

1- PdfFileReader() needs to be opened in read-binary mode by passing 'rb' as the second argument to open(). Likewise, the File object passed to PyPDF2. PdfFileWriter() needs to be opened in write-binary mode with 'wb'.

2- First, import the PyPDF2 module. Then open meetingminutes.pdf in read binary mode and store it in pdfFileObj. To get a PdfFileReader object that represents this PDF, call PyPDF2.PdfFileReader() and pass it pdfFileObj. Store this PdfFileReader object in pdfReader.

The total number of pages in the document is stored in the numPages attribute of a PdfFileReader object ❶. The example PDF has 19 pages, but let's extract text from only the first page.

To extract text from a page, you need to get a Page object, which represents a single page of a PDF, from a PdfFileReader object. You can get a Page object by calling the getPage() method ❷ on a PdfFileReader object and passing it the page number of the page you're interested in—in our case, 0.

3- The total number of pages in the document is stored in the numPages attribute of a PdfFileReader object ❶

```
4- >>> import PyPDF2
>>> pdfReader = PyPDF2.PdfFileReader(open('encrypted.pdf', 'rb'))
❶ >>> pdfReader.isEncrypted
True
>>> pdfReader.getPage(0)
❷ Traceback (most recent call last):
  File "<pyshell#173>", line 1, in <module>
    pdfReader.getPage()
--snip--
  File "C:\Python34\lib\site-packages\PyPDF2\pdf.py", line 1173, in getObject
    raise utils.PdfReadError("file has not been decrypted")
PyPDF2.utils.PdfReadError: file has not been decrypted
❸ >>> pdfReader.decrypt('swordfish')
1
>>> pageObj = pdfReader.getPage(0)
```

5- This method involves using shortcuts to rotate a PDF. While viewing a PDF document, press the "Ctrl+Shift+L" keys or "Ctrl+Shift+R" keys to rotate the current page to the left or right.

6- Each Paragraph object also has a runs attribute that is a list of Run objects. Run objects also have a text attribute, containing just the text in that particular run. Let's look at the text attributes in the second Paragraph object, 'A plain paragraph with some bold and some italic' .]

8- Run object

9- True / False are used where the field takes a boolean value. None is used where the field takes an Optional[List] .

10- The first step is to install this third-party module python-docx. You can use pip “pip install python-docx” or download the tarball from here. ...
After installation import “docx” NOT “python-docx”.
Use “docx. Document” class to start working with the word document.

11- by using append file handling.

12- The integer 0 makes the heading the Title style, which is used for the top of the document. Integers 1 to 4 are for various heading levels, with 1 being the main heading and 4 the lowest subheading.