

AVISIA

An aerial photograph of a city street, likely Paris, showing buildings and a road. A faint network of white lines and dots is overlaid on the image, suggesting a digital or data theme.

DATA, DIGITAL
& TECHNOLOGY

Catalogue de formation

Détail du contenu des formations

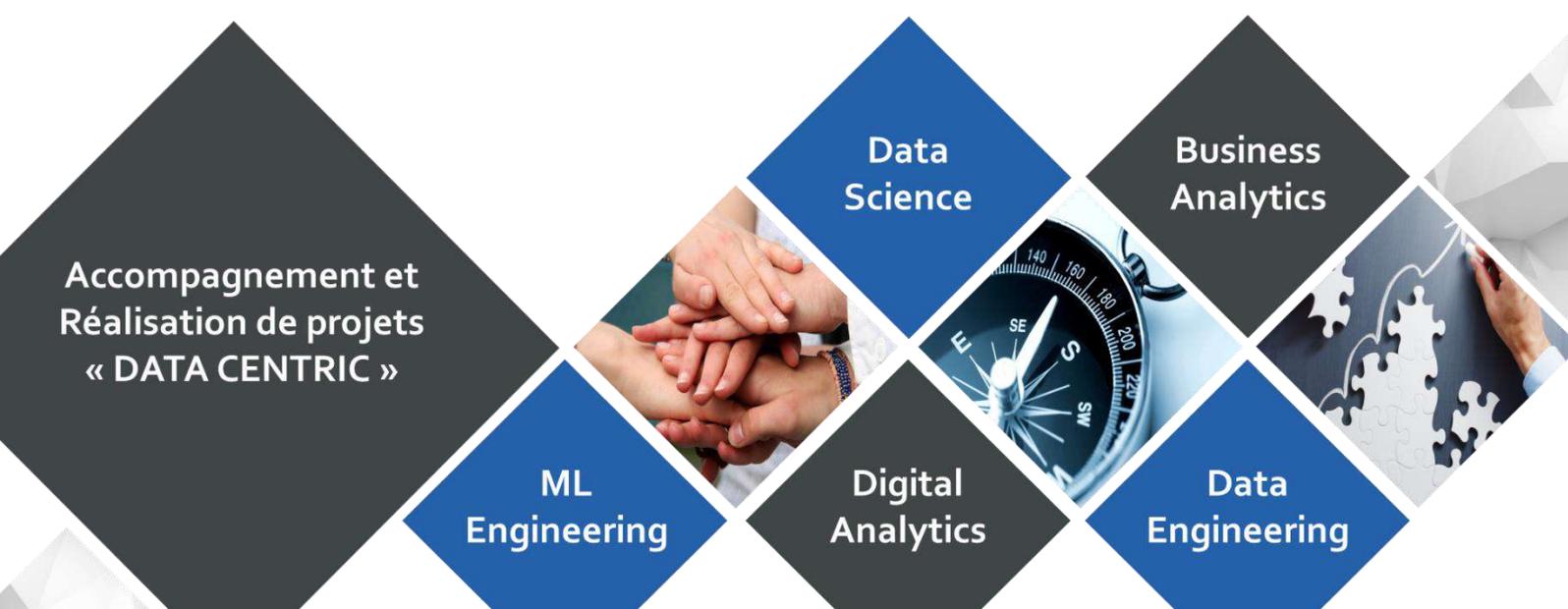
Référence : AVI\2021\FORMATIONS

QUI SOMMES NOUS ?

Fondé en 2007, AVISIA est devenu au fil des ans un acteur de référence dans le domaine de la Data avec une ambition simple :

Faire un **métier de service**
 en tant que **partenaire des Hommes & Femmes**
 qui sont **mobilisés dans la révolution de l'IA**

Dans nos interventions au quotidien, nos consultants mettent ainsi au service de nos clients des compétences à haute valeur ajoutée sur les métiers suivants :

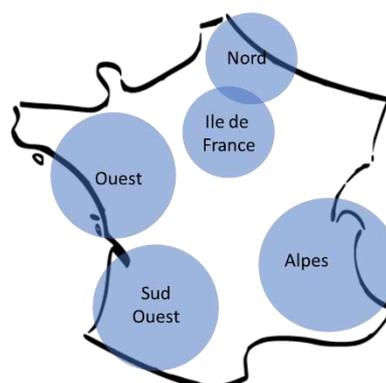


Au travers de notre Académie de Formation, nous souhaitons partager et transmettre cette expertise à nos clients pour les rendre les plus autonomes possibles. Cela passe par 2 aspects :

Pour labeliser la qualité de notre expertise, nous sommes engagés dans des partenariats éditeurs et nos formateurs sont ainsi certifiés :



Le positionnement stratégique de nos agences nous permet d'animer des formations pour vos équipes sur l'ensemble du territoire :





NOTRE VISION DE LA FORMATION

« 30% des compétences acquises sont obsolètes au bout de 2 ans. Sans formation, chaque personne peut ainsi être dépassée au bout de 6 ans »

C'est par ce prisme que nous assurons des formations Expertes, Concrètes, Adaptées & Captivantes :

Une offre adaptée aux problématiques terrain

Nos cycles de formations sont enrichis de nombreux exemples concrets issus de cas développés par nos consultants et inspirés de leurs expériences sur le terrain.

CONCRETES

EXPERTES

Une expertise qui répond à vos besoins

Nous avons une capacité d'appréhender toutes les problématiques et les retranscrire dans les supports de cours.

ADAPTÉES

Des formules flexibles qui s'adaptent à vous

Sessions inter ou intra entreprises, dans nos locaux ou chez vous ou à distance. Possibilité de création de cours personnalisés avec vos propres données...

Des formateurs passionnés par la Data

Nos consultants / formateurs interviennent en mission de secteurs très variés, et sauront s'adapter à chaque niveau et contexte.

Leur volonté est de transmettre leur savoir en permettant aux participants de se l'approprier.

CAPTIVANTES

De l'identification de vos attentes jusqu'à la réalisation des formations, nous vous intégrons pleinement dans chaque phase :



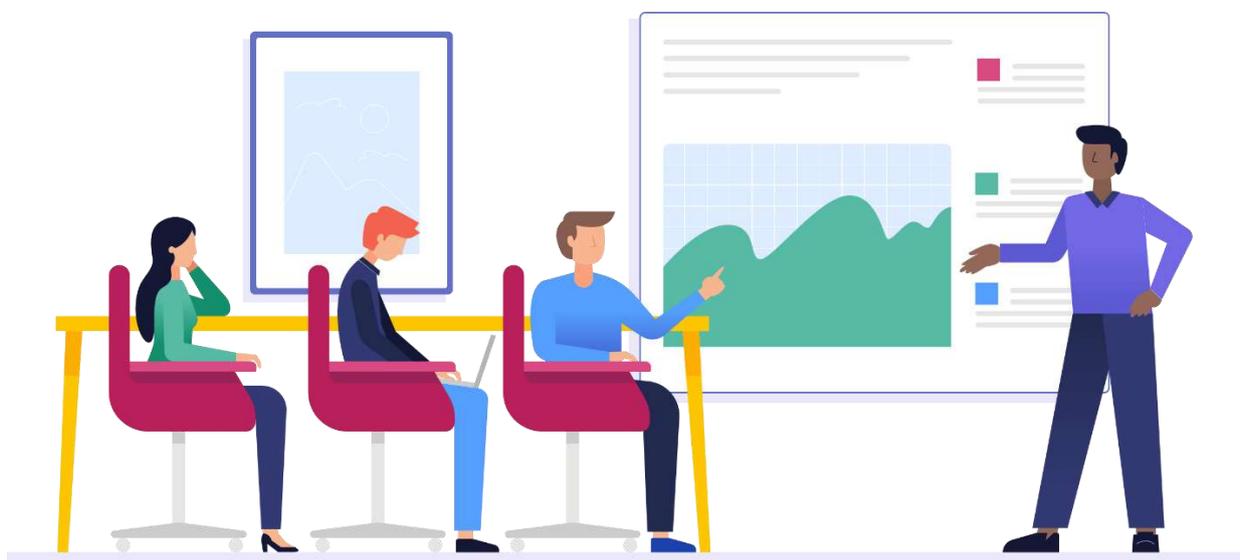
MODALITE D'APPRENTISSAGE

Toutes les agences Avisia proposent des classes virtuelles ou des formations dans les locaux du client.

