**Bank Noted Analysis**

**The Purpose of the project:**

The purpose of this project is to make an analysis of the Bank Noted Authentic and Forged from the given dataset by OpenML. After make an analysis over the given data, we will make a recommendation for the client based on the graph by Jupyter notebook.

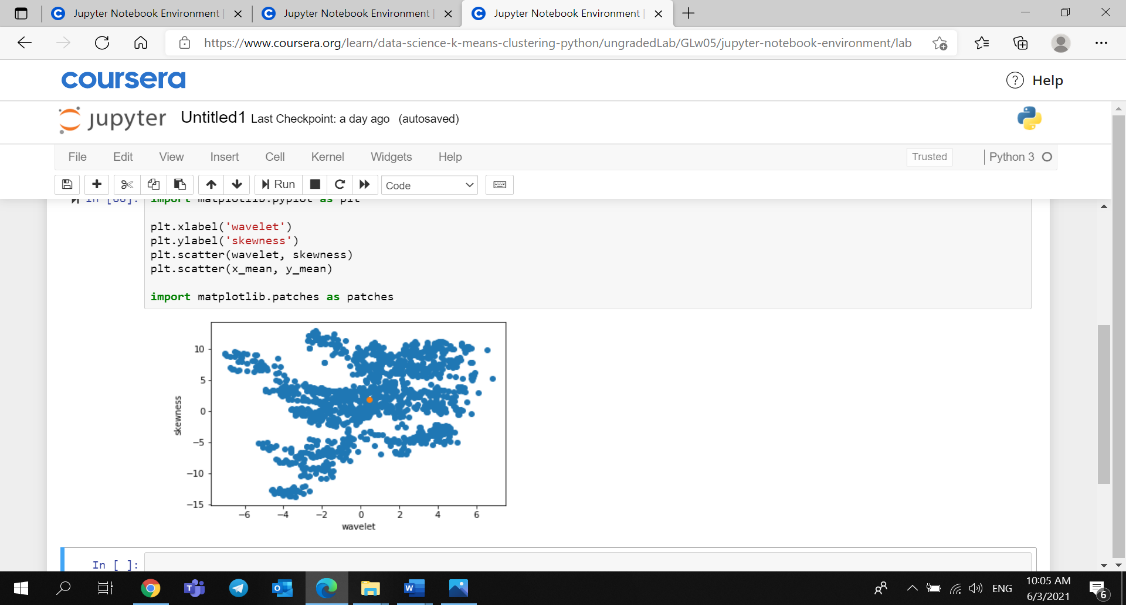
**Description of Data**

The data that we have gather from is OpenML which is the most reputable source of data provider for Data science and Machine Learning. We gather CSV to load on Python that the data has the description as below:

* V1: variance of wavelet transformed image
* V2: variance of skewness of wavelet transformed image

**Methods: how the data were analyzed**

We analyzed data by using the K-mean Clustering. K-means has five basic steps and works as follows. **Step one.** **First**, we **select the number of clusters** we want to look for. This is the k in K-means. Here we choose k equals two (Rich and Poor). The algorithm then randomly selects k points on our data axis. Note that it doesn't matter that these points do not necessarily correspond to existing data. These points are called our data **centers or centroids**. **Step two.** The distance from each data point to each of **our k centroids is calculated**. In this case, distance is simply a measure of the difference in income between the points. **Step three. Clusters are formed** by assigning each data point to either centroid one or two, depending on which is closest. **Step four.** This is the **update step.** The average value calculated over the members of each cluster is then set as the new centroid. We ignore and dispense with the previous centroid value. **Finally, step five.** We then recursively **run steps two to four recalculating the centroids until eventually the centroid positions do not change.** When the centroids remain stable like this, the algorithm is said to have **converged.**

** Summary of the results**

Based on the graph we can see that the V1 which is the variance of wavelet transformed image and V2 is variance of skewness of wavelet transformed image, has got the same mean. Furthermore, we can see that the distance from the mean to each variable was within the ranges. This so us that each variable got the similar result of tested on V1 which is the variance of wavelet transformed image and V2 is variance of skewness of wavelet transformed image, suggest us that the note on wavelet and skewness was accurate.

**Recommendations**

As the wavelet and skewness was accurate based on our tested above, we recommend that the client could use this experiment as the basis for identifying Authentic and Forged bank note.