



Rapport de traitement des données

Titre du projet	<i>Collectez des données en respectant les normes RGPD</i>
------------------------	--

Version	Auteur	Description	Date
V1	Liodenot Julien	Création du document	15.02.2025

Introduction
<p>Après la plainte à la CNIL d'un ancien client de Dev'Immediat, sur la trop longue conservation de ses anciennes données, la société est sanctionnée d'une limitation temporaire des traitements. La finalité de ce projet est donc de s'assurer que les règles du RGPD sont parfaitement appliquées et intégrées dans les processus de la société pour demander la levée de cette sanction.</p> <p>Il est également demandé de mettre à disposition de l'équipe de performance commerciale une extraction des données du CRM pour qu'ils puissent continuer à travailler.</p>



Contenu du Rapport :

- **Préconisations** en lien avec les règles RGPD permettant de garantir à l'avenir le respect du RGPD
 - o **Minimisation des données**

Ne collecter que les données nécessaires au métier de courtier en assurance automobile, permettant de réaliser les devis, les contrats etc..

Des données sensibles sur la santé sont actuellement collectées (numéro de sécurité social, groupe sanguin) et ne doivent plus l'être.
 - o **Conservation des données**

En l'absence de conclusion de contrat, les données récoltées devront être mise à l'écart, par exemple dans une base différente. Après trois ans il faudra les supprimer.

On ne gardera dans la base active que les données des clients actifs le temps de leur traitement et utilité, puis on les passera en archivage intermédiaire. Dès qu'un client part, on mettra ses données sur une base à l'écart.
 - o **Transparence aux utilisateurs**

Sur le site de l'agence, mettre en place un lien permettant d'en savoir plus sur le traitement des données. Les explications doivent être claires et concises.

Il est également possible de fractionner les explications en les rendant disponible au moment de leur collectes . Par exemple si la première page d'un formulaire récupère les données liées au statut véhicule/permis, mettre un encart 'en savoir plus sur la récupération de ces données', faire de même pour les formulaires suivants.
 - o **Pseudonymisation des données**

Au moment de l'archivage des données, il faudra les Pseudonymiser, afin qu'elles ne puissent pas directement permettre une identification de la personne. Si besoin de désarchiver ces données il sera toujours possible de revenir en arrière, l'identification indirecte étant toujours possible.

De même si besoin de fournir les données à une équipe non habilitée, comme demandé ici pour l'équipe de performance commerciale, la pseudonymisation, voir l'anonymisation (identification même indirecte impossible), seront utilisés.
 - o **la sécurité des données**

Il faut garantir la sécurité des données. Seuls les tiers habilités doivent avoir accès aux données (en sectorisant les Droits), et il faut bien séparer la collecte dans différentes bases. Une base active sera physiquement séparée des bases d'archivage intermédiaires.



- **Documentation** des traitements effectués :

o **Requête SQL d'extraction**

- On ne va sélectionner que les données pertinentes
- On va filtrer les dossier complets et l'année 2022

```
select date_demande,sexe,date_naissance,enfant_conduite_accompagne,nombre_enfants,
revenus,usage_vehicule,type_vehicule ,age_vehicule,type_conduite,formule,tarif_devis,
points_perdus from base_client
where etat_dossier='complet' and date_demande like'2022%';
```

o **Application dans SQLITE et export des données**

- 1158 lignes à extraire

Requête

Historique

1

select date_demande,sexe,date_naissance,enfant_conduite_accompagne,nombre_enfants,revenus,

2

usage_vehicule,type_vehicule ,age_vehicule,type_conduite,formule,tarif_devis,points_perdus

3

from base_client


4


where etat_dossier='complet' and date_demande like '2022%';


5


Table


Formulaire














1








Nombre de lignes chargées : 1158


	date demande	sexe	date naissa	enfant co	nombre enfants	revenus	usage vehic
1	2022-02-25 16:13:38	F	1999-09-25	0	1	125301	Commercial
2	2022-05-24 21:07:34	F	1932-01-28	1	1	50815	Private

- Utilisation de la fonction export de SQLite vers fichier CSV

←  Exporter

Exporter formatset options

Format d'exportation
CSV

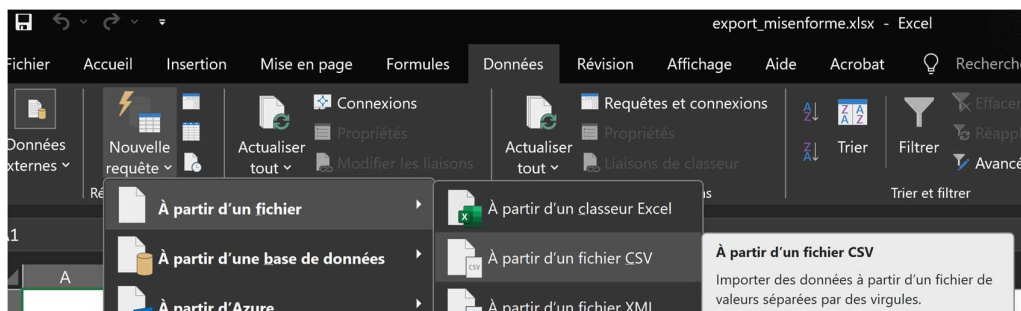
Sortie
☒ Fichier C:/cours/Projet 4/prepa livrables/donnees_SQLITE.csv 
☐ presse-papier

Texte encodé exporté : UTF-8

Exporter options de format
☒ Noms des colonnes dans la 1ère ligne
Séparateur de colonnes : ; (point virgule)
Exporter la valeur NULL comme : Chaîne de caractères vide

○ Import via PowerQuery

- Sur un nouveau fichier Excel on va ouvrir le fichier csv via PowerQuery. Pour cela, dans données on va faire **Nouvelle requête/A partir d'un fichier/A partir d'un fichier CSV** et sélectionner notre fichier de données.



- On vérifie dans l'aperçu les données puis on charge.

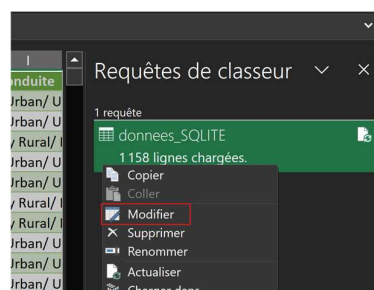
donnees_SQLITE.csv

Origine du fichier: 1252: Europe de l'Ouest (Windows) | Délimiteur: Point-virgule | Détection du type de données: Selon les 200 premières lignes

date_demande	sexe	date_naissance	enfant_conduite_accompagne	nombre_enfants	revenus	usage_vehicule	t
25/02/2022 16:13:38	F	25/09/1999	0	1	125301	Commercial	S
24/05/2022 21:07:34	F	28/01/1932	1	1	50815	Private	M
24/01/2022 08:17:21	F	15/09/1978	0	2	21204	Private	F
20/04/2022 03:46:00	F	17/02/2021	0	0	0	Private	Z
23/03/2022 01:32:40	M	14/06/1956	0	0	null	Commercial	F
23/03/2022 22:20:50	F	11/04/1992	0	1	79606	Private	Z
13/03/2022 01:10:59	F	18/03/1941	2	2	130540	Commercial	M
09/03/2022 03:43:53	M	06/03/1956	0	0	51278	Private	Z
20/01/2022 10:17:10	M	10/03/1956	0	0	115011	Private	M

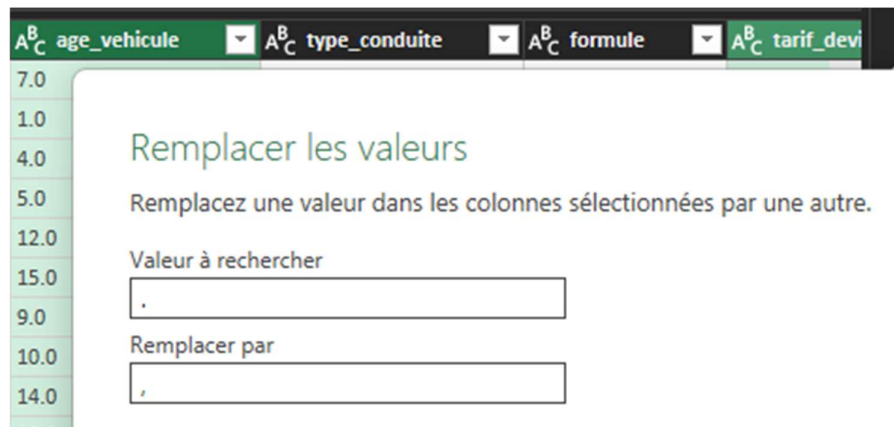
Charger | Transformer les données | Annuler

- On peut accéder à Power Query via le volet « requête de classeur »

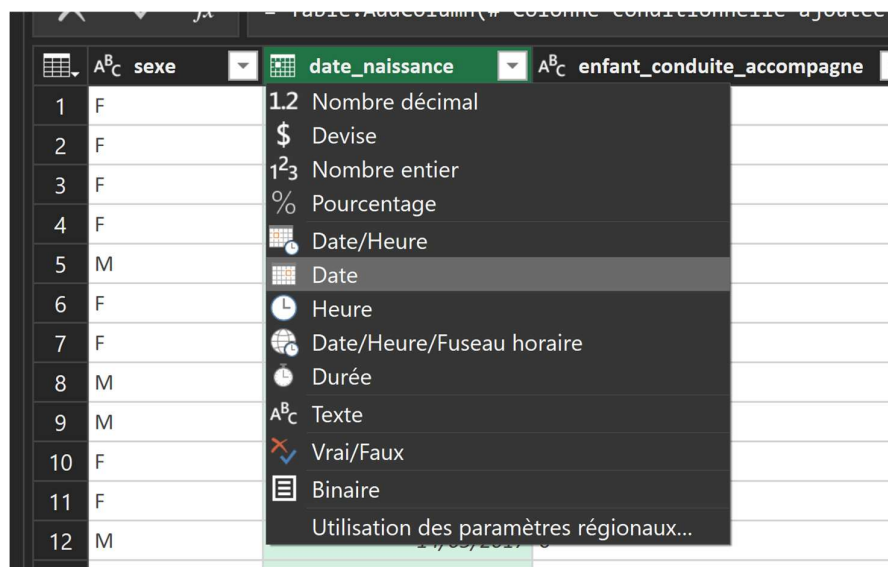


○ Power Query : Modification des données

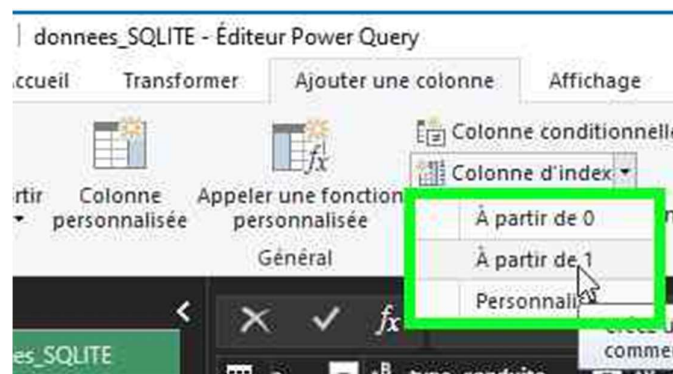
- Nombres décimaux : Pour deux champs (age_vehicule et Tarif_devis), les valeurs ont bien été importées en nombre, cependant avec un séparateur décimal en point. On va changer le séparateur « point » par un séparateur « virgule », étant le type de séparateur en vigueur, via le menu transformer/Remplacer les valeurs.



- Type de données : Tous les champs ont été importés en données de type texte. On va préciser pour certaines colonnes le type de données plus précisément (champs date, nombre entier etc...):



- **Id_client** : on va créer une nouvelle colonne qui va remplacer les identifiants d'origine que l'on n'a pas rapatriés.



- **Age** : On va calculer l'âge des personnes à partir de leur date de naissance. Pour cela on ajoute une nouvelle colonne personnalisée où l'on va calculer l'année actuelle - l'année de naissance, tout en retranchant une année si l'anniversaire de la personne n'est pas passé

Nouveau nom de colonne

Age

Formule de colonne personnalisée

```
= let
    aujourd'hui = Date.From(DateTime.LocalNow()),
    age = Date.Year(aujourd'hui) - Date.Year([date_naissance])
-
    (if Date.AddYears([date_naissance], Date.Year
(aujourd'hui) - Date.Year([date_naissance])) > aujourd'hui then
1 else 0)
in
    age
```

Le principe de la formule est la création de deux variables :

- *aujourd'hui* qui est l'année actuelle
- *age* qui est le calcul final et correspond à l'année actuelle Moins l'année de naissance moins 1 an si l'anniversaire de la Personne n'est pas passé (condition IF)

- Age on va ensuite créer une seconde colonne qui répartit l'âge par tranches (la colonne âge sera ensuite supprimée).

Ajouter une colonne conditionnelle

Ajoutez une colonne conditionnelle calculée en fonction des autres colonnes ou valeurs.

Nouveau nom de colonne

Tranche d'âge

	Nom de la colonne	Opérateur	Valeur ①		Sortie ①
Si	Age	est inférieur ou é...	ABC 123 17	Alors	ABC 123 Mineur
Autre...	Age	est supérieur à	ABC 123 65	Alors	ABC 123 [+65]
Autre...	Age	est supérieur à	ABC 123 49	Alors	ABC 123 [50-65]
Autre...	Age	est supérieur à	ABC 123 28	Alors	ABC 123 [29-49]
Autre...	Age	est supérieur à	ABC 123 17	Alors	ABC 123 [18-28]

Ajouter une clause

Autre ①

ABC 123 [N/A]

- Enfant conduite accompagnée : remplacement de valeurs 0 et 1 par Non et Oui. Pour cela on va créer une nouvelle colonne conditionnelle :

Ajouter une colonne conditionnelle

Ajoutez une colonne conditionnelle calculée en fonction des autres colonnes ou valeurs.

Nouveau nom de colonne

Enfant conduite accompagnée

	Nom de la colonne	Opérateur	Valeur ①		Sortie ①
Si	enfant_conduite_...	égal à	ABC 123 0	Alors	ABC 123 non
Autre...	enfant_conduite_...	est différent de	ABC 123 0	Alors	ABC 123 oui

Ajouter une clause

Autre ①

ABC 123 [N/A]

- Nombre enfant : On applique la même action de colonne conditionnelle, pour avoir oui ou non.

- Age du véhicule : Remplacement par une nouvelle colonne contenant l'âge par tranche.

Ajouter une colonne conditionnelle

Ajoutez une colonne conditionnelle calculée en fonction des autres colonnes ou valeurs.

Nouveau nom de colonne

Age du vehicule

	Nom de la colonne	Opérateur	Valeur ①			Sortie ①	
Si	age_vehicule ▾	égal à ▾	ABC 123 ▾	null	Alors	ABC 123 ▾	[N/A]
Autre...	age_vehicule ▾	est supérieur ou é... ▾	ABC 123 ▾	20	Alors	ABC 123 ▾	[20 ans+]
Autre...	age_vehicule ▾	est inférieur à ▾	ABC 123 ▾	10	Alors	ABC 123 ▾	[Moins de 10 ans]
Autre...	age_vehicule ▾	est supérieur ou é... ▾	ABC 123 ▾	10	Alors	ABC 123 ▾	[10-20]

Ajouter une clause

Autre ①

ABC 123 ▾ [N/A]

A noter que sur age_vehicule les champs vides avaient été remplis en *null* par Powerquery. La première condition est donc ici de transformer ce *null* en N/A, sinon la valeur error apparaîtra.

- Tarif du devis : Il a été explicitement demandé de ne pas mettre à disposition les tarifs exacts pratiqués. Aussi nous allons créer des tranches de prix pour remplacer cette colonne.

Ajouter une colonne conditionnelle

Ajoutez une colonne conditionnelle calculée en fonction des autres colonnes ou valeurs.

Nouveau nom de colonne

Tranche Tarif

	Nom de la colonne	Opérateur	Valeur ①		Sortie ①
Si	tarif_devis	est inférieur à	ABC 123 300	Alors	ABC 123 [Moins de 300]
Autre...	tarif_devis	est supérieur ou é...	ABC 123 800	Alors	ABC 123 [800+]
Autre...	tarif_devis	est supérieur ou é...	ABC 123 600	Alors	ABC 123 [600-799]
Autre...	tarif_devis	est supérieur ou é...	ABC 123 300	Alors	ABC 123 [300-599]

Ajouter une clause

Autre ①

ABC 123 N/A

- Revenus : De la même manière on va regrouper les revenus par tranche.

Tranche Revenus

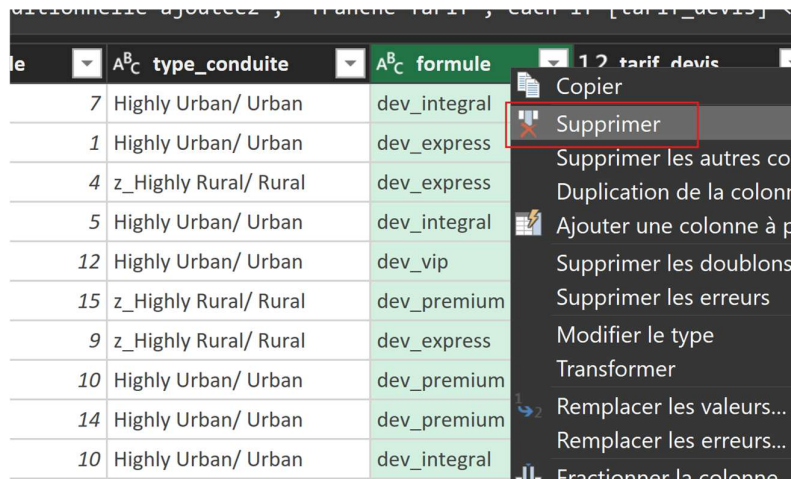
	Nom de la colonne	Opérateur	Valeur ①		Sortie ①
Si	revenus	égal à	ABC 123 null	Alors	ABC 123 [N/A]
Autre...	revenus	est inférieur à	ABC 123 50000	Alors	ABC 123 [-50k]
Autre...	revenus	est supérieur ou é...	ABC 123 150000	Alors	ABC 123 [+150k]
Autre...	revenus	est supérieur ou é...	ABC 123 100000	Alors	ABC 123 [100k-149k]
Autre...	revenus	est supérieur ou é...	ABC 123 50000	Alors	ABC 123 [50k-99k]

Ajouter une clause

Autre ①

ABC 123 [N/A]

- Le cas de la colonne Formule : bien qu'à la base cette colonne faisait sens dans le fait de la récupérer, après réflexion elle pourrait permettre aux équipes commerciales de lier une fourchette de prix avec une formule. J'ai donc décidé de ne pas la garder. Pour cela dans Power Query il suffit de supprimer la colonne.



	AB_C type_conduite	AB_C formule
7	Highly Urban/ Urban	dev_integral
1	Highly Urban/ Urban	dev_express
4	z_Highly Rural/ Rural	dev_express
5	Highly Urban/ Urban	dev_integral
12	Highly Urban/ Urban	dev_vip
15	z_Highly Rural/ Rural	dev_premium
9	z_Highly Rural/ Rural	dev_express
10	Highly Urban/ Urban	dev_premium
14	Highly Urban/ Urban	dev_premium
10	Highly Urban/ Urban	dev_integral

- Type de conduite : simplification des valeurs de la colonne

Remplacer les valeurs

Remplacez une valeur dans les co Remplacez une valeur dans l

Valeur à rechercher

Highly Urban/ Urban

Remplacer par

Urban

Valeur à rechercher

z_Highly Rural/ Rural

Remplacer par

Rural

- Type_vehicule : correction de la valeur SUV

Valeur à rechercher

z_SUV

Remplacer par

SUV

Export du fichier final



Une fois toutes ces actions effectuées sur le fichier Excel, nous allons enregistrer les données au format CSV. Cela aura pour effet de générer un fichier complet à l'attention des équipes, sans le paramétrage Powerquery.

A	B	C	D	E	F	G	H	I	J
ID_Client	sexe	Tranche d'age	points_perdus	Enfant conduite accompagnée	usage_vehicule	type_vehicule	Age du véhicule	type_conduite	Tranche Tarif
1	F	[18-28]	0	Non	Commercial	Sports Car	[Moins de 10 ans]	Urban	[300-599]
2	F	[65+]	2	Oui	Private	Minivan	[Moins de 10 ans]	Urban	[Moins de 300]
3	F	[29-49]	0	Non	Private	Pickup	[Moins de 10 ans]	Rural	[Moins de 300]
4	F	Mineur	3	Non	Private	SUV	[Moins de 10 ans]	Urban	[300-599]
5	M	[65+]	2	Non	Commercial	Panel Truck	[10-20]	Urban	[600-799]
6	F	[29-49]	2	Non	Private	SUV	[10-20]	Rural	[300-599]
7	F	[65+]	0	Non	Commercial	Minivan	[Moins de 10 ans]	Rural	[Moins de 300]
8	M	[65+]	1	Non	Private	SUV	[10-20]	Urban	[300-599]
9	M	[65+]	0	Non	Private	Minivan	[10-20]	Urban	[300-599]

Conclusion

Toutes ces manipulations ont mené le projet a terme : Les équipes commerciales ont récupéré un fichier Anonymisé mais complet en informations afin de continuer à travailler pendant le blocage de la CNIL. Les recommandations délivrées permettront, une fois mise en place, de lever le blocage et de travailler en conformité avec la réglementation en vigueur.