

# DATA ANALYST

Formation visant la Certification Fitec : Big Data et processus décisionnel

**Nombre de places disponibles par session : 20**  
**Rémunération moyenne en sortie de formation \* :**  
32 à 45 K€/an\* selon profil et expérience  
**Durée et lieu de la formation :** 57 jours à Nanterre (92) – ou à distance (matériel de formation fourni par Devenez)

\* moyenne constatée des recruteurs participant avec une variation possible selon expérience et compétences



**Prochaines dates de formation**  
*Consulter notre site web*

<https://www.devenez.fr/espace-candidats/les-offres-en-cours/data-analyst/>

## PROFIL RECHERCHE

- Bac+5 : diplômé de l'enseignement supérieur en informatique ou filières scientifiques, Ingénieurs (Universitaire ou autre),
- Vous avez une première expérience dans le domaine de l'analyse de données, la modélisation, et/ou le développement (HTML, JavaScript, Python, C++, Matlab..),
- Vous avez occupé des postes dans le domaine du Décisionnel, Datamining, Statistiques, ou de DBA

## DESCRIPTIF METIER ET OBJECTIFS OPERATIONNELS

Vos missions consisteront à :

- Formaliser les cas d'usages,
- Concevoir et mettre en œuvre des plateformes basées sur des technologies Big Data,
- Installer et déployer des clusters logiciels,
- Concevoir et mettre en œuvre des flux d'intégration de données,
- Optimiser la performance et la qualité logicielle,
- Mettre en œuvre la stratégie d'exploitation des plateformes Big Data (gestion des sauvegardes, – procédures de récupération des données, mises à jour et montées de version logicielles,...),
- Préconiser des outils et/ou technologies adaptés.

A l'issue de la formation, le stagiaire sera capable de :

- Prendre la responsabilité de l'analyse des données issues de l'activité de l'entreprise pour les traiter et les interpréter afin de soumettre des recommandation pertinentes et stratégiques à la direction

## CERTIFICATION

- **CERTIFICATION FITEC** : Big Data et processus décisionnel

# DATA ANALYST

Formation visant la Certification Fitec : Big Data et processus décisionnel

## PROGRAMME DE FORMATION

MODULES	DUREE EN JOURS	DUREE EN HEURES
<b>Introduction</b>	<b>1</b>	<b>1</b>
Présenter l'écosystème Big Data du point de vue technique, économique Détailer les différentes missions d'un Data Analyst en illustrant par des cas concrets métier Cloud & Big Data. Positionner le Cloud et ces enjeux pour un Data Analyst		
<b>Langages Python</b>	<b>5</b>	<b>5</b>
Introduire le langage Python par comparaison aux autres langages usuels Connaître les possibilités et limites de Python Maîtriser la syntaxe essentielle de Python. Être capable de créer et maintenir des applications complètes en Python Découvrir les fonctions avancées du langage Python et les domaines d'applications variés Programmer pour le web avec Python.		
<b>NO SQL Cassandra</b>	<b>4</b>	<b>4</b>
Comprendre l'architecture et le fonctionnement du SGBD NoSQL Cassandra Connaître les apports de Cassandra par rapport à d'autres solutions existantes Savoir installer et configurer le SGDB Cassandra Créer une base de données et manipuler ses objets Découvrir les principes de l'optimisation et du développement Se familiariser avec les outils d'administration et monitoring Cassandra Manipuler le CQL (Cassandra Query Langage) Apprendre à administrer et sécuriser un cluster Cassandra		
<b>Management de projet avec la méthode Agile</b>	<b>1</b>	<b>1</b>
Comprendre les fondements de l'agilité Identifier les étapes d'une transformation agile Différencier les rôles, les fonctions et les livrables au sein d'une gestion d'un projet en agilité Savoir échanger sur un projet sur les apports de l'agilité et sa mise en place		
<b>Ecosystème Spark</b>	<b>4</b>	<b>4</b>
Présenter Spark et les modules de son écosystème : Shark (Spark SQL) et spark streaming. Décrire les fondamentaux de Spark Exploiter les APIs de base de Spark pour manipuler des données Concevoir et implémenter des cas d'usage typiques de Spark Construire des pipelines de données et requêter de larges jeux de données grâce à Spark SQL et aux DataFrames Analyser les jobs Sparks à l'aide des interfaces d'administration et des logs au sein des environnements Databricks Créer des jobs de type Structured Streaming Découvrir les bases du fonctionnement interne de Spark Découvrir le pattern Deltalake		
<b>RGPD dans les projet BIGDATA</b>	<b>1</b>	<b>1</b>
Rappeler les principes généraux de la RGPD Identifier les spécificités du Big Data selon le RGPD Apprendre à analyser les données pour en estimer la conformité avec la RGPD Concevoir des stratégies dans la chaîne de valeur de l'analyse du Big Data		

