

# Devenir Data Analyst

MaCarrière

m2i  
Formation

Document mis à jour le 25/04/2025

Dispositif de formation : **Préparation Opérationnelle à l'Emploi (POE)**  
Date de formation : **Du 19 mai au 8 août 2025**  
Réunion d'Information : **Le 11 avril 2025**  
Date limite d'envoi de candidature : **Le 6 mai 2025**

Dans le cadre d'une **Préparation Opérationnelle à l'Emploi Collective** de 57 jours, qui se déroulera du **19 mai 2025 au 08 août 2025** en distanciel, M2i Formation, en partenariat avec l'**OPCO Atlas Auvergne-Rhône-Alpes**, vous propose de vous former au métier de Data Analyst.

La science des données (ou Data Science en anglais) est devenue au cours de ces dernières années, un élément indispensable pour permettre aux entreprises d'innover, performer et de se différencier.

Les directions métiers, tous secteurs confondus, utilisent déjà ces sciences et cette tendance est amenée à augmenter exponentiellement au cours des prochaines années.

L'objectif visé est de devenir un spécialiste en science des données, acquérir la capacité d'avoir une vision globale des projets, de les concevoir et de les diriger dans leur ensemble.

*Exemple de métiers visés :*

Data Scientist, Data Engineer, Data Analyst  
Machine Learning Engineer  
Développeur Full-Stack Python (ou Software Engineer)  
Consultant Intelligence Artificielle (ou Consultant Data)  
Product Manager Data (ou Product Owner ou chef de produit Data)  
Chef de Projet Data (ou Project Manager Data)  
Ingénieur d'étude ou Business Intelligence Directeur  
Directeur de l'Organisation et des Systèmes d'Information -DOSI  
Attaché d'études statistiques



## OBJECTIFS DE FORMATION

A l'issue de cette formation, vous serez capable de :

- Collecter des quantités importantes de données structurées et non structurées
- Normaliser les données en conservant leur intégrité (Qualité de données)
- Manipuler des structures de données à l'aide d'un langage de programmation
- Utiliser les statistiques descriptives afin de modéliser les données et en faire émerger des informations pertinentes
- Exploiter des données massives en s'appuyant sur les outils d'analyse de données du Big Data afin d'extraire des indicateurs utiles aux décideurs
- Présenter les résultats sous forme graphique (Dash Board)
- Aider à la prise de décision en se basant sur l'analyse effectuée

## **PRÉREQUIS**

---

- Diplômés en études supérieures (Bac +4/+5 requis) scientifiques
- Bon niveau en mathématiques et statistiques
- Connaissance d'un langage de programmation et des fondamentaux de l'infrastructure Windows et/ou Linux
- Capacité d'analyse et de synthèse
- Rigueur et sens de la méthode
- Connaissance de l'anglais est un plus

## **PUBLIC CONCERNÉ ET FINANCEMENT**

---

**Cette formation financée à 100%\* est à destination des demandeur-se-s d'emploi inscrit-e-s à France Travail en région Auvergne Rhône Alpes.**

Après validation de votre candidature par l'entreprise, elle sera soumise à votre conseiller France Travail pour valider la cohérence entre votre profil et votre projet de formation.

