

PLANNING DU CERTIFICAT DATA SCIENTIST (SESSION INTENSIVE BDI 5)

MODULES	PERIODE
Module 1	Du 22 au 26 octobre 2018
1. Présentation du métier de data science ; rappels R	22 octobre 2018
2. Manipulation des données avec R ; data table, dply, ggplot ; Atelier	23 octobre 2018
3. Rappel des bases de données (SQL, no SQL, connexion avec R)	24 octobre 2018
4. Hadoop, principes et usages	25 octobre 2018
5. Bases de données avec Hive, introduction à Spark	26 octobre 2018
Module 2	Du 10 au 14 décembre 2018
1. Réduction de dimension et Classification	10 décembre 2018
2. Visualisation ; Shiny	11 décembre 2018
3. Machine learning I : MCO, Anova, Régression logistique	12 décembre 2018
4. Machine learning I : Régression sous contrainte (ridge, lasso)	13 décembre 2018
5. Atelier	14 décembre 2018
Module 3	Du 4 au 8 février 2019
1. Machine learning II : agrégation de modèles, boosting, bagging, SVM	4 février 2019
2. Machine learning II : introduction à l'optimisation, réseaux de neurones	5 février 2019
3. Introduction au text mining et au graph mining	6 février 2019
4. Environnement juridique du big data	7 février 2019 (matin)
5. Enjeux de la sécurité informatique	7 février 2019 (après-midi)
6. Atelier	8 février 2019

L'**examen** aura lieu le **25 mars 2019** en présentiel à l'Ensaie-Ensaie Formation continue (Cepe)

Des conférences seront organisées en fin de journée après certaines formations