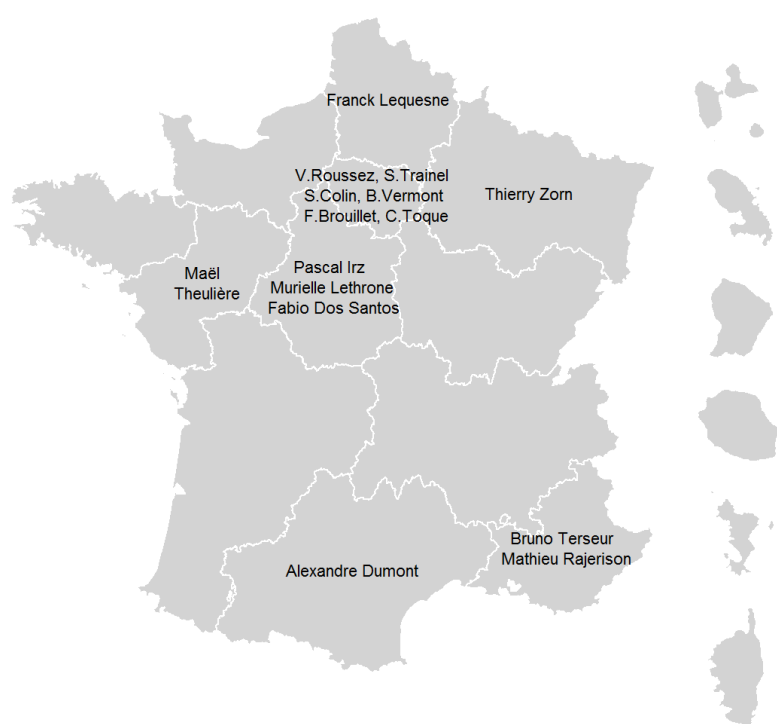


Chapitre 2 R, son écosystème et ses possibilités

2.1 Le groupe de référents

- Un groupe pour structurer une offre de formations sur R
- Un réseau d'entraide



Pour rejoindre le groupe des utilisateurs du ministère, vous pouvez écrire à la liste SYMPA : labo.communaute-r@developpement-durable.gouv.fr

2.2 R c'est quoi ?

- Un langage de programmation interprété exécutable pas à pas → lignes de code
- Un logiciel interactif permettant de traiter divers types de données

2.3 Un peu d'histoire

- A l'origine existait le langage S, servant à programmer avec des données (*Bell laboratory*, années 70)
- Au milieu des années 90, mélange de S et du langage Lisp pour le dédier à l'analyse de données et aux graphiques → création de R. Ihaka R, Gentleman R (1996) R: A Language for Data Analysis and Graphics. *Journal of Computational and Graphical Statistics*, 5(3), 299–314
- S'est d'abord développé dans le monde académique, puis a essaimé vers d'autres domaines professionnels, voire associatifs
- Système toujours d'actualité - Mise à jour tous les 6 mois
- Avenir dans la statistique publique : Adopté comme standard par l'Insee (à l'horizon 2025) et divers SSM. Des packages pour :
 - charger les données sdmx de l'Insee,
 - charger les formats des organisations internationales (Eurostat, OCDE, ONU, etc.),
 - gérer les millésimes du référentiel communal, etc.

2.4 Un logiciel libre

- Disponible en libre téléchargement sur [le site officiel du CRAN](#)
- Installable sur la plupart des systèmes d'exploitation
- Utilisé en recherche, en enseignement et en entreprise
- Une communauté très active :
 - Forum des [utilisateurs de R en français](#)
 - Communauté française des utilisateurs de R (avec des packages spécifiques, qui permettent de gérer les changements de géographie communale par exemple) : [Frrrenchies](#)
 - Slack de ces utilisateurs (pour les questions/réponses, actualités...) [Slack](#)
 - Forum des utilisateurs de R en français : <http://forums.cirad.fr/logiciel-R/>
 - R-bloggers : <https://www.r-bloggers.com/>
 - Stackoverflow : <https://stackoverflow.com/questions/tagged/r>

C'est un logiciel libre : les utilisateurs ont la liberté d'exécuter, copier, distribuer, étudier, modifier et améliorer ce logiciel. R fait partie de la "galaxie" GNU (GNU's Not Unix). R est donc libre, gratuit et multiplateforme.

