# Banknote Authentication Analysis Using Data Science Technology

## Why use this technology?

Forged banknotes are a worldwide issue and kind of hard to solve and identify, we have developed a system that can identify forged notes by their features and predict them.

## Our purpose

To show how did we analyze the given data about the banknotes and identifying them as real or forged, we are going to use several statistics and computer tools, we will visualize the results and making a final diagnostic.

The main objective and goal are easily doing the former and training the machine so it can detect the notes by itself in the future.

## Data

The data that will be used is from a lot of images of banknotes, we used a tool to extract data from them so we can identify which one are real or forged, it will not be immediate though the tool that was used doesn't tell you exactly if it's real or forged, statistics are used to divided them into 2 groups. We'll see that topic in more detail later.

## Methods

Using real and forged banknotes we extracted information by transforming images into data, we have an input of 1370 banknotes with their respective information, the technique used here to identify them It's called clustering, it can easily detect forged and real banknotes based on their data.

The first step was to visualize the whole data without filters and sorting, as you can see it's a mess and you can't identify anything much less make a conclusion or decisions with it.

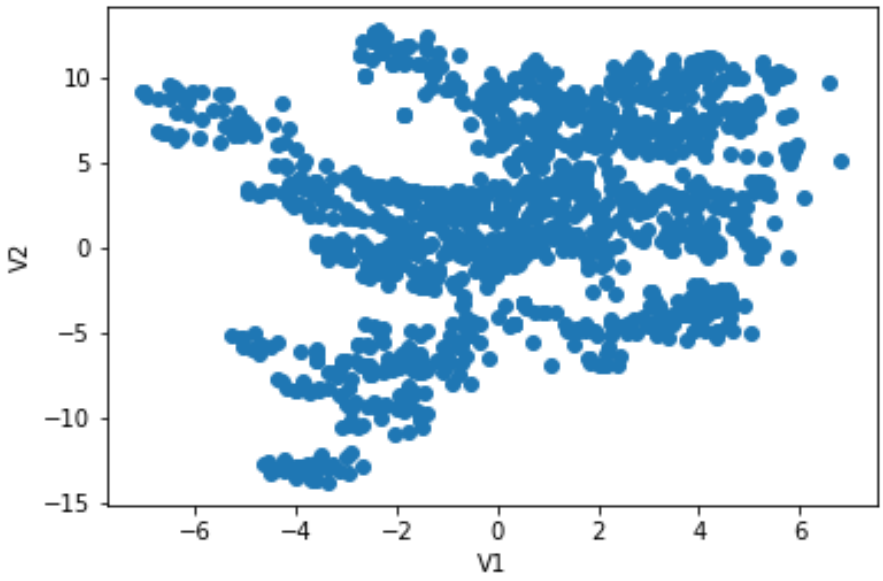


Fig 1. Scatter plot of the raw data.

After having the data displayed and ready, we can now apply the algorithm and divide the data into groups. Ideally forged and genuine.

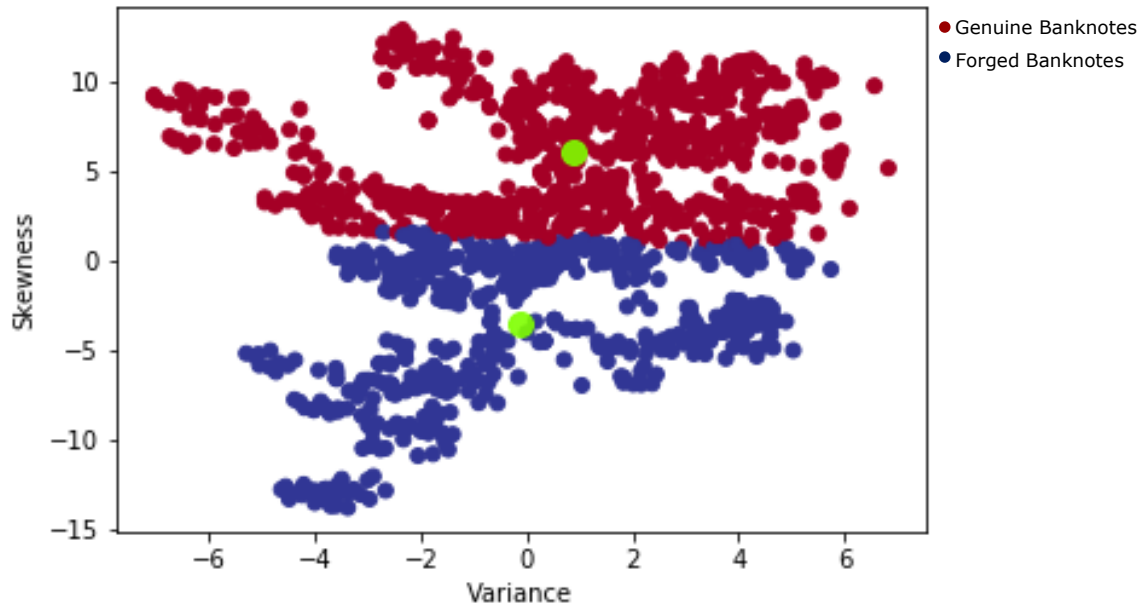


Fig . Scatter plot of the organized data.

After we applied the algorithm, we can see 2 big, separated regions, one red and one blue, being the first one the genuine notes and the latter the false ones. This is just one example of how great the grouping algorithm can be, we will not only use the visual information, but the resulting data can also be entered into more techniques and machines to further detect which are the false notes, we will not only use the visual information, but the resulting data can also be entered into more algorithms and techniques to detect even more details about the false notes.

The more information is being input the better the result will be. Also, we recommend assuring the data of your end, quantity is nothing without quality, with a mix of both even greater results can be achieved pair with an excellent accuracy.