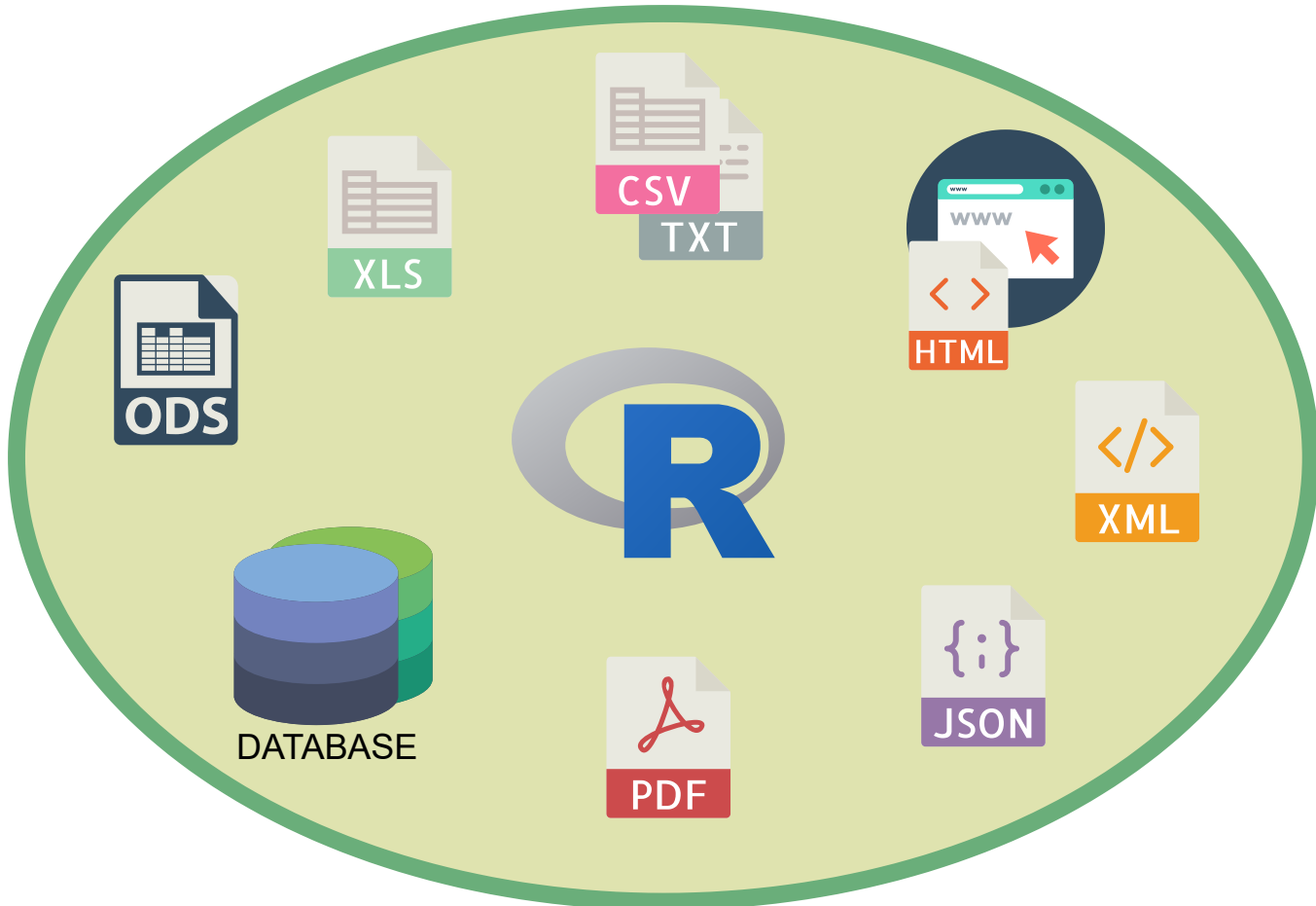


Importation de données ou le choix du format

Benjamin Louis

15/10/2019 (MàJ: 18/11/2021)

Une multitude de formats



+ shapefiles, images, API, SPSS, Stata ...

