SMDE FIRST ASSIGNEMENT (20% OF THE FINAL MARK, INDIVIDUAL)

SECOND QUESTION: ANOVA (25% OF THE FIRST ASSIGMENT).

Generate three populations that follow a normal distribution, **using your own algorithm**. As an example, the first is a population that follows a normal distribution with a parameter mean=0, the second with mean=10, and the third with mean=0. Select the SAME variance for the three distributions at your convenience (a value >0).

We want to analyze using an ANOVA if these three populations are different (or not) depending on the parameter selected.

Analyze and explain the results obtained. Justify your answers.

Remember to test the ANOVA assumptions. What do you expect on the assumptions?

Once you finish the analysis and you are familiar with ANOVA test: on the dataset contained on Kaggle, named "Red and White Wine Quality", we want to analyze if in both (type or quality) affects some properties of the wine. After combining the two datasets (one for red wines and one for white wines), you should create two variables. First, "type" that identifies if the wine is red or white, and second, wine quality categorized in three groups: <5 (low), 5-6 (medium) and >6 (high). Once you complete preprocessing steps, please answer to the following questions applying appropriate statistical techniques:

1) Which of the chemical properties influence the quality of the wines?
2) Which of the chemical properties are related with type of the wines?
3) How does type and quality of wines affect (separately and together) percentage of alcohol present in the wine?
4) Detail the results of Two-Way ANOVA considering as dependent variable "fixed acidity", and independent variable "type" and "quality".