

**Dcid - Data Science Consulting**

Partenaire de vos projets de valorisation de données et d'aide à la décision



# CATALOGUE FORMATION 2019

---

## VALORISEZ VOS DONNÉES



Acquérir une culture data et comprendre les enjeux du numériques



S'initier à la data science et développer ses programmes et algorithmes



Maîtriser les techniques d'intelligence artificielle et de big data analytics



# LA DATA SCIENCE, DÉCOUVERTE DE L'ANALYTICS

La science des données (data science) est l'extraction de connaissance à partir d'ensembles de données. Elle s'appuie sur des bases statistique, informatique et technologique. Concrètement, la science des données vous permet d'explorer et d'analyser vos données brutes pour les transformer en informations précieuses et générer de la valeur de façon créative et innovante.



## Durée

- 1 jour



## Tarif

- Inter : 500€ HT / participant (6 participants maximum)
- Intra : nous consulter



## Lieu

- Locaux de Dcid Consulting (Orléans)
- Locaux de votre entreprise (Région Centre et Paris)



## Public

Tout public souhaitant acquérir une culture data et adopter une démarche 'data-driven'

- Analyse de données ☆☆☆☆☆
- Statistique ☆☆☆☆☆
- Programmation ☆☆☆☆☆
- Base de données ☆☆☆☆☆
- Curiosité ☆☆☆☆☆



## Renseignements

- [contact@dcid-consulting.fr](mailto:contact@dcid-consulting.fr)
- 02 38 25 33 90



## Prérequis

- Aucun



## Ressources utiles

- [Le big data.fr](http://Le big data.fr)
- [Wikipedia](http://Wikipedia)
- [Open Classrooms](http://Open Classrooms)



## Objectifs pédagogiques

- Comprendre les enjeux de la révolution numérique
- Acquérir une culture data
- Adopter une démarche d'innovation 'data-driven'
- Evaluer les possibilités offertes/promises
- Identifier de nouveaux leviers de croissance



## Programme

- La data science, c'est quoi ?  
Introduction, présentation des concepts  
Les bonnes pratiques et la gouvernance des données
- La data science pour qui ?  
Les objectifs à atteindre et les défis à relever  
Les environnements et contextes favorables
- Quelles ressources, quels acteurs et quels outils ?  
Panorama des ressources nécessaires  
Les acteurs impliqués dans un projet de data science  
Panorama des outils disponibles
- Le déroulement d'un projet de data science  
Idéation, design, création, test et mise en production
- Des cas d'usage démonstratifs
- Perspectives: mettre en œuvre son propre projet de data science



## Méthodes et moyens pédagogiques

- Alternance d'exposés et d'applications pratiques
- Supports de formation remis aux participants
- Formation dispensée par un data scientist expérimenté



## BIG DATA, PANORAMA ET ENJEUX

Le big data (ou mégadonnées, données massives) désigne des ensembles de données devenus si volumineux qu'ils dépassent les capacités humaines d'analyse et des outils informatiques classiques. Le big data vous aide à réduire vos risques, faciliter vos prises de décision, créer votre différence grâce à l'analyse prédictive.



### Durée

- 1 jour



### Tarif

- Inter : 500€ HT / participant (6 participants maximum)
- Intra : nous consulter



### Lieu

- Locaux de Dcid Consulting (Orléans)
- Locaux de votre entreprise (Région Centre et Paris)



### Public

Tout public souhaitant acquérir une culture data et adopter une démarche 'data-driven'

- Analyse de données ☆☆☆☆☆
- Statistique ☆☆☆☆☆
- Programmation ☆☆☆☆☆
- Base de données ☆☆☆☆☆
- Curiosité ☆☆☆☆☆



### Renseignements

- [contact@dcid-consulting.fr](mailto:contact@dcid-consulting.fr)
- 02 38 25 33 90



### Prérequis

- Aucun



### Ressources utiles

- [Wikipedia](#)
- [Le big data.fr](#)
- [Open Classrooms](#)
- [You Tube](#)



### Objectifs pédagogiques

- Comprendre les enjeux de la révolution numérique
- Acquérir une culture data
- Adopter une démarche d'innovation 'data-driven'
- Evaluer les possibilités offertes/promises
- Identifier de nouveaux leviers de croissance



