

# Les frictions de la cohabitation R-Excel

Swann Floc'hlay<sup>1</sup>

## Résumé

Vous rappelez-vous vos premières commandes R ? Que ça remonte à l'époque de Great Pumpkin ou à hier soir, on est nombreux à avoir depuis délaissé Excel pour le tidyverse. Sauf que, quand il s'agit de partager nos ressources avec des collègues non initiés, c'est parfois compliqué. La dernière friction en date, c'est l'adoption du format parquet. Une aubaine pour nous autres codeurs, un casse-tête pour celui qui redoute encore le chiffre 1 048 576<sup>2</sup>. Afin de faciliter cette transition, nous proposons une application Shiny pour explorer en autonomie et sans connaissance R préalable les ressources à haute volumétrie de l'Insee. Cette application vise à être didactique, robuste et flexible aux différents usages utilisateurs.

**Mots-clefs (3 à 5) :** Shiny - Parquet - Accessibilité - Collectivité - REX

## Développement

Cette soumission aborde les enjeux liés à la collaboration d'une équipe mixte Excel-R. Avec la généralisation des données volumineuses dans les instituts publics (e.g. catalogues Insee et Cerema), les collectivités se tournent vers les outils programmatiques. En parallèle, le format parquet se démocratise rapidement et favorise également la transition d'Excel à R.

Cette transition n'est cependant pas instantanée, ni adaptée à tous les corps de métier. Il est donc intéressant de comprendre les enjeux d'une cohabitation d'usage Excel-R, et les leviers possibles de transfert entre ces deux outils.

Pour éclairer ce sujet, nous présenterons un cas d'usage concret concernant l'analyse des données Insee de migrations résidentielles. Ce jeu de données décrit les flux de résidents entre chaque commune du territoire. Au total, il comporte plus de 12 millions de lignes, ce qui le rend difficilement exploitable par une approche tableur classique. Afin de permettre l'exploration et l'extraction de données par tous les agents de la collectivité, nous avons développé une application Shiny servant d'interface entre l'utilisateur et le fichier parquet.

L'objectif de cette application est double :

- créer des extractions de données exploitables par Excel
- réaliser des agrégations simples à la volée sur l'ensemble du fichier parquet

À terme, nous souhaitons élargir les usages de cette application sur d'autres ressources de données ouvertes.

---

<sup>1</sup>Département de Loire-Atlantique, Direction Prospective, swann.flochlay@loire-atlantique.fr

<sup>2</sup>Nombre de lignes maximales sur une feuille de calcul Excel