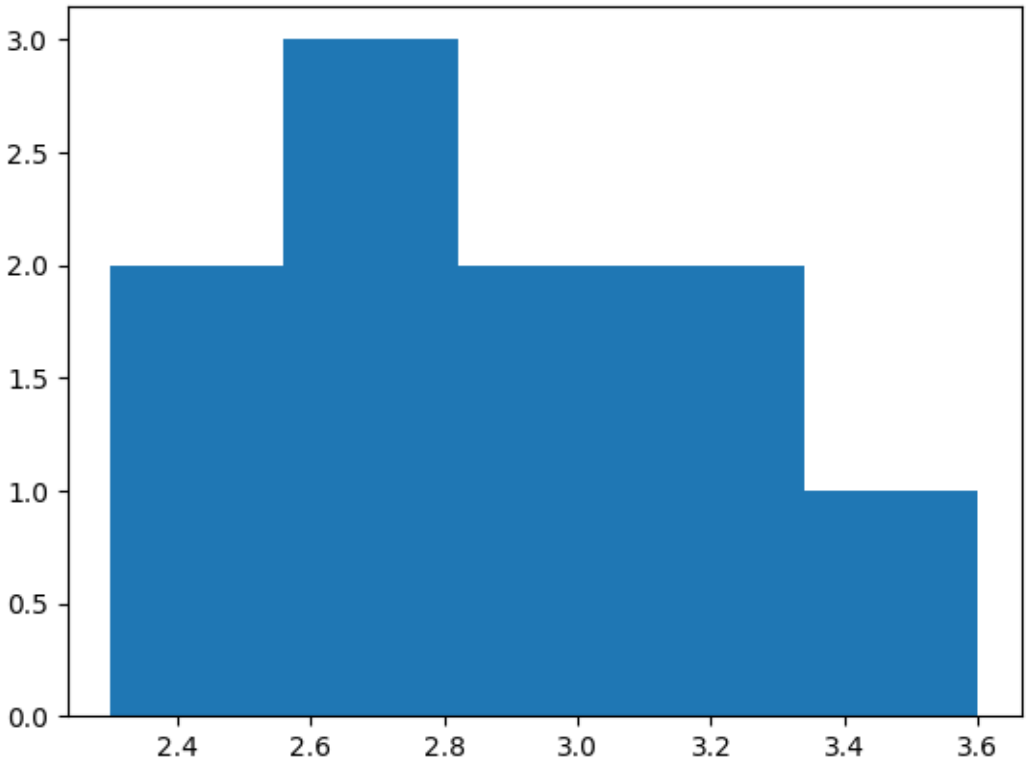
sta-a-1

August 17, 2024

[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
x=np.array([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)
plt.show()
```



[30]:
```
# 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=[4, 5, 8, 2, 4, 2, 5]
y=[5, 6, 3, 8, 3, 7, 8]
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: 4.285714285714286
Mean of y is: 5.714285714285714
Median of x is: 4.285714285714286
Median of y is: 5.714285714285714
Mode of x is : ModeResult(mode=2, count=2)
Mode of y is : ModeResult(mode=3, count=2)
Variance of x is : 3.6326530612244894
Variance of y is : 3.918367346938776
Standard deviation of x is : 1.9059520091609048
Standard deviation of y is : 1.979486637221574
Range of x is : 6
Range of y is : 5
Interquartile range of x is : 2.0
Interquartile range of y is : 3.5
Skewness of x is : 0.5785710902698393
Skewness of y is : -0.26386711521557127
Kurtosis of x is : -0.34282287589950755
Kurtosis of y is : -1.4541015625000002
```

[44]:
```
# 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
```

```
d=pd.read_csv('/content/sample_data/california_housing_train.csv')
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         -114.31     34.19                15.0       5612.0          1283.0
1         -114.47     34.40                19.0       7650.0          1901.0
2         -114.56     33.69                17.0        720.0           174.0
3         -114.57     33.64                14.0       1501.0           337.0
4         -114.57     33.57                20.0       1454.0           326.0
…             …         …                   …            …               …
16995     -124.26     40.58                52.0       2217.0           394.0
16996     -124.27     40.69                36.0       2349.0           528.0
16997     -124.30     41.84                17.0       2677.0           531.0
16998     -124.30     41.80                19.0       2672.0           552.0
16999     -124.35     40.54                52.0       1820.0           300.0

       population  households  median_income  median_house_value
0          1015.0       472.0         1.4936             66900.0
1          1129.0       463.0         1.8200             80100.0
2           333.0       117.0         1.6509             85700.0
3           515.0       226.0         3.1917             73400.0
4           624.0       262.0         1.9250             65500.0
…             …           …              …                  …
16995       907.0       369.0         2.3571            111400.0
16996      1194.0       465.0         2.5179             79000.0
16997      1244.0       456.0         3.0313            103600.0
16998      1298.0       478.0         1.9797             85800.0
16999       806.0       270.0         3.0147             94600.0

