

Rapport TP Open Food Facts

Ce rapport présente nos choix ainsi que les différentes étapes de la conception d'une partie d'un système décisionnel.

L'objectif est de modéliser puis développer deux "cubes" Mondrian, la base de diffusion adaptée à ceux-ci à partir d'une base d'intégration existante et enfin les tester avec différentes requêtes.

Modélisation

Pour le choix des cubes a été fait en analysant les requêtes demandées.

- Un cube 'Produits' pour la mesure 'Nombre de produits' (= nb. de code-barres différents).
- Un cube 'Versions Produits' pour la mesure 'Nombres de versions' (nb. d'entrées de la table).
- Une dimension 'Date' utilisée pour la 'Date de création' d'un produit ainsi que la 'Date de version' pour la "version" d'un produit.
- Une dimension 'Nutri-score'.
- Une dimension 'Contributeur'. Une dimension 'Type de contributeur' a été envisagée** mais finalement non retenue car manque d'informations dans la base d'intégration.
- Et une dimension 'Catégorie' qui va servir à créer de nouvelles requêtes.

Une dimension 'Produit' a été envisagée mais finalement pas retenue car jugée non nécessaire pour le travail demandé.

Schéma relationnel de la base diffusion

Une table pour chaque dimension et une table de fait pour chaque cube avec les différentes contraintes de clé étrangères. Aucune dimension n'a été dégradée car nous avons jugé que cela n'était pas nécessaire.

Remplissage des tables de faits et dimensionnelles

L'extraction d'informations pour le remplissage des tables dimensionnelles a été faite à partir de SQL (Noeud Extraction depuis une table) car jugée simple et efficace. Concernant la dimension 'Nutriscore', les valeurs NULL (absence de nutri-score) ont été remplacées par la chaîne de caractère "Non renseigné".

L'extraction d'informations de la table de fait pour le cube "Versions Produits" a été faite également avec SQL suivi de plusieurs étapes intermédiaires de recherche pour mettre en relation la table des faits et les tables dimensionnelles associées.

Concernant la table de fait 'Produit', les mêmes étapes que pour la table de fait précédente avec une étape d'agrégation en plus sur les code-barres.

Nous avons remarqué des codes barres similaires avec des dates de création différentes ce qui nous semblait anormal (voir capture d'écran). Nous avons donc fait le choix lors de l'étape d'agrégation de garder seulement le code-barres avec la date de création la plus récente car plus à jour d'après nos observations.

Requêtes MDX

1. Nombre de produits par contributeur (rangées) et par année de création (colonnes)

```
select Hierarchize({[Date de création].[Toutes années],
[Date de création].[Année].Members}) ON COLUMNS,
[Contributeur].Members ON ROWS
from [Produits]
```

2. Nombre de versions de produits par mois et année (lignes) selon la présence ou non du Nutri-score (colonnes).

```
select [Nutri-score].Members ON COLUMNS,
NON EMPTY Hierarchize({[Date de version].[Toutes années],
[Date de version].[Année].Members,
[Date de version].[Mois].Members}) ON ROWS
from [Versions produits]
```

3. Nombre de produits par type de boisson/beverages (lignes) par année de création (colonnes)

```
select Hierarchize({[Date de création].[Toutes années],
[Date de création].[Année].Members}) ON COLUMNS,
Hierarchize({[Catégorie].Beverages,[Catégorie].Beverages.Children}) ON ROWS
from [Produits]
```

4. Nombre de produits par pnns1 (lignes) par années de création (colonnes) enregistrés par le contributeur "carrefour"

```
select NON EMPTY Hierarchize({[Date de création].[Toutes années],
[Date de création].[Année].Members}) ON COLUMNS,
[Catégorie].[Pnns1].Members ON ROWS
from [Produits]
where [Contributeur].carrefour
```

Annexe

Code SQL produits avec le même code-barres mais des dates de création différentes

```
SELECT barcode, DATE(date_creation) AS creation_date
FROM OFF_2_version_produit
WHERE barcode IN (
    SELECT barcode
```

```
FROM OFF_2_version_produit
GROUP BY barcode
HAVING COUNT(DISTINCT DAY(date_creation)) > 1
      OR COUNT(DISTINCT MONTH(date_creation)) > 1
      OR COUNT(DISTINCT YEAR(date_creation)) > 1
);
```