

## PLANNING DU CERTIFICAT DATA SCIENTIST (BD 16)

MODULES	PERIODE	DUREE
<b>Module 1</b>	<b>Du 23 au 25 mars 2020</b>	3 jours
1. Présentation du métier de data science et Rappels de R	23 mars 2020	
2. Architecture de stockage des données : SQL. NoSQL	24 mars 2020	
3. Distribution stockage et calcul ; Hadoop. Hive ; Utilisation du Cloud	25 mars 2020	
<b>Module 2</b>	<b>Du 27 au 29 avril 2020</b>	3 jours
1. R avancé	27 avril 2020	
2. Machine learning 1 : Régressions MCO et logistique	28 avril 2020	
3. Régressions sous contraintes, arbres et forêts aléatoires	29 avril 2020	
<b>Module 3</b>	<b>Du 25 au 27 mai 2020</b>	3 jours
1. Clustering et réduction dimensions	25 mai 2020	
2. Visualisation et Shiny	26 mai 2020	
3. Projet tutoré (matin) Sécurité informatiques (après-midi)	27 mai 2020	
<b>Module 4</b>	<b>Du 29 juin au 1<sup>er</sup> juillet 2020</b>	3 jours
1. Machine learning 2 : Boosting. Svm	29 juin 2020	
2. Machine learning 2 : Optimisation. Réseau de neurones. Deep learning	30 juin 2020	
3. Machine learning 2 Projet tutoré (après-midi)	1 juillet 2020	
<b>Module 5</b>	<b>Du 7 au 9 septembre 2020</b>	3 jours
1. Introduction au text mining	7 septembre 2020	
2. Passage à l'échelle	8 septembre 2020	
3. Projet tutoré (matin) Droit et éthique de la donnée (après-midi)	9 septembre 2020	
<b>Module 6</b>	<b>Du 12 au 14 octobre 2020</b>	3 jours
1. Introduction au graph mining	12 octobre 2020	
2. Pipeline de traitement Python	13 octobre 2020	
3. Projet tutoré (matin) Data camp (après-midi)	14 octobre 2020	

L'examen aura lieu le **16 novembre 2020** en présentiel à l'Ensaie-Ensaie Formation Continue (Cepe).

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