# DATA ANALYST

Formation visant la Certification Fitec : Big Data et processus décisionnel

<b>Ecosystème Hadoop</b>	<b>4</b>	<b>4</b>
Découvrir les différents éléments de l'écosystème Hadoop et leurs rôles respectifs : HDFS pour le stockage, YARN pour les calculs, et les fonction de stockage, outils d'extraction, de conversion, ETL, analyse, Comprendre l'architecture des applicatifs Hadoop et savoir quels sont les apports et les cas d'usage des solutions Hadoop Présenter le rôle et interactions des différents composants d'une architecture Hadoop Rechercher des données complexes S'initier aux calculs distribués sur des clusters Hadoop		
<b>Data ingestion</b>	<b>3</b>	<b>3</b>
Apprendre à utiliser Talend Data Ingestion pour traiter les données en masse issues des Raw DATA Savoir Collecter des données de flux en continue. Configurer, organiser et simplifier les flux de données Connecter les flux de données avec d'autres systèmes Utiliser Talend Open Studio for Data Integration Créer, exécuter, dupliquer et exporter un Job Extraire des données d'une base de données et de formats de fichiers communs Transformer et intégrer les données dans des cibles.		
<b>Machine Learning LIB Spark</b>	<b>4</b>	<b>4</b>
Apprendre à utiliser le Machine Learning Library avec spark dans un contexte smart data. Être capable de mettre en œuvre Spark ML pour de l'apprentissage machine et l'analyse de données Comprendre comment récupérer les données et les préparer pour un traitement efficace Utiliser DataFrames pour organiser la structure de ses données Savoir modéliser les données et optimiser les modèles Connaître les types d'algorithmes d'apprentissage automatique les plus couramment utilisés : clustering, classification, régression et recommandations Être en mesure d'industrialiser et de déployer une application de prédiction		
<b>Data Visualization</b>	<b>3</b>	<b>3</b>
Comprendre le rôle et les principe de base de la Data visualisation : BDD, API, architecture) Se familiariser au Processus de la data visualisation : méthodes et bonnes pratiques de la réalisation à la diffusion Être capable de concevoir un modèle de documents répondant aux attentes Connaître la grammaire graphique et savoir sélectionner le bon graphique pour représenter la bonne donnée Savoir bâtir un schéma qui met en valeur la pertinence des analyses effectuée Utilisation de Tableau Software et Kibana pour restituer en temps réel le rendu des données collectées		
<b>INDEXATION DES DONNEES Elastic / Elassandra / Kibana</b>	<b>3</b>	<b>3</b>
Apprendre à maîtriser l'écosystème Elastic - Mise en place de connecteur Beat, la solution. Comprendre le fonctionnement d'ElasticSearch Savoir installer et configurer ElasticSearch Être capable d'indexer des volumes importants de données Comprendre comment administrer le système et le surveiller afin de garantir sa disponibilité Savoir utiliser Elasticsearch, logstash et Kibana pour indexer, chercher et visualiser des données et des documents		
<b>STREAMING AVEC KAFKA</b>	<b>4</b>	<b>4</b>
Exploiter la puissance de Kafka et kafka connect avec Spark Streaming pour mettre en place une architecture LAMBDA / KAPPA. Être capable de comprendre le fonctionnement de Kafka Publier et récupérer des données vers un cluster Kafka Savoir configurer Kafka pour intégrer les données de différents formats et de sources différentes Créer des Consumers et Producers pour envoyer des données vers Kafka et lire des données depuis Kafka Ecrire des applications de streaming avec Kafka Streams		

# DATA ANALYST

Formation visant la Certification Fitec : Big Data et processus décisionnel

<b>SCRAPING DE DONNEES AVEC PYTHON</b>	<b>3</b>	<b>3</b>
Définir les méthodes et techniques de scraping Collecter des données de différentes sources web ou fichiers et les traiter. Réaliser du scraping de données avec Python Faire les actions d'ingestion nécessaires pour alimenter un Data Lake. Connaitre La base du traitement par lot Savoir Analyser le JSON, le CSV et l'XML Parcourir des données sur le Web et stoker les résultats Savoir gérer la performance du système : Multiprocessing, ram .... Manipuler différentes API Utiliser un outil de scraping automatique		
<b>MICROSOFT CLOUD AZURE</b>	<b>5</b>	<b>5</b>
Connaitre les concepts généraux du Cloud et les différents services proposés par Microsoft dans Azure Présenter les services fondamentaux de Cloud Azure: services, produits et solutions Comprendre l'Architecture AZURE et les opérations système telles que l'infrastructure globale, les services principaux et la sécurité des comptes Présenter et manipuler des briques Azure : Azure datalake Manipuler les outils de sécurité, de surveillance et de confidentialité dans Azure Savoir créer et dimensionner des machines virtuelles Être capable d'implémenter des solutions de stockage Mettre en oeuvre des applications Web et des conteneurs Sauvegarder et partager des données S'initier au Devops Azure depuis Azure Devops Services Mettre en place des Pipelines et administrer les projets		
<b>ECOSYSTEME HADOOP NIFI</b>	<b>2</b>	<b>2</b>
Décrire l'architecture de NiFi : configurations autonomes et en clusters Savoir installer et configurer NiFi Naviguer, configurer les flux de données et utiliser les informations avec l'interface utilisateur NiFi Organiser et simplifier les flux de données Connecter les flux de données avec d'autres systèmes, tels que Kafka Gérer et automatiser des flux de données entre plusieurs systèmes informatiques		
<b>Efficacité personnelle : Posture du consultant</b>	<b>1</b>	<b>1</b>
Etre à l'aise dans sa relation client Connaitre les principes fondamentaux d'une communication verbale ou non verbale efficace Adapter sa communication aux personnes et aux situations Savoir véhiculer l'image et les valeurs de son Entreprise Savoir se présenter « à 1 client ou à une équipe projet » et définir sa mission de manière concise Actualiser son CV de manière professionnel en intégrant les compétences de la formation Utiliser une communication collaborative tournée vers le client Savoir valoriser sa prestation et formuler des préconisations Gérer les tâches quotidiennes et les priorités Planifier ses tâches et activités en fonction des priorités de sa fonction		
<b>Efficacité personnelle : Préparer et animer une présentation</b>	<b>1</b>	<b>1</b>
Rappeler les fondamentaux d'une présentation efficace : gestion de l'écrit, gestion des visuels et valeur ajoutée de l'oral Gérer son « trac » et réguler ces émotions Impliquer et gérer son auditoire Aborder les principaux facteurs clés de succès permettant de réussir une bonne présentation Identifier le contexte et l'objectif à atteindre Organiser le contenu de la présentation avec un esprit synthétique Savoir démarrer son intervention, conserver le fil de son déroulement, effectuer des transitions pertinentes Se préparer à réagir positivement aux aléas, aux questions, aux contradictions Conserver la maîtrise du timing et savoir conclure positivement.		

# DATA ANALYST

Formation visant la Certification Fitec : Big Data et processus décisionnel

<b>Mise en situation professionnelle</b>	<b>8</b>	
<p>Etre capable d'appliquer les méthodes, utiliser les outils et gérer un projet à partir de l'ensemble des connaissances et compétences acquises dans la formation et au travers une mise en situation générale basée sur un cas concret du milieu professionnel</p> <p>Vous aurez à mettre en place d'une architecture Big Data avec une exploration continue de données Intelligentes collectées à chaud et effectuer les prédictions pour les prises de décisions.</p> <p>A partir d'une expression de besoin ou d'un cahier des charges, vous devrez mettre tout en œuvre pour comprendre le contexte, identifier les attentes et la demande, auditer l'auteur du cahier des charges pour obtenir des précisions, mettre en place les actions fonctionnelles et techniques pour composer la réponse à apporter à votre client ou directeur de projet</p> <p>Mettre en place l'équipe projet, organiser les tâches et formaliser les résultats attendus</p> <p>Préparer les éléments de synthèse de son travail et prévoir un reporting à présenter oralement en réunion avec son client ou directeur de projet</p>		
<b>Nombre total de jours</b>	<b>57J</b>	<b>399H</b>

