

PLANNING DU CERTIFICAT DATA SCIENTIST (BD 17)

MODULES	PERIODE	DUREE
Module 1	Du 21 au 23 octobre 2019	3 jours
1. Présentation du métier de data science. Rappels R	21 octobre 2019	
2. Architecture de stockage des données. SQL. NoSQL	22 octobre 2019	
3. Distribution stockage et calcul. Hadoop. Hive. Cloud	23 octobre 2019	
Module 2	Du 2 au 4 décembre 2019	3 jours
1. R et connexion aux bases de données	2 décembre 2019	
2. Machine learning 1 : Régressions. Régressions logistiques	3 décembre 2019	
3. Machine learning 1 : Régressions sous contraintes	4 décembre 2019	
Module 3	Du 22 au 24 janvier 2020	3 jours
1. Clustering et réduction dimensions	22 janvier 2020	
2. Visualisation. Shiny	23 janvier 2020	
3. Projet tutoré (matin). Sécurité informatique (après-midi)	24 janvier 2020	
Module 4	Du 26 au 28 février 2020	3 jours
1. Machine learning 2 : Boosting. Svm	26 février 2020	
2. Machine learning 2 : Optimisation. Réseau de neurones. Deep learning	27 février 2020	
3. Machine learning 2 : Réseau de neurones avec Word2Vec. Projet tutoré (après-midi)	28 février 2020	
Module 5	Du 18 au 20 mars 2020	3 jours
1. Introduction au text mining	18 mars 2020	
2. Passage à l'échelle	19 mars 2020	
3. Projet tutoré (matin). Environnement juridique (après-midi)	20 mars 2020	
Module 6	Du 20 au 22 avril 2020	3 jours
1. Introduction au graph mining	20 avril 2020	
2. Pipeline de traitement Python	21 avril 2020	
3. Projet tutoré (matin). Data camp (après-midi)	22 avril 2020	

L'examen aura lieu le **19 mai 2020** en présentiel à l'Ensaie-Ensaie Formation Continue (Cepe).

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