**ANONYMIZING DATA FROM GIVEN SAMPLE DATA (PATIENT ID only)**

**NOTE- The following document is divided in three parts.**

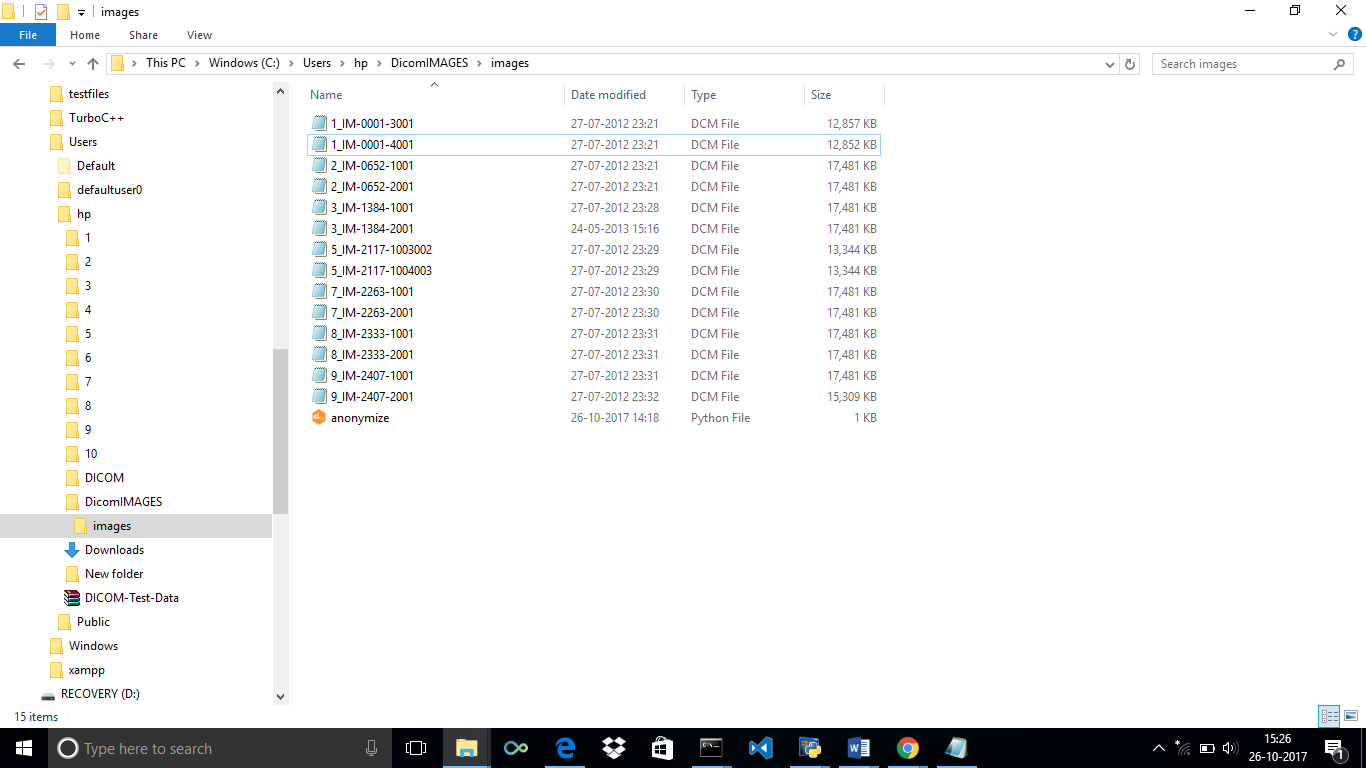
Part-1 Anonymizing Patient ID from given sample data.

Part-2 Error in the files not Anonymized.

Part-3 Reading the dcm images and data.

**PART-1**

* Download the zip file and extract it.
* Copy all the images in a folder (which are currently in different folders 1,2,3…)
* Create a python file anonymize.py inside the same folder, which contains the code to anonymize the patient ID of all dcm images placed within the same folder.



* Creating anonymize.py :-

Code-

import dicom

import glob

import os, sys

from os.path import join, expanduser

dicomPath = sys.argv[0]

for file in glob.glob("\*.dcm"):

print("### Now reading file:", file)

df = dicom.read\_file(file,'rb')

print(file)

df[0x10,0x20].value="XXXX";

print(df[0x10,0x20]).value

df.save\_as("modified"+file);

print("end")

* Open command prompt and go to the location where the above images and file is placed:-

C:\Users\hp\DicomIMAGES\images>python anonymize.py

* OUTPUT-

('### Now reading file:', '1\_IM-0001-3001.dcm')

1\_IM-0001-3001.dcm

XXXX

('### Now reading file:', '1\_IM-0001-4001.dcm')

1\_IM-0001-4001.dcm

XXXX

('### Now reading file:', '2\_IM-0652-1001.dcm')

2\_IM-0652-1001.dcm

XXXX

('### Now reading file:', '2\_IM-0652-2001.dcm')

2\_IM-0652-2001.dcm

XXXX

('### Now reading file:', '3\_IM-1384-1001.dcm')

3\_IM-1384-1001.dcm

XXXX

('### Now reading file:', '3\_IM-1384-2001.dcm')

3\_IM-1384-2001.dcm

XXXX

('### Now reading file:', '5\_IM-2117-1003002.dcm')

5\_IM-2117-1003002.dcm

XXXX

('### Now reading file:', '5\_IM-2117-1004003.dcm')

5\_IM-2117-1004003.dcm

XXXX

('### Now reading file:', '7\_IM-2263-1001.dcm')

7\_IM-2263-1001.dcm

XXXX

('### Now reading file:', '7\_IM-2263-2001.dcm')

7\_IM-2263-2001.dcm

XXXX

('### Now reading file:', '8\_IM-2333-1001.dcm')

8\_IM-2333-1001.dcm

XXXX

('### Now reading file:', '8\_IM-2333-2001.dcm')

