

## PLANNING DU CERTIFICAT DATA SCIENTIST (SESSION INTENSIVE BDI 6)

MODULES	PERIODE
<b>Module 1</b>	<b>Du 21 au 25 octobre 2019</b>
1. Présentation du métier de data science ; rappels R	21 octobre 2019
2. Manipulation des données avec R ; data table, dply, ggplot ; Atelier	22 octobre 2019
3. Rappel des bases de données (SQL, no SQL, connexion avec R)	23 octobre 2019
4. Hadoop, principes et usages	24 octobre 2019
5. Bases de données avec Hive, introduction à Spark	25 octobre 2019
<b>Module 2</b>	<b>Du 2 au 6 décembre 2019</b>
1. Réduction de dimension et Classification	2 décembre 2019
2. Visualisation ; Shiny	3 décembre 2019
3. Machine learning I : MCO, Anova, Régression logistique	4 décembre 2019
4. Machine learning I : Régression sous contrainte (ridge, lasso)	5 décembre 2019
5. Atelier	6 décembre 2019
<b>Module 3</b>	<b>Du 3 au 7 février 2020</b>
1. Machine learning II : agrégation de modèles, boosting, bagging, SVM	3 février 2020
2. Machine learning II : introduction à l'optimisation, réseaux de neurones	4 février 2020
3. Introduction au text mining et au graph mining	5 février 2020
4. Environnement juridique du big data	6 février 2020 (matin)
5. Enjeux de la sécurité informatique	6 février 2020 (après-midi)
6. Atelier	7 février 2020

L'**examen** aura lieu le **20 mars 2020** en présentiel à l'Ensaie-Ensaie Formation continue (Cepe).

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