1. In what modes should the PdfFileReader() and PdfFileWriter() File objects will be opened?

PdfFileReader() - the file object needs to be opened in ‘rb’ read-binary mode

PdfFileWriter() - the file object needs to be opened in ‘wb’ write-binary mode

2. From a PdfFileReader object, how do you get a Page object for page 5?

import PyPDF2

Pdfnew = open(“inputfile.pdf”,’rb’)

pdfread = PyPdf.pdfnew

pdfread.getPage(4)

#getPage(4) - to be used to get the Page 5

3. What PdfFileReader variable stores the number of pages in the PDF document?

import PyPDF2

Pdfnew = open(“inputfile.pdf”,’rb’)

pdfread = PyPdf.pdfnew

numpage = pdfread.numPages

#numpage will store the number of pdf pages

4. If a PdfFileReader object’s PDF is encrypted with the password swordfish, what must you do before you can obtain Page objects from it?

import PyPdf2

pdfnew = open(“inputencryptpdf.pdf”,”rb”)

pdfread = PyPdf.pdfnew

pdfread.decrypt(‘swordfish’)

pagezero = pdfread.getpage(0)

#using decrypt using the password ‘swordfish’

5. What methods do you use to rotate a page?

import PyPdf2

pdfnew = open(“inputencryptpdf.pdf”,”rb”)

pdfread = PyPdf.pdfnew

pagezero = pdfread.getpage(0)

pdfzero.rotateClockwise(90)

pdfzero.rotateCounterClosewise(90)

#using rotateClosewise(‘degree’) & rotateCounterClosewise we can rotate a pdf file

6. What is the difference between a Run object and a Paragraph object?

A run object is a continuous string with the same styles.

For example : let say we have text paragraphs as below

This is text paragraph to show styles like **bold**,*italic*,colors,fonts inside it.

In this example , “This is text paragraph to show styles like “ has the same kind of attributes in terms of fonts , styles or color. so they are 1 run object followed by another run object in bold , then italic and color.

for a change in style we need a new run object

while a paragraph object will not have these attributes inside it. each paragraph object will have text attributes without any styling attributes in it. each paragraph can contain multiple run objects inside it.

7. How do you obtain a list of Paragraph objects for a Document object that’s stored in a variable named doc?

import Docx

doc = docx.Document(“worddocument.docx”)

listpara = []

for para in doc.paragraphs:

listpara.append(para.text)

‘/n’.join(listpara)

#by using doc.paragraphs we can get a list of paragraph object for a document stored in variable doc

8. What type of object has bold, underline, italic, strike, and outline variables?

The run object has these kind of variables

9. What is the difference between False, True, and None for the bold variable?

False in run attributes means the attribute will be disabled for the run text

True in run attributes means the attribute will be enabled for the run text

None in run attributes means the attribute will be set as per the default value set

10. How do you create a Document object for a new Word document?

doc = docx.Document() #will create a new empty word document

11. How do you add a paragraph with the text 'Hello, there!' to a Document object stored in a variable named doc?

doc = docx.Document()

doc.add\_paragraph(‘Hello,there!’)

#using doc.add\_paragraph will add the required text into the word document

12. What integers represent the levels of headings available in Word documents?

0,1,2,3 and 4 are the different level of heading in word document