8\_IM-2333-2001.dcm

XXXX

('### Now reading file:', '9\_IM-2407-1001.dcm')

9\_IM-2407-1001.dcm

XXXX

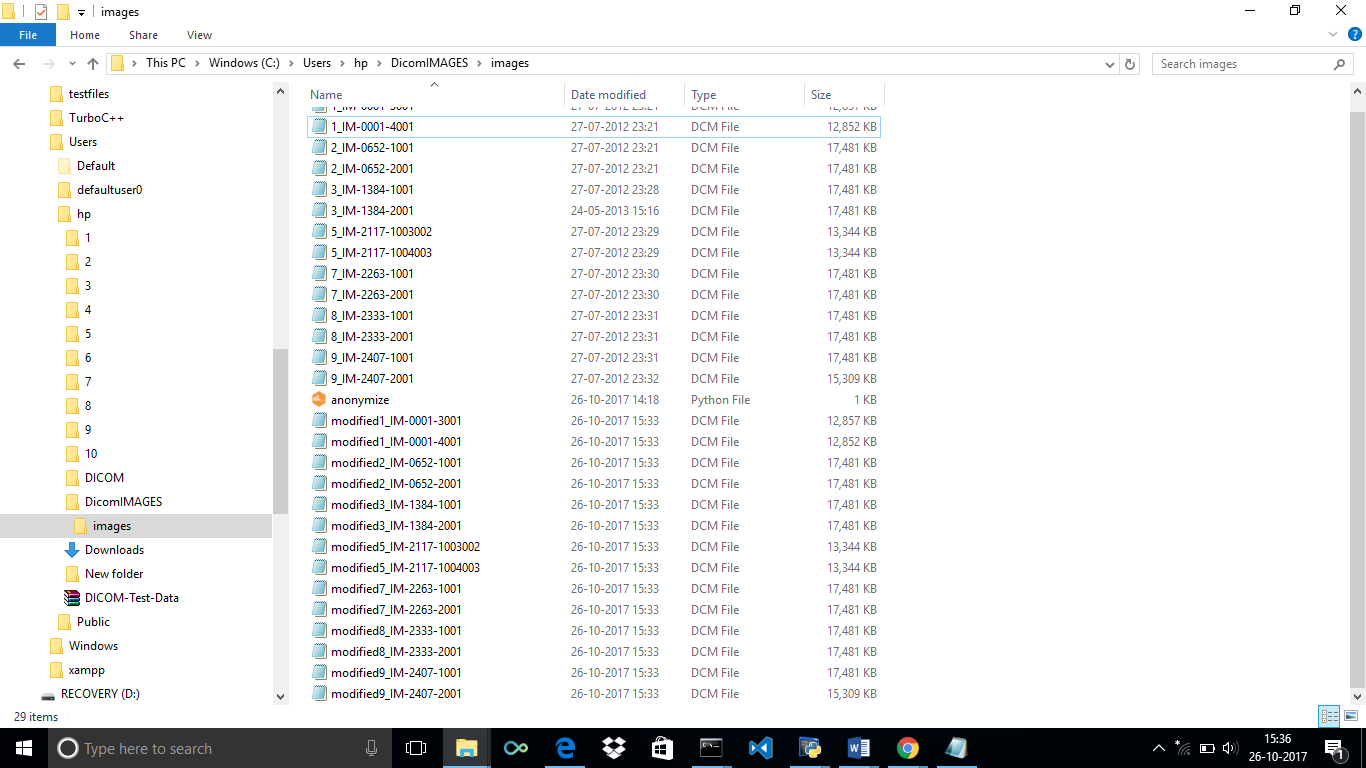
('### Now reading file:', '9\_IM-2407-2001.dcm')

9\_IM-2407-2001.dcm

XXXX

end

**The folder now contains the anonymized patient Id dcm file with Prefix modified added in file name: -**



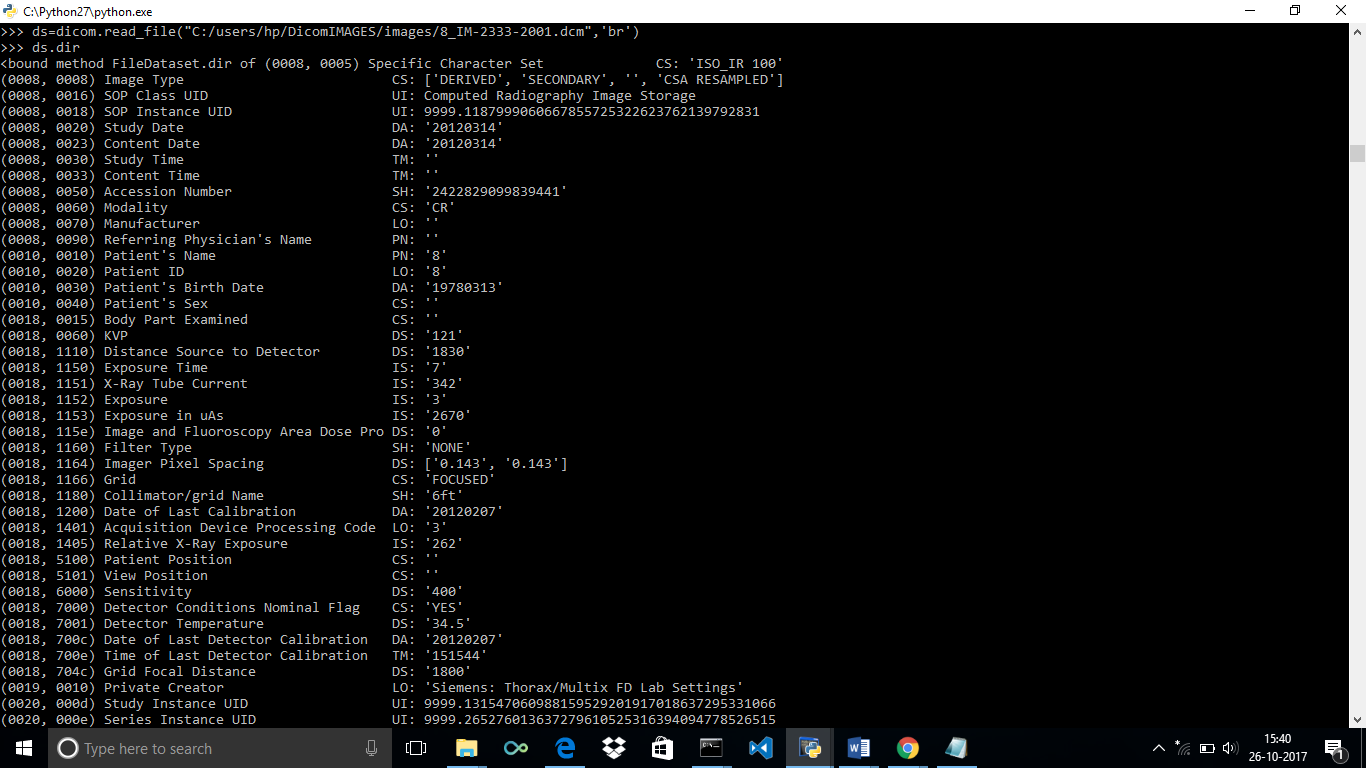
* Go to python.exe to see the anonymized PATIENT ID: -
* We will check first the ID before Anonymized and then after it.

