SAKIRA

TUTORIEL 32

Nettoyer un fichier Excel (doublons, formats de dates)

© SAKIRA

© SAKIRA. Reproduction et diffusion interdites sans accord écrit.
SAKIRA SAS · 6 rue d'Armaille, 75017 Paris · SIREN 108 493 933

TUTORIEL 32 · NIVEAU DÉBUTANT

Nettoyer un fichier Excel (doublons, formats de dates)

Un fichier alimenté par plusieurs personnes finit avec des clients en double et des dates écrites de cinq façons. Une IA donne les formules, c'est vous qui décidez des lignes qui disparaissent.

La règle à respecter

Travaillez sur une copie et comptez les lignes avant et après. Les erreurs habituelles : supprimer les doublons sans les avoir regardés (deux lignes « Entreprise A », à Lyon et à Lille, sont peut-être deux clients), croire qu'Excel garde la ligne la plus récente alors qu'il garde la première rencontrée, convertir des dates en laissant passer « 07/04 » qui se lit le 7 avril ou le 4 juillet, et recopier des formules sans les avoir contrôlées sur trois lignes. Un nombre de lignes qui ne tombe pas juste veut dire qu'on s'arrête et qu'on reprend la copie. Ne donnez à l'assistant aucune ligne de clients : il n'a besoin que des titres de colonnes et de quelques exemples de dates.

AVANT DE COMMENCER

Il vous faut

Votre fichier Excel, et un assistant d'IA (ChatGPT ou un autre). Dans notre exemple : 480 lignes de clients, colonne A le nom, B la ville, C la date de dernière commande, D le montant en euros hors taxes. Les titres sont en ligne 1, les données de la ligne 2 à la ligne 481.

Résultat

Un fichier sans doublons, avec de vraies dates, et deux nombres de lignes (avant et après) qui s'expliquent.

01 Faire la copie et compter

Dupliquez le fichier dans l'explorateur de fichiers ou dans le Finder, ajoutez « copie » à son nom et travaillez uniquement dans cette copie. Dans une cellule libre à droite des données, par exemple J2, tapez `=NBVAL(A2:A481)` : elle compte les noms remplis. Écrivez le résultat sur papier ou dans un message à vous-même : 480 dans notre exemple. Si le résultat n'est pas celui que vous attendiez, il y a des lignes vides ou des lignes en plus, à comprendre avant d'aller plus loin.

GARDE-FOU

Cette formule compte des cellules remplies, pas des clients. Un client dont le nom est vide n'est pas compté. Comparez avec la dernière ligne remplie de la colonne A, en descendant jusqu'au bout de la colonne : le résultat doit être le numéro de cette ligne moins 1.

© SAKIRA. Reproduction et diffusion interdites sans accord écrit. · sakira.fr
SAKIRA SAS · 6 rue d'Armaille, 75017 Paris · SIREN 108 493 933

02 Débutant : les noms et les doublons

Le prompt ne contient aucune donnée de client : seulement la description des colonnes et un exemple. Adaptez les numéros de ligne et les lettres de colonne à votre fichier. Les formules sont écrites pour la version française d'Excel, avec des points-virgules : dans une version anglaise, les noms de fonctions changent et les points-virgules deviennent des virgules.

À COPIER

```
J'ai un fichier Excel de 480 lignes avec la liste des clients d'une entreprise de négoce. Colonnes : A Nom du client, B Ville, C Date de dernière commande, D Montant de la dernière commande en euros hors taxes. La ligne 1 contient les titres, les données commencent en ligne 2 et finissent en ligne 481.

Certains clients apparaissent deux fois, avec des majuscules ou des espaces en trop. Exemple : « Entreprise A » en ligne 12 et « ENTREPRISE A  » (deux espaces à la fin) en ligne 305.

Donne-moi deux formules pour la version française d'Excel, avec des points-virgules :
1. Une formule à mettre en E2 qui recopie le nom de A2 sans espaces en trop et avec une seule façon d'écrire les majuscules.
2. Une formule à mettre en F2 qui compte combien de fois le nom de E2 apparaît dans la colonne E.

Pour chaque formule : la cellule où la mettre, le format de cellule, comment la recopier jusqu'à la ligne 481, et ce que je dois voir. Dis aussi ce que la formule ne rattrape pas, par exemple deux orthographes différentes du même nom. Ne supprime rien : je supprimerai moi-même, après avoir regardé. Aucune donnée de client ne t'est donnée, écris seulement les formules et les gestes. Ne cite aucun nom de menu : décris le geste, avec le raccourci pour Windows et pour Mac s'il y en a un. Réponds-moi en me vouvoyant. Écris en texte simple, sans astérisques ni dièses, chaque formule sur sa propre ligne. Réponds en moins de 250 mots.
```

