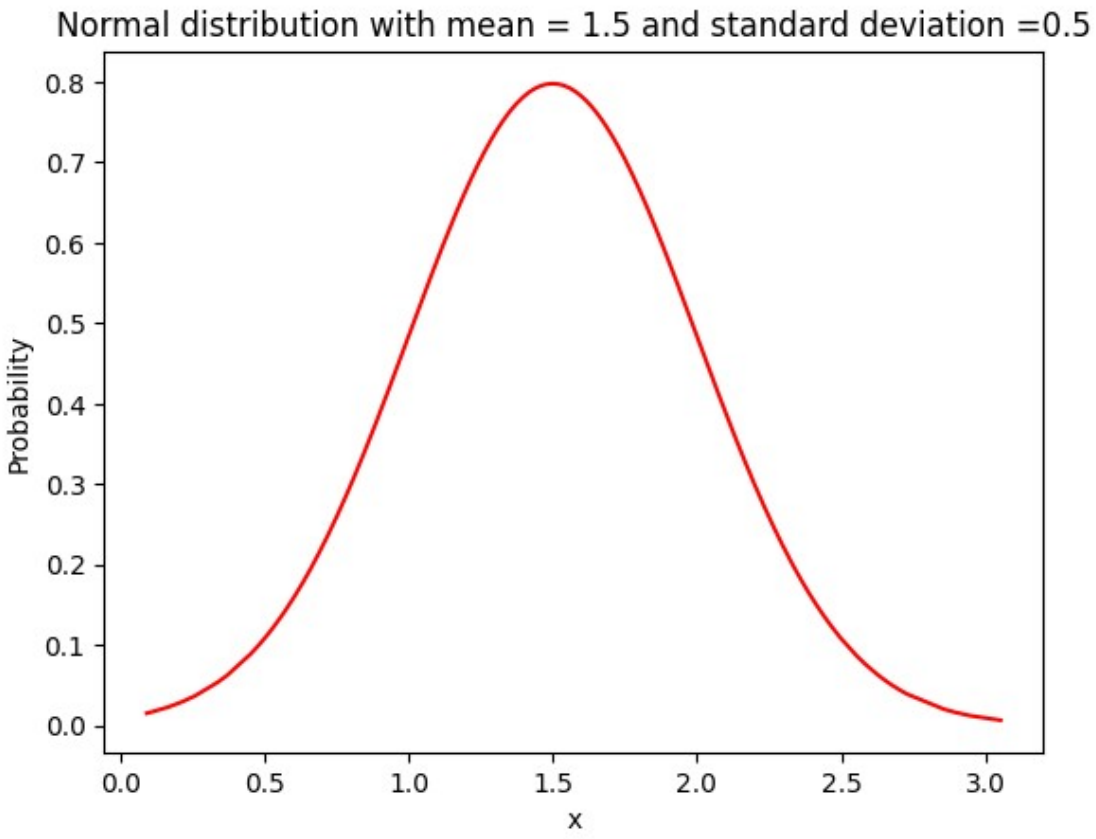
```
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sb
from astroML import stats as asts
import pandas as pd
import scipy
from scipy import stats
```

Question 1

```
n=1000
mean = 1.5
std_deviation = 0.5

nDist = stats.norm(mean,std_deviation)
draws = nDist.rvs(n)
draws.sort()
pdf = nDist.pdf(draws)

plt.title("Normal distribution with mean = 1.5 and standard deviation
=0.5")
sb.lineplot(x=draws, y=pdf , color = 'red')
plt.xlabel('x')
plt.ylabel('Probability')

Text(0, 0.5, 'Probability')
```

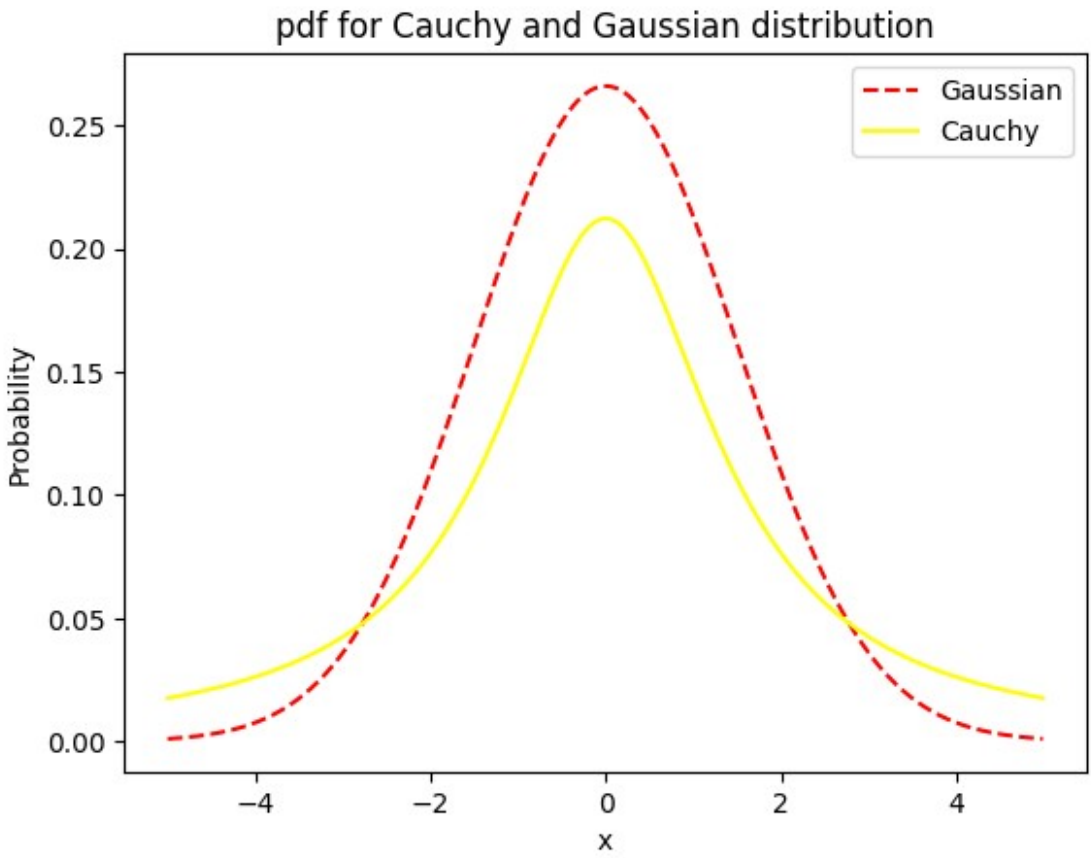


```
mean_value = np.mean(draws)
variance_value = np.var(draws)
skewness = stats.skew(draws)
kurtosis = stats.skew(draws)
median = np.median(draws)

temp = np.zeros(np.size(draws))
i=0;
for xi in draws:
    temp[i] = abs(xi-median)
    i=i+1
MAD = np.median(temp)
std_deviation_MAD = 1.482*MAD
std_deviation_sigmaG = asts.sigmaG(draws)

print("Mean is :" , mean_value)
print("Variance is :" , variance_value)
print("skewness is :" , skewness)
print("kurtosis is :",kurtosis)
print("standard deviation using MAD is :",std_deviation_MAD)
print("standard deviation using sigma g :",std_deviation_sigmaG)
```

```
Mean is : 1.4969886497835856
Variance is : 0.2653880904569996
skewness is : 0.040065558217233004
kurtosis is : 0.040065558217233004
standard deviation using MAD is : 0.5399555826992642
standard deviation using sigma g : 0.5364333165555728
```

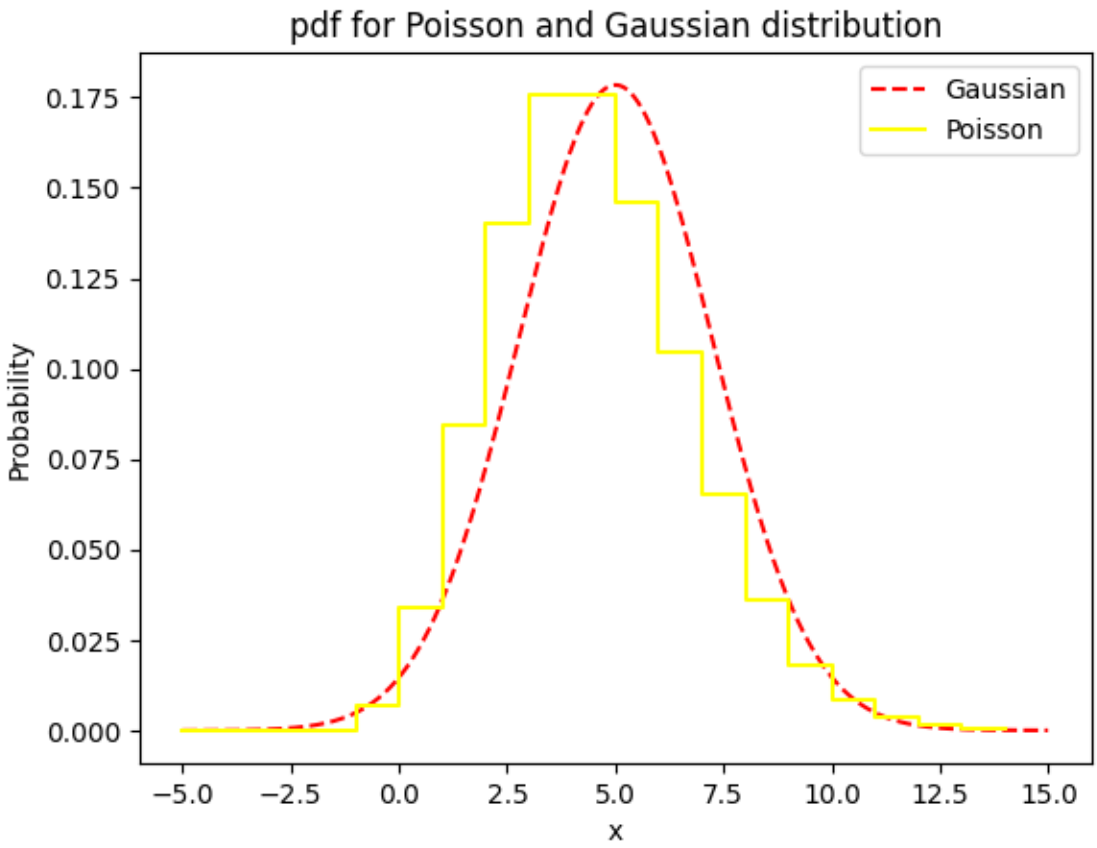
Question 2

```
mean = 0
std_deviation = 1.5
draws = np.arange(-5,5,0.01)
cauchy_dist = stats.cauchy(mean,std_deviation)
gaussian_dist = stats.norm(mean,std_deviation)
pdf_gaussian = gaussian_dist.pdf(draws)
pdf_cauchy = cauchy_dist.pdf(draws)

plt.plot(draws, pdf_gaussian , color = 'red', ls ="--")
plt.plot(draws, pdf_cauchy , color = 'yellow')
plt.title("pdf for Cauchy and Gaussian distribution")
plt.xlabel('x')
plt.ylabel('Probability')
plt.legend(labels=["Gaussian","Cauchy"])
plt.show()
```



Question 3

```
mean =5
std_deviation = np.sqrt(5)
data_poisson = np.arange(-5,15,1)
data_gaussian = np.arange(-5,15,0.01)

gaussian_dist = stats.norm(mean,std_deviation)
pdf_gaussian = gaussian_dist.pdf(data_gaussian)
poisson_dist = stats.poisson(mean)
pmf_poisson = poisson_dist.pmf(data_poisson)

plt.plot(data_gaussian, pdf_gaussian , color = 'red', ls ="--")
plt.step(data_poisson, pmf_poisson , color = 'yellow')
plt.title("pdf for Poisson and Gaussian distribution")
plt.xlabel('x')
plt.ylabel('Probability')
plt.legend(labels=["Gaussian","Poisson"])
plt.show()
```



Question 4

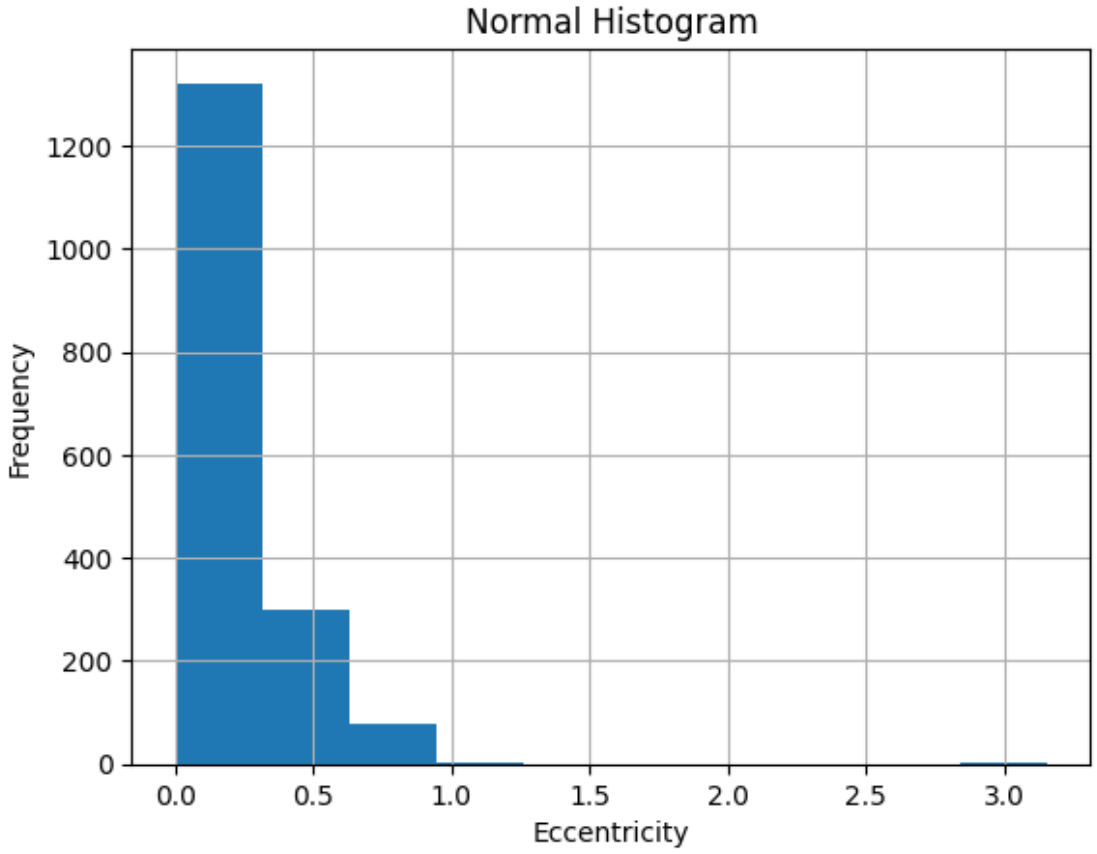
```
weighted_mean = 0
uncertainty = 0
mean_life = [0.8920 , 0.881 , 0.8913 , 0.9837 , 0.8958]
error_in_mean_life = [0.00044, 0.009, 0.00032, 0.00048, 0.00045]
for i in error_in_mean_life:
        uncertainty += 1/(i*i)
for i in range(0,len(mean_life)):
         weighted_mean
+=(mean_life[i]/(error_in_mean_life[i]*error_in_mean_life[i]))/uncerta
inty
uncertainty = pow(1/uncertainty , 0.5)

print("Weighted mean lifetime in 10^-10s is  :" , weighted_mean)
print("Uncertainty of the mean is :",uncertainty)
```

```
Weighted mean lifetime in 10^-10s is  : 0.9089185199574896
Uncertainty of the mean is : 0.00020318737026848627
```

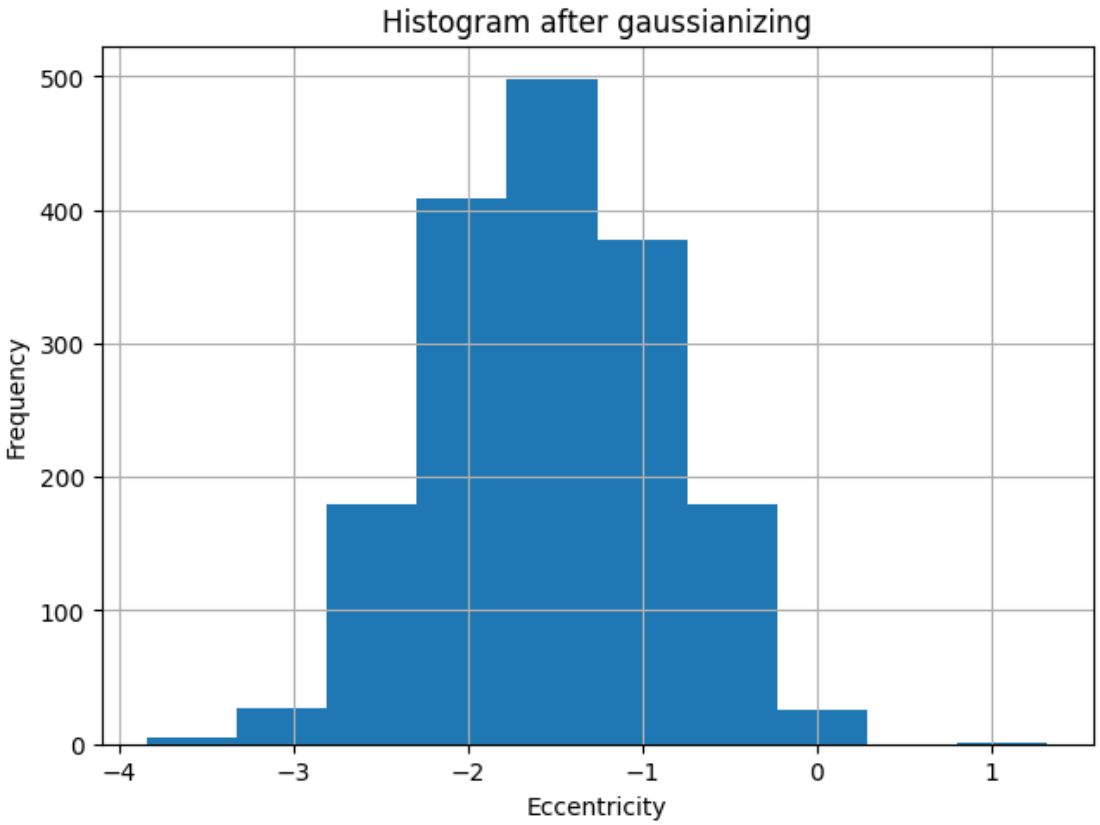
Question 5

```
data = pd.read_csv('exoplanet.eu_catalog.csv')
df = pd.DataFrame(data)
df=df.loc[:,"eccentricity"]
df1 =[]
```

```
for row in df:
    if (row>0) :
        df1.append(row)
df2 = pd.DataFrame(df1)
df2.head()
df2.hist( )
plt.xlabel("Eccentricity")
plt.ylabel("Frequency")
plt.title("Normal Histogram")
plt.show()
```




```
Gauusianized_data,temp  =scipy.stats.boxcox(df1)
Gauusianized_data = pd.DataFrame(Gauusianized_data)
Gauusianized_data.hist()
plt.tight_layout()
plt.xlabel("Eccentricity")
plt.ylabel("Frequency")
plt.title("Histogram after gaussianizing")
plt.show()
```


Histogram after gaussianizing
Frequency
Eccentricity
500
400
300
200
100
0
−4
−3
−2
−1
0
1