**Before-**

* Write the following command in Python.exe-

**>>> ds=dicom.read\_file("C:/users/hp/DicomIMAGES/images/8\_IM-2333-2001.dcm",'br')**

**>>> ds.dir**



Taking few data here-

(0008, 0070) Manufacturer LO: ''

(0008, 0090) Referring Physician's Name PN: ''

(0010, 0010) Patient's Name PN: '8'

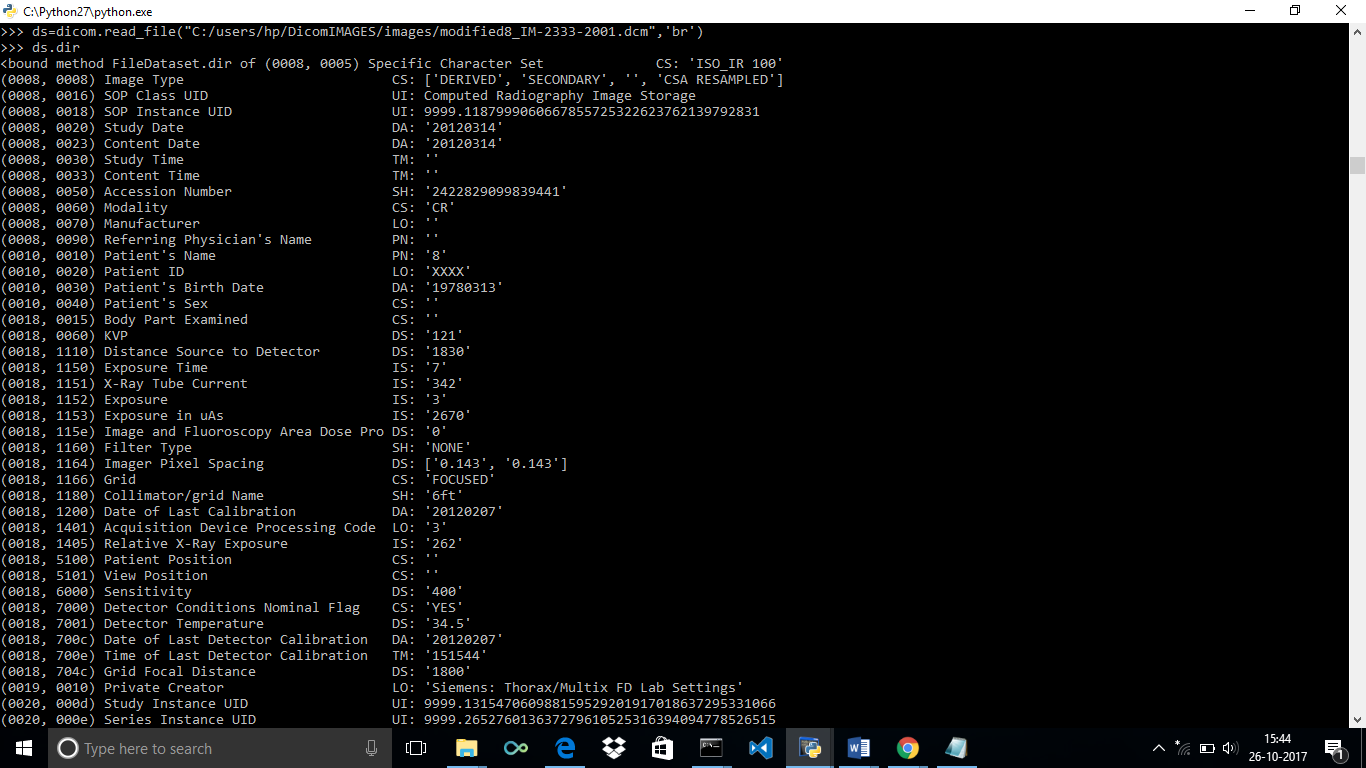
(0010, 0020) Patient ID LO: '8'

(0010, 0030) Patient's Birth Date DA: '19780313'

After-

**>>> ds=dicom.read\_file("C:/users/hp/DicomIMAGES/images/modified8\_IM-2333-2001.dcm",'br')**

**>>> ds.dir**



* (Taking few data here)

(0008, 0070) Manufacturer LO: ''

(0008, 0090) Referring Physician's Name PN: ''

(0010, 0010) Patient's Name PN: '8'

