

# Démonstration technique – Backend / Data Engineering

Bienvenue dans cette démonstration technique proposée par Marwan BENLKHIR. Je souhaiterais ici évaluer ma capacité à comprendre une structure de données existante, proposer une structuration plus claire, analyser des corrélations et formuler des hypothèses, proposer une méthode d'enrichissement ou de valorisation des données.

## Objectifs évalués :

- Capacité à structurer et modéliser proprement une donnée brute.
- Clarté dans l'analyse technique
- Qualité des hypothèses et rigueur de la méthode
- Curiosité, prise d'initiative et logique data-driven

## Stack technique imposée :

Je peux choisir d'utiliser Python ou TypeScript (Node.js)

## Données fournies :

Ce fichier contient un extrait des données fournies par Soundcharts, relatif à l'audience et au catalogue musical d'artistes.

[https://drive.google.com/file/d/1x\\_DZWBrvLPzkvLuBzUK1Dltp6xyreV8y/view?usp=sharing](https://drive.google.com/file/d/1x_DZWBrvLPzkvLuBzUK1Dltp6xyreV8y/view?usp=sharing)

## Rendu attendu :

Un dépôt GitHub public contenant :

- Un schéma de BDD (image ou export)
- Un fichier .md ou .pdf expliquant ta proposition technique, ton analyse et ta méthode
- Les scripts réalisés pendant le test technique

## Étape 1 – Modélisation de la donnée

Étudie le fichier fourni pour comprendre l'organisation et la nature des données collectées.

Représente visuellement la structure de ces données à l'aide d'un outil de ton choix (exemples : tldraw, diagrams.net, etc.).

Indiquer les champs que je conserverais, ceux à renommer, regrouper ou normaliser.

Proposer un diagramme relationnel (ERD) prêt à être implémenté sous PostgreSQL, permettant de structurer proprement les données fournies.

## Étape 2 – Analyse et enrichissement

J'ai carte blanche sur cette partie, mais voici quelques orientations à explorer :

### Analyse :

- Proposer une méthode d'analyse de la corrélation entre popularity\_score et d'autres variables disponibles dans le fichier (par exemple : nombre de streams, plateformes, etc.).
- Je peux utiliser des visualisations, des tests statistiques, ou des métriques simples pour appuyer ton raisonnement.

### Enrichissement :

Suggérer des moyens d'enrichir la donnée localement (explique brièvement comment tu procédera).

- Proposer des fonctions ou scripts permettant :

- De calculer des moyennes par artiste, par plateforme, ou par chanson
- De récupérer, sur une plage de temps donnée, l'évolution des streams en valeur absolue (métrique) et en pourcentage

Je peux t'appuyer sur du code Python, TypeScript ou du pseudocode pour illustrer ma logique.

Remarque : Il n'y a pas de mauvaise réponse ici — l'objectif est avant tout d'évaluer ma motivation, ma logique de réflexion et ma capacité à aborder un problème concret de manière structurée.