

# Outils numériques de recherche

Adrien Cloutier

19 septembre 2024

# Section 1

## Les langages de programmation

# Les langages de programmation



## Section 2

### Retour sur le cours 2

# Le logiciel libre et le code source ouvert

- 1 Quelle est la différence entre logiciel libre et le code source ouvert?
- 2 Sont-ce des outils?
- 3 Des exemples de logiciels libres?
- 4 Méthodes vs outils?
- 5 Qu'est-ce que la License publique général (GNU)?

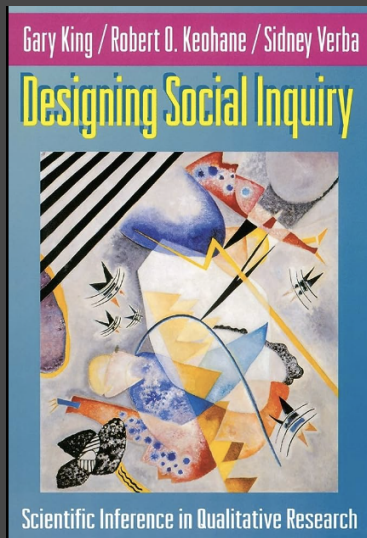
# License publique général (GNU)

- ❶ Création en 1983 par Stallman.
- ❷ Permet l'arrivée de Firefox, OpenOffice, Linux. . .
- ❸ Valeur: entraide, collaboration, partage.
- ❹ Pour protéger les droits des utilisateurs, pas les droits des auteurs.

# Citation inspirante

« Seulement de cette façon, l'utilisateur peut contrôler le logiciel, plutôt que de se faire contrôler par ce dernier (Stallman, 1986). »

# Designing Social Inquiry, King et al. (2021)

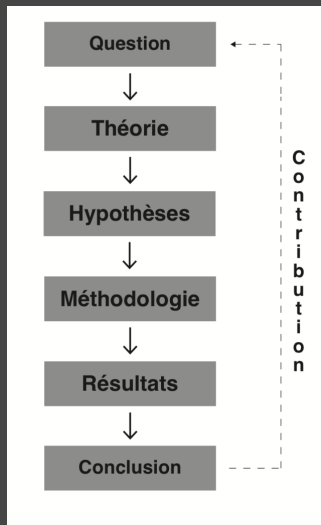




# 4 critères de la recherche scientifique selon KKV

- ❶ Le but est l'inférence;
- ❷ Les procédures sont publiques;
- ❸ Les conclusions sont incertaines;
- ❹ Le contenu est la méthode.

# Cours basé sur le cycle de la recherche



# Avant la recherche

- ❶ « Dépôt » GitHub pour rendre son code accessible;
- ❷ Enregistrement des hypothèses de recherche.

# Pendant la recherche

- 1 Cueillette de données (R)
- 2 Analyse de données (R)
- 3 Visualisation de données (R)
- 4 Rédaction (LaTeX, Zotero)

- ❶ « Open science » (science ouverte)
  - Partage des données
  - Article accessible gratuitement à tous

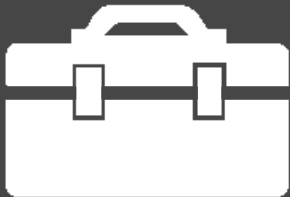
# Critères de sélection des outils du cours

- ➊ Accessibilité (Gratuit ou peu dispendieux)
- ➋ Existence d'une communauté d'utilisateurs
- ➌ Popularité dans le champ
- ➍ Compatibilité avec d'autres outils
- ➎ Transparence et répliquabilité
- ➏ Adaptabilité et flexibilité

## Section 3

# R comme outil de recherche

# R comme outil de recherche





# Datacamp (5%!)

The screenshot displays the Datacamp website interface. At the top, the navigation bar includes the Datacamp logo and links for Home, Learn, Workspace, Certification, Jobs, and Groups. A search bar is located on the right. The left sidebar contains a 'Progress' section with links to Progress, Bookmarks, Leaderboard, and Assignments. Below this is the 'LEARN' section with links to Tracks, Courses, Practice, Assessments, and Tutorials (marked as 'NEW'). The 'APPLY' section includes links to Projects, Case Studies, Code Alongs, and Competitions. At the bottom of the sidebar are 'POPULAR TOPICS' (marked as 'NEW') for Artificial Intelligence and Data Engineering, and a 'Getting Started (0/4)' button.

The main content area features a dark blue header for the 'Introduction to R' course, labeled as an 'INTERACTIVE COURSE'. It includes a 'Start Course' button and a 'Bookmark' button. Below the header, course details are listed: 'Beginner', '4 hours', '0 videos', '62 exercises', '2,631,849 participants', and '6200 XP'. The 'Course Description' section states that the course covers the basics of R, including vectors, factors, lists, and data frames. A 'Read More' link is provided. The 'PREREQUISITES' section indicates there are no prerequisites. The 'PART OF THESE TRACKS' section lists 'Data Analyst', 'Data Scientist', and 'Data Scientist Professional'. A 'Start Chapter' button is visible. A chat bubble from 'Adrien' is present, offering assistance. A 'Create Course Notes' button is also shown.

# Qu'est-ce?

- 1 Une fonction?
- 2 Un package?
- 3 base R?
- 4 Tidyverse?
- 5 La différence entre R et RStudio?

Posez beaucoup de questions!!

