

# Data Mining et Data visualisation

Cours 2 : Introduction au package dplyr  
aux cartes leaflet et Rmarkdown

Laurent POLITIS  
politis.laurent@gmail.com  
ESSCA 2023



# **Le package dplyr**

# Le package dplyr

- Une alternative à la manipulation des données
- Son propre lot de fonctions appelées « verbs »
- Classe de variable utilisée : tibble et non data frame
- Considéré plus adapté aux études statistiques

<https://cran.r-project.org/web/packages/dplyr/vignettes/dplyr.html>

# **LES CARTES**

# Les cartes

- Les cartes sont utiles si vous avez des données géolocalisées
- Pour les metadata (localisation des points d'intérêt, rayon de l'étude) ou données ponctuelles facilement lisibles, le package leaflet est idéal et produit des cartes interactives

Quelques exemples : <https://rstudio.github.io/leaflet/>

- Pour des données « continues », très denses spatialement, il y a d'autres méthodes : <https://cran.r-project.org/doc/contrib/intro-spatial-rl.pdf>

# Rmarkdown

# Les rapports Rmarkdown

- Langage Markdown de base (HTML simplifié)
- Possibilité d'intégrer des éléments issus de calculs R, parfois interactifs, comme les cartes Leaflet mais aussi tableaux, plots, images...
- Rapport HTML ou PDF immuable, transférable par mail, ou posté sur internet et personnalisable à souhait

<https://openclassrooms.com/fr/courses/1304236-redigez-en-markdown>

<https://rmarkdown.rstudio.com/lesson-1.html>