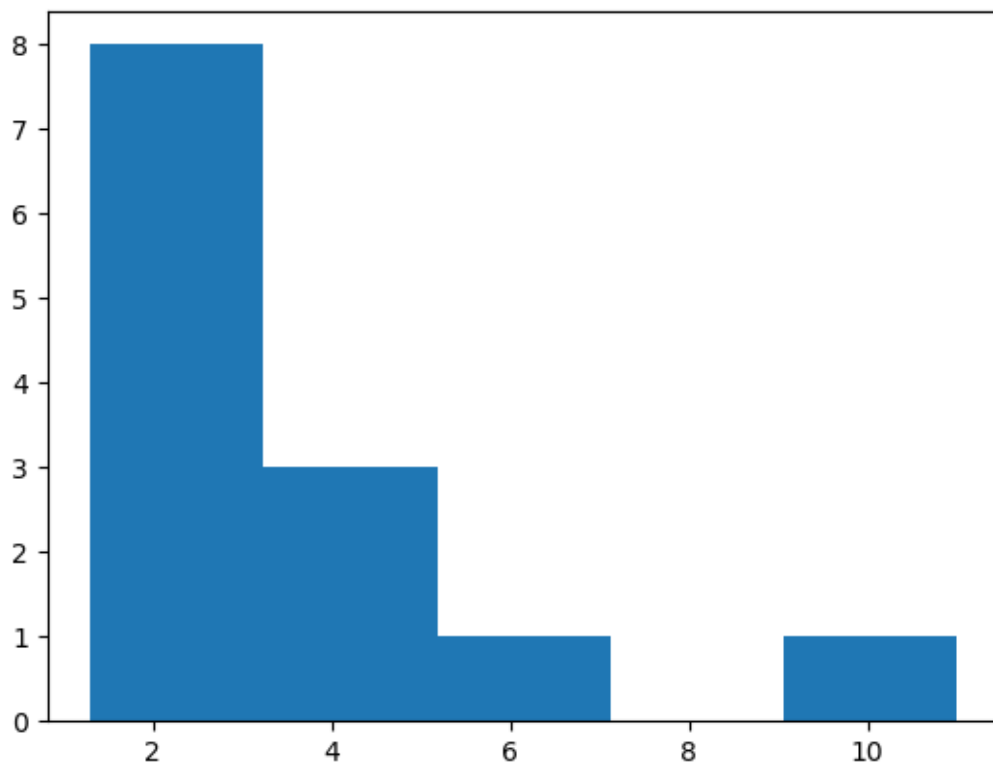


sta-2-a

August 17, 2024

```
[ ]: # 1) Given the following dataset:  $X=\{1.3,1.5,2.6,2.6,3.2,3.9,4.2,3.7,3.10,3.0,11,6.7,1.9\}$ . Compute the density of the data using a histogram with 5 bins.
```

```
import numpy as np
import matplotlib.pyplot as plt
x=np.array([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])
plt.hist(x, bins=5)
plt.show()
```



```
[ ]: # 2) Find Mean, Median, Mode , Variance , Standard deviation , Range ,  
      Interquartile Range (IQR), Skewness , Kurtosis of  $x$  and  $y$  .
```

```

import numpy as np
from scipy import stats as s
import matplotlib.pyplot as plt
x = [12,13,56,23,41,28,51]
y = [51,63,37,88,33,72,81]
print("Mean of x is:",np.mean(x))
print("Mean of y is:",np.mean(y))
print("Median of x is:",np.mean(x))
print("Median of y is:",np.mean(y))
print("Mode of x is :",s.mode(x))
print("Mode of y is :",s.mode(y))
print("Variance of x is :",np.var(x))
print("Variance of y is :",np.var(y))
print("Standard deviation of x is :",np.std(x))
print("Standard deviation of y is :",np.std(y))
print("Range of x is :",max(x)-min(x))
print("Range of y is :",max(y)-min(y))
print("Interquartile range of x is :",np.percentile(x,75)-np.percentile(x,25))
print("Interquartile range of y is :",np.percentile(y,75)-np.percentile(y,25))
print("Skewness of x is :",s.skew(x))
print("Skewness of y is :",s.skew(y))
print("Kurtosis of x is :",s.kurtosis(x))
print("Kurtosis of y is :",s.kurtosis(y))

