

PLANNING DU CERTIFICAT DATA SCIENTIST (BD 20)

MODULES	PERIODE	DUREE
Module 1	Du 30 novembre au 2 décembre 2020	3 jours
1. Présentation du métier de data science et Rappels de R	30 novembre 2020	
2. Architecture de stockage des données : SQL. NoSQL	1 décembre 2020	
3. Distribution stockage et calcul ; Hadoop. Hive ; Utilisation du Cloud	2 décembre 2020	
Module 2	Du 18 au 20 janvier 2021	3 jours
1. R avancé	18 janvier 2021	
2. Machine learning 1 : Régressions MCO et logistique	19 janvier 2021	
3. Régressions sous contraintes, arbres et forêts aléatoires	20 janvier 2021	
Module 3	Du 15 au 17 février 2021	3 jours
1. Clustering et réduction de dimensions	15 février 2021	
2. Visualisation et Shiny	16 février 2021	
3. Projet tutoré (matin) Sécurité informatique (après-midi)	17 février 2021	
Module 4	Du 15 au 17 mars 2021	3 jours
1. Machine learning 2 : Boosting. Svm	15 mars 2021	
2. Machine learning 2 : Optimisation. Réseau de neurones. Deep learning	16 mars 2021	
3. Machine learning 2 Projet tutoré (après-midi)	17 mars 2021	
Module 5	Du 12 au 14 avril 2021	3 jours
1. Introduction au text mining	14 avril 2021	
2. Passage à l'échelle	15 avril 2021	
3. Projet tutoré (matin) Droit et éthique de la donnée (après-midi)	16 avril 2021	
Module 6	Du 3 au 5 mai 2021	3 jours
1. Introduction au graph mining	3 mai 2021	
2. Pipeline de traitement Python	4 mai 2021	
3. Projet tutoré (matin) Data camp (après-midi)	5 mai 2021	

L'examen aura lieu le **7 juin 2021** en présentiel à l'Ensaë-Ensaï Formation Continue (Cepe).

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