

Visualisation de données multivariées: réimplémentation des fonctionnalités graphiques du package ade4

A. Julien-Laferrière et **S. Dray**

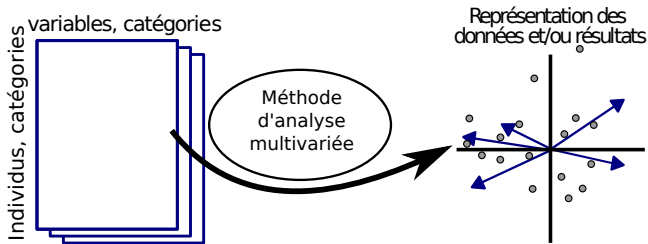
Laboratoire de biométrie et biologie évolutive (UMR CNRS 5558)

CNRS - Université Lyon 1

1 juillet 2012

ade4

Package R pour l'analyse de un à K-tableaux de données.

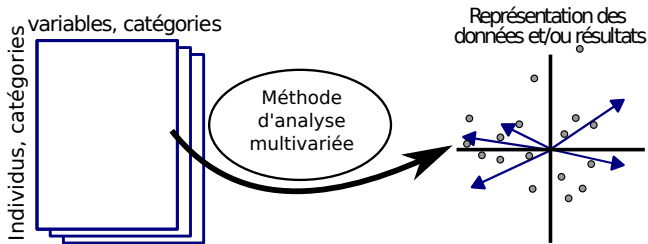


Automatisation des collectes de données → données massives.

Besoin de méthodes d'analyse et de représentation des données adaptées.

ade4

Package R pour l'analyse de un à K-tableaux de données.



Automatisation des collectes de données → données massives.

Besoin de méthodes d'analyse et de représentation des données adaptées.

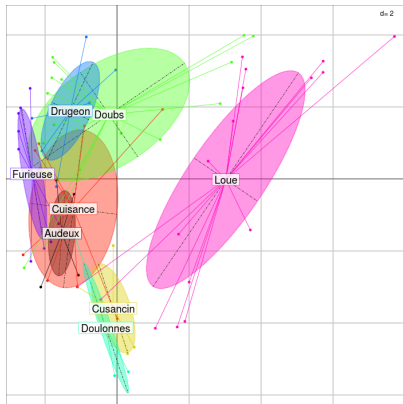
Objectifs:

- Création de nouvelles représentations pour de gros volumes de données
- Flexibilité des graphiques et souplesses de représentation

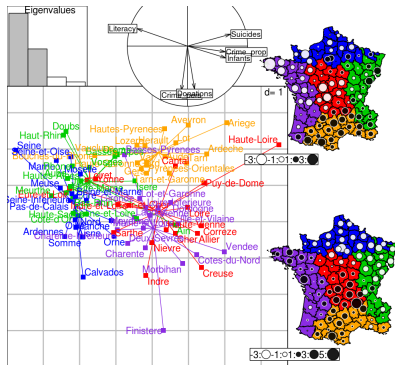
Choix de développement

- Programmation objet S4
- Utilisation du package *lattice*
 - Thèmes graphiques (N&B, palette de couleurs, etc)
 - Coordination possible avec d'autres packages tel *phylobase*

Deux grandes classes d'objets:



adeG: une seule méthode de représentation des données



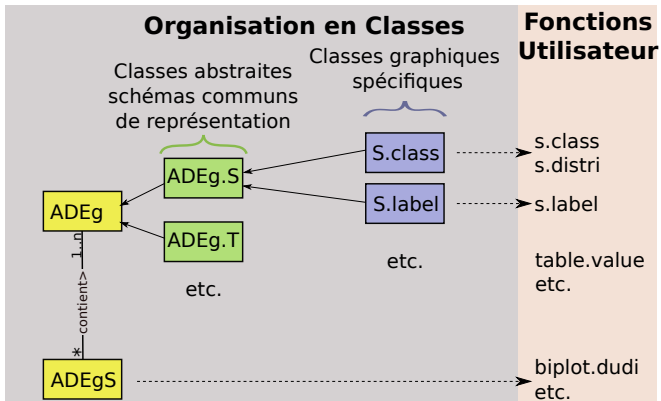
adeGs : superposition/juxtaposition d'adeG

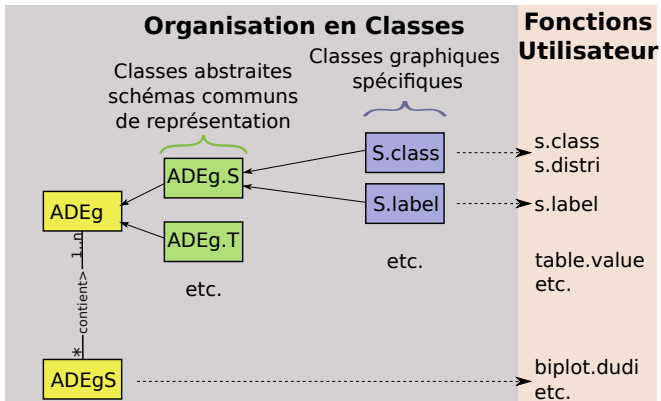
En pratique:

