EXPERIMENT NO: 4

OUTLIER DETECTION FOR GIVEN DATASET

Aim:

To write a Python program to understand and detect outliers in a given dataset using IQR method.

Algorithm:

1. Import NumPy and generate a random array of integers.
2. Calculate mean and percentiles (Q1, Q2, Q3) using np.percentile.
3. Define a function outDetection() to compute IQR, LR, and UR.
4. Use LR and UR to identify and filter out outliers from the array.
5. Visualize the original and filtered data using Seaborn's displot.
6. Reapply outlier detection to confirm no further outliers remain.

Program:

[1]:
```
import numpy as np
array=np.random.randint(1,100,16)
array
```

[1]:
```
array([ 6, 78, 22, 85,  9, 59, 92,  3, 32, 21,  8, 47, 20, 85, 39, 74],
      dtype=int32)
```

[2]:
```
array.mean()
```

[2]:
```
np.float64(49.5625)
```

[3]:
```
np.percentile(array,25)
```

[3]:
```
np.float64(28.0)
```

[4]:
```
np.percentile(array,50)
```

[4]:
```
np.float64(54.5)
```

[5]:
```
np.percentile(array,75)
```

[5]:
```
np.float64(77.25)
```

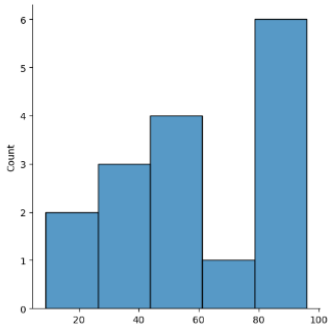
[6]:
```
np.percentile(array,100)
```

[6]:
```
np.float64(86.0)
```

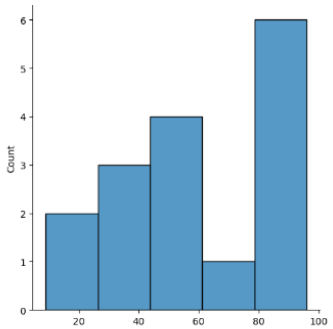
[7]:
```
def outdetect(array):
    sorted(array)
    q1,q3=np.percentile(array,[25,75])
    IQR=q3-q1
    lr=q1-(1.5*IQR)
    ur=q3+(1.5*IQR)
    return lr,ur
lr,ur=outdetect(array)
lr,ur
```

[7]:
```
(np.float64(-45.875), np.float64(151.125))
```

```
[8]: import seaborn as sns
     %matplotlib inline
     import matplotlib.pyplot as plt
     sns.displot(array)
     plt.show()
```



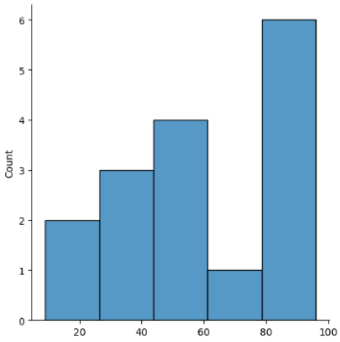

```
[9]: sns.displot(array)
     plt.show()
```




```
[10]: newarray=array[(array>lr)&(array<ur)]
      newarray
```

```
[10]: array([47, 65,  9, 91, 96, 87, 86, 11, 54, 38, 56, 40, 86, 34, 95, 50],
            dtype=int32)
```

```
[11]: sns.displot(newarray)
      plt.show()
```

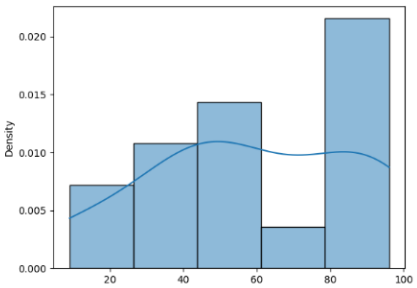



```
lr1,ur1=outdetect(newarray) lr1,ur1
```

```
[14]: final_array=newarray[(newarray>lr)&(newarray<ur)]
      final_array
```

```
[14]: array([47, 65,  9, 91, 96, 87, 86, 11, 54, 38, 56, 40, 86, 34, 95, 50],
            dtype=int32)
```

```
[15]: sns.histplot(final_array, kde=True, stat="density", bins="auto")
      plt.show()
```



Result:

Thus, the outliers have been detected and the experiment executed successfully.