

# Le logiciel de statistiques



PRÉFET  
DE LA RÉGION  
GRAND EST

*Liberté  
Égalité  
Fraternité*

et son utilisation au sein  
de la DREAL Grand Est



# Sommaire

- Généralités – historique
- Logiciel modulaire
- Qui utilise R ?
- Que permet R ?
- Utilisation de R au sein de la DREAL Grand Est



# Logiciel modulaire

- R est un logiciel modulaire :
  - Il existe un grand nombre de « packages » qui ajoutent des fonctionnalités très diverses :  
géocodage d'adresses avec {BanR},  
cartographie dynamique avec {leaflet}...
  - Ce qui peut expliquer son utilisation dans des domaines très variés

# Qui utilise R ?

- Longtemps restreint au domaine de la recherche, R s'est « démocratisé » dans les années 2010,
  - entre autres suite à l'apparition de RStudio, un EDI (environnement de développement intégré) puissant et convivial
  - R peut être utilisé depuis Jupyter (EDI de Python)
- de plus en plus utilisé dans les entreprises privées, les banques, les assurances, et les organismes publics (INSEE, CEREMA, DDT, DREAL, ....), => orientation ministérielle.

# Que permet R ?

- Grâce à sa structure modulaire, R s'applique à de nombreux domaines et permet :
- des calculs statistiques au sens large :  
(stats descriptives, modèles linéaires, analyses multivariées, analyse textuelle, tests, contrôle de qualité, ML ...)
- représentations graphiques diversifiées :  
<https://r-graph-gallery.com/index.html>
- cartes et analyse spatiale :  
<https://r-graph-gallery.com/map.html>

# Import de données performant

- Import de données de format très variés :  
*CSV, XLS, ODS, SAS, SHP, XML, JSON*
- Fichiers locaux ou téléchargement direct,
- accès à des bases de données (requêtes SQL)  
*PostgreSQL / PostGIS*
- Récupération des données via des API (et autres webservice)
- Moissonnage de données dans des pages internet  
*ex : Le prix du loyer sur LeBonCoin*

# Que permet encore R ?

- Rapports reproductibles Markdown
  - Fichiers de scripts élaborés
  - Inclusion de code R et de texte libre (format Markdown) : commentaires du code, analyse des données
  - Permettent lisibilité et traçabilité d'un traitement de données
  - Évolutifs et adaptables
  - convertibles en HTML, DOCX, PDF
- => Démo

# Que permet encore R ?

- Applications Internet (package Shiny) :  
*exemples : <https://shiny.rstudio.com/gallery/>*
- *Outil portrait mobilité (détaillé plus loin)*

# R au sein de la DREAL Grand Est (1)

- Premiers pas en R en 2016, en autoformation
- Intégration au groupe des référents R national
  - Participation à la création des valises de formation du MTES  
<https://mtes-mct.github.io/parcours-r/>
  - Actions de formations régulières (CVRH)

# R au sein de la DREAL Grand Est (2)

- Outil Portrait Mobilité (OPM)
  - Application Shiny (web) qui génère un portrait lié à la mobilité, sur un territoire donné, avec possibilité de saisir des commentaires personnalisés
  - V1 en 2019, code peu optimisé, données incorporées manuellement,
  - V2 en 2020, code meilleur, développement à 2 (collaboratif, mode Agile), avec scripts d'import facilitant la mise à jour des données

Accès par <https://ssm-ecologie.shinyapps.io/opm2020/>

Code disponible sur <https://gitlab.com/drealge/opm2020>

# R au sein de la DREAL Grand Est (3)

- Portraits d'agglomération : rapports Markdown automatisés (2018)  
<https://www.grand-est.developpement-durable.gouv.fr/portraits-d-agglomerations-a18306.html>
- Perspectives d'évolution de l'OPM
  - Automatisation de la collecte de données de l'application (par API, requêtes SQL, depuis entrepôt de données, ...),
  - Ajout de nouveaux indicateurs
- Développer l'utilisation d'API dans les différents projets à venir
- Développer d'autres outils de *dataviz*

Me  ci

pour votre attention

**Avez vous des questions ?**