

Figure 12-4. Barplot with seaborn

Histogram/Density Plots

Histogram and density plots are essential for examining the statistical distribution of a variable. For a simple histogram, we'll create a set of 100,000 points from the normal distribution. The outputs with matplotlib and seaborn are shown in Figure 12-5 and Figure 12-6, respectively.

```
# create 100000 data points from the normal distributions
data = np.random.randn(100000)
# create a histogram plot
plt.hist(data)
plt.show()
# create a density plot using seaborn
my_fig = sns.distplot(data, hist=False)
plt.show()
```

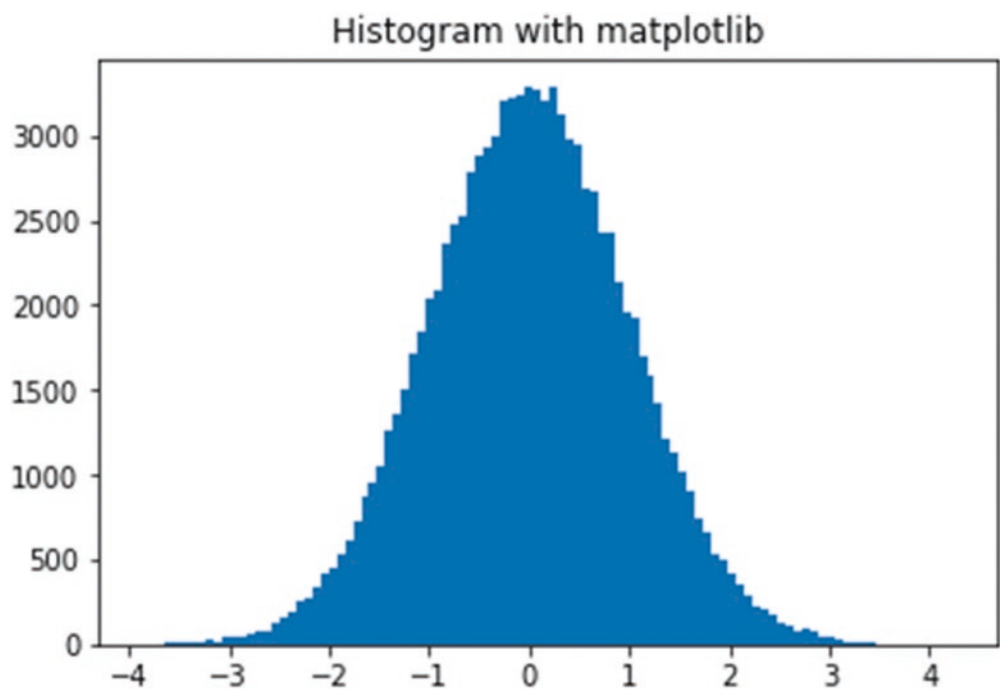


Figure 12-5. *Histogram with Matplotlib*

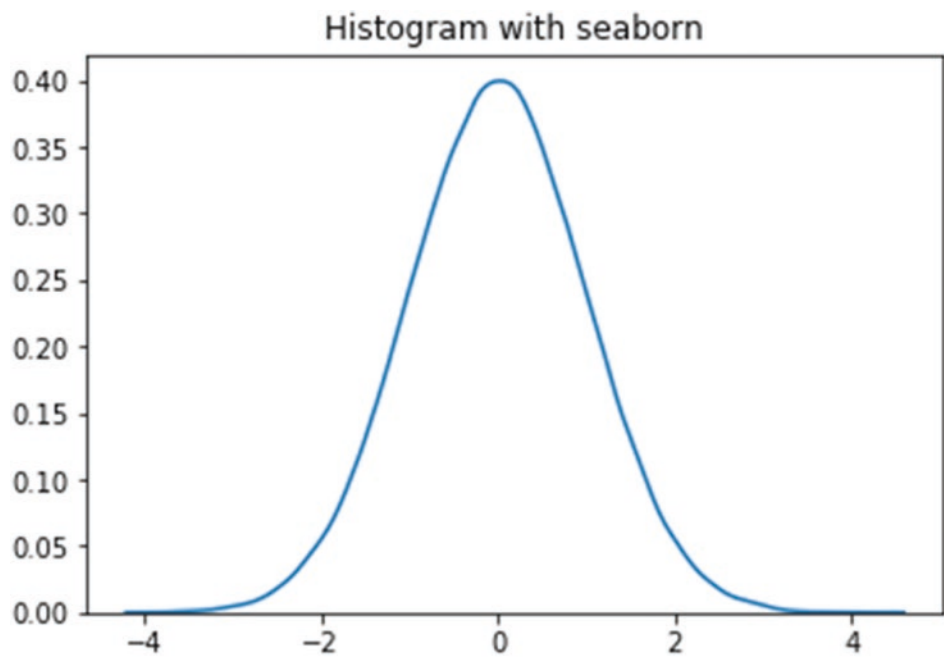


Figure 12-6. *Histogram with seaborn*

Box and Whisker Plots