A Paris, nous pouvons également accueillir les stagiaires dans une salle dédiée à la formation équipée d'ordinateurs dans lesquels seront installés les outils nécessaires à la formation (Machines virtuelles, langage de programmation, Cluster Big Data, ...).

Ainsi les stagiaires manipulent l'outil et réalisent des exercices.

Le formateur présente le cours via un système vidéo. Les supports de formation sont composés de théorie et de cas pratiques.



Dans le cadre de situation handicap, la formation se fera à distance ou dans les locaux du client. Notre structure de formation n'est pas adaptée aux stagiaires en situation de handicap.

ILS NOUS ONT FAIT CONFIANCE

Rejoignez une communauté de plus de 100 clients que nous avons déjà accompagné :



Et la vôtre c'est pour quand ?



Table des matières

<i>QUI SOMMES NOUS ?</i>	2
<i>NOTRE VISION DE LA FORMATION</i>	3
<i>MODALITE D'APPRENTISSAGE</i>	5
<i>ILS NOUS ONT FAIT CONFIANCE</i>	6
1. DATAIKU	8
1.1. DATAIKU : DSS Newcomer.....	9
1.2. DATAIKU : DSS Designer	10
1.3. DATAIKU : DSS Deployer	11
1.4. DATAIKU : DSS Admin	12
2. SAS	13
2.1. SAS Enterprise Guide : Utilisation de l'interface	14
2.2. Programmation SAS I : Fondamentaux	15
2.3. SAS Visual Analytics 1 pour SAS Viya : les fondamentaux	16
2.4. Programmation SAS Macro.....	17
2.5. SAS SQL : Utilisation du SQL dans SAS.....	18
2.6. Programmation SAS II : gestion des données	19
2.7. Programmation SAS III : Techniques avancées.....	20
3. PYTHON	21
3.1. Initiation à Python	22
3.2. Initiation à Python pour le DataManagement	23
3.3. Machine Learning avec Python.....	24
3.4. Moteur de Recommandation avec Python	25
4. LANGAGE R	26
4.1. Langage R pour la Data Science.....	27
5. DATA SCIENCE	28
5.1. Fondamentaux de la DataScience.....	29
6. BIG DATA	30
6.1. Initiation au Big Data avec Hadoop, Hive et Spark	31
7. POWER BI	32
7.1 Initiation à POWER BI.....	33
8. DIGITAL	34
8.1 Initiation à Google Analytics	35
8.2 Initiation à Google Tag Manager	36
8.3 Initiation à Tealium	37

1. DATAIKU

AVISIA est partenaire de DATAIKU



Les formations Dataiku se composent de 4 cycles de formation :

#DEBUTANT

1

NEWCOMER

Découvrir les principales fonctionnalités de l'outil

1 jour

#CONFIRME

2

DESIGNER

Découvrir les fonctionnalités avancées

1 jour

#EXPERT

3

DEPLOYER

Déployer en production des produits data

1 jour

#ADMIN

4

ADMIN

Gestion des profils, sécurité des ressources

2 jours



1.1. DATAIKU : DSS NEWCOMER



Objectifs

Découvrir les principales fonctionnalités offertes par l'outil : Datamanagement, Data Visualisation et modélisation.

Les stagiaires apprendront, sur un exemple de projet et à l'aide des *recipés* visuelles, à importer et manipuler des données (filtres, jointures, agrégations) et à réaliser des graphiques.



Compétences développées

- » Prise en main de la plateforme Dataiku DSS
- » Utilisation des *recipés* visuelles pour créer un workflow de données



Prérequis

Aucun

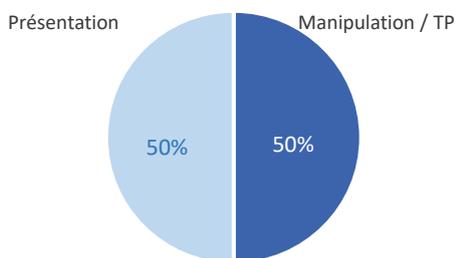


Durée

1 jour (soit 7 heures)



Répartition des activités pédagogiques



Plan de Formation

- » Introduction à Data Science Studio (DSS)
 - Qu'est-ce que DSS ?
 - Les principaux concepts de DSS
 - Architecture
 - Les datasets
 - Les *recipés*
 - Les projets et flow
 - Créer un projet
- » Le Datamanagement avec DSS
 - Importer des jeux de données, et visualiser les données
 - La *recipé* « Prepare » pour nettoyer / transformer les données
 - La *recipé* « Sample/Filter » pour filtrer des datasets
 - La *recipé* « Stack » pour combiner des datasets
 - La *recipé* « Group » pour agréger les données
 - La *recipé* « Join with » pour fusionner des datasets
- » Les graphiques avec DSS
 - Les graphiques de base (Boxplots, histogrammes, pie, nuages de points)
 - Les cartes
- » La modélisation via le moteur visuel sur un cas de classification :
 - Le mode lab pour créer des modèles :
 - En mode automatique
 - En mode expert
 - Les *recipés* « Predict », « Cluster » et « Evaluate » pour appliquer les modèles.

1.2. DATAIKU : DSS DESIGNER



Objectifs

Les stagiaires apprendront à entrainer des modèles et les déployer via l'interface visuelle.

Ils découvriront aussi les fonctionnalités avancées de DSS, apprendront à utiliser du code R et Python dans DSS, et à développer des plugins et macros.



Compétences développées

- » Utilisation des fonctionnalités avancées de DSS
- » Développer et utiliser du code R ou Python dans DSS à l'aide des notebooks
- » Développer un plugin sous DSS



Prérequis

- » Avoir suivi la formation *Dataiku DSS - Newcomer*
- » Avoir des notions en programmation Python, R et SQL

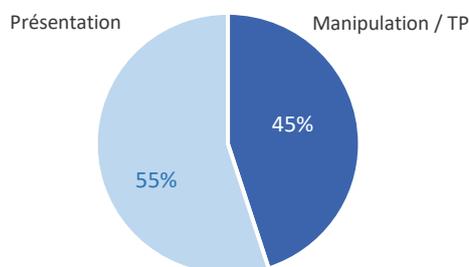


Durée

1 jour (soit 7 heures)



Répartition des activités pédagogiques



Plan de Formation

- » Les *recipe* visuelles avancées (Pivot, Window)
- » Présentation des fonctionnalités avancées dans DSS : le code
- » L'API interne de Dataiku pour le code R et Python
- » La gestion des environnements virtuels R et Python
- » Les notebooks dans DSS
- » Les notebooks SQL, Hive et Impala dans DSS
- » Les variables et folders
- » Utilisation et Ecriture de plugins
- » Les macros



1.3. DATAIKU : DSS DEPLOYER



Objectifs

Cette formation s'adresse à ceux qui déploient en production des produits data. Elle vous apportera une bonne compréhension des différents nœuds de DSS et de leur interopérabilité pour mettre en production vos *use cases* en mode batch ou temps réel.



