

Présentation & Programme de la filière

»» Data Analyst

Filière

57 jours

| 399 heures

Version 2026 2.1

Présentation de la filière

Distanciel

5 586 € HT / participant

57 jours – 399 heures

La durée et le prix correspondent à une inscription en inter-entreprises. Toute demande intra-entreprise fait systématiquement l'objet d'un devis sur-mesure, devant être approuvé pour acceptation.

Objectifs

- Maîtriser les fondamentaux de la BI et du Big Data.
- Savoir alimenter les entrepôts de données à l'aide d'ETL.
- Requête, analyser et restituer des données avec des outils de Dataviz.
- Mettre en place les concepts SOA à l'aide de BUS et de solutions EAI.
- Acquérir le savoir être du consultant.

Public

Demandeurs d'emploi souhaitant exercer le métier de la Data et du Digital Business.

Prérequis

Bac+ 2 minimum, candidats intéressés par les activités numériques.

Modalités et délais d'accès

Les postulants devront passer une série d'entretiens pour intégrer la formation. Ils seront informés de leur inscription au plus tard 15 jours avant le début de la session.

Accessibilités aux personnes en situation de handicap

Les personnes en situation de handicap sont invitées à nous communiquer leurs besoins spécifiques. Nous ferons tout pour les mettre dans les meilleures conditions de suivi de la formation possibles (compensation, accessibilité...).

Modalités d'évaluation

À l'issue de chaque module, le formateur évaluera chacun des participants en fonction des cas pratiques et exercices effectués. La fin de la formation sera consacrée à un projet final reprenant l'ensemble des acquis de la formation. Les apprenants participeront à une soutenance pour présenter leur projet devant un jury et démontrer leurs nouvelles compétences.

Attestation/certification

Une attestation de fin de stage sera remise à tous les participants à l'issue de leur parcours.

Présentation de la filière

Distanciel

5 586 € HT / participant

57 jours – 399 heures

La durée et le prix correspondent à une inscription en inter-entreprises. Toute demande intra-entreprise fait systématiquement l'objet d'un devis sur-mesure, devant être approuvé pour acceptation.

Méthodes mobilisées

Alternance d'exercices, cas pratiques, QCM et de notions théoriques, projet Fil Rouge avec une répartition du temps de travail : 40% théorie, 60% pratique. Des présentations théoriques des concepts clés illustrés par des démonstrations du formateur (Ex : code live...) seront suivies de mises en pratique des apprenants.

Évaluations régulières et retour du formateur sur les points moins bien assimilés. ; les apprenants réaliseront tout au long de la formation des exercices, QCM, mises en situation, TP, TD qui seront corrigés pour faciliter l'acquisition de compétences.

En classe virtuelle, accès à notre plateforme à distance, à des machines virtuelles en local ou dans le cloud contenant les logiciels utiles et les supports de cours en français seront mis à disposition via notre la plate-forme de téléchargement AJC Classroom.

Accès à notre plateforme à distance de Classe Virtuelle : mêmes possibilités et interactions avec votre formateur que lors d'une formation présentielle: votre formation se déroulera en connexion continue 7h/7 :

- Échanges directs avec le formateur et l'équipe pédagogique à travers la visioconférence, les forums et chats ;
- Vérification de l'avancement de votre travail et évaluation par votre formateur à l'aide d'exercices et de cas pratiques ;
- Suivi pédagogique et conseils personnalisés pendant toute la formation.

Vous recevrez les informations de connexion par mail dès votre inscription. En cas de problème de connexion, vous pourrez joindre notre équipe à tout moment (avant ou même pendant la formation) au 01 82 83 72 41 ou par mail (hotline@ajc-formation.fr).

En présentiel, mise à disposition d'ordinateurs portables (16Go RAM, SSD); nos salles sont équipées de matériels pédagogiques (Tableau blanc, vidéo projecteur, tableau tactile...) et informatiques.

Contenu pédagogique

Comportemental	Rôle et comportement du consultant	2 jours
	Travail en équipe	1 jour
Fondamentaux	Modélisation normalisée	1 jour
	Introduction aux outils décisionnels	1 jour
	Modélisation décisionnelle	2 jours
Projet	Projet Modélisation	1 jour
Base de données	SQL / SQL avancé	4 jours
	PL / SQL	2 jours
Projet	Projet Base de Données	1 jour
ETL	Talend Studio	4 jours
Projet	Projet ETL	2 jours
Dataviz, analyse et design	SAP BO BI 4	3 jours
	Qlik Sense – Designer et Developer	3 jours
	Power BI	3 jours
	MicroStrategy (Web, Desktop, Archi)	3 jours
	Tableau Software	3 jours
Projet	Projet Dataviz	3 jours

Contenu pédagogique

Big Data	Présentation du Big Data	5 jours
	Spark	4 jours
Projet	Projet Big Data	2 jours
Comportemental	Présenter ses nouvelles compétences	1 jour
	Rédiger mon CV et ma lettre de motivation	1 jour
Projet	Projet Final & Soutenance - Data Analyst	5 jours
		57 jours

Projet BI

Projet Fil Rouge

Énoncé :

- Le stagiaire choisira un sujet avec une réalisation d'une modélisation d'un entrepôt de données. Il doit réaliser la conception avec une modélisation MCD, MLD et MPD.
- Une fois la modélisation de l'entrepôt de données est terminée il faudra créer deux ou trois modélisation OLAP en étoile ou flocon pour le datamart.
- Les créations et l'intégration des données se fera via des scripts PL/SQL et/ou ETL Talend.
- Les cubes OLAP seront visualisés avec un ou plusieurs outils de datavisualisation.
- Cette solution pourra éventuellement compléter une solution Big Data en récupérant quelques sources d'un Data Lake existant.

Programme détaillé

» Comportemental

Rôle et comportement du consultant

Distanciel

2 jours – 14 heures

Objectifs

- La communication interne et externe au sein de l'entreprise
- Adapter et maîtriser les différents types de communication pour accroître son efficacité personnelle

Contenu de programme

Pourquoi s'intéresser aux comportements en tant que consultant ?

