

Catalogue de données pour le projet d'analyse et recommandation des films dans la plateforme Netflix

Description des données :

Nous avons un ensemble de données provenant d'un script de génération de fausse données.

Elles comprennent :

1. Utilisateurs :

Contenant des informations des utilisateurs à savoir :

- userid : Unique identifiant de l'utilisateur
- name : Nom complet de l'utilisateur
- email : Adresse email
- location : L'emplacement de l'utilisateur
- preferences : Les préférences alimentaires de l'utilisateur
- dob : Date de naissance de l'utilisateur

2. Restaurants :

Contient des informations à propos des restaurants

- restaurantid : Unique identifiant du restaurant
- name : Nom du restaurant
- location : L'emplacement du restaurant
- cuisine : La cuisine offerte par le restaurant
- operationalhours : Heures de travail du restaurant

3. Reviews :

Contient les revues des utilisateurs par rapport aux restaurants

- reviewid : Identifiant unique de la revue
- reviewtext : Le message laissé par l'utilisateur
- rating : Note donnée par l'utilisateur
- reviewdate : Date et temps de la revue

Données utilisées dans le projet :

Toutes les données ci-dessus seront utilisées dans le projet et seront traitées et stockées conformément aux règles en vigueur.

Comment on a utilisé les données :

Implémentation:

- Génération des données depuis le script, cryptage des données personnelles de l'utilisateurs (name, email, location, dob)
- Sauvegarde des données cryptées et l'ensemble dans Cosmos DB
- Déclancher un traitement sur Azure Functions à chaque nouveau document inséré dans Cosmos DB
- Analyse du sentiment de la revue avec Azure AI Services. Retour du score
- Azure Functions ajoute le score dans le document initial et envoi un message d'alerte dans Slack en cas de sentiment négatif.
- Azure envoi le document mis à jour à Logic App via des requetes HTTP
- Logic App met à jour le document dans Cosmos DB.

Date de publication : 25/12/2023

Signé : SEFDINE Nassuf