

## Exercice 1 (6 points)

1. Expliquer le résultat suivant :

```
TRUE + FALSE + TRUE
```

```
## [1] 2
```

2. Soit la liste :

```
l <- list(pays = c("France", "Canada"),
          population = c(66616416, 35540419),
          langues_officielles = list("Français",
                                    c("Anglais", "Français")))
```

- (a) extraire le premier élément de la liste `l` en y accédant dans un premier temps par indice, puis dans un second temps par nom,
  - (b) à l'aide de la fonction appropriée, chercher la position de la chaîne de caractères "Canada" dans le vecteur `pays` de la liste `l`, et stocker le résultat dans un objet,
  - (c) à l'aide de l'objet obtenu à la question précédente, extraire le vecteur de langues officielles uniquement pour le Canada.
3. Quelle fonction permet d'effectuer une régression linéaire? Si `reg_lin` est le résultat à l'issue de l'estimation par cette fonction, comment peut-on extraire les résidus?
  4. Rajouter la colonne `rang` à `mtcars` (`data.frame` contenu dans le *package* `datasets`), indiquant le rang de l'observation dans le `data.frame`;
  5. Soit la date "2014-12-05". Indiquer les instructions qui permettent de transformer cette chaîne en un objet `POSIXct`, en utilisant une fonction du *package* `lubridate` (on supposera que le *package* `lubridate` n'est pas installé sur la machine).

## Exercice 2 (5 points)

Supposons que les adresses e-mails des étudiants de l'Université de Rennes 1 sont constituées de la manière suivante : le prénom et le nom de famille séparés par un point, le symbole arobase et le enfin le nom de domaine. Supposons de plus que les étudiants ont un seul prénom, et aucune particule au nom de famille. La syntaxe des adresses e-mail est donc comme suit : `nom.prenom@etudiant.univ-rennes1.fr`.

1. ewen.gallic[at]univ-rennes1.fr

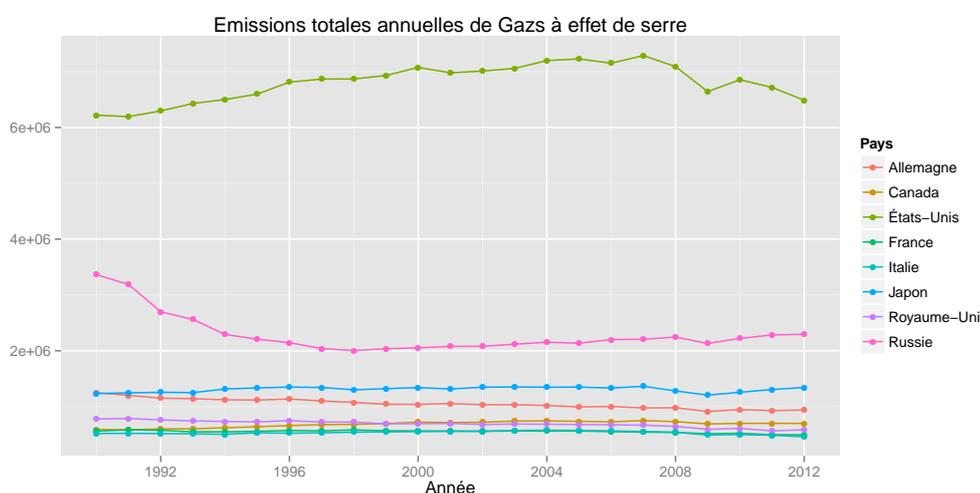
```
emails <- c("marie.petit@etudiant.univ-rennes1.fr",
            "jean.dupont@etudiant.univ-rennes1.fr",
            "isabelle.martinez@etudiant.univ-rennes1.fr",
            "pierre.moreau@etudiant.univ-rennes1.fr")
```

1. Créer une fonction, qui à partir d'une adresse e-mail d'un étudiant, retourne un `data.frame` contenant trois variables : le prénom, le nom et l'adresse e-mail de cet étudiant ;
2. Appliquer la fonction à chaque élément du vecteur `emails`, de manière à obtenir un `data.frame` à trois colonnes (prénom, nom, e-mail), et dont chaque ligne correspond à un étudiant, en utilisant deux méthodes différentes :
  - (a) à l'aide d'une boucle,
  - (b) à l'aide d'une fonction vectorielle.

