12 Containerization & Packages

Key terms: venv pip3

Exercise: Use the PyPI faceted search to identify a package of interest to your goals.

Reading: Pradyun Gedam's "Thoughts on the Python packaging ecosystem"²⁸

Reading: Python Packaging User Guide: Installing Packages²⁹

Users can share short, discrete programs between themselves as single .py files, but more complicated projects demand better tooling for distribution. For better or worse, in the field of software packaging, the Python ecosystem has many tools with overlapping goals. A single Python Packaging Authority co-manages these various tools and maintains a glossary.³⁰

Users on hills can install packages inside a virtual environment, venv, which is an lightweight container with its own configuration and installed programs. The principal reasons to use one are installing diverse modules that are not needed or desired site-wide, and creating a standard, reproducible context. A sample creation, activation, and deactivation of a virtual environment follows.

[yourname@hills]\$ python3 -m venv 456 [yourname@hills]\$ source 456/bin/activate (456) [yourname@hills]\$ # Any work to be done in the environment. (456) [yourname@hills]\$ deactivate [yourname@hills]\$

Often we must install Python software that was not packaged with the language. Packages can be distributed in a "built" form that is ready to install or a "source" form that requires building later. The Python Package Index is the principal of free third-party modules.³¹ Each entry is annotated with a legal license (e.g. GPL, MIT), a Development Status (e.g., Alpha, Production, Inactive), and metadata about its dependencies. There are many experimental or one-off packages in the database that you will have to ignore. The resources cover every subfield.

pip is Python's package manager, an application bundled with CPython. It performs dependency resolution, downloads, setup, and reporting tasks. You will be installing packages with the pip install command. Afterwards, it can tell us about them:

```
(456) [yourname@hills ~]$ pip show nltk
Name: nltk
Version: 3.9.1
Summary: Natural Language Toolkit
Home-page: https://www.nltk.org/
Author: NLTK Team
Author-email: nltk.team@gmail.com
License: Apache License, Version 2.0
Location: /tmp/456/lib/python3.10/site-packages
Requires: click, joblib, regex, tqdm
Required-by:
```

You may see a warning about pip's version, which in the case of hills, we are obliged to ignore.

Next week we'll review some of the features of the relational database SQLite.

³¹https://pypi.org/

²⁸https://pradyunsg.me/blog/2023/01/21/thoughts-on-python-packaging/

²⁹https://packaging.python.org/en/latest/tutorials/installing-packages/

³⁰https://packaging.python.org/en/latest/glossary/