

Introduction to Data Science

Data Science Essentials



Goals for today

- **conda environments**
- **Create a choropleth**



Review last session coding tasks

week3_review notebook



conda

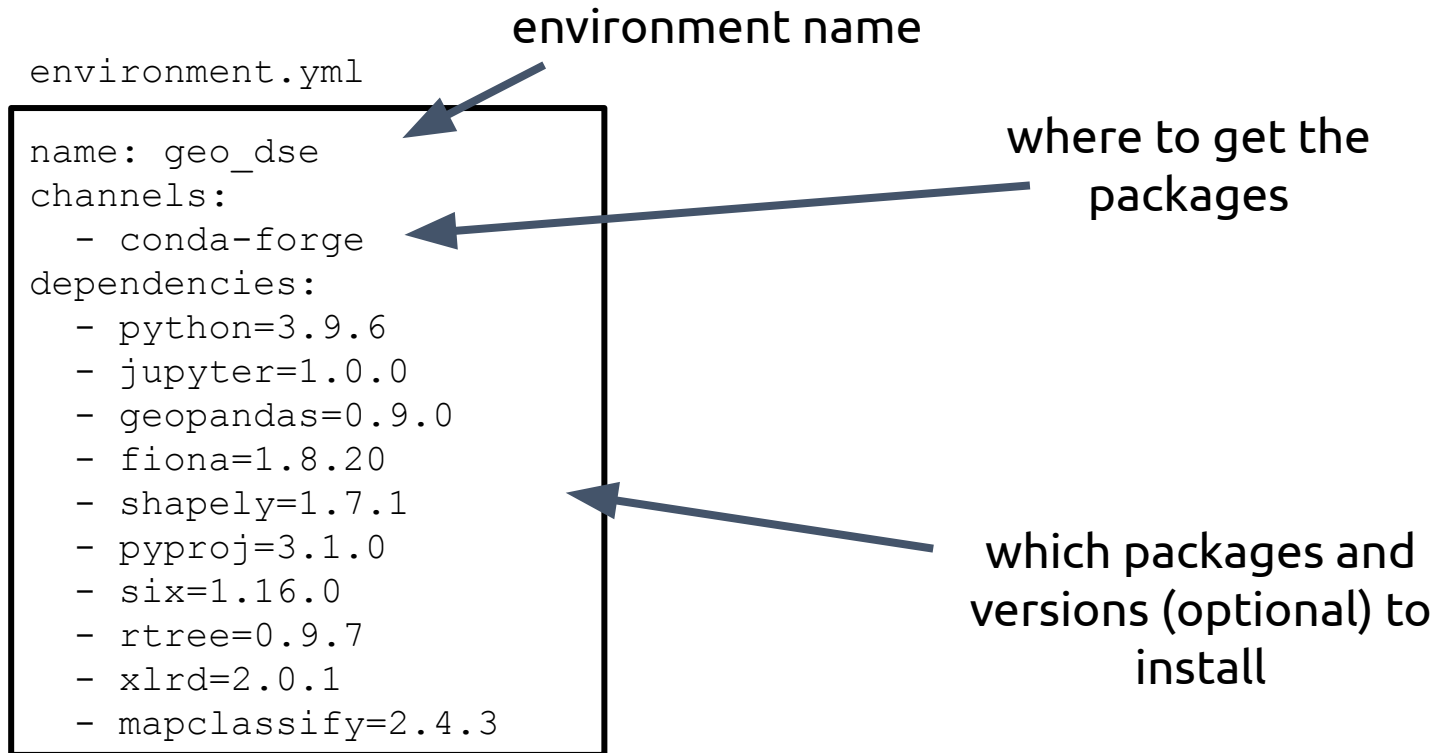
conda is not just a package manager, but is also an *environment* manager, meaning that you can create separate environments containing files, packages, their dependencies, and their own versions of the Python interpreter.