[17000 rows x 9 columns]
1.Mean of Longitude is =  -119.5621082352941
```

```
2.Mean of Latitude is =  35.62522470588235
3.Mean of Housing Median Age is =  28.58935294117647
4.Mean of Total Rooms is =  2643.664411764706
5.Mean of Total Bedrooms is =  539.4108235294118
6.Mean of Population is =  1429.5739411764705
7.Mean of Households is =  501.2219411764706
8.Mean of Median Income is =  3.8835781000000007
9.Mean of Median House_value is =  207300.91235294117
1.Median of Longitude is =  -118.49
2.Median of Latitude is =  34.25
3.Median of Housing Median Age is =  29.0
4.Median of Total Rooms is =  2127.0
5.Median of Total Bedrooms is =  434.0
6.Median of population is =  1167.0
7.Median of Households is =  409.0
8.Median of Median Income is =  3.5446
9.Median of Median House Value is = 180400.0
1.Mode of Longitude is =  ModeResult(mode=-118.31, count=136)
2.Mode of Latitude is =  ModeResult(mode=34.06, count=205)
3.Mode of Housing Median Age is = ModeResult(mode=52.0, count=1052)
4.Mode of Total Rooms is =  ModeResult(mode=1582.0, count=16)
5.Mode of Total Bedrooms is =  ModeResult(mode=280.0, count=48)
6.Mode of Population is =  ModeResult(mode=891.0, count=23)
7.Mode of Households is =  ModeResult(mode=306.0, count=48)
8.Mode of Median Income is =  ModeResult(mode=3.125, count=41)
9.Mode of Median House Value is = ModeResult(mode=500001.0, count=814)
1.Variance of Longitude is =  4.020455814167474
2.Variance of Latitude is =  4.56795267891903
3.Variance of Housing Median_age is = 158.42166311072666
4.Variance of Total Rooms is =  4751889.69496877
5.Variance of Total Bedrooms is = 177651.3369887336
6.Variance of Population is =  1317488.9119444669
7.Variance of Households is =  147847.57962446712
8.Variance of Median Income is =  3.6408471185562723
9.Variance of Median House Value is = 13451442293.56867
1.Standard deviation of Longitude is = 2.005107432076265
2.Standard deviation of Latitude is =  2.1372769307974644
3.Standard deviation of Housing Median Age is = 12.58656677218719
4.Standard deviation of Total Rooms is = 2179.8829544195187
5.Standard deviation of Total Bedrooms is = 421.4870543548563
6.Standard deviation of Population is = 1147.819198281884
7.Standard deviation of Households is = 384.50953125308496
8.Standard deviation of Median Income is = 1.9081003953032116
9.Standard deviation of Median House Value is = 115980.35304985354
1.Range of Longitude is =  10.039999999999992
2.Range of Latitude is =  9.410000000000004
3.Range of Housing Median Age is = 51.0
4.Range of Total Rooms is =  37935.0
```

```
5.Range of Total Bedrooms is =  6444.0
6.Range of Population is =  35679.0
7.Range of Households is =  6081.0
8.Range of Median Income is =  14.5002
9.Range of Median House Value is = 485002.0
1.Interquartile range of Longitude is = 3.7900000000000063
2.Interquartile range of Latitude is = 3.789999999999999
3.Interquartile range of Housing Median Age is = 19.0
4.Interquartile range of Total Rooms is = 1689.25
5.Interquartile range of Total Bedrooms is = 351.25
6.Interquartile range of Population is = 931.0
7.Interquartile range of Households is = 323.25
8.Interquartile range of Median Income is = 2.2006250000000005
9.Interquartile range of Median House_value is = 145600.0
1.Skewness of Longitude is  -0.3039761523070951
2.Skewness of Latitude is  0.4717594898275316
3.Skewness of Housing Median Age is 0.06488830685009067
4.Skewness of Total Rooms is  4.002376807943256
5.Skewness of Total Bedrooms is  3.322343534496144
6.Skewness of Population is  5.1867541718637655
7.Skewness of Households is  3.3423734139781573
8.Skewness of Median Income is  1.6265495626993867
9.Skewness of Median House Value is 0.9729507751946058
1.Kurtosis of Longitude is  -1.3222937000181343
2.Kurtosis of Latitude is  -1.1122523149717827
3.Kurtosis of Housing Median Age is  -0.8009436273526562
4.Kurtosis of Total Rooms is  19.686605656520893
5.Kurtosis of Total Bedrooms is  29.50685142754095
6.Kurtosis of Population is  80.83786311937729
7.Kurtosis of Households is  20.686206054602916
8.Kurtosis of Median Income is 4.762390906200881
9.Kurtosis of Median House_value is 0.303555274502215
```

[ ]:
```
from google.colab import drive
drive.mount('/content/drive')
```