# Aide-mémoire Openrefine 3.4 : principales commandes

Encadré bleu : commandes accessibles depuis la page d’accueil ; Encadré blanc : commandes accessibles depuis un projet.

# Exporter

> Exporter

## Un projet Openrefine

> Exporter le projet

## Les données du projet

…export complet

> Valeurs séparées par des tabulations

> Valeurs séparées par des virgules

> Table HTML

> Excel / Excel 2007

> Opendocument (Libreoffice)

> Data package

…export personnalisé

> Export tabulaire

> Export SQL

> Modélisation

> QuickStatements (Wikidata)

> Publication sur Wikidata

# Importer

## Un projet Openrefine

> Importer un projet

## Des données

(CSV, TSV et variantes, Excel, Libreoffice, Google Data, texte brut, XML et JSON avec limitations, RDF, Wikitext)

> Créer un projet

… depuis un ou plusieurs fichiers locaux :

> Cet ordinateur

…par saisie directe :

> Presse-papier

… depuis un Google Doc

> Google Data

… depuis un ou plusieurs fichiers en ligne :

> Adresses web

… depuis un « data package » en ligne ou local

> Data packages

… depuis une base de données

> Database

# Gérer les projets

## Ouvrir un projet

> Ouvrir un projet

Supprimer un projet

Modifier le nom, ajouter des sujets, des étiquettes, une description

# Explorer

## Filtres de colonnes

> Filtrer le texte (Regex possible)

## Facettes de colonnes

> Facette

> Textuelle

> Numérique

> Chronologique

> Textuelle personnalisée (formules)

> Numérique personnalisée

> Personnalisée

> par mots

> par doublons (doublons calculés sur tout le jeu de données, sans prise en compte des filtres et facettes activés)

> par longueur de texte

> par erreur

> par valeur nulle

> par chaîne vide

> par valeur vide

## Facettes globales (col. Toutes)

> Facette par étoile

> Facette par marque

> Facette par valeur vide

> Valeurs non vides par colonnes

> Entrées non vides par colonnes

# Mode lignes et mode entrées

Une entrée regroupe une ou plusieurs lignes en fonction du contenu de leur 1re colonne.

Modifie le fonctionnement des facettes et des fonctions de scissions et de vidage/remplissage

Pour prendre en compte les entrées : Voir en … entrées

Pour prendre en compte les lignes : Voir en … lignes

# Personnaliser l’affichage

## Afficher les valeurs *null (*col Toutes)

## > Afficher/masquer les valeurs nulles dans les cellules

## Masquer ou afficher des colonnes

> Aperçu

# Agir sur les lignes

## Trier

> Trier sur une ou plusieurs colonnes

> Retrier de façon permanente ou Supprimer le tri

## Supprimer les lignes affichées (col. Toutes )

> Editer les lignes

> Supprimer les lignes

## Signaler des lignes (col. Toutes )

> Editer les lignes

> étoiler/ désétoiler

> marquer / démarquer

## Scinder des cellules sur plusieurs lignes et vice versa

> Editer les cellules

> Diviser les cellules multivaluées

Possible de personnaliser le séparateur (regex possible) ; diviser par longueur de champ

> Joindre les cellules multivaluées

# Gérer les extensions

> Ouvrir un projet

> Parcourir le dossier de l’espace de travail

Déposer le dossier contenant l’extension dans le dossier extension.

# Gérer les options

> Langue

> Préférences

|  |
| --- |
| ui.browsing.listFacet.limit |

# Agir sur le contenu des cellules

## Transformations courantes

> Editer les cellules > Transformations courantes

…supprimer les espaces superflus

> Supprimer les espaces de début et de fin

> Rassembler les espaces consécutifs

…harmoniser la casse

> En initiale majuscules

> En majuscules

> En minuscules

…convertir

> date et texte en nombre

> texte en dates

> nombre et date en texte

…effacer

> en valeurs nulles

> en chaînes vides

## Remplacer

> Editer les cellules

> Remplacer (Regex possible)

## Vider les valeurs répétées / recopier les valeurs dans les cellules vides

> Editer les cellules

> Vider les valeurs répétées…

> Recopier les valeurs répétées…

Prend en compte les limites des entrées si le mode entrée est activé

# Agir sur les colonnes

## Renommer

> Editer la colonne

> Renommer cette colonne

## Supprimer

> Editer la colonne

> Supprimer cette colonne

## Déplacer

> Editer la colonne

> Déplacer la colonne (en 1er, en dernier, à droite, à gauche)

## Supprimer ou déplacer plusieurs colonnes (Col. Toutes )

> Retrier/supprimer les colonnes

## Scinder une colonne en plusieurs

> Editer la colonne

> Diviser en plusieurs colonnes

Possible de personnaliser le séparateur (regex possible) ; diviser par longueur de champ ; limiter le nombre de divisions ; conserver la colonne d’origine ; désactiver la détection du type de données

## Joindre plusieurs colonnes

> Editer la colonne

> Joindre des colonnes

## Ajouter une nouvelle colonne basée sur une colonne existante

> Editer la colonne

> Ajouter une colonne en fonction de cette colonne

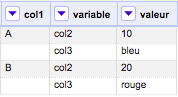
## Ajouter une nouvelle colonne basée sur une requête HTTP

> Editer la colonne

> Ajouter une colonne en moissonnant une URL

Options utiles : délai ; conserver les erreurs ; mettre les réponses en cache (rapidité) ; entête HTTP personnalisé si le serveur permet la négociation de contenu (ex : application/json ou text/xml)

Pas de prise en compte des redirections y compris https -> http.



# Appliquer une transformation personnalisée

## Appliquer une transformation personnalisée à une colonne

> Transformer

## Appliquer une transformation personnalisée à plusieurs colonne

Col. Toutes >

> Transformer

## Langage : GREL, Python (2,7), Clojure

## Saisie de l’expression

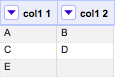
## Prévisualisation du résultat

## Historique des dernières formules

## Formules favorites (étoilées)

## Aide (limitée mais utile)

## Options en cas d’erreur : conserver l’original, vider la cellule, afficher l’erreur



# Transposer lignes et colonnes

> Transposer

## Passer d’un format « large » à un format « long »

> Transposer les cellules de plusieurs colonnes en lignes

## Transposer les cellules en colonnes séparées

> Transposer les cellules en colonnes séparées

## Passer d’un format « long » à un format « large »

> Convertir en liste les colonnes de clé/valeur

# Réconcilier depuis une source de données extérieure

## Réconcilier depuis une source de données extérieure

> Réconcilier > Démarrer la réconciliation

## Choisir le service ou en installer un nouveau (par défaut : Wikidata)

## Valider le type d’entité cherchée

## Pour certains services possibilité d’utiliser d’autres colonnes pour préciser la requête

## Dans la grille, valider les correspondances exactes