**Analysis Banknote**

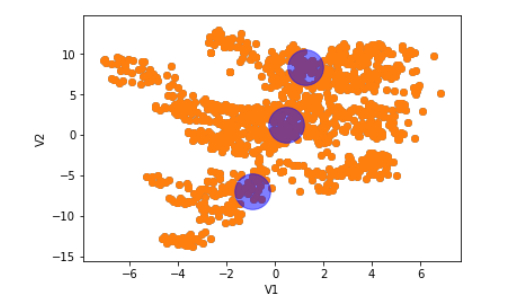
**Purpose**

In this analysis, working with de archive “Banknote-authentication-dataset-.csv”, we will prove if the data in V1 and V2 is sufficient to prove the banknotes legitimacy.

**Description**

1. Organize de data:

Graph 1



We use KMeans to be able to visualize the data is grouped.

1. This information is obtained:

Min. V1 -7.0421

Max. V1 6.8248

Mean v1 0.43373525728862977

Standard deviation: 2.841726405206097

Min. V2 -13.7731

Max. V2 12.9516

Mean v2: 1.9223531209912554

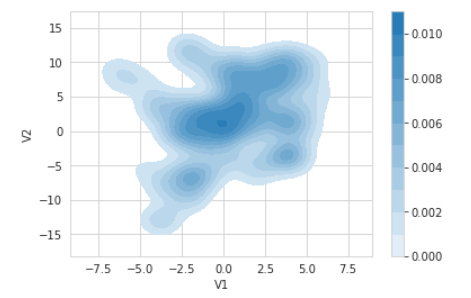
Standard deviation: 5.866907488271993

In the information above we can see the following relationships:

* Min. V2 is double V1
* Max. V2 is double V1
* Mean V2 is 4.43 times V1
* Standard deviation V2 is double V1

1. Using Kdeplot we can see most clairely how the data is grouped

Graph 2



**Summary**

Darker areas can be seen, which correspond to areas with a greater amount of data; as this parameter changes, the area become clearer. The darker areas are coincident with the means plotted in graph 1.

**Recommendations**

Looking at the information obtained, the recommendations would be:

* Continue to collect data in order to obtain more accurate readings.
* Conclude, according the data, which banknotes correspond to counterfeit banknotes and which are not; in order to remove from circulation these first.
* Place as an attachment to each data reading, the ticket number corresponding to said reading.