(0010, 0020) Patient ID LO: 'XXXX'

(0010, 0030) Patient's Birth Date DA: '19780313'

(0010, 0040) Patient's Sex CS: ''

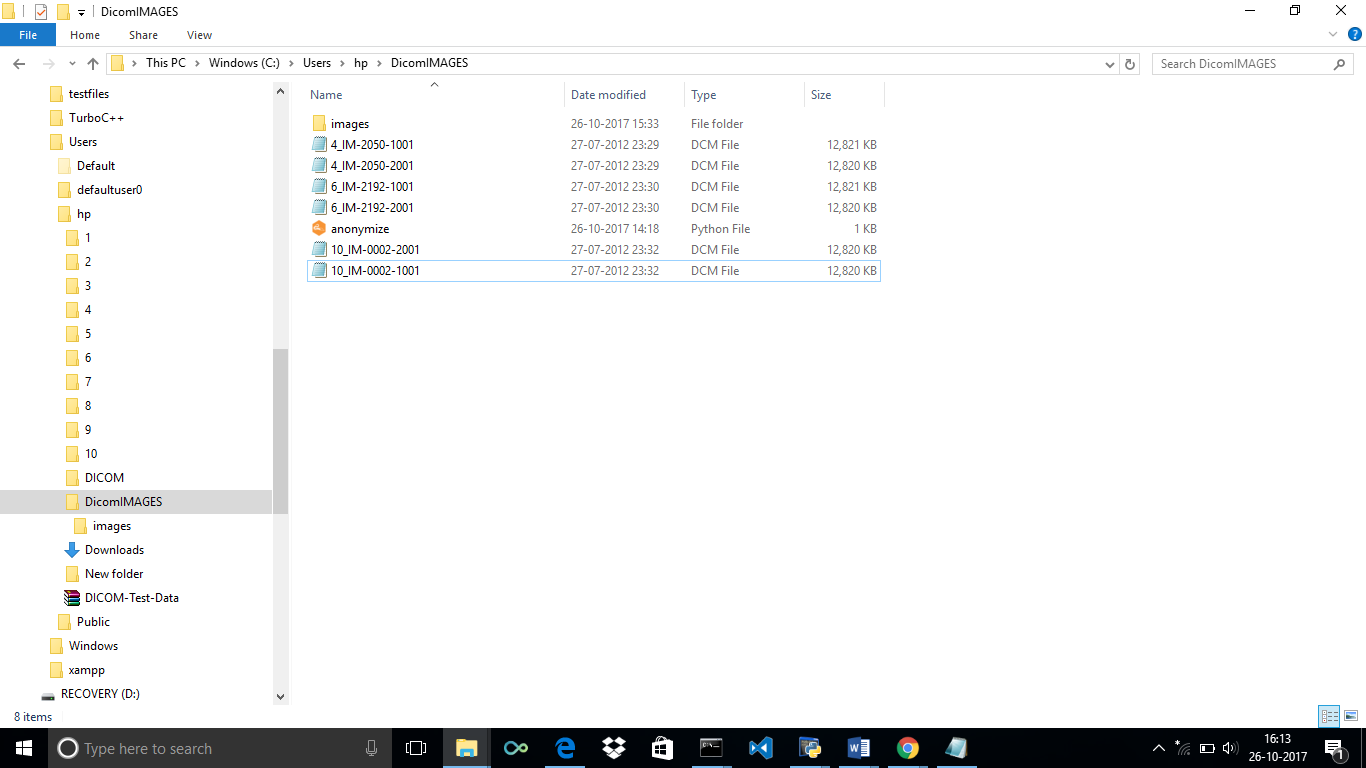
(0018, 0015) Body Part Examined CS: ''

Similarly, we can see for different images also.

**PART -2**

ERROR in 6 images-

These are the 6 images (which are placed inside folders 4,6,10 inside the zip.)



This is the error while just reading these 6 files-

**>>>ds=dicom.read\_file("C:/users/hp/DicomIMAGES/10\_IM-0002-2001.dcm",'br')**

**>>> ds.dir**

Traceback (most recent call last):

File "<stdin>", line 1, in <module>

File "C:\Python27\lib\site-packages\pydicom-0.9.8-py2.7.egg\dicom\dataset.py", line 514, in \_\_str\_\_

return self.\_pretty\_str()

File "C:\Python27\lib\site-packages\pydicom-0.9.8-py2.7.egg\dicom\dataset.py", line 446, in \_pretty\_str

strings.append(dataset.\_pretty\_str(indent + 1))

File "C:\Python27\lib\site-packages\pydicom-0.9.8-py2.7.egg\dicom\dataset.py", line 441, in \_pretty\_str

for data\_element in self:

File "C:\Python27\lib\site-packages\pydicom-0.9.8-py2.7.egg\dicom\dataset.py", line 315, in \_\_iter\_\_

yield self[tag]