© SAKIRA. Reproduction et diffusion interdites sans accord écrit. · sakira.fr
SAKIRA SAS · 6 rue d'Armaille, 75017 Paris · SIREN 108 493 933

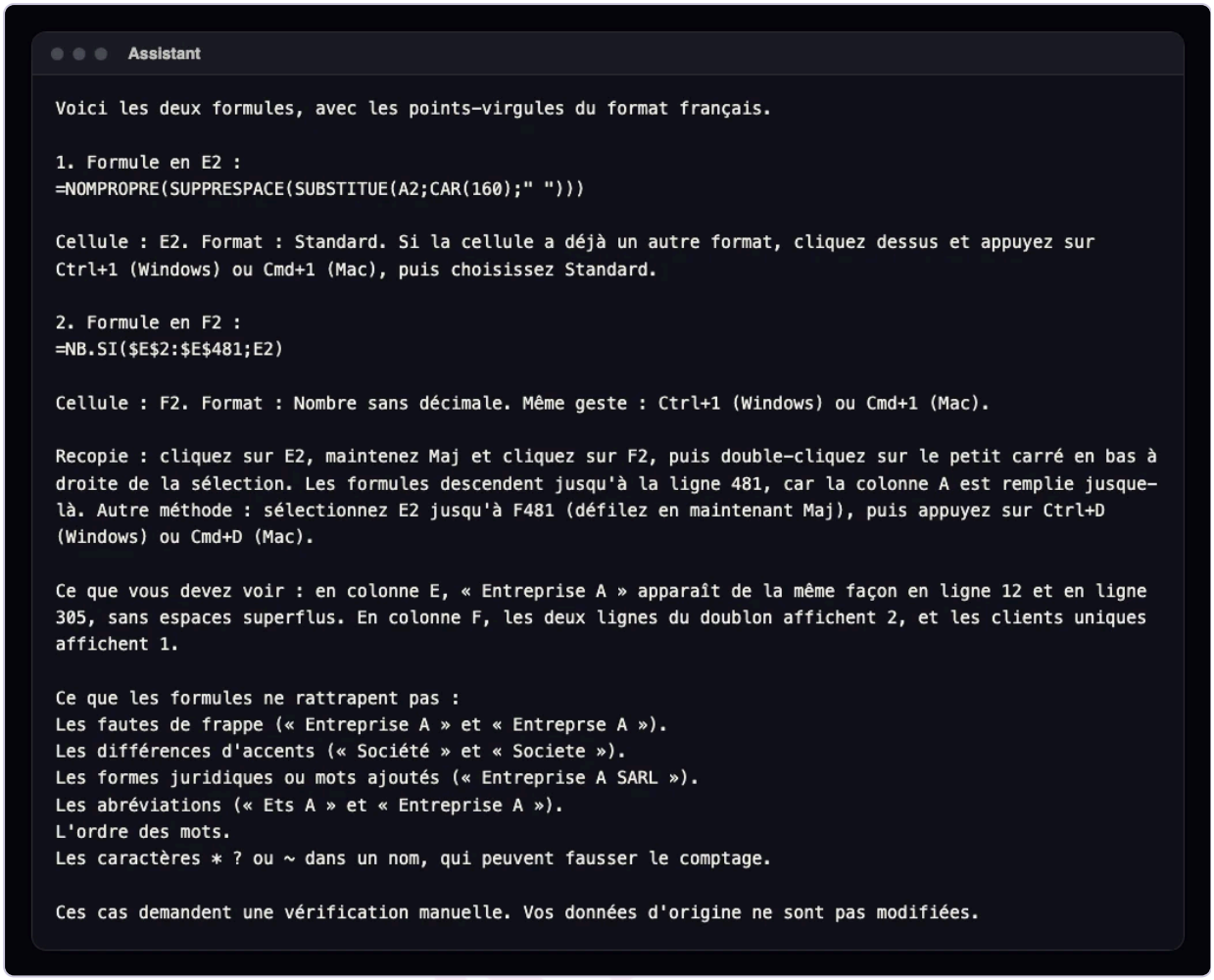


Figure 1.Réponse sur l'exemple : une formule pour nettoyer le nom, une pour compter les apparitions, et la liste de ce que ces formules ne rattrapent pas.

Mettez les formules en E2 et F2, contrôlez-les sur trois lignes dont celles de votre exemple, puis recopiez jusqu'à la ligne 481. Dans la colonne F, un nombre supérieur à 1 signale un nom qui revient. Filtrez la colonne F sur « supérieur à 1 » et lisez les lignes une par une.