### Programme

- Le big data, c'est quoi ?  
Introduction, présentation des concepts  
Les bonnes pratiques et la gouvernance des données
- Le big data pour qui ?  
Les objectifs à atteindre et les défis à relever  
Les environnements et contextes favorables
- Quelles ressources, quels acteurs et quels outils ?  
Panorama des ressources nécessaires  
Les acteurs impliqués dans un projet big data  
Panorama des outils disponibles
- Le déroulement d'un projet big data  
Idéation, design, création, test et mise en production
- Des cas d'usage démonstratifs
- Perspectives: mettre en œuvre son propre projet big data



### Méthodes et moyens pédagogiques

- Alternance d'exposés et d'applications pratiques
- Supports de formation remis aux participants
- Formation dispensée par un data scientist expérimenté



## OPEN DATA, ENJEUX ET USAGES

Les données ouvertes (open data) sont des données dont l'accès, l'exploitation et l'utilisation sont libres de droit. De multiples opportunités s'offrent alors pour améliorer la connaissance et créer de nouveaux produits et services de qualité. Comprendre les enjeux et les usages est donc une véritable stratégie d'innovation.



### Durée

- 1 jour



### Tarif

- Inter : 500€ HT / participant (6 participants maximum)
- Intra : nous consulter



### Lieu

- Locaux de Dcid Consulting (Orléans)
- Locaux de votre entreprise (Région Centre et Paris)



### Public

Tout public souhaitant acquérir une culture data et adopter une démarche 'data-driven'

- Analyse de données ★★★★★
- Statistique ★★★★★
- Programmation ★★★★★
- Base de données ★★★★★
- Curiosité ★★★★★



### Renseignements

- [contact@dcid-consulting.fr](mailto:contact@dcid-consulting.fr)
- 02 38 25 33 90



### Prérequis

- Aucun



### Ressources utiles

- [Wikipedia](#)
- [Le big data.fr](#)
- [Site gouvernemental](#)



### Objectifs pédagogiques

- Comprendre les enjeux de l'open data
- Acquérir une culture data
- Adopter une démarche d'innovation 'data-driven'
- Evaluer les possibilités offertes/promises
- Identifier de nouveaux leviers de croissance



### Programme

- L'open data, c'est quoi ?  
Introduction, présentation des concepts  
Les bonnes pratiques et la gouvernance des données
- L'open data pour qui ?  
Les objectifs à atteindre et les défis à relever  
Les environnements et contextes favorables
- Quelles ressources, quels acteurs et quels outils ?  
Panorama des ressources nécessaires  
Les acteurs impliqués dans un projet d'open data  
Panorama des outils disponibles
- Le déroulement d'un projet d'open data  
Surveiller la qualité et la fiabilité des données  
Idéation, design, création, test et mise en production
- Des cas d'usage démonstratifs
- Perspectives: mettre en œuvre son propre projet d'open data



### Méthodes et moyens pédagogiques

- Alternance d'exposés et d'applications pratiques
- Supports de formation remis aux participants
- Formation dispensée par un data scientist expérimenté

## INTRODUCTION À R

R est un logiciel open source de statistique qui est porté par une communauté scientifique internationale de plus de 3 millions d'utilisateurs. Il fournit une large variété d'outils statistiques et graphiques qui le place comme un des langages de référence des data scientists. Maîtriser la programmation en R est maintenant indispensable pour tous projets d'analyse et de science des données.



### Durée

- 2 jours



### Tarif

- Inter : 950€ HT / participant (6 participants maximum)
- Intra : nous consulter



### Lieu

- Locaux de Dcid Consulting (Orléans)
- Locaux de votre entreprise (Région Centre et Paris)



### Public

Tout public ayant déjà quelques notions en programmation et/ou analyse de données

- Analyse de données ★★★★★
- Statistique ★★★★★
- Programmation ★★★★★
- Base de données ★★★★★
- Curiosité ★★★★★



### Renseignements

- [contact@dcid-consulting.fr](mailto:contact@dcid-consulting.fr)
- 02 38 25 33 90



### Prérequis

- Ordinateur individuel
- Connexion internet



### Ressources utiles

- [R, le site officiel](#)
- [R Studio, le site officiel](#)
- [Vincent Goulet, guide d'introduction](#)
- [Open Classrooms](#)



### Objectifs pédagogiques

- Comprendre le fonctionnement de R et de son écosystème
- Maîtriser les notions principales de la programmation en R
- Être autonome dans la conduite d'un projet simple d'analyse de données avec R



