

PLANNING DU CERTIFICAT DATA SCIENTIST (BD 15)

MODULES	PERIODE	DUREE
Module 1	Du 11 au 13 décembre 2019	3 jours
1. Présentation du métier de data science et Rappels de R	11 décembre 2019	
2. Architecture de stockage des données : SQL. NoSQL	12 décembre 2019	
3. Distribution stockage et calcul ; Hadoop. Hive ; Utilisation du Cloud	13 décembre 2019	
Module 2	Du 8 au 10 janvier 2020	3 jours
1. R avancé	8 janvier 2020	
2. Machine learning 1 : Régressions MCO et logistique	9 janvier 2020	
3. Régressions sous contraintes, arbres et forêts aléatoires	10 janvier 2020	
Module 3	Du 12 au 14 février 2020	3 jours
1. Clustering et réduction de dimensions	12 février 2020	
2. Visualisation et Shiny	13 février 2020	
3. Projet tutoré (matin) Sécurité informatiques (après-midi)	14 février 2020	
Module 4	Du 11 au 13 mars 2020	3 jours
1. Machine learning 2 : Boosting. Svm	11 mars 2020	
2. Machine learning 2 : Optimisation. Réseau de neurones. Deep learning	12 mars 2020	
3. Machine learning 2 Projet tutoré (après-midi)	13 mars 2020	
Module 5	Du 22 au 24 avril 2020	3 jours
1. Introduction au text mining	22 avril 2020	
2. Passage à l'échelle	23 avril 2020	
3. Projet tutoré (matin) Droit et éthique de la donnée (après-midi)	24 avril 2020	
Module 6	Du 13 au 15 mai 2020	3 jours
1. Introduction au graph mining	13 mai 2020	
2. Pipeline de traitement Python	14 mai 2020	
3. Projet tutoré (matin) Data camp (après-midi)	15 mai 2020	

L'examen aura lieu le **12 juin 2020** en présentiel à l'Ensaie-Ensaie Formation Continue (Cepe).

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