

■ Bases indispensables avant de commencer le Data Engineering

1. ■ Logique de programmation (Python)

Langage recommandé : Python. Apprends à :

- Utiliser les variables, listes, dictionnaires
- Faire des conditions (if/else)
- Créer des boucles (for, while)
- Définir et appeler des fonctions
- Lire et écrire des fichiers

2. ■ Mathématiques de base

Comprends bien :

- Les opérations de base (somme, moyenne, pourcentage)
- Les proportions et les ratios
- La logique booléenne (ET, OU, NON)
- Les statistiques simples (moyenne, écart type)

3. ■■■■ Bases de données relationnelles & SQL

Commence avec les concepts suivants :

- Tables, colonnes, clés primaires/étrangères
- Requêtes : SELECT, WHERE, ORDER BY
- Jointures : INNER JOIN, LEFT JOIN
- GROUP BY et fonctions d'agrégation
- Manipulation des données : INSERT, UPDATE, DELETE
- Création de tables : CREATE TABLE

4. ■ Utilisation du terminal (ligne de commande)

- Naviguer entre les dossiers (cd, ls, mkdir)
- Exécuter un script (python monscript.py)
- Installer un outil avec pip

5. ■ Git et Organisation de projet

- Créer un projet bien structuré (dossiers /data, /scripts, /docs)
- Commandes Git de base : git init, git add, git commit
- Publier sur GitHub

■ Ressources gratuites recommandées :

- Cours Python - OpenClassrooms
- Cours SQL complet - sql.sh
- Apprendre Git de façon interactive
- Khan Academy - Mathématiques