Hardware, Software, and Data Management

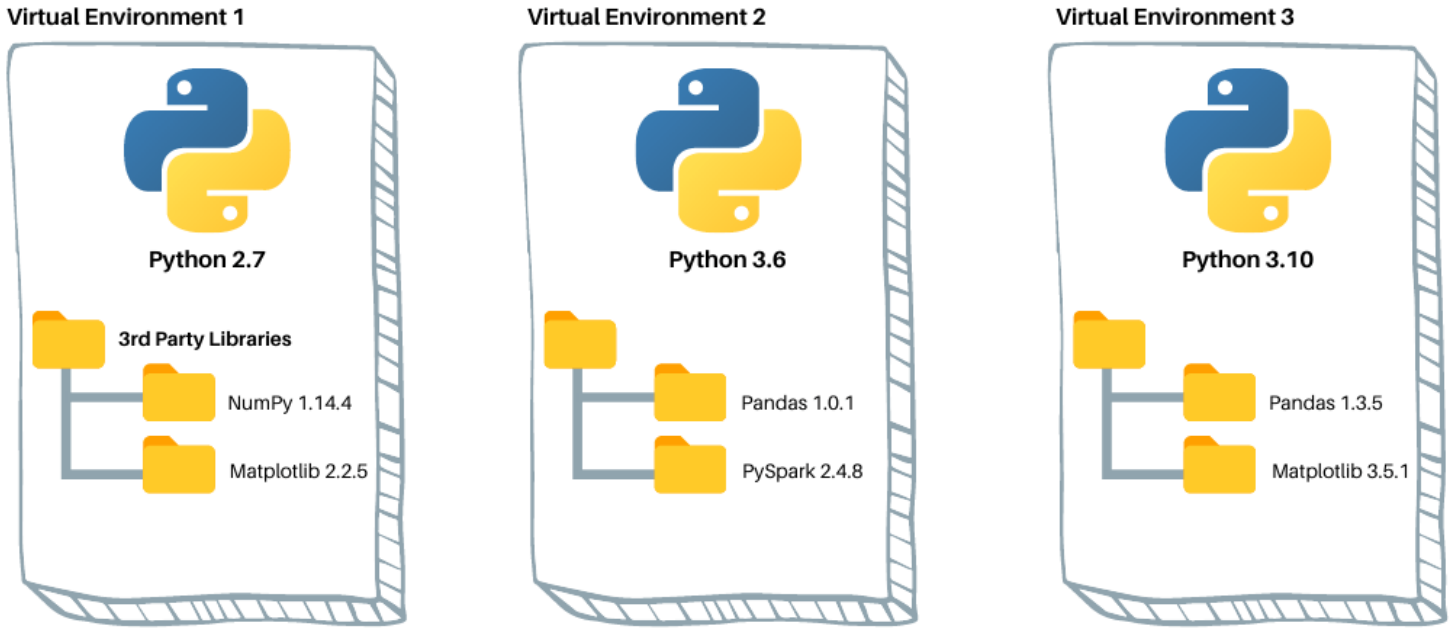
Tasks for This Week

- Understanding computing hardware
- Organizing and managing projects
- Managing environments
- Working with Jupyter lab/notebooks
- Working in Google Colab
- Introduction to SQL

File Management

Virtual Environments

Virtual Environments



dataquest.io

Virtual Environments

1. Clean

Virtual Environments

1. Clean
2. Dependency conflicts

Virtual Environments

1. Clean
2. Dependency conflicts
3. Reproducibility

Let’s set one up!

Getting started

As a rule of thumb, avoid installing to your base environment!

Windows

- Open Anaconda Prompt
- Dir: list files
- cd: change directory

macOS

- Open the terminal
- ls: list files
- cd: change directory

Linux

- Open the terminal
- ls: list files
- cd: change directory

Add Conda Forge

conda config -add channels conda-forge

conda config -set channel_priority strict

Creating an Environment

conda create -n ENVNAME python=x.x

conda activate ENVNAME

conda deactivate ENVNAME

conda install PACKAGE

conda list

conda env export > ENVIRONMENT.yml

Conda env create -n ENVNAME -file ENVIRONMENT.yml

Create a new environment with the following: Python 3.8, jupyterlab, pandas, matplotlib, and seaborn

Vocab

- Virtual/Conda/Computing environment
- Dependency
- RAM
- CPU
- GPU
- Hard drive/storage