

Introduction to Package Management

What is package management?

In software development, a set of organization, indexing and distribution standards of installable units of software

What is package management?

Many package management systems throughout open-source software projects, notably in Linux

Low-level OS package management tools: **rpm, dpkg**

High-level OS tools: **apt, aptitude, yum, zypper**

MacOS: **homebrew**

Ruby: **gem** Python: **pip**

Building packages

PyPI is the recommended source for python packages.

Using PyPI offers:

Easy installation and management of packages using the **pip** tool

Automatic installation and tracking of dependencies

No-fuss package removal

Building packages

Near the end of 2021, PyPI managed over 500,000 users offering over 330,000 projects

**Find, install and publish Python packages
with the Python Package Index**

Search projects



Or [browse projects](#)

338,784 projects

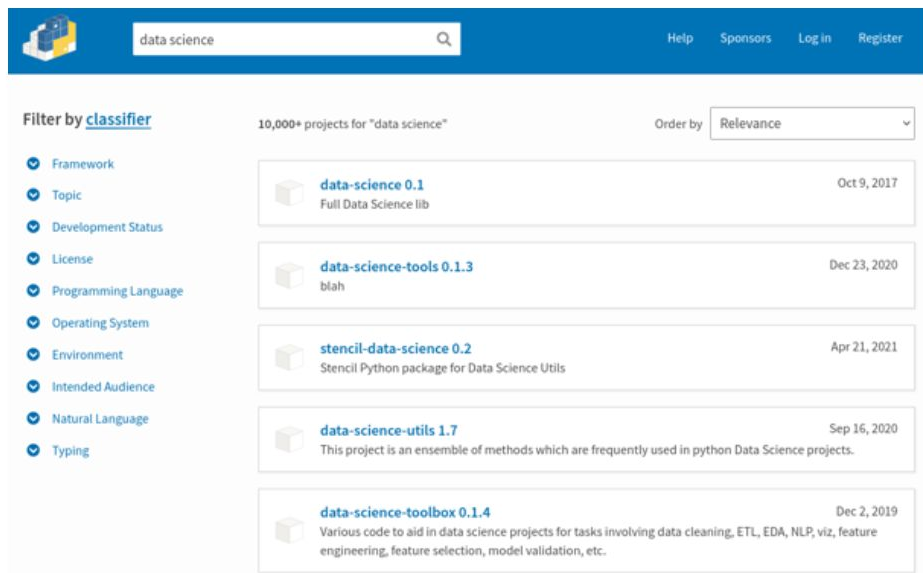
3,014,017 releases

5,143,412 files

550,052 users

Building packages

For example, searching by topic for “data science” generates a list of projects that offer tools that may apply to your project



The screenshot shows a search results page for "data science" projects. The page has a blue header with a search bar containing "data science" and navigation links for Help, Sponsors, Log in, and Register. Below the header, there's a filter section on the left with various categories like Framework, Topic, Development Status, License, Programming Language, Operating System, Environment, Intended Audience, Natural Language, and Typing, all of which are checked. The main content area shows 10,000+ projects for "data science", ordered by Relevance. Five project cards are visible, each with a cube icon, a title, a description, and a date.

| Project Name | Description | Date |
|-----------------------------------|--|--------------|
| data-science 0.1 | Full Data Science lib | Oct 9, 2017 |
| data-science-tools 0.1.3 | blah | Dec 23, 2020 |
| stencil-data-science 0.2 | Stencil Python package for Data Science Utils | Apr 21, 2021 |
| data-science-utils 1.7 | This project is an ensemble of methods which are frequently used in python Data Science projects. | Sep 16, 2020 |
| data-science-toolbox 0.1.4 | Various code to aid in data science projects for tasks involving data cleaning, ETL, EDA, NLP, viz, feature engineering, feature selection, model validation, etc. | Dec 2, 2019 |

Building packages

Organize your project in preparation for packaging and distribution to PyPI

Packages can be built manually and uploaded to PyPI

Or use tools such as Github Workflows to package and publish project files automatically when changes are committed