

### EXERCICE n°1 : Vecteurs, Opérateurs logiques

Déterminer quelles ont été les années bissextiles depuis 1880 et les afficher.

Rappel : condition pour qu'une année soit bissextile :

- soit elle est divisible par 4 mais non divisible par 100
- soit elle est divisible par 400.

.

### EXERCICE n°2 : Le théorème central limite par l'exemple

- 1) Générer sous R, suivant une loi uniforme  $U[0,1]$ , 1000 vecteurs de dimension 12 et stocker les données dans une matrice 12x1000.
- 2) Tracer un histogramme des données.
- 3) Calculer les moyennes des 1000 vecteurs et tracer l'histogramme des moyennes. Interpréter le résultat.

### EXERCICE n°3 : Pratique de l'objet data .frame

- 1) A partir du jeu de données iris disponible par défaut sous R (taper `data(iris)` pour charger les données dans une `data.frame` nommée `iris`), calculer les statistiques de base des variables contenues dans l'objet `iris`.
- 2) Créer un nouveau `data.frame` nommé `iris2` comportant uniquement les données de la modalité `versicolor` de la variable `Species`.
- 3) Trier par ordre décroissant les données de `iris2` en fonction de la variable `Sepal.Length` (exploiter la fonction `order`).

<b>EXERCICE n°4 : Importation de données, data.frame</b>
--

**Objectif** : manipuler la structure data.frame

- Copier dans votre répertoire de travail, le fichier ozone.txt disponible sur moodle
- Charger les données dans une data.frame nommée data
- Afficher les statistiques de base relatives aux différentes variables
- Afficher les valeurs de maxO3 associées aux valeurs de T15 > 30
- Créer un sous-jeu de données data2 issu de data et contenant les lignes correspondant à la modalité "Sec" de la variable qualitative pluie
- Trier par ordre décroissant les données de data2 en fonction de la variable T12
- Tracer dans une même fenêtre, en exploitant la data.frame data, l'histogramme et la boîte à moustaches relatifs à la variable Ne9
- Afficher les déciles de la variable maxO3

<b>EXERCICE n°5 : Etude climatologique des stations de Toulouse et d'Agen</b>
---

**Objectifs :** Décrire statistiquement le paramètre TX pour les deux stations (Toulouse-Blagnac, indicatif INSEE 31069001 et Agen-La-Garenne, indicatif INSEE 47091001), à partir d'une archive de 30 ans de mesures quotidiennes (fichier CLIM.txt décrit en annexe et disponible sur moodle). Coder et exploiter une fonction sous R.

- Charger les données dans une data.frame.
  - Ajouter à votre data.frame deux variables supplémentaires de type numeric, AN et MOIS, élaborées à partir de la variable DATE.
  - Extraire les données des stations dans deux data.frames distinctes nommées 'toul' et 'agen' (Utiliser deux méthodes d'extraction différentes).
  - Afficher pour les deux stations la valeur maximale de TX et la date associée.
  - Représenter dans une même fenêtre graphique les histogrammes de TX des 2 stations. Ajouter des titres, utiliser 2 couleurs différentes, imposer des classes d'amplitude 2°C entre -9° et 41°C pour les 2 histogrammes.
  - Ecrire une fonction qui renvoie, pour une variable d'une data.frame, une matrice de dimension (n,13) contenant dans les 12 premières colonnes les moyennes mensuelles pour chacune des n années et dans la 13<sup>ème</sup> colonne la moyenne annuelle. Les arguments de la fonction seront le nom de la data.frame et l'indice de colonne correspondant à la variable traitée.
  - Utiliser cette fonction pour créer la matrice des moyennes de TX pour chacune des deux stations. Donner pour chaque station la moyenne mensuelle maximale de TX ainsi que la moyenne annuelle maximale.
  - Tracer sur un même graphique les boîtes à moustaches des moyennes annuelles des TX pour les deux stations.
-

**Annexe : description du fichier de données**

**Nom du fichier** : CLIM.txt

**Période** : du 01/01/1971 au 31/12/2000

**Séparateurs** : séparateur de données : le point-virgule ; séparateur décimal : la virgule ;  
noms des variables en 1ere ligne

**Colonnes** :

- indicatif INSEE du poste
  - date (format AAAAMMJJ)
  - RR hauteur de précipitations quotidienne (mm et  $1/10^{\text{ème}}$ )
  - TN température minimale sous abri ( $^{\circ}\text{C}$  et  $1/10^{\text{ème}}$ )
  - TX température maximale sous abri ( $^{\circ}\text{C}$  et  $1/10^{\text{ème}}$ )
  - FXY vitesse vent quotidien maxi moyenné sur 10 mn (m/s et  $1/10^{\text{ème}}$ )
  - DXY direction vent quotidien maxi moyenné sur 10 mn (rose 360)
  - GLOT rayonnement global quotidien ( $\text{J}/\text{cm}^2$ )
  - NEIG occurrence de neige (binaire)
  - BROU occurrence de brouillard (binaire)
  - ORAG occurrence d'orage (binaire)
-