### Programme

- Découvrir R, son environnement et son écosystème  
Présentation et chargement de R, des librairies principales  
Présentation de R Studio et de son écosystème
- Comprendre les notions principales de programmation  
Les différents objets dans R  
Les opérateurs, boucles, conditions et fonctions
- Importer des données  
Importation de données web, fichier, base de données  
Exportation de données
- Manipuler des données  
Extraction, filtrage, requête sur le jeu de données  
Modification des données
- Réaliser des analyses statistiques  
Réalisation de tests et analyses statistiques simples  
Exemple de fonctionnalités avancées
- Visualiser un jeu des données  
Création de graphiques simples  
Exemple de fonctionnalités avancées
- Et après ? Perspectives et cas d'usage



### Méthodes et moyens pédagogiques

- Alternance d'exposés et d'applications pratiques
- Possibilité de travailler sur les jeux de données des participants
- Codes source et supports de formation remis aux participants
- Formation dispensée par un data scientist expérimenté (pratique de R depuis 7 ans minimum)

# IMPORTATION ET MANIPULATION DE DONNÉES AVEC R

Le traitement de données est une étape incontournable de tout projet de data science. Grâce à ses librairies 'dplyr' et 'data.table' (notamment), R est un des langages de référence des data scientists pour le traitement de données et requiert des compétences méthodologiques et informatiques. Maîtriser R fera de vous un acteur incontournable pour le traitement et la valorisation des données de votre entreprise.



## Durée

- 1 jour



## Tarif

- Inter : 550€ HT / participant (6 participants maximum)
- Intra : nous consulter



## Lieu

- Locaux de Dcid Consulting (Orléans)
- Locaux de votre entreprise (Région Centre et Paris)



## Public

Tout public ayant déjà quelques notions en programmation et/ou analyse de données

- Analyse de données ★★★★★
- Statistique ★★★★★
- Programmation ★★★★★
- Base de données ★★★★★
- Curiosité ★★★★★



## Renseignements

- [contact@dcid-consulting.fr](mailto:contact@dcid-consulting.fr)
- 02 38 25 33 90



## Prérequis

- Ordinateur individuel
- Connexion internet



## Ressources utiles

- [R, le site officiel](#)
- [R Studio, le site officiel](#)
- [Dplyr dans le tidyverse](#)



## Objectifs pédagogiques

- Maîtriser l'importation et la manipulation de données de sources variées (numérique, catégorielle, textuelle, temporelle)
- Savoir préparer un jeu de données en vue d'une utilisation ultérieure (dataviz, machine learning, analyses statistiques)
- Savoir extraire, filtrer et résumer un jeu de données
- Savoir visualiser et explorer un jeu de données



## Programme

- Découvrir R, son environnement et son écosystème  
Présentation et chargement de R, des librairies principales  
Présentation de R Studio et son écosystème
- Importer des données de sources variées  
Chargement de fichiers de données  
Connexion et requête sur une base de données  
Extraction de données depuis un site web
- Manipuler les données de sources variées  
Données numériques, catégorielles, textuelles et temporelles
- Manipuler et préparer un jeu de données  
Extraction d'un sous ensemble du jeu de données  
Sélection, filtrage, jointure des données  
Résumé statistique des données  
Modification du format du jeu de données
- Visualiser et explorer un jeu des données  
Représentation graphique simple  
Présentation d'outils et méthodes pour l'exploration
- Et après ? Perspectives et cas d'usage



## Méthodes et moyens pédagogiques

- Alternance d'exposés et d'applications pratiques
- Possibilité de travailler sur les jeux de données des participants
- Codes source et supports de formation remis aux participants
- Formation dispensée par un data scientist expérimenté (pratique de R depuis 7 ans minimum)

# ANALYSE STATISTIQUE ET MODÉLISATION AVEC R

L'analyse de données est une étape essentielle à toute démarche de valorisation de données. Grâce à sa grande variété d'outils statistiques et graphiques, R est un des langages de référence des data scientists pour l'analyse de données et requiert des compétences techniques et informatiques. Maîtriser R fera de vous un acteur incontournable pour l'analyse et la valorisation des données de votre entreprise.



