

Devoir– R

Dans le cadre de ce travail, vous effectuerez des manipulations de données et produirez un tableau de bord en R à partir des données contenues dans le fichier :

- « `credit_frais_de_garde_denfants.csv` »

Les données proviennent du site : <https://www.donneesquebec.ca/fr/> et ont été transformées en `.csv`. Ces données contiennent des informations sur les crédits de frais de garde d'enfants selon l'âge, le genre, l'état civil et la région administrative pour les années d'imposition 2014 à 2018.

1. **(2 points)** Importez les données du fichier `credit_frais_de_garde_denfants.csv` dans un objet de classe `dataframe` nommé `frais_garde` :
 - a. À partir de la variable `Age`, créez une nouvelle variable `Age2` qui sera de type numérique et qui prendra les valeurs 1 à 6 selon le groupe d'âge chronologique (1 = Moins de 30 ans, ..., 6 = 50 ans et plus);
 - b. Créez ensuite une nouvelle variable `genre_age` qui sera la concaténation du premier caractère de la variable `Genre` et de la variable `Age2` (créée à 1a.). Un « - » devra séparer les deux informations (par exemple : H-1).

Astuce : Si vous éprouvez des problèmes d'encodage avec les données caractères du fichier, la spécification de l'encodage **UTF-8** à la lecture du fichier sera utile.

2. **(3 points)**
Attention : vous devez réaliser cette question deux fois :

1. En utilisant uniquement des fonctions de base dans R
2. En utilisant **dplyr**

Créez deux nouveaux objets `frais_garde_2014` et `frais_garde_2015`, qui seront des agrégations pour leur année respective (2014 et 2015 respectivement) du nombre de participants en fonction des variables `genre_age` et `Region_administrative`. Effectuez ensuite une juxtaposition (jointure) des deux objets selon les variables `genre_age` et `Region_administrative` et qui conservera l'information sur le nombre de particuliers de chaque année. Renommez vos variables de façon à faire référence à l'année dans le nom de colonne.

3. **(3 points)** Construisez un vecteur comprenant toutes les régions administratives sans doublon. À partir de ce vecteur et en utilisant obligatoirement le concept de REGEX abordé au chapitre 4 des notes de cours, construisez un vecteur comprenant tous les noms composés de région.

Le vecteur obtenu devra être le suivant :

```
[1] "Bas-Saint-Laurent" "Îles-de-la-Madeleine" "Lac-Saint-Jean"
[4] "Capitale-Nationale" "Abitibi-Témiscamingue" "Côte-Nord"
[7] "Nord-du-Québec" "Chaudière-Appalaches" "Centre-du-Québec"
```

4. **(2 points)** Créez une fonction qui vous permettra, à partir de la table de données *frais_garde*, d'afficher dans la console le nombre de particuliers total selon la prise en compte de deux arguments :
- avec ou sans conjoint (*Avec_conjoint*)
 - selon l'année d'imposition (*Annee_imposition*)

Testez votre fonction avec deux combinaisons de critères différentes.

Tableau de bord :

5. **(5 points)** À partir des données fournies, réalisez un tableau de bord d'une taille équivalente à une à deux pages (d'un format comparable au format lettre ou légal) présentant les informations que vous jugez les plus pertinentes en regard des données fournies. Vous êtes complètement libres dans cette question : présentation des résultats, format du livrable produit ...

Le seul élément obligatoire :

- Utiliser uniquement R pour réaliser le tableau de bord

Indice : Pour réaliser un tableau de bord uniquement en R, considérez les librairies *RMarkdown*, *flexdashboard*, *shiny* ou une solution similaire.

Rappel des livrables à déposer :

- Code(s) .r
- Tableau de bord produit