

# Analyse et Traitement Avancé des Données avec R

PEHAN Boré Hermann

## Introduction à R et Rstudio

### R

- Logiciel open-source
- Langage orienté vers le traitement, l'analyse des données et que la visualisation des données
- Rédaction de rapport (word, powerpoint, pdf et html)
- Installation
  - Sous windows: [Télécharger ici](#)
  - Sous Mac OS X: [Télécharger ici](#)

### Rstudio

- Environnement de développement intégré pour R

Installation: [Télécharger ici](#)

### Importation de fichier

#### Importation des fichiers textes (.txt,.csv)

#### Importation graphique

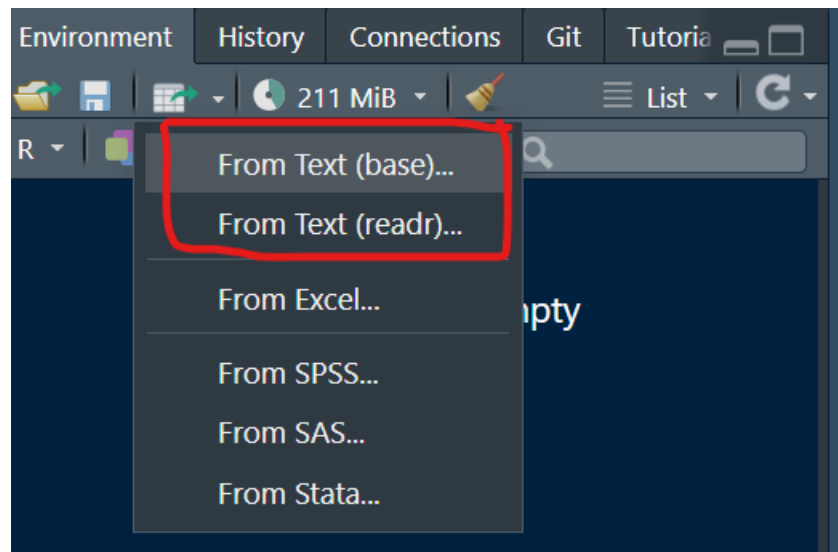


Figure 1: Charger un fichier texte par l'interface

## Utilisation d'un script

```
library(readr)
table_texte <- read_delim("data/exemple_texte_table.txt", delim = "\t", quote = "")
communes <- read_csv("data/communes-01012019.csv")
```

## Importation des fichiers excel (.xls ou .xlsx)

Pour l'importation des fichiers excel, on peut utiliser les packages `readxl` ou `xlsx`

## Introduction à la gestion avancée des données avec R

### Lecture des données

### Concepts avancés de manipulation des données avec tidyverse et data.table

### Concepts avancés de manipulation des données avec tidyverse

### Concepts avancés de manipulation des données avec data.table

**Optimisation du traitement des grands ensembles de données**

**Programmation fonctionnelle et manipulation efficace des données**

**Analyse exploratoire et modélisation**

**Techniques avancées de visualisation avec ggplot2 et plotly**

**Méthodes statistiques et modélisation prédictive avec caret et randomForest**

**Analyse des séries temporelles avec forecast**

**Traitement des données et automatisation**

**Automatisation du nettoyage des données avec tidymodels**

**Optimisation des performances d'analyse avec parallel et bigmemory**

**Interaction entre R et Python avec reticulate**