|  |  |
| --- | --- |
|  | Chapter 13 Practice Questions – Working with PDF and word Documents |
| **Q1** | **A string value of the PDF filename is not passed to the PyPDF2.PdfFileReader() function. What do you pass to the function instead?** |
| A | You pass it a pdfFileObj that is stored from opening the pdf file name  Example:  pdfFileObj = open(‘pdfdocument.pdf’, ‘rb’)  pdfReader = PyPDF2.PdfFileReader(pdfFileObj)  \*rb means open in read binary mode |
| **Q2** | **What modes do the File objects for PdfFileReader() and pdfFileWriter() need to be opened in?** |
| A | For PdfFileReader() function mode needs to be in read binary (rb)  For pdfFileWriter() function mode needs to be in write binary (wb) |
| **Q3** | **How do you acquire a Page object for About This Book from a PdfFileReader object?** |
| A | A page object is a single page of a PDF. You can get a page object by calling the getPage() method on a PdfFileReader object and passing it the page number of the page you’re interested in. |
| **Q4** | **What PdfFileReader variable stores the number of pages in the PDF document?** |
| A | The total number of pages in the document is stored in the numPages attribute of a PdfFileReader object  numPages (used after the pdfReader object like: pdfReader.numPages |
| **Q5** | **If a PdfFileReader object’s PDF is encrypted with the password swordfish, what must you do before you can obtain Page objects from it?** |
| A | To read an encrypted PDF, call the decrypt() function and pass the password as a string  Example:  pdfReader.decrypt(‘swordfish’) |
| **Q6** | **What methods do you use to rotate a page?** |
| A | The pages of a PDF can also be rotated in 90-degree increments with the rotateClockwise() and rotateCounterClockwise() methods. |
| **Q7** | **What method returns a Document object for a file named demo.docx?** |
| A | We open a .docx file in Python, call docx.Document(), and pass the filename demo.docx.  Example:  Doc = docx.Document(‘demo.docx’) |
| **Q8** | **What is the difference between a Paragraph object and a Run object?** |
| A | A paragraph object begins whenever a user presses ENTER or RETURN. A paragraph object can contain a list of Run objects (A Run object is a contiguous run of text with the same style) |
| **Q9** | **How do you obtain a list of Paragraph objects for a Document object that’s stored in a variable named doc?** |
| A | Use the paragraphs attribute on a document object such as doc. If we use the len function on doc.paragraphs it would return the list of paragraphs in the document object |
| **Q10** | **What type of object has a bold, underline, italic, strike, and outline variables?** |
| A | Runs can be further styled using text attributes and each attribute can be set to one of three values: True, False or None.  Example:  doc.paragraphs[1].runs[0].underline = True |
| **Q11** | **What is the difference between setting the bold variable to True, False, or None?** |
| A | True makes a run bold. False would unbold a run. None removes any text attributes on the run |
| **Q12** | **How do you create a Document object for a new Word document?** |
| A | doc = docx.Document()  Where doc is the new document object |
| **Q13** | **How do you create a Document with the text ‘Hello there!’ to a Document object stored in a variable named doc?** |
| A | Create the document object and then add the text as a paragraph and save.  Example:  doc = docx.Document()  doc.add\_paragraph(‘Hello world!’)  doc.save(‘helloworld.docx’) |
| **Q14** | **What integers represent the levels of headings available in Word documents?** |
| A | Integers 0 – 4 represent the levels of headings available in Word  Example:  doc.add\_heading(‘Header 0’, 0)  The arguments to add\_heading() are a string of the heading text and an integer from 0 to 4 |