

Formation Data Analyst

Se former à l'analyse de données est essentiel aujourd'hui pour plusieurs raisons clés, car la donnée est devenue l'un des actifs les plus précieux des entreprises et des organisations.

Amélioration de la Prise de Décision

L'analyse de données permet de remplacer les suppositions et l'intuition par des faits concrets et des preuves.

- **Décisions Basées sur les Preuves** : Vous apprenez à extraire des insights (informations exploitables) des ensembles de données pour orienter les stratégies d'entreprise, les politiques publiques ou les décisions opérationnelles.
- **Prédiction des Tendances** : Les techniques d'analyse permettent de modéliser les tendances passées et actuelles pour prédire les événements futurs (ventes, défaillances d'équipement, comportement des clients, etc.).

Demande Croissante sur le Marché du Travail

Le rôle d'analyste de données (ou de scientifique des données) est l'un des plus demandés et des mieux rémunérés.

- **Rareté des Compétences** : Il y a une pénurie de professionnels capables de manier et d'interpréter correctement les grands volumes de données (Big Data).
- **Polyvalence Sectorielle** : Les compétences en analyse de données sont recherchées dans tous les secteurs : finance, santé, marketing, commerce de détail, technologie, gouvernement, etc.

Optimisation et Efficacité

La formation en analyse de données vous donne les outils pour identifier les domaines d'amélioration de l'efficacité et de réduction des coûts.

- **Identification des Gaspi** : Vous pouvez détecter les inefficacités dans les processus opérationnels, la chaîne d'approvisionnement ou les dépenses inutiles.
- **Personnalisation Client** : En analysant les données clients, vous aidez les entreprises à offrir des produits et services plus pertinents, augmentant ainsi la satisfaction et la fidélisation (par exemple, les systèmes de recommandation).

Développement de Compétences Essentielles

Se former à l'analyse de données développe un ensemble de compétences transférables et hautement valorisées :

- **Pensée Critique** : La capacité à poser les bonnes questions aux données et à interpréter les résultats avec scepticisme et rigueur.
- **Compétences Techniques** : Maîtrise d'outils et de langages clés comme SQL, Python (avec des bibliothèques comme Pandas ou NumPy), R, et des logiciels de visualisation de données (Tableau, Power BI).

En résumé, se former à l'analyse de données est un investissement qui offre une sécurité professionnelle élevée, un impact direct sur la performance des organisations, et le développement d'une mentalité analytique indispensable à l'ère numérique.



Nous contacter pour rédiger un dossier détaillé d'apprentissage des différents modules :

Programmation en Python,
 Numpy, Pandas, Scikit-learn
 Data visualisation : Matplotlib, Seaborn, Plotly
 API
 Web scraping
 SQL
 Machine Learning
 Business Intelligence (Power BI)

Projet à mener pour la formation complète de Data Analyst.

Pour chaque module, un micro-projet est à faire par le ou le groupe de stagiaires.

La correction du micro-projet aura lieu via Zoom selon la date fixée avec les apprentis.

Module	Durée en jours	Prix € HT
1. Programmation en Python	3	1800
2. Maîtriser la librairie Numpy	1	600
3. Maîtriser la librairie Pandas	2	1200
4. Apprendre la Data visualisation	5	3000
5. Maîtriser les API pour extraire de la data	2	1200
6. Apprendre le Web scraping	2	1200
7. Maîtriser le langage SQL	3	1800
8. Le machine learning	5	3000
9. Business Intelligence (Power BI)	4	2400

Contenus des modules de formation

Module 1 : Programmation en Python

Installation et première prise en main de l'environnement de travail

Introduction

Qu'est-ce que la Data Science ?

Présentation et installation d'Anaconda

Découverte d'Anaconda

Maîtrise du notebook Jupyter : les bases

Maîtrise du notebook Jupyter : enregistrer son travail

Les bases du langage Python

Introduction

Concept de variable et déclaration d'une variable

Concept de type de variable

Les chiffres et les opérateurs mathématiques

Les strings et les opérateurs

Vérifier et modifier le type d'une variable

Maîtrise de jupyter notebook : session 2

Présentation des exercices

Corrections des exercices

Les conditions

Les conditions logiques

La structure des conditions

Présentation des exercices

Corrections des exercices

Les listes

Concept et déclaration d'une liste en Python

Consultation et analyse d'une liste

Manipulations d'une liste

TP : Analyse des données du Titanic !