Compétences développées

- » Déploiement et industrialisation des modèles de machine learning dans DSS
- » Utilisation de l'*automation Node* ou de l'*API Node* selon les cas d'usages
- » Mise en place de contrôles de qualité de données



Prérequis

- » Avoir suivi la formation *Dataiku DSS – Newcomer*
- » Avoir des notions en programmation Python, R et SQL

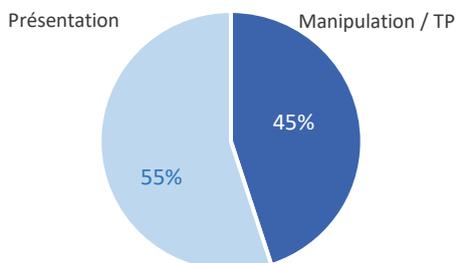


Durée

1 jour (soit 7 heures)



Répartition des activités pédagogiques



Plan de Formation

- » Le partitionnement de Datasets dans DSS
- » Les checks et métriques (*y/c custom*)
- » La mise en place de scénarios en utilisant les recettes visuelles, les déclencheurs, et reportings pour automatiser l'enrichissement de données et le réapprentissage de modèle
- » Le déploiement d'un projet sur l'*Automation Node* avec les *bundles*
- » L'*API Node* pour exposer des modèles visuels ou custom sous python via un *API endpoint*

1.4. DATAIKU : DSS ADMIN



Objectifs

Cette formation s'adresse à ceux qui déploient en production des produits data. Cette formation vous apportera une bonne compréhension des différents nœuds de DSS, et de leur interopérabilité pour mettre en production vos use cases en mode batch ou temps réel.

Découvrez ces concepts en déployant un modèle d'IA en utilisant le *score node*.



Compétences développées

- » Administration de la plateforme DSS
- » Gestions des profils
- » Management de la sécurité et des ressources



Prérequis

- » Avoir suivi la formation *Dataiku DSS – Newcomer*
- » Avoir une expérience d'administration de logiciels en entreprise

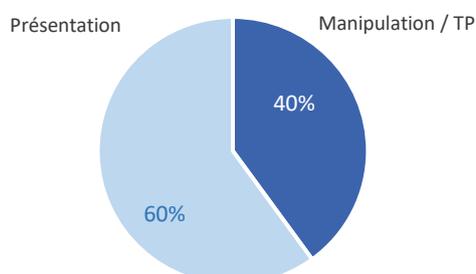


Durée

2 jours (soit 14 heures)



Répartition des activités pédagogiques



Plan de Formation

Vous découvrirez comment fonctionne l'architecture de DSS, comment manager la sécurité de la plateforme, et comment résoudre les différents problèmes fréquemment rencontrés.

Les principaux concepts abordés sont :

- » L'architecture technique et fonctionnelle de DSS
- » La sécurité : les users, les groupes, la gestion des connections
- » Installation, configuration et upgrade : guidelines
- » Les logs, les problèmes fréquents
- » Les *code environments*, le monitoring, les macros
- » L'API publique de DSS
- » Le management des ressources



2. SAS

AVISIA est partenaire de SAS



Les formations SAS se composent de plusieurs cycles de formation. Vous trouverez ci-dessous les principales :

#DEBUTANT

1

SAS ENTERPRISE GUIDE

Produire en quelques clics et sans écrire aucune ligne de code, différents types d'états (rapport, graphique...) 2 jours

#CONFIRME

2

SAS MACRO

Automatiser les traitements de manière à pouvoir les réexécuter 2 jours

#DEBUTANT

1

SAS I : FONDAMENTAUX

Se familiariser avec la syntaxe et les concepts fondamentaux de la programmation 3 jours

#CONFIRME

2

SQL DANS UNE SESSION SAS

SQL, langage très répandu, permettant d'interroger des bases de données relationnelles (SGBD) 2 jours

#DEBUTANT

1

SAS VA 1 POUR SAS VIYA

Introduction à la préparation de données, l'exploration des données et à la création de rapport en utilisant SAS Visual Analytics 2 jours

#CONFIRME

2

SAS II : GESTION DES DONNEES

Découvrir les techniques avancées de la manipulation des données 2 jours

2.1. SAS ENTERPRISE GUIDE : UTILISATION DE L'INTERFACE



Objectifs

Vous êtes analyste ou créateur de rapports, et souhaitez produire en quelques clics et sans écrire aucune ligne de code, différents types d'états (rapport, tableau statistique, graphique, etc ...) ?

Cette formation s'adresse aux utilisateurs qui souhaitent accéder, analyser, gérer et résumer des données de différentes sources, présenter des résultats en utilisant l'interface SAS Enterprise Guide (SEG).



Compétences développées

Utilisation du logiciel SAS Enterprise Guide en mode clic bouton pour charger et transformer des données



Prérequis

- » Avoir des notions de bases de données
- » Aucune connaissance de SAS ou d'un langage de programmation n'est nécessaire

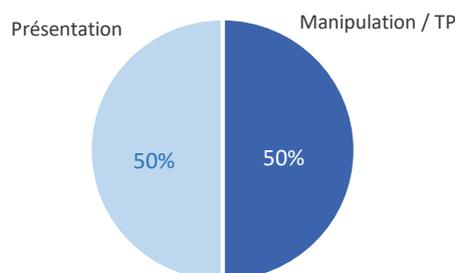


Durée

2 jours (soit 14 heures)



Répartition des activités pédagogiques



Plan de Formation

Jour 1

- » Accéder aux données
 - Comprendre la structure des données SAS
 - Accéder à des tables SAS et des tables de SGBD
 - Importer des fichiers de données externes
- » Utiliser les tâches
 - Introduction aux tâches et aux assistants
 - Créer un rapport de fréquence
 - Générer des rapports HTML, PDF, et RTF
 - Créer un rapport sous forme de tableau
 - Filtrer les données dans une tâche
 - Créer un graphique
- » Créer des requêtes simples
 - Filtrer et trier les données
 - Créer de nouvelles colonnes
 - Grouper et sommer des données
 - Joindre des tables

Jour 2

- » Créer des sorties synthétisées
 - Générer des statistiques de synthèse
 - Créer un rapport de synthèse avec les assistants
- » Utiliser les paramètres dans les tâches et les requêtes
 - Paramétrer un projet
 - Créer et utiliser un paramètre dans une tâche
 - Créer et utiliser un paramètre dans une requête
- » Personnaliser et organiser son projet
 - Combiner des résultats
 - Mettre à jour et organiser ses projets
- » Ecrire et soumettre un programme SAS



2.2. PROGRAMMATION SAS I : FONDAMENTAUX



Objectifs

Après avoir découvert les environnements de travail SAS, vous vous familiariserez avec la syntaxe et les concepts fondamentaux de la programmation.

Au terme de la formation, vous serez en mesure de créer vos propres tables de données, de les étudier et de générer des rapports au format HTML, PDF ou encore RTF.