Boxplot, also popularly called box and whisker plot, is another useful visualization technique for gaining insights into the underlying data distribution. The boxplot draws a box with the upper line representing the 75th percentile and the lower line the 25th percentile. A line is drawn at the center of the box indicating the 50th percentile or median value. The whiskers at both ends give an estimation of the spread or variance of the data values. The dots at the tail end of the whiskers represent possible outlier values. The outputs with matplotlib and seaborn are shown in Figure 12-7 and Figure 12-8, respectively.

```
# create data points
data = np.random.randn(1000)
## box plot with matplotlib
plt.boxplot(data)
plt.show()
## box plot with seaborn
sns.boxplot(data)
plt.show()
```

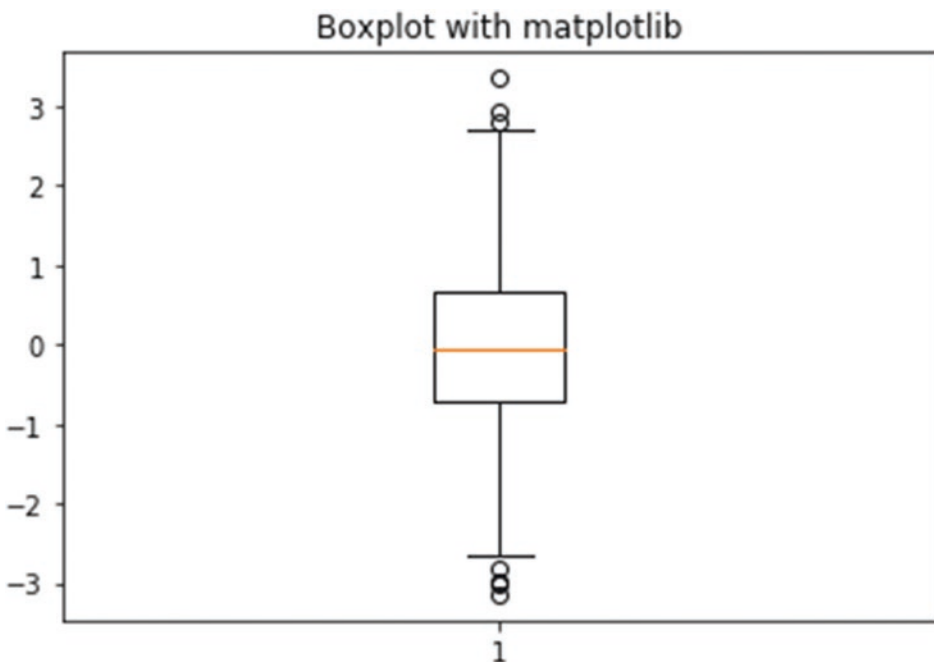


Figure 12-7. Boxplot with Matplotlib