IMAGE VISUALIZATION

April 1, 2019

```
In [1]: import pandas as pd
        import os
        import seaborn as sb
        import matplotlib.pyplot as plt
        from PIL import Image
        import numpy as np

In [2]: train_path = 'dataset/train/'
        test_path = 'dataset/test/'

In [3]: train = os.listdir(train_path)
        test = os.listdir(test_path)
```

0.1 Took very few image datas to my CPU to visualize the datas

```
In [4]: print("Total no. of train images: ",len(train_path))
        print("Total no. of test images: ",len(test_path))

Total no. of train images:  14
Total no. of test images:  13
```

0.2 Given CSV File

```
In [5]: labels = pd.read_csv(f'train_labels.csv')
```

0.3 Total number of training images

```
In [6]: labels.shape

Out[6]: (220025, 2)

In [7]: print(train[0:5])
        print(labels.loc[0:5])

['19960cfe73fd1cdea6ea8b095daf127cf97f606d.tif', '4bb590dd2da8fc0ef50b1f5b6e0ff76a4c16022d.tif
                                id  label
0  f38a6374c348f90b587e046aac6079959adf3835      0
1  c18f2d887b7ae4f6742ee445113fa1aef383ed77      1
```
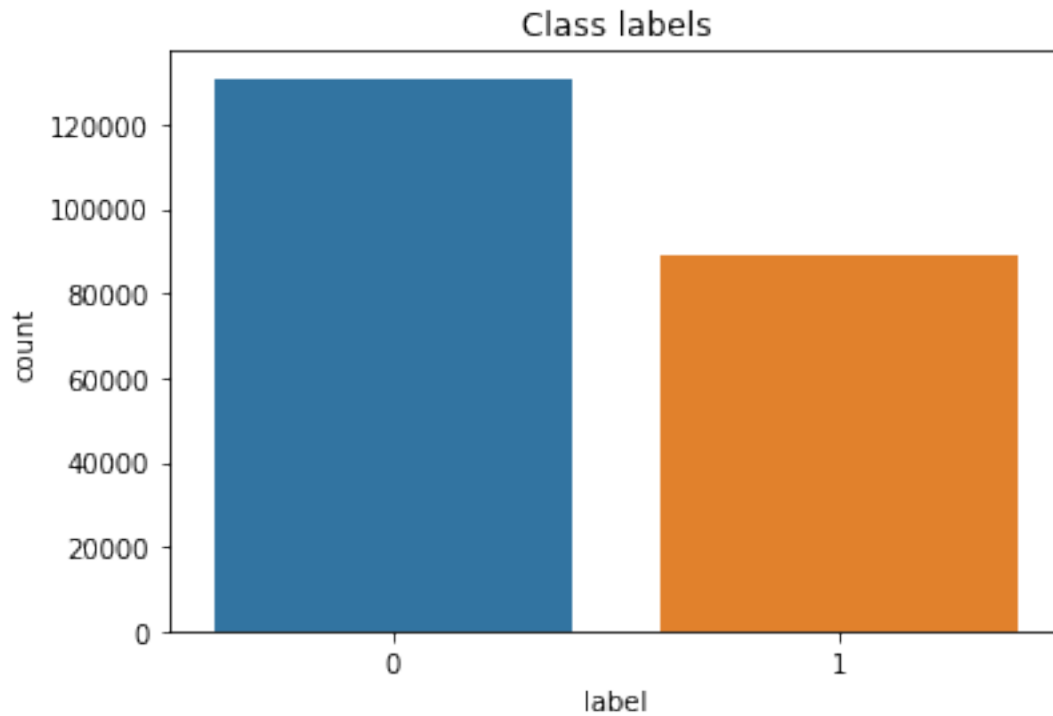
```
2   755db6279dae599ebb4d39a9123cce439965282d        0
3   bc3f0c64fb968ff4a8bd33af6971ecae77c75e08        0
4   068aba587a4950175d04c680d38943fd488d6a9d        0
5   acfe80838488fae3c89bd21ade75be5c34e66be7        0
```

```
In [8]: sb.countplot('label',data = labels).set_title('Class labels')
```

```
Out[8]: Text(0.5, 1.0, 'Class labels')
```



```
In [9]: fig = plt.figure(figsize=(25, 4))
        # display 20 images
        for idx, img in enumerate(np.random.choice(train, 20)):
            ax = fig.add_subplot(2, 20//2, idx+1, xticks=[], yticks=[])
            im = plt.imread(f'{train_path}' + img)
            lab = labels.loc[labels['id'] == img.split('.')[0], 'label'].values[0]
            ax.set_title(f'Label: {lab}')
            ax.set_xlabel(im.shape[0])
            ax.set_ylabel(im.shape[1])
            plt.imshow(im)
```