

Ce travail est à réaliser individuellement. Vous devez remettre les documents suivants :

- votre rapport au format PDF
- votre code **R** ou un fichier Rmarkdown
- un fichier CSV avec vos regroupements (voir instructions en bas de page)

Utilisez la convention de nomenclature `d6_matricule.extension`, où

- `matricule` est votre matricule et
- `extension` est un de `csv`, `pdf`, `R` ou `Rmd`.

Note Retéléchargez le paquet `hecmulti` via `remotes::install_github("lbelzile/hecmulti")` (version \geq 2022.11.22) et redémarrez **R** avant de commencer le devoir.

Les données `electionsqc` du paquet `hecmulti` contiennent les résultats électoraux (en pourcentage) pour les cinq principaux partis (CAQ, QS, PQ, PLQ, PCQ) ainsi que le taux de participation pour chacune des 125 circonscriptions aux élections générales de 2018 et 2022. La base de données contient aussi plusieurs variables sociodémographiques.

Vous devez réaliser une analyse de regroupement avec les résultats électoraux pour faire ressortir des informations pertinentes.

Voici quelques balises :

1. Effectuez si besoin des transformations préalables des variables explicatives (manipulations, créations de nouvelles variables, traitement des valeurs manquantes).
2. Déterminez si vous devez standardiser les données et comment.
3. Choisissez au moins **trois** méthodes (différentes dissemblances, algorithmes, etc.) et faites l'analyse de regroupement. Choisissez adéquatement vos hyperparamètres en vous basant sur les critères couverts en classe.
4. Choisissez la solution qui vous semble la plus satisfaisante parmi celles obtenus en justifiant adéquatement votre choix.
5. Procédez à l'interprétation des prototypes à une échelle interprétable et calculez les statistiques descriptives des variables sociodémographiques. Commentez sur le résultat de votre segmentation au vu de votre connaissance du système politique.

Vous devez retourner en plus de votre rapport et de votre code un fichier CSV avec deux colonnes :

- la première colonne pour les noms de circonscriptions (`circonscription`),
- la deuxième colonne pour les étiquettes de groupe (`regroupements`).