

Install pip PyPdf3

```
pip install pypdf3

Requirement already satisfied: pypdf3 in /usr/local/lib/python3.10/dist-packages (1.0.6)
Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-packages (from pypdf3) (4.66.1)
```

Import Now!

```
import PyPDF3 as MyPdf

from google.colab import files
uploaded = files.upload()

Choose Files Test_Pillow.pdf
• Test_Pillow.pdf(application/pdf) - 1240756 bytes, last modified: 10/19/2021 - 100% done
Saving Test_Pillow.pdf to Test_Pillow (1).pdf
```

Extracting Pdf document information (Title, Author....)

```
#First we have to open the file !
Mypdf_File = open("Test_Pillow.pdf" , "rb")

#Now we have 2 opt (read or write) here i will read
pdf_reader = MyPdf.PdfFileReader(Mypdf_File , strict = False) #strict = True means show the error if have

#get data from pdf file
print(pdf_reader.getNumPages())

doc_info = pdf_reader.getDocumentInfo()
print(doc_info)

#-----Go Through All Data----->
for item in doc_info:
    print("Item" , item)
    print("Value" ,doc_info[str(item)])

#-----
FullInfo = str(item).removeprefix("/")+"."+doc_info[str(item)]
print(FullInfo)

26
{'/Author': 'Rec User', '/CreationDate': "D:20211019145802-03'00'", '/ModDate': "D:20211019145802-03'00'", '/Producer': 'Microsoft: Print To PDF', '/Title': '39 .pillow.rtf'}
Item /Author
Value Rec User
Item /CreationDate
Value D:20211019145802-03'00'
Item /ModDate
Value D:20211019145802-03'00'
Item /Producer
Value Microsoft: Print To PDF
Item /Title
Value 39 .pillow.rtf
Title:39 .pillow.rtf
```

Colab paid products - [Cancel contracts here](#)

✓ 0s completed at 7:49 PM

