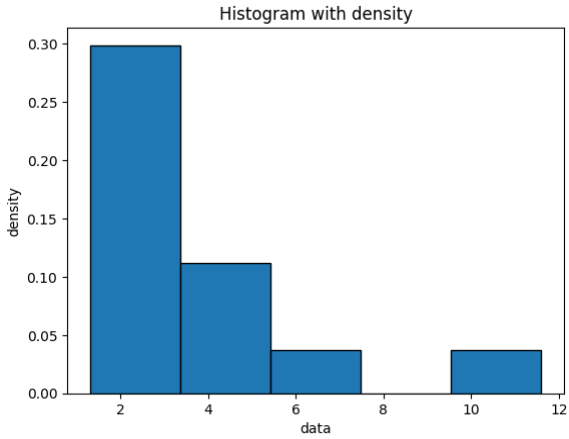
Assignment - 2

B.Rithwik
2303A52330
Batch - 35

#QUESTION-1

```
import numpy as np
import matplotlib.pyplot as plt

#Given dataset
x=[1.3,1.5,2.6,2.6,3.2,3.9,4.2,3.7,3.10,3.0,11,6.7,1.9]
#number of bins
num_bin=5
#compute the histogram
counts,bin_edges=np.histogram(x,bins=num_bin,density=True)
#calculate the width of eachbin
bin_width=np.diff(bin_edges)
#calculate the density of eachbin
density=counts*bin_width
#print the binedges ,counts and density
print("bin edges:",bin_edges)
print("counts(density values):",counts)
print("density values adjusted for bin width:",density)
#plot the histogram
plt.hist(x,bins=num_bin,density=True,edgecolor='black')
plt.xlabel('data')
plt.ylabel('density')
plt.title('Histogram with density')
plt.show()
```



#OUTPUT

```
bin edges: [ 1.3   3.36  5.42  7.48  9.54 11.6 ]
counts(density values): [0.2987304 0.1120239 0.0373413 0.        0.0373413]
density values adjusted for bin width: [0.61538462 0.23076923 0.07692308 0.         0.07692308]
```

#QUESTION-2

```
import numpy as np
import pandas as pd
from scipy import stats
import matplotlib.pyplot as plt
dependent_variables=[12,13,56,23,41,28,51]
indedent_variables=[51,63,37,88,33,72,81]
#mean
mean_x=np.mean(indedent_variables)
mean_y=np.mean(dependent_variables)
print("mean of x:",mean_x)
print("mean of y:",mean_y)
#median
median_x=np.median(indedent_variables)
median_y=np.median(dependent_variables)
print("median of x:",median_x)
print("median of y:",median_y)
```

```
#mode
mode_x=stats.mode(indedent_variables)
mode_y=stats.mode(dependent_variables)
print("mode of x:",mode_x)
print("mode of y:",mode_y)
#variance
var_x=np.var(indedent_variables)
var_y=np.var(dependent_variables)
print("variance of x:",var_x)
print("variance of y:",var_y)
#standard deviation
std_x=np.std(indedent_variables)
std_y=np.std(dependent_variables)
print("standard deviation of x:",std_x)
print("standard deviation of y:",std_y)
#range
range_x=np.max(indedent_variables)-np.min(indedent_variables)
range_y=np.max(dependent_variables)-np.min(dependent_variables)
print("range of x:",range_x)
print("range of y:",range_y)
#interquartile(IQR)
q1_x=np.percentile(indedent_variables,25)
q3_x=np.percentile(indedent_variables,75)
iqr_x=q3_x-q1_x
print("interquartile range of x:",iqr_x)
q1_y=np.percentile(dependent_variables,25)
q3_y=np.percentile(dependent_variables,75)
iqr_y=q3_y-q1_y
print("interquartile range of y:",iqr_y)
#skewness
skew_x=stats.skew(indedent_variables)
skew_y=stats.skew(dependent_variables)
print("skewness of x:",skew_x)
print("skewness of y:",skew_y)
#kurtosis
kurt_x=stats.kurtosis(indedent_variables)
kurt_y=stats.kurtosis(dependent_variables)
print("kurtosis of x:",kurt_x)
print("kurtosis of y:",kurt_y)
```

#OUTPUT

```
mean of x: 60.714285714285715
mean of y: 32.0
median of x: 63.0
median of y: 28.0
mode of x: ModeResult(mode=33, count=1)
mode of y: ModeResult(mode=12, count=1)
variance of x: 387.6326530612245
variance of y: 268.0
standard deviation of x: 19.68838878784205
standard deviation of y: 16.3707055437449
range of x: 55 range of y: 44
interquartile range of x: 32.5
interquartile range of y: 28.0
skewness of x: -0.10160487226954186
skewness of y: 0.18755223867066048
```

kurtosis of x: -1.4265295827972655
kurtosis of y: -1.47684021258314

#QUESTION-3

```
import numpy as np
import pandas as pd
from scipy import stats
import matplotlib.pyplot as plt
x=[1.3,1.5,2.6,2.6,3.2,3.9,4.2,3.7,3.10,3.0,11,6.7,1.9]
#mean
mean_x=np.mean(x)
print("mean of x:",mean_x)
#median
median_x=np.median(x)
print("median of x:",median_x)
#mode
from scipy import stats
mode_x=stats.mode(x)
print("mode of x:",mode_x)
#variance
var_x=np.var(x)
print("variance of x:",var_x)
#standard deviation
std_x=np.std(x)
print("standard deviation of x:",std_x)
#range
range_x=np.max(x)-np.min(x)
print("range of x:",range_x)
#interquatile(IQR)
q1_x=np.percentile(x,25)
q3_x=np.percentile(x,75)
iqr_x=q3_x-q1_x
print("interquartile range of x:",iqr_x)
#skewness
skew_x=stats.skew(x)
print("skewness of x:",skew_x)
#kurtosis
kurt_x=stats.kurtosis(x)
print("kurtosis of x:",kurt_x)
```

#OUTPUT

mean of x: 3.746153846153846
median of x: 3.1
mode of x: ModeResult(mode=2.6, count=2)
variance of x: 6.147100591715977
standard deviation of x: 2.4793347074802097
range of x: 9.7
interquartile range of x: 1.2999999999999998
skewness of x: 1.8747042061161974
kurtosis of x: 2.9478483416827777