Great Expectations & Soda : Automatiser les Contrôles de Qualité des Données dans vos Bases

Lien : https://innov-systems.com/formation/great-expectations-soda-automatiser-les-controles-de-qualite-des-donnees-dans-vos-bases

DURÉE
5 jours (35h)

RÉFÉRENCE
BDI97

CATÉGORIE
Big Data : fondamentaux et architecture

Innov Systems

OBJECTIFS DE LA FORMATION

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

- Comprendre les enjeux de la qualité des données
- Maîtriser Great Expectations et Soda
- Automatiser les contrôles de données dans les pipelines
- Détecter les anomalies et incohérences de données
- Améliorer la fiabilité des systèmes data
- Intégrer la data quality dans les workflows ETL/ELT
- Mettre en place des règles de validation métier
- Déployer des systèmes de monitoring de données
- Réduire les erreurs analytiques et opérationnelles
- Améliorer la gouvernance des données
- Gérer les incidents liés à la qualité des données
- Étudier des cas concrets de data observability

POUR QUI ?

- Data engineers
- Data analysts
- Data scientists
- Analytics engineers
- DevOps data / MLOps engineers
- Architectes data
- BI engineers
- Administrateurs bases de données
- Product managers data
- Consultants data
- Responsables data governance
- Toute personne impliquée dans la qualité et la fiabilité des données

Innov Systems

Programme détaillé

1 / Introduction à la data quality

- Importance de la qualité des données en entreprise
- Risques liés aux données incorrectes ou incomplètes
- Cas d’usage en data engineering et analytics

2 / Fondements du data testing

- Concepts de validation et de contrôle des données
- Différence entre tests unitaires et tests de données
- Data observability et data reliability

3 / Présentation de Great Expectations

- Architecture et principes de fonctionnement
- Définition des “expectations”
- Intégration dans les pipelines data

4 / Présentation de Soda (Soda Core / Soda Cloud)

- Approche déclarative de la data quality
- Monitoring et alerting des données
- Comparaison avec Great Expectations

5 / Mise en place de règles de qualité des données

- Détection des valeurs nulles et incohérences
- Contraintes de schéma et validation métier
- Tests sur données structurées et semi-structurées

Innov Systems

6 / Intégration dans les pipelines data

- Validation dans les workflows ETL/ELT
- Intégration avec Airflow, dbt et orchestrateurs
- Automatisation des contrôles qualité

7 / Data observability et monitoring

- Suivi continu de la qualité des données
- Alertes et tableaux de bord
- Détection des anomalies en temps réel

8 / Gestion des incidents data

- Identification des erreurs de données
- Workflow de correction et de reprocessing
- Collaboration data engineering / business

9 / Gouvernance et conformité des données

- Standards de qualité et data contracts
- Traçabilité et audit des données
- Alignement avec les exigences réglementaires

10 / Cas pratiques et déploiement

- Mise en place d’un pipeline de data quality
- Tests automatisés avec Great Expectations et Soda
- Retours d’expérience en production

Innov 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

07 au 11 Déc. 2026	Distanciel
22 au 26 Fév. 2027	Présentiel - Casablanca
19 au 23 Avr. 2027	Présentiel - Casablanca
14 au 18 Juin 2027	Présentiel - Casablanca
09 au 13 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