

Transformer un document en tableau

Fixer les colonnes, conserver les lignes sources et contrôler les cellules ajoutées

Module DO4 - FAIRE - 6 min - version 1.1

À vérifier à l'écran avant de suivre ce module

- Export tableur : [À VÉRIFIER : un tableau produit peut-il être exporté vers un tableur, dans quel format, et dans quelle offre ?]

Volatilité V3 - ce point reste ouvert jusqu'à observation réelle de l'outil visé.

Prix, limites et noms de menus changent sans prévenir. Si ce que tu vois à l'écran ne correspond pas, c'est l'écran qui a raison.

Transformer du texte en tableau paraît mécanique, mais un tableau propre peut masquer trois défauts : un élément oublié, deux éléments fusionnés, ou une cellule remplie sans matière source.

Le contrôle le plus robuste n'est pas le nombre annoncé par le modèle. C'est la correspondance entre chaque ligne du tableau et un élément identifié dans la source.

Ce que tu sauras faire en sortant

- Définir les colonnes avant la conversion.
- Numéroté ou identifier les éléments sources avant d'envoyer le texte.
- Exiger un identifiant source par ligne du tableau.
- Laisser vide ce que la source ne renseigne pas.
- Rechercher omission, fusion et ajout sans provenance.

L'essentiel, en trente secondes

Avant la conversion, transforme ta matière en éléments repérables :

`text S01 | mardi | remplacement de 3 tuiles S02 | mardi | dépannage en fin de journée S03 | mercredi | reprise zinguerie 4 m `

Puis impose les colonnes et garde source_id dans le résultat :

| source_id | date | intervention | quantité | |---|---|---|---| | S01 | mardi | remplacement de tuiles | 3 | | S02 | mardi | dépannage | | | S03 | mercredi | reprise zinguerie | 4 m |

Une cellule vide signifie seulement « cette information n'est pas présente dans l'élément source ». Elle ne signifie pas que l'information n'existe nulle part ailleurs.

Le geste

`text Transforme les éléments ci-dessous en tableau avec exactement ces colonnes : source_id | [COLONNE 1] | [COLONNE 2] | [COLONNE 3]

Règles : - une ligne de tableau par source_id ; - recopie source_id sans le modifier ; - ne fusionne jamais deux source_id ; - laisse une cellule vide si l'élément source ne fournit pas l'information ; - n'ajoute aucune valeur déduite ; - après le tableau, liste les source_id absents ou dupliqués.

Source : [ÉLÉMENTS NUMÉROTÉS] `

Le diagnostic absents ou dupliqués produit par le modèle est utile comme alerte, mais tu fais toi-même la vérification finale en comparant la liste des source_id source et résultat.

Micro-défi

Source : S01, S02, S03, S04.

Tableau : S01, S02, S02, S04.

Quelles anomalies sont certaines ?

Attendu : S03 est absent et S02 est dupliqué. Cela ne dit pas encore pourquoi : omission, fusion ou erreur de copie. Le contrôle suivant consiste à revenir aux deux éléments concernés.

Trois contrôles séparés

1. Exhaustivité des éléments

Compare la liste des identifiants source à la colonne source_id du tableau.

- identifiant source manquant ligne à reprendre ;
- identifiant dupliqué tableau à examiner ;
- identifiant inconnu ajout sans provenance.

2. Provenance des cellules

Choisis les cellules qui portent des nombres, dates, noms ou valeurs importantes. Pour chacune : peux-tu montrer le fragment source qui la justifie ?

3. Cases vides

Une case vide est acceptable si l'information n'existe pas dans cet élément source. Ne demande pas au modèle de « compléter raisonnablement » pour rendre le tableau plus homogène.

Aucun de ces trois contrôles n'est remplacé par le compte de lignes annoncé par le modèle : ce serait demander au même système de certifier sa propre transformation.

Le cas où ça se passe mal

Cas reconstitué.

