

PLANNING DU CERTIFICAT DATA SCIENTIST (BD 27)

MODULES	PERIODE	DUREE
Module 1 1. Présentation du métier de data science et Rappels de R 2. Machine learning 1 : MCO et régressions pénalisées 3. Machine Learning 1 Régression logistique, arbres et forêt	Du 28 au 30 novembre 2022 28 novembre 2022 29 novembre 2022 30 novembre 2022	3 jours
Module 2 1. R avancé et SQL 2. R avancé et introduction au NoSQL 3. Webscraping et pipeline de traitements	Du 19 au 21 décembre 2022 19 décembre 2022 20 décembre 2022 21 décembre 2022	3 jours
Module 3 1. Méthodes d'agrégation : Bagging, Boosting - Support Vector Machine 2. Visualisation et Shiny 3. Graph Mining (à confirmer) Projet tutoré (après-midi)	Du 16 au 18 janvier 2023 16 janvier 2023 17 janvier 2023 18 janvier 2023	3 jours
Module 4 1. Clustering et réduction de dimension 2. Tidymodels : révision de ML et pipeline de traitements 3. Data Camp (matin) Projet tutoré (après-midi)	Du 13 au 15 février 2023 13 février 2023 14 février 2023 15 février 2023	3 jours
Module 5 1. Deep Learning 2. Text Mining 3. Text Mining (matin) Projet tutoré (après-midi)	Du 20 au 22 mars 2023 20 mars 2023 21 mars 2023 22 mars 2023	3 jours
Module 6 1. Passage à l'échelle : L'écosystème Hadoop 2. Passage à l'échelle : Spark 3. Projet tutoré (matin) Droit et éthique de la donnée (après-midi)	Du 17 au 19 avril 2023 17 avril 2023 18 avril 2023 19 avril 2023	3 jours

L'examen aura lieu le **12 mai 2023** en présentiel à l'Ensaie-Ensaie Formation Continue (Cepe).

L'ordre et le détail des modules peuvent être amenés à évoluer légèrement