**E.D.A:**

From the exploratory data analysis, we resume some extra tasks done. Like the count of null values that we have already seen in class, so that we can check in a better way, how by mean replacing this number goes to zero.

A basic plot representing the unbalancing cero type presence.

A better manner to represent the relationship between a discrete and a continuous variable.

**[pass of diapositive]**

And lastly, the exchange of values from countries to a specific number assignation.

It is important to mention that we divide whole dataset into samples instead of doing it per country, due to the low number of data that we would pick up if we do it by this way, as it can be seen in Afghanistan example or French.

**[pass of diapositive]**

**4.1:**

In this first section, we found previous week the best parameters for decision trees thanks to grid method and so on. Here you can see the results from, and a graphical representation of what features have more importance in the model for the regression.

But now, we focus on k neighbours, so we consider apply a loop for each case of k value that we considered and apply its root mean square error in order to know which is the best. Finally repeat the process but with the variance of formula to compute the weights.

**[pass of diapositive]**

Here are the plots that show the variance of the performance metric

**[end]**