

# Devenir Data Analyst

MaCarrière

m<sup>2</sup>i  
Formation

Document mis à jour le 19/05/2025

Dispositif de formation : **Préparation Opérationnelle à l'Emploi (POE)**  
Date de formation : **Du 16 juin au 16 septembre 2025**  
Date limite d'envoi de candidature : **Le 9 juin 2025**

Dans le cadre d'une **Préparation Opérationnelle à l'Emploi Collective** de 57 jours, qui se déroulera du **16 juin 2025 au 16 septembre 2025** en distanciel, M2i Formation, en partenariat avec l'**OPCO Atlas Normandie**, vous propose de vous former au métier de Data Analyst (POEC de 15 personnes).

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

- Diplômé d'un BAC +3 scientifique minimum
- Bon niveau d'anglais
- Une expérience dans les domaines de la programmation, mathématiques, statistiques ou algorithmie sont un plus
- Capacité d'analyse et de synthèse
- Rigueur et sens de la méthode
- Mobilité en Normandie

## **PUBLIC CONCERNÉ ET FINANCEMENT**

---

**Cette formation financée à 100% par l'Opco Atlas est à destination des demandeur·se·s d'emploi inscrit·e·s à France Travail en Normandie.**

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

# COMPÉTENCES ET TECHNOLOGIES ABORDÉES

Catégorie	Module	Durée en j	Durée en h
	<p><b>Présentation</b></p> <p>Présentation du cursus, des plateformes pédagogiques et du projet fil rouge - Présentation du métier - Activités brise-glace pour mettre en confiance les apprenants.</p>	0,5	3,5
Méthode	<p><b>Revoir les bases de l'algorithmie :</b></p> <p>Introduction - Algorithmie - Variables et types de données - La "grammaire" d'un langage - Structurer son programme.</p>	1,5	10,5
Fondamentaux	<p><b>Avoir une vue d'ensemble du Big Data :</b></p> <p>Identifier l'ensemble des enjeux et facteurs à prendre en compte pour réussir l'intégration du Big Data dans la vision large du SI - Evaluer et sélectionner les outils appropriés dans le cadre d'un plan de mise en œuvre du Big Data.</p>	2	14
Base de données	<p><b>Interroger et manipuler les objets de bases de données avec le langage SQL :</b></p> <p>Présenter une vue d'ensemble du SGBDR PostgreSQL, de ses fonctionnalités et des différentes sources d'informations utiles - Décrire les concepts de l'algèbre relationnelle utilisés dans le langage SQL - Interroger une base de données avec la clause SQL SELECT - Utiliser les commandes SQL de mise à jour des données - Identifier les commandes SQL de début et fin de transaction BEGIN, COMMIT et ROLLBACK - Présenter les concepts de gestion des privilèges systèmes et objets avec les commandes SQL GRANT et REVOKE - Créer, modifier et supprimer certaines catégories d'objets (table, index, vues...) avec CREATE, ALTER et DROP.</p>	3	21
Validation des acquis intermédiaire	<p><b>Travaux pratiques :</b></p> <p>Interroger la base de données support du projet fil rouge.</p>	1	7
Compétence transverse	<p><b>Technique de Recherche d'emploi : Concevoir son CV et sa lettre de motivation :</b></p> <p>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.</p>	1	7
Développement	<p><b>Maîtriser le langage Python pour gérer des données :</b></p> <p>Connaître les usages courants du langage - Maîtriser les fondamentaux du langage - Structurer son code en fonction, classes et modules - Utiliser des modules existants - Manipuler et gérer des tableaux de données - Interroger, manipuler et ordonner des jeux de données- Maîtriser la programmation objet en Python - Gérer la persistance des données</p>	6	42
Scraping	<p><b>Extraire et gérer des données (scraping) :</b></p> <p>Introduction à HTML/CSS - Comprendre l'architecture d'un site Web - Réaliser du scraping de données à partir du web - Appliquer le scraping sur google - Faire les actions d'ingestion nécessaires pour alimenter un Data Lake.</p>	3	21
Analyse	<p><b>Analyser des données avec Python :</b></p> <p>Créer et manipuler des tableaux avec Numpy - Explorer des données avec Pandas - Créer des graphiques avec Matplotlib et pandas plot -</p>	3	21