- Qu'est-ce qu'un comportement ? Qu'est-ce qu'un rôle ?
- En quoi les comportements peuvent faire la différence ?
- Pourquoi choisit-on d'adopter un comportement ? Le processus d'apprentissage d'un «savoir-être»
- *Jeu pédagogique : nouvellement recruté dans une entreprise, vous sollicitez un entretien pour vous faire préciser votre rôle*

Adopter la meilleure stratégie de coopération pour mieux travailler en équipe

- Comment agir pour développer des relations positives et durables ?
- La théorie CRP
- *Jeu pédagogique : 2 équipes doivent atteindre le score le plus élevé possible*

Savoir communiquer et éviter les malentendus

- Pourquoi la communication passe-t-elle mal : les filtres, le cadre de référence ?
- *Travaux pratiques : Test sur les croyances limitantes*
- Savoir utiliser l'écoute active : questionnement ouvert et reformulation
- *Jeu pédagogique : utiliser l'écoute active pour dessiner un objet sans le voir en questionnant et demandant des précisions à votre interlocuteur*
- Savoir convaincre : comment influencer positivement les échanges
- *Travaux pratiques : racontez la journée type d'un consultant qui mettrait en œuvre tous les leviers à sa disposition pour convaincre ses interlocuteurs*

Comment faire évoluer ses comportements

- Qu'est-ce qui conditionne nos comportements ?
- Sur quel levier agir pour ajouter des «cordes à son arc»

Comprendre sa personnalité et mieux cerner celle des autres

- Savoir se situer et comprendre en quoi notre personnalité se traduit à travers nos comportements
- Situer les autres et comprendre leur mode de fonctionnement pour mieux coopérer
- *Test de personnalité*

Rôle et comportement du consultant (suite)

Distanciel

2 jours / 14 heures

Contenu du programme

Développer son intelligence émotionnelle pour modifier ses comportements

- Qu'est-ce que l'intelligence émotionnelle ?
- En quoi notre QE est-il déterminant par rapport à nos comportements

Apprendre à gérer son stress pour éviter les comportements inadaptés

- Le stress : de quoi parle-t-on ?
- Comment prévenir le stress et le gérer ?
- *Test d'auto-évaluation QE*

Appréhender le rôle des croyances et de l'éducation dans nos comportements

- Qu'est-ce qu'une croyance ?
- Pourquoi conditionnent-elles nos comportements ?
- *Test : qu'est-ce qui caractérise votre personnalité et en quoi cela influence-t-il vos comportements*
- *Vidéo : comment naît un comportement collectif*

L'assertivité et l'empathie pour mieux travailler en équipe

- Qu'est-ce que l'assertivité ? Qu'est-ce que l'empathie ?
- La notion de respects des besoins et de gagnant-gagnant
- Savoir recadrer un comportement qui ne nous convient pas et renouer avec des relations positives
- *Jeu de rôles*

Le travail en équipe

Distanciel

1 jour / 7 heures

Objectifs

- Comprendre la dynamique d'une équipe
- Susciter la participation et l'engagement
- Utiliser les techniques et les outils appropriés pour agir en équipe
- S'organiser au sein d'une équipe
- Communiquer efficacement quel que soit son rôle

Contenu du programme

Le travail en équipe

- Définition
- La dynamique de groupe
- La structuration de l'équipe de travail, La taille de l'équipe
- Les facteurs d'influence
- Les comportements, Les styles de leadership
- Les points clés de réussite du travail en équipe.

La dynamique de groupe

- Les facteurs de cohésion et de dissociation
- La vie affective du groupe et son évolution dans le temps
- *Travaux pratiques : Comprendre comment naît une habitude collective après l'expérience de Asch*

La structuration de l'équipe

- Sa mission
- Ses objectifs
- Les ressources et les moyens
- L'information et le suivi d'activité

Les facteurs d'influence

- Les facteurs de démoralisation
- Les facteurs de cohésion

Les comportements

- Individuels et de groupe

Le travail en équipe (suite)

Distanciel

1 jour / 7 heures

Contenu du programme

Les points clés de réussite du travail en équipe

- Savoir écouter et s'exprimer
- Savoir accepter le consensus
- Savoir négocier.
- Respecter les autres.
- Savoir mettre en œuvre une méthode de travail qui vise à atteindre les objectifs fixés

Cas pratique de synthèse :

- *Mettre en œuvre un plan d'actions suite à un brief client délicat en utilisant les techniques de créativité*

Programme détaillé

»» Fondamentaux

Modélisation normalisée

Distanciel

1 jour / 7 heures

Objectifs

- Comprendre les modèles conceptuels et physique de la méthode

Contenu du programme

Introduction

- Historique
- Bases de données

Introduction à Merise

- L'avènement de Merise
- Pourquoi Merise?
- Les origines
- Points forts, Points faibles

Merise /2

- La modélisation des données aujourd'hui
- Vocabulaire
- Organisation et système
- Le système d'information
- Modèle, Méthode, Analyse

Les principes de base de Merise

- Les principes de base
- Introduction
- Les cycles de la démarche Merise
- La technique de la démarche
- *Travaux pratiques : Création d'un scénario avec un cycle.*

Le MCD

- Introduction
- M.C.D
- Entité
- Entité : Propriété
- Entité : Identifiant
- Entité : Occurrence

Modélisation normalisée (suite)

Distanciel

1 jour / 7 heures

Contenu du programme

Le MCD (suite)

- Association
- Compléter le modèle
- Les cardinalités
- Les dimensions
- *Travaux pratiques : Création d'un besoin avec son MCD complet.*

Les compléments du modèle

- Compléter le modèle
- Les dimensions
- Les formes normales

Contraintes inter-associations

- Différentes situations
- Partition
- Exclusion
- Totalité
- Inclusion
- Contraintes sur les propriétés

Spécialisation/Généralisation d'entités-types

- Contraintes sur spécialisation

Passage du MCD au modèle physique

- Introduction
- Le MLD, Le MPD
- Une table exemple
- Formalisme
- Passage du MCD au MLD
- Le reverse-engineering
- Les outils de modélisation et de génération
- Migration de bases
- *Travaux pratiques : Transformer le MCD de l'ancien besoin en MLD et MR en suite en MPD.*

Introduction aux outils décisionnels

Distanciel

1 jour / 7 heures

Objectifs

- Connaître les fonctions, forces et faiblesses des différents outils du marché décisionnel

Contenu du programme

Introduction

- Objectifs et définitions
- Architecture générale
- Architecture technique
- Le stockage des données
- La conception du Datawarehouse
- L'analyse du besoin : Exercice
- Les limites du SI opérationnel
- La modélisation des données
- La base de données
- Optimisation du modèle
- L'administration des données
- *Travaux pratiques : Création d'une base de données à partir d'un MLD.*