## Durée

- 2 jours



## Tarif

- Inter : 950€ HT / participant (6 participants maximum)
- Intra : nous consulter



## Lieu

- Locaux de Dcid Consulting (Orléans)
- Locaux de votre entreprise (Région Centre et Paris)



## Public

Tout public ayant déjà quelques notions en programmation et/ou analyse de données

- Analyse de données ★★★★★
- Statistique ★★★★★
- Programmation ★★★★★
- Base de données ★★★★★
- Curiosité ★★★★★



## Renseignements

- [contact@dcid-consulting.fr](mailto:contact@dcid-consulting.fr)
- 02 38 25 33 90



## Prérequis

- Ordinateur individuel
- Connexion internet



## Ressources utiles

- [R, le site officiel](#)
- [R Studio, le site officiel](#)
- [C. Genolini, guide R](#)



## Objectifs pédagogiques

- Maîtriser l'importation et la manipulation des données
- Être autonome dans la réalisation de ses propres fonctions et programmes
- Savoir réaliser des analyses descriptives et des tests statistiques
- Savoir utiliser les principales bibliothèques d'analyse de données (analyse factorielle, classification et régression)



## Programme

- Découvrir R, son environnement et son écosystème  
Présentation et chargement de R, des bibliothèques principales  
Présentation de R Studio et son écosystème
- Comprendre les notions principales de programmation  
Les différents objets dans R  
Les opérateurs, boucles, conditions et fonctions
- Importer, manipuler et exporter des données de sources variées  
Chargement, extraction, filtrage, résumé du jeu de données  
Export et sauvegarde du jeu de données
- Réaliser des analyses statistiques de base  
Résumé statistique et analyses descriptives  
Tests d'hypothèses
- Réaliser des analyses de données multivariées  
Analyses factorielles  
Classification et régression
- Visualiser et explorer un jeu des données  
Représentation graphique simple  
Présentation d'outils et méthodes pour l'exploration
- Et après ? Perspectives et cas d'usage



## Méthodes et moyens pédagogiques

- Alternance d'exposés et d'applications pratiques
- Possibilité de travailler sur les jeux de données des participants
- Codes source et supports de formation remis aux participants
- Formation dispensée par un data scientist expérimenté (pratique de R depuis 7 ans minimum)

## LA DATAVIZ AVEC R

La visualisation de données est une étape essentielle à toute démarche de valorisation de données. Grâce à sa grande variété d'outils statistiques et graphiques, R est un des langages de référence des data scientists pour la dataviz. Maîtriser R et ses bibliothèques graphiques (ggplot2, plotly, highcharter, etc.) fera de vous un acteur incontournable pour la visualisation et la valorisation des données de votre entreprise.



### Durée

- 2 jours



### Tarif

- Inter : 950€ HT / participant (6 participants maximum)
- Intra : nous consulter



### Lieu

- Locaux de Dcid Consulting (Orléans)
- Locaux de votre entreprise (Région Centre et Paris)



### Public

Tout public ayant déjà quelques notions en programmation et/ou analyse de données

- Analyse de données ★★★★★
- Statistique ★★★★★
- Programmation ★★★★★
- Base de données ★★★★★
- Curiosité ★★★★★



### Renseignements

- [contact@dcid-consulting.fr](mailto:contact@dcid-consulting.fr)
- 02 38 25 33 90



### Prérequis

- Ordinateur individuel
- Connexion internet



### Ressources utiles

- [Le big data.fr](http://Le big data.fr)
- [R Studio, le site officiel](#)
- [ggplot2 dans le tidyverse](#)
- [Highcharter, le site officiel](#)



### Objectifs pédagogiques

- Maîtriser R, son fonctionnement et son écosystème
- Maîtriser les principes de la visualisation de données dans R
- Appliquer les bonnes pratiques en matière de dataviz
- Savoir utiliser 'ggplot2', la bibliothèque graphique incontournable de R
- Savoir créer des graphiques interactifs avec plotly, highcharter, etc.



### Programme

- Découvrir R, son environnement et son écosystème  
Présentation et chargement de R, des bibliothèques principales  
Présentation de R Studio et son écosystème
- Comprendre les principes et les enjeux de la dataviz  
Quelles données ? Pour quels objectifs ?  
Les bonnes pratiques à appliquer
- Connaître les différents types de graphique et leurs usages  
Les graphiques de type nuage de points, courbe, histogramme, boîte à moustaches, etc.  
Les usages prévus selon le type de graphique
- Utiliser 'ggplot2', outil puissant et incontournable de l'écosystème de R  
Présentation de 'ggplot2' et de ses principales fonctions  
Réalisation de graphiques
- Réaliser des explorations spatiales grâce à la cartographie
- Utiliser les bibliothèques interactives pour offrir plus de liberté aux utilisateurs
- Communiquer ses réalisations graphiques, finalité de la dataviz
- Et après ? Perspectives et cas d'usage



### Méthodes et moyens pédagogiques

- Alternance d'exposés et d'applications pratiques
- Possibilité de travailler sur les jeux de données des participants
- Codes source et supports de formation remis aux participants
- Formation dispensée par un data scientist expérimenté (pratique de R depuis 7 ans minimum)



## R POUR LA DATA SCIENCE

Explorer et analyser les données passées et présentes pour prédire des données futures, tel est l'enjeu de la data science. Grâce à sa large variété d'outils statistiques et graphiques, R est un langage de référence pour les data scientists. Maîtriser la programmation en R fera de vous un acteur complet dans la réalisation de projets de data science.



### Durée

- 2 jours



### Tarif

- Inter : 950€ HT / participant (6 participants maximum)
- Intra : nous consulter



### Lieu

- Locaux de Dcid Consulting (Orléans)
- Locaux de votre entreprise (Région Centre et Paris)



### Public

Tout public ayant déjà quelques notions en programmation et/ou analyse de données

- Analyse de données ★★★★★
- Statistique ★★★★★
- Programmation ★★★★★
- Base de données ★★★★★
- Curiosité ★★★★★



### Renseignements

- [contact@dcid-consulting.fr](mailto:contact@dcid-consulting.fr)
- 02 38 25 33 90



### Prérequis

- Ordinateur individuel
- Connexion internet



### Ressources utiles

- [R, le site officiel](#)
- [Le big data.fr](#)
- [La fabrique du net.fr](#)



### Objectifs pédagogiques

- Maîtriser R, son fonctionnement et son écosystème
- Maîtriser le processus complet d'un projet de data science, de la collecte des données jusqu'à la communication des insights
- Analyser et exploiter des données structurées et non structurées avec R
- Elaborer des programmes d'analyses avec R



### Programme

- Découvrir la science des données, en comprendre ses enjeux
- Découvrir R, son environnement et son écosystème  
Les librairies, la programmation, les différents objets dans R  
Les opérateurs, boucles, conditions et fonctions
- Importer, manipuler et exporter des données  
Chargement de données de sources variées (texte, csv, Excel, SQL, web, etc.)  
Préparation, extraction, filtrage et jointure du jeu de données
- Analyser et exploiter des données multivariées  
Régression linéaire et régression multiple  
Analyses factorielles et classification
- Analyser et exploiter des données textuelles  
Import et manipulation de données textuelles  
Les principaux concepts : tokenization, lemmatisation, stop-words, etc.
- Modéliser et prédire l'information  
Apprentissage supervisé et non supervisé  
Introduction au machine learning
- Visualiser et communiquer ses résultats  
Les principes de la dataviz, les différents type de graphique et leurs usages  
La librairie ggplot2 et les librairies interactives
- Et après ? Perspectives et cas d'usage



### Méthodes et moyens pédagogiques

- Alternance d'exposés et d'applications pratiques
- Possibilité de travailler sur les jeux de données des participants
- Codes source et supports de formation remis aux participants
- Formation dispensée par un data scientist expérimenté (pratique de R depuis 7 ans minimum)



## LE MACHINE LEARNING AVEC R

Le machine learning, sous domaine de l'intelligence artificielle, est une manière de modéliser des phénomènes (sous forme de données), dans le but de prendre des décisions stratégiques. R dispose de nombreux algorithmes de machine learning ainsi que de la fameuse librairie d'apprentissage caret. Maitriser le machine learning avec R vous fera basculer dans l'univers du prédictif.



### Durée

- 1 jour (possibilité 2 jours)



### Tarif

- Inter : 550€ HT / participant (6 participants maximum)
- Intra : nous consulter



### Lieu

- Locaux de Dcid Consulting (Orléans)
- Locaux de votre entreprise (Région Centre et Paris)



### Public

Tout public ayant déjà des connaissances en programmation et/ou analyse de données

- Analyse de données ★★★★★
- Statistique ★★★★★
- Programmation ★★★★★
- Base de données ★★★★★
- Curiosité ★★★★★



### Renseignements

- [contact@dcid-consulting.fr](mailto:contact@dcid-consulting.fr)
- 02 38 25 33 90



### Prérequis

- Ordinateur individuel
- Connexion internet



### Ressources utiles

- [Datacamp](#)
- [Machine learning mastery](#)
- [Laurent Gatto, introduction au ML](#)
- [Open Classrooms](#)



### Objectifs pédagogiques

- Maitriser R, son fonctionnement et son écosystème
- Maitriser le processus complet d'un projet de machine learning, de la sélection des algorithmes jusqu'à l'évaluation de leurs performances
- Connaître et savoir utiliser les algorithmes de machine learning
- Elaborer des programmes complets de machine learning avec R



