

PLANNING DU CERTIFICAT DATA SCIENTIST (BD 14)

MODULES	PERIODE	DUREE
Module 1	Du 11 au 13 septembre 2019	3 jours
1. Présentation du métier de data science. Rappels R	11 septembre 2019	
2. Architecture de stockage des données. SQL. NoSQL	12 septembre 2019	
3. Distribution stockage et calcul. Hadoop. Hive. Cloud	13 septembre 2019	
Module 2	Du 30 sept. au 2 oct. 2019	3 jours
1. R et connexion aux bases de données	30 septembre 2019	
2. Machine learning 1 : Régressions. Régressions logistiques	1er octobre 2019	
3. Machine learning 1 : Régressions sous contraintes	2 octobre 2019	
Module 3	Du 18 au 20 novembre 2019	3 jours
1. Clustering et réduction dimensions	18 novembre 2019	
2. Visualisation. Shiny	19 novembre 2019	
3. Projet tutoré (matin). Sécurité informatique (après-midi)	20 novembre 2019	
Module 4	Du 18 au 20 décembre 2019	3 jours
1. Machine learning 2 : Boosting. Svm	18 décembre 2019	
2. Machine learning : Optimisation. Réseau de neurones. Deep learning	19 décembre 2019	
3. Machine learning 2 : Réseau de neurones avec Word2Vec. Projet tutoré (après-midi)	20 décembre 2019	
Module 5	Du 15 au 17 janvier 2020	3 jours
1. Introduction au text mining	15 janvier 2020	
2. Passage à l'échelle	16 janvier 2020	
3. Projet tutoré (matin). Environnement juridique (après-midi)	17 janvier 2020	
Module 6	Du 19 au 21 février 2020	3 jours
1. Introduction au graph mining	19 février 2020	
2. Pipeline de traitement Python	20 février 2020	
3. Projet tutoré (matin). Data camp (après-midi)	21 février 2020	

L'examen aura lieu le **27 mars 2020** en présentiel à l'Ensaie-Ensaie Formation Continue (Cepe).

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