```

```

Mean of x is: 32.0
Mean of y is: 60.57142857142857
Median of x is: 32.0
Median of y is: 60.57142857142857
Mode of x is : ModeResult(mode=12, count=1)
Mode of y is : ModeResult(mode=33, count=1)
Variance of x is : 268.0
Variance of y is : 387.10204081632645
Standard deviation of x is : 16.3707055437449
Standard deviation of y is : 19.67490891507065
Range of x is : 44
Range of y is : 55
Interquartile range of x is : 28.0
Interquartile range of y is : 32.5
Skewness of x is : 0.18755223867066048
Skewness of y is : -0.08021495675726419
Kurtosis of x is : -1.47684021258314
Kurtosis of y is : -1.4248824876045976

```

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

```

import pandas as pd

```

```

import numpy as np
from scipy import stats as s
import matplotlib.pyplot as plt
data=pd.read_csv('/content/sample_data/california_housing_test.csv')
print(data)
d=pd.DataFrame(data)
print(d)

print("1.Mean of Longitude is",np.mean(d['longitude']))
print("2.Mean of Latitude is",np.mean(d['latitude']))
print("3.Mean of Housing Median Age is",np.mean(d['housing_median_age']))
print("4.Mean of Total Rooms is",np.mean(d['total_rooms']))
print("5.Mean of Total Bedrooms is",np.mean(d['total_bedrooms']))
print("6.Mean of Population is",np.mean(d['population']))
print("7.Mean of Households is",np.mean(d['households']))
print("8.Mean of Median Income is",np.mean(d['median_income']))
print("9.Mean of Median House_value is",np.mean(d['median_house_value']))

print("1.Median of Longitude is = ",np.median(d['longitude']))
print("2.Median of Latitude is = ",np.median(d['latitude']))
print("3.Median of Housing Median Age is = ",np.median(d['housing_median_age']))
print("4.Median of Total Rooms is = ",np.median(d['total_rooms']))
print("5.Median of Total Bedrooms is = ",np.median(d['total_bedrooms']))
print("6.Median of population is = ",np.median(d['population']))
print("7.Median of Households is = ",np.median(d['households']))
print("8.Median of Median Income is = ",np.median(d['median_income']))
print("9.Median of Median House Value is = ",np.median(d['median_house_value']))

print("1.Mode of Longitude is = ",s.mode(d['longitude']))
print("2.Mode of Latitude is = ",s.mode(d['latitude']))
print("3.Mode of Housing Median Age is = ",s.mode(d['housing_median_age']))
print("4.Mode of Total Rooms is = ",s.mode(d['total_rooms']))
print("5.Mode of Total Bedrooms is = ",s.mode(d['total_bedrooms']))
print("6.Mode of Population is = ",s.mode(d['population']))
print("7.Mode of Households is = ",s.mode(d['households']))
print("8.Mode of Median Income is = ",s.mode(d['median_income']))
print("9.Mode of Median House Value is = ",s.mode(d['median_house_value']))

print("1.Variance of Longitude is = ",np.var(d['longitude']))
print("2.Variance of Latitude is = ",np.var(d['latitude']))
print("3.Variance of Housing Median_age is = ",np.var(d['housing_median_age']))
print("4.Variance of Total Rooms is = ",np.var(d['total_rooms']))
print("5.Variance of Total Bedrooms is = ",np.var(d['total_bedrooms']))
print("6.Variance of Population is = ",np.var(d['population']))
print("7.Variance of Households is = ",np.var(d['households']))
print("8.Variance of Median Income is = ",np.var(d['median_income']))
print("9.Variance of Median House Value is = ",np.var(d['median_house_value']))

