On s'intéresse à classer les Pokémons afin de permettre aux joueurs de se faire une équipe compétitive. Les variables que nous avons jugées intéressantes sont :

* Les statistiques : hp, attack, defense, sp\_attack, sp\_defense et speed jugeant l'aptitude au combat du Pokémon.
* Les types des Pokémon : repérer des groupes entre les types associés aux variables de combat.
* Les against des Pokémon : repérer des groupes entre les against associés aux variables de combat.

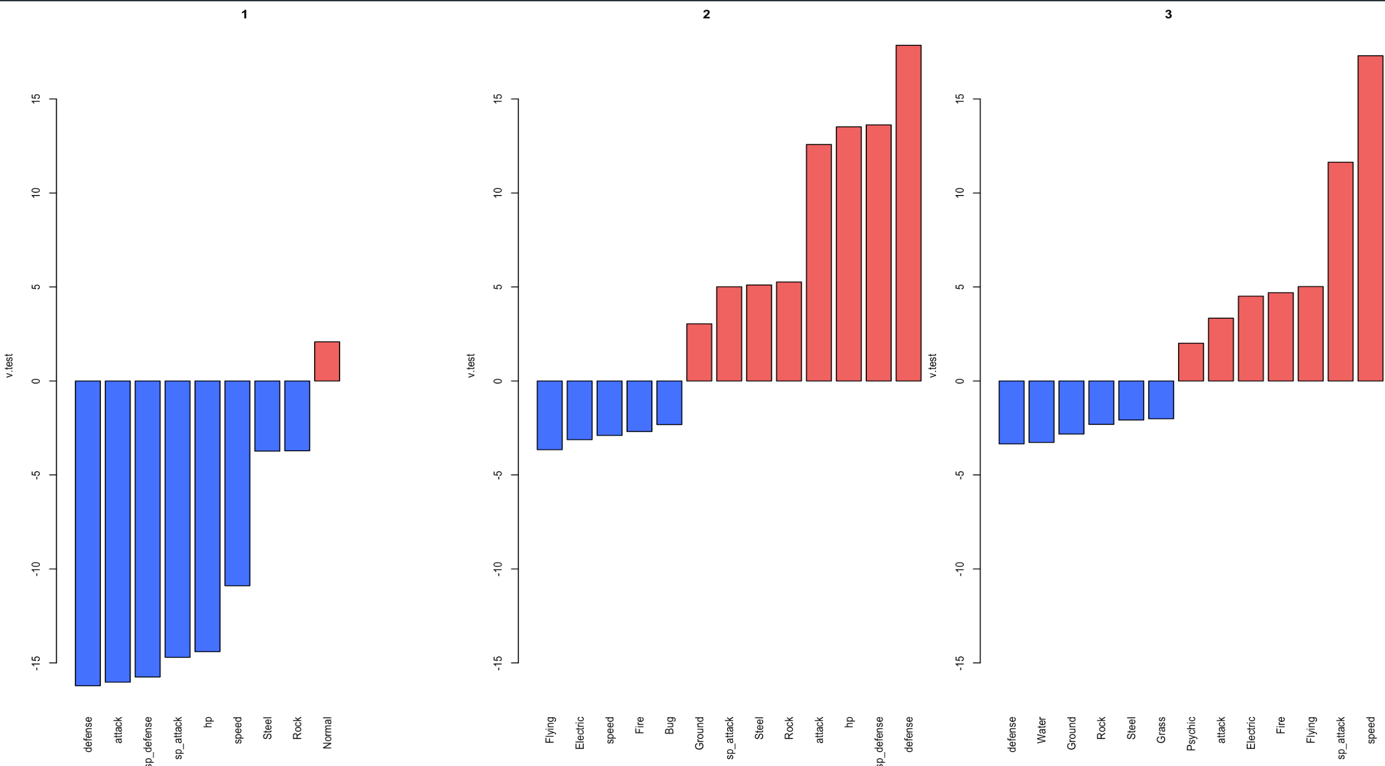
Le fait de prendre des variables qualitatives plutôt que quantitatives nous permettra de pondérer les distances entre individus.

