

stockage clef/valeur et stockage colonne

Semaine B5

Modélisons des données

Soit les données suivantes :

- <https://s3.amazonaws.com/tripdata/index.html>

Télécharger les données les plus récentes. Les modéliser.

Les questions suivantes peuvent être posées :

- Quels sont les trajets partant de telle zone, arrivant dans telle zone ?
- Quels sont les trajets qui partent tel jour à telle heure ?
- Quelles sont les zones les plus actives le lundi ?

Stockage clef/valeur.

1. Modéliser un ou des magasins clef valeur pour pouvoir répondre à ces questions en un nombre minimal de requêtes.
2. Implémenter ces magasins clef/valeurs avec `lmdb`.
3. Charger les données dans iceux.

Stockage orienté colonnes.

Oubliez les données.

En utilisant `lmdb` pour stocker le magasin clef/valeur principal, et du `json` pour les magasins inclus, implémentez une famille de colonne sous forme de magasin de magasin.

Votre classe python devra être utilisable pour récupérer tous les enregistrements correspondants à une clef de partitionnement.

Modélisation des données en orienté colonne

Modéliser différentes familles de colonnes, pour répondre aux questions en un nombre minimal de requête.

Insérer les données, et fournissez le code répondant aux questions en un nombre minimal de requêtes.

Fin

Bon courage.