Formation Fabric Powered by MYPE
29 rue des Morillons 75015 Paris
Email: contact@mype-consulting.com
Tel: 01 85 53 82 58

FORMATION FABRIC – Parcours complet DP-600 (4 jours)

Microsoft Fabric est un logiciel très puissant, conçu pour mener à bien des projets d'analyse de données, depuis l'ingestion des données jusqu'à l'élaboration de rapports Power BI, en passant par de nombreuses transformations possibles des données. Dans Microsoft Fabric, les outils disponibles sont nombreux : on peut utiliser des pipelines de données, des dataflows, on peut travailler sur les données à l'aide d'un moteur SQL ou d'un moteur Power Query, et on peut créer des rapports Power BI depuis Fabric, entre autres.

Fabric a pour vocation de regrouper sur un même logiciel les tâches de tous les métiers autour de la data : les *data engineers* peuvent ingérer les données depuis diverses sources et les transformer, les *analytics engineers* peuvent créer et gérer des modèles sémantiques, les *data analysts* peuvent créer des rapports Power BI et les *data scientists* peuvent utiliser des blocs-notes Apache Spark pour entraîner des modèles de *machine learning*.

Cette formation s'adresse aux professionnels de la data, qui souhaitent monter en compétences sur cet outil multi-usages qu'est Fabric. La formation vous prépare à l'examen Microsoft DP-600.

Durée : 4 jours / 28 heures

Profils des stagiaires

- Data scientists
- Data analysts
- Data engineers
- Analytics engineers
- Consultants
- Développeurs Python ou SQL
- Autodidactes

Prérequis

Aucun prérequis n'est strictement nécessaire, mais il est conseillé d'avoir des bases en Business Intelligence, pour comprendre la portée des outils que l'on explorera.

Objectifs pédagogiques

- Comprendre les différences entre lakehouses, entrepôts et lacs de données
- Coder en PySpark dans un bloc-notes Apache Spark

- Savoir utiliser un pipeline de données pour ingérer des données
- Savoir utiliser le SQL endpoint et en comprendre les particularités
- Savoir utiliser Power Query dans Fabric
- Savoir effectuer une analyse d'impact des dépendances
- Savoir utiliser DAX Studio, le Tabular Editor et le VertiPaq Analyzer
- Savoir assurer la sécurité des données

Contenu de la formation

Jour 1

1. Concepts utiles
   - Les capacités
   - Le processus ETL (Extract, Transform, Load)
   - Architecture en médaillon
   - Les dimensions lentement changeantes (SCD)
   - Pipelines de déploiement
   - SQL analytics endpoint
2. Les lakehouses
   - Création d'un lakehouse
   - Shortcuts, files, tables
   - Format Parquet et tables Delta
3. Modèles sémantiques
   - Transformation d'un schéma en étoile en modèle sémantique
4. Blocs-notes Apache Spark
   - Création de tables en PySpark
   - Processus ETL dans un bloc-notes Apache Spark
   - Visualisation de données dans un bloc-notes Apache Spark
5. Pipelines de données ('data pipelines')
   - Activité 'Copy data'
   - Dataflows Gen2
   - Orchestration de l'exécution d'un bloc-notes Spark
   - Activités 'Lookup' et 'Get Metadata'
6. Utilisation de Power Query et Power BI
   - Visual Query Editor
7. Analyse d'impact des dépendances en aval
   - A partir de lakehouses, d'entrepôts de données (data warehouses), de flux de données (dataflows) et de modèles sémantiques
8. Optimisation des performances
   - Identification des 'bottlenecks'

- DAX Studio
- Tabular Editor
- Utilisation du cache de requêtes, du VertiPaq Analyzer, du pliage de requêtes, des tables d'agrégation

9. Sécurité et gouvernance
   - Au niveau de l'espace de travail, RLS, CLS, OLS, étiquettes de confidentialité

Jours 2 à 4

Etude interactive de 169 questions du DP-600 : En petit groupe (autour de 5 apprenants par session), l'un de nos formateurs vous laisse quelques minutes de réflexion pour chaque question, vous montre la réponse et vous explique individuellement tous les points de la justification qui peuvent poser problème.

Organisation de la formation

Equipe pédagogique

Nos formateurs en CDI sont des experts sur Fabric.

Ils sont titulaires du DP600.

Moyens pédagogiques et techniques

- Accueil des stagiaires dans une salle dédiée à la formation.
- Documents supports de formation projetés.
- Exposés théoriques.
- Etude de cas concrètes.
- Quiz et/ou exercices en salle.
- Mise à disposition en ligne de documents supports à la suite de la formation.

Dispositif de suivi de l'exécution de l'évaluation des résultats de la formation

- Feuilles de présence.
- Questions orales ou écrites (QCM).
- Mises en situation.
- Formulaires d'évaluation de la formation.