A.5. Pour l’intervalle de confiance des différentes corrélation, je créer une fonction pour calculer directement les intervalles de confiance sur la matrice de corrélation pour avoir tous les intervalles d’un coup :

Une image contenant texte

Description générée automatiquement

Ce qui donne quand on appelle la fonction avec la matrice de corrélation :

Une image contenant texte

Description générée automatiquement

La première matrice correspondant à l’intervalle gauche et la deuxième à l’intervalle droit. Par exemple, l’IC à 95% de la corrélation entre sepal\_length et petal\_length est [-0.55 ; 0.99] ou encore l’IC à 95% de la corrélation entre petal\_length et petal\_width est [0.02 ; 0.99].

On obtient bien des intervalles compris entre -1 et 1 car la corrélation est comprise entre -1 et 1. De plus, on remarque que quand il y a une corrélation entre deux données, l’IC à 95% est plus petit alors que quand il n’y a pas de corrélation, l’IC à 95% à tendance à être centré sur 0.

B. Données multivariées : anthropométrie



Une image contenant texte

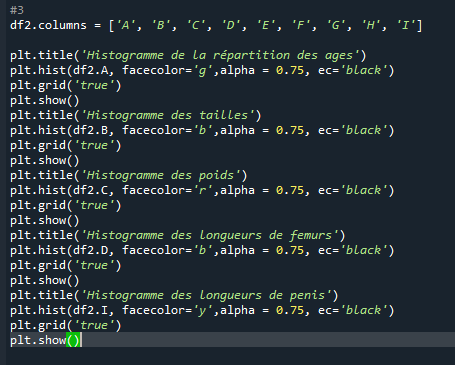
Description générée automatiquement

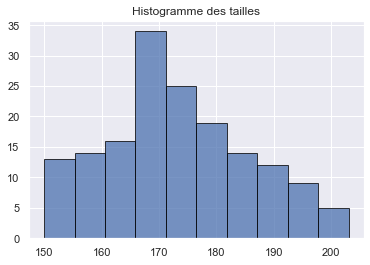
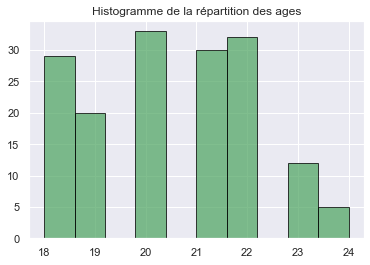
1. En utilisant la fonction describe() on obtient :

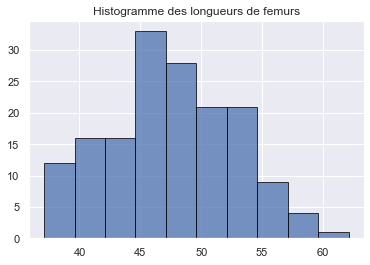
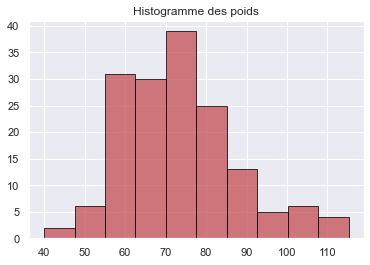
Une image contenant photo, assis, écran, tenant

