



# Rapport de traitement des données

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

Version	Auteur	Description	Date
VI	Roxane RIVIERE	Documentation des actions de conformité au RGPD : processus de collecte et de traitement des données. Rédaction des préconisations à destination de la direction.	06/2024

Introduction
<p>Dev'Immediat subit une sanction de la CNIL, à la suite d'une plainte d'un client et d'une non-conformité au Règlement Général sur la Protection des Données. Vous devez poursuivre votre activité de prospection commerciale malgré la restriction imposée des données.</p> <p>L'équipe ressent une forte pression puisqu'elle a besoin d'exploiter les données pour piloter la performance commerciale. Vous souhaitez donc résoudre cette situation rapidement et efficacement.</p> <p>L'enjeu est de lever la sanction en prouvant que Dev'Immediat respecte les règles du RGPD et qu'elle dispose des procédures internes adéquates pour le faire. La mise en conformité permettra de retrouver la confiance de vos clients sur la protection de leurs données.</p>



## Contenu du rapport :

### - Préconisations en lien avec les règles RGPD

#### ▪ Collecte des données

- ✓ **Minimisation des données** : ne collectez que les données strictement nécessaires à votre prospection commerciale.
- ✓ **Droits des personnes** : assurez-vous que les droits des personnes sont respectés (accès, rectification, suppression des données personnelles) et identifiez les interlocuteurs à privilégier pour répondre aux demandes.
- ✓ **Transparence** : informez vos clients de manière claire et complète sur la collecte, l'utilisation et la conservation de leurs données personnelles.

#### ▪ Traitement des données

- ✓ **Conservation limitée des données** : définissez les durées après lesquelles les données doivent être archivées, anonymisées ou supprimées.
- ✓ **Sécurisation des données** : protégez les données contre les cyberattaques en les chiffrant et en sensibilisant vos équipes sur ce type de risque.

Nous avons établi ces préconisations après avoir exploré votre base client. En les respectant, vous disposerez de pratiques durables de gestion des données.

### - Documentation des traitements effectués

#### ▪ Requête SQL

La requête SQL suivante a permis l'extraction des données des clients, dont l'état du dossier est complet pour l'année 2022 :

```
SELECT sexe, date_naissance, id_client, 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 >= "2022-01-01"  
AND etat_dossier = "complet"
```



Seules les données strictement nécessaires ont été conservées :

#### Informations sur le client

- **sexe** : pour la personnalisation des offres
- **date\_naissance** : pour la vérification de l'âge
- **id\_client** : pour l'identification unique
- **enfant\_conduite\_accompagne** : pour l'évaluation des risques
- **nombre\_enfants** : pour la personnalisation des offres
- **revenus** : pour l'évaluation de la capacité financière

#### Informations sur le véhicule

- **usage\_vehicule** : pour l'établissement du devis
- **type\_vehicule** : pour l'établissement du devis
- **est\_rouge** : pour l'évaluation des risques
- **points\_perdus** : pour l'évaluation des risques
- **age\_vehicule** : pour l'établissement du devis
- **type\_conduite** : pour l'établissement du profil de conduite

#### Informations sur la demande de devis

- **date\_demande** : pour la vérification de la période de demande
- **etat\_dossier** : pour la vérification de la complétude du dossier
- **formule** : pour l'établissement du devis
- **tarif\_devis** : pour l'analyse des performances commerciales

Ces données ont été exportées au format CSV depuis SQLite Studio et les valeurs vides ont été remplacées par des zéros.

### ▪ Etapas de traitement dans Power Query

#### 1. Importation des données

L'extraction des données CSV est importée dans Power Query à l'aide d'Excel.

#### 2. Transformation des données

Cette étape permet d'agréger les données et préserve leur confidentialité.

- **date\_naissance** :
  - Transformer les dates de naissance en âge ;
  - Extraire l'année de naissance : Ajouter une colonne => Extraire => Derniers caractères.



Insérer les derniers caractères

Entrez le nombre de caractère de fin à conserver.

Nombre

OK Annuler

- Modifier le type de donnée en nombre entier pour pouvoir faire des calculs ;
- Calculer l'âge du client à l'aide d'une colonne personnalisée.

Colonne personnalisée

Ajoutez une colonne calculée à partir des autres colonnes.

Nouveau nom de colonne

Formule de colonne personnalisée ⓘ

Colonnes disponibles

- sexe
- date\_naissance
- id\_client
- enfant\_conduite\_accompagne
- nombre\_enfants
- revenus
- usage\_vehicule

<< Insérer

En savoir plus sur les formules Power Query

✓ Aucune erreur de syntaxe n'a été détectée.