Compétences développées

- » Développement de programmes SAS grâce aux fonctionnalités de base
- » Chargement et transformation de données avec l'étape *DATA*
- » Transformation de données à l'aide des procédures *PROC*



Prérequis

- » Avoir de préférence des notions en programmation
- » Aucune expérience sur SAS n'est requise

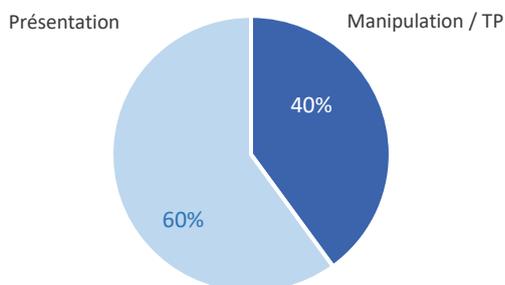


Durée

3 jours (soit 21 heures)



Répartition des activités pédagogiques



Plan de Formation

Jour 1

- » Les bases et les concepts du langage SAS®
 - Syntaxe
 - Exécution
 - Débogage
- » Phases de compilation et d'exécution des programmes SAS®
- » Accès aux données : bibliothèques et structure des tables
- » Lecture de fichiers :
 - données brutes (délimités, standards ou non)
 - feuilles Microsoft Excel,
 - tables SAS ou issues d'un SGBD,...

Jour 2

- » Exploration et description des données
- » Manipulation de données
 - Filtres
 - Sélection et création de variables
 - Traitement conditionnel
 - Utilisation de fonctions SAS® pour manipuler les dates
 - Utilisation et création de formats personnalisés

Jour 3

- » Combinaison de tables
 - Concaténation et imbrication
 - Fusion des correspondances et des non correspondances
- » Création et mise en forme de rapports de détail ou agrégés
 - Les procédures *FREQ*, *MEANS*, *UNIVARIATE*
 - L'Output Delivery System (*ODS*)

2.3. SAS VISUAL ANALYTICS 1 POUR SAS VIYA : LES FONDAMENTAUX



Objectifs

Cette formation vous propose une introduction à la préparation de données, à l'exploration des données et à la création de rapport en utilisant SAS Visual Analytics 8.x.



Compétences développées

- » Interagir avec la page d'accueil
- » Visualiser et naviguer dans un rapport avec le Viewer
- » Accéder aux données
- » Préparer les données avec SAS Data Studio
- » Explorer et analyser les données avec SAS Visual Analytics
- » Créer des rapports interactifs à l'aide de SAS Visual Analytics 8.x



Prérequis

Aucun prérequis

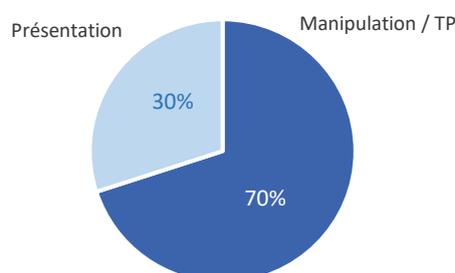


Durée

2 jours (soit 14 heures)



Répartition des activités pédagogiques



Plan de Formation

Jour 1

- » Débuter avec SAS Visual Analytics
 - Introduction à SAS Visual Analytics
 - Présentation de l'environnement de formation
 - Visualiser des rapports avec le viewer
- » Préparer les données avec SAS Visual Analytics
 - Accéder aux données
 - Nettoyer les données en utilisant SAS Data Studio

Jour 2

- » Explorer les données avec SAS Visual Analytics
 - Travailler les éléments de données
 - Explorer les données à l'aide de représentations graphiques
 - Créer de nouveaux éléments de données et appliquer des filtres
 - Lier les graphiques
- » Créer des rapports avec SAS Visual Analytics
 - Rapports simples
 - Rapports interactifs
 - Utiliser des règles d'affichages



2.4. PROGRAMMATION SAS MACRO



Objectifs

Vous souhaitez automatiser vos traitements de manière à pouvoir les réexécuter en ne modifiant que quelques paramètres et ainsi éviter de devoir les modifier physiquement ?

L'objectif de cette formation est de comprendre comment fonctionne le macro processeur au sein du langage de programmation SAS.

Vous pourrez alors automatiser vos *reportings*, rendre lisibles et exploitables vos programmes SAS, et dégager du temps pour des analyses à valeur ajoutée.



Compétences développées

- » Automatisation de traitements à l'aide des macro variables et macro programmes



Prérequis

- » Connaissance du langage SAS
- » Ou avoir suivi la formation « Programmation SAS I : Fondamentaux »

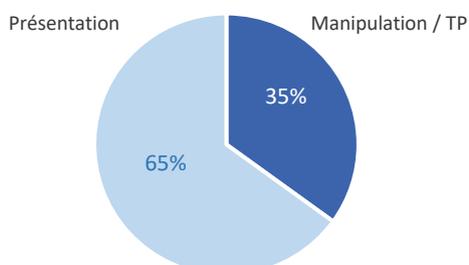


Durée

2 jours (soit 14 heures)



Répartition des activités pédagogiques



Plan de Formation

Jour 1

- » Introduction au langage Macro
- » Les Macros variables :
 - Présentation
 - Macro-variables automatiques et utilisateurs
 - Référencement de Macros variables
 - Délimitation des noms de Macros variables
 - Fonctions Macro
- » Définition de programmes Macros :
 - Définition et appel de programmes Macro
 - Macros programmes avec paramètres

Jour 2

- » La programmation Macro : traitement conditionnel et itératif
- » L'étape DATA et l'interface SQL :
 - Création de Macros variables durant l'étape DATA (*CALL SYMPUTX*)
 - Référence indirecte de Macros variables
 - Récupération de la valeur d'une Macro variable dans l'étape DATA (*SYMGET*)
 - Création des Macros variables avec du SQL

2.5. SAS SQL : UTILISATION DU SQL DANS SAS



Objectifs

Cette formation vous permettra de découvrir le *Structured Query Language* (SQL), un langage très répandu permettant d'interroger des bases de données relationnelles (SGBD) ou directement des tables SAS.

Elle vous permettra d'opter pour la technique de manipulation de données la plus appropriée à vos besoins. Vous verrez, comment l'utilisation des fonctions SAS peut être combinée avec le langage SQL.



Compétences développées

- » Développement de scripts SQL dans l'environnement SAS



Prérequis

- » Connaissance du langage SAS
- » Ou avoir suivi la formation « Programmation SAS I : Fondamentaux »

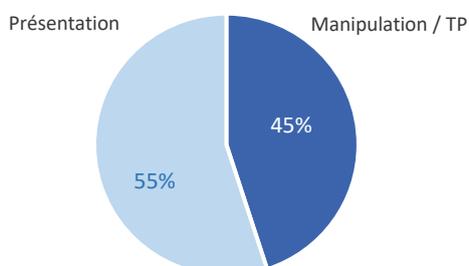


Durée

2 jours (soit 14 heures)



Répartition des activités pédagogiques



Plan de Formation

Jour 1

- » Présentation du langage SQL
- » Rédaction et structure de requêtes standard et agrégées
- » Utilisation de fonctions SAS dans les procédures SQL
- » Création de requêtes imbriquées
- » Personnalisation des sorties générées

Jour 2

- » Manipulation de plusieurs tables à l'aide des opérateurs
 - De jointures internes et externes
 - D'union, d'intersection...
- » Création de tables et de vues
- » Gestion et maintenance des tables via le langage SQL
- » Les options SQL
- » L'utilisation du dictionnaire de métadonnées



2.6. PROGRAMMATION SAS II : GESTION DES DONNEES



Objectifs

Cette formation permet de devenir autonome en découvrant les techniques liées à la manipulation et au contrôle des données créées dans vos tables grâce à l'étape DATA.



