

Certificat de Data Scientist

Session intensive en partenariat avec Berkeley

Du 24 avril au 7 juillet 2017

Attention : Le programme ci-dessous est prévisionnel. Il est susceptible d'évoluer.

MODULES	PERIODE
MODULE 1 - Du 24 au 28 avril 2017 (5 jours)	
1. Introduction à la data science avec R et atelier	24 avril 2017
2. Bases de données SQL et R	25 avril 2017
3. Bases de données NoSQL	26 avril 2017
4. MCO, Anova, régression logistique	27 avril 2017
5. Régression sous contrainte: ridge, lasso, elasticnet et régression non paramtrique	28 avril 2017
MODULE 2 – Du 29 mai au 2 juin 2017 (5 jours)	
1. Réduction de dimension et classification	29 mai 2017
2. Visualisation et shiny	30 mai 2017
3. Machine learning (SVM, Boosting, Bagging..)	31 mai 2017
4. HADOOP, Hive	1 ^{er} juin 2017
5. Atelier	2 juin 2017
MODULE 3 – Du 3 au 7 juillet 2017 (5 jours)	
1. Spark	3 juillet 2017
2. Atelier	4 juillet 2017
3. Introduction au textmining et graphmining	5 juillet 2017
5. Introduction à l'optimisation	6 juillet 2017
6. Examen	7 juillet 2017

Plusieurs conférences seront organisées de 17h à 18h après la formation :

- Jeudi 27 avril : *Le big data et la finance*
- Mardi 30 mai : *L'environnement juridique du Big Data*
- Mardi 4 juillet, mercredi 5 juillet ou jeudi 6 juillet : *Les enjeux de la sécurité informatique*

Une soirée de remise des certificats est prévue le 7 juillet 2017.