Correction des travaux pratiques

Les dictionnaires

Qu'est-ce qu'un dictionnaire ?

Navigation et manipulation sur un dictionnaire

Les dictionnaires multi-valeurs

TP : Analyse des données du Titanic !

Correction des travaux pratiques

Les boucles

Concept des boucles avec While

La boucle For

Création de listes de chiffres

Les fonctions

Le concept de fonction

Les paramètres de fonctions

“Return” dans les fonctions

Les paramètres optionnels

TP : les boucles et les fonctions d’abord

Analyse des données du Titanic

Correction des travaux pratiques

Les librairies

Importer et utiliser une librairie

Utiliser le gestionnaire PIP

Module 2 : Maîtriser la librairie Numpy

Créer un tableau et sélectionner des valeurs

Faire des opérations dans un tableau

Analyser un tableau

Manipuler un tableau

La fonction “Arrange”

TP : Analyse des boissons délivrées dans les Starbucks

Correction des travaux pratiques

Module 3 : Maîtriser la librairie Pandas

Rappel : Une première approche de la statistique descriptive

Les types de variables

Les statistiques de base

Qu’est-ce que la loi normale

Calculer la corrélation entre 2 variables

Pandas :

Créer votre premier Dataframe

Comprendre la composition d’un Dataframe

Navigation dans un Dataframe

Faire une boucle dans un Dataframe

Le Data cleaning

Les premières analyses avec Pandas

Introduction au feature modeling

Rassembler plusieurs jeux de données dans un Dataframe

La gestion des dates

Modifier le paramétrage de Pandas

TP : faire du shopping sur un site e-commerce !

Correction des travaux pratiques

Module 4 : Apprendre la Data visualisation

- La librairie Matplotlib
- Créer son premier graphique
- Apprendre à gérer le multi graph
- Le bar chart
- L'histogramme de répartition
- Le nuage de points et scatter matrix
- La librairie Seaborn
- TP : Quel est ton Digimon préféré ?
- Corrections des exercices

Module 5 : Maîtriser les API pour extraire de la data

- Qu'est-ce qu'une API ?
- Qu'est-ce que le format Json ?
- Comment requêter une API en Python ?
- Transformer du Json en Dataframe
- Dynamiser le paramétrage de vos appels API
- TP : Récupérez des Data sur les pays du monde !
- Correction des travaux pratiques

Module 6 : Apprendre le Web scraping

- Qu'est-ce que le web scraping ?
- La librairie "Beautifulsoup"
- TP : Crawler un site e-commerce pour récupérer des données sur les produits !
- Correction des travaux pratiques

Module 7 : Maîtriser le langage SQL

- Introduction au SQL
- Faire des filtres du votre base de données
- Faire des tris
- Faire des calculs en SQL
- Group by et les Alias
- Les concepts de clé primaire et étrangère
- Faire des jointures en SQL
- TP : analyser les parcours des 'velib' de San Fransisco
- Correction des travaux pratiques

Module 8 : Le machine learning

- Le concept du machine learning
- Comprendre le type de problème à résoudre
- Le data cleaning : analyse de la valeur à prédire
- Le data cleaning : vérifier le format des colonnes
- Le data cleaning : gestion des valeurs manquantes
- Le data cleaning : corriger problème de logique de certaines colonnes
- Le data modeling
- Concept de la régression linéaire
- Splitter le jeu de données entre entraînement et test
- Créer son premier modèle uni-varié
- La normalisation des variables
- Mise en pratique de la normalisation des variables
- Le dilemme biais-Variance
- La gestion des variables avec catégories
- TP : Estimer les frais de santé d'une personne en fonction de ses caractéristiques
- Correction des travaux pratiques

Module 9 : Business Intelligence (Power BI)

- Connexion aux sources de données
- Data Visualisation avec Power Query
- Relations et Modélisation de données transformées
- Introduction au langage DAX
- Création d'un tableau de bord approfondi et étude de cas