### Programme

- Découvrir le machine learning, en comprendre ses enjeux
- Découvrir R, son environnement et son écosystème  
La programmation, les différents objets dans R, les opérateurs  
Les librairies de machine learning
- Comprendre et utiliser l'apprentissage supervisé  
Les régressions linéaires, multivariées, logistiques  
Les arbres et forêts aléatoires  
Le Gradient Boosting Machine
- Comprendre et utiliser l'apprentissage non-supervisé  
Le clustering (k-means, k-mode)  
La détection de pattern (association rules)
- Evaluer les performances des algorithmes d'apprentissage
- Utiliser le machine learning dans un workflow
- Et après ? Perspectives et cas d'usage



### Méthodes et moyens pédagogiques

- Alternance d'exposés et d'applications pratiques
- Possibilité de travailler sur les jeux de données des participants
- Codes source et supports de formation remis aux participants
- Formation dispensée par un data scientist expérimenté (pratique de R depuis 7 ans minimum)



## BIG DATA ANALYTICS AVEC R

La variété et la quantité de données (ou big data) représentent un véritable enjeu : comment les traiter ? comment les exploiter ? pour quelle valeur ajoutée ? R rend possible l'exploration et l'analyse de ces mégadonnées (ou big data) et fera de vous un acteur incontournable pour prendre les bonnes décisions.



### Durée

- 1 jour (possibilité 2 jours)



### Tarif

- Inter : 550€ HT / participant (6 participants maximum)
- Intra : nous consulter



### Lieu

- Locaux de Dcid Consulting (Orléans)
- Locaux de votre entreprise (Région Centre et Paris)



### Public

Tout public ayant déjà des connaissances en programmation et/ou analyse de données

- Analyse de données ★★★★★
- Statistique ★★★★★
- Programmation ★★★★★
- Base de données ★★★★★
- Curiosité ★★★★★



### Renseignements

- [contact@dcid-consulting.fr](mailto:contact@dcid-consulting.fr)
- 02 38 25 33 90



### Prérequis

- Ordinateur individuel
- Connexion internet



### Ressources utiles

- [Le big data.fr](http://Le big data.fr)
- [Open Classrooms, initiation au ML](#)
- [R Studio, webinar big data](#)



### Objectifs pédagogiques

- Maîtriser R, son fonctionnement et son écosystème
- Savoir créer des programmes fonctionnels et performants
- Maîtriser la manipulation et l'exploitation de données massives, de données textuelles et de réseau
- Connaître et savoir utiliser les algorithmes adaptés aux données massives et variées



### Programme

- Connaître et appliquer les bonnes pratiques  
Programmation avancée, programmation fonctionnelle et calcul parallèle
- Manipuler et exploiter des données massives  
Import et traitement des données  
Représentation graphique des données  
Méthode de régression et classification
- Manipuler et exploiter des données textuelles  
Import et traitement des données  
Analyse, modélisation et clustering
- Manipuler et exploiter des données de réseau  
Import et traitement des données  
Représentation graphique des réseaux  
Analyse et clustering
- Et après ? Perspectives et cas d'usage



### Méthodes et moyens pédagogiques

- Alternance d'exposés et d'applications pratiques
- Possibilité de travailler sur les jeux de données des participants
- Codes source et supports de formation remis aux participants
- Formation dispensée par un data scientist expérimenté (pratique de R depuis 7 ans minimum)



# LES SÉRIES TEMPORELLES AVEC R

Apprendre à représenter, décrire et modéliser des données temporelles (ou chronologiques) pour pouvoir prédire de l'information, tels sont les enjeux de ce stage. Vous découvrirez aussi comment extraire de l'information utile et appliquer des techniques de machine learning grâce aux nombreux packages R dédiés aux séries temporelles.



## Durée

- 1 jour (possibilité 2 jours)



## Tarif

- Inter : 550€ HT / participant (6 participants maximum)
- Intra : nous consulter



## Lieu

- Locaux de Dcid Consulting (Orléans)
- Locaux de votre entreprise (Région Centre et Paris)



## Public

Tout public ayant déjà des connaissances en programmation et/ou analyse de données

- Analyse de données ★★★★★
- Statistique ★★★★★
- Programmation ★★★★★
- Base de données ★★★★★
- Curiosité ★★★★★



## Renseignements

- [contact@dcid-consulting.fr](mailto:contact@dcid-consulting.fr)
- 02 38 25 33 90



## Prérequis

- Ordinateur individuel
- Connexion internet



## Ressources utiles

- [Forecasting : Principles and Practice](#)
- [R project, time series](#)
- [Statmethods](#)
- [A little book of R for time series](#)