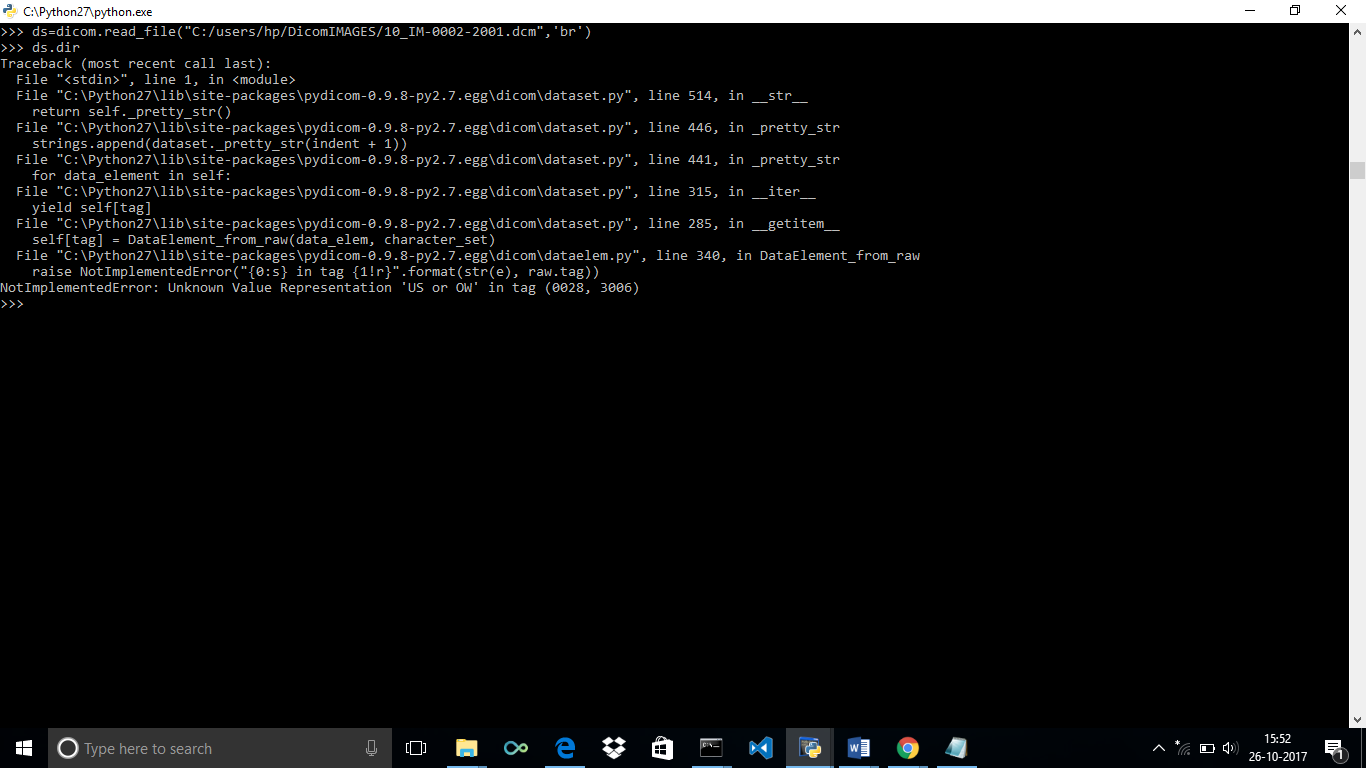
File "C:\Python27\lib\site-packages\pydicom-0.9.8-py2.7.egg\dicom\dataset.py", line 285, in \_\_getitem\_\_

self[tag] = DataElement\_from\_raw(data\_elem, character\_set)

File "C:\Python27\lib\site-packages\pydicom-0.9.8-py2.7.egg\dicom\dataelem.py", line 340, in DataElement\_from\_raw

raise NotImplementedError("{0:s} in tag {1!r}".format(str(e), raw.tag))

NotImplementedError: Unknown Value Representation 'US or OW' in tag (0028, 3006)



**ERROR**

While executing the anonymize python file.

C:\Users\hp\DicomIMAGES>python anonymize.py

('### Now reading file:', '10\_IM-0002-1001.dcm')

10\_IM-0002-1001.dcm

XXXX

Traceback (most recent call last):

File "anonymize.py", line 12, in <module>

df.save\_as("modified"+file);

File "C:\Python27\lib\site-packages\pydicom-0.9.8-py2.7.egg\dicom\dataset.py", line 467, in save\_as

dicom.write\_file(filename, self, WriteLikeOriginal)

File "C:\Python27\lib\site-packages\pydicom-0.9.8-py2.7.egg\dicom\filewriter.py", line 344, in write\_file

write\_dataset(fp, dataset)

File "C:\Python27\lib\site-packages\pydicom-0.9.8-py2.7.egg\dicom\filewriter.py", line 196, in write\_dataset

write\_data\_element(fp, dataset[tag], dataset\_encoding)

File "C:\Python27\lib\site-packages\pydicom-0.9.8-py2.7.egg\dicom\filewriter.py", line 160, in write\_data\_element

writers[VR](fp, data\_element, encoding=encoding)

File "C:\Python27\lib\site-packages\pydicom-0.9.8-py2.7.egg\dicom\filewriter.py", line 207, in write\_sequence

write\_sequence\_item(fp, dataset, encoding)

File "C:\Python27\lib\site-packages\pydicom-0.9.8-py2.7.egg\dicom\filewriter.py", line 217, in write\_sequence\_item

write\_dataset(fp, dataset, parent\_encoding=encoding)

File "C:\Python27\lib\site-packages\pydicom-0.9.8-py2.7.egg\dicom\filewriter.py", line 196, in write\_dataset

write\_data\_element(fp, dataset[tag], dataset\_encoding)

File "C:\Python27\lib\site-packages\pydicom-0.9.8-py2.7.egg\dicom\dataset.py", line 285, in \_\_getitem\_\_