<b>Validation des acquis intermédiaire</b>	<b>Travaux pratiques :</b> Extraire les données du projet « fil rouge » depuis le Web et les analyser	2	14
<b>Dataviz</b>	<b>Rendre visible l'invisible avec la Data Visualisation :</b> Maîtriser et personnaliser des rendus graphiques variés –Maîtriser les bonnes pratiques de la data visualisation – Utiliser les bibliothèques Python Matplotlib et Seaborn. - Produire des graphiques interactifs et les intégrer dans une page Web avec Bokeh	3	21
<b>Analyse</b>	<b>Traiter les données avec l'écosystème Spark :</b> Les bases de Spark – Spark et Hadoop - Manipulation des dataframes et des schémas - Analyser des données avec des requêtes sur dataframes - Les RDD : Structure fondamentale de Spark - Transformer les données avec des RDD - Agrégation de données avec les RDD de paires - Requêtage 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.	3	21
<b>Validation des acquis intermédiaire</b>	<b>Travaux pratiques :</b> Manipulation des données du projet fil rouge.	1	7
<b>Base de données</b>	<b>Mettre en œuvre l'ingestion de données avec NiFi :</b> Introduction NiFi - Différence entre données froides et données chaudes - Présentation des outils et technologies Big Data - Hadoop (HDFS et MapReduce) - Installation et configuration de NiFi - Vue d'ensemble de l'architecture NiFi - Approches de développement - Outils de développement d'applications et état d'esprit - Extraction, transformation et chargement (ETL) des outils et de la mentalité	2	14
<b>Validation des acquis intermédiaire</b>	<b>Travaux pratiques :</b> Mise en œuvre d'un processus d'ingestion de données.	1	7
<b>Base de données</b>	<b>Découvrir les bases de données NoSQL :</b> Caractéristiques NoSQL - Principaux acteurs - Mise en œuvre - Introduction au Big Data et à Cassandra - Installation, configuration et mise en œuvre locale de Cassandra - Introduction aux différents outils proposés par Cassandra - Architecture interne de Cassandra - Découvrir MongoDB - Découvrir HBASE.	2	14
<b>Analyse</b>	<b>Prendre en main et développer avec Elasticsearch :</b> Situer Elasticsearch dans un écosystème Big Data - Identifier les enjeux et les cas d'utilisation d'un moteur de recherche - Indexer des données - Faire des recherches simples et complexes - Manipuler les agrégations et des recherches évoluées.	2	14
<b>Compétence transverse</b>	<b>Technique de Recherche d'emploi : Faire une simulation d'entretien d'embauche pendant un job dating :</b> 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.	1	7
<b>Machine learning</b>	<b>Comprendre la Data Science et le Machine Learning :</b> Identifier les enjeux de la Data Science et de l'IA - Schématiser le cycle d'un projet Data Science - Appliquer les meilleures pratiques en matière de nettoyage et de préparation de vos données avant l'analyse - Modéliser un problème de Data Science.	2	14

<b>Analyse</b>	<b>Traiter des données avec PySpark :</b> Utiliser l'API PySpark pour interagir avec Spark en Python - Mettre en œuvre les méthodes de Machine Learning avec la librairie MLlib de Spark - Traiter les flux de données avec Spark Streaming - Manipuler les données avec Spark SQL.	3	21
<b>Text Mining</b>	<b>Utiliser les technologies de fouilles de texte (Text Mining) :</b> Comprendre comment fonctionne le text mining - Se repérer dans le périmètre technologique - Mettre en œuvre l'extraction des caractéristiques de données textuelles - Choisir un algorithme de classification	2	14
<b>Méthode</b>	<b>Concevoir et piloter un projet Big Data :</b> Bien identifier les challenges dans la mise en place d'un projet Big Data - Mesurer l'impact politique et stratégique sur l'entreprise et la pertinence d'un reporting régulier - Mettre en place une bonne gestion des conflits et de la conduite du changement.	2	14
<b>Compétence transverse</b>	<b>Développer ses compétences comportementales :</b> Les nouveaux contextes de travail - Faire le point sur sa cartographie de compétences - Communication - Intelligence émotionnelle - Se placer dans une dynamique positive - mieux gérer son temps - Gérer la pression.	2	14
<b>BI</b>	<b>Modéliser, analyser et visualiser des données avec Power BI :</b> Décrire le cycle de création d'un rapport Power BI - Vous connecter à des sources de données - Transformer, nettoyer et combiner des sources - Structurer un modèle de données - Créer des indicateurs - Afficher les indicateurs dans des rapports PBI - Utiliser les options avancées du langage DAX - Partager en ligne vos tableaux de bords et rapports - Utiliser des visualisations interactives.	4	28
<b>Validation des acquis finale (Phase 1)</b>	<b>Travaux pratiques :</b> Travaux pratiques de synthèse : Finalisation du projet fil rouge	3	21
<b>Validation des acquis finale (Phase 2)</b>	<b>Révisions et préparation de la certification TOSA Python :</b> Obtenir des informations, des astuces et des conseils relatifs au déroulement de l'examen - Analyser des exemples d'études de cas - Passer en revue chaque section de l'examen et leurs principaux concepts - Renforcer les connaissances et identifier les lacunes/domaines d'approfondissement - Passage de la certification.	1	7

## CERTIFICATION

TOSA Python

## 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

