

Evaluation 1: statistique descriptive avec le Logiciel R**Prof. Armel Yodé****Exercice 1 Etude sur l'indice de masse corporelle**

Un échantillon de dossiers d'enfants a été saisi. Ce sont des enfants vus lors d'une visite en 1re section de maternelle en 1996-1997 dans des écoles de Bordeaux (Gironde, France). L'échantillon présenté ici est constitué de dix enfants âgés de 3 ou 4 ans. Les données disponibles pour chaque enfant sont :

- le sexe : F pour fille et G pour garçons;
- le fait que leur école soit située en ZEP (zone d'éducation prioritaire, c'est-à-dire réunissant plusieurs indices de précarité sociale) : O pour oui et N pour non ;
- l'âge en années et en mois à la date de la visite (deux variables, une pour le nombre d'années et une pour le nombre de mois) ;
- le poids en kg arrondi à 100 g près ;
- la taille en cm arrondie à 0.5 cm près.

Prenom	Sexe	ZEP	Poids	An	Mois	Taille
Erika	F	O	16	3	5	100.0
Celia	F	O	14	3	10	97.0
Eric	G	O	13.5	3	5	95.5
Eve	F	O	15.4	4	0	101.0
Paul	G	N	16.5	3	8	100.0
Jean	G	O	16	4	0	98.5
Adam	G	N	17	3	11	10.3
Louis	G	O	14.8	3	9	98.0
Jules	G	O	17	4	1	101.5
Leo	G	O	16.7	3	3	100.5

En statistique, il est très important de connaître le type des variables étudiées: qualitatives ou quantitatives. R permet de spécifier explicitement ce type au moyen des fonctions.

Voilà les manipulations à effectuer avec R.

- (1) Choisissez la fonction R appropriée pour enregistrer les données de chacune des variables précédentes dans des vecteurs que vous nommerez Individus, Poids, Taille et Sexe.
- (2) Calculez la moyenne des variables pour lesquelles cela est possible.
- (3) Calculez l'IMC des individus et regroupez les valeurs obtenues dans un vecteur nommé IMC (faites attention aux unités).
- (4) Regroupez ces variables dans la structure R qui vous paraît la plus adaptée.

- (5) Construisez le digramme en barres et le diagramme en secteurs des variables "Sexe" (couleur bleu pour G et couleur rose pour F) et "ZEP". Pensez à fournir un titre à votre graphique et à annoter l'axe des ordonnées.
- (6) Tracez le nuage de points du Poids en fonction de la Taille. Pensez à fournir un titre à votre graphique et à annoter vos axes.
- (7) Existe-il une liaison linéaire entre le Poids et la Taille? Si oui Déterminer la droite de régression.

Exercice 2 Vous faites une enquête dans une maternité auprès de 60 femmes et vous étudiez l'âge de la mère à la date de naissance du premier enfant.

23 24 18 19 35 26 28 24 22 19 19 17 22 26 31 28 29 21 20 22 23 18 20 27 29 24 24 22 23 23 32
29 27 21 22 23 24 28 32 30 25 26 23 20 29 35 38 19 20 22 24 23 31 26 27 20 21 22 23 28

- (1) Quelle est la population étudiée ?
- (2) Déterminer le caractère étudié. Ce caractère est-il discret ou continu ?
- (3) Calculer l'âge moyen à la naissance du premier bébé.
- (4) Construire un tableau statistique en regroupant la série en classe d'amplitude 3 et faire apparaître les effectifs correspondants, les fréquences et les fréquences cumulées.
- (5) Tracer l'histogramme des fréquences.
- (6) Tracer la boîte à moustaches. Interpréter.
- (7) Tracer sur le même graphique la courbe des fréquences cumulées croissantes et celle des fréquences cumulées décroissantes. Que représente le point d'intersection des deux courbes?