EDA

Análisis exploratorio de los Datos (EDA) es entender los datasets sumarizando sus principales características y usualmente presentándolas visualmente.

In [ ]:

```
#folder_path = Params.path_dcm
folder_path = "./DCM/"

ds_list=[]
for p in os.listdir(folder_path):
    for q in os.listdir(os.path.join(folder_path, p)):
        dcm_path = os.path.join(folder_path, p, q)
        images_path = os.listdir(dcm_path)
        for n, image in enumerate(images_path):
            ds = pydicom.read_file(os.path.join(dcm_path, image), force=True)
            ds_list.append(ds)
```

In [ ]:

```
plt.imshow(ds_list[-10].pixel_array)
```

In [ ]:

```
print(ds_list[1].pixel_array.shape)
```

In [ ]:

```
len(ds_list)
```

In [ ]:

```
len(info_image['shape'])
```

In [ ]:

```
info_image = {}
info_image['shape'] = []
info_image['mean_pixel'] = []
info_image['min_pixel'] = []
info_image['max_pixel'] = []
info_image['std_pixel'] = []

for i in ds_list:
    info_image['shape'].append(i.pixel_array.shape)
    info_image['mean_pixel'].append(i.pixel_array.mean())
    info_image['min_pixel'].append(i.pixel_array.min())
    info_image['max_pixel'].append(i.pixel_array.max())
    info_image['std_pixel'].append(i.pixel_array.std())
```

In [ ]:

```
info_imagesDF=pd.DataFrame(info_image)
print(info_imagesDF)
```

In [ ]:

```
info_imagesDF['shape'].value_counts()
```

In [ ]:

```
info_imagesDF['shape'] = info_imagesDF['shape'].astype('str')
info_imagesDF = info_imagesDF[info_imagesDF['shape']=="(512, 512)"]
print(info_imagesDF)
```

In [ ]:

```
print(info_imagesDF.describe())
```

In [ ]:

```
info_imagesDF.hist()
```