Compétences développées

- » Comprendre et contrôler le traitement de l'étape DATA
- » Créer une colonne cumulative et traiter les données par groupe
- » Manipuler les données à l'aide des fonctions
- » Convertir le type de colonnes
- » Créer des formats personnalisés
- » Concaténer et fusionner les tables
- » Etablir un traitement itératif des données
- » Restructurer les tables



Prérequis

Toutes personnes ayant suivi le cours « Programmation SAS I : les fondamentaux »

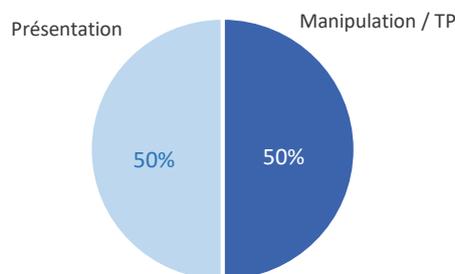


Durée

2 jours (soit 14 heures)



Répartition des activités pédagogiques



Plan de Formation

- » Contrôle de l'étape DATA
- » Fonctionnement (compilation/exécution)
- » Contrôle des sorties
- » Agrégation des données
 - Création et Utilisation de colonnes cumulatives
 - Traitement par groupes dans l'étape DATA
- » Manipulation des données à l'aide de fonctions
 - Comprendre les fonctions SAS et les CALL routines
 - Utilisation des fonctions SAS : numérique / date / caractères
 - Conversion du type des colonnes à l'aide des fonctions spécifiques
- » Création de formats personnalisés
 - Création et utilisation d'un format personnalisé
 - Création d'un format personnalisé à partir d'une table
- » Combinaison de tables
 - Concaténation de tables
 - Jointure de tables
 - Identification des données concordantes et non concordantes
- » Traitement itératif
 - Utilisation des boucles DO
 - Utilisation des boucles DO conditionnelles
- » Restructuration des Tables
 - Restructurer une table avec l'étape DATA
 - Restructurer une table avec une PROC TRANSPOSE

2.7. PROGRAMMATION SAS III : TECHNIQUES AVANCEES



Objectifs

Vous découvrez les techniques de programmation avancées et les ressources machine nécessaires à l'exécution de vos programmes. Vous saurez apprécier l'adéquation entre les techniques proposées et les ressources consommées pour obtenir le programme le plus optimisé.



Compétences développées

- » Comprendre et contrôler le traitement de l'étape DATA
- » Utiliser des fonctions avancées, des fonctions d'expressions régulières PRX et créer des fonctions personnalisées.
- » Traiter le code répétitif, faire pivoter les données et effectuer des recherches de données à l'aide des tableaux.
- » Effectuer des recherches de données à l'aide d'objets de hachage et d'itérateurs de hachage.
- » Créer des modèles de formats numériques en utilisant la procédure FORMAT.



Prérequis

Toutes personnes ayant suivi les cours Programmmations SAS I et SAS II. Mise en pratique des fonctionnalités abordées lors de cette formation (plusieurs mois d'expérience de programmation en SAS sont conseillés).

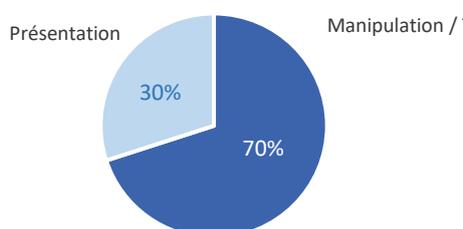


Durée

3 jours (soit 21 heures)



Répartition des activités pédagogiques



Plan de Formation

- » Révision du fonctionnement de l'étape DATA.
- » Utilisation des fonctions avancées
 - LAG, FINDC/FINDW et COUNT/COUNTC/COUNTW
 - Expressions régulières Perl (fonction PRX)
- » Manipulation des données à l'aide des tableaux
 - Définition et référencement des tableaux unidimensionnels (Arrays).
 - Utilisation avancée des tableaux unidimensionnels.
 - Définition et référencement des tableaux bidimensionnels.
- » Traitement des données en mémoire à l'aide des objets de hachage
 - Déclaration des objets de hachage.
 - Définition des objets de hachage.
 - Récupération des valeurs clés dans un objet de hachage.
 - Écriture des données dans une table à partir d'un objet de hachage.
 - Utilisation des objets d'itération de hachage.
- » Utilisation de procédures avancées
 - Création et utilisation des formats PICTURE avec la procédure FORMAT.
 - Création et utilisation des fonctions personnalisées à l'aide de la procédure FCMP.
- » Méthodes d'optimisation du programme
 - Création et utilisation des index
 - Techniques de jointure et combinaison efficaces des tables avec l'étape DATA et la procédure SQL



3. PYTHON

Les formations Python se composent de 4 cycles de formation :

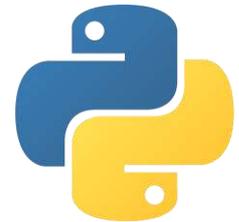
#DEBUTANT

1

INITIATION

Maitriser les bases du langage Python (boucles, fonctions, objets)

1 jour



#CONFIRME

2

DATAMANAGEMENT

Maitriser les différentes méthodes de la librairie pandas

1 jour

#EXPERT

3

MACHINE LEARNING

Découvrir la librairie scikit-learn et apprendre à développer des modèles de Data Science

1 jour

#EXPERT ++

4

RECOMMANDATIONS

Découvrir des moteurs de recommandations

1 jour

3.1. INITIATION A PYTHON



Objectifs

Maîtriser les bases du langage Python (boucles, fonctions, objets).



Compétences développées

- » Développement de scripts python
 - A l'aide des objets standards
 - De pipelines de transformations de données (module *pandas*)



Prérequis

- » Aucuns prérequis

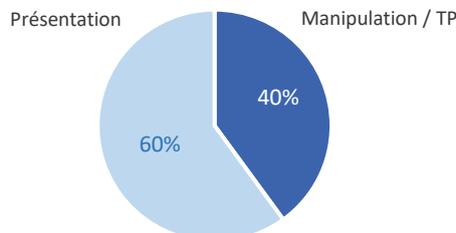


Durée

1 jour (soit 7 heures)



Répartition des activités pédagogiques



Plan de Formation

- » Introduction au langage Python :
 - Qu'est-ce que le langage Python ?
 - Spécificités du langage
 - Comment exécuter du code Python ?
L'environnement de travail
- » Les types de données de base :
 - Les données numériques
 - Les booléens
 - Les chaînes de caractères
 - Les Listes
 - Les tuples
 - Les dictionnaires
- » Les structures conditionnelles
- » Les boucles :
 - La boucle for
 - La boucle while
- » Les fonctions en python :
 - La définition d'une fonction
 - Les fonctions Lambda
- » Le langage modulaire
- » Initiation au langage objet :
 - Les classes
 - Les méthodes
 - Les méthodes des objets String, Listes, Tuples et Dictionnaires
 - L'héritage



3.2. INITIATION A PYTHON POUR LE DATAMANAGEMENT



Objectifs

Maitriser les différentes méthodes de la librairie pandas pour le DataManagement en manipulant l'objet DataFrame (tableaux structurés).



