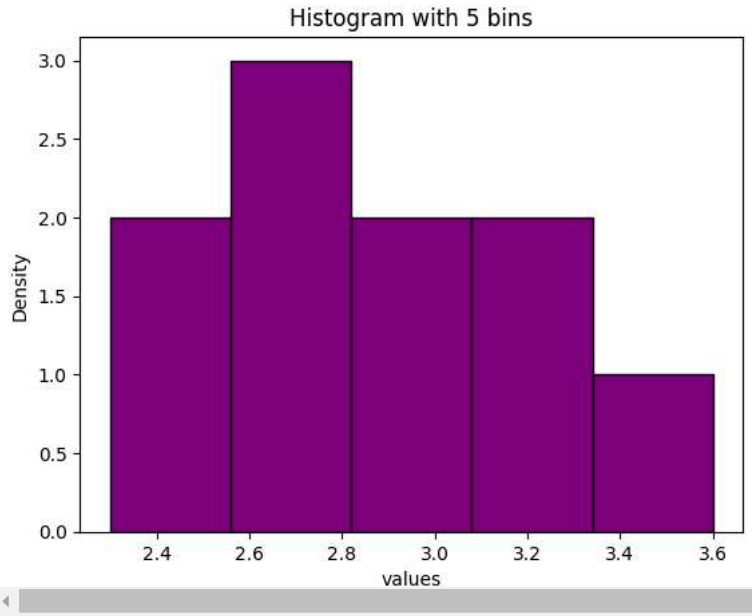
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt


db=pd.read_csv('/content/housing.csv')
db
```

	longitude	latitude	housing_median_age	total_rooms	total_bedrooms	population	households	median_income	median_house_value	
0	-122.23	37.88	41.0	880.0	129.0	322.0	126.0	8.3252	452600.0	
1	-122.22	37.86	21.0	7099.0	1106.0	2401.0	1138.0	8.3014	358500.0	
2	-122.24	37.85	52.0	1467.0	190.0	496.0	177.0	7.2574	352100.0	
3	-122.25	37.85	52.0	1274.0	235.0	558.0	219.0	5.6431	341300.0	
4	-122.25	37.85	52.0	1627.0	280.0	565.0	259.0	3.8462	342200.0	
...	...	...	...	...	...	...	...	...	...	
20635	-121.09	39.48	25.0	1665.0	374.0	845.0	330.0	1.5603	78100.0	
20636	-121.21	39.49	18.0	697.0	150.0	356.0	114.0	2.5568	77100.0	
20637	-121.22	39.43	17.0	2254.0	485.0	1007.0	433.0	1.7000	92300.0	
20638	-121.32	39.43	18.0	1860.0	409.0	741.0	349.0	1.8672	84700.0	
20639	-121.24	39.37	16.0	2785.0	616.0	1387.0	530.0	2.3886	89400.0	

20640 rows × 10 columns

```
X=[2.3, 2.5, 3.6, 2.8, 3.1, 2.9, 3.2, 2.7, 2.8, 3.0]
plt.hist(X,bins=5,color="purple",edgecolor="black")
plt.ylabel('Density')
plt.xlabel('values')
plt.title('Histogram with 5 bins')
plt.show()
```




```
import numpy as np
from scipy import stats

x = [4, 5, 8, 2, 4, 2, 5]
y = [5, 6, 3, 8, 3, 7, 8]

def mean(data):
    return np.mean(data)

def median(data):
    return np.median(data)

def mode(data):
```

```
    return stats.mode(data)

def variance(data):
    return np.var(data)

def stddev(data):
    return np.std(data)

def range(data):
    return max(data) - min(data)

def iqr(data):
    q1, q3 = np.percentile(data, [25, 75])
    return q3 - q1

def skewness(data):
    return stats.skew(data)

def kurtosis(data):
    return stats.kurtosis(data)


print("Statistics for x:")
print("Mean:", mean(x))
print("Median:", median(x))
print("Mode:", mode(x))
print("Variance:", variance(x))
print("Standard deviation:", stddev(x))
print("Range:", range(x))
print("Interquartile range (IQR):", iqr(x))
print("Skewness:", skewness(x))
print("Kurtosis:", kurtosis(x))


print("\nStatistics for y:")
print("Mean:", mean(y))
print("Median:", median(y))
print("Mode:", mode(y))
print("Variance:", variance(y))
print("Standard deviation:", stddev(y))
print("Range:", range(y))
print("Interquartile range (IQR):", iqr(y))
print("Skewness:", skewness(y))
print("Kurtosis:", kurtosis(y))
```

```
Statistics for x:
Mean: 4.285714285714286
Median: 4.0
Mode: ModeResult(mode=2, count=2)
Variance: 3.6326530612244894
Standard deviation: 1.9059520091609048
Range: 6
Interquartile range (IQR): 2.0
Skewness: 0.5785710902698393
Kurtosis: -0.34282287589950755

Statistics for y:
Mean: 5.714285714285714
Median: 6.0
Mode: ModeResult(mode=3, count=2)
Variance: 3.918367346938776
Standard deviation: 1.979486637221574
Range: 5
Interquartile range (IQR): 3.5
Skewness: -0.26386711521557127
Kurtosis: -1.4541015625000002
```

```
import pandas as pd
from scipy.stats import skew, kurtosis
Y = [1, 2, 3, 4, 5]
X_independent = [5, 4, 3, 2, 1]
XY_df = pd.DataFrame({'Y': Y, 'X': X_independent})
XY_stats = XY_df.describe().T
XY_stats['IQR'] = XY_stats['75%'] - XY_stats['25%']
XY_stats['Skewness'] = XY_df.apply(lambda x: skew(x), axis=0)
XY_stats['Kurtosis'] = XY_df.apply(lambda x: kurtosis(x), axis=0)
print("\nStatistical measures for the features in the given table:")
print(XY_stats)
```

```
Statistical measures for the features in the given table:
   count  mean       std  min  25%  50%  75%  max  IQR  Skewness  Kurtosis
Y    5.0   3.0  1.581139  1.0  2.0  3.0  4.0  5.0  2.0       0.0      -1.3
X    5.0   3.0  1.581139  1.0  2.0  3.0  4.0  5.0  2.0       0.0      -1.3
```

Done By KODAM SHISHIR BHAGATH [2303A52164]