# COMPÉTENCES ET TECHNOLOGIES ABORDÉES

Compétences	Modules	Durée (Jours)	Durée (Heures)
Les fondamentaux	<b>Présentation du cursus</b> , des outils pédagogiques et du projet « fil rouge ».	0,5	3,5
	<b>Les fondamentaux des commandes Linux :</b> Utiliser interactivement le Shell et connaître les commandes essentielles - Gérer les fichiers et les dossiers - Editer un fichier - Reconnaître les métacaractères et les expressions régulières	2,5	17,5
	<b>Révision - Algorithmique :</b> Introduction - Algorithmie - Variables et types de données - La "grammaire" d'un langage - Structurer son programme	1	7
	<b>Gestion de projets Data :</b> Introduction aux méthodes Agiles - Bien identifier les challenges dans la mise en place d'un projet Data - mesurer l'impact stratégique sur l'entreprise et la pertinence d'un reporting régulier - Mettre en place une bonne gestion de la conduite du changement	2	14
Stockage et manipulation de données	<b>Modéliser, concevoir et interroger une base de données structurée en SQL:</b> Les fondamentaux des bases de données relationnelles - Ecrire des instructions SELECT pour extraire des colonnes d'une ou de plusieurs tables - Trier et filtrer des données sélectionnées - Utiliser des fonctions intégrées pour renvoyer des valeurs de données - Créer des groupes de données et agréger les résultats - Modifier des données en utilisant INSERT, UPDATE, DELETE et MERGE.	3	21
	<b>Base de données NoSQL :</b> Découvrir les bases de données NoSQL - Caractéristiques NoSQL - Principaux acteurs - Mise en œuvre - Introduction à Cassandra ou Mongo DB- Installation, configuration et mise en œuvre locale - Introduction aux différents outils proposés - Architecture interne de Cassandra ou MongoDB	3	21
	<b>ETL Talend Open Studio :</b> Extraction , Transformation et chargement des données à partir des sources hétérogènes.	4	28
	<b>Appréhender l'écosystème Hadoop</b> L'écosystème d'Hadoop - Manipulation des données dans un Cluster Hadoop - Les particularités des distributions - Importer des données dans un cluster Hadoop	3	21
	<b>Ecosystem Spark :</b> Les bases de spark - Manipulation des dataframes et des schemas - Analyser des données avec des requêtes sur dataframes - Les rdd – structure fondamentale de spark - Transformer les données avec des rd - Agrégation de données avec les rdd de paires - Requête de tables et de vues avec spark sql - Travailler avec les datasets spark en scala - Ecrire, configurer et lancer des applications spark - Le traitement distribué avec spark - Persistance de la donnée distribuée - Introduction à spark structured streaming – Introduction à PySpark	4	28
	<b>Rechercher, analyser et visualiser des données avec la suite Elasticsearch</b> Introduction à la recherche d'information - Maîtriser l'utilisation du moteur de recherche - Ingérer des données provenant d'une multitude de sources avec Logstash - Analyser et visualiser des données avec Kibana - Mettre en place une architecture avancée (scalabilité et clustering) - Administrer et déployer la suite Elastic	3	21

	<b>L'ingénierie des données avec Microsoft Azure :</b> Introduction au cloud et à la solution Azure - Expérimenter les options de calcul et de stockage pour les workloads de Data Engineering dans Azure - Exécuter des requêtes interactives à l'aide de pools SQL sans serveur - Effectuer l'exploration et la transformation des données dans Azure Databricks - Explorer, transformer et charger des données dans le Data Warehouse à l'aide d'Apache Spark - Ingérer et charger des données dans le Data Warehouse - Transformer les données avec Azure Data Factory ou Azure Synapse Pipelines - Intégrer les données à partir des notebooks avec Azure Data Factory ou Azure Synapse Pipelines - Prendre en charge le traitement analytique transactionnel hybride (HTAP) avec Azure Synapse Link - Assurer la sécurité end-to-end avec Azure Synapse Analytics - Effectuer un traitement de flux en temps réel avec Stream Analytics - Créer une solution de traitement de flux avec Event Hubs et Azure Databricks.	4	28
Programmation Python	<b>Les fondamentaux du langage Python :</b> Connaître les usages courants du langage - Maîtriser le scripting en Python - Structurer son code en fonction, classes et modules - Utiliser des modules existants - Maîtriser la programmation objet en Python	4	28
	<b>Scrapping de données avec Python:</b> Comprendre l'intérêt et le processus du scraping, - Architecture d'internet, - Où et comment valoriser les données web, - Présentation et explication d'exemples concrets - Extraction de base de données via les réseaux sociaux (extraction linkedin, facebook, twitter, ...), - Codage et automatisation	2	14
	<b>Python pour l'analyse de données :</b> Traiter des bases de données grâce à des analyses statistiques descriptives et inférentielles via des bibliothèques de programmation avec Numpy et Pandas , pour les organiser et les nettoyer afin de les normaliser par rapport à la population étudiée.	2	14
Analyse et visualisation	<b>Visualiser les données avec Python :</b> Présenter le résultat d'une analyse statistique de données structurées, massives ou non, grâce à des bibliothèques de programmation avec python pour synthétiser ce résultat devant un public profane afin de faciliter la prise de décisions et appuyer leurs déclinaisons opérationnelles	2	14
	<b>Analyse statistique avec R :</b> Installer et utiliser l'environnement d'analyse R - Manipuler des données avec R - Importer et exporter des données - Reconnaître les différents types d'objets de R - Créer des programmes d'analyse avec R - Réaliser des analyses statistiques basiques avec R - Restituer des résultats à l'aide de graphiques.	2	14
	<b>Modéliser et visualiser les données avec Power BI :</b> Traiter et analyser des ensembles de données importants - Se connecter aux sources de données et effectuer des transformations de données, modéliser et visualiser des données à l'aide de Microsoft Power BI Desktop - Mettre en pratique le langage DAX - Configurer des tableaux de bord à l'aide du service Power BI	4	28
Soft skills	<b>Technique de recherche d'emploi :</b> Aborder de manière sereine les outils de candidatures - Savoir rédiger et construire de manière personnalisée et impactante des outils de candidature (CV et lettre de motivation) - Avoir une posture d'offreur de services et non de demandeur d'emploi. - Préparer et réussir un entretien d'embauche - Vous positionner et avoir l'état d'esprit de réussite d'un entretien ou d'un job dating - Préparer les étapes d'un entretien (avant, pendant et après) - Savoir anticiper et pouvoir répondre à tous types de questions - Connaître un cadrage de réponses pour des questions d'entretiens..	2	14
Validation des acquis	<b>Travaux pratiques :</b> Fil rouge : intégration et traitement de données avec Talend, Hadoop et Spark – Mise en place d'une solution en environnement Microsoft Azure – Atelier de visualisation avec Power BI – Analyse et visualisation avec Python	8	56
	<b>Préparation à la certification Power BI</b>	1	7