Compétences développées

- » Développement de scripts python
 - A l'aide des objets standards
 - De pipelines de transformations de données (module *pandas*)



Prérequis

- » Connaitre les bases de Python
- » Ou avoir participé à la formation initiation Python

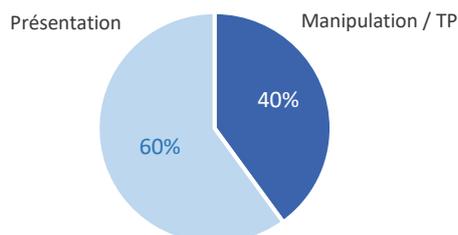


Durée

1 jour (soit 7 heures)



Répartition des activités pédagogiques



Plan de Formation

- » Introduction au module *numpy* :
 - L'objet Array
 - Les méthodes et fonctions sur les array
- » Introduction au module *pandas* – types d'objets « Series » et « DataFrames »
- » Visualiser et accéder aux données
 - Afficher un DataFrame / les méthodes head() et tail()
 - Afficher les caractéristiques d'un DataFrame
 - Les méthodes de sélections de colonnes / de lignes
 - Les index
- » Modifier les données :
 - L'ajout de variables
 - Accéder et modifier des valeurs
 - Les tris
 - La suppression de variables
 - Le traitement des valeurs manquantes
 - Concaténation, fusion de plusieurs tables
 - Exporter un Dataframe

3.3. MACHINE LEARNING AVEC PYTHON



Objectifs

Découvrez le module scikit-learn et apprenez à développer des modèles de Data Science tels que les Random Forest et/ou Gradient Boosting.



Compétences développées

- » Développement de scripts python pour la mise en place de pipelines de *machine learning*



Prérequis

- » Avoir des connaissances en programmation Python
- » Ou avoir suivi la formation « Initiation à Python pour le DataManagement »

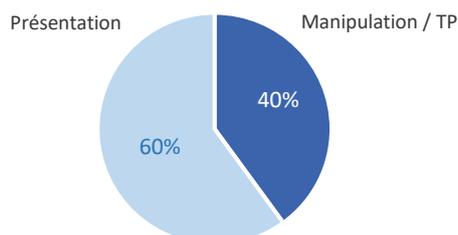


Durée

1 jour (soit 7 heures)



Répartition des activités pédagogiques



Plan de Formation

- » La méthodologie de *scikit learn* (méthodes fit / predict / transform)
- » Le preprocessing sous scikit-learn
- » Les principales méthodes de classification et métriques
 - Régressions logistiques
 - Arbres de décision
 - Random Forest
 - Gradient Boosting
- » Les principales méthodes de régressions et métriques
 - Régressions linéaires
 - Arbres de décision
 - Random Forest
 - Gradient Boosting
- » Les méthodes non supervisées
 - ACP
 - Clustering (Kmeans, CAH)
- » Tuning des hyperparamètres et cross validation
- » Les pipelines, la sauvegarde des modèles



3.4. MOTEUR DE RECOMMANDATION AVEC PYTHON



Objectifs

Introduire les principes théoriques des moteurs de recommandation et pouvoir les mettre en pratique sur Python grâce à scikit-learn ou des packages spécialisés.



Compétences développées

Moteurs de recommandation, moteurs de recommandation implémentés sur Python



Prérequis

- » Avoir des connaissances en programmation Python
- » Ou avoir suivi la formation « Initiation à Python pour le DataManagement »
- » Ou avoir suivi la formation « Machine Learning avec Python »

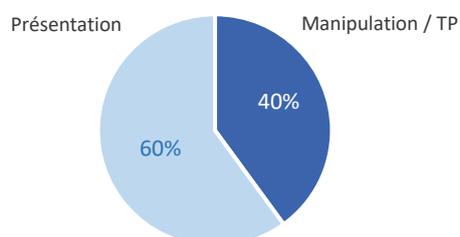


Durée

1 jour (soit 7 heures)



Répartition des activités pédagogiques



Plan de Formation

- » Introduction
- » Les méthodes content based
- » Les méthodes collaborative filtering (Focus méthodes memory based, Focus méthodes model based)
- » Les méthodes hybrides
- » Comparaison, choix et évaluation de son algorithme de recommandation, Conclusion

4. LANGAGE R

La formation sur le Langage R possède 1 cycle de formation :

#DEBUTANT

1

NEWCOMER

Découvrir les principales fonctionnalités de l'outil

3 jours





4.1. LANGAGE R POUR LA DATA SCIENCE



Objectifs

Développer vos connaissances sur les principales bibliothèques modernes.

Manipulation des données avec R et mise en place des démarches de Data Science.



Compétences développées

- » Développement de scripts R
- » Être en mesure de manipuler des données (Dataframe) et faire du datamanagement
- » Entraîner des modèles de machine learning grâce au logiciel R



Prérequis

- » Avoir des notions en développement informatique ainsi qu'en manipulation de données

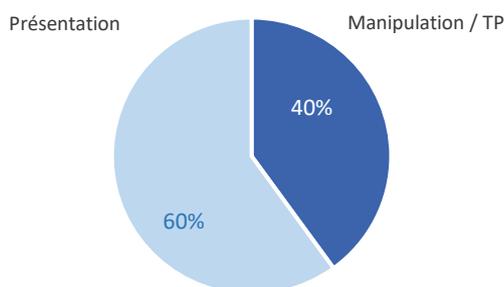


Durée

3 jours (soit 21 heures)



Répartition des activités pédagogiques



Plan de Formation

Jour 1

- » Introduction au langage R :
 - Qu'est-ce que le langage R ?
 - Spécificités du langage
 - Prise en main de R Studio
- » Les structures conditionnelles
- » Les boucles (la boucle for et la boucle while)
- » Les fonctions en R :
 - La définition d'une fonction
 - Le langage objet et R

Jour 2

- » Découverte des packages R de la collection *Tidyverse*
- » Package *readr* : Importation des données
- » Package *Dplyr* / *Tidyr* et *data.table*
 - Visualiser et accéder aux données
 - Les méthodes de sélections de colonnes / de lignes
 - Modifier les données
 - L'ajout de variables
 - Accéder et modifier des valeurs
 - Les tris
 - La suppression de variables
 - Le traitement des valeurs manquantes
 - Les opérations sur les *DataFrame*
 - Les méthodes statistiques (sum, mean, quantile)
 - L'agrégation de données
 - Concaténation, fusion de plusieurs tables
 - Exporter un *Dataframe*
- » Réaliser des premiers graphiques : *ggplot2*
- » Optimisation du code R

Jour 3

- » Méthodologie de machine learning sous R (package *caret*, *xgboost*, *rminer*):
 - Le preprocessing sous R
 - Les principales méthodes de classification et métriques
 - Les principales méthodes de régressions et métriques

5. DATA SCIENCE

La formation sur le Data Science possède 1 cycle de formation :

#DEBUTANT

1

INITIATION

Découvrir les concepts
fondamentaux de la Data Science

2 jours



**DATA
SCIENCE**



5.1. FONDAMENTAUX DE LA DATA SCIENCE



Objectifs

Cette formation a pour objectif de revoir les fondamentaux de la *DataScience* et de découvrir les nouveautés en matière d'algorithmes de *machine learning*.

Dans un premier temps, un rappel des principales définitions sera réalisé, avant de présenter les différents algorithmes et leur cas d'utilisation.

La seconde journée de la formation sera dédiée à la présentation des étapes clé des projets de *DataScience*.



