

Documentation

Le script reparateur.py

Introduction

Ce script permet de modifier le fichier de sortie du logiciel ARX afin de le rendre compatible avec les metrics d'évaluation.

Il prend en paramètre un nom de fichier, sans l'extension .csv et qui doit se trouver dans le répertoire csv/, et créer un fichier dans ce même répertoire de la forme : "fichier_rep.csv".

Utilisation

1. Mettez le fichier que vous souhaitez réparer dans le dossier **csv/**.
2. Réglez les paramètre de réparation en modifiant le fichier **reparateur.conf**.
3. Lancez le script **reparateur.py** en précisant le nom du fichier que vous voulez réparer (sans l'extension).
 - Exemple : pour le fichier **anonyme.csv**
 - `$ python reparateur.py anonyme`
4. Attendez quelques ~~longues~~ secondes.
5. Après cela, vous trouverez la version réparée de votre fichier dans le répertoire **csv/** sous la forme suivante : **fichier_rep.csv**.

Le fichier **reparateur.conf**

Ce fichier contient les différentes options de réparation du script. Il se présente de la manière suivante :

```
ordonate:non;
id_user:hash;
date:mediane;
hours:fixe;
id_item:hash;
price:mediane;
qty:moyenne;
```

Les options :

- ordonate
 - oui -> Les lignes du fichier à réparer **commencent** par un numéro qui permet de les garder dans l'ordre après le passage d'ARX.
 - non -> Les lignes du fichier à réparer **ne commencent pas** par ce numéro.
- id_user
 - clair -> L'id des client **n'est pas hashé**.
 - hash -> L'id des client **est hashé**.
- date
 - clair -> La date **n'est pas changée**.
 - fixe -> La date est **fixée à 15**.
 - mediane -> La date est changée au profit de la **date médiane du mois**.
- hours
 - clair -> L'heure **n'est pas changée**.
 - fixe -> L'heure est **fixée à 13:37**.
- id_item
 - clair -> L'id des produits **n'est pas hashé**.
 - hash -> L'id des produits **est hashé**.
- price
 - moyenne -> Le prix est calculé comme étant la **moyenne de l'intervale de prix** donné par ARX.
 - mediane -> Le prix est changée au profit du **prix médian de l'intervale**.
- qty
 - moyenne -> La quantité est calculé comme étant la **moyenne de l'intervale de quantité** donné par ARX.