

- Analyse de données -
TP 1 Analyse en composantes principales

1 Budget-temps

Il s'agit d'une enquête sur les Budgets-temps (temps passé dans différentes activités au cours de la journée).

Le tableau comprend 10 variables numériques et 4 variables catégorisées.

Les 10 variables numériques sont le temps passé en : PROFession, TRANsport, MéNAge, ENFANTS, COURses, TOILette, REPAs, SOMMeil Téli, LOISirs.

Les 4 variables catégorisées sont : Le sexe (1=Hommes, 2=Femmes), l'activité (1=Actifs, 2=Non Act., 9=Non précisé), l'état civil (1=Célibataires, 2=Mariés, 9=Non précisé), le Pays (1=USA, 2=Pays de l'Ouest, 3=Pays de l'Est, 4=Yougoslavie).

Le code suivant est utilisé pour identifier les lignes : H : Hommes, F : Femmes, A : Actifs, N : Non Actifs(ves), M : Mariés, C : Célibataires, U : USA, W : Pays de l'Ouest sauf USA, E : Est sauf Yougoslavie, Y : Yougoslavie.

Les temps sont notés en centièmes d'heures. La première case en haut et à gauche du tableau (HAU) indique que les Hommes Actifs des USA passent en moyenne 6 heures et 6 minutes (6 heures + 10/100 d'heure) en activité PROFessionnelle.

Le total d'une ligne (sur ces 10 variables numériques) est 2400 (24 heures).

2 Questions

1. Peut-on calculer le temps moyen passé au travail par les hommes aux Etats Unis ?
2. Quelles variables peut-on utiliser pour une analyse en composantes principales ?
3. Les autres apportent-elles une information supplémentaire par rapport au tableau où leur colonnes auraient été retirées ?
4. On réalise une ACP. Combien y a-t-il de valeurs propres ? Calculer les valeurs propres à l'aide de dudi.pca. On remarque qu'il manque une valeur propre. La calculer. Quelle conséquence sur les variables peut-on en déduire ?
5. Combien faut-il en retenir ?
6. Calculer les composantes principales. Pour chacune des 4 premières composantes principales, donner la liste des individus qui contribuent à l'axe de manière significative.
7. En utilisant la description des noms des individus donnés en introduction, donner une interprétation des deux premiers axes (au moins).