**Acme Money Analysis and Prediction Enterprises**

Executive Summary

Problem Statement:

The following details are provided regarding the counterfeit notes in 1372 samples of data. This classification of data is to be used for further analysis and to perform the statistical analysis on the data. The degree of dependency of each variable on the others and the variable to be predicted (Counterfeit / Genuine) is to be determined using the statistical tools correlation and cross - covariance.

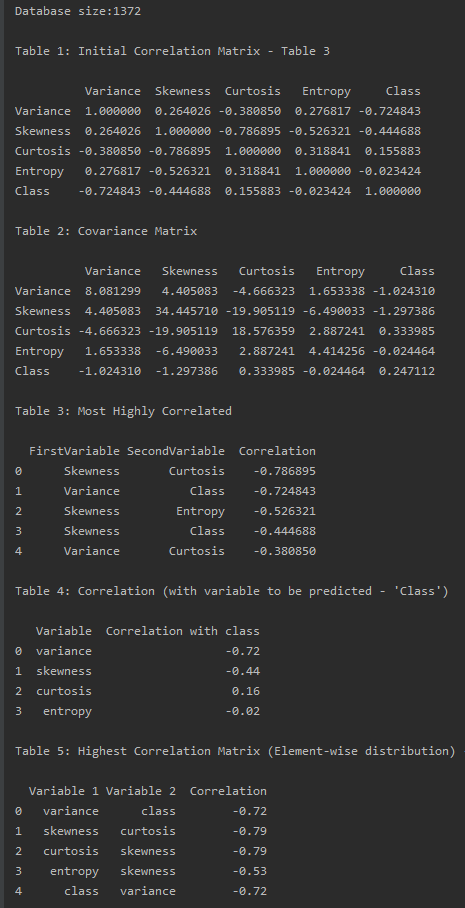
The following data are given regarding the counterfeit bill:

1. Variance of wavelength transformed image (continuous)
2. Skewness of wavelength transformed image (continuous)
3. Curtosis of wavelength transformed image (continuous)
4. Entropy of image (continuous)
5. Class (integer)

Objective:

Create a machine learning model that learns based on the given information on the bills with their classification information. This model will be used for developing an app for the US Treasury to aid in the detection of counterfeit bills. Also, perform an analysis to choose the variables of significance from the 5 entities listed above that will have a profound impact in the design of the machine learning model.

Observations:



Inferences:

1. Table 1 and Table 2 give the correlation and the cross co - variance between all the variables within the dataset.
2. From Table 3, the top 5 correlation values and their respective variables are taken into consideration. Here, the absolute values of the correlation represent the degree of correlation between the 2 variables within the data sample. So, skewness is considered to be one of the best parameters with good correlation with curtosis, entropy and class.