

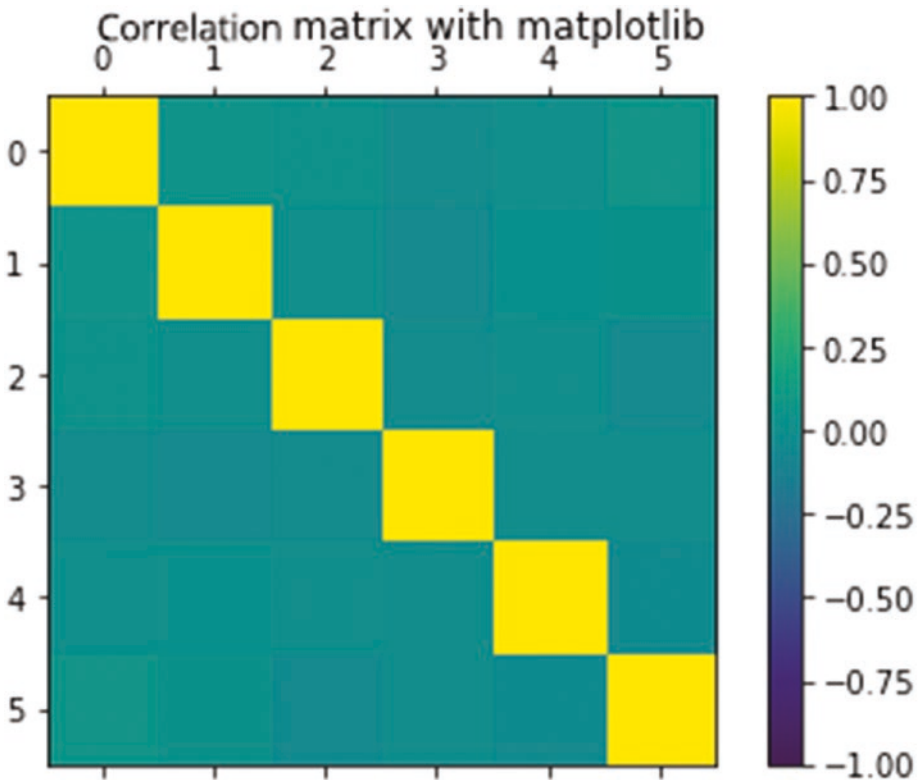
*Figure 12-12. Pairwise scatter plot with seaborn*

## Correlation Matrix Plots

Again, correlation shows how much relationship exists between two variables. By plotting the correlation matrix, we get a visual representation of which variables in the dataset are highly correlated. Remember that parametric machine learning methods such as logistic and linear regression can take a performance hit when variables are

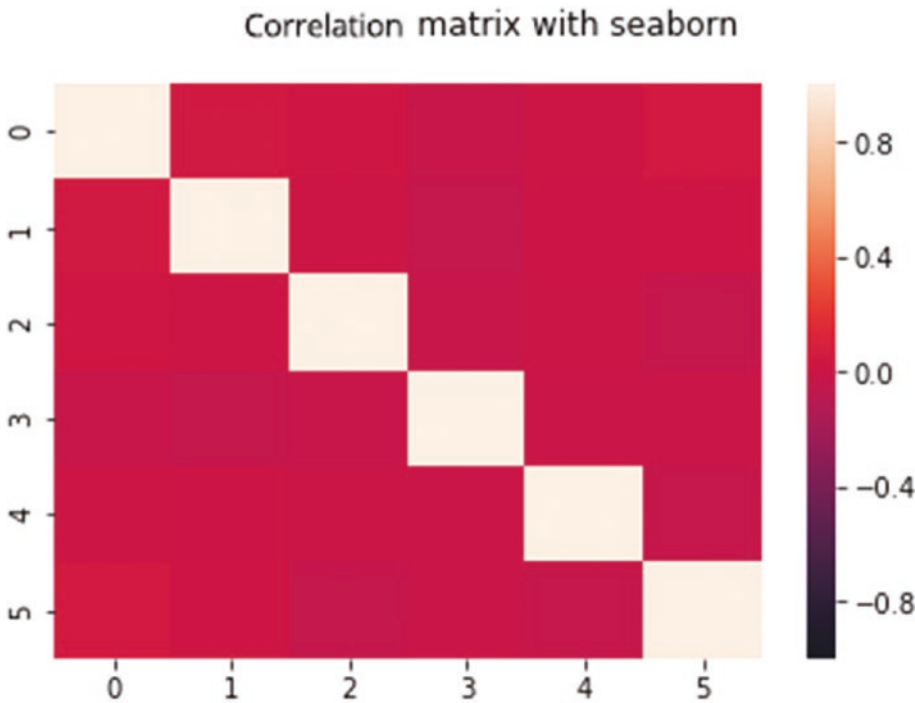
highly correlated. Also, in practice, the correlation values that are greater than  $-0.7$  or  $0.7$  are for the most part highly correlated. The outputs with matplotlib and seaborn are shown in Figure 12-13 and Figure 12-14, respectively.

```
# create the dataset
data = np.random.random([1000,6])
# plot covariance matrix using the Matplotlib matshow function
fig = plt.figure()
ax = fig.add_subplot(111)
my_plot = ax.matshow(pd.DataFrame(data).corr(), vmin=-1, vmax=1)
fig.colorbar(my_plot)
plt.show()
```



**Figure 12-13.** *Correlation matrix with Matplotlib*

```
# plot covariance matrix with seaborn heatmap function
sns.heatmap(pd.DataFrame(data).corr(), vmin=-1, vmax=1)
plt.show()
```



**Figure 12-14.** *Correlation matrix with seaborn*

## Images

Matplotlib is also used to visualize images. This process is utilized when visualizing a dataset of image pixels. You will observe that image data is stored in the computer as an array of pixel intensity values ranging from 0 to 255 across three bands for colored images.

```
img = plt.imread('nigeria-coat-of-arms.png')
# check image dimension
img.shape
'Output': (232, 240, 3)
```