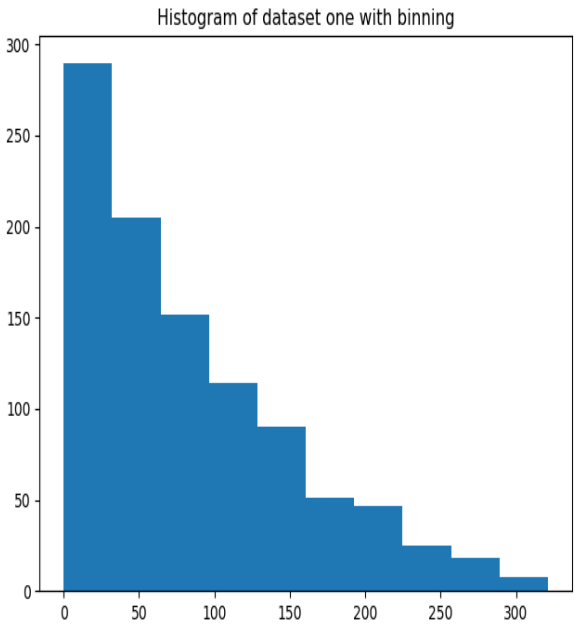
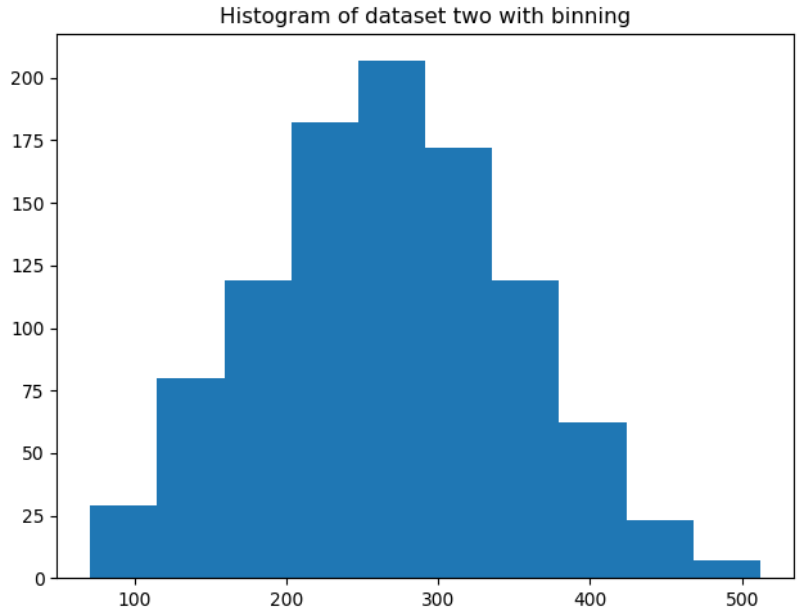
Q1) Determine the distributions of the vector component values for both datasets. For each dataset, randomly pick up 10 samples and report the distribution parameters for each of the 10 samples.

Ans –

1) First I have plotted the dataset graph with five different techniques.
   First, I plotted the histogram with a built-in function that gave me normal distribution for those 10 random vectors.
   After that I plotted Q-Q and boxplot again they gave me the uniform plot for those 10 random samples as well.
   For the confirmation, I plotted the frequency plot of the data by making the dictionary of keys and values. In which key is my data sample whereas values are the number of frequency of that data. Which confirms that dataset one is following **Uniform distribution**.
2) For the second dataset, I repeated the process that I used for the first dataset. In this as well I found out the Q-Q plot, Boxplot, Histogram and the plot based on the frequency distribution of random ten vectors.
   From these plots, I gained the similarity in terms of pattern which is having **a Gaussian distribution.**





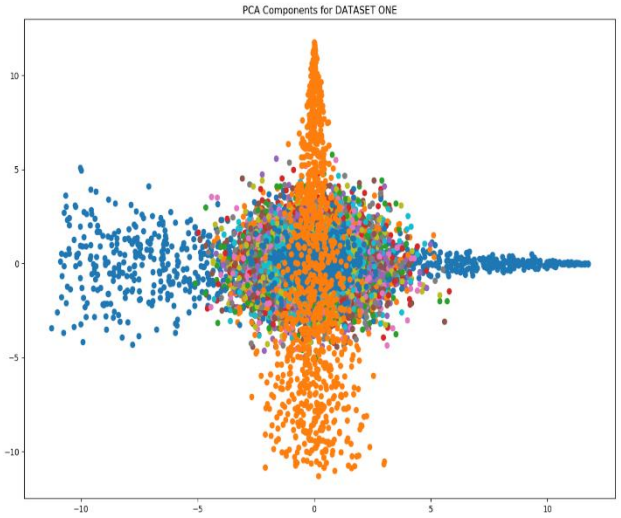
Q2) Implement PCA and DCT methods and apply them for feature extraction to the two datasets, respectively. Report the principle you have proposed to truncate the dimensionality and the reduced dimensionalities for the two datasets after the feature extraction for PCA and DCT, respectively.

Ans –

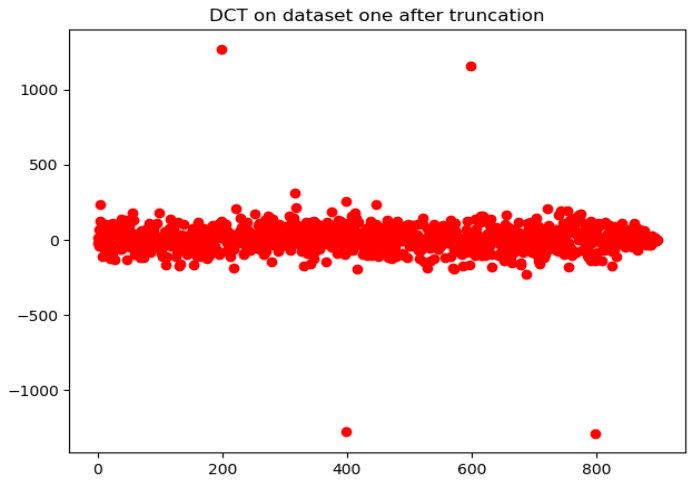
1) For the feature extraction, I first found out the covariance and correlation matrix. Covariance gave me the change into how data samples move with each other and correlation gave me the relationship between two data samples.
2) Then I got the distribution by each sample into variance, so for 100 components in data gave me a total variance of 100. From which my first 53 components gave me ~81% of the original data. Because of this, I truncated dataset one with 53 components for PCA.
3) Then I applied the same strategy for the dataset two again I got the 72 components which provided me the ~91% accuracy in the data. So for dataset two with PCA analysis, I truncated with 72 components.
4) For the implementation of DCT, I used DCT-1 technique on both datasets. First, I traversed through the matrix in zigzag way to get the one-dimensional list out from the matrix. In which I find out the threshold of the data through the absolute threshold method. First I got the minimum value form the DCT applied data. Then I got the value of 201 after which I truncated the rest of the data. Which provided me the 81.71% of data. For second dataset I applied the same strategy which gave me 1001 value. With this I got 91.08% data and I truncated graph after that.

Q3. Compare the feature extraction results between the two methods for the two datasets, respectively, and report your comparison conclusion.

1) For comparison purposes I truncated dataset one at ~81% because after DCT-1 operation I got ~81% data then I truncated the PCA at the same data variance as well. That gave me a fair analysis of data for two different methods.
2) DCT is basically used for compression purposes. So when we do DCT dimension of the data remains constant but data is compressed into DCT feature extraction.
3) So in PCA we have got 53 components for dataset1 and plotted it on the graph. In DCT we plotted one-dimensional data.
4) We apply the same method for the second dataset where we got ~91% of data with 72 components after PCA analysis.
5) I applied DCT-1 through formula.

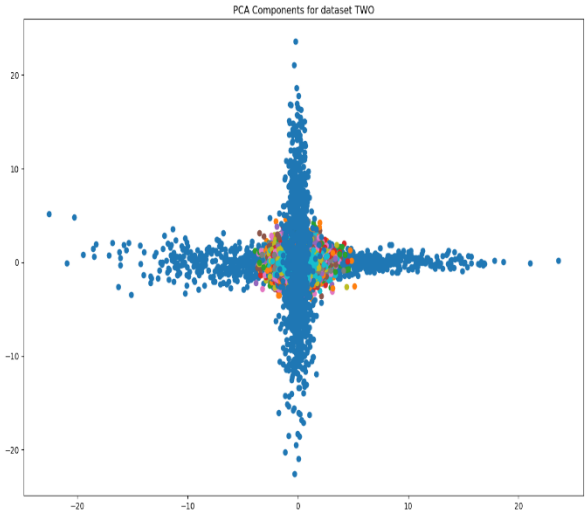


PCA at Dataset 1

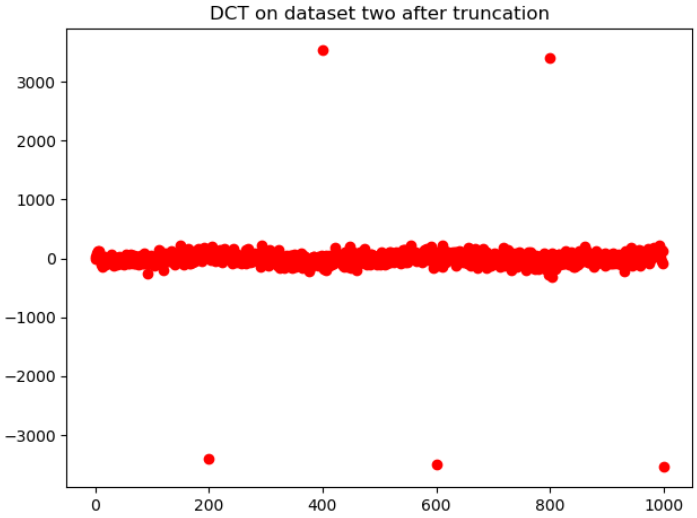


DCT at Dataset 1

PCA at Dataset 2



DCT at Dataset 2



6) Above are the results for both datasets with PCA and DCT implementation.

Q4) Read the literature on Independent Component Analysis (ICA) and implement ICA. Then apply ICA to the two datasets, respectively. Report your comparison studies on the two datasets between PCA and ICA on feature extraction.

1) ICA is known as Independent component analysis, where the data is distributed at the components and the data is situated at one of the original components of the data. On the other hand, PCA gives data at the different components on the basis of the Eigenvalues and the Eigenvectors.

2) ICA and PCA both uses a similar number of components. In dataset one they both take 53 components and in dataset two they both take 72 components for the respective feature extraction.
3) The ICA provides a better result then PCA at uniformly distributed data whereas PCA generates better results at the gaussian distribution.