

All of you must be familiar with what PDFs are. In fact, they are one of the most important and widely used digital media. PDF stands for Portable Document Format. It uses .pdf extension. It is used to present and exchange documents reliably, independent of software, hardware, or operating system. Invented by Adobe, PDF is now an open standard maintained by the International Organization for Standardization (ISO). PDFs can contain links and buttons, form fields, audio, video, and business logic. In this article, we will learn, how we can do various operations like:

Extracting text from PDF

Rotating PDF pages

Merging PDFs

Splitting PDF

Adding watermark to PDF pages

Installation: Using simple python scripts!

We will be using a third-party module, PyPDF2.

[PyPDF2](#) is a python library built as a PDF toolkit. It is capable of: