



Rapport de traitement des données

Titre du projet	Collectez des données en respectant les normes RGPD
------------------------	---

Version	Auteur	Description	Date
V1	Antoine Vatin	Rapport	17.01.2025



Introduction

Contexte

Environnement réglementaire

Le RGPD impose des normes strictes pour la collecte, le traitement et la conservation des données personnelles. Chaque traitement doit être fondé sur une finalité légitime et respecter des principes tels que la minimisation des données, la transparence, et la sécurité.

Situation actuelle

L'entreprise gère un volume important de données clients à travers son CRM, ce qui implique des responsabilités majeures pour garantir la confidentialité et éviter tout risque de violation des règles.

Objectifs du projet

1. **Respect de la conformité** : garantir que l'ensemble des données soient gérées dans le respect des exigences RGPD.
2. **Optimisation des processus** : établir des flux de traitement des données sécurisés et transparents.
3. **Valorisation des données** : permettre une exploitation analytique des données tout en respectant leur anonymisation.



Contenu du rapport :

Extraction des données

1. Objectif de la requête

L'objectif de cette requête est d'extraire des informations spécifiques depuis la table « **base_client** » pour les clients :

- Dont la **date de demande** est dans l'année **2022**.
- Dont l'**état du dossier** est marqué comme « **complet** ».

Les colonnes sélectionnées incluent des informations pertinentes pour l'analyse des profils clients, comme les caractéristiques personnelles, les informations sur leur véhicule et les données liées à leur demande d'assurance.

2. Exécuter la requête

1. Accéder à l'éditeur de requêtes :

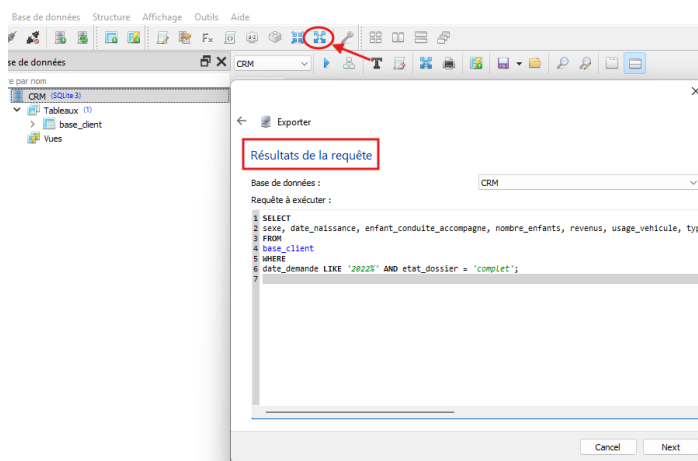
- Dans SQLite Studio, ouvrez l'onglet **Éditeur de requêtes**.

```
SELECT  
sexe, date_naissance, enfant_conduite_accompagne, nombre_enfants, revenus, usage_vehicule,  
type_vehicule, est_rouge, points_perdus, age_vehicule, type_conduite, date_demande, etat_dossier,  
formule, tarif_devis  
FROM  
base_client  
WHERE  
date_demande LIKE '2022%' AND etat_dossier = 'complet';
```

2. Exporter les données

1. Si vous souhaitez analyser les données dans un autre outil (Excel, Power BI, etc.) :

- Cliquez sur **Exporter**.
- Choisissez un format approprié (CSV recommandé).
- Spécifiez un nom et un emplacement pour le fichier exporté.



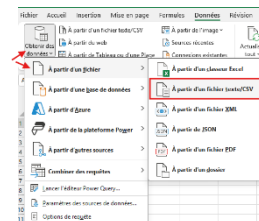
Importation et traitement dans Power Query

Étape 1 : Importation du fichier dans Power Query

1. Ouvrir Excel

2. Accéder à Power Query :

- Aller dans l'onglet **Données**.
- Cliquer sur **Obtenir des données** > **À partir d'un fichier** > **À partir d'un fichier texte/CSV**.



3. Sélectionner le fichier : Naviguez jusqu'à l'emplacement de votre fichier et sélectionnez « **Vatin_Antoine_2_donnees_012025.csv** ».

