Assignment 12 Solutions

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

ANS: For **pdfFileReader( )** file objects should be opened in **rb** -> read binary mode,

Whereas for **pdfFileWriter( )** file objects should be opened in **wb** -> write binary mode.

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

ANS: pdfFileReader class provides a method called **getPage(page_no)** to get a page object.

```python
#Example Code:
from PyPDF2 import PdfFileReader
pdf_reader = PdfFileReader(file_path)
for page in pdf_reader.getNumPages():
    pdf_reader.getPage(page)
```

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

ANS: getNumPages() method of **PdfFileReader** class stores the no pages in a PDF document

```python
#Example Code:
from PyPDF2 import PdfFileReader
pdf_reader = PdfFileReader(file_path)
print(pdf_reader.getNumPages())
```

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?

ANS: If a **PdfFileReader** object's PDF is encrypted with the password **swordfish** and you're not aware of it.first read the Pdf using the PdfFileReader Class.PdfFileReader class provides a attribute called **isEncrypted** to check whether a pdf is encrypted or not.the method returns true if a pdf is encrypted and vice versa. if pdf is encrypted use the **decrypt( )** method provided by PdfFileReader class first then try to read the contents/pages of the pdf,else PyPDF2 will raise the following error

PyPDF2.utils.PdfReadError: file has not been decrypted

```python
#Example Code:
from PyPDF2 import PdfFileReader
pdf_reader = PdfFileReader(file_path)
if pdf_reader.isEncrypted:
    pdf_reader.decrypt("swordfish")
for page in pdf_reader.pages:
    print(page.extractText())
```

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

ANS: PyPDF2 Package provides 2 methods to rotate a page:

1.**rotateClockWise( )** -> For Clockwise rotation

2.**rotateCounterClockWise( )** -> For Counter Clockwise rotation

The PyPDF2 package only allows you to rotate a page in increments of 90 degrees.You will receive an AssertionError otherwise.

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

ANS: The structure of a document is represented by three different data types in **python-Docx**.At the highest level,a **Document** object represents the entire document.The Document object contains a list of **Paragraph** objects for the paragraphs in the document.(A new paragraph begins whenever the user presses **ENTER** or **RETURN** while typing in a Word document.)Each of these Paragraph objects contains a list of one or more **RUN** objects.

The text in a Word document is more than just a string.It has font,size,color,and other styling information associated with it.A style in Word is a collection of these attributes.A Run object is needed whenever the text style changes.

7. How do you obtain a list of Paragraph objects for a Document object that's stored in a variable

named doc?

```
# Example Program
from docx import Document
doc = Document("sample_file.docx")
print(doc.paragraphs)
for paragraph in doc.paragraphs:
    print(paragraph.text)
```

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

ANS: Run object has bold,underline,italic,strike,and outline variables.The text in a Word document is more than just a string.It has font,size,color,and other styling information associated with it.

A style in Word is a collection of these attributes.A Run object is a contiguous run of text with the same style.A new Run object is needed whenever the text style changes.

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

```
bold = True
bold = False
bold = None
```

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

```
# Example Program
from docx import document
document = Document()
document.add_paragraph("iNeuron Full Stack DataScience Course")
document.save('mydocument.docx')
```

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

```
# Example Program
from docx import Document
doc = Document()
doc.add_paragraph('Hello, there!')
doc.save('hello.docx')
```

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

ANS: 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 starting 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.

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js