

## **Projet Pratique N°6**

*Les données de ce projet Proviennent du portail de la Banque Africaine du Développement*

### **Instructions :**

#### **A : Data Step**

1. Créer une librairie nommée '**P\_6**' qui sera la librairie de référence de tout le projet.
2. Importer le fichier de données '**Data\_P6**' en utilisant un **Data Step** SAS et en respectant les types de données des colonnes.
3. Effectuer les changements nécessaires pour que les colonnes soient conformes aux règles ci-dessous :
  - **Internet\_Users** : ne doit pas avoir des chiffres après la virgule.
  - **Phone lines per 1000 Person** : doit avoir 2 chiffres uniquement après la virgule.
  - **Milk & White Bread** : doivent être en \$. Ils s'agissent du prix d'un litre du lait et du prix de 200 grammes du pain.
  - **Region & Country** : Doivent être en majuscule.
4. Utiliser **Proc Format** pour changer les formats des colonnes conformément au fichier source.
5. Utiliser **Proc SQL** pour créer une table ayant uniquement les données agrégées par Région.  
(indication : Utiliser la fonction **SUM** avec le **GROUP BY**)

#### **B : Etude Démographique**

1. En utilisant **Proc Means**, faites une comparaison entre les populations de chaque région.
2. Donner l'**histogramme** et le descriptif des mesures de variabilité de la variable Population.
3. Donner le **Boxplot** relatif à la variable Population par région.
4. Donner le descriptif du **Boxplot** relatif à la région **North Africa**.
5. Compléter la comparaison des régions par un **Pie Chart** comparant la distribution des langues par région.
6. Est-ce qu'on peut considérer les valeurs aberrantes données par **Proc Univariate** comme des anomalies ? Expliquer votre réponse.

#### **C : Etude Economique**

1. En utilisant **Proc Tabulate**, réaliser un tableau comparatif des prix du lait et du pain par région. Le tableau doit contenir les moyennes des prix, et aussi les prix max et min de chaque région.
  2. Rajouter la variable **population** à cette table en calculant le revenu total mensuel de vente du lait et du pain par région, si on suppose que chaque personne consomme 12 litres du lait et 3 kg du pain par mois.  
(Indication : il est possible de créer une nouvelle table de données pour calculer les prix mensuels)
  3. Interpréter les résultats trouvés en donnant des commentaires sur la variation des prix par région.
  4. Donner le nuage des points de la variable **population** par région avec la nouvelle variable : **revenu\_mensuel\_vente\_lait\_pain**.  
Quels sont les pays qui consomment le plus de lait par mois ?
  5. Avec **Proc Univariate**, étudier la distribution des utilisateurs d'internet et le nombre de lignes téléphoniques utilisées par région. Donner les **histogrammes** et **Boxplot** relatifs à ces 2 variables par région.
- 

Le travail rendu doit être :

**1-** Les réponses aux questions des parties A B et C.

**2-** Le code SAS complet avec des commentaires donnant le numéro de chaque question traitée et des explications sur le choix des procédures. **1** Dans un **rapport au format Word** avec les prénoms, noms et groupes des étudiants qui ont réalisé le travail.), et **2** dans un fichier texte avec le code SAS commenté par question.

**Travail en groupes de 3 max (Projets Solo ne sont pas autorisés)**

– à rendre par mail (1 mail par groupe) avant le vendredi 11/12/2015 avant **23 h** à l'adresse suivante : **SAS.ESSAI@GMAIL.COM**