Note : Ce document n'est pas contractuel et peut faire l'objet de modifications afin de répondre à des impératifs d'ordre pédagogique

# DATA ANALYST

Formation visant la Certification Fitec : Big Data et processus décisionnel

## TARIFS ET FINANCEMENTS (possibilité de prise en charge partielle ou intégrale)

Tarif public des formations de 57 jours : 7500 € HT / stagiaire (tarif indicatif)

**Votre formation peut être prise en charge par un organisme de financement. Dans le cadre d'un FNE, POE, AIF, CSP, n'hésitez pas à nous contacter à [recrutement@devenez.fr](mailto:recrutement@devenez.fr)**

En fonction de votre statut au moment de l'entrée en formation (salarié ou demandeur d'emploi), cette prise en charge pourra être **partielle ou intégrale**.

Plus d'informations sur les financements : <https://www.devenez.fr/financements/>

## MODALITES D'INSCRIPTION ET DELAIS D'ACCES

Délai d'inscription variable selon le mode de financement.

### INSCRIPTION SUR NOTRE SITE INTERNET :

<https://www.devenez.fr/espace-candidats/les-offres-en-cours/data-analyst/>

## APPROCHE PEDAGOGIQUE ET MODALITES D'EVALUATION

- Formation présentielle et/ou en distanciel alternant présentation théoriques et TP – sur le site de FITEC NANTERRE

### FORMATEUR ET PÉDAGOGIE

Chaque formateur à FITEC doit avoir une double casquette celle d'expert et de pédagogue. C'est cette double compétence qui garantit une acquisition de compétences optimum pour chaque stagiaire.

Tous les formateurs ont suivi une formation et maîtrisent l'outil Teams. Ils ont également été formés quant à l'organisation de la formation à distance et cela dès sa mise en place.

### QUALITÉ DES CONDITIONS TECHNIQUES

Un formateur anime le cours sur un **groupe Teams**, partage sa **vidéo** et son **écran**, diffuse ses **supports**, met une **base documentaire** à disposition, **interagit** avec les stagiaires, lance les TP sur des outils et plateformes cloud, peut **prendre la main** sur les postes.

### RYTHME ET SÉQUENÇAGE D'APPRENTISSAGE

Les formations à distance sont donc des cours **synchrones** qui requièrent la présence de tous comme une salle de formation en virtuel.

Le **rythme** est de 9h à 13h et 14h à 17h par jour.

# DATA ANALYST

Formation visant la Certification Fitec : Big Data et processus décisionnel

## INTERACTIVITÉ

A tout moment, les stagiaires peuvent **interagir** avec le formateur, poser des **questions orales** grâce au micro ou par écrit sur le **chat** de Teams.

Les micros des stagiaires sont coupés par défaut et chacun peut l'activer pour intervenir.

## ACCOMPAGNEMENT

Pour chaque session de formation, en plus du formateur, un **modérateur** membre de l'équipe pédagogique veille à la bonne marche de la formation à distance. Il reste connecté à la formation, vérifie les connexions, répond aux questions ou difficultés des stagiaires.

Le formateur peut ainsi se concentrer sur l'animation pédagogique de son cours.

Un **questionnaire en ligne de satisfaction** est rempli par les stagiaires.

## MODALITES D'EVALUATION - VALIDATION DE LA MISE EN OEUVRE DE LA COMPETENCE

20% du temps est consacré au cours théorique et 80% du temps est consacré aux travaux pratiques, exercices et projets.

Les cursus complets de Reskilling sont validés par une mise en situation professionnelle finale et une soutenance en ligne ou en présentiel face à un jury de professionnels du métier.

Chaque partie, réalisation technico-fonctionnelle et restitution orale, sont associées à des formulaires de notation d'un ensemble de critères sur le Savoir, Savoir-Faire et Savoir-être

## ACCESSIBILITE :

Pour tous besoins spécifiques liés à une situation de handicap ou besoins en compensation, nous vous invitons à contacter notre référente handicap pour étudier la faisabilité de votre projet de formation :

Mme BOURG Sandrine – [referenthandicap@groupefitec.fr](mailto:referenthandicap@groupefitec.fr) .

Nos salles de formation sont accessibles aux personnes à mobilité réduite.

## POUR TOUTE DEMANDE DE RENSEIGNEMENT :

Contactez-nous à [recrutement@devenez.fr](mailto:recrutement@devenez.fr)

Version mise à jour le 18/11/2022