L'alimentation du Datawarehouse

- Les ETL
- Architecture de divers ETL
- Présentation d'un ETL
- *Travaux pratiques : Alimentation d'un fichier de données (txt ou csv) dans une base de données avec un ETL.*

Les outils de restitution

- Présentation de différents outils
- La gestion de projet
- Le décisionnel et la législation
- Les perspectives d'évolutions

Conclusion

- Connaître les logiciels et outils du décisionnel
- Avoir une vue d'ensemble de ces outils
- Connaître leurs historiques et leurs places sur le marché du décisionnel
- *Travaux pratiques : Créer une représentation graphique de la base de données alimentée précédemment.*

Modélisation décisionnelle

Distanciel

2 jours / 14 heures

Objectifs

- Créer une architecture de données adaptée aux besoins décisionnels des utilisateurs.
- Aborder pour aborder un projet décisionnel en se basant sur les fondements
- Modéliser par les objets
- Mesurer les enjeux et impacts d'un projet décisionnel
- Concevoir et modéliser un entrepôt de données
- Découvrir le rôle des différents outils de l'informatique décisionnelle
- Connaître les pièges à éviter lors de la mise en œuvre

Contenu du programme

Introduction au contexte

- Historique de l'informatique décisionnelle
- Modélisation dimensionnelle

Les limites du système d'information opérationnel

- Systèmes transactionnels
- Il ne permet pas les analyses
- Information de contrôle vs Information de conduite

L'objectif d'un l'entrepôt des données

- Définition de l'entrepôt de données
- L'objectif de l'entrepôt de données
- Les défis d'un entrepôt de données
- Centraliser les informations en une base unique
- Intégrer des données élémentaires comme des données complexes
- Faciliter la « gestion » des données centralisées
- Respecter les contraintes de temps
- Offrir un outil sur mesure permanent
- *Travaux pratiques : Création d'un besoin ou scénario d'archivage de données dans un entrepôt de données.*

La Modélisation

- Introduction
- *Travaux pratiques : Modélisation de l'entrepôt de données (MCD, MLD) du scénario précédent.*

Modélisation décisionnelle (suite)

Distanciel

2 jours / 14 heures

Contenu du programme

Méthodologie de modélisation

- Les différents types de modélisation

Modélisation en Étoile

- La table de Fait
- Les tables de dimension
- Remarque sur les performances
- Le modèle en étoile

Modélisation en flocon

Modélisation en Constellation

- Le modèle en constellation
- Modélisation de type « SnowStorm »

Modélisation en Matrice

- Modélisation : Axiomes
- *Travaux pratiques : Modélisation OLAP avec un modèle en étoile.*

Comment modéliser ?

- Les dimensions
- Schéma à parents multiples
- Dimensions Multi-Références
- Volume des tables de dimension
- *Travaux pratiques : Modélisation OLAP avec un modèle en flocon.*

Programme détaillé

»» **Projet**

Projet Modélisation

Distanciel

1 jour / 7 heures

Objectifs

- Mettre en application les acquis de la formation sur un projet Modélisation

Contenu de programme

- Les stagiaires travaillent en toute autonomie, en binôme. Ils sont libres d'effectuer les choix adaptés, de développer les parties dont ils jugent avoir le plus besoin et d'apporter leurs propres solutions aux problèmes posés.
- Le formateur encadre les stagiaires par sa présence et répond aux questions. Il intervient pour épauler un binôme en difficulté ou pour faire le point à l'ensemble du groupe sur des notions non acquises. Il peut être amené à approfondir ou compléter certaines connaissances.

Programme détaillé

» Base de données

SQL / SQL avancé

Distanciel

4 jours / 28 heures

Objectifs

- Mettre en place des requêtes SQL avec jointures, regroupements et sous-requêtes
- Utiliser les nouveaux ordres SQL et les nouvelles fonctionnalités des jointures, regroupements et sous-requêtes
- Suivre les plans d'exécution de requêtes

Contenu du programme

Introduction

Les requêtes simples

- Connaître la syntaxe de l'ordre SELECT
- Connaître les prédicats simples
- Savoir écrire des requêtes sur une seule table
- Comprendre ce qu'est une jointure
- Savoir écrire des requêtes sur plusieurs tables
- Connaître les fonctions utilisées dans les requêtes

Les requêtes avec regroupement

- Comprendre le regroupement
- Savoir écrire une requête utilisant le regroupement
- *Travaux pratiques : Utilisation des requêtes avec des fonctions d'agrégation et le Group BY*

Les requêtes ensemblistes

- Connaître les opérateurs ensemblistes et savoir écrire une requête utilisant des opérateurs ensemblistes

Les requêtes imbriquées

- Connaître les opérateurs utilisés pour les sous requêtes
- Savoir écrire des requêtes imbriquées
- *Travaux pratiques : Utilisation des requêtes imbriquées avec des Union, Intersect et Except...*

Le langage de manipulation des données

- Savoir insérer, modifier ou supprimer une ligne dans une table (fonctions l'insert, update et delete)
- *Travaux pratiques : Création, modification et suppression des enregistrements d'une table*

SQL / SQL avancé (suite)

Distanciel

4 jours / 28 heures

Contenu du programme

Le langage de définition de données

- Syntaxe de création des différents types de tables
- Définition de contraintes (clé primaire, clé étrangère ...)
- Index B-tree, bitmap, états des lieux des index
- *Travaux pratiques : Création d'un MPD avec son exécution dans une base de données.*

Clauses avancées

- Opérateurs de comparaison modifiés par ANY, ALL et SOME
- Limiter le nombre d'enregistrement avec TOP et WITH TIES
- Élargir une recherche avec les caractères génériques (%, _ , [] , [^])
- Les fonctions RANK et DENSE RANK
- *Travaux pratiques : Création de requête avec les clauses ANY, ALL et SOME en utilisant des fonctions analytiques.*

Synthétiser les données

- Utilisation de la clause COMPUT
- Les clauses ROLLUP et CUBE
- La table PIVOT

Jointures avancées

- Jointure interne
- Jointures multiples
- Corréler des sous-requêtes
- La commande MERGE
- *Travaux pratiques : Création de requête avec des jointures multiples INNER JOIN.*

Analyse de l'exécution et du plan d'exécution des requêtes

- Exécution des requêtes et optimiseur interne
- Analyser les plans d'exécution
- Process de requête adaptatif
- Mise en cache des plans et recompilation

Utiliser des tables temporaires

- Les différents types de tables temporaires
- Quand utiliser des tables temporaires ?

