

## TP : Régression linéaire

Le fichier **DataHeight.txt** contient respectivement l'âge et la taille de 50 enfants. L'objectif de ce travail est de montrer qu'il existe une relation linéaire entre l'âge et la taille des enfants. Pour y parvenir, il est vous ai demandé de procéder aux étapes suivantes :

- 1- Charger les données dans une variable nommée **tab**.
- 2- Charger le vecteur **age** dans une variable nommée **x**.
- 3- Charger le vecteur **height** dans une variable nommée **y**.
- 4- Donner la distribution de y en fonction de x (utiliser les fonctions **plot** et **qplot**).
- 5- On regardant le graphique. Peut-on dire qu'il y a une corrélation entre les deux variables. Justifiez.
- 6- Prouvez d'une manière statistique (rigoureuse) que x et y sont soit corrélées soit non ?
- 7- Quelles sont les valeurs des paramètres ( $\theta_0$  et  $\theta_1$ ).
- 8- Quelle est la valeur de l'erreur résiduelle.
- 9- Donnez l'équation de la droite de régression.
- 10- Dessiner la droite de régression.
- 11- Selon le modèle trouvé prédire les poids de 4 enfants ayant respectivement 3, 7, 9 et 12 ans.
- 12- Prédire la taille d'un enfant de 10 ans avec un intervalle de confiance de 95%.