### **ThinkR**

50 rue Arthur Rimbaud 93300 Aubervilliers Email: digiforma@thinkr.fr





# R Niveau 1 - Utilisateur - Analyse de données - 100% distanciel

Certification du répertoire spécifique de France Compétences RS6193 - Analyse statistique de données avec le langage R

Le titulaire de la certification sera capable, dans une démarche de data science, d'intégrer et de traiter des données, et de produire un rapport en utilisant de manière appropriée le langage de programmation R.

Le titulaire de la certification sera capable de mettre en place un processus de traitement des données sous R qui soit reproductible avec différents type de jeux de données

**Durée:** 35.00 heures (5.00 jours)

#### Profils des stagiaires

- Utilisateurs d'Excel pour l'analyse de données
- Utilisateurs de logiciels de statistiques propriétaires
- Utilisateur de R nécessitant une remise à niveau

#### **Prérequis**

- Affinité pour la programmation
- Affinité pour les données
- Aisance informatique

## Objectifs pédagogiques

- Connaître et comprendre l'environnement R
- Importer et exporter des données avec R
- Concevoir des opérations de manipulation de données (filtrer, trier, agréger, transformer...) comme un flux de traitement organisé dans R
- Dessiner le graphique approprié à la représentation souhaitée avec R
- Manipuler les dates dans R
- Manipuler les chaines de caractères dans R
- Maîtriser le paramétrage de rapports mis en forme d'analyse de données avec R

#### Organisation de la formation

#### Moyens pédagogiques et techniques

- Formation 100% en distanciel synchrone (par demi-journées)
- Documents supports de formation projetés
- Mise à disposition en ligne de documents supports à la suite de la formation
- Plateforme d'apprentissage intégrée en ligne
- Exposés théoriques
- Travaux dirigés
- Quiz en salle
- Mise à disposition d'une base de ressources documentaires

### **ThinkR**

50 rue Arthur Rimbaud 93300 Aubervilliers Email: digiforma@thinkr.fr

Tel: 0658744151



#### Dispositif de suivi de l'exécution de l'évaluation des résultats de la formation

- Emargement en ligne
- Évaluations formatives régulières sous forme de QCM
- Évaluations sommatives régulières sous forme de travaux dirigés
- Formulaires d'évaluation de la formation, à chaud et à froid
- Examen certifiant en fin de cursus
- Alternance régulière entre théorie et pratique

#### Contenu de la formation

- Présentation de R, premiers pas dans la console
  - Démystifier la console, une «super-calculatrice» qui exécute du code
  - Prendre ses marques dans RStudio, créer son premier projet
  - Comprendre ce qu'est un vecteur
  - Se familiariser avec les packages et le moyen de les utiliser
  - Utiliser Quarto/Rmarkdown pour la prise de notes
- Visualiser des données
  - Choisir le graphique adapté
  - Construire, modifier et exporter des graphiques avec le package (ggplot2)
- Importer et exporter des données
  - Importer des fichiers plats de type .csv et .xls/.xlsx avec les packages {readr} et {readxl}
  - Vérifier la qualité de l'import
  - Typer les données (déterminer si chaque colonne est une date, une valeur numérique, un texte...)
  - Exporter les données vers les formats .csv, .xlsx et .RDS
- Manipuler et nettoyer des données
  - Se familiariser avec les packages {dplyr} et {tidyr}
  - Concevoir les opérations de manipulation de données comme un flux de traitement organisé
  - Chainer les opération avec l'opérateur %>%
  - Manipuler, transformer et résumer des données avec {dplyr}
  - Remodeler les données sous un format adéquat avec {tidyr}
- Produire des rapports d'analyse
  - Appréhender le concept de recherche reproductible
  - Compiler des rapports commentés imbriqués de code R avec Quarto/Rmarkdown
  - Exporter les rapports en docx, en html et en pdf
  - Mettre en forme les tableaux avec {DT}
  - Paramétrer finement les rapports d'analyse réalisés avec Quarto/Rmarkdown
- Décrire des données
  - Choisir des paramètres descriptifs pertinents pour décrire des données (effectifs, moyennes, médianes, écart-types...)
  - Produire des tableaux de statistique descriptives à l'aide du package {skimr}
  - Manipuler et traiter les données manquantes
- Bonnes pratiques
  - Trouver de l'aide face à un problème rencontré sur R
  - Concevoir un exemple reproductible
  - Respecter les conventions de style
- Réaliser des jointures de données
  - Réaliser une jointure simple
  - Réaliser une jointure complexe
- Manipuler des dates et des chaines de caractères
  - Formater, manipuler et opérer des calculs sur des dates avec le package {lubridate}
  - Comprendre, appréhender et utiliser les formats dates dans R
  - Manipuler les chaînes de caractères avec le package {stringr}

# **ThinkR**

50 rue Arthur Rimbaud 93300 Aubervilliers Email: digiforma@thinkr.fr

Tel: 0658744151



#### Modalité de Certification

Examen en autonomie en temps limité de 2 heures.

Série de questions nécessitant d'utiliser les notions abordées pendant la certification (nettoyage de données, manipulation de données, visualisation de données, etc.).

Rendu sous la forme d'un rapport mis en forme selon des spécifications précises.