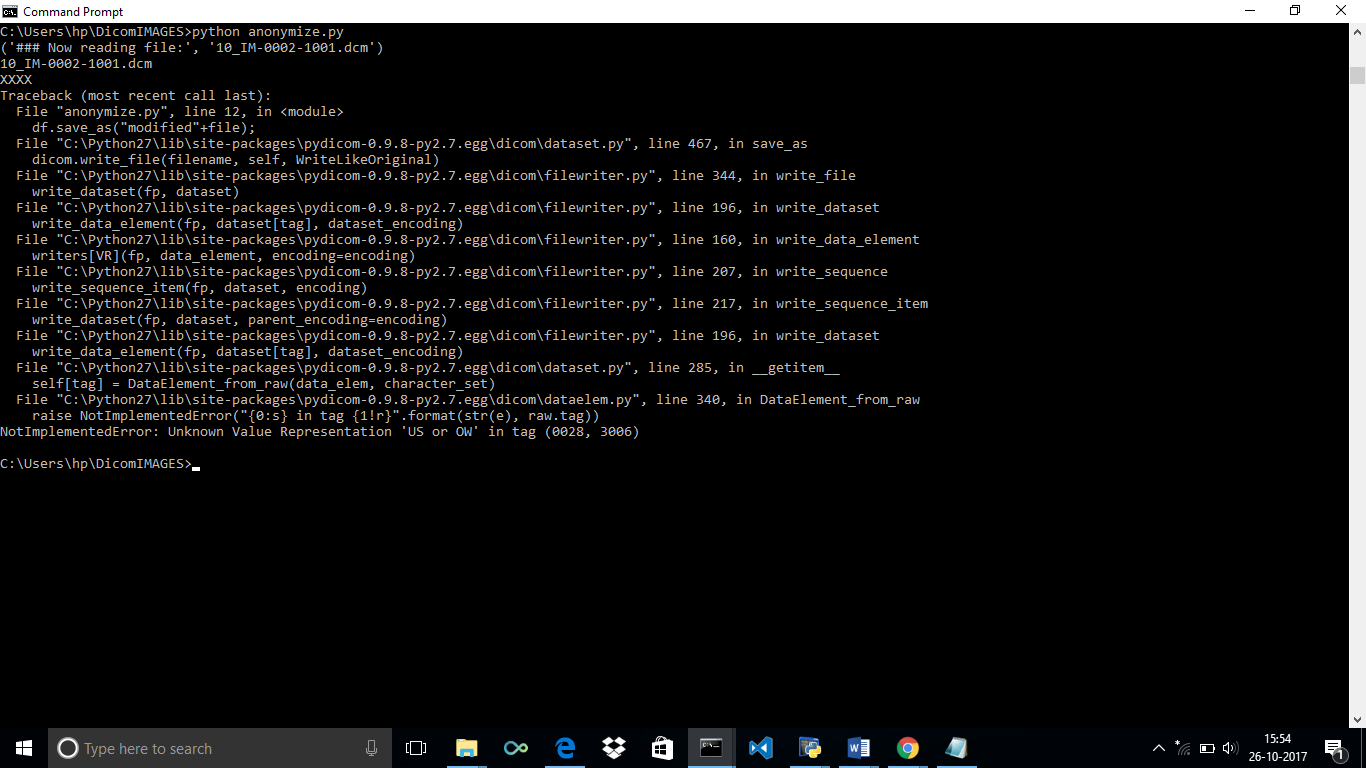
self[tag] = DataElement\_from\_raw(data\_elem, character\_set)

File "C:\Python27\lib\site-packages\pydicom-0.9.8-py2.7.egg\dicom\dataelem.py", line 340, in DataElement\_from\_raw

raise NotImplementedError("{0:s} in tag {1!r}".format(str(e), raw.tag))

NotImplementedError: Unknown Value Representation 'US or OW' in tag (0028, 3006)

C:\Users\hp\DicomIMAGES>

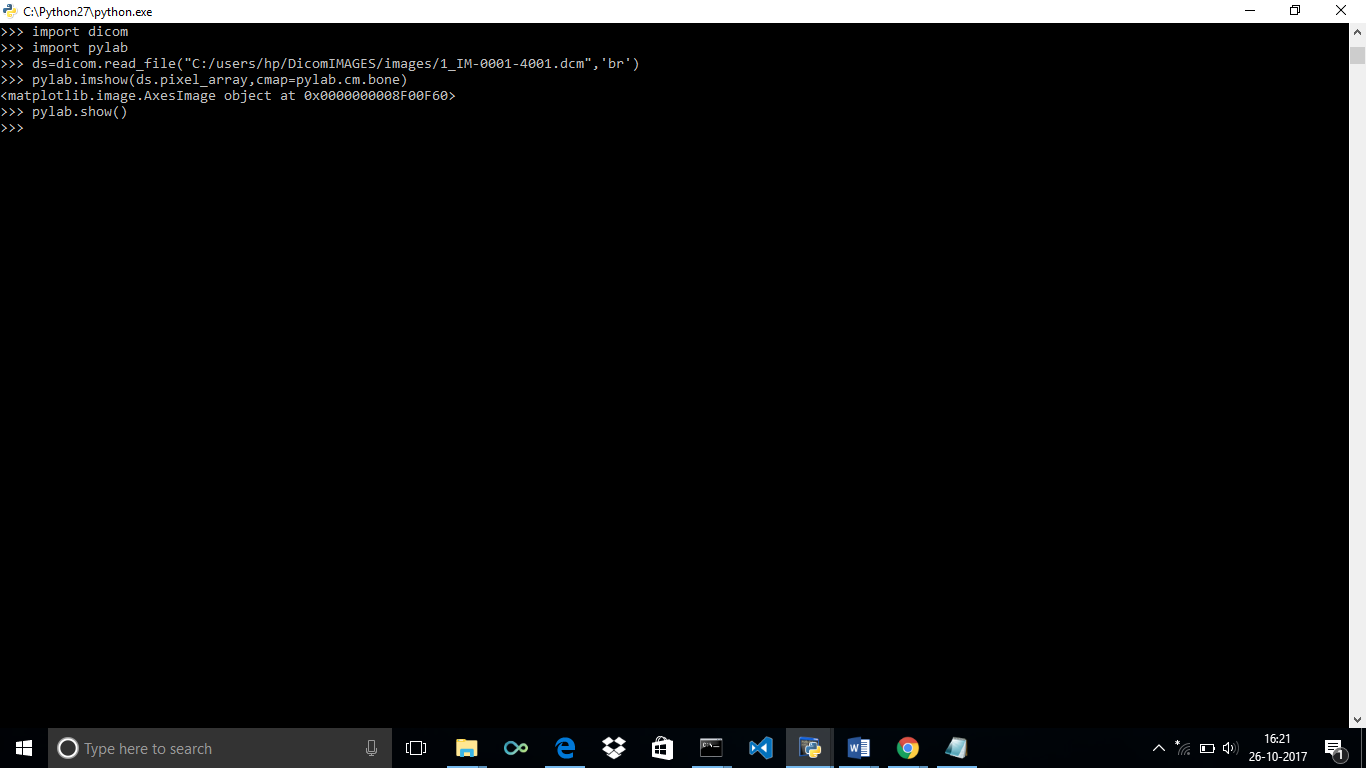


**ERROR**

**PART-3**

**READING OF IMAGES-**

* Download the zip file and extract it.
* Copy all the images in a folder (which are currently in different folders 1,2,3…)
* Go to python.exe and add the following code( or create a file and run python filename.py)