This serves two purposes:

1. Isolates your projects
2. Makes it easier to share your work and allows for reproducibility.

conda Environments

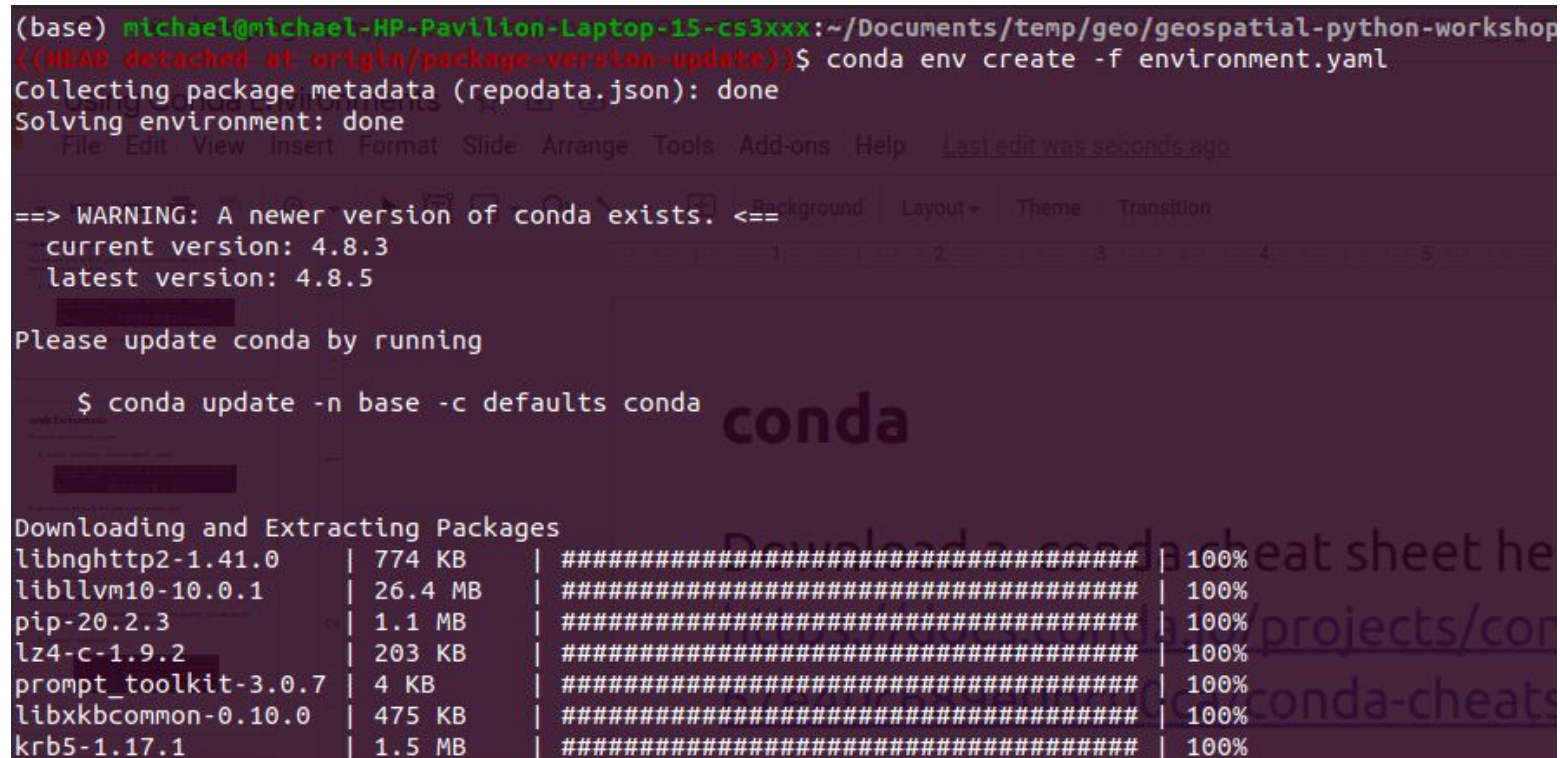
We're going to create an environment from a YAML file which lists which packages and versions should be installed.



conda Environments

To create an environment from a YAML file, run this from the folder containing the environment file:

```
$ conda env create -f environment.yaml
```

A terminal window with a dark background and light-colored text. The prompt is '(base) michael@michael-HP-Pavilion-Laptop-15-cs3xxx:~/Documents/temp/geo/geospatial-python-workshop'. The command executed is '\$ conda env create -f environment.yaml'. The output shows 'Collecting package metadata (repodata.json): done' and 'Solving environment: done'. A warning message states: 'WARNING: A newer version of conda exists. current version: 4.8.3 latest version: 4.8.5'. It then says 'Please update conda by running' followed by '\$ conda update -n base -c defaults conda'. Below this, a section titled 'Downloading and Extracting Packages' shows a list of packages being installed with their sizes and progress bars. The packages listed are libnghttp2-1.41.0, libllvm10-10.0.1, pip-20.2.3, lz4-c-1.9.2, prompt_toolkit-3.0.7, libxkbcommon-0.10.0, and krb5-1.17.1, all showing 100% completion.

```
(base) michael@michael-HP-Pavilion-Laptop-15-cs3xxx:~/Documents/temp/geo/geospatial-python-workshop
((HEAD detached at origin/package-version-update))$ conda env create -f environment.yaml
Collecting package metadata (repodata.json): done
Solving environment: done

==> WARNING: A newer version of conda exists. <==
  current version: 4.8.3
  latest version: 4.8.5

Please update conda by running

  $ conda update -n base -c defaults conda

Downloading and Extracting Packages
libnghttp2-1.41.0 | 774 KB | ##### | 100%
libllvm10-10.0.1 | 26.4 MB | ##### | 100%
pip-20.2.3 | 1.1 MB | ##### | 100%
lz4-c-1.9.2 | 203 KB | ##### | 100%
prompt_toolkit-3.0.7 | 4 KB | ##### | 100%
libxkbcommon-0.10.0 | 475 KB | ##### | 100%
krb5-1.17.1 | 1.5 MB | ##### | 100%
```

conda Environments

You can see a list of all of your conda environments along with the one currently active by typing

```
$ conda env list
```

```
(base) michael@michael-HP-Pavilion-Laptop-15-cs3xxx ~ $ conda env list
# conda environments:
#
base                *  /home/michael/anaconda3
geospatial          /home/michael/anaconda3/envs/geospatial
```

This will display a * next to the active environment.

conda Environments

To switch environments, type

```
$ conda activate <environment name>
```

```
(base) michael@michael-HP-Pavilion-Laptop-15-cs3xxx ~ $ conda activate geospatial
(geospatial) michael@michael-HP-Pavilion-Laptop-15-cs3xxx ~ $ conda env list
# conda environments:
#
base                /home/michael/anaconda3
geospatial          * /home/michael/anaconda3/envs/geospatial
```

If you need to return to the base environment, type

```
$ conda deactivate
```


conda Environments

Then (as long as your current environment includes jupyter), you can launch jupyter in the current environment by typing

```
$ jupyter notebook
```

```
(geospatial) michael@michael-HP-Pavilion-Laptop-15-cs3xxx ~ $ jupyter notebook
[I 10:24:09.366 NotebookApp] The port 8888 is already in use, trying another port.
[I 10:24:09.371 NotebookApp] Serving notebooks from local directory: /home/michael
[I 10:24:09.371 NotebookApp] The Jupyter Notebook is running at:
[I 10:24:09.371 NotebookApp] http://localhost:8889/?token=dc72fd8e78a41802290ae2c77e5c8b143c37adab3c7ef971
[I 10:24:09.371 NotebookApp] or http://127.0.0.1:8889/?token=dc72fd8e78a41802290ae2c77e5c8b143c37adab3c7ef971
[I 10:24:09.371 NotebookApp] Use Control-C to stop this server and shut down all
```

