# Image processing:

## 1.Read a jpeg image and print the image file

In [1]:

```
from PIL import Image
img = Image.open('book.jpg')
print(img.format)
```

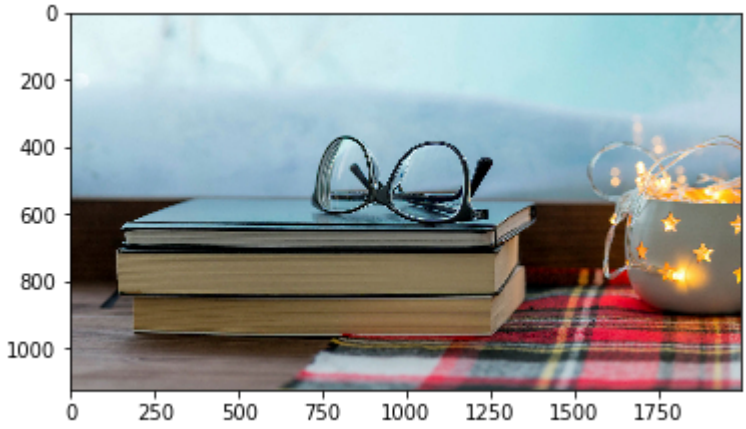
```
JPEG
```

In [3]:

```
import matplotlib.pyplot as plt
plt.imshow(img)
```

Out[3]:

```
<matplotlib.image.AxesImage at 0x24e1b622608>
```




## 2. Merge two pdf files using python script

In [5]:

```
pip install PyPDF2
```

```
Collecting PyPDF2
  Downloading PyPDF2-1.26.0.tar.gz (77 kB)
Building wheels for collected packages: PyPDF2
  Building wheel for PyPDF2 (setup.py): started
  Building wheel for PyPDF2 (setup.py): finished with status 'done'
  Created wheel for PyPDF2: filename=PyPDF2-1.26.0-py3-none-any.whl size=610
87 sha256=bec1bb49deb7a5e5449f53eced831170ca17aba89b01caafe4e4d57691fa939e
  Stored in directory: c:\users\vinit\appdata\local\pip\cache\wheels\80\1a\2
4\648467ade3a77ed20f35cfd2badd32134e96dd25ca811e64b3
Successfully built PyPDF2
Installing collected packages: PyPDF2
Successfully installed PyPDF2-1.26.0
Note: you may need to restart the kernel to use updated packages.
```

In [13]:

```
import PyPDF2
import sys

inputs = sys.argv[1:]

def pdfmerge(pdf_list):
    merger = PyPDF2.PdfFileMerger()
    for pdf in pdf_list:
        print(pdf)
        merger.append(pdf)
    merger.write('super.pdf')
```

In [15]:

```
pdf_list=['Python GUI Programming Cookbook.pdf','data communication.pdf']
pdfmerge(pdf_list)
```

```
Python GUI Programming Cookbook.pdf
data communication.pdf
```

In [16]:

```
pdf3 = open("super.pdf",'rb')
```

In [17]:

```
pdf3Reader = PyPDF2.PdfFileReader(pdf3)
```

In [18]:

```
print(pdf3Reader)
```

```
<PyPDF2.pdf.PdfFileReader object at 0x0000024E1BA602C8>
```

In [19]:

```
pdf3.close()
```

# 3. Scrape a website and store the data into DB.

In [21]:

```
import requests
from bs4 import BeautifulSoup
```

In [22]:

```
import mysql.connector as sql
conn=sql.connect(
    host="localhost",
    user="root",
    passwd="Vinita@28",
)
print(conn)
```

```
<mysql.connector.connection.MySQLConnection object at 0x0000024E1FB59988>
```

In [23]:

```
import mysql.connector as sql
conn=sql.connect(
    host="localhost",
    user="root",
    passwd="Vinita@28",
    database="employee_mangement"
)
curr= conn.cursor()
```

In [24]:

```
url_to_scrape = 'https://www.flipkart.com/audio-video/headphones/pr?sid=0pm%2Cfcn&otracker=
plain_html_text = requests.get(url_to_scrape)
soup = BeautifulSoup(plain_html_text.text, "html.parser")
```

In [25]:

```
print(soup.prettify())
```

```
<!DOCTYPE html>
<html lang="en">
 <head>
  <link href="https://rukminim1.flixcart.com" rel="preconnect"/>
  <link href="//static-assets-web.flixcart.com/www/linchpin/fk-cp-zion/cs
s/app.chunk.3e7df7.css" rel="stylesheet"/>
  <meta content="text/html; charset=utf-8" http-equiv="Content-type"/>
  <meta content="IE=Edge" http-equiv="X-UA-Compatible"/>
  <meta content="102988293558" property="fb:page_id"/>
  <meta content="658873552,624500995,100000233612389" property="fb:admin
s"/>
  <meta content="noodp" name="robots"/>
  <link href="https://static-assets-web.flixcart.com/www/promos/new/201505
28-140547-favicon-retina.ico" rel="shortcut icon"/>
  <link href="/osdd.xml?v=2" rel="search" type="application/opensearchdesc
ription+xml"/>
  <meta content="website" property="og:type"/>
  <meta content="Flipkart.com" name="og_site_name" property="og:site_nam
e"/>
```

In [ ]: