

PLANNING DU CERTIFICAT DATA SCIENTIST BD33 (R et Python)

MODULES	PERIODE	DUREE
Module 1 1. R avancé et SQL 2. R avancé et introduction au NoSQL 3. Webscraping et pipeline de traitements	Du 9 au 11 octobre 2024 9 octobre 2024 10 octobre 2024 11 octobre 2024	3 jours
Module 2 1. Machine learning : MCO & régressions pénalisées 2. Machine learning : Régression logistique 3. Machine learning : Arbres et random forest	Du 13 au 15 novembre 2024 13 novembre 2024 14 novembre 2024 15 novembre 2024	3 jours
Module 3 1. Machine learning : Boosting & SVM 2. Réduction de dimension : Clustering & Analyse factorielle 3. Cloud Computing (matin) 4. Data Camp (après-midi)	Du 9 au 11 décembre 2024 9 décembre 2024 10 décembre 2024 11 décembre 2024	3 jours
Module 4 1. Visualisation 2. Deep Learning 3. Deep Learning (matin) 4. Projet tutoré (après-midi)	Du 13 au 15 janvier 2025 13 janvier 2025 14 janvier 2025 15 janvier 2025	3 jours
Module 5 1. Deep Learning : Application aux textes ou images 2. Deep Learning : Application aux textes ou images (matin) 3. Projet tutoré (après-midi) 4. Pipeline de traitements Python	Du 3 au 5 février 2025 3 février 2025 4 février 2025 5 février 2025	3 jours
Module 6 1. Big Data Processing avec Spark (1/2) 2. Big Data Processing avec Spark (2/2) 3. Projet tutoré (matin) 4. Droit et éthique de la donnée (après-midi)	Du 10 au 12 mars 2025 10 mars 2025 11 mars 2025 12 mars 2025	3 jours

L'examen aura lieu le **31 mars 2025** en présentiel à l'Ensaë-Ensaï Formation Continue (Cepe).

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