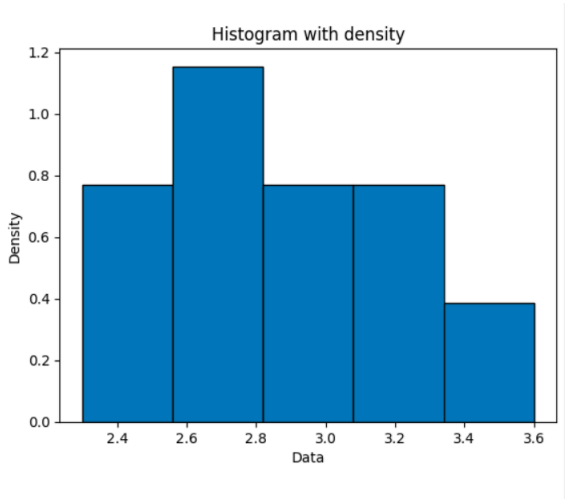
Assignment - 1

B.Rithwik
2303A52330
Batch - 35

Question 1:

1. Given the following dataset:
X={2.3,2.5,3.6,2.8,3.1,2.9,3.2,2.7,2.8,3.0}.
Compute the density of the data using a histogram with 5 bins.

```
import numpy as np
import matplotlib.pyplot as plt
#given dataset
X=[2.3,2.5,3.6,2.8,3.1,2.9,3.2,2.7,2.8,3.0]
#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 bin edges,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: [2.3 2.56 2.82 3.08 3.34 3.6 ]
Counts(density values): [0.76923077 1.15384615 0.76923077
0.76923077 0.38461538]
Density values adjusted for bin width: [0.2 0.3 0.2 0.2 0.1]
```

Question 2:

(dependent variables) (independent variables)

4 5
5 6
8 3
2 8
4 3
2 7
5 8

2. Find Mean, Median, Mode , Variance , Standard deviation , Range ,
Interquartile Range (IQR), Skewness ,Kurtosis of x and y .

```
import numpy as np
import matplotlib.pyplot as plt
dependent_variable=[4,5,8,2,4,2,5]
independent_variable=[5,6,3,8,3,7,8]
#mean
mean_x=np.mean(independent_variable)
mean_y=np.mean(dependent_variable)
print("mean of x:",mean_x)
print("mean of y:",mean_y)
#median
median_x=np.median(independent_variable)
median_y=np.median(dependent_variable)
print("median of x:",median_x)
print("median of y:",median_y)
#mode
from scipy import stats
mode_x=stats.mode(independent_variable)
mode_y=stats.mode(dependent_variable)
print("mode of x:",mode_x)
print("mode of y:",mode_y)
#variance
var_x=np.var(independent_variable)
var_y=np.var(dependent_variable)print("variance of x:",var_x)
print("variance of y:",var_y)
#standard deviation
std_x=np.std(independent_variable)
std_y=np.std(dependent_variable)
print("standard deviation of x:",std_x)
print("standard deviation of y:",std_y)
#Range
range_x=np.max(independent_variable)-np.min(independent_variable)
range_y=np.max(dependent_variable)-np.min(dependent_variable)
print("range of x:",range_x)
print("range of y:",range_y)
#interQuartile(IQR)
q1_x=np.percentile(independent_variable,25)
q3_x=np.percentile(independent_variable,75)
iqr_x=q3_x-q1_x
print("interquartile range of x:",iqr_x)
q1_y=np.percentile(dependent_variable,25)
q3_y=np.percentile(dependent_variable,75)
iqr_y=q3_y-q1_y
print("interquartile range of y:",iqr_y)
#Skewness
skew_x=stats.skew(independent_variable)
skew_y=stats.skew(dependent_variable)
print("skewness of x:",skew_x)
```

```
print("skewness of y:",skew_y)
#Kurtosis
kurt_x=stats.kurtosis(independent_variable)
kurt_y=stats.kurtosis(dependent_variable)
print("kurtosis of x:",kurt_x)
print("kurtosis of y:",kurt_y)
#OUTPUT:
```
mean of x: 5.714285714285714
mean of y: 4.285714285714286
median of x: 6.0
median of y: 4.0
mode of x: ModeResult(mode=3, count=2)
mode of y: ModeResult(mode=2, count=2)
variance of x: 3.918367346938776
variance of y: 3.6326530612244894
standard deviation of x: 1.979486637221574
standard deviation of y: 1.9059520091609048
range of x: 5
range of y: 6
range of x: 3.5interquartile range of y: 2.0
skewness of x: -0.26386711521557127
skewness of y: 0.5785710902698393
kurtosis of x: -1.4541015625000002
kurtosis of y: -0.34282287589950755

Question 3:
From the above data Find Mean, Median, Mode, Variance, Standard deviation,
Range , Interquartile Range (IQR), Skewness ,Kurtosis of each feature.

```
import numpy as np
import matplotlib.pyplot
x=[2.3,2.5,3.6,2.8,3.1,2.9,3.2,2.7,2.8,3.0]
#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: 2.89
median of x: 2.8499999999999996
mode of x: ModeResult(mode=2.8, count=2)
variance of x: 0.12090000000000005
standard deviation of x: 0.3477067730142743
range of x: 1.3000000000000003
interquartile range of x: 0.3500000000000001
skewness of x: 0.2980178611012123
kurtosis of x: -0.18689850931906538
```