Here’s a step-by-step guide to turn your data/file handling code into a reusable Python package with utility modules:

**1. Create a Package Directory**

Create a directory for your utilities, e.g. dataplay/utils/.

**2. Add \_\_init\_\_.py**

Add an empty \_\_init\_\_.py file in dataplay/ and dataplay/utils/ to make them Python packages.

**3. Move Utility Code to Module(s)**

Create a file like dataplay/utils/parquet\_utils.py and move your reusable functions there. For example:

* read\_parquet\_folder
* read\_parquet\_zip
* merge\_parquet\_parts
* etc.

**4. Refactor Code into Functions**

Wrap your logic into functions with clear arguments and docstrings. Example: