1. The module PdfFileReader() should be open in ‘rb’ i.e. read binary, and the module RdfFileWriter() should be open in ‘wb’ write binary.

2. It is done as the following,

from PyPDF2 import PdfFileReader

pdf\_reader = PdfFileReader(‘filename.pdf’, ‘rb’)

pdf\_reader.getPage(5)

3. To get the number of pages in the pdf file, we can do the following.

from PyPDF2 import PdfFileReader

pdf\_reader = PdfFileReader(‘filename.pdf’, ‘rb’)

numb\_pages = pdf\_reader.getNumPages()

4. If the pdf file is encrypted with a password ‘swardfish’, we must decrypt it befor we obtain the pages of the pdf file

from PyPDF2 import PdfFileReader

pdf\_reader = PdfFileReader(‘filename.pdf’, ‘rb’)

pdf\_reader.decrypt(‘swardfish’)

once the pdf is decrypted, we can use it’s content.

5. There are main two methods to rotate a page

rotateClockWise(), rotate the page in click wise direction for 90 degree

rotateCounterClockWise(), rotate the page in counter clock wise direction for 90 degree

6. The document contains one or more Paraghraph object, and a Paragraph contains one or more Run object, when we press enter new paragraph is created, while when we make change in font style, we need to call run object.

7. We can get all the paraghraphs on the document by using the following piece of code,

from docx import Document

doc = Document("sample\_file.docx")

print(doc.paragraphs)

for paragraph in doc.paragraphs:

print(paragraph.text)

8. RUN object has all the variable that can do the styling of the text.

9. Bold variable can be set True means set the text to bold, False means set the text to not bold, and None means not applicable.

10. We can create a new document object

from docx import Document

document = Document()

11. We can add a paraghraph in a document object

from docx import Document

doc = Document()

doc.add\_paragraph('Hello, there!')

12. The levels for a heading in a word document can be specified by using the level attribute inside the add\_heading method. There are a total of 5 levels statring for 0 to 4. where level 0 makes a headline with the horizontal line below the text, whereas the heading level 1 is the main heading. Similarly, the other headings are sub-heading with their's font-sizes in decreasing order.