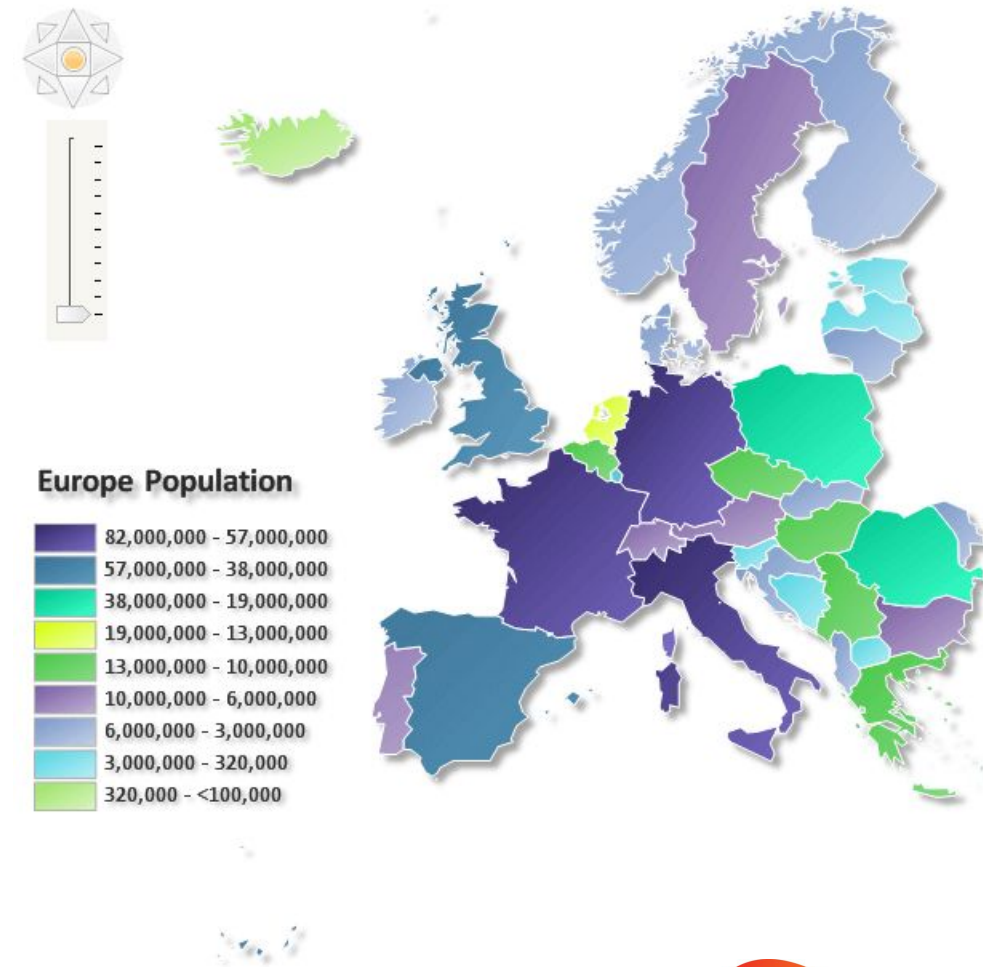
Choropleths

A choropleth is a map where areas are colored or shaded according to the value of some aggregate statistic for that area (eg. average income, population density, unemployment rate, etc.)

We will create choropleths in Python by using the *geopandas* library, which you will most likely need to install.

To install geopandas, open the Terminal (Mac) or Anaconda Prompt (Windows) and type

\$ `conda install geopandas -c conda-forge`



Building a choropleth

Choropleth_Tutorial notebook



Questions?

