

Data Essentials : maîtriser les fondamentaux de la Data et de l'IA

Date et durée
Code formation : DATA010FR Durée : 9 jours Nombre d'heures : 63 heures
Description
<p>La donnée est le nouveau carburant des entreprises, mais savoir l'exploiter reste rare. Cette formation Data Essentials vous permettra de mieux comprendre le monde des données et l'intelligence artificielle. En 9 jours, vous passerez d'un fichier Excel à des prédictions exploitables, sans jamais devoir écrire une ligne de code complexe.</p> <p>Grâce à la méthode « Learning by doing » (70 % de pratique), vous apprenez directement par l'action. Vous utiliserez SQL pour extraire des données et construirez un tableau de bord pour les rendre visuelles et compréhensibles. Enfin, vous maîtriserez Python (via Pandas et Scikit-learn en mode recettes clés en main) pour créer vos propres modèles de machine learning (churn, scoring, recommandation).</p> <p>À l'issue de ces cours, vous repartirez avec un portfolio de 3 projets réels, déployés sur GitHub ou via l'outil Tableau. C'est la preuve tangible pour votre employeur ou les recruteurs que vous savez transformer des données brutes en décisions stratégiques... et que la data n'a plus de secrets pour vous.</p>
Objectifs
<p>À l'issue de cette formation, vous atteindrez les objectifs de compétences suivants :</p> <ul style="list-style-type: none">• extraire et manipuler des données avec le langage SQL et comprendre les infrastructures cloud ;• réaliser des tableaux de bord interactifs et visuels avec Tableau Public (Data Visualization) ;• analyser des jeux de données et mener des tests statistiques (A/B testing) avec Python ;• créer et entraîner des modèles de Machine Learning (régression, classification) pour prédire des tendances ;• restituer des résultats d'analyse de manière claire pour orienter la prise de décision.
Points forts
<ul style="list-style-type: none">• Pédagogie pratique (70%) : vous passerez la majorité du temps à pratiquer en petits groupes sur des cas concrets, loin des cours magistraux théoriques.• Stack technique moderne : vous apprendrez à utiliser des outils utilisés par des experts en Data science comme Python, SQL, Tableau et Scikit-Learn).• Portfolio professionnel : vous repartirez avec des projets aboutis à présenter en entretien pour crédibiliser votre profil immédiatement.• Accessibilité totale : un programme de formation construit spécifiquement pour démocratiser la Data, accessible sans aucune connaissance préalable en programmation.
Modalités d'évaluation
Travaux Pratiques

Pré-requis

Suivre cette formation ne nécessite au prérequis technique ou d'expérience professionnelle particulière.

Public

Cette formation est le tremplin idéal pour entrer dans le monde de la Tech. Le public inclut notamment :

- les **professionnels en reconversion** qui souhaitent acquérir rapidement des compétences en Data sans passer par un cursus long ;
- les **analystes et chargés d'études** qui désirent moderniser leurs outils (passer d'Excel à SQL/Python) pour traiter de plus gros volumes de données ;
- les **managers et chefs de projet** qui veulent comprendre le langage de leurs équipes techniques et piloter des projets Data.

Programme

Jours 1 & 2 : Maîtriser la data visualisation avec Tableau public

- La découverte de l'interface de Tableau et la connexion aux différentes sources de données.
- La mise en forme des données pour faciliter l'analyse visuelle.
- La création de tableaux de bord interactifs pour piloter l'activité.
- L'exploration poussée des données (Data Mining) pour extraire des tendances cachées.

Jours 3 & 4 : Exploiter les bases de données (SQL) et le cloud

- La compréhension du fonctionnement d'une infrastructure Data (Data Warehouse, Data Lake, ETL).
- L'écriture de requêtes SQL basiques : filtrer (WHERE), trier (ORDER BY) et dédoublonner (DISTINCT).
- La construction de requêtes SQL complexes : jointures (JOIN), agrégations (SUM, AVG) et segmentation (GROUP BY, HAVING).
- L'initiation aux services de cloud computing via Google Cloud Platform.

Jours 5 & 6 : Analyser la donnée avec Python et les statistiques

- L'acquisition des bases de l'algorithmique en Python : variables, conditions, boucles.
- La manipulation et le nettoyage des jeux de données avec la librairie Pandas.
- La maîtrise des fondamentaux statistiques : moyenne, médiane, écart-type, Z-Score.
- La conception et l'analyse des A/B tests pour optimiser des sites web ou applications.

Jours 7 & 8 : Déployer des modèles de Machine Learning (IA)

- La compréhension des concepts clés et du cycle de vie d'un projet de Machine Learning.
- Le développement de modèles de régression linéaire avec la librairie Scikit-Learn.
- La mise en œuvre d'algorithmes de classification, d'arbres de décision et de Random Forest.
- La visualisation des performances des modèles avec Matplotlib.

Jour 9 : Projet final et soutenance (Hackathon)

- Sélectionner un dataset réel (ex.: Marathon de Boston, Spotify ou Recrutement).
- Mener une analyse exploratoire et nettoyer les données en autonomie.
- Entraîner un modèle prédictif pour répondre à une problématique business.
- Restituer les résultats sous forme de storytelling visuel devant le jury.

Tableau® est une marque déposée de Tableau Software, LLC (www.tableau.com).

Python™ est une marque déposée de la Python Software Foundation (www.python.org).

*Google Cloud Platform™ est une marque de Google LLC (cloud.google.com).
Les autres noms de produits et d'entreprises mentionnés sont la propriété de leurs détenteurs respectifs.*