```

```

print("1.Standard deviation of Longitude is = ",np.std(d['longitude']))
print("2.Standard deviation of Latitude is = ",np.std(d['latitude']))
print("3.Standard deviation of Housing Median Age is = ",np.
    ↳std(d['housing_median_age']))
print("4.Standard deviation of Total Rooms is = ",np.std(d['total_rooms']))
print("5.Standard deviation of Total Bedrooms is = ",np.
    ↳std(d['total_bedrooms']))
print("6.Standard deviation of Population is = ",np.std(d['population']))
print("7.Standard deviation of Households is = ",np.std(d['households']))
print("8.Standard deviation of Median Income is = ",np.std(d['median_income']))
print("9.Standard deviation of Median House Value is = ",np.
    ↳std(d['median_house_value']))

print("1.Range of Longitude is = ",max(d['longitude'])-min(d['longitude']))
print("2.Range of Latitude is = ",max(d['latitude'])-min(d['latitude']))
print("3.Range of Housing Median Age is_
    ↳=",max(d['housing_median_age'])-min(d['housing_median_age']))
print("4.Range of Total Rooms is =",max(d['total_rooms'])-min(d['total_rooms']))
print("5.Range of Total Bedrooms is_
    ↳=",max(d['total_bedrooms'])-min(d['total_bedrooms']))
print("6.Range of Population is = ",max(d['population'])-min(d['population']))
print("7.Range of Households is = ",max(d['households'])-min(d['households']))
print("8.Range of Median Income is_
    ↳=",max(d['median_income'])-min(d['median_income']))
print("9.Range of Median House Value is_
    ↳=",max(d['median_house_value'])-min(d['median_house_value']))

print("1.Interquartile range of Longitude is = ",np.
    ↳percentile(d['longitude'],75)-np.percentile(d['longitude'],25))
print("2.Interquartile range of Latitude is = ",np.
    ↳percentile(d['latitude'],75)-np.percentile(d['latitude'],25))
print("3.Interquartile range of Housing Median Age is = ",np.
    ↳percentile(d['housing_median_age'],75)-np.
    ↳percentile(d['housing_median_age'],25))
print("4.Interquartile range of Total Rooms is = ",np.
    ↳percentile(d['total_rooms'],75)-np.percentile(d['total_rooms'],25))
print("5.Interquartile range of Total Bedrooms is = ",np.
    ↳percentile(d['total_bedrooms'],75)-np.percentile(d['total_bedrooms'],25))
print("6.Interquartile range of Population is = ",np.
    ↳percentile(d['population'],75)-np.percentile(d['population'],25))
print("7.Interquartile range of Households is = ",np.
    ↳percentile(d['households'],75)-np.percentile(d['households'],25))
print("8.Interquartile range of Median Income is = ",np.
    ↳percentile(d['median_income'],75)-np.percentile(d['median_income'],25))

```

```

print("9.Interquartile range of Median House_value is = ",np.
↳percentile(d['median_house_value'],75)-np.
↳percentile(d['median_house_value'],25))

print("1.Skewness of Longitude is ",s.skew(d['longitude']))
print("2.Skewness of Latitude is ",s.skew(d['latitude']))
print("3.Skewness of Housing Median Age is ",s.skew(d['housing_median_age']))
print("4.Skewness of Total Rooms is ",s.skew(d['total_rooms']))
print("5.Skewness of Total Bedrooms is ",s.skew(d['total_bedrooms']))
print("6.Skewness of Population is ",s.skew(d['population']))
print("7.Skewness of Households is ",s.skew(d['households']))
print("8.Skewness of Median Income is ",s.skew(d['median_income']))
print("9.Skewness of Median House Value is ",s.skew(d['median_house_value']))

print("1.Kurtosis of Longitude is ",s.kurtosis(d['longitude']))
print("2.Kurtosis of Latitude is ",s.kurtosis(d['latitude']))
print("3.Kurtosis of Housing Median Age is ",s.
↳kurtosis(d['housing_median_age']))
print("4.Kurtosis of Total Rooms is ",s.kurtosis(d['total_bedrooms']))
print("5.Kurtosis of Total Bedrooms is ",s.kurtosis(d['total_rooms']))
print("6.Kurtosis of Population is ",s.kurtosis(d['population']))
print("7.Kurtosis of Households is ",s.kurtosis(d['households']))
print("8.Kurtosis of Median Income is ",s.kurtosis(d['median_income']))
print("9.Kurtosis of Median House_value is ",s.
↳kurtosis(d['median_house_value']))