1. CAH AVEC QUANTITATIVES + TYPES

Tout d’abord on a séparé les données qualitatives des données quantitatives. Puis on a standardisé les données quantitatives afin de créer la matrice de distance pour ces données. Ensuite, on a créé un tableau de contingence pour les types. On ne distingue donc plus les types 1 des types 2. Les différentes variables de la table deviennent donc uniquement les différents types (ayants comme valeurs 1 si le Pokémon est de ce type, 0 sinon ; un Pokémon peut donc avoir comme valeur 1 pour minimum 1 type et pour maximum 2). Puis on a calculé la matrice de distances pour cette table. À la suite de cela, on a créé notre matrice de distances finale en sommant les deux matrices de distance précédentes et en pondérant. La pondération est arbitraire, on a choisi de donner plus de poids aux variables quantitatives ici.

En appliquant la méthode de Ward, on a choisi comme nombre de groupes k=3.

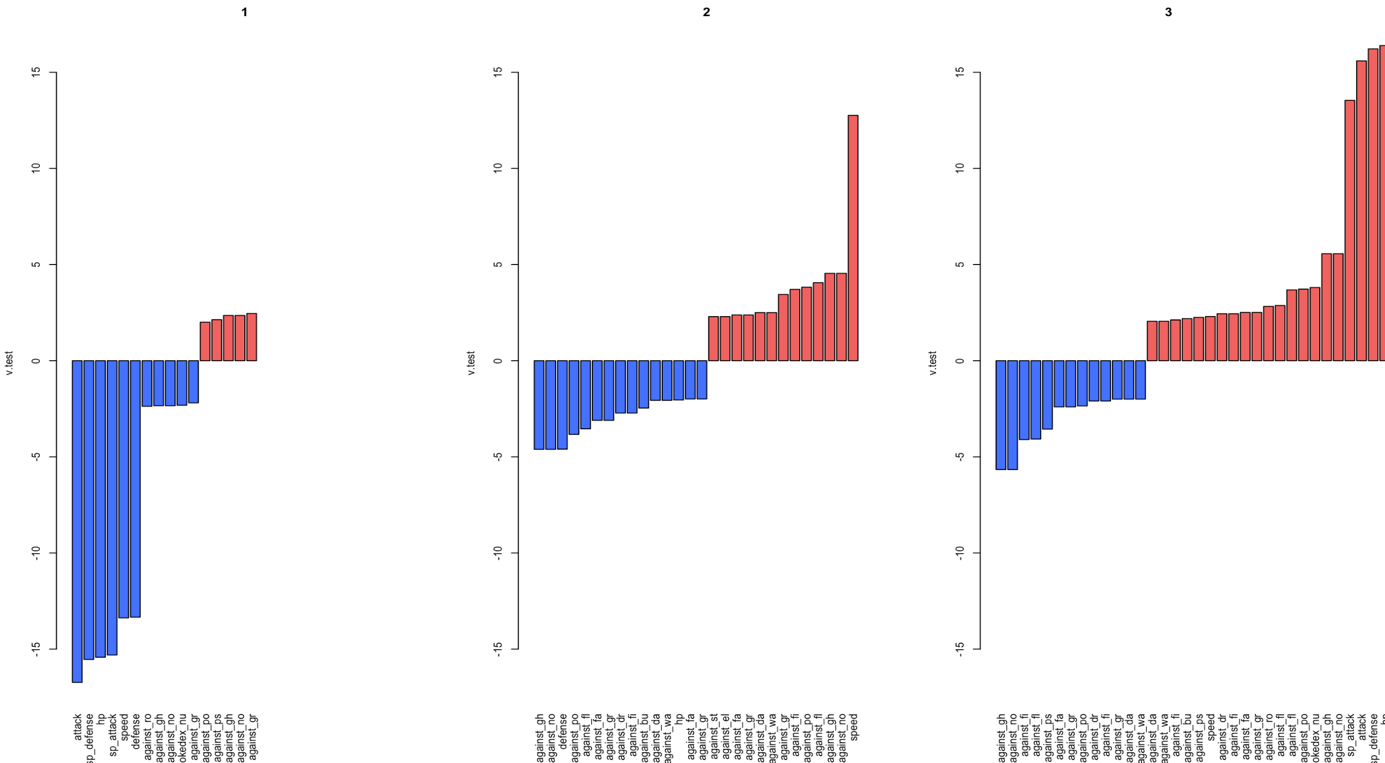
On distingue donc les trois groupes suivants ainsi que leurs caractéristiques :



1. CAH AVEC QUANTITATIVES + AGAINST

Pour cette partie, on utilise le même procédé que la partie précédente avec les mêmes variables quantitatives, mais en utilisant les variables « against » transformées en facteurs. À la suite de la création de notre matrice de distances finale pour ces variables ainsi que l’utilisation de la méthode de Ward, on a également choisi k=3 groupes.