## Objectifs pédagogiques

- Maîtriser les fondamentaux théoriques des séries temporelles
- Maîtriser l'usage des principaux packages (lubridate, zoo, xts, etc.)
- Savoir modéliser des séries temporelles et prédire de l'information
- Savoir extraire de l'information et réaliser du clustering de séries temporelles
- Savoir représenter graphiquement les données temporelles



## Programme

- Découvrir les séries temporelles et ses principes fondamentaux
  - Définition et description
  - Représentation graphique simple
- Manipuler les données au format "times series"
  - Découverte et prise en main des packages "lubridate", "zoo" et "xts"
- Modéliser et prédire des séries temporelles
  - Les régressions simples et multiples
  - La décomposition et le lissage
- Réduire la dimension et extraire de l'information
  - Les séries temporelles en grande dimension
  - Le clustering et l'extraction d'information
- Représenter les séries temporelles
  - Représentation graphique avancée
  - Les packages de représentation dynamique
- Et après ? Perspectives et cas d'usage



## Méthodes et moyens pédagogiques

- Alternance d'exposés et d'applications pratiques
- Possibilité de travailler sur les jeux de données des participants
- Codes source et supports de formation remis aux participants
- Formation dispensée par un data scientist expérimenté (pratique de R depuis 7 ans minimum)



## LE TEXT MINING AVEC R

Le text mining (ou fouille de textes) est un sous domaine de l'intelligence artificielle. Cette technique permet d'extraire de l'information dans des données textuelles grâce à des méthodes et algorithmes d'analyse et d'interprétation. R dispose de nombreux packages dédiés au text mining. Cette formation fera de vous un 'text miner' aguerri.



### Durée

- 1 jour (possibilité 2 jours)



### Tarif

- Inter : 550€ HT / participant (6 participants maximum)
- Intra : nous consulter



### Lieu

- Locaux de Dcid Consulting (Orléans)
- Locaux de votre entreprise (Région Centre et Paris)



### Public

Tout public ayant déjà des connaissances en programmation et/ou analyse de données

- Analyse de données ★★★★★
- Statistique ★★★★★
- Programmation ★★★★★
- Base de données ★★★★★
- Curiosité ★★★★★



### Renseignements

- [contact@dcid-consulting.fr](mailto:contact@dcid-consulting.fr)
- 02 38 25 33 90



### Prérequis

- Ordinateur individuel
- Connexion internet



### Ressources utiles

- [Wikipedia](#)
- [R project, NLP](#)
- [Edutech wiki, U. de Genève](#)
- [Bradley Boehmke, U. de Cincinnati](#)



### Objectifs pédagogiques

- Maîtriser les fondamentaux théoriques des données textuelles
- Maîtriser l'import et la manipulation de données au format texte
- Savoir prétraiter les données textuelles (nettoyage, tokenisation, stemming, etc.) et les représenter graphiquement
- Modéliser des données textuelles : analyse sémantique et topic modeling



### Programme

- Découvrir les données textuelles et ses principes fondamentaux  
Définition, description et représentation
- Importer et manipuler les données textuelles  
Les sources de données texte  
Traitement des chaînes de caractères et expression régulière
- Prétraiter les données textuelles  
Le nettoyage, la ponctuation et les 'stopwords'  
La tokenisation ou unité basique lexicale  
Le stemming ou l'extraction du radical  
La transformation en 'term-document matrix'
- Modéliser les données textuelles  
Analyse sémantique  
Topic modeling et clustering
- Représenter graphiquement les données textuelles  
Représentation graphique avancée  
Les packages de représentation dynamique
- Et après ? Perspectives et cas d'usage



### Méthodes et moyens pédagogiques

- Alternance d'exposés et d'applications pratiques
- Possibilité de travailler sur les jeux de données des participants
- Codes source et supports de formation remis aux participants
- Formation dispensée par un data scientist expérimenté (pratique de R depuis 7 ans minimum)

## APPLI WEB AVEC R & SHINY

Créer sa propre appli dynamique pour le web, partager ses plus beaux supports de dataviz, donner de la liberté à vos utilisateurs. Tout cela est possible grâce à la librairie Shiny de R Studio. Ce stage vous permettra de maîtriser les fondamentaux de Shiny et de créer vos premières applications dynamiques.