OK Annuler

- **id\_client :**

- Ajouter une colonne indexée pour anonymiser les identifiants des clients.

Ajouter une colonne d'index

Ajoutez une colonne d'index et indiquez un index de départ et un incrément.

Index de départ

Incrément

OK Annuler

- Ajouter une colonne personnalisée qui va attribuer aléatoirement, à chaque ligne, un numéro entre 1 et 1158.



Colonne personnalisée

Ajoutez une colonne calculée à partir des autres colonnes.

Nouveau nom de colonne  
num\_aleatoire

Formule de colonne personnalisée ⓘ  
= Number.RoundDown(Number.RandomBetween(1,1158))

Colonnes disponibles  
sexe  
date\_naissance  
id\_client  
enfant\_conduite\_accompagne  
nombre\_enfants  
revenus  
usage\_vehicule

<< Insérer

En savoir plus sur les formules Power Query

✓ Aucune erreur de syntaxe n'a été détectée.

OK Annuler

- Trier les lignes dans l'ordre croissant : on peut retrouver le même numéro plusieurs fois ;
- Ajouter, à nouveau, une colonne indexée : chaque numéro est unique et va de 1 à 1158 => les lignes sont réparties de façon aléatoire ;
- Supprimer les colonnes intermédiaires.

```
= Table.RemoveColumns("#Index client aléatoire",{ "Index", "num_aleatoire" })
```

- **enfant\_conduite\_accompagne et nombre\_enfants :**

- Remplacer le nombre d'enfants par des variables binaires.

Ajouter une colonne conditionnelle

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

Nouveau nom de colonne  
enfants

	Nom de la colonne	Opérateur	Valeur ⓘ		Sortie ⓘ
Si	nombre_enfants	égal à	ABC 123 0	Alors	ABC 123 Non
Autre...	nombre_enfants	est supérieur à	ABC 123 0	Alors	ABC 123 Oui

Ajouter une clause

Autre ⓘ  
ABC 123 null

OK Annuler

- **revenus :**

- Regrouper les revenus par tranches de 20k€ avec une colonne personnalisée.



Ajouter une colonne conditionnelle

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

Nouveau nom de colonne  
tranches\_revenus

	Nom de la colonne	Opérateur	Valeur	Sortie
Si	revenus	est inférieur ou é...	20000	Alors [ < 20k€ ]
Autre...	revenus	est inférieur ou é...	40000	Alors [ 20k€ - 40k€ ]
Autre...	revenus	est inférieur ou é...	60000	Alors [ 40k€ - 60k€ ]
Autre...	revenus	est inférieur ou é...	80000	Alors [ 60k€ - 80k€ ]
Autre...	revenus	est inférieur ou é...	100000	Alors [ 80k€ - 100k€ ]
Autre...	revenus	est supérieur à	100000	Alors [ > 100k€ ]

Ajouter une clause

Autre  
null

OK Annuler

- **tarif\_devis :**

- Modifier le type texte en type nombre décimal ;
- Remplacer les virgules par des points pour éviter des erreurs.

Remplacer les valeurs

Remplacez une valeur dans les colonnes sélectionnées par une autre.

Valeur à rechercher  
.

Remplacer par  
.

Options avancées

OK Annuler

- Regrouper les tarifs par tranches de 200€ à l'aide d'une colonne personnalisée.

Ajouter une colonne conditionnelle

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

Nouveau nom de colonne  
tranches\_devis

	Nom de la colonne	Opérateur	Valeur	Sortie
Si	tarif_devis	est inférieur à	200	Alors [ < 200€ ]
Autre...	tarif_devis	est inférieur ou é...	400	Alors [ 200€ - 400€ ]
Autre...	tarif_devis	est inférieur ou é...	600	Alors [ 400€ - 600€ ]
Autre...	tarif_devis	est inférieur ou é...	800	Alors [ 600€ - 800€ ]
Autre...	tarif_devis	est inférieur ou é...	1000	Alors [ 800€ - 1000€ ]
Autre...	tarif_devis	est supérieur à	1000	Alors [ > 1000€ ]