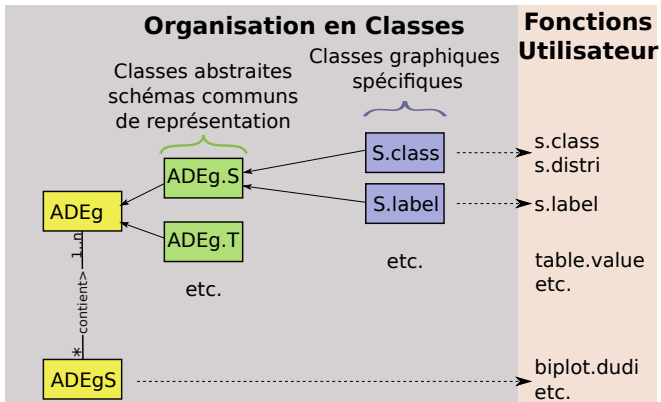
Chaque graphique → un objet **S4** .

Manipulation d'**objets** graphiques:

- A la création
 - Création objet, stockage des données et paramètres
 - Calculs préliminaires
 - Préparation de l'appel graphique
- Manipulations possibles *a posteriori*:
 - Affichage
 - Modifications de paramètres graphiques
 - Superposition ou juxtaposition d'un ou plusieurs autres graphiques



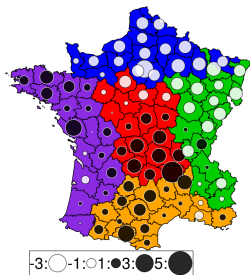




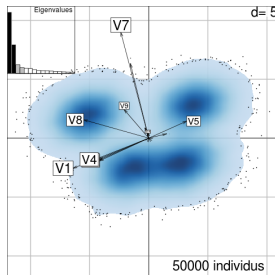
Hiéarchisation des graphiques et programmation orientée objet

- Modularité des représentations graphiques
- Ajout de nouvelles méthodes de représentation graphique facilité

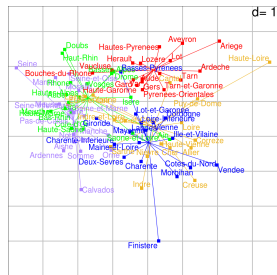
Quelques représentations possibles



(a) Utilisation de données spatiales (package sp)



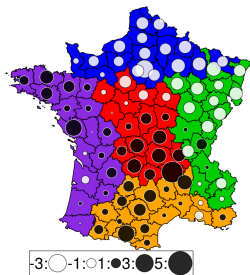
(b) Aide à l'affichage de données massives



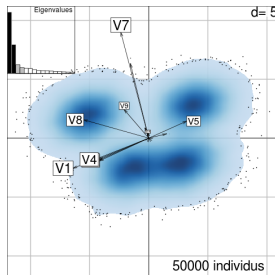
(c) Optimisation du placement des labels

Personnalisations facilitées par un grand nombre de paramètres graphiques

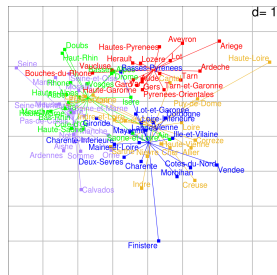
Quelques représentations possibles



(a) Utilisation de données spatiales (package sp)



(b) Aide à l'affichage de données massives

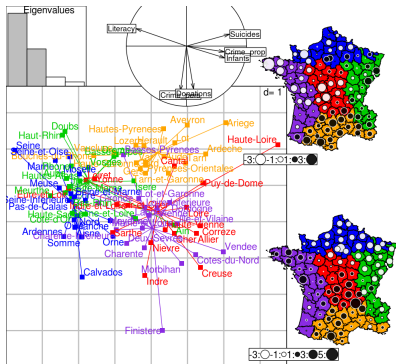


(c) Optimisation du placement des labels

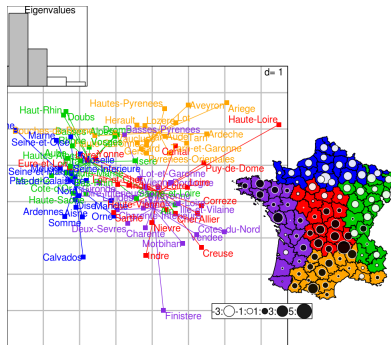
Personnalisations facilitées par un grand nombre de paramètres graphiques

Manipulation d'ADEgS

print(analysis)



print(analysis[1:4], newpos)



Perspectives

- En cours d'implémentation
- Développement de thèmes multiples
- Multipanel ("facettes") pour des groupes d'individus/variables
- Ajout d'appel avec des formules
- Graphiques interactifs
- Nouvelles représentations graphiques