S02 dit seulement « dépannage en fin de journée ». Le tableau lui attribue 2 h parce que la colonne durée est renseignée pour les autres lignes.

Correction : la cellule durée de S02 reste vide. Le tableau n'a pas pour rôle de rendre la source plus complète qu'elle ne l'est.

Calculs et tris

Ce module porte sur la structuration, pas sur le calcul.

- Un total doit être calculé et contrôlé dans l'outil approprié.
- Un classement ou une priorité est une opération supplémentaire, avec ses propres critères ; ne la laisse pas apparaître implicitement pendant la conversion.
- Un export natif peut être préférable à une transformation par IA lorsqu'il existe, mais le module ne prétend pas qu'un export est « toujours exact » : contrôle aussi la source et le fichier obtenu selon

ton usage.

Données réelles

Si les notes nomment des clients, salariés ou autres tiers, le tri F3 précède tout transfert à un service tiers. Pour apprendre le geste, utilise des éléments fictifs et des identifiants S01, S02...

Ce que ce module ne règle pas

- Il ne calcule pas ou ne valide pas les totaux.
- Il ne garantit pas qu'un tableau ayant le bon nombre de lignes soit exact.
- Il ne prouve pas qu'une cellule vide correspond à une absence dans le document complet.
- Il ne choisit pas les critères d'un tri ou d'un classement.
- Il ne garantit aucune capacité d'export produit : le marqueur reste ouvert.

Quand tu t'arrêtes

- Un source_id manque, est dupliqué ou apparaît sans source.
- Une cellule importante n'a pas de fragment source identifiable.
- Le texte réel n'a pas passé F3.
- Tu es sur le point d'utiliser un total ou un classement apparu pendant la conversion sans contrôle séparé.

À toi

Crée dix lignes fictives S01 à S10, dont :

- deux lignes sans quantité ;
- deux lignes le même jour ;
- une ligne très courte.
- Choisis trois colonnes métier + source_id.
- Lance la conversion.
- Compare les dix identifiants source au tableau.
- Vérifie la provenance de toutes les cellules numériques.
- Confirme que les informations manquantes restent vides.

C'est réussi si tu peux reconstruire, pour chaque ligne du tableau, l'élément source qui l'a produite.

Vérifie ton geste

- Le nombre de lignes est correct, mais S05 manque et S04 est dupliqué. Tableau incorrect malgré le bon total.
- Une quantité apparaît alors qu'elle n'est pas dans S07. Valeur sans provenance ; retirer.
- Deux interventions du même jour sont distinctes. Elles gardent deux source_id, donc deux lignes.
- Une cellule est vide. Acceptable si Sxx ne fournit pas l'information ; aucune conclusion d'absence globale.
- Le texte contient des données de tiers. F3 avant transfert.

Seuil de réussite : 4 sur 5, questions 1 et 2 obligatoires.

Selon ce que tu obtiens

- Tous les IDs correspondent passe au contrôle des cellules importantes.
- ID manquant/dupliqué reviens à la source ; ne corrige pas seulement le nombre total.
- Beaucoup de cellules inventées renforce la règle laisser vide et réduit les colonnes aux informations réellement présentes.
- Bloqué commence avec quatre éléments fictifs et deux colonnes.

Fiche mémo

- Numérote les éléments source.
- Fixe les colonnes avant la conversion.
- Garde source_id dans le tableau.
- Compare les IDs, pas seulement le nombre de lignes.
- Cellule importante = fragment source identifiable.
- Manque = case vide, jamais valeur déduite.

Ce que tu viens de pratiquer

- Tu as rendu chaque ligne traçable jusqu'à sa source.
- Tu as séparé exhaustivité, provenance des cellules et données manquantes.
- Tu ne demandes plus au modèle de certifier lui-même son propre compte.

Sources et statut : docs/bloc-documents-PROVENANCE.md`. DO4 reste brouillon ; le marqueur d'export reste ouvert.