

Gestion et analyse de données : Contrôle continu

---

*Ouvrir le logiciel STATA®, puis importer le fichier master\_microdata2018.dta.*

1. Créer un indicateur composite, nommé **hh\_wealth**, qui synthétiserait la richesse matérielle du ménage en faisant la somme des différents actifs qu'il possède. De quelle nature est cette variable ? Comment interpréter cet indicateur ?
2. Unir les modalités *graduate* et *postgraduate* de la variable **education**. Pourquoi cette procédure est-elle nécessaire ?
3. Identifier les labels des deux variables **AP8** et **AP5**, puis les renommer en conséquence.
4. Créer la variable **BMI** qui correspond au *body-mass index* (indice de masse corporelle) à l'aide de la formule suivante :

$$BMI = \frac{weight(kg)}{height(cm)^2} * 10\ 000$$

5. A partir de la variable **BMI**, créer une variable catégorielle permettant de distinguer les individus en situation de dénutrition ( $13 \leq BMI < 18,5$ ), ceux ayant un poids normal ( $18,5 \leq BMI < 25$ ) et ceux en état de surpoids ( $25 \leq BMI < 45$ ) selon l'Organisation Mondiale de la Santé (OMS). Nommer cette variable **BMI\_class**, puis penser également à nommer chaque modalité (*underweight*, *normalweight* et *overweight*), de même qu'éditer un label pour cette variable (*classification of BMI*).
6. Exclure les individus ayant moins de 18 ans.
7. Quels sont les taux de dénutrition, de poids-normal et de surpoids par zone d'habitat (rurale et urbaine) ?
8. Représenter ces taux dans des diagrammes en secteurs circulaires (camemberts). Interpréter.
9. Comparer et commenter à l'aide d'un diagramme en tuyaux d'orgue les taux de dénutrition, de poids normal et de surpoids en fonction du niveau d'éducation. Modifier manuellement le graphique pour le rendre clair et lisible.