```

	longitude	latitude	housing_median_age	total_rooms	total_bedrooms	\
0	-122.05	37.37	27.0	3885.0	661.0	
1	-118.30	34.26	43.0	1510.0	310.0	
2	-117.81	33.78	27.0	3589.0	507.0	
3	-118.36	33.82	28.0	67.0	15.0	
4	-119.67	36.33	19.0	1241.0	244.0	
...	
2995	-119.86	34.42	23.0	1450.0	642.0	
2996	-118.14	34.06	27.0	5257.0	1082.0	
2997	-119.70	36.30	10.0	956.0	201.0	
2998	-117.12	34.10	40.0	96.0	14.0	
2999	-119.63	34.42	42.0	1765.0	263.0	

	population	households	median_income	median_house_value
0	1537.0	606.0	6.6085	344700.0
1	809.0	277.0	3.5990	176500.0
2	1484.0	495.0	5.7934	270500.0
3	49.0	11.0	6.1359	330000.0
4	850.0	237.0	2.9375	81700.0

...
2995	1258.0	607.0	1.1790	225000.0
2996	3496.0	1036.0	3.3906	237200.0
2997	693.0	220.0	2.2895	62000.0
2998	46.0	14.0	3.2708	162500.0
2999	753.0	260.0	8.5608	500001.0

[3000 rows x 9 columns]

	longitude	latitude	housing_median_age	total_rooms	total_bedrooms	\
0	-122.05	37.37	27.0	3885.0	661.0	
1	-118.30	34.26	43.0	1510.0	310.0	
2	-117.81	33.78	27.0	3589.0	507.0	
3	-118.36	33.82	28.0	67.0	15.0	
4	-119.67	36.33	19.0	1241.0	244.0	
...	
2995	-119.86	34.42	23.0	1450.0	642.0	
2996	-118.14	34.06	27.0	5257.0	1082.0	
2997	-119.70	36.30	10.0	956.0	201.0	
2998	-117.12	34.10	40.0	96.0	14.0	
2999	-119.63	34.42	42.0	1765.0	263.0	

	population	households	median_income	median_house_value
0	1537.0	606.0	6.6085	344700.0
1	809.0	277.0	3.5990	176500.0
2	1484.0	495.0	5.7934	270500.0
3	49.0	11.0	6.1359	330000.0
4	850.0	237.0	2.9375	81700.0
...
2995	1258.0	607.0	1.1790	225000.0
2996	3496.0	1036.0	3.3906	237200.0
2997	693.0	220.0	2.2895	62000.0
2998	46.0	14.0	3.2708	162500.0
2999	753.0	260.0	8.5608	500001.0

[3000 rows x 9 columns]

1. Mean of Longitude is -119.58919999999999
2. Mean of Latitude is 35.635389999999994
3. Mean of Housing Median Age is 28.845333333333333
4. Mean of Total Rooms is 2599.5786666666667
5. Mean of Total Bedrooms is 529.95066666666666
6. Mean of Population is 1402.7986666666666
7. Mean of Households is 489.912
8. Mean of Median Income is 3.8072717999999997
9. Mean of Median House_value is 205846.275
1. Median of Longitude is = -118.485
2. Median of Latitude is = 34.27
3. Median of Housing Median Age is = 29.0
4. Median of Total Rooms is = 2106.0