Description générée automatiquement

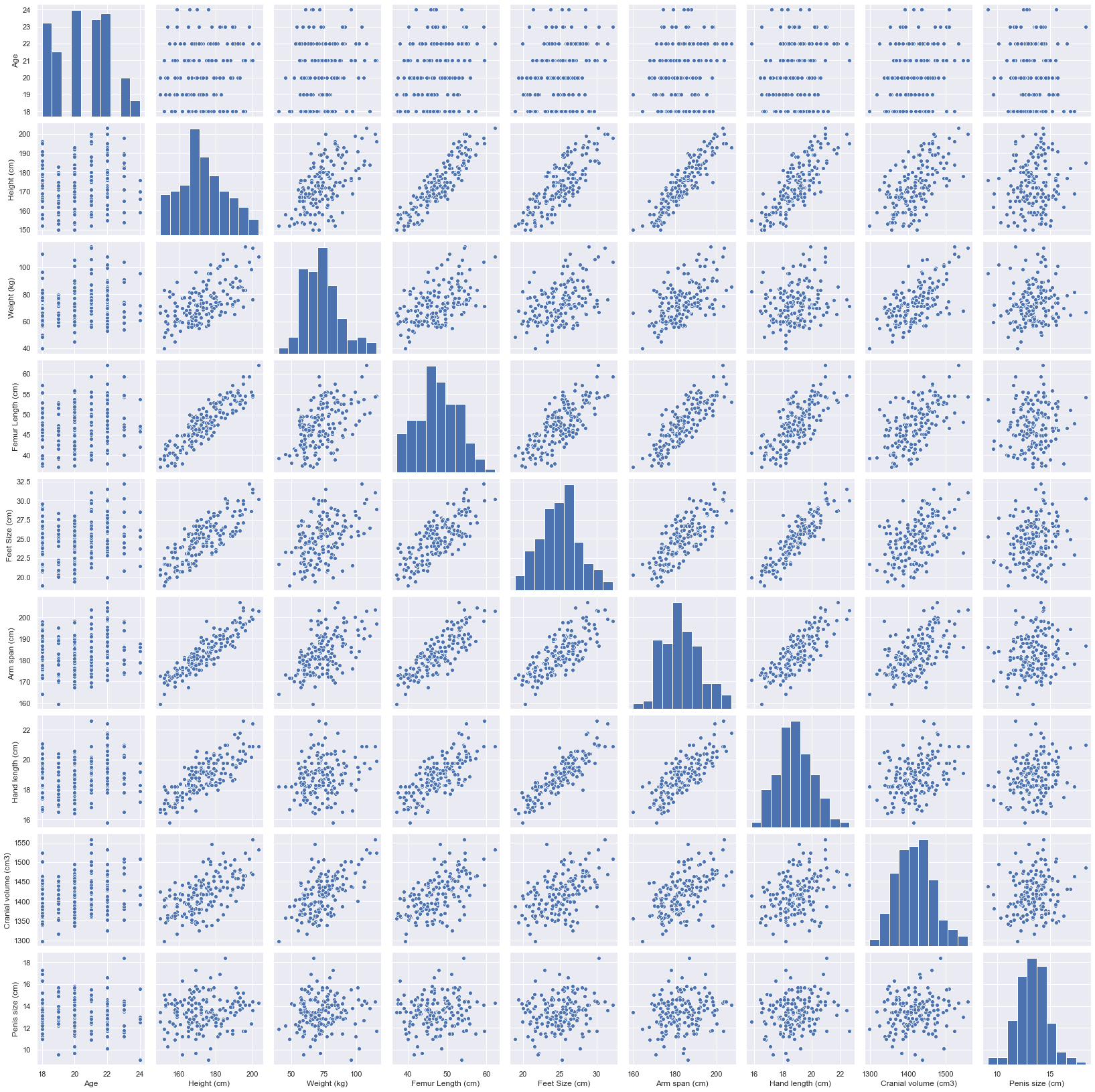
Cette fonction sert donc à calculer et afficher directement le nombre de ligne, la moyenne, l’écart-type, la valeur min/max ainsi que le 1er, 2ème et 3ème quartile de notre tableau de données.







When we analyze the different histograms, we can begin to see, thanks to the similarity between the histograms, a link between the size and the length of the femur that will be confirmed later with the correlations.



With pairplot (), we see that there is a lot of correlation between our data. Especially between the length of the femur and the size of the individual, where we notice a strong correlation. Hence the use in archeology of the length of the femur to predict the size of an individual.

5) For the calculation of the confidence intervals, we use the create function in the previous part which allows us to obtain all the 95% confidence intervals of the correlation matrix.

Une image contenant texte

Description générée automatiquement

Une image contenant texte

Description générée automatiquement

We thus obtain two confidence interval matrixes (the first is the left interval and the second is the right interval of the CI). We therefore obtained all the 95% CIs of all the correlations in our data. For example, we have the 95% CI of the correlation between age and the size of the femur equal to [-0.52; 0.76].