SQL / SQL avancé (suite)

Distanciel

4 jours / 28 heures

Contenu du programme

Manipulation de données avancée

- Utilisation de la clause OUTPUT avec INSERT, UPDATE et DELETE
- La commande TRUNCATE TABLE
- *Travaux pratiques : Création de requête avec Manipulation de données avancée.*

Intégrité des données

- Contraintes uniques
- Contrainte CHECK
- Contrainte DEFAULT
- *Travaux pratiques : Création d'une table avec des colonnes en CHECK plus DEFAULT.*

Traitement des données

- Point de sauvegarde d'une transaction SQL
- Régler le niveau d'isolation de la transaction
- *Travaux pratiques : Utilisation de commit et rollback*

PL / SQL

Distanciel

2 jours / 14 heures

Objectifs

- Comprendre le traitement transactionnel dans Oracle
- Maîtriser les éléments du langage PL/SQL
- Savoir accéder une base de données Oracle
- Créer des procédures et de fonctions stockées
- Utiliser les différents types des triggers

Contenu du programme

Introduction et rappels

- Traitement transactionnel dans Oracle.
- Objets et dictionnaire d'une base Oracle.
- SQL et PL/SQL. Vue globale de PL/SQL.
- Outils utilisés. Documentation Oracle.

Éléments du langage

- Environnements hôtes. Introduction syntaxique. Déclarations de variables et de constantes PL/SQL.
- Types de données et conversion. Fonctions prédéfinies.
- Instructions de contrôle : IF-ELSE, la boucle LOOP, la boucle FOR.
- Types composites : records, tables indexées.
- Types objets. Définition des sous-programmes.
- Paramétrage des sous-programmes.
- *Travaux pratiques : Création d'un script avec des instructions conditionnelles et des boucles.*

Accès à la base

- La base et la notion de curseur.
- Accès via curseur implicite : update, delete, insert, select. Utiliser les clauses RETURNING et BULK COLLECT.
- Collections et FORALL.
- Boucles FOR à curseur. Curseurs en mise à jour.
- Curseurs explicites. Variables curseur.
- *Travaux pratiques : Création d'un script avec deux curseurs implicites.*

PL / SQL (suite)

Distanciel

2 jours / 14 heures

Contenu du programme

Traitements stockés

- Gestion de sous-programmes stockés, de bibliothèques stockées et de paquetages.
- Gestion des traitements stockés : droits d'accès, vues du dictionnaire et dépendance.
- Utilisation des packages.
- *Travaux pratiques : Création d'un package.*

Gestion des erreurs

- Traitement d'une exception.
- Utilisation de RAISE_APPLICATION_ERROR.
- Déclenchement d'une exception via RAISE.
- Propagation et gestion transactionnelle.
- *Travaux pratiques : Création d'un script complet en ajoutant les traitements des exceptions.*

Triggers

- Définition et types de triggers Événement/Action. Triggers Ordre.
- Lignes (utilisation de : NEW et : OLD), E/C/A (Événement/Condition/Action).
- « INSTEAD OF » sur les vues, sur les événements système.
- Utilisation des transactions autonomes.
- *Travaux pratiques : Création d'un trigger avant et après insertion d'enregistrement.*

Programme détaillé

»» **Projet**

Projet Base de Données

Distanciel

1 jour / 7 heures

Objectifs

- Mettre en application les acquis de la formation sur un projet Base de Données

Contenu de programme

- Les stagiaires travaillent en toute autonomie, en binôme. Ils sont libres d'effectuer les choix adaptés, de développer les parties dont ils jugent avoir le plus besoin et d'apporter leurs propres solutions aux problèmes posés.
- Le formateur encadre les stagiaires par sa présence et répond aux questions. Il intervient pour épauler un binôme en difficulté ou pour faire le point à l'ensemble du groupe sur des notions non acquises. Il peut être amené à approfondir ou compléter certaines connaissances.

Programme détaillé

» ETL

Talend Studio

Distanciel

4 jours / 28 heures

Objectifs

- Mettre en place et paramétrer un environnement de développement stable
- Modéliser ses besoins
- Maîtriser la bibliothèque de composants.
- Implémenter, déboguer et déployer ses jobs.
- Maîtriser les composants avancés de Talend.

Contenu de programme

Talend open studio : présentation générale

- La problématique liée à Talend
- Enterprise Resource Planning (ERP) et Entreprise Application Integration (EAI)
- Être capable d'installer Talend Open Studio
- Débuter avec TOS (Talend Open Studio)
- Avoir une vision globale des composants TOS

Réalisation de ses premiers flux de données sur TOS

- Comprendre le Business Model
- Assimiler les bonnes pratiques
- Connaître les composants de base
- Manipuler des fichiers Excel (Lire, écrire)
- Être capable de manipuler des emails
- Connaître les différents types de liaisons existantes entre composants
- Comprendre le job : unité de traitement principale
- Savoir exécuter un job
- *Travaux pratiques : Création d'un Business Model et d'un job qui lit et affiche un fichier plat.*

Réalisation de job : fichiers et bases

- Principe du flux et schéma de données
- Job designer : de bonnes pratiques
- Générer et sortir des données
- Exécuter un job
- Gérer les accès fichiers
- Gérer les accès aux BD et les autres composants

Talend Studio (suite)

Distanciel

4 jours / 28 heures

Contenu de programme

Appréhender les métadonnées et le traitement des données

- Connaître les différents types de traitement de données (tri, filtre, agrégateur...)
- Introduction aux métadonnées (noeud Metadata)
- Appréhender les Métadonnées pour Access et Excel
- *Travaux pratiques : Création d'un job plus complet avec plusieurs transformations et à partir des métadonnées.*

Interactions avec les bases de données

- Connexion à une base de données avec Talend
- Lecture et écriture dans une base de données
- *Travaux pratiques : Création d'un job avec deux connexions via métadonnée et à deux bases de données différentes. Réaliser quelques transformations et migration de données.*

