

## PLANNING DU CERTIFICAT DATA SCIENTIST (BD 15)

MODULES	PERIODE	DUREE
<b>Module 1</b>	<b>Du 11 au 13 décembre 2019</b>	3 jours
1. Présentation du métier de data science. Rappels R	11 décembre 2019	
2. Architecture de stockage des données. SQL. NoSQL	12 décembre 2019	
3. Distribution stockage et calcul. Hadoop. Hive. Cloud	13 décembre 2019	
<b>Module 2</b>	<b>Du 8 au 10 janvier 2020</b>	3 jours
1. R et connexion aux bases de données	8 janvier 2020	
2. Machine learning 1 : Régressions. Régressions logistiques	9 janvier 2020	
3. Machine learning 1 : Régressions sous contraintes	10 janvier 2020	
<b>Module 3</b>	<b>Du 12 au 14 février 2020</b>	3 jours
1. Clustering et réduction dimensions	12 février 2020	
2. Visualisation. Shiny	13 février 2020	
3. Projet tutoré (matin). Sécurité informatique (après-midi)	14 février 2020	
<b>Module 4</b>	<b>Du 11 au 13 mars 2020</b>	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 : Réseau de neurones avec Word2Vec. Projet tutoré (après-midi)	13 mars 2020	
<b>Module 5</b>	<b>Du 22 au 24 avril 2020</b>	3 jours
1. Introduction au text mining	22 avril 2020	
2. Passage à l'échelle	23 avril 2020	
3. Projet tutoré (matin). Environnement juridique (après-midi)	24 avril 2020	
<b>Module 6</b>	<b>Du 13 au 15 mai 2020</b>	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.