

# Travaux pratiques

Master 1 Actuariat ISF

Ibrahima SY

10/20/2021

## **Partie II : Manipulation de vecteurs**

Considérons le vecteur suivant :

$$x = [1, 2, 3, 4, 5]$$

1. Créer ce vecteur dans R et le stocker dans un objet que l'on appellera `xx`
2. Afficher le mode de `x`, puis sa longueur ;
3. Extraire le premier élément, puis le dernier ;
4. Extraire les trois premiers éléments et les stocker dans un vecteur que l'on nommera `a` ;
5. Extraire les éléments en position 1, 3, 5 ; les stocker dans un vecteur que l'on nommera `b` ;
6. Additionner le nombre 10 au vecteur `x`, puis multiplier le résultat par 2 ;
7. Effectuer l'addition de `a` et `b`, commenter le résultat ;
8. Effectuer l'addition suivante : `x+a`, commenter le résultat, puis regarder le résultat de `a+x` ;
9. Multiplier le vecteur par le scalaire `c` que l'on fixera à 2 ;
10. Effectuer la multiplication de `a` et `b`, commenter le résultat ;
11. Effectuer la multiplication suivante : `x*a`, commenter le résultat ;
12. Récupérer les positions des multiples de 2 et les stocker dans un vecteur que l'on nommera `ind`, puis conserver uniquement les multiples de 2 de `x` dans un vecteur que l'on nommera `mult_2` ;
13. Afficher les éléments de `x` qui sont multiples de 3 *et* multiples de 2 ;
14. Afficher les éléments de `x` qui sont multiples de 3 *ou* multiples de 2 ;
15. Calculer la somme des éléments de `x` ;
16. Remplacer le premier élément de `x` par un 4 ;
17. Remplacer le premier élément de `x` par la valeur `NA`, puis calculer la somme des éléments de `x` ;
18. Lister les objets en mémoire dans la session R ;
19. Supprimer le vecteur ;
20. Supprimer la totalité des objets de la session.

## **Partie II : Manipulation de listes**

1. évaluer le code suivant : `TRUE+FALSE+TRUE*4` et le commenter ;
2. évaluer les expressions suivantes : `c(1, 4, TRUE)`, et `c(1, 4, TRUE, "bonjour")`, commenter ;
3. Créer une liste que l'on appellera `l` et qui contient les éléments 1, 4 et `TRUE` en première, deuxième et troisième positions, respectivement ;
4. Extraire le premier élément de la liste `l`, et afficher son mode. En faire de même avec le troisième élément, et commenter ;
5. Ajouter un quatrième élément à la liste `l` : `"bonjour"`, puis afficher la structure de `l` ;
6. Retirer le troisième élément de la liste `l` ;

7. Créer une liste de trois éléments : votre nom, votre prénom, et votre année de naissance. Ces trois éléments de la liste devront être nommés respectivement "nom", "prenom" et "année de naissance". Stocker la liste ainsi créée dans un objet nommé `moi` ;
8. Extraire le prénom de la liste `moi` de deux manières : en utilisant l'indice, et en utilisant le nommage ;
9. Créer une liste avec la même structure que celle de `moi`, en la remplissant avec les informations d'une autre personne et la nommer `toi`. Puis, créer la liste `personnes`, qui contiendra les listes `toi` et `moi` ;
10. Extraire la liste `toi` de `personnes` (en première position) ; 11 Extraire directement depuis `personne` le prénom de l'élément en première position.

## Partie III : Manipulation de données

1. À l'aide de la fonction `read_excel()` du *package* `readr`, importer le contenu de la feuille intitulée `notes_2012` du fichier Excel disponible à l'adresse suivante : et le stocker dans une variable que l'on nommera `notes_2012` ;
2. Afficher les 6 premières lignes du jeu de données, puis les dimensions du tableau ;
3. Conserver uniquement la colonne `note_stat` du tableau de données `notes_2012` dans un objet que l'on appellera `tmp` ;
4. Conserver uniquement les colonnes `num_etudiant`, `note_stat` et `note_macro` dans l'objet `tmp` ;
5. Remplacer le contenu de `tmp` par les observations de `notes_2012` pour lesquelles l'individu a obtenu une note de stat supérieure (strictement) à 10 ;
6. Remplacer le contenu de `tmp` par les observations de `notes_2012` pour lesquelles l'individu a obtenu une note comprise dans l'intervalle (10,15)(10,15) ;
7. Regarder s'il y a des doublons dans le tableau de données `notes_2012` ; le cas échéant, les retirer du tableau ;
8. Afficher le type des données de la colonne `num_etudiant`, puis afficher le type de toutes les colonnes de `notes_2012` ;
9. Ajouter au tableau `notes_2012` les colonnes suivantes :
  - `note_stat_maj` : la note de stat (`note_stat`) majorée d'un point,
  - `note_macro_maj` : la note de macro (`note_macro`) majorée de trois points (le faire en deux étapes : d'abord deux points en plus, puis un point) ;
10. Renommer la colonne `year` en `annee` ;
11. Depuis le fichier `notes_etudiants.xlsx` (c.f. question 1), importer le contenu des feuilles `notes_2013`, `notes_2014` et `prenoms` et le stocker dans les objets `notes_2013`, `notes_2014` et `prenoms` respectivement ;
12. Empiler le contenu des tableaux de données `notes_2012`, `notes_2013` et `notes_2014` dans un objet que l'on nommera `notes` ;
13. Fusionner les tableaux `notes` et `prenoms` à l'aide d'une jointure gauche, de manière à rajouter les informations contenues dans le tableau `prenoms` aux observations de `notes`. La jointure doit se faire par le numéro d'étudiant et l'année, l'objet final viendra remplacer le contenu de `notes` ;
14. Trier le tableau `notes` par années croissantes et notes de macro décroissantes ;
15. Changer le type des colonnes `annee` et `sexe` en facteur ;
16. Créer une colonne `apres_2012` qui prend la valeur `TRUE` si l'observation concerne une note attribuée après 2012 ;
17. à l'aide de la fonction `summarize()` du *package* `{dplyr}`, calculer :
  - la moyenne et l'écart-type annuels des notes pour chacune des deux matières,
  - la moyenne et l'écart-type annuels et par sexe des notes pour chacune des deux matières ;
18. En utilisant la fonction `pivot_longer()` du *package* `{tidyr}`, créer un tableau dans lequel chaque ligne renseigne le numéro d'étudiant, l'année, le prénom, le sexe, l'enseignement (macro ou stat) et la note ;
19. En repartant de l'objet obtenu à la question précédente, utiliser la fonction `pivot_wider()` du *package* `{tidyr}` pour retomber sur le même tableau que `notes`.