On distingue donc les trois groupes suivants ainsi que leurs caractéristiques :

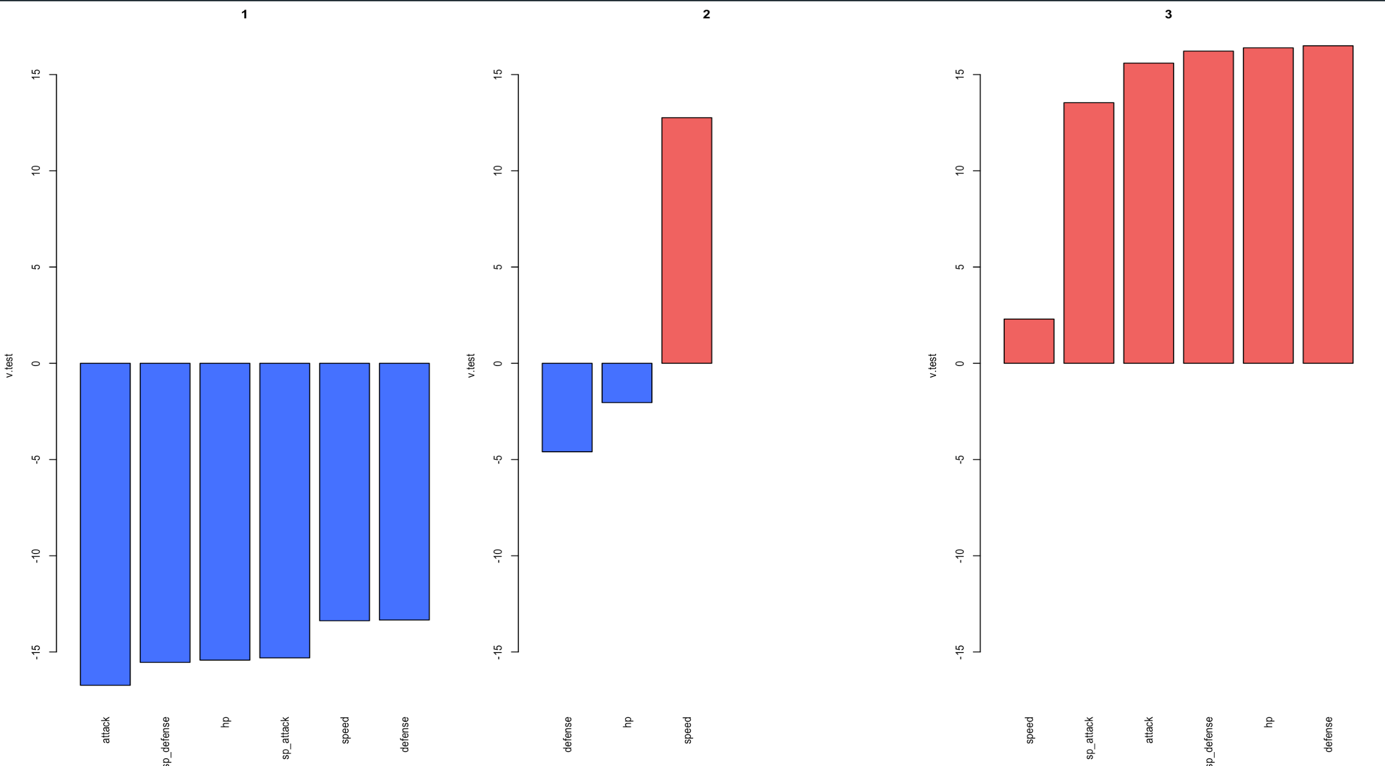


1. CAH AVEC QUE LES QUANTI

Ici, on ne s’intéresse qu’aux statistiques des Pokémons (les 6 variables présentées plus haut).

En utilisant la méthode de Ward on choisit k=3 groupes.

On distingue donc les trois groupes suivants ainsi que leurs caractéristiques :



1. KMEANS sur les variables quantitatives (3 groupes)