4. Prévisualisation des données : Une fenêtre s'ouvre avec un aperçu des données. Vérifiez que le délimiteur (par exemple, une virgule) est correctement détecté. Cliquez sur **Transformer les données**.

Vatin_Antoine_2_donnees_012025.csv

Origine du fichier

1252: Europe de l'Ouest (Windows)

Délimiteur

Virgule

Détection du type de données

Selon les 200 premières lignes

sexe	date_naissance	enfant_conduite_accompagne	nombre_enfants	revenus	usage_vehicule	type_vehicule	est_rouge	point
F	25/09/1999	0	1	125301	Commercial	Sports Car	no	
F	28/01/1932	1	1	50815	Private	Minivan	no	
F	15/09/1978	0	2	21204	Private	Pickup	no	
F	17/02/2021	0	0	0	Private	z_SUV	no	
M	14/06/1956	0	0	null	Commercial	Panel Truck	no	
F	11/04/1992	0	1	79606	Private	z_SUV	no	
F	18/03/1941	2	2	130540	Commercial	Minivan	no	
M	06/03/1956	0	0	51278	Private	z_SUV	no	
M	18/02/1956	0	0	115044	Private	Minivan	no	
F	07/03/1974	0	0	null	Commercial	Panel Truck	yes	
F	13/02/1967	1	2	24060	Private	Minivan	no	
M	14/05/2017	0	4	20822	Private	z_SUV	no	
F	20/12/1950	0	0	74426	Private	z_SUV	no	
M	26/07/2017	0	0	0	Commercial	Van	no	
M	07/10/1945	0	1	38042	Private	Minivan	no	
M	12/02/1938	0	0	45787	Private	Minivan	no	
F	21/03/1973	0	0	108152	Private	z_SUV	no	
F	01/02/1936	0	0	53244	Commercial	Van	yes	
F	04/04/1955	0	0	76954	Private	Minivan	yes	
F	23/05/1940	0	0	43290	Private	z_SUV	no	

Charger

Transformer les données

Annuler

Charger Transformer les données Annuler

Étape 2 : Transformer les données

Ajouter une colonne calculée pour l'âge

1. Accéder à l'option de calcul :

- Dans le ruban, cliquez sur **Ajouter une colonne > Colonne personnalisée**.

2. Nommer la colonne :

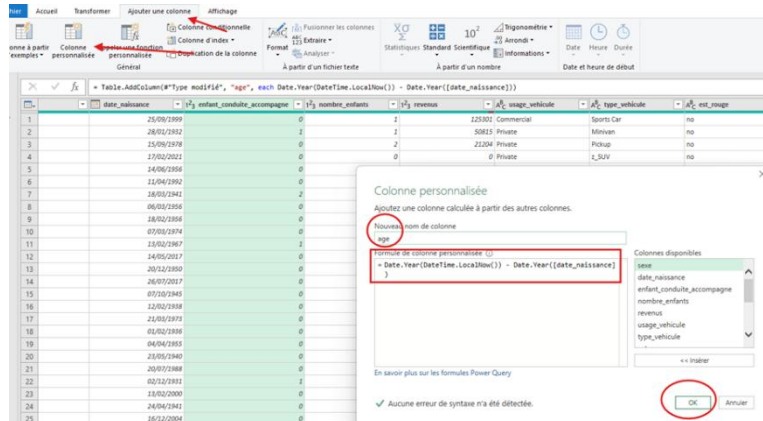
- Donnez-lui un nom pertinent, comme « **age** ».

3. Entrer la formule :

- Utilisez la formule suivante pour calculer l'âge :

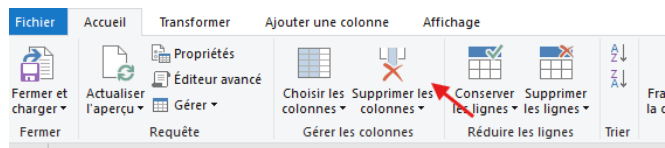
Date.Year(DateTime.LocalNow()) - Date.Year([date_naissance])

Cela soustrait l'année de naissance de l'année en cours.



Supprimer la colonne originale

- Si la date de naissance n'est plus nécessaire :
 - Sélectionnez la colonne « date_naissance », cliquez avec le bouton droit, puis choisissez **Supprimer les colonnes**.



Requêtes	Table	Colonne	Valeur
1	F	date_naissance	25/09/1999
2	F	date_naissance	28/01/1932
3	F	date_naissance	15/09/1978
4	F	date_naissance	17/02/2021
5	M	date_naissance	14/06/1956
6	F	date_naissance	11/04/1992

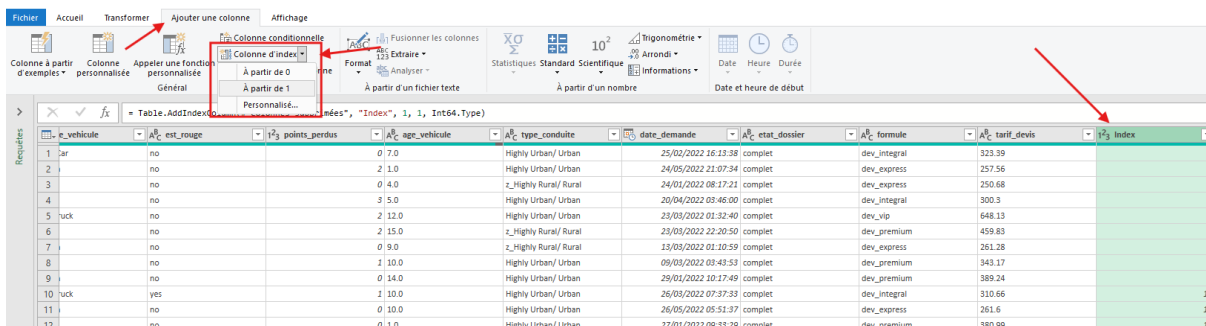
Ajouter une colonne d'index

1. Accéder à l'option d'ajout de colonne d'index :

- Dans Power Query, allez dans le ruban **Ajouter une colonne**.
- Cliquez sur **Ajouter une colonne d'index** :
 - Choisissez **À partir de 1** ou **À partir de 0**, selon vos besoins.

2. Nommer la colonne :

- Par défaut, la colonne sera nommée « Index ».



The screenshot shows the Power Query interface. The ribbon 'Ajouter une colonne' is active, with 'Colonne d'index' selected and 'À partir de 1' chosen. Below, a table of 12 rows is shown with a new 'Index' column added at the end, containing values from 1 to 12.

	id_vehicule	est_rouge	points_perdus	age_vehicule	type_conduite	date_demande	etat_dossier	formule	tarif_devis	Index
1	ar	no	0	7.0	Highly Urban/ Urban	25/03/2022 16:13:58	complet	dev_integral	323.39	1
2		no	2	1.0	Highly Urban/ Urban	24/05/2022 21:07:34	complet	dev_express	257.56	2
3		no	0	4.0	z_Highly Rural/ Rural	24/01/2022 08:17:21	complet	dev_express	250.68	3
4		no	3	5.0	Highly Urban/ Urban	20/04/2022 08:46:00	complet	dev_integral	300.3	4
5	uck	no	2	12.0	Highly Urban/ Urban	23/03/2022 01:32:40	complet	dev_vip	648.13	5
6		no	2	15.0	z_Highly Rural/ Rural	23/03/2022 22:20:50	complet	dev_premium	459.83	6
7		no	0	9.0	z_Highly Rural/ Rural	13/03/2022 01:10:59	complet	dev_express	261.28	7
8		no	1	10.0	Highly Urban/ Urban	09/03/2022 03:43:53	complet	dev_premium	343.17	8
9		no	0	14.0	Highly Urban/ Urban	29/01/2022 10:17:49	complet	dev_premium	389.24	9
10	uck	yes	1	10.0	Highly Urban/ Urban	26/03/2022 07:37:33	complet	dev_integral	310.66	10
11		no	0	10.0	Highly Urban/ Urban	26/05/2022 05:51:37	complet	dev_express	261.6	11
12		no	0	1.0	Highly Urban/ Urban	27/01/2022 09:33:29	complet	dev_premium	380.99	12

Ajouter une composante aléatoire

1. Créer une colonne personnalisée :

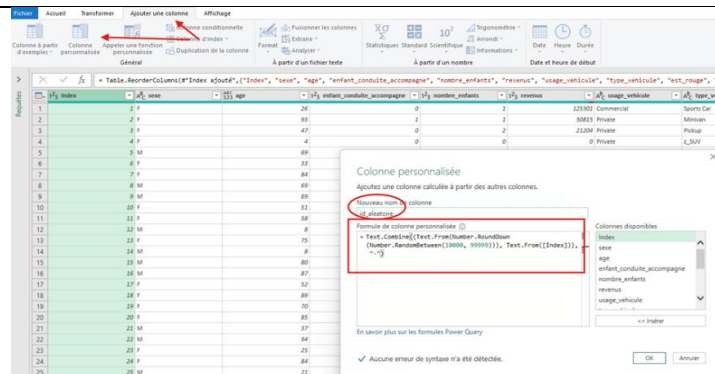
- Dans le ruban **Ajouter une colonne**, cliquez sur **Colonne personnalisée**.
- Donnez un nom explicite à la colonne, par exemple, ID_aléatoire.

2. Utiliser la fonction aléatoire :

- Dans le champ de formule, entrez l'expression suivante :

```
Text.Combine({Text.From(Number.RoundDown(Number.RandomBetween(10000, 99999))),
              Text.From([Index])}, "-")
```

- Cette formule génère un ID aléatoire de 5 chiffres combiné à l'index unique. Vous pouvez ajuster la plage aléatoire si besoin.



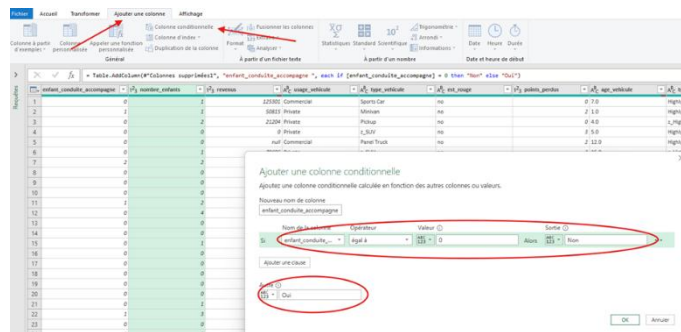
Modifier la colonne « enfant_conduite_accompagne »

1. Sélectionner la colonne :

- Dans Power Query, sélectionnez la colonne « **enfant_conduite_accompagne** »

2. Créer une colonne conditionnelle :

- Allez dans l'onglet **Ajouter une colonne**, puis cliquez sur **Colonne conditionnelle**.
- Remplissez les champs comme suit :
 - **Nom de la nouvelle colonne** : Donnez-lui un nom, par exemple « Enfants_conduite_accompagne ».
 - **Condition** :
 - Colonne : « **enfant_conduite_accompagne** »
 - Opérateur : **est égal à**
 - Valeur : **0**
 - Résultat : **Non**
 - Autres : **Oui**



3. Valider :

- Cliquez sur **OK** pour appliquer la transformation.
- Une nouvelle colonne sera créée avec les valeurs "Oui" ou "Non".



Modifier la colonne « nombre_enfants »

1. Suivre le même processus :

- Répétez les étapes ci-dessus pour la colonne « nombre_enfants ».
- Nommez la nouvelle colonne, par exemple, Enfants.

2. Conditions :

- Si nombre_enfants = 0, alors Non.
- Si nombre_enfants > 0, alors Oui.

Modification des revenus

Ajouter une colonne conditionnelle pour les tranches

1. Accéder à l'option de colonne conditionnelle :

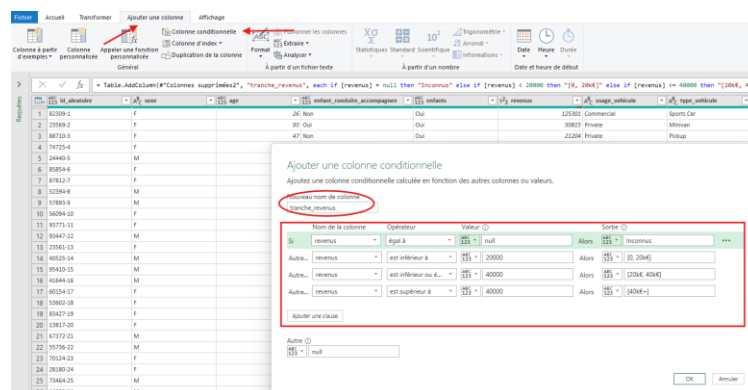
- Allez dans le ruban **Ajouter une colonne** et sélectionnez **Colonne conditionnelle**.

2. Configurer les tranches :

- Donnez un nom à la nouvelle colonne, par exemple, Tranches_Revenus.
- Ajoutez les conditions suivantes :
 - **Condition 1** : Si revenus est **null**, alors Inconnus.
 - **Condition 2** : Si revenus est **inférieur ou égal à 20000**, alors [0, 20k€].
 - **Condition 3** : Si revenus est **supérieur à 20000** et **inférieur ou égal à 40000**, alors [20k€, 40k€].
 - **Condition 4** : Si revenus est **supérieur à 40000**, alors [40k€+].

3. Valider :

- Cliquez sur **OK** pour appliquer les transformations.

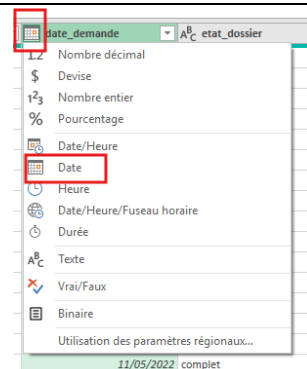


Changement du format de « date_demande »



Changer le type de données en Date :

- Avec la colonne sélectionnée, cliquez sur l'icône du type de données à gauche du nom de la colonne.
- Sélectionnez **Date** dans la liste.



Transformer la colonne « *tarif_devis* » au format décimal

Remplacer les points par des virgules

1. Utiliser l'option Remplacer des valeurs :

- Sélectionnez la colonne « **tarif_devis** ».
- Allez dans le ruban **Transformer** et cliquez sur **Remplacer des valeurs**.
- Configurez comme suit :
 - **Valeur à rechercher** : « . » (Le point).
 - **Valeur de remplacement** : « , » (La virgule).
- Cliquez sur **OK**.



Convertir la colonne en format Décimal

1. Changer le type de données en Décimal :

- Avec la colonne « **tarif_devis** » sélectionnés, cliquez sur l'icône à gauche du nom de la colonne.
- Sélectionnez **Nombre décimal** dans la liste.



Conclusion

En conformité avec les normes du Règlement Général sur la Protection des Données (RGPD), ce rapport met en évidence les ajustements nécessaires pour garantir la protection des données personnelles tout en optimisant leur exploitation. Le respect des principes de minimisation, de finalité légitime, de transparence et de sécurité a guidé les étapes de ce projet, depuis l'extraction des données jusqu'à leur anonymisation partielle.

Le fichier de données élaboré dans ce cadre est spécifiquement destiné au service commercial et a été conçu pour répondre exclusivement à ses besoins opérationnels et analytiques. Il est important de noter que d'autres services de l'entreprise pourraient avoir des attentes différentes et nécessiter des ajustements spécifiques des données pour répondre à leurs propres objectifs, toujours dans le respect des exigences RGPD.

Par ailleurs, pour renforcer la conformité, une recommandation clé est l'ajout d'une colonne calculée dans la base de données permettant de déterminer l'ancienneté des données. Cette démarche facilite la gestion des durées de conservation, un élément essentiel pour garantir une politique d'archivage, de suppression ou d'anonymisation adaptée.

Enfin, il est primordial de maintenir une vigilance continue et d'effectuer des audits réguliers pour ajuster les processus en fonction des évolutions technologiques et des recommandations des autorités compétentes, notamment la CNIL. Ces efforts contribueront à préserver la confiance des clients et à renforcer la réputation de l'entreprise dans la gestion éthique et sécurisée des données.