Compétences développées

- » Compréhension et culture générale sur les algorithmes de *machine learning*
- » Compréhension sur le cycle de vie d'un projet de *machine learning*



Prérequis

- » Notions en statistiques et mathématiques

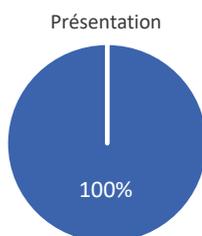


Durée

2 jours (soit 14 heures)



Répartition des activités pédagogiques



Plan de Formation

Jour 1

- » Rappels sur les fondamentaux, définitions
- » Algorithmes supervisés ou non supervisés
- » Les problèmes de classification et de régression
- » L'échantillonnage
- » Le paramétrage des algorithmes et les grilles de paramètres
- » Sur et sous-apprentissage
- » Les erreurs et mesure de performance
- » Les algorithmes non supervisés (l'analyse en composantes principales, la classification hiérarchique, l'algorithme Kmeans)
- » Les algorithmes supervisés
 - Les régressions linéaires : classiques, Ridge, Lasso, ElasticNet
 - La régression logistique
 - Les arbres de décision
 - Les modèles ensemblistes (bagging, les Random Forests, boosting et Gradient Boosting Trees)
- » Introduction aux moteurs de recommandation (les algorithmes : user based, content based, collaborative filtering)

Jour 2 - les fondamentaux des projets datascience

- » Identifier le problème à résoudre
- » La collecte de la donnée et la constitution des indicateurs
- » Les analyses descriptives préalables
- » Feature engineering et sélections (traitement des variables catégorielles, valeurs manquantes, la normalisation des variables)
- » L'échantillonnage
- » Le choix des algorithmes à tester selon le cas d'usage
- » Le paramétrage des algorithmes pour identifier le plus performant sur l'apprentissage
- » La comparaison des algorithmes sur l'échantillon de test et la sélection du meilleur modèle
- » L'industrialisation du meilleur modèle



6.1. INITIATION AU BIG DATA AVEC HADOOP, HIVE ET SPARK



Objectifs

Cette formation s'adresse à ceux qui souhaitent découvrir les concepts fondamentaux du Big Data sous Hadoop. Après vous avoir présenté les concepts clés sur lesquels reposent la technologie Hadoop, vous découvrirez la solution Hive, entrepôts de données stockées sous HDFS, permettant de requêter de très grands volumes de données rapidement grâce à un langage très proche du SQL.

La seconde journée est consacrée au moteur Spark, framework de calcul distribué permettant de développer des applications performantes sur de très gros volumes de données.



Plan de Formation

Jour 1

- » Introduction à Hadoop
 - Le paradigme MapReduce
 - Le stockage avec HDFS
 - Le manager de ressources Yarn
 - Introduction à Hive et son langage HiveQL

Jour 2

- » Introduction au moteur Spark pour le datamanagement
 - Les concepts de Spark
 - La notion de dataframe et les méthodes usuelles sur cet objet
 - Le SQL sous Spark
 - Introduction aux RDD



Compétences développées

- » Initiation à Hadoop
- » Développements de requêtes *HiveQL*
- » Développements de scripts *PySpark* pour manager des données



Prérequis

- » Connaissances en programmation python
- » Connaissances en programmation SQL
- » Notions en système UNIX

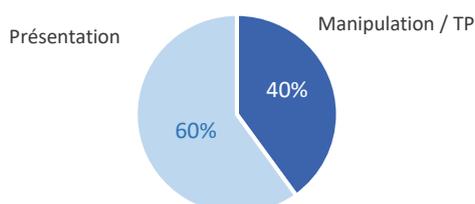


Durée

2 jours (soit 14 heures)



Répartition des activités pédagogiques



7. POWER BI

La formation Power BI se compose d'un cycle de formation :

	POWERBI	
	Découvrir les fonctionnalités de Power BI	<u>3 jours</u>





7.1 INITIATION A POWER BI



Objectifs

- » Découvrir l'outil Power BI
- » Apprendre à utiliser Power BI et être autonome pour créer un report.



Compétences développées

- » Apprendre à préparer les données avec Power Query
- » Créer un modèle de données
- » Visualiser les données
- » Analyser les données
- » Partager ses reports/Dashboard



Prérequis

- » Télécharger PBI Desktop sur son ordinateur
- » Un minimum d'expérience dans le traitement des données.

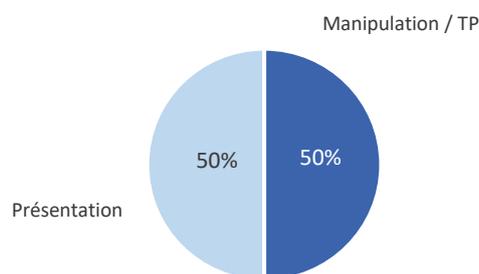


Durée

3 jours (soit 21 heures)



Répartition des activités pédagogiques



Plan de Formation

Jour 1

- » Présentation de Power BI
- » Préparation des données
- » Modélisation des données

Jour 2

- » Visualisation des données
- » Analyse des données
- » Déployer et maintenir

Jour 3

- » Hackathon

8. DIGITAL

AVISIA est partenaire de Google Cloud Platform :