**CODE**-

>>> import dicom

>>> import pylab

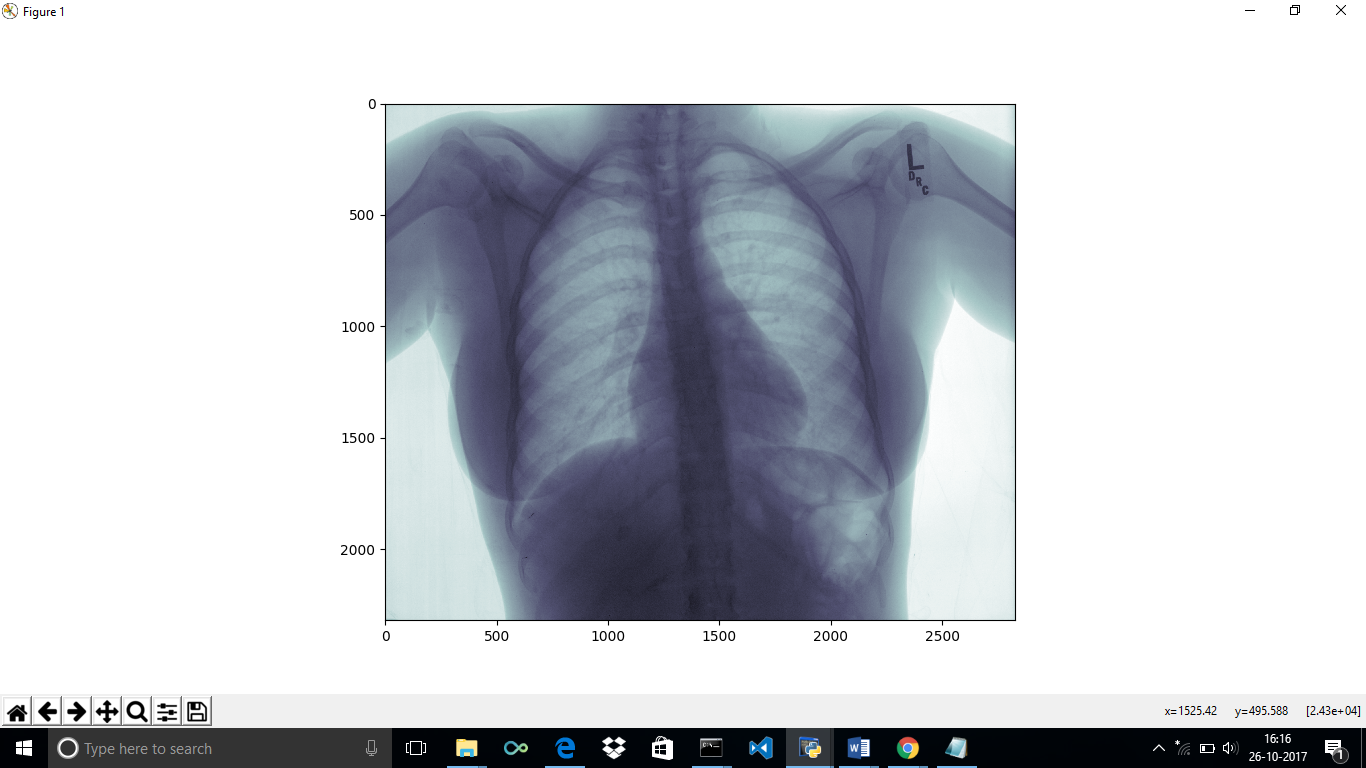
>>> ds=dicom.read\_file("C:/users/hp/DicomIMAGES/images/1\_IM-0001-4001.dcm",'br')

>>> pylab.imshow(ds.pixel\_array,cmap=pylab.cm.bone)

<matplotlib.image.AxesImage object at 0x0000000008F00F60>

>>> pylab.show()

The dcm image is read using python.