Fichiers text (.csv)



Importation avec



➔ **read_delim()** (base R : **read.delim()** ou **read.table()**)

```
read_delim(file = "chemin/vers/fichier.csv", delim = ";",  
            na = c("", "NA"), col_types = NULL,  
            locale = locale(decimal.mark = "."))
```

➔ **Autres fonctions avec des valeurs fixes pour `delim`**

- **read_csv** pour **delim = ","** (base R : **read.csv()**)
- **read_csv2** pour **delim = ";"** (base R : **read.csv2()**)
- **read_tsv** pour **delim = "\t"** (tabulation)

Importation avec



➔ Argument `col_types` =

- `NULL` : le type des colonnes est deviné
- `cols(nom_col1 = col_double(), nom_col2 = col_date(), ...)` :
il faut préciser pour **toutes** les colonnes !
- `cols_only(nom_col1 = col_double(), nom_col2 = col_date())` :
il faut préciser pour **certaines** colonnes !
- `c("d", "c", "?", "-", ...)` il faut préciser pour **toutes** les colonnes !

```
# Pour plus d'information
?read_delim
?cols
?cols_only
vignette("readr")
```

Fichiers excel (.xls, .xlsx)



Importation avec



➔ `read_excel()`

```
read_excel(path = "chemin/vers/fichier.xls", sheet = NULL,  
           range = NULL, na = "", col_types = NULL)
```

- **sheet** : nom de la feuille (e.g. **sheet = "feuille1"**)
- **range** : plage des cellules (e.g. **range = "B3:D87"**)

➔ Autres fonctions si on connaît le format

- **read_xls** si l'extension est **.xls**
- **read_xlsx** si l'extension est **.xlsx**

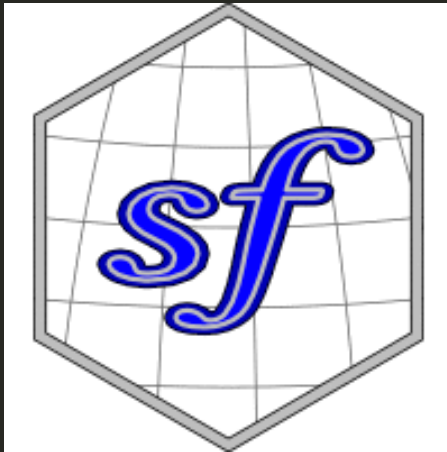
Importation avec

➔ Argument `col_types` =

- `NULL` : le type des colonnes est deviné
- `c("numeric", "text", "skip", "guess" ...)` il faut préciser pour **toutes** les colonnes !

```
# Pour plus d'information  
?read_excel
```


Shapefiles (.shp & co), geopackage (.gpkg)



Importation avec

➔ `st_read()`

```
read_excel(dsn = "chemin/vers/fichier.shp")  
read_excel(dsn = "chemin/vers/fichier.gpkg")
```

- `geometry_column` : nom de la colonne des coordonnées (si plusieurs)

➔ Attention !

- Pour les shapefiles, le fichier .shx au moins doit être présent mais il est recommandé d'avoir tous les autres (.prj, .dat, fichier de données)
- Fonctions utiles : `st_zm()`, `st_transform()`

Autres formats

Autres formats

readODS : <https://cran.r-project.org/web/packages/readODS/index.html>



(DBI)

<https://dbplyr.tidyverse.org/>

pdftools <https://docs.ropensci.org/pdftools/>



<https://rvest.tidyverse.org/>

jsonlite <https://github.com/jeroen/jsonlite>

...

Crédits

Files icons (except ods) made by **Smashicons** from www.flaticon.com

Webpage icon made by **Vectors Market** from www.flaticon.com

ods icon made by **Freepik** from www.flaticon.com