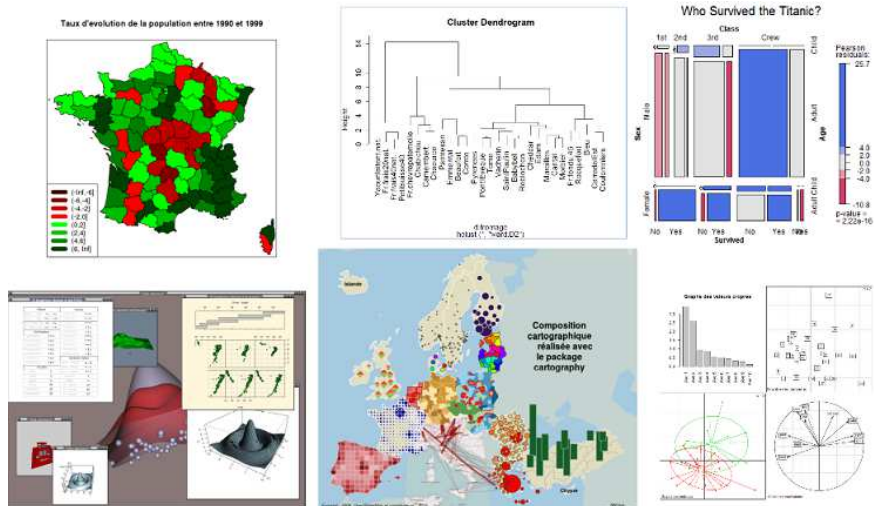
Le site [CRAN = Comprehensive R Archive Network](#) est le site de référence pour R. Il contient non seulement les installateurs du logiciel mais également de nombreuses ressources : documentations, FAQ, tutoriels...

2.5 Que peut-on faire avec R ?

- Lire des données de formats divers → attention, il peut être gourmand en ressources
- Les manipuler (prétraitements)
- Les enrichir à partir de données externes
- Les analyser, les modéliser
- Présenter les résultats (tableaux, graphiques, cartes)
- Publier ...

... et ce au moyen d'une palette de méthodes extrêmement diversifiée.

Privilégie la flexibilité et le découpage des traitements en parties simples.



2.6 Un logiciel modulaire

- Le package `base` contient nombreuses fonctionnalités
- Extensible au moyen de **packages** (1 package = 1 ensemble de nouvelles fonctions)
- Nombreuses extensions spécialisées
- Proposition d'une version installable de R + ensemble de packages utiles

De base, R permet déjà de faire un très grand nombre de choses avec son socle commun : le package `base`. Pour afficher la liste des fonctions disponibles dans le package `base` :

```
library(help = 'base')
```

Sur ce package `base` vient se greffer un ensemble de packages (plugins, modules) complémentaires. Un package est une bibliothèque de fonctions. La plupart de ceux qui font référence sont mis à disposition de tous sur le dépôt CRAN.

Cela permet d'étendre à l'infini les possibilités de R : celles-ci sont seulement limitées par la volonté de développement des utilisateurs.

De même que les packages sont dépendants de fonctions du module principal `base`, certains packages peuvent être dépendants de fonctionnalités d'autres packages → notion d'arbre des dépendances (géré par R de façon transparente).

Structure modulaire → il existe de multiples façons d'effectuer une même tâche. Il suffit d'une méthode maîtrisée pour aller au résultat.

Les fonctionnalités utilisables dans R, sont en perpétuelle évolution et permettent d'intégrer les méthodes les plus récentes dans des domaines qui évoluent rapidement : modélisation, analyse spatiale, dataviz,...

La proposition du groupe de référents est de fournir une version installable de R intégrant un ensemble de packages estimés indispensables pour travailler avec ce logiciel. Vous pouvez trouver cette version à [cette adresse](#)

2.7 Des exemples de ce qu'on peut faire avec R :

- Le support de cette formation
- [Visualisation du compte du logement](#)
- [Les pesticides dans les cours d'eau des Pays de la Loire](#)
- [Des exemples de ce qu'on peut faire avec shiny](#)