5. Median of Total Bedrooms is = 437.0
 6. Median of population is = 1155.0
 7. Median of Households is = 409.5
 8. Median of Median Income is = 3.4871499999999997
 9. Median of Median House Value is = 177650.0
 1. Mode of Longitude is = ModeResult(mode=-118.26, count=26)
 2. Mode of Latitude is = ModeResult(mode=34.02, count=35)
 3. Mode of Housing Median Age is = ModeResult(mode=52.0, count=173)
 4. Mode of Total Rooms is = ModeResult(mode=907.0, count=5)
 5. Mode of Total Bedrooms is = ModeResult(mode=314.0, count=15)
 6. Mode of Population is = ModeResult(mode=870.0, count=7)
 7. Mode of Households is = ModeResult(mode=273.0, count=12)
 8. Mode of Median Income is = ModeResult(mode=15.0001, count=9)
 9. Mode of Median House Value is = ModeResult(mode=500001.0, count=125)
 1. Variance of Longitude is = 3.9784442266666673
 2. Variance of Latitude is = 4.5339804479
 3. Variance of Housing Median_age is = 157.58541155555554
 4. Variance of Total Rooms is = 4645033.750478223
 5. Variance of Total Bedrooms is = 172710.9642328889
 6. Variance of Population is = 1061664.8941315555
 7. Variance of Households is = 133489.24558933332
 8. Variance of Median Income is = 3.4380673509780935
 9. Variance of Median House Value is = 12791798338.666039
 1. Standard deviation of Longitude is = 1.9946037768606244
 2. Standard deviation of Latitude is = 2.1293145488396026
 3. Standard deviation of Housing Median Age is = 12.553302814620363
 4. Standard deviation of Total Rooms is = 2155.2340361265233
 5. Standard deviation of Total Bedrooms is = 415.58508663436044
 6. Standard deviation of Population is = 1030.3712409280238
 7. Standard deviation of Households is = 365.3618009443972
 8. Standard deviation of Median Income is = 1.854202618641796
 9. Standard deviation of Median House Value is = 113100.83261703266
 1. Range of Longitude is = 9.6900000000000012
 2. Range of Latitude is = 9.36
 3. Range of Housing Median Age is = 51.0
 4. Range of Total Rooms is = 30444.0
 5. Range of Total Bedrooms is = 5417.0
 6. Range of Population is = 11930.0
 7. Range of Households is = 4928.0
 8. Range of Median Income is = 14.5002
 9. Range of Median House Value is = 477501.0
 1. Interquartile range of Longitude is = 3.79000000000000063
 2. Interquartile range of Latitude is = 3.7599999999999998
 3. Interquartile range of Housing Median Age is = 19.0
 4. Interquartile range of Total Rooms is = 1728.0
 5. Interquartile range of Total Bedrooms is = 345.0
 6. Interquartile range of Population is = 962.75
 7. Interquartile range of Households is = 324.25

8.Interquartile range of Median Income is = 2.112475
9.Interquartile range of Median House_value is = -121200.0
1.Skewness of Longitude is -0.29770868312210047
2.Skewness of Latitude is 0.45958599688301055
3.Skewness of Housing Median Age is 0.01850386331347812
4.Skewness of Total Rooms is 4.165553250335494
5.Skewness of Total Bedrooms is 3.861461224097534
6.Skewness of Population is 2.9481956505974214
7.Skewness of Households is 3.5579732876776733
8.Skewness of Median Income is 1.697662360984838
9.Skewness of Median House Value is 0.9890670635093471
1.Kurtosis of Longitude is -1.3625007669256197
2.Kurtosis of Latitude is -1.1244987241902247
3.Kurtosis of Housing Median Age is -0.8044440474433037
4.Kurtosis of Total Rooms is 28.48753189248459
5.Kurtosis of Total Bedrooms is 32.044466669077906
6.Kurtosis of Population is 16.41387618265221
7.Kurtosis of Households is 26.18366680809118
8.Kurtosis of Median Income is 5.614812216724314
9.Kurtosis of Median House_value is 0.3927409720063544