

Lesson 11 (notMNIST Dataset) The file `notMNIST.train_100.npz` contains grey scale images of letters.

- (a) How many images are in the data set? What is the resolution of the images in pixels?
- (c) Select a random image and display the image using the commands below.
- (d) Print the label name corresponding to the image above. You may find the commands below useful.
- (b) Is the data set “balanced?” That is, are the different classes of letters well represented in the data set? (See commands below.)
- (e) Choose an image that you think would be easy to classify. Show the image in your notebook.
- (f) Choose an image that you think would be hard to classify. Show the image in your notebook.
- (g) How many different “representations” of the letter H can you find in the data set? Show the images in your notebook.

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline

notMNIST = np.load('../Data/notMNIST/notMNIST_train_100.npz')
notMNIST.keys()

plt.imshow(your_image, cmap='Greys_r')
plt.show()

label_names = np.array(['A', 'B', 'C', 'D', 'E', 'F', 'G', 'H', 'I', 'J'])
pd.Series(label_names[labels]).value_counts()
```
