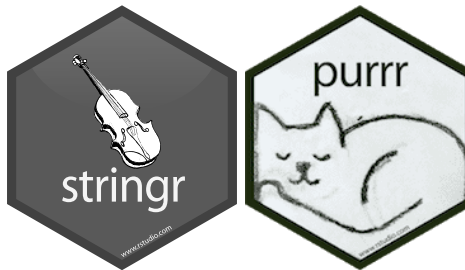


Chaînes de caractères et fichiers

La manipulation de chaînes de caractères est une compétence essentielle pour nettoyer et préparer des données avec R. Cette formation détaille les techniques usuelles avec 'stringr' et 'stringi', approfondit les expressions régulières, puis aborde la gestion des chemins et le traitement de fichiers en série.

Technologies et packages utilisés :



À PROPOS DE LA FORMATION

Formation disponible à distance et dans vos locaux.

Nous pouvons adapter la formation pour les besoins spécifiques de votre organisation.

Durée : 1j — 7h

Prix : 600 € HT par participant en inter-entreprise

Accessibilité aux personnes handicapées : Pour toute inscription de personnes en situation de handicap, il convient de nous prévenir. Nous vous proposerons un entretien afin de prendre en compte les spécificités de votre handicap et de déterminer les éventuels aménagements ou dispositions à mettre en place.

Effectifs par groupe : de 3 à 8 participants.

Horaires : En présentiel et en classe à distance, les cours ont lieu de 9h30 à 12h30 et de 13h30 à 17h30. Les participants sont accueillis à partir de 9h15.

Accessibilité et délais d'accès : Vous pouvez vous inscrire jusqu'à un jour ouvré avant la formation, il est toutefois recommandé de le faire au moins 1 mois avant pour avoir le temps de contracter ou de valider le financement avec votre organisme de financement.

PUBLIC CONCERNÉ

- » Les utilisateurs du tidyverse qui souhaitent maîtriser la manipulation de chaînes de caractères
- » Les utilisateurs de R qui doivent nettoyer du texte et traiter des fichiers en série

PRÉREQUIS

Il est demandé de connaître R pour suivre cette formation.

- » connaître les bases du langage R
- » savoir manipuler des données avec le tidyverse

OBJECTIFS PÉDAGOGIQUES

À l'issue de la formation, le participant sera en mesure de :

- » Produire et composer des chaînes de caractères
- » Transformer et nettoyer du texte
- » Utiliser les expressions régulières
- » Gérer les chemins et traiter des fichiers en série

PROGRAMME DE LA FORMATION

Cette formation s'adresse aux utilisateurs du tidyverse qui souhaitent maîtriser les techniques de manipulation de chaînes de caractères avec R. Les techniques les plus fréquentes sont détaillées, puis approfondies avec les expressions régulières. Enfin, la gestion des chemins et le traitement de fichiers en série sont abordés.

Le plan de formation est détaillé ci-dessous :

- » Les packages de gestion des chaînes de caractères
 - » base R, 'stringi' et 'stringr'
- » Produire des chaînes de caractères
 - » Concaténation et composition ('paste', 'glue')
 - » Formatage des valeurs numériques et des dates ('formatC', 'sprintf', 'format')
- » Transformer des chaînes de caractères
 - » Découpage, casse, nettoyage des espaces et des accents
 - » Normalisation des noms de variables
 - » Lire et écrire du texte ('readLines', 'writeLines')
- » Les expressions régulières
 - » Identifier, extraire, localiser, remplacer et découper
 - » Correspondance des fonctions base R et 'stringr'
 - » Syntaxe : opérateurs et classes de caractères
- » Gérer les chemins et les fichiers
 - » Chemins absolus, relatifs et UNC ; le package 'here'

- » Lister et inspecter les fichiers ('list.files', vérification d'existence)
- » Traitement de fichiers en série (imports / exports multiples, 'purrr', zip)

MODALITÉS

Modalités pédagogiques

Notre formation s'articule sur une alternance d'apports théoriques et de pratiques :

- » Exposés théoriques réalisés par le formateur
- » Apport d'expériences de le formateur
- » Echanges et retours d'expérience entre les participants.
- » Exercices et ateliers d'application permettant aux stagiaires de pratiquer concrètement.

Modalités d'évaluation initiale

Une évaluation préalable à l'entrée en formation est réalisée dès la demande d'inscription afin d'identifier les connaissances et les besoins de chaque futur participant.

Modalités d'évaluation finale

- » Évaluation sous forme d'autoévaluation
- » Questionnaire d'évaluation
- » Test de positionnement en début et fin de la formation pour valider les compétences acquises

Documentation remise aux stagiaires

- » Les supports de cours au format PDF
- » Ces mêmes supports contiennent les exercices et leurs corrections

Outils utilisés lors de la formation

- » R et Rstudio
- » Navigateur moderne (Chrome, FireFox, Edge, Safari)