# Les logiciels statistiques

- SPSS (1968)
- SAS (1976)
- STATA (1985)

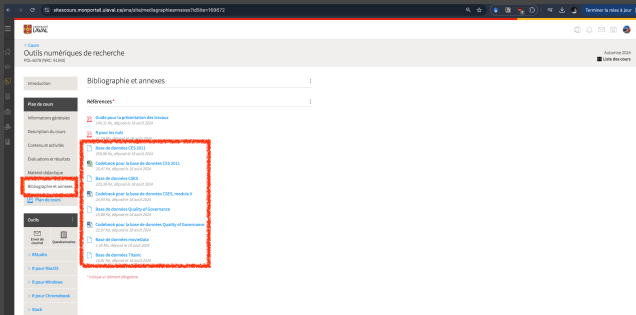
# Émergence du langage R

- Avec la philosophie du logiciel libre et de l'open source
- Université d'Auckland, Nouvelle-Zélande
- Construit sur le langage S
- Gratuit, flexible, extensible

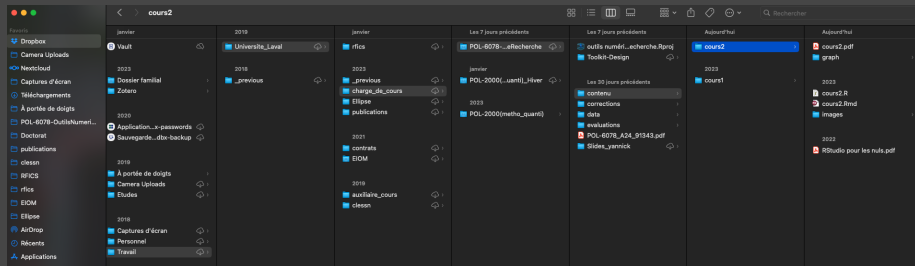
- Comprehensive R Archive Network (CRAN)
- Entrepôt de packages
- Politique stricte

## Télécharger les données

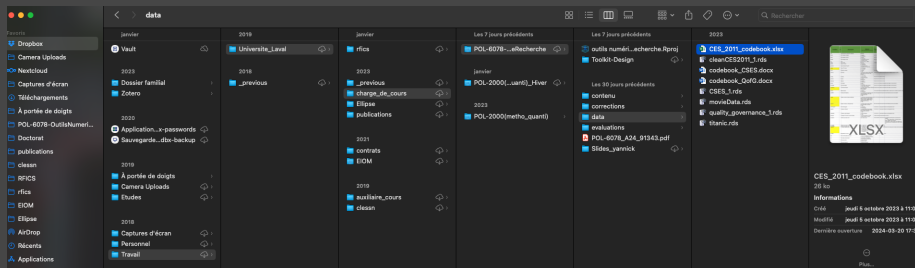
- 1 Créez un nouveau dossier nommé « POL6078 » sur votre ordinateur (je conseille de toujours tout stocker dans un cloud, comme Dropbox)
- 2 Créez un sous-dossier nommé « données » ou « data » (toujours mieux sans accents et sans espace)
- 3 Téléchargez movieData.rds, et autres rds (et les codebooks!)
- 4 Une fois téléchargés, placez-les dans le sous-dossier « data »



# Chemin d'arborescence (le fameux « path »)



# Stocker ses données





# Qu'est-ce qu'un codebook?

- Le « dictionnaire » qui comprend toutes les informations sur le jeu de données
- ESSENTIEL à consulter et à bien comprendre pour un aperçu rapide des données

# Vos meilleurs amis

- Google
- Stackoverflow
- r-bloggers
- DataCamp
- Chaines Slack r\_et\_datacamp
- ... ChatGPT..!

# Prochain cours

- Outils de gestion de projet
- Datacamp: Introduction to R: Data frames + Lists
- Installation de Notion.so!
- Création de votre compte GitHub!
- Lectures:
  - Chapitre 2: Les langages de programmation
  - R for Data Science de Hadley Wickam, chapitre 4 (Workflow: code style)

# On ouvre RStudio

The screenshot shows the RStudio environment with a presentation file open. The source editor on the left contains R code for a presentation, including comments in French and image inclusions. The right pane shows the presentation preview, which is currently empty. The bottom pane shows the console with the command to process the presentation file.

```
547  
548 - ## L'ordinateur  
549  
550 \begin{center}  
551 \includegraphics[width=0.6\textwidth]{images/annexe_ordi.png}  
552 \end{center}  
553  
554  
555 - ## Logiciel libre et code source ouvert  
556  
557  
558 \begin{center}  
559 \includegraphics[width=0.6\textwidth]{images/chapitre1.png}  
560 \end{center>  
561  
562  
563 - ## Panique?  
564  
565 - Un cours à la fois  
566 - Je m'occupe de rendre ça concret  
567 - En échange: lecture + relecture + Datacamp + R + installations + pratique  
568 - Autodidaxie et indépendance  
569  
570  
571 - ## À la semaine prochaine!  
572  
573 \begin{center}  
574 \includegraphics[width=0.6\textwidth]{images/rstudio.png}  
575 \end{center>  
576  
577
```

Environment is empty

Files Plots Packages Help Viewer Presentation

Name	Size	Modifié
previous		
Rhistory	0 B	May 3,
cours1.Rmd	19.1 KB	Sep 5,
images		
cours1.pdf	17.4 MB	Sep 5,

processing file: cours1.Rmd