Introduction aux manipulations avancées de Talend

- Présentation des différents types de conditions et d'itérations
- Les boucles, les itérations, les conditions
- Gestion des logs et des erreurs
- Variables de contexte
- Cascade de jobs
- Les routines

Automatisation et documentation

- Documenter ses jobs
- Générer des scripts
- *Travaux pratiques : Création de plusieurs jobs en externalisant les variables de contexte et avec la gestion des log et erreurs.*

Programme détaillé

»» **Projet**

Projet ETL

Distanciel

2 jours / 14 heures

Objectifs

- Mettre en application les acquis de la formation sur un projet ETL

Contenu de programme

- Les stagiaires travaillent en toute autonomie, en binôme. Ils sont libres d'effectuer les choix adaptés, de développer les parties dont ils jugent avoir le plus besoin et d'apporter leurs propres solutions aux problèmes posés.
- Le formateur encadre les stagiaires par sa présence et répond aux questions. Il intervient pour épauler un binôme en difficulté ou pour faire le point à l'ensemble du groupe sur des notions non acquises. Il peut être amené à approfondir ou compléter certaines connaissances.

Programme détaillé

» Dataviz, analyse et design

SAP BO BI 4

Distanciel

3 jours / 21 heures

Objectifs

- Créer une requête simple à partir des données de l'univers
- Créer des tableaux de données et les mettre en forme
- Filtrer et regrouper les données à restituer
- Créer et formater des diagrammes sur les données
- Utiliser des calculs, alerteurs et formules pour analyser et mettre en valeur des données
- Organiser la diffusion des rapports

Contenu de programme

L'environnement : Infoview

- Web Intelligence : généralités sur les rapports Web Intelligence
- InfoView
- La page d'accueil Infoview
- Généralités sur l'utilisation des objets Web Intelligence
- Définir les préférences de l'utilisateur
- L'interface Infoview

Interroger : la gestion des requêtes sous InfoView

- Vue d'ensemble
- Généralités sur l'utilisation des requêtes
- Construction d'une requête Web Intelligence
- Utilisation des filtres
- Compléments

Reporter : l'utilisation des rapports

- Modes de travail avec un rapport
- Présentation Types de tableaux
- Structure et résultat d'un document Web Intelligence
- Gestion des tableaux dans un rapport
- Ajout d'informations et de calculs à un rapport
- Ajouter des éléments décoratifs ou informationnels
- Utilisation des sections
- Utilisation des ruptures
- Mise en forme de rapports
- Mode plan (Plier/Déplier)
- *Travaux pratiques : Création d'un rapport avec des sections et avec des ruptures.*

SAP BO BI 4 (suite)

Distanciel

3 jours / 21 heures

Contenu de programme

Reporter : les diagrammes

- Présentation
- Créer un diagramme
- Manipuler un diagramme
- Mise en forme des diagrammes

Analyser les données

- Les formules
- Les variables
- Quelques fonctions Web Intelligence
- Utilisation des tris
- Classement de données
- Utilisation des alerteurs
- Filtres de rapport

Explorer les données

- Exploration des diagrammes
- *Travaux pratiques : Application et ajout des fonctions Web Intelligence dans un rapport.*

Utilisation de plusieurs fournisseurs de données

- Présentation
- Limitations
- Ajouter des fournisseurs de données
- Manipuler les fournisseurs de données correspondant à une requête
- Purger des documents
- Fusion de dimensions

Problèmes et solutions : fonctions avancées

- Utilisation des sous-requêtes
- Combinaison de requêtes
- Contexte de calcul
- *Travaux pratiques : Ajout des sous-requêtes et des contextes de calcul dans un rapport.*

SAP BO BI 4 (suite)

Distanciel

3 jours / 21 heures

Contenu de programme

Trucs et astuces

- Fonctions Deski non encore intégrées dans BI4 et solutions de contournement
- Extensions de certaines fonctionnalités

Qlik Sense – Designer et Developer

Distanciel

3 jours / 21 heures

Objectifs

- Découvrir Qlik Sense Desktop, Qlik Sense Cloud, Qlik Sense Entreprise
- Comprendre l'architecture d'une application Qlik Sense
- Naviguer dans une application Qlik Sense
- Créer une application en mode assistant
- Charger et transformer les données
- Créer des visualisations et des mises en récit
- Se connecter à différentes sources de données
- Manipuler des données
- Utiliser tous les graphiques proposés par Qlik Sense
- Collaborer et gérer la sécurité d'accès aux données

Contenu de programme

Présentation de Qlik Sense

- Qlik Sense en bref.
- Qlik Sense versus QlikView.
- Architecture d'une application Qlik Sense.
- Navigation dans l'interface utilisateur : le hub, mode feuille, mode mise en récit.
- Gestionnaire des données. Éditeur de chargement de données. Visionneur de modèle de données.
- A quoi servent Qlik Sense Cloud et Qlik Sense Entreprise
- Éditeur de chargement de données.
- Visionneur de modèle de données.

QLIK SENSE DESIGNER

Gestion des données avec Qlik Sense

- Créer une connexion vers une source de données.
- Préparation des données avant chargement dans l'application.
- Gestion des associations des données (via l'assistant).
- Structuration et Modélisation des données. Bonne pratique de modélisation. Exemple d'une modélisation en étoile.
- Création d'une dimension et d'une mesure.
- *Travaux pratiques : Chargement d'un fichier csv dans une application et création d'une mesure.*

Qlik Sense – Designer et Developer (suite)

Distanciel

3 jours / 21 heures

Contenu de programme

QLIK SENSE DÉVELOPPEUR

Manipulation des données via le script

- Création d'une connexion et ajout des données.
- Syntaxe des scripts : instruction de contrôle, préfixe et variable.
- Comment optimiser la performance d'une application ?
- Association des tables via l'instruction Join et Keep.
- Utilisation de l'instruction MAPPING et INTERVALMATCH.
- Utilisation de fichiers de données Qlik (.QVD)
- Démonstration : exemple d'une modélisation en étoile.
- *Travaux pratiques : Création d'une application avec une association entre deux ou plusieurs fichiers en entrées (modèle en étoile).*

Création de visualisations

- Structuration d'une application.
- Mesures & dimensions.
- Création des visualisations et les différents cas d'utilisation.
- Utilisation des expressions & variables dans les graphiques.
- Utilisation des "set analysis".
- Mise en place d'une carte géographique.
- *Travaux pratiques : Création d'une feuille avec plusieurs graphiques.*

Collaboration et Partage dans Qlik Sense

- Mise en récit des données.
- Publication d'application, feuille et récit.
- Extraction des données.
- Utilisation de Qlik Sense Cloud.
- Gestion de la sécurité dans une application.
- Conclusion : que faut-il retenir ?
- *Travaux pratiques : Création d'un récit d'une analyse de rapport déjà réalisé avec des commentaires.*

Power BI

Distanciel

3 jours / 21 heures

Objectifs

- Se connecter, interroger et associer des sources de données hétérogènes
- Modéliser et préparer ses données
- Créer et mettre en forme des rapports avec Power BI Desktop
- Concevoir des mesures avec le langage DAX
- Découvrir les services Power BI Online

Contenu de programme

Introduction

- Présentation de l'offre BI Microsoft.
- La Suite Power BI : Power BI Desktop, le service Power BI et les applications Power BI Mobile.

Obtenir les données

- Extraire, transformer et charger des données dans Power BI.
- Utiliser des fichiers plats, Excel, bases de données relationnelles, sources SSAS, Web.
- Choisir les colonnes, choisir les lignes, filtrer, trier et supprimer les doublons.
- *Travaux pratiques : Manipulation des fichiers plats via l'outil.*

Transformer les données

- Nettoyer et compléter les jeux de données.
- Fractionner les colonnes, les formater et définir le type de données, assembler et fusionner des tables.
- Créer des colonnes calculées. Empiler et fusionner des requêtes.

Définir le modèle de données

- Utiliser la vue Diagramme pour définir des relations entre les tables.
- Concevoir des mesures avec les fonctions statistiques.
- Fonctions DAX pour concevoir des mesures élaborées : CALCULATE, ALL, ALLEXCEPT, FILTER, RANKX etc.
- Utiliser les fonctions DAX Time Intelligence.
- *Travaux pratiques : Manipulation et transformation d'un jeu complet de données en appliquant les fonctions DAX.*

Power BI (suite)

Distanciel

3 jours / 21 heures

Contenu de programme

Conception de rapport Power BI

- Création de rapports, pages, tableaux de bord.
- Afficher des données : Table, Matrice, Carte.
- Exploiter les paramètres de données et de formatage.
- Insérer des éléments visuels. Exploiter les visualisations cartographiques.
- Importer des éléments visuels (.pbviz) depuis la galerie : jauges, images etc.
- Ajouter des outils de filtrage, des segments, des KPI. Mise en page pour écran ou smartphone.
- *Travaux pratiques : Création d'un tableau de bord avec des éléments graphiques en incluant des filtres.*

Présentation du service Power BI Online

- Présentation des fonctionnalités du service Power BI, stockage, présentation, partage.
- Épingler les visualisations dans un tableau de bord.
- Présentation de conception de rapport avec le service Power BI.
- Quels jeux de données dans Power BI ?
- *Travaux pratiques : Réaliser un tableau de bord complet.*

MicroStrategy (Web, Desktop, Archi)

Distanciel

3 jours / 21 heures

Objectifs

- Comprendre le fonctionnement de MicroStrategy Web
- Comprendre les concepts de la Data Visualisation
- Savoir créer des analyses visuelles sous MicroStrategy Visual Insight
- Comprendre le fonctionnement de MicroStrategy Architect

Contenu de programme

Création de rapports Web

- Introduction à MicroStrategy BI
- L'architecture de MicroStrategy BI
- Introduction aux principes de Reporting

Utilisation de l'interface MicroStrategy Web

- Présentation de l'interface MicroStrategy Web
- Rapports et documents

Les rapports sous MicroStrategy Web

- Qu'est-ce qu'un rapport ?
- Création de rapports
- Sauvegarder des modèles et des filtres
- Options de sauvegarde avancées
- *Travaux pratiques : Création d'un premier rapport et manipulation de l'outil.*

Manipulation de rapports sous MicroStrategy Web

- Manipulation de rapports
- Souscription, impression et exportation de rapports et de documents

Le formatage de rapport sous MicroStrategy Web

- Formatage de grilles
- Formatage de graphiques

Les filtres sous MicroStrategy Web

- Création de filtres de rapport
- Réglage des opérateurs
- *Travaux pratiques : Création d'un rapport web complet avec exportation de rapport.*

MicroStrategy (Web, Desktop, Archi) (suite)

Distanciel

3 jours / 21 heures

Contenu de programme

Les invités sous MicroStrategy Web

- L'éditeur d'invite
- Création d'invite
- Ajout d'invite à un rapport
- Options de sauvegarde avancée

Les groupes personnalisés sous MicroStrategy Web

- Création de groupes personnalisés
- Ajouter des éléments de groupes personnalisés
- *Travaux pratiques : Création d'invite de groupes personnalisés avec attribution aux rapports.*

Les mesures sous MicroStrategy Web

- L'éditeur de fonctions
- L'éditeur de formules

Introduction à Rapport Services sous MicroStrategy Web

- Création de documents
- Création de tableaux de bords

Les services OLAP sous MicroStrategy Web

- Qu'est-ce que OLAP Services ?
- Création de rapport par cubes intelligents
- Importation de données
- Éléments dérivés
- *Travaux pratiques : Création d'un tableau de bords en s'appuyant sur une analyse OLAP.*

Visual Insight sous MicroStrategy Web

- Qu'est-ce que Visual Insight ?
- Introduction à la Data Discovery
- La création des jeux de données
- Les différentes visualisations
- La création des analyses

MicroStrategy (Web, Desktop, Archi) (suite)

Distanciel

3 jours / 21 heures

Contenu de programme

Services de distribution sous MicroStrategy Web

- Qu'est-ce que Distribution Services ?
- Souscription aux services de distribution

MicroStrategy Desktop

Le module Report Services

- Création de documents
- Création de tableaux de bords

Introduction à la Data Visualisation

- Qu'est-ce que la Data Visualisation ?
- Qu'est-ce que le Data Wrangling ?
- Qu'est-ce que le Data Blending ?
- Présentation de MicroStrategy Visual Insight ?
- Présentation de MicroStrategy Analytics Desktop ?

Création de jeux de données

- Import des données
- Data Wrangling
- Data Blending

Les différentes visualisations

- Grille
- Graphique
- Carte de chaleur
- Cartes
- Réseau
- Open sources (D3)

La création des analyses

- Ajouter des visualisations
- Les fonctionnalités
- Les fonctionnalités avancées

MicroStrategy (Web, Desktop, Archi) (suite)

Distanciel

3 jours / 21 heures

Contenu de programme

Le partage des analyses

- Partager une analyse
- Exporter une analyse
- Conversion en document

Services de distribution

- Qu'est-ce que Distribution Services ?
- Souscription aux services de distribution
- *Travaux pratiques : Création d'un tableau de bords complet avec le Desktop.*

MicroStrategy Architect

Introduction

- L'architecture de MicroStrategy BI
- Les 3 composants de l'architecture de MicroStrategy BI :
- Les sources de données
- Les métadonnées
- Les applications MicroStrategy
- Introduction aux principes de Reporting

Rôle de MicroStrategy Architect

- Utilisation de MicroStrategy Architect
- Initialiser les métadonnées
- Créer les objets de schéma

Le cycle de création de projet

- Configurer les métadonnées
- Configurer la connectivité du projet
- Créer les objets de schéma
- Mettre à jour le schéma du projet

MicroStrategy (Web, Desktop, Archi) (suite)

Distanciel

3 jours / 21 heures

Contenu de programme

Les modèles de données de MicroStrategy Architect

- Le modèle logique
- Attributs
- Faits
- Hiérarchies
- Le modèle physique : Tables, Colonnes

La création de projets sur MicroStrategy Architect

- L'accès aux données
- Les interfaces de création de projet

La couche sémantique du projet sur MicroStrategy Architect

- Définitions
- Création de la couche sémantique
- Les faits
- Les attributs des hiérarchies
- Les mesures

Administration de projet sur MicroStrategy Architect

- La documentation de projet
- La sécurité
- La gestion des utilisateurs
- La maintenance des utilisateurs
- *Travaux pratiques : Création d'un tableau de bords complet avec MicroStrategy Architect avec gestion des utilisateurs.*

Tableau Software

Distanciel

3 jours / 21 heures

Objectifs

- Concevoir, créer et partager des visualisations (tableau de bord interactif, reporting...)
- Acquérir une méthodologie de mise en œuvre visualisations

Contenu de programme

Introduction à Tableau Desktop

- Présentation
- Tableau software, l'éditeur
- Tableau software, du concept à la mise en œuvre
- *Travaux pratiques : Manipulation de l'outil.*

Se connecter aux données

- Les sources de données
- Les types de connexions

Créer ses premières visualisations

- Interface
- Données discrètes ou continues
- Graphiques et tables
- *Travaux pratiques : Création d'une connexion à une base de données et réaliser un graphique visuel.*

Enrichir les données

- Hiérarchies, groupes
- Champs calculés

Enrichir les visualisations

- L'étagère repères
- Tester des graphiques
- Améliorer les tables
- Cartes
- Maîtriser filtres, paramètres, tris, ensembles. Assembler le tableau de bord et l'histoire
- *Travaux pratiques : Création d'un rapport complet avec plusieurs types de graphiques en incluant les filtres et les calculs.*

Programme détaillé

»» **Projet**

Projet Dataviz

Distanciel

3 jours / 21 heures

Objectifs

- Mettre en application les acquis de la formation sur un projet Dataviz

Contenu de programme

- Les stagiaires travaillent en toute autonomie, en binôme. Ils sont libres d'effectuer les choix adaptés, de développer les parties dont ils jugent avoir le plus besoin et d'apporter leurs propres solutions aux problèmes posés.
- Le formateur encadre les stagiaires par sa présence et répond aux questions. Il intervient pour épauler un binôme en difficulté ou pour faire le point à l'ensemble du groupe sur des notions non acquises. Il peut être amené à approfondir ou compléter certaines connaissances.

Programme détaillé

»» **Big Data**

Présentation du Big Data

Distanciel

5 jours / 35 heures

Objectifs

- Appréhender l'ensemble des enjeux et facteurs à prendre en compte pour réussir l'intégration du Big Data dans la vision large du SI

Contenu de programme

Introduction

- Introduction au Big Data : de quoi s'agit-il ?
- Exemples pratiques
- Les technologies concernées
- Les outils
- Les langages : Hortonworks, MapR, Cloudera, IBM Watson

Démystification du Big Data

- Les acteurs principaux
- Les différents métiers du Big Data (Administrateur, Architecte, Développeur, Data Analyst, Data Scientist, DPO, Data Owner, CDO...)

Enjeux et évolutions du Big Data

- La qualité des données
- Le traitement des données
- La protection des données
- L'image de la donnée
- L'humanité des données

Enjeux stratégiques et organisationnels

- Le challenge technique
- Investissement dans la capacité de stockage
- Investissement dans l'analyse
- Le Web sémantique
- Les nouvelles techniques de Data mining
- L'enjeu économique
- L'impact organisationnel
- La conduite du changement
- L'apparition de nouveaux métiers
- *Travaux pratiques : Définir un scénario avec un sujet Big Data à réaliser par un prestataire.*

Présentation du Big Data (suite)

Distanciel

5 jours / 35 heures

Contenu de programme

Les technologies utilisées dans des projets Big Data

- Les outils de stockage
- L'écosystème Hadoop (Apache, Hortonworks, Cloudera, MapR, IBM, Oracle)

Focus sur les outils

- Le paradigme MapReduce.
- Le système de gestion des fichiers distribués HDFS
- Échange de données via Sqoop, Flume, NIFI, Kafka
- Analyse des données avec Pig et Hive
- Les bases NoSQL
- Ordonnanceur Hadoop : Oozie
- Moteur de recherche : Solr, ElasticSearch
- Spark : framework de calcul distribué in memory
- Les solutions de visualisation (MicroStrategy, Tableau, QlikView...)

