Data Engineering et Analytics sur Microsoft Azure

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

DURÉE
5 jours (35h)

RÉFÉRENCE
BSI530

CATÉGORIE
Data Science et Data Engineering

OBJECTIFS DE LA FORMATION

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

- Comprendre les différentes options de calcul et de stockage pour les charges de travail Data Engineering sur Azure
- Maîtriser l’exploration, la transformation et le traitement des données à grande échelle avec Azure Databricks et Synapse
- Développer des pipelines ETL/ELT robustes et sécurisés pour un data warehouse moderne
- Appliquer les meilleures pratiques pour le traitement en temps réel et le traitement hybride analytique et transactionnel (HTAP)
- Garantir la sécurité et la conformité des données dans un environnement Azure

POUR QUI ?

- Ingénieurs Data et Data Engineers
- Développeurs spécialisés en données
- Administrateurs Azure et BI
- Analystes techniques souhaitant automatiser les flux de données

Innov Systems

Programme détaillé

1 / INTRODUCTION AUX ARCHITECTURES DATA SUR AZURE

- Présentation des services Azure pour Data Engineering et Analytics
- Comprendre les concepts de Data Lake, Delta Lake et Data Warehouse
- Identifier les bonnes pratiques d’architecture pour batch, streaming et HTAP

2 / EXPLORATION ET TRANSFORMATION DE DONNÉES AVEC AZURE DATABRICKS

- Manipulation avancée des DataFrames avec Apache Spark
- Gestion des fichiers Parquet, CSV et JSON dans un Data Lake
- Techniques de nettoyage, agrégation et enrichissement des données
- Optimisation des performances via caching et partitionnement

3 / REQUÊTES INTERACTIVES ET SECURITÉ AVEC AZURE SYNAPSE

- Utilisation des serverless SQL pools pour l’exploration de données
- Création d’objets de métadonnées et gestion des vues
- Mise en œuvre de la sécurité : RBAC, ACL, Data Masking, RLS et CLS
- Intégration avec Azure Active Directory et Key Vault

4 / CONSTRUCTION DE PIPELINES ET INTÉGRATION DE DONNÉES

- Orchestration de pipelines ETL/ELT avec Azure Data Factory et Synapse Pipelines
- Transformation code-free et Mapping Data Flows
- Chargement efficace dans les dedicated SQL pools et gestion de workloads
- Intégration de multiples sources de données et gestion des erreurs

5 / TRAITEMENT EN TEMPS RÉEL ET ANALYSE STREAMING

- Ingestion de données en temps réel avec Event Hubs et Stream Analytics

Innov Systems

- Traitement de flux avec Azure Databricks Structured Streaming
- Utilisation de fenêtres glissantes, watermarking et partitionnement pour l'optimisation
- Analyse et stockage des flux en temps réel dans Synapse Analytics

6 / INGESTION ET TRANSFORMATION AVANCÉES AVEC APACHE SPARK

- Techniques avancées de transformation et jointure entre Data Lake et Data Warehouse
- Utilisation de Spark Pools pour le traitement batch et streaming
- Meilleures pratiques pour la gestion des fichiers volumineux et hiérarchiques

7 / HYBRID TRANSACTIONAL ANALYTICAL PROCESSING (HTAP) AVEC SYNAPSE LINK

- Connexion et synchronisation entre Azure Cosmos DB et Synapse Analytics
- Requêtes analytiques en temps réel sur les données transactionnelles
- Optimisation des performances et cas d'usage HTAP

8 / SÉCURITÉ ET CONFORMITÉ DES DONNÉES

- Gestion des secrets et clés avec Azure Key Vault
- Implémentation de contrôles d'accès et masquage dynamique des données
- Conformité aux normes et régulations sur les données sensibles

9 / PROJET PRATIQUE FINAL

- Conception et implémentation d'un pipeline complet d'ingestion, transformation et stockage
- Combinaison des traitements batch, streaming et HTAP
- Présentation des résultats et recommandations pour production

10 / OPTIMISATION ET BONNES PRATIQUES

- Optimisation des performances des pipelines et requêtes SQL
- Gestion des coûts et monitoring des workloads Azure
- Approches avancées pour Data Engineering scalable et résilient

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

12 au 18 Jan. 2027	Présentiel - Casablanca
08 au 12 Mars 2027	Présentiel - Casablanca
03 au 07 Mai 2027	Présentiel - Casablanca
28 Juin au 02 Juil. 2027	Présentiel - Casablanca
23 au 27 Août 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