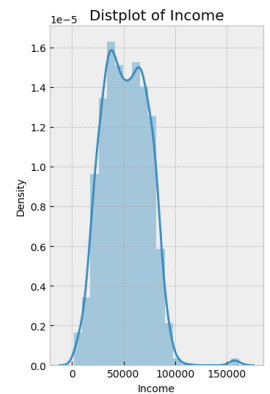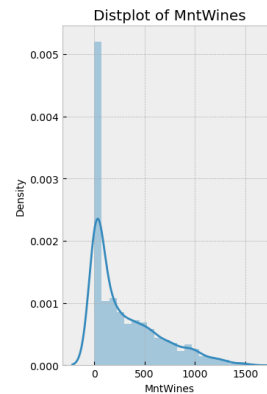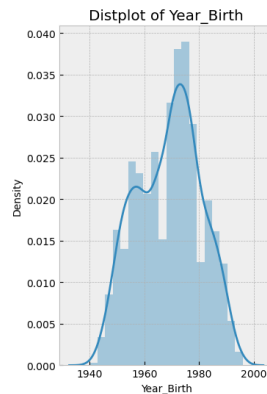# Customer Segmentation and Analysis

These histograms will provide a general view of the distribution of the 3 variables that we will take in consideration to create a segmentation of our clients:
The year of birth, the amount of wines they buy and the annual income.
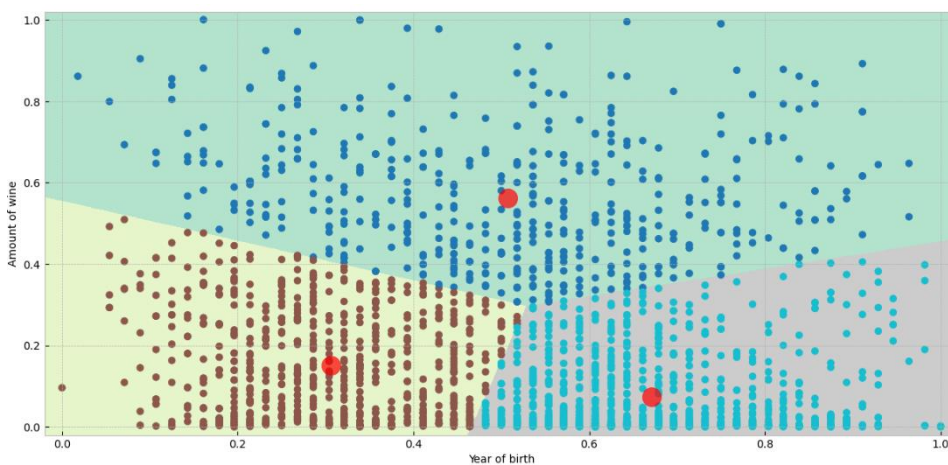
This is a clusterization related to the amount of wine and the year of birth, it split the client in a first green cluster with high consume of wine, and then split in 2 clients with low consume of wine in younger and older.

We choose 3 as a number of cluster based from inertia, because we have less decrease in inertia after 3.
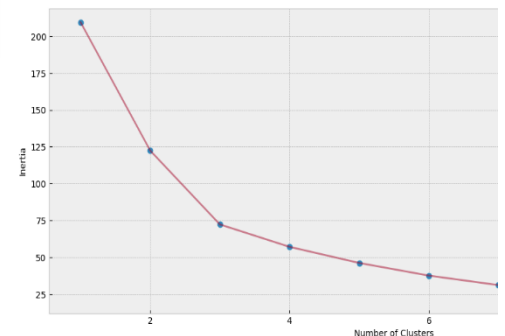
In the last graph we see the 3 cluster created considering all 3 variables, we can see that the year of birth is not a factor influencing the annual income, and that the cluster as similar as before, we have a high income cluster (blue) and two with low amount of wine (yellow and pink) separated by different year of birth.
But we can notice as expected correlation with high income and high consume of wine.
So in case we need to create an advertising campaign based on our project we will have 3 different segmentation of our clients.