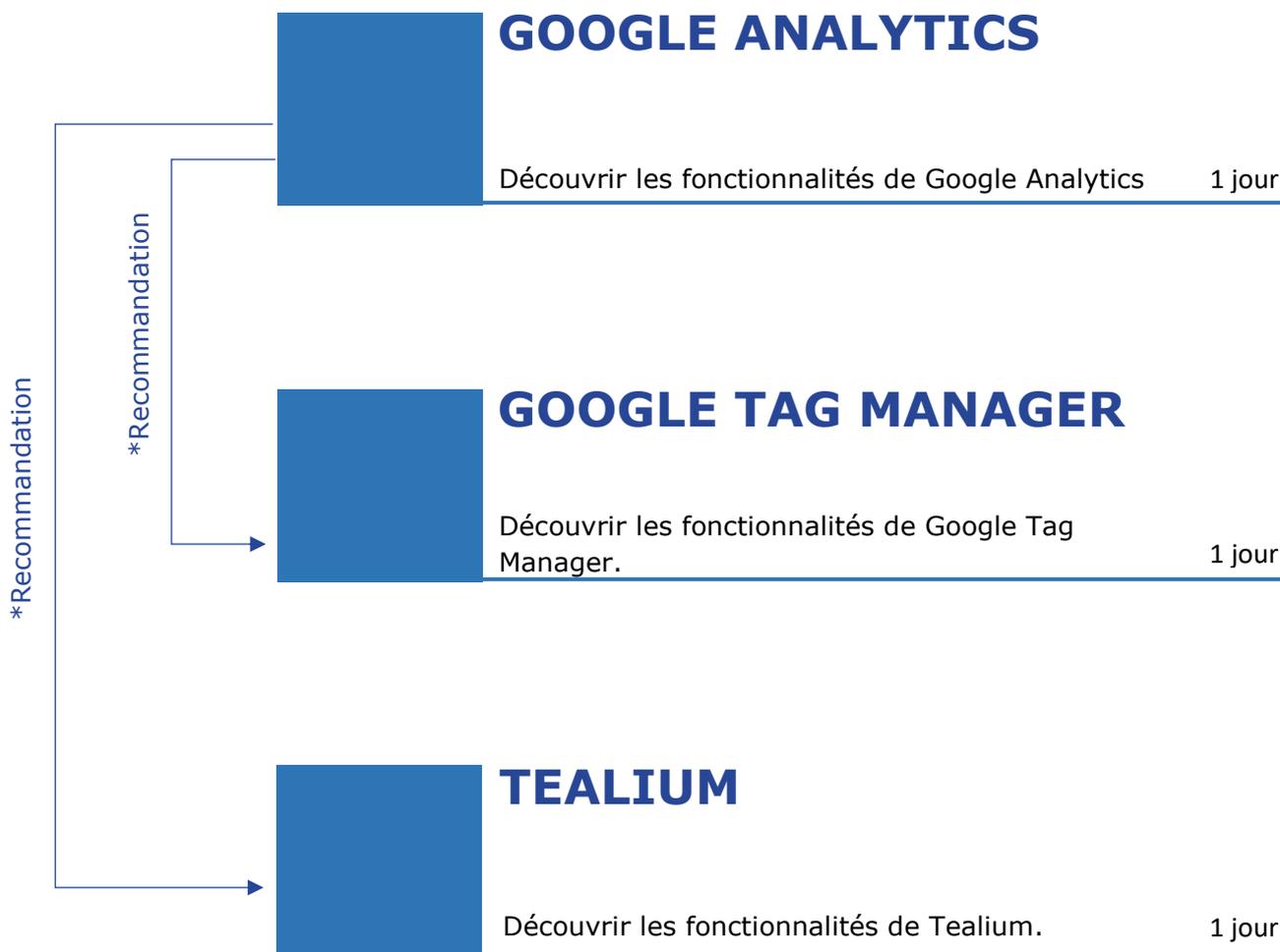
Google Analytics



Google Tag Manager



Les formations digitales se composent de plusieurs cycles de formation. Vous trouverez ci-dessous les 3 premiers :



*Recommandation: Dans le cas où vous souhaiteriez suivre 2 formations ou plus, nous vous recommandons de suivre en premier la formation Google Analytics puis la formation Google Tag Management ou la formation Tealium.



8.1 INITIATION A GOOGLE ANALYTICS



Objectifs

Initiation aux principes et au fonctionnement d'Universal Google Analytics à travers

- Une présentation complète de l'administration
- Une présentation complète des fonctionnalités
- Une présentation détaillée des rapports disponibles



Compétences développées

Configuration, utilisation et analyse des données comportementales Google Analytics



Prérequis

Connaissances digitales

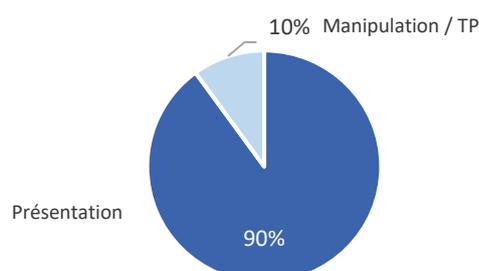


Durée

1 jour (soit 7 heures)



Répartition des activités pédagogiques



Plan de Formation

- » Les concepts clés de Google Analytics
 - Rappel, liaison dataLayer TMS / Plan de marquage
 - Utilisation Desktop Vs Mobile
 - Approche standard versus Server-side
 - GA4 (Analytics Application+ Web)
- » Création et gestion des comptes
 - Gestion des comptes, propriétés, vues et utilisateurs
 - Utilisation des filtres
 - Présentation du script de suivi standard
 - Associations de produits
 - Historisation des modifications
 - Suppression des données
- » Notions clés de Google Analytics
 - Utilisation des dimensions et métriques
 - Définition des visites, visiteurs, taux de rebond...
- » Présentation de l'interface Google Analytics
 - Vue globale de l'interface
 - Utilisation des fonctionnalités transverses (I)
 - Présentation de la section audience et la section acquisition dédiée au suivi du trafic entrant
 - Présentation de la section comportement
 - Présentation de la section conversion
 - Utilisation des fonctionnalités transverses (II)
 - Présentation de la section temps réel
 - Focus mobile - Présentation des spécificités
 - Présentation de la section Analyse
 - Exploitation des données
 - Accès aux données via les API Google Analytics
 - Utilisation d'outils de dataviz
 - Overview sur Google Data Studio & Power BI
 - Exemples + cas pratiques + best practices

8.2 INITIATION A GOOGLE TAG MANAGER



Objectifs

Initiation aux principes et au fonctionnement de Google Tag Manager à travers une présentation complète des fonctionnalités et de l'administration de l'outil pour la mise en place de plans de marquage Analytics et de scripts digitaux.



Compétences développées

Configuration et exploitation de Google Tag Manager pour le déploiement de tags digitaux (Analytics, Media,...)



Prérequis

- » Connaissances digitales
- » Une appétence aux sujets digitaux techniques est un plus.

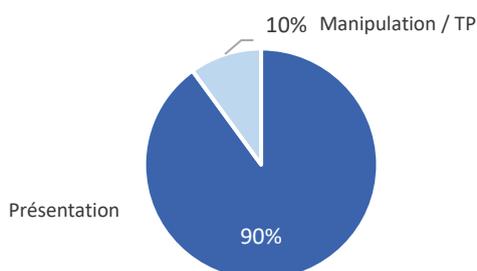


Durée

1 jour (soit 7 heures)



Répartition des activités pédagogiques



Plan de Formation

- » Les concepts clés de Google Tag Manager
 - Les différentes fonctionnalités (variables, déclencheurs, balises)
 - Orchestration des différentes fonctionnalités pour le paramétrage
 - Liaison dataLayer / plan de marquage (la complémentarité)
 - Utilisation Desktop Vs Mobile
 - Approche GTM Web versus Server-side
- » Création et paramétrage d'un compte
 - Gestion des conteneurs
 - Gestion des utilisateurs
 - Importation / Exportation...
 - Environnements
- » Présentation de l'interface Google Tag Manager
 - Tableau de bord / Workspace
 - Utilisation des variables :
 - Les déclencheurs
 - Les balises
 - Gestion des versions
 - Publication des versions
 - Utilisation des dossiers et des zones
- » Google Tag Manager pour Firebase
 - L'approche Firebase + GTM
 - Limitations inhérentes au contexte technique
 - Différences liées à la collecte
 - Méthode de debug (plus complexe)
 - Délai lors des publications



8.3 INITIATION A TEALIUM



Objectifs

Initiation aux principes et au fonctionnement de TealiumIQ à travers une présentation complète des fonctionnalités et de l'administration de l'outil pour la mise en place de plans de marquage Analytics et de scripts digitaux.



Compétences développées

Configuration et utilisation de TealiumIQ



Prérequis

Des bases de compréhension sur les TMS

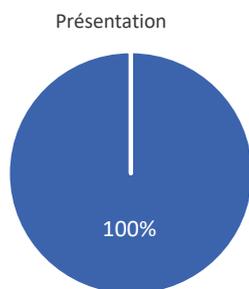


Durée

1 jour (soit 7 heures)



Répartition des activités pédagogiques



+ démonstration de l'outil



Plan de Formation

- » La suite Tealium
 - Présentation globale
 - Les différents outils de la suite
 - Tealium IQ (avec mention de la CMP intégrée)
- » Les concepts clés de TealiumIQ
 - La structure (comptes/profils)
 - Les différentes fonctionnalités (variables, déclencheurs, tags, extensions)
 - Le socle technique
- » L'administration
 - Les comptes, profils, librairies
 - Gestion des utilisateurs, permissions
 - Gestion des environnements (dev/qa/prod)
- » Présentation de l'interface TealiumIQ
 - Dashboard
 - Data Layer (variables)
 - Load Rules
 - Tags
 - Extensions
- » Versions
 - Save/Publish
 - Environnements
 - Présentation de l'arbre des versions
- » **Bonus** : aperçu de la CMP intégrée dans Tealium
 - Fonctionnement général
 - Interaction avec autodata
 - Privacy Manager et gestion des tags