*Statistical analysis for Silvestrum*

**Objective:**

Create a statistical analysis to find the appropriate number of sample plots in order to get an approximate calculation of total carbon per tree in Mg per hectare.

The ideal scenario would be to find the number of plots and number of trees to find a statistical analysis.

**Background information:**

Forest mangrove sizes does not seem to control how densely carbon (C) is stored in these mangroves, sampling at different locations within each forest revealed more about the distribution of C. Fringe areas had relatively low and consistent C densities, while in the interiors of forests we observed some of the greatest C densities. <https://datamares.org/stories/gulf-mangroves-store-carbon-by-the-ton/>

**Summary and insights from the A/R Methodological Tool**

The tool is based in the targeted precision for the biomass stock estimate, the variability of the biomass stock in different strata within the projects area, and the number of possible sample plots within the project boundary. The number of sample plots is calculated using a formula that considers the estimated standard deviation of the biomass stock in each stratum, the relative weight of each stratum, and the acceptable margin of error. The tool assumes that the approximate area of each stratum and the approximate variance of the biomass stock in each stratum are known.

**Steps in my end**

1. Calculate the variance of biomass stock from each stratum.
   1. Have the SD of each plot
   2. Create assumptions and results for that variance. Which plot has the largest concentrated pool of carbon? And why?
      1. The less trees in the plot the more concentrated carbon we have per hectare.
   3. Maybe create a matrix with the correlations of the variables?

Chart, treemap chart

Description automatically generated

1. Create a function to run the AR Methodological Tool for the sample plot calculations.
   1. The formula is iterated until the final value of the number of sample plots is obtained.

**Assumptions:**

* I’m just using the plots of 154m2
* I’m merging both data sets so we can analyze more data
* We are using the power package to calculate the power

**Questions:**

1. Should I create each plot as a stratum? Should I randomly stratify the same number of trees and calculate the mean of carbon per tree?
2. When we look at the density of trees per plot. We can see that the density is very variable and there is a trend of reducing SD as you increase density.

**Things to do next**

* Check the power calculation column showing the number of trees to have a 90% variance for each plot
* Check the power calculation column showing the number of plots to have a 90% variance for each year
* Create better plot names in the general data frames with a for loop
* Create the equation function and test it once you have the table variable.