

PLANNING DU CERTIFICAT DATA SCIENTIST (BD29)

MODULES	PERIODE	DUREE
Module 1	Du 2 au 4 octobre 2023	3 jours
1. Présentation du métier de data science et Rappels de R	2 octobre 2023	
2. Machine learning 1 : MCO et régressions pénalisées	3 octobre 2023	
3. Machine Learning 1 Régression logistique, arbres et forêt	4 octobre 2023	
Module 2	Du 13 au 15 novembre 2023	3 jours
1. R avancé et SQL	13 novembre 2023	
2. R avancé et introduction au NoSQL	14 novembre 2023	
3. Webscraping et pipeline de traitements	15 novembre 2023	
Module 3	Du 4 au 6 décembre 2023	3 jours
1. Méthodes d'agrégation : Boosting - Support Vector Machine	4 décembre 2023	
2. Visualisation et Shiny	5 décembre 2023	
3. Cloud Computing (matin) Projet tutoré (après-midi)	6 décembre 2023	
Module 4	Du 8 au 10 janvier 2024	3 jours
1. Clustering et réduction de dimension	8 janvier 2024	
2. Tidymodels : révision de ML et pipeline de traitements	9 janvier 2024	
3. Data Camp (matin) Projet tutoré (après-midi)	10 janvier 2024	
Module 5	Du 5 au 7 février 2024	3 jours
1. Deep Learning	5 février 2024	
2. Text Mining	6 février 2024	
3. Text Mining (matin) Projet tutoré (après-midi)	7 février 2024	
Module 6	Du 11 au 13 mars 2024	3 jours
1. Passage à l'échelle : L'écosystème Hadoop	11 mars 2024	
2. Passage à l'échelle : Spark	12 mars 2024	
3. Projet tutoré (matin) Droit et éthique de la donnée (après-midi)	13 mars 2024	

L'examen aura lieu le **3 avril 2024** en présentiel à l'Ensaе-Ensaі Formation Continue (Cepe).
L'ordre et le détail des modules peuvent être amenés à évoluer légèrement