Ajouter une clause

Autre  
null

OK Annuler

Anonymisation des données - Éditeur Power Query

Table.TransformColumnTypes(#"Valeur remplacée1",{"age\_vehicule", Int64.Type}, {"tranches\_devis", type text})

	index_aléatoire	sexe	age_client	enfant_conduite_accompagnee	enfants	tranches_revenus	usage_vehicule
1	2	M	24	Non	Non	[40k€ - 60k€]	Privé
2	2	M	5	Non	Non	[80k€ - 100k€]	Privé
3	3	M	69	Non	Non	[40k€ - 60k€]	Comm
4	4	F	71	Oui	Oui	[< 20k€]	Privé
5	5	M	27	Non	Oui	[20k€ - 40k€]	Comm
6	6	M	83	Non	Oui	[80k€ - 100k€]	Comm
7	7	F	55	Non	Non	[80k€ - 100k€]	Privé
8	8	F	61	Non	Non	[20k€ - 40k€]	Comm
9	9	M	27	Oui	Oui	[20k€ - 40k€]	Comm
10	10	F	57	Non	Non	[> 100k€]	Comm
11	11	M	38	Non	Oui	[> 100k€]	Comm
12	12	M	51	Non	Non	[< 20k€]	Comm
13	13	F	41	Non	Non	[80k€ - 100k€]	Comm
14	14	F	35	Non	Non	[< 20k€]	Privé
15	15	F	59	Oui	Oui	[< 20k€]	Privé
16	16	M	50	Non	Non	[40k€ - 60k€]	Comm
17	17	M	69	Oui	Oui	[20k€ - 40k€]	Privé
18	18	M	20	Non	Non	[40k€ - 60k€]	Privé
19	19	M	69	Non	Non	[60k€ - 80k€]	Comm
20	20	M	90	Oui	Oui	[60k€ - 80k€]	Privé
21	21	M	28	Non	Non	[> 100k€]	Privé
22	22	M	9	Non	Oui	[40k€ - 60k€]	Privé
23	23	M	77	Non	Non	[80k€ - 100k€]	Privé
24	24	F	49	Non	Oui	[< 20k€]	Privé
25	25	F	8	Non	Oui	[80k€ - 100k€]	Privé
26	26	F	83	Non	Oui	[40k€ - 60k€]	Comm
27	27						

16 COLONNES, 999+ LIGNES - Profilage de la colonne en fonction des 1000 premières lignes

APERÇU TÉLÉCHARGÉ À 11:09

## Intégralité des étapes – Power Query

### 3. Nettoyage des données

Après avoir transformé les données brutes, il est important de supprimer les informations superflues ou sensibles.

- Supprimer toutes les colonnes qui ne sont plus utiles ;
- Déplacer éventuellement les colonnes anonymisées ;
- Modifier les types de données.

```
Table.TransformColumnTypes(#"Valeur remplacée1",{"age_vehicule", Int64.Type}, {"tranches_devis", type text})
```

Les colonnes suivantes ont été conservées après anonymisation :

- index\_aléatoire
- sexe
- age\_client
- enfant\_conduite\_accompagnee
- enfants
- tranches\_revenus
- usage\_vehicule
- type\_vehicule
- est\_rouge
- points\_perdus
- age\_vehicule
- type\_conduite
- date\_demande
- etat\_dossier
- formule
- tranches\_devis



#### 4. Vérification de la conformité et de la qualité des données

Les données sont chargées dans Excel et enregistrées au format CSV.

Dans la nouvelle extraction, les données sensibles et personnelles ont été supprimées, les colonnes ont été regroupées par catégorie et les informations essentielles ont été conservées. Les données sont bien anonymisées et il n'est pas possible pour l'équipe commerciale d'identifier les individus.

Dev'Immediat peut donc poursuivre son activité en conformité avec le RGPD tout en répondant à ses besoins opérationnels.

Toutefois, des anomalies ou des incohérences ont été repérées.

- **age\_client** : < 18 ans
- **enfant\_conduite\_accompagnee** : incohérent selon l'âge déclaré
- **enfants** : incohérent selon l'âge déclaré
- **type\_vehicule** : données débutant par «Z\_ »
- **type\_conduite** : données débutant par «Z\_ »

Il est important de veiller à la qualité des données recueillies et de les fiabiliser. Leur analyse et les tendances décelées seront également de meilleures qualités.

### Conclusion

La conformité au RGPD n'est pas qu'une obligation légale, c'est aussi une exigence éthique qui assure la protection des données personnelles de vos clients. Il est nécessaire de mettre à jour continuellement vos processus internes et de former régulièrement vos équipes sur l'importance que revêt la sécurisation des données.

En respectant nos préconisations et en suivant les étapes documentées dans ce rapport, Dev'Immediat peut demander la levée de la sanction de la CNIL. Vous préserverez votre réputation et éviterez des sanctions d'ordres financiers.