

Data science : programmation en langage R

RS5074 · RS · Publiée

L'essentiel

Code(s) NSF	326
Formacode(s)	31026, 30820, 31067, 31052
Type d'enregistrement	Enregistrement sur demande
Date d'enregistrement	27/01/2020
Date d'échéance	27/01/2023

Certificateur(s)

THINKR · SIRET : 81006451900020

Résumé de la certification

OBJECTIFS ET CONTEXTE

Dans une démarche de data science menée par son organisation, le titulaire de la certification est capable de créer une fonction en langage R, sous forme de package, en respectant les bonnes pratiques de développement pour permettre son usage par d'autres utilisateurs dans l'environnement R.

Cette certification permet à des professionnels du traitement de données (Big Data) de créer leurs propres outils dans un environnement et un langage libre de droit.

Pour les développeurs, cette certification permet d'acquérir un nouveau langage et de maîtriser les outils liés au traitement de grands volumes de données (Big Data).

Cette certification permet aux entreprises, organisations et administrations d'accéder à des compétences de traitement et d'exploitation de grands volumes de données sur leurs cœurs de métiers et ce de façon industrielle. Ils peuvent ainsi valoriser leur capital « data » via une usine logicielle et accéder à de nouvelles informations pour la prise de décision stratégique.

COMPÉTENCES ATTESTÉES

1. Identifier l'opportunité de factoriser un ensemble de lignes de codes pour en faire une ou plusieurs fonctions intégrées dans un package en langage R
2. Développer les fonctions d'un package en langage R en respectant les protocoles de versionnage et de documentation, dans un environnement logiciel R adapté, et procéder à des tests fonctionnels pour chaque fonction, afin que le package soit utilisable par des tiers et de faciliter les évolutions futures.
3. Définir et réaliser l'ensemble des tests fonctionnels globaux d'un package développé en langage R (périmètre des tests, écriture des lignes de codes permettant les tests), puis compiler le package afin de permettre sa livraison aux utilisateurs finaux.

Voies d'accès

MODALITÉS D'ACCÈS

Formation continue

PRÉREQUIS À L'ENTRÉE EN FORMATION

Etre capable d'importer, nettoyer, visualiser des données et produire un rapport reproductible dans l'environnement R. Avoir déjà rédigé une ligne de code (dont formule Excel) et connaître les concepts de moyenne, médiane et écart-type.

COMPOSITION DU JURY — FORMATION CONTINUE

Le jury est composé du directeur de l'organisme certificateur, d'un formateur agréé et d'un professionnel spécialisé

dans le traitement des grands volumes de données.

Base légale

Date de décision	27/01/2020
Durée de l'enregistrement	3 ans
Date d'échéance de l'enregistrement	27/01/2023

Liens avec d'autres certifications

NOUVELLE(S) CERTIFICATION(S)
RS6194

Pour plus d'informations

LIEN VERS LE DESCRIPTIF
<https://thinkr.fr/r-niveau-2-creation-de-packages/>

Fiche officielle (France Compétences)

<https://www.francecompetences.fr/recherche/rs/5074/>

Entrées et sorties de formation

Entrées en formation	4
Parcours menés à terme	3
Taux d'achèvement	75 %
Période	03/2022 – 02/2023

Entrées et sorties — Caisse des Dépôts (open data), tous organismes confondus.

Demande CPF — formations engagées

Dossiers CPF engagés	8
Montant total engagé	22 678 €
Coût moyen par dossier	2 835 €
Durée moyenne	35 h
Période	12/2020 – 10/2022
France compétences	92 %
Reste à charge titulaire	8 %

Formations engagées via le CPF — Caisse des Dépôts (open data), tous organismes confondus.