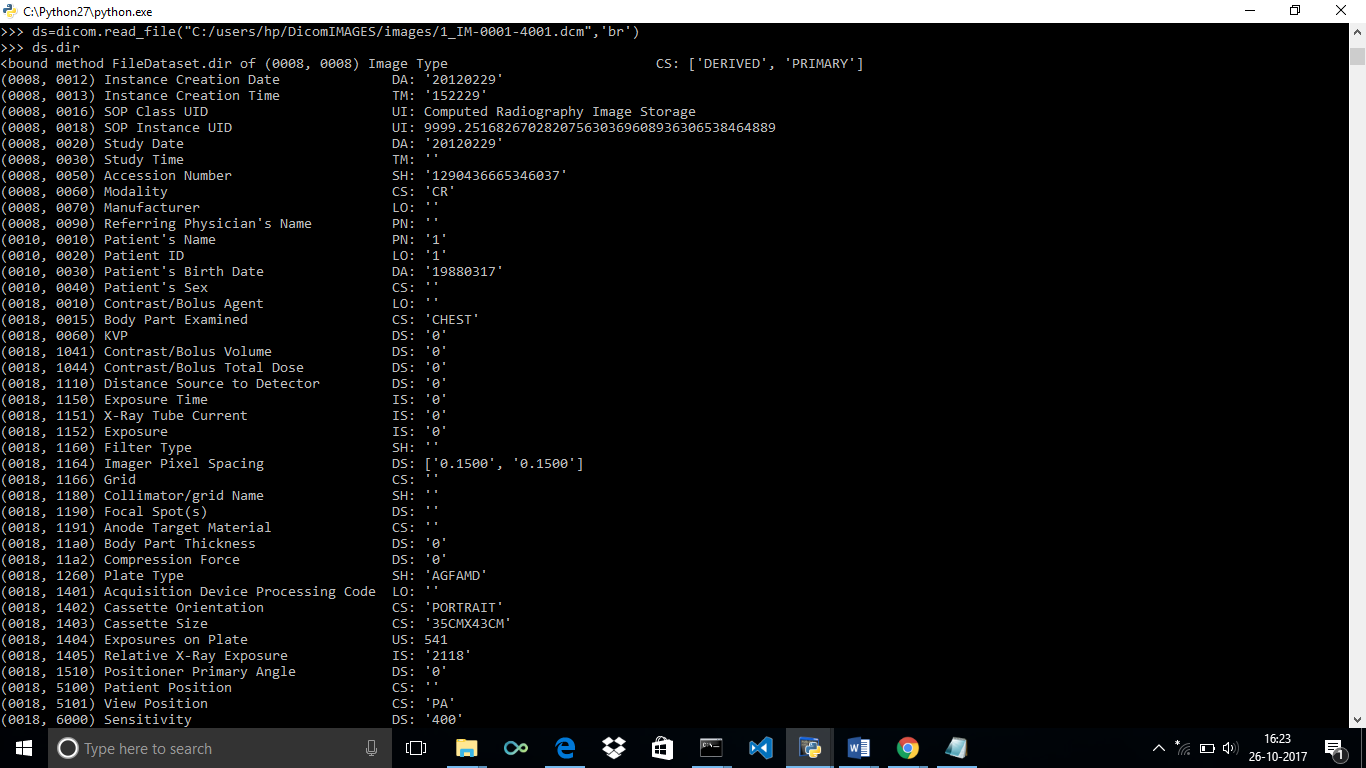


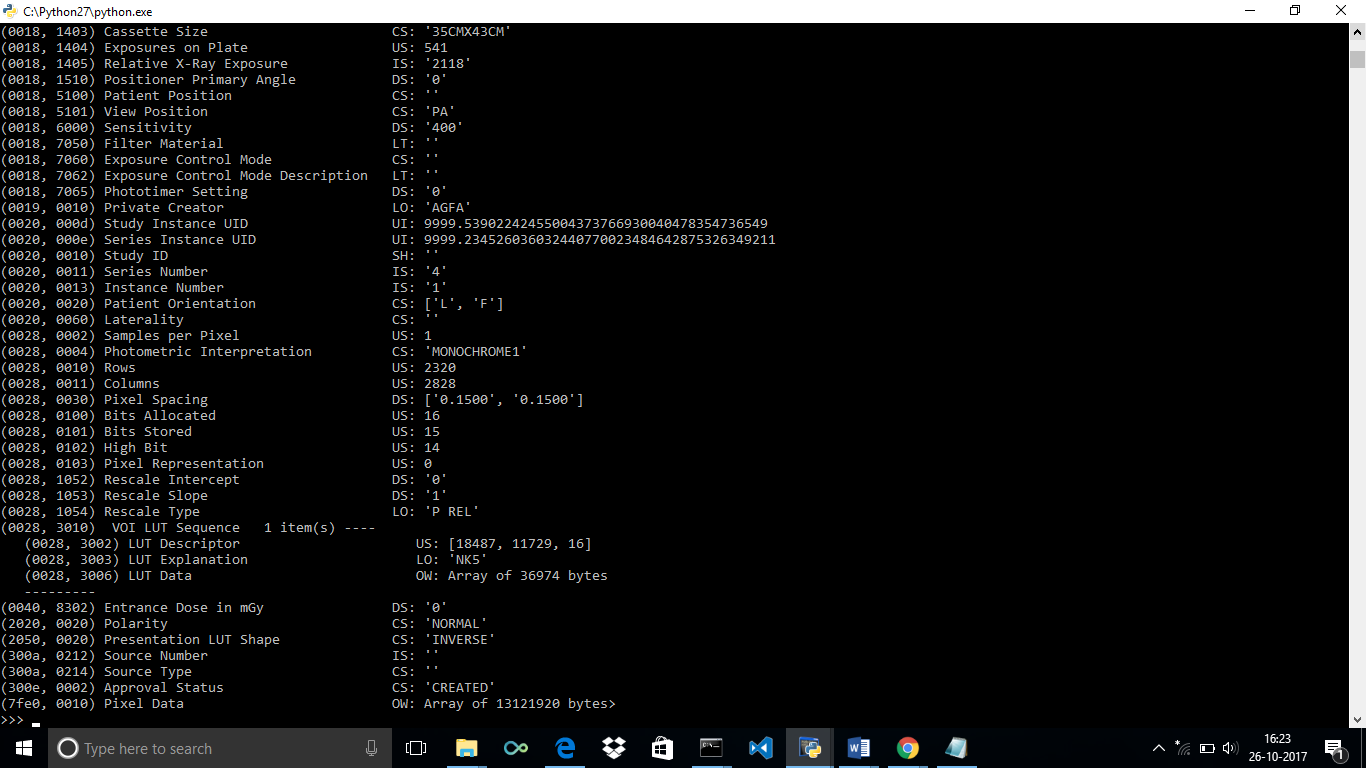
Similarly, we can read all the images.

Reading the data –

**>>> ds=dicom.read\_file("C:/users/hp/DicomIMAGES/images/1\_IM-0001-4001.dcm",'br')**

**>>> ds.dir**





Similarly we can read all the images except those in which error is coming(6).

**Thanks**

**Vithika Gupta**