Ingénierie Big Data et Analytics sur Microsoft Azure

Lien : https://innov-systems.com/formation/ingenierie-big-data-et-analytics-sur-microsoft-azure

DURÉE	RÉFÉRENCE	CATÉGORIE
5 jours (35h)	**VSC327**	**Amazon Web Services**

OBJECTIFS DE LA FORMATION

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

- Comprendre les principes et l’architecture du Big Data dans l’écosystème Azure
- Concevoir et déployer des pipelines de données performants et sécurisés
- Mettre en œuvre le traitement de flux en temps réel avec Azure Stream Analytics
- Exploiter Azure Data Lake et Azure Synapse Analytics pour l’analyse avancée des données
- Intégrer des fonctions personnalisées et de Machine Learning dans les traitements Big Data
- Automatiser les flux de données avec Azure Data Factory et orchestrer les transformations à grande échelle

POUR QUI ?

- Architectes Cloud et Data
- Ingénieurs et Développeurs Big Data
- Responsables Data et BI
- Administrateurs de systèmes Cloud souhaitant maîtriser les environnements Data Azure

Innov Systems

Programme détaillé

1 / PRINCIPES ET ARCHITECTURES DU BIG DATA

- Comprendre les concepts fondamentaux du Big Data et des systèmes distribués
- Identifier les scénarios d'usage (batch, temps réel, hybride)
- Présentation des services Azure pour la donnée (Data Lake, Synapse, Data Factory, Stream Analytics)

2 / CONCEPTION D'UNE ARCHITECTURE BIG DATA SUR AZURE

- Composants d'une architecture Data moderne sur le Cloud
- Modèles d'intégration : ingestion, stockage, transformation et analyse
- Bonnes pratiques d'architecture pour la performance et la sécurité

3 / PRISE EN MAIN DE L'ENVIRONNEMENT AZURE

- Création et configuration d'un espace de travail Big Data
- Gestion des identités, accès et ressources via Azure Portal et CLI
- Mise en place des premières ressources Data Lake et Synapse

4 / DÉCOUVERTE D'AZURE STREAM ANALYTICS

- Architecture et cas d'usage de Stream Analytics
- Sources et sorties compatibles : IoT Hub, Event Hub, Power BI, Data Lake
- Écriture et exécution de requêtes SQL temps réel

5 / CONCEPTION DE JOBS DE TRAITEMENT DE FLUX

- Définir et tester des jobs de traitement continu
- Agrégations et fenêtres temporelles
- Gestion des erreurs et des performances en streaming

Innov Systems

6 / EXTENSION AVANCÉE DE STREAM ANALYTICS

- Ajout de fonctions personnalisées avec JavaScript et .NET
- Intégration de modèles de Machine Learning pour l’analyse prédictive
- Optimisation et supervision des jobs temps réel

7 / STRUCTURATION ET GOUVERNANCE DES DONNÉES DANS AZURE DATA LAKE

- Architecture et principes de stockage hiérarchique
- Bonnes pratiques de structuration, nommage et sécurité
- Gestion des droits d’accès et suivi des activités

8 / TRAITEMENTS ET ANALYSES AVANCÉES AVEC DATA LAKE ANALYTICS

- Introduction au langage U-SQL et ses usages
- Écriture de requêtes de transformation, filtrage, agrégation et jointure
- Optimisation des performances des jobs U-SQL

9 / INTÉGRATION DE FONCTIONNALITÉS PERSONNALISÉES

- Création de fonctions et d’extensions personnalisées dans U-SQL
- Gestion des ressources et suivi des performances
- Scénarios d’intégration avec d’autres services Azure (Machine Learning, Power BI)

10 / INTRODUCTION À AZURE SYNAPSE ANALYTICS (EX SQL DATA WAREHOUSE)

- Concepts d’architecture MPP et différence avec les bases relationnelles classiques
- Définir les besoins en stockage et calcul
- Création et configuration d’un entrepôt de données

11 / CONCEPTION DE TABLES ET OPTIMISATION DES REQUÊTES

- Stratégies de distribution et d’indexation pour la performance
- Chargement de données depuis Data Lake et Data Factory

Inov Systems

- Meilleures pratiques pour le partitionnement et la maintenance

12 / ANALYTIQUE ET VISUALISATION DES DONNÉES

- Requêtes analytiques avancées avec T-SQL
- Connexion de Synapse à Power BI pour la visualisation
- Sécurisation et surveillance de la performance

13 / INTRODUCTION À AZURE DATA FACTORY

- Rôle d'ADF dans les architectures de données modernes
- Concepts de pipeline, dataset et linked services
- Création et planification de pipelines de données

14 / TRANSFORMATION ET ENRICHISSEMENT DES DONNÉES

- Activités de transformation avec Data Flow
- Intégration de fonctions personnalisées et scripts Python/SQL
- Gestion des erreurs et reprise automatique des traitements

15 / SUPERVISION ET OPTIMISATION DES PIPELINES

- Surveillance en temps réel et alertes
- Analyse des coûts et optimisation des performances
- Bonnes pratiques d'orchestration et d'automatisation de bout en bout

Inov Systems

Approche pédagogique

- Support Ecrit et Projection
- Exposés Interactifs, Podcasts et Vidéos
- Brainstorming et Jeux de Rôle
- Cas Pratiques et Labs inclus pour leur impact opérationnel
- Test de Validation des Acquis des Connaissances

Prochaines dates programmées

25 au 29 Jan. 2027	Présentiel - Casablanca
22 au 26 Mars 2027	Présentiel - Casablanca
17 au 21 Mai 2027	Présentiel - Casablanca
12 au 16 Juil. 2027	Présentiel - Casablanca
06 au 10 Sep. 2027	Présentiel - Casablanca

Autres dates possibles sur demande. Contactez-nous pour organiser une session intra-entreprise.

Réservation & Renseignements

Téléphone : +212 522 247 210
Email : contact@innov-systems.com
Web : https//www.innov-systems.com

Innov Systems — Tous droits réservés