

TP 3

(datasets, contents, format, print, sort, means univariate tabulate iml)

Les fichiers se trouvent dans l'archive *donnees_tp2.zip*. Vous les placerez dans un répertoire et allouerez la librairie SAS correspondante.

1. Concaténez les deux tables permanentes `total_expenditure1` et `total_expenditure2`. Concaténez les deux tables permanentes `public_expenditure1` et `public_expenditure2`. On utilisera la PROC DATASETS. A l'issue de cette première étape, on aura deux tables permanentes `total_expenditure1` et `public_expenditure1`.

Ces tables permanentes contiennent des données de l'OCDE sur les dépenses de santé des pays membres.

2. Supprimez les doublons dans chacune des tables permanentes `total_expenditure1` et `public_expenditure1`.
3. Renommez les tables permanentes : `total_expenditure1` sera renommée `total_expenditure` et `public_expenditure1` sera renommée `public_expenditure`.
4. Fusionnez les deux tables `total_expenditure` et `public_expenditure` en une table temporaire `health_expenditure`. Supprimez les tables permanentes `total_expenditure` et `public_expenditure`.
5. Copiez la table `health_expenditure` dans la librairie SAS.
6. Éditez le dictionnaire des données de la table permanente `health_expenditure`.
7. Examinez la distribution des variables `total_expen_02` et `public_expen_02` et déterminez notamment les quartiles de chacune des variables.
8. Définissez les formats de découpage en 4 classes d'effectifs égaux des variables `total_expen_02` et `public_expen_02`. Définissez le format de regroupement par continent de la variable `country`. Ces formats seront stockés dans la librairie `myfiles`.

Visualisez le codage des formats en l'éditant dans la fenêtre Output.

9. Créez une table temporaire `health_expenditure` qui contiendra les variables de la table permanente `health_expenditure` ainsi qu'une variable `continent`.
10. Imprimez dans la fenêtre Output le fichier `health_expenditure` en sommant les variables `total_expen_02` et `public_expen_02` par continent. Les noms des pays serviront d'identifiant des observations (au lieu du numéro d'observation).
11. En utilisant les formats définis pour les variables `total_expen_02` et `public_expen_02`, éditez le tableau des fréquences de ces deux variables regroupées par modalités (on doit obtenir environ 25% des effectifs pour chaque classe) et le tableau croisé de `total_expen_02` par `public_expen_02`.

12. Editez le tableau suivant :

	Public_expen_02		Total_expen_02	
	Mean	Std	Mean	Std
continent				
America	1553.67	1140.87	2917.00	2357.03
Asia	1107.50	832.26	1536.50	764.38
Europe	1587.23	703.72	2110.91	855.73
Oceania	1577.50	184.55	2180.50	457.50
Tous	1550.00	710.96	2157.87	1026.67

13. Créez la table temporaire stat_expenditure, qui contiendra la moyenne, l'écart-type, le minimum et le maximum des dépenses publiques et totales de santé par continent.
14. Renommez la table myfiles.crops1 durwheat et la table myfiles.crops2 softwheat.
15. Créez la table wheat qui contiendra la somme pour chaque année des surfaces cultivées pour les deux types de blé. On utilisera la proc iml et on ne fusionnera pas les fichiers.