### Durée

- 1 jour (possibilité 2 jours)



### Tarif

- Inter : 550€ HT / participant (6 participants maximum)
- Intra : nous consulter



### Lieu

- Locaux de Dcid Consulting (Orléans)
- Locaux de votre entreprise (Région Centre et Paris)



### Public

Tout public ayant déjà des connaissances en programmation et/ou analyse de données

- Analyse de données ★★★★★
- Statistique ★★★★★
- Programmation ★★★★★
- Base de données ★★★★★
- Curiosité ★★★★★



### Renseignements

- [contact@dcid-consulting.fr](mailto:contact@dcid-consulting.fr)
- 02 38 25 33 90



### Prérequis

- Ordinateur individuel
- Connexion internet



### Ressources utiles

- [Shiny, site officiel](#)
- [Shiny gallery, site officiel](#)
- [Shiny tutorial, site officiel](#)
- [Victor Perrier, Shiny sous le capot](#)



### Objectifs pédagogiques

- Maîtriser les concepts des applications Shiny
- Savoir développer sa propre application
- Savoir personnaliser l'apparence de son appli
- Savoir restituer des éléments de type table de données, graphique et image



### Programme

- Découvrir les concepts des applications Shiny  
Commençons par quelques exemples  
Comment ça marche ?  
Installation
- Développer une première appli  
Définition du besoin et choix du design  
Paramétrage de l'affichage et de la restitution : table, graphique et analyse
- Personnaliser l'apparence de son appli  
La disposition de l'appli, les éléments HTML et composants CSS  
Les différents widgets
- Personnaliser les éléments de restitution  
Restitution de tables, images, graphiques, textes, etc.  
Intégration de scripts R
- Découvrir des fonctions avancées  
Gestion des événements  
Partage de son application  
Exemples d'applications avancées
- Et après ? Perspectives et cas d'usage



### Méthodes et moyens pédagogiques

- Alternance d'exposés et d'applications pratiques
- Possibilité de travailler sur les jeux de données des participants
- Codes source et supports de formation remis aux participants
- Formation dispensée par un data scientist expérimenté (pratique de R depuis 7 ans minimum)



## TRAITEMENT DE DONNÉES IOT AVEC R

212 milliards d'objets connectés d'ici 2020 (source IDC), mais une donnée collectée et stockée ne dégage aucune valeur pour l'entreprise. Dans ce stage, vous apprendrez à utiliser R pour importer et manipuler les données issues de l'IoT. Vous découvrirez également comment extraire des séquences d'information, des motifs récurrents (patterns), et ainsi dégager une vraie valeur des objets connectés.



### Durée

- 1 jour (possibilité 2 jours)



### Tarif

- Inter : 550€ HT / participant (6 participants maximum)
- Intra : nous consulter



### Lieu

- Locaux de Dcid Consulting (Orléans)
- Locaux de votre entreprise (Région Centre et Paris)



### Public

Tout public ayant déjà des connaissances en programmation et/ou analyse de données

- Analyse de données ★★★★★
- Statistique ★★★★★
- Programmation ★★★★★
- Base de données ★★★★★
- Curiosité ★★★★★



### Renseignements

- [contact@dcid-consulting.fr](mailto:contact@dcid-consulting.fr)
- 02 38 25 33 90



### Prérequis

- Ordinateur individuel
- Connexion internet



### Ressources utiles

- L'Usine digitale
- Forecasting : Principles and Practice



### Objectifs pédagogiques

- Maîtriser les fondamentaux des données de l'IoT
- Savoir importer les données issues des objets connectés
- Maîtriser la manipulation et le traitement des données de l'IoT
- Savoir extraire de l'information, réaliser du clustering, découvrir des patterns récurrents
- Savoir représenter graphiquement les données de l'IoT



### Programme

- Découvrir les données de l'IoT et ses principes fondamentaux  
Définition et description  
Représentation graphique simple
- Importer les données  
Connexion à une base de données
- Manipuler et prétraiter les données  
Analyse du timestamp, évaluation de la qualité des données  
Traitement des données manquantes et aberrantes  
Normalisation des données
- Modéliser et prédire les données  
Les méthodes de régression  
La décomposition et le lissage
- Réduire la dimension et extraire de l'information  
Les données temporelles en grande dimension  
Le clustering, l'extraction d'information et de patterns récurrents
- Représenter les données de l'IoT  
Représentation graphique avancée  
Les packages de représentation dynamique
- Et après ? Perspectives et cas d'usage



### Méthodes et moyens pédagogiques

- Alternance d'exposés et d'applications pratiques
- Possibilité de travailler sur les jeux de données des participants
- Codes source et supports de formation remis aux participants
- Formation dispensée par un data scientist expérimenté (pratique de R depuis 7 ans minimum)