> **GARDE-FOU**
> Un nom qui revient n'est pas forcément un doublon. Deux lignes « Entreprise A », l'une à Lyon et l'autre à Lille, peuvent être deux clients, ou deux sites du même. Regardez la ville, le montant, la date. Les accents, les abréviations (« Ets », « Établissements ») et les fautes de frappe échappent à ces formules : repérez-les en triant la colonne E.

03 Les dates

Même démarche, même fichier. Les cinq exemples du prompt sont ceux de votre colonne C : remplacez-les par cinq lignes que vous y avez vues.

© SAKIRA. Reproduction et diffusion interdites sans accord écrit. · sakira.fr
SAKIRA SAS · 6 rue d'Armaille, 75017 Paris · SIREN 108 493 933

À COPIER

```
Dans le même fichier de 480 lignes, la colonne C contient des dates que Excel traite
comme du texte. Voici cinq exemples de ce qu'on y trouve :
12/03/2025
2025-03-14
3 mars 2025
21/04/2025 (avec un espace devant)
07/04/2025

Donne-moi la façon la plus courte de les convertir en vraies dates dans la colonne G,
pour la version française d'Excel, avec des points-virgules : une formule à mettre en
G2, le format de cellule à choisir, comment la recopier jusqu'à la ligne 481.

Donne aussi une formule à mettre en H2 qui affiche « à vérifier » quand la conversion
a échoué, et une formule, dans une cellule libre, qui compte combien de lignes sont à
vérifier.

Dis-moi quelles dates peuvent être lues de deux façons (jour et mois inversés) et
comment je peux m'en rendre compte. Ne modifie pas la colonne C. Aucune donnée ne
t'est donnée, écris seulement les formules et les gestes. Ne cite aucun nom de menu :
décris le geste, avec le raccourci pour Windows et pour Mac s'il y en a un. Réponds-
moi en me vouvoyant. Écris en texte simple, sans astérisques ni dièses, chaque formule
sur sa propre ligne. Réponds en moins de 250 mots.
```