# À PROPOS DES CERTIFICATIONS

## Certification éditeur :

Les certifications éditeurs dépendent des éditeurs uniquement (PEGA, SAP, Salesforce, Microsoft, Red Hat...). Chaque éditeur a sa propre façon d'évaluer son candidat.

Cela se fait souvent par le biais d'un QCM chronométré. Sur cette base, il faut généralement obtenir un pourcentage de succès supérieur à 70%. Merci de prendre le temps de demander à vos interlocuteurs M2i le cadre et les conditions de ce passage.

## Validation des acquis M2i :

La validation des acquis M2i se fait soit par un QCM chronométré, soit par un examen de fin de parcours ou les candidats passent par groupes de 2 ou 3 devant un jury qui déterminera vos acquis suite à votre cursus de formation. Généralement cette soutenance est précédée de 3 jours de travaux pratiques en groupe afin de préparer cette soutenance. Les équipes M2i pourront vous guider dans votre projet.

# LES PLUS DE M2I FORMATION

Microsoft Teams Education, un outil de suivi et d'animation en présentiel et à distance :

- Espace de stockage collaboratif pour accéder aux différents supports de cours et cahiers d'exercices
- Fonctionnalités pour gérer des sessions à distance

La playlist e-learning : tous les apprenants ont accès avant, pendant et après le cursus à notre plateforme e-learning M2i Learning pour :

- Acquérir les connaissances prérequis avant de démarrer (sous condition de l'existence des modules pour combler les lacunes)
- Utiliser les modules conseillés par les formateurs pour faire de l'ancrage mémoriel sur des sujets abordés pendant la formation
- Revenir sur un sujet après la formation pour continuer à s'auto-former

# MODALITÉS, MOYENS ET MÉTHODES PÉDAGOGIQUES

Formation délivrée en présentiel et/ou distanciel (e-learning, classe virtuelle, présentiel à distance).

Le formateur alterne entre des méthodes démonstratives et actives (via des travaux pratiques et/ou des mises en situation).

La validation des acquis peut se faire via des études de cas, des quiz et/ou une certification.

# MODALITÉS D'ACCÈS

Nos équipes accorderont toute leur attention au traitement de votre candidature et s'engagent à vous faire un premier retour dans un délai de 7 jours.

# ADMISSION

- Dossier de candidature
- Test de vérification des prérequis
- Entretien
- Validation et éligibilité France Travail (dans le cas de certains dispositifs)

Le groupe M2i s'engage pour faciliter l'accessibilité de ses formations. Les détails de l'accueil des personnes en situation en handicap sont consultables sur la page [Accueil PSH](#).

## **POURQUOI CHOISIR M2I FORMATION ?**

- Un apprentissage métier proactif axé sur l'emploi et basé sur le faire avec l'accompagnement de nos formateurs tout au long du parcours.
- Un accès à des experts : bénéficiez de l'expertise de nos formateurs.
- En présentiel ou à distance : accès individuel aux ressources de formation et progression personnalisée si besoin.
- Outils de suivi collectif et individuels (espaces d'échanges et de partage en ligne, salles virtuelles si en présentiel ou à distance, supports de cours, TP, exercices).

## **MODALITÉS D'ÉVALUATION**

Exemples de validation des acquis de formation :

- Travaux dirigés dans chaque module
- Mise en situation via des cas pratiques et un mini projet
- Certification (si prévue dans le programme de formation)
- Soutenance devant un jury de 30 à 40 min