### Exercice 3 (9 points)

Des données d'émissions de gaz à effets de serre sont proposées par l'OCDE<sup>2</sup>. Une copie de ces données est disponible à l'adresse suivante : [http://editerna.free.fr/Enseignement/R/AIR\\_GHG.csv](http://editerna.free.fr/Enseignement/R/AIR_GHG.csv).

1. Charger les données dans la session R (les charger depuis [http://editerna.free.fr/Enseignement/R/AIR\\_GHG.csv](http://editerna.free.fr/Enseignement/R/AIR_GHG.csv)) ;
2. Limiter les observations aux pays suivants : Canada, France, Allemagne, Italie, Japon, Russie, Royaume-Uni, États-Unis ;
3. Se limiter également aux observations dont la variable `Variable` vaut "Emissions totales exluant l'UTCATF" ;
4. Calculer et afficher les moyennes d'émissions de gaz à effet de serre sur la totalité de la période (ignorer les valeurs non disponibles) par pays et polluant, en ordonnant par ordre alphabétique des polluants puis des valeurs décroissantes des moyennes.
5. Ajouter une variable `year`, collant l'année de chaque observation à la chaîne de caractères "-01-01", puis transformer cette variable en objet `Date` (la modification doit être effectuée à l'intérieur du `data.frame`) ;
6. Reproduire le graphique suivant, et donner une instruction R pour le sauvegarder au format pdf :

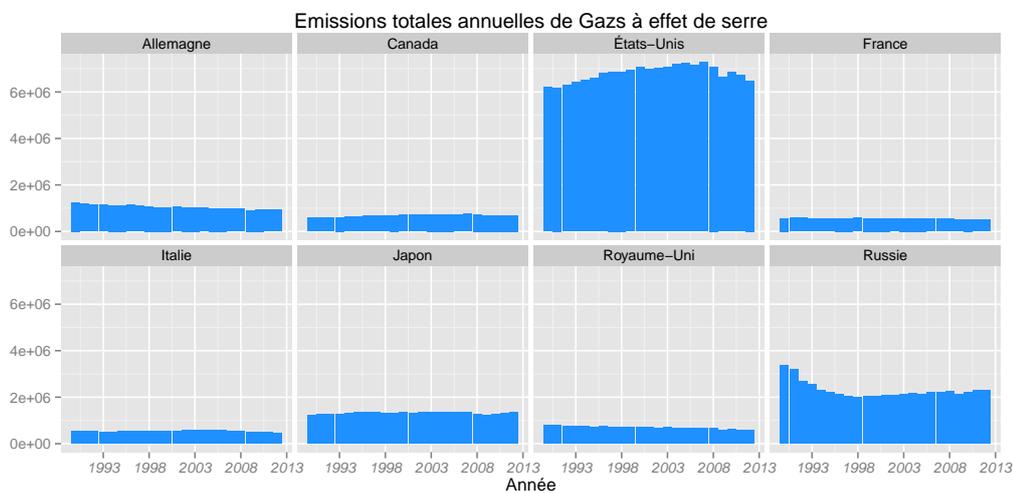


2. Statistiques de l'OCDE sur l'environnement (2014) (base de données), DOI : [10.1787/data-00594-fr](https://doi.org/10.1787/data-00594-fr)

Quelques indications :

- émission de gaz à effets de serre (valeur prise par la variable `Polluant`) en fonction des années,
- affichage des dates en abscisse tous les 4 ans ;

7. Reproduire le graphique suivant, et donner une instruction R pour le sauvegarder au format pdf :



Quelques indications :

- mêmes données que la question précédente,
- représentation par barres au lieu des lignes et des points,
- affichage des dates en abscisse tous les 5 ans,
- les étiquettes des valeurs en abscisse doivent être en italique,
- facetage par pays.