

```

import os

from zipfile import ZipFile

direccion = 'C:/Users/johnv/python_books.zip'

# Descomprimir el archivo ZIP

with ZipFile(direccion) as archivo:

    archivo.extractall("C:/Users/johnv/")


import fitz

documento1 = fitz.open('C:/Users/johnv/python_books/Python - AWS.pdf')

print("Número de páginas: ", documento1.pageCount)

print("Metados: ", documento1.metadata)

```

```

pagina = documento1.loadPage(0)

texto = pagina.getText("text")

texto.count("")

```

```

Número de páginas:  1
Metados:  {'format': 'PDF 1.3', 'title': '', 'author': '', 'subject': '',
'keywords': '', 'creator': 'Adobe InDesign CS6 (Windows)', 'producer': 'm
acOS Version 11.2.1 (Build 20D74) Quartz PDFContext', 'creationDate': "D:
20210225223227Z00'00'", 'modDate': "D:20210225223227Z00'00'", 'trapped':
'', 'encryption': None}

```

```

Out[27]:
1821

```

```

import fitz

documento1 = fitz.open('C:/Users/johnv/python_books/Python  Data Science
Cookbook.pdf')

print("Número de páginas: ", documento1.pageCount)

print("Metados: ", documento1.metadata)

```

```
pagina = documento1.loadPage(0)
texto = pagina.getText("text")
texto.count("")
```

```
Número de páginas: 1
Metados: {'format': 'PDF 1.3', 'title': 'Python Data Science Cookbook',
'author': 'Gopi Subramanian', 'subject': '', 'keywords': '', 'creator': '
calibre 2.35.0 [http://calibre-ebook.com]', 'producer': 'macOS Version 11
.2.1 (Build 20D74) Quartz PDFContext', 'creationDate': "D:20210225223143Z
00'00'", 'modDate': "D:20210225223143Z00'00'", 'trapped': '', 'encryption
': None}
```

Out[29]:

3372

```
import fitz
```

```
documento1 = fitz.open('C:/Users/johnv/python_books/Python - Finance.pdf')
```

```
print("Número de páginas: ", documento1.pageCount)
```

```
print("Metados: ", documento1.metadata)
```

```
pagina = documento1.loadPage(0)
```

```
texto = pagina.getText("text")
```

```
texto.count("")
```

```
Número de páginas: 1
Metados: {'format': 'PDF 1.3', 'title': 'Python for Finance', 'author':
'Yuxing Yan', 'subject': '', 'keywords': '', 'creator': 'Adobe InDesign C
S6 (Windows)', 'producer': 'macOS Version 11.2.1 (Build 20D74) Quartz PDF
Context', 'creationDate': "D:20210225223246Z00'00'", 'modDate': "D:202102
25223246Z00'00'", 'trapped': '', 'encryption': None}
```

Out[30]:

1346

Respuesta: El texto con mayor cantidad de palabras es Python Data Science Cookbook.pdf