

## PLANNING DU CERTIFICAT DATA SCIENTIST (BD 17)

MODULES	PERIODE	DUREE
<b>Module 1</b>	<b>Du 21 au 23 octobre 2019</b>	3 jours
1. Présentation du métier de data science et Rappels R	21 octobre 2019	
2. Architecture de stockage des données : SQL. NoSQL	22 octobre 2019	
3. Distribution stockage et calcul ; Hadoop. Hive ; Utilisation du Cloud	23 octobre 2019	
<b>Module 2</b>	<b>Du 2 au 4 décembre 2019</b>	3 jours
1. R avancé	2 décembre 2019	
2. Machine learning 1 : Régressions MCO et logistique	3 décembre 2019	
3. Régressions sous contraintes, arbres et forêts aléatoires	4 décembre 2019	
<b>Module 3</b>	<b>Du 22 au 24 janvier 2020</b>	3 jours
1. Clustering et réduction de dimensions	22 janvier 2020	
2. Visualisation et Shiny	23 janvier 2020	
3. Projet tutoré (matin) Sécurité informatique (après-midi)	24 janvier 2020	
<b>Module 4</b>	<b>Du 26 au 28 février 2020</b>	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 Projet tutoré (après-midi)	28 février 2020	
<b>Module 5</b>	<b>Du 18 au 20 mars 2020</b>	3 jours
1. Introduction au text mining	18 mars 2020	
2. Passage à l'échelle	19 mars 2020	
3. Projet tutoré (matin). Droit et éthique de la donnée (après-midi)	20 mars 2020	
<b>Module 6</b>	<b>Du 20 au 22 avril 2020</b>	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-Ensa Formation Continue (Cepe).

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