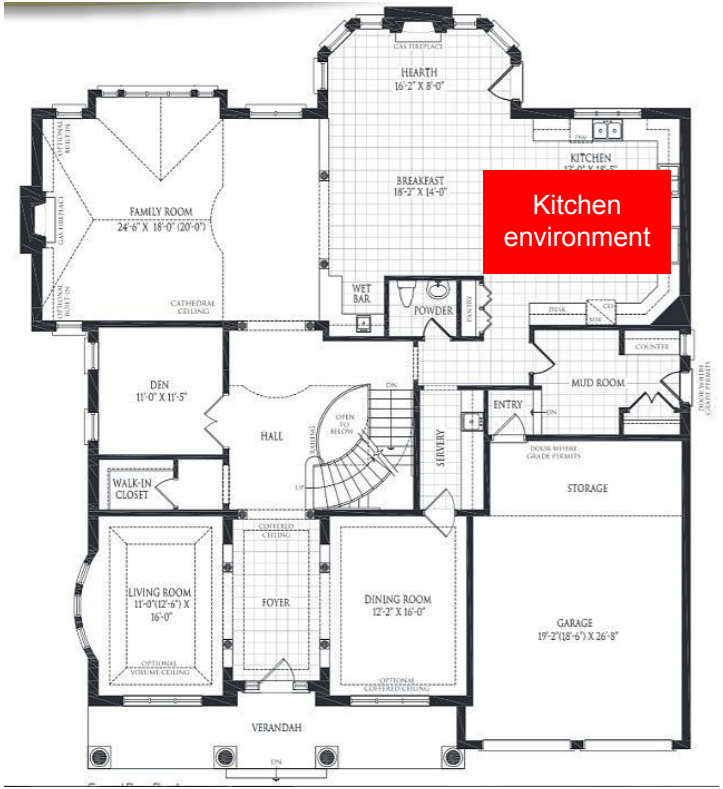
Step 4: Use our conda environment!

Stay tuned for phylogenetics tutorial (and others)!

Overview: Kitchen analogy

Overview: Kitchen analogy

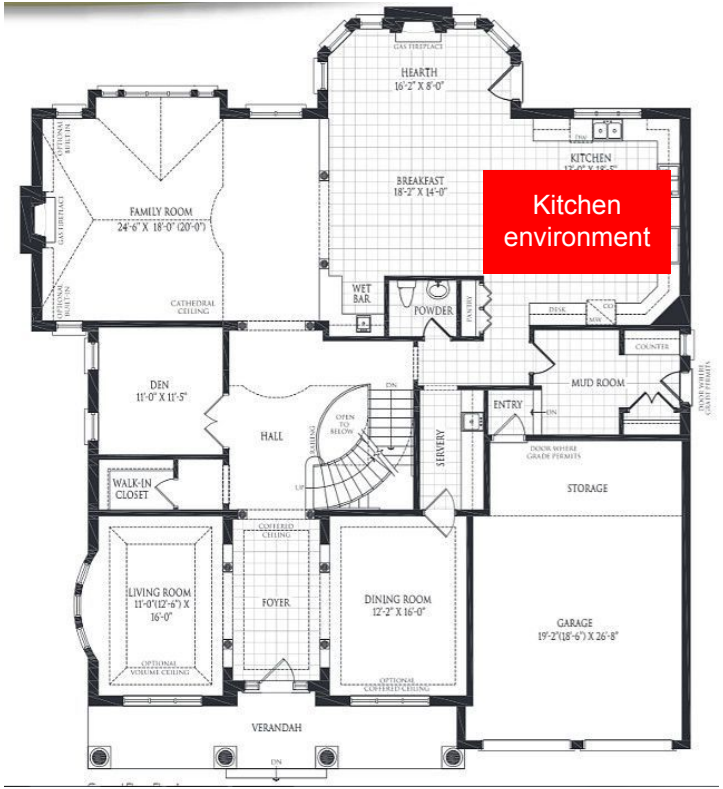
Step 1: Build/create conda environment (build kitchen)



Kitchen environment

Overview: Kitchen analogy

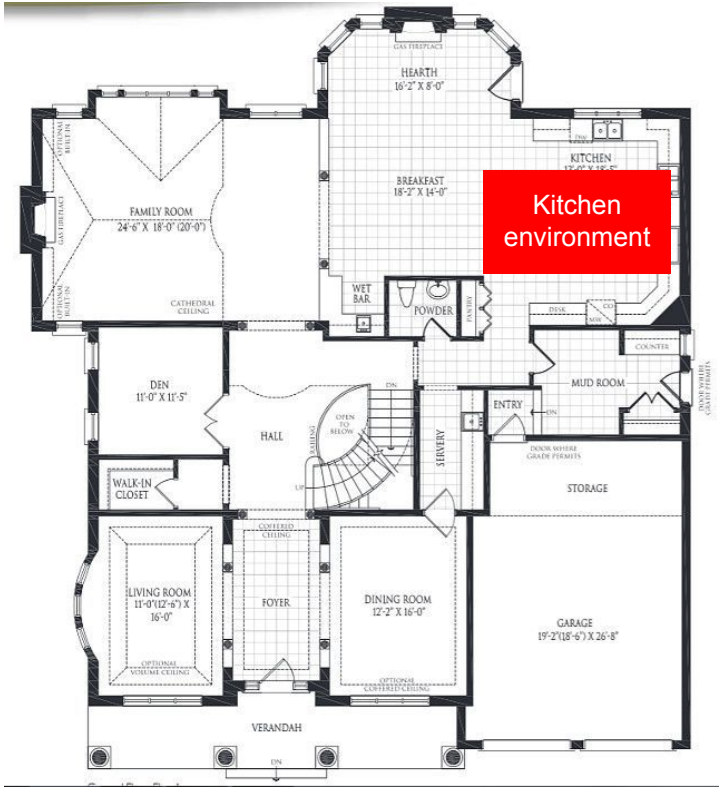
Step 1: Build/create conda environment (build kitchen)



Step 2: Enter conda environment (enter kitchen)

Overview: Kitchen analogy

Step 1: Build/create conda environment (build kitchen)



Kitchen environment

Step 2: Enter conda environment (enter kitchen)

Step 3: Install program for use (install kitchen appliances for kitchen use)

Overview: Kitchen analogy

Step 1: Build/create conda environment (build kitchen)

Kitchen environment

Now you can use the kitchen as intended!

Step 2: Enter conda environment (enter kitchen)

nstall kitchen appliances for kitchen use)