Sécurité, éthique et enjeux juridiques

- Assurer la protection des données
- L'anonymisation d'une donnée
- Le contrôle d'intégrité
- Le chiffrement d'une donnée
- Qu'est-ce que la blockchain
- Cas d'usage : Bitcoin
- Introduction au Machine Learning

Travaux pratiques :

- *Réaliser une architecture Big Data avec l'écosystème Hadoop du scénario réalisé précédemment*

Spark

Distanciel

4 jours / 28 heures

Objectifs

- Appréhender le fonctionnement et l'architecture de Spark
- Développer des applications avec Apache Spark
- Optimiser une application Spark
- Utiliser Spark SQL, les dataframes et les datasets
- Faire de l'analyse en temps réel avec Spark streaming

Contenu de programme

Introduction et motivations pour Apache Spark

- L'écosystème Spark
- Spark vs. Hadoop
- Obtenir et installer Spark
- La console Spark, et SparkContext

Mise en pratique

- Mettre en place l'environnement de lab
- Démarrer l'interpréteur Scala
- Premiers pas avec Apache Spark
- Premiers pas avec la console Spark

Introduction des RDDs

- Les concepts de RDD, de cycle de vie, et de l'évaluation paresseuse.
- Travailler avec des RDDs : création et transformations (map, filter, etc.)
- Partitionnement et transformation des RDDs
- Transformations avancées (flatMap, explode, et split)

Introduction des DataFrames et DataSets

- Le concept de SparkSession
- Création et inférence de schéma
- Identification des formats supportés (dont JSON, CSV, Parquet, Text...)
- Travailler avec l'API DataFrame
- Travailler avec l'API DataSet
- Transformations via des requêtes SQL (Spark SQL)

Spark (suite)

Distanciel

4 jours / 28 heures

Contenu de programme

Comparaison entre les DataSets, DataFrames et RDDs

- Mise en pratique
- Les bases des RDD
- Opérations sur de multiples RDDs
- Les formats de données
- Les bases de Spark SQL
- Transformation de DataFrames
- L'API typée des DataSets
- Fractionner les données

Optimisations

- Shuffling, dépendances larges et étroites, et leur impact sur la performance
- L'optimiseur de requêtes Catalyst
- L'optimiseur Spark Tungsten (format binaire, gestion du cache...)
- Le caching Spark (concept, type de cache, recommandations)
- Minimiser le shuffling pour améliorer la performance
- Utilisation de la diffusion de variables et de l'accumulateur

Recommandations globales de performances

- L'interface Spark UI
- Les transformations efficaces
- Stockage de données
- Monitoring

Mise en pratique

- Comprendre le Shuffling
- Explorer l'optimiseur de requête Catalyst
- Explorer l'optimiseur Tungsten
- Travailler avec la mise en cache, le shuffling et la diffusion de variables
- Recommandations générales sur le broadcast

Spark (suite)

Distanciel

4 jours / 28 heures

Contenu de programme

Les applications Spark

- Configurer et créer une SparkSession
- Construire et lancer des applications
- Cycle de vie des applications (Driver, Executors, et Tasks)
- Les modes d'exécutions (Standalone, YARN, Mesos)
- Logging et Debugging

Introduction au temps réel

- **Spark Streaming (Spark 1.0+)**
- DStreams, Receivers, Batching
- Transformations Stateless
- Transformations Windowed
- Transformations Stateful
- **Structured Streaming (Spark 2+)**
- Applications en continu
- Le paradigme de Table, et de Result Tables
- Les étapes du structured streaming
- Les sources et puits
- Introduction de Kafka
- Consommer des données Kafka
- Le Structured Streaming au format «kafka»
- Traitement du Stream

Mise en pratique

- Déclencher des jobs Spark

Capacités additionnelles de Spark

- Spark Streaming
- Spark Structured Streaming
- Spark Structured Streaming avec Kafka

Programme détaillé

»» **Projet**

Projet Big Data

Distanciel

2 jours / 14 heures

Objectifs

- Mettre en application les acquis de la formation sur un projet Big Data

Contenu de programme

- Les stagiaires travaillent en toute autonomie, en binôme. Ils sont libres d'effectuer les choix adaptés, de développer les parties dont ils jugent avoir le plus besoin et d'apporter leurs propres solutions aux problèmes posés.
- Le formateur encadre les stagiaires par sa présence et répond aux questions. Il intervient pour épauler un binôme en difficulté ou pour faire le point à l'ensemble du groupe sur des notions non acquises. Il peut être amené à approfondir ou compléter certaines connaissances.

Programme détaillé

» Comportemental

Présenter ses nouvelles compétences

Distanciel

1 jour / 7 heures

Objectifs

- Se présenter en entretien tout en mettant en valeur ses nouvelles compétences en les considérant acquises

Contenu du programme

Les bases de la communication

- Écoute active
- Le questionnement
- Reformulation et feedback
- Interviews croisées

La communication verbale et non verbale

- Importance de la communication non verbale
- Savoir se présenter à l'oral
- Postures – Attitudes – Discours
- *Travaux pratiques : jeu des devinettes et jeu de mime*

Les profils comportementaux

- Les 4 profils
- Auto-évaluation
- Développer son adaptabilité relationnelle
- *Travaux pratiques : Construire sa matrice de compétences*

Développer son Capital Talents

- Définition d'un talent
- Talent vs points forts
- 5 stratégies pour gérer ses points faibles

Exercice de synthèse :

- *Présenter ses 3 compétences fortes*

Rédiger mon CV et ma lettre de motivation

Distanciel

1 jour / 7 heures

Objectifs

- Rédiger et optimiser son CV
- Rédiger une lettre de motivation en cohérence avec le CV
- Savoir se comporter lors d'un entretien d'embauche
- Savoir utiliser les réseaux sociaux pour gagner en visibilité

Contenu du programme

- Comment rédiger son CV
- Comment rédiger sa lettre de motivation
- Se préparer à un entretien d'embauche
- Le Social Networking (LinkedIn)

Programme détaillé

»» **Projet**

Projet Final & Soutenance Data Analyst

Distanciel

5 jours / 35 heures

Objectifs

- Mettre en application les acquis de la formation en complétant les mini projets réalisés dans tout le cursus

Contenu de programme

- Les stagiaires travaillent en toute autonomie, en binôme. Ils sont libres d'effectuer les choix adaptés, de développer les parties dont ils jugent avoir le plus besoin et d'apporter leurs propres solutions aux problèmes posés.
- Le formateur encadre les stagiaires par sa présence et répond aux questions. Il intervient pour épauler un binôme en difficulté ou pour faire le point à l'ensemble du groupe sur des notions non acquises. Il peut être amené à approfondir ou compléter certaines connaissances.

À votre disposition
pour discuter du futur

»» Contact

33 avenue Sainte-Foy
92200 Neuilly-sur-Seine

Tél. : 01 87 58 00 00

eclairezmoi@ajc-formation.fr

www.ajc-formation.fr

www.unjourunjob.fr