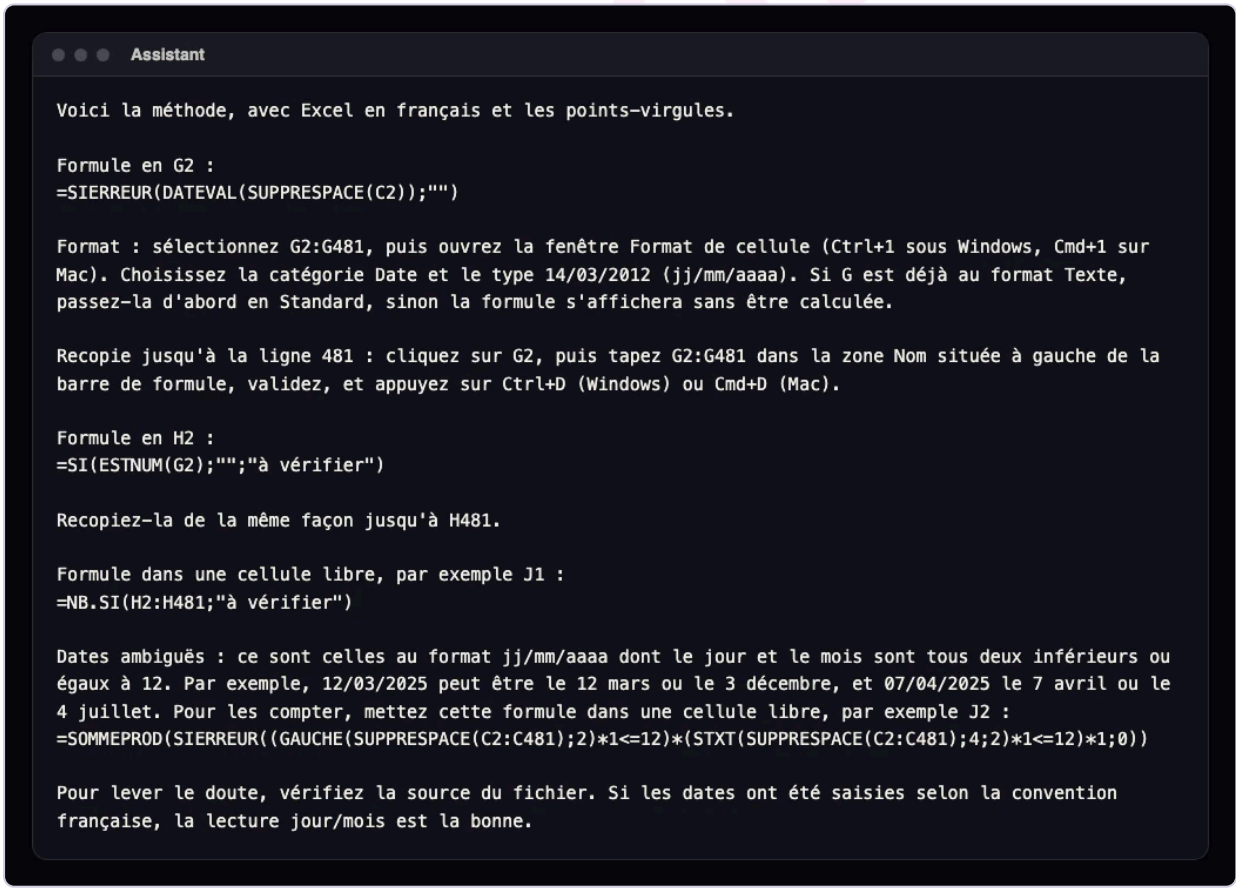


Figure 2. Réponse sur l'exemple : la formule de conversion en G, le signalement des échecs en H, le compte des lignes à vérifier, et les dates lisibles de deux façons avec une formule qui les compte.

© SAKIRA. Reproduction et diffusion interdites sans accord écrit. · sakira.fr
SAKIRA SAS · 6 rue d'Armaille, 75017 Paris · SIREN 108 493 933

> **GARDE-FOU**
> Les formules marchent pour les formats que vous avez cités, pas pour les autres. Les lignes marquées « à vérifier » se corrigent à la main. Pour les dates comme « 07/04/2025 », Excel suit les réglages de votre ordinateur, qui sont jour puis mois en France : si le fichier vient d'un logiciel américain, elles sont peut-être inversées. Si votre Excel est en anglais, redemandez la formule avec les noms de fonctions anglais et des virgules.

04 Supprimer les doublons sur une feuille à part

Quand les lignes « à vérifier » sont corrigées (à la main, dans la colonne G), copiez les colonnes utiles vers une nouvelle feuille avec un collage de valeurs, dans cet ordre : nom nettoyé en A, ville en B, date convertie en C, montant en D : les formules deviennent des nombres et du texte fixes. Sur cette feuille, triez par date, de la plus récente à la plus ancienne : l'outil « Supprimer les doublons » de l'onglet Données garde la première ligne rencontrée, donc la plus récente. Cochez le nom et la ville seulement. Excel annonce combien de doublons il a supprimés : notez ce nombre.

05 Recompter et comparer

Sur la feuille propre, comptez à nouveau les lignes avec `=NBVAL(A2:A481)` : la fin de la plage est maintenant vide, et la formule ne compte que les cellules remplies. Le résultat doit être le nombre de départ moins le nombre de doublons annoncé par Excel : 480 moins 37 doublons donnent 443, par exemple. Vérifiez aussi que la colonne des dates ne contient que des dates : `=NB(C2:C481)`, qui compte les cellules numériques, doit donner le même nombre de lignes. Si les nombres ne se recoupent pas, ne corrigez pas à la main : repartez de la copie.

Qui fait quoi

L'IA fait	Vous faites
Donne les formules de nettoyage et de conversion	Faites la copie et comptez les lignes avant
Dit quels cas ses formules ne rattrapent pas	Regardez chaque doublon avant de le supprimer
Signale les dates qui peuvent être lues de deux façons	Corrigez à la main les lignes « à vérifier »
Ne voit aucune ligne de votre fichier	Comptez après et comparez avec le départ

Ce que ce guide ne fait pas

Il ne rapproche pas deux orthographes différentes d'un même client, ni les noms écrits avec des abréviations : ces cas demandent une lecture. Il ne traite pas les cellules fusionnées ni les fichiers de plusieurs milliers de lignes avec plusieurs onglets. Les formules sont écrites pour la version française d'Excel. Si le fichier contient des

© SAKIRA. Reproduction et diffusion interdites sans accord écrit. · sakira.fr
SAKIRA SAS · 6 rue d'Armaille, 75017 Paris · SIREN 108 493 933

données personnelles, vérifiez avec votre délégué à la protection des données avant de le partager hors de l'entreprise.

© SAKIRA

© SAKIRA. Reproduction et diffusion interdites sans accord écrit. · sakira.fr
SAKIRA SAS · 6 rue d'Armaille, 75017 Paris · SIREN 108 493 933