



Modèle de rapport de traitement des données

Titre du projet	<i>Collecte de données en respectant le RGPD</i>
------------------------	--------------------------------------------------

Version	Auteur	Description	Date
V1	<i>Samuel.O</i>	<i>Rapport</i>	<i>15.11.2024</i>

Introduction
<p><i>Suite à une sanction de la CNIL, l'entreprise Dev'Immediat ne peut plus traiter les données de ces clients.</i></p> <p><i>Hors elle a besoin de ces traitements pour son activité.</i></p> <p><i>C'est l'occasion pour l'entreprise de revoir ses process en matière de RGPD et éviter d'autres problèmes à l'avenir.</i></p>



Contenu du rapport :

- **Préconisations** en lien avec les règles RGPD permettant de garantir à l'avenir le respect du RGPD
 - Notamment:
 - collecte (communication d'information vers les clients : transparence, droit des personnes, etc.).
 - traitement des données
- **Documentation** des traitements effectués :
 - Notamment :
 - documentation de la requête SQL
 - documentation des étapes de traitement du fichier Power Query
 - documentation des traitements effectués sur la base de données brutes
 - étapes pour retravailler le jeu de données et les explications associées
-

Conclusion

Ouverture sur les exigences, contraintes et enjeux de la confidentialité



Préconisations :

1. Limiter la collecte des données aux informations strictement nécessaires

Conformément à la règle de minimisation des données, on recommande de revoir la base de données et de restreindre la collecte aux informations essentielles pour les analyses. Par exemple, le groupe sanguin et le numéro de sécurité sociale ne sont pas indispensables pour les contrats d'assurance, il faut les supprimer.

Cette approche permet de réduire les risques associés au traitement de données sensibles.

2. Définir clairement les finalités des données récoltées

Chaque donnée collectée doit être associée à une finalité explicite et légitime.

Définissez ces finalités, et communiquez-les clairement aux clients, en précisant pourquoi chaque information est requise et comment elle sera utilisée.

Par exemple, les informations de revenus et de résidence peuvent être justifiées pour évaluer les risques en assurance, mais pas pour des usages promotionnels non prévus initialement.

3. S'assurer le consentement pour les données sensibles

Pour les données sensibles (santé, informations financières), il convient d'obtenir un consentement explicite des clients. Ce consentement doit être libre, spécifique et informé. Il est alors utile de prévoir des mécanismes de traçabilité de ce consentement pour prouver la conformité en cas de contrôle.

4. Mettre en place une politique de conservation des données

Limitez la durée de conservation des données en fonction de leur finalité. Par exemple, les informations relatives aux devis non conclus peuvent être supprimées ou anonymisées après un délai défini (ex. 2 ans), tandis que les informations de clients actifs peuvent être conservées tant que le contrat est en d'actualité.

5. Renforcer les mesures de sécurité et les contrôles d'accès

Conformément à l'obligation de sécurité, mettez en place des mesures comme le chiffrement des données sensibles (numéro de sécurité sociale, adresse, informations financières) et des contrôles d'accès stricts pour garantir que seules les personnes autorisées peuvent accéder aux données personnelles (Groupes AD, RLS etc.).



Documentation des traitements effectués :

Requête SQL :

```
SELECT
SUBSTR(date_demande, 1, 4) as annee,
id_client,
date_naissance,
nombre_enfants,
enfant_conduite_accompagne,
sexe,
adresse,
revenus,
est_rouge,
usage_vehicule,
type_vehicule,
age_vehicule,
formule,
tarif_devis
FROM base_client
WHERE etat_dossier = "complet" AND SUBSTR(date_demande, 1, 4) = "2022"
;
```

Dans cette requête SQL, on ne sélectionne que les données qui peuvent potentiellement servir dans le cadre de l'activité de l'entreprise et par la suite à des fins d'analyse commerciale.

Les données ne concernent que l'année 2022 et les dossiers considérés comme « complet ».

Afin d'être anonymisés, les données seront de nouveau traitées avec le logiciel Power Query.































Traitements via Power Query :

Nom

base_client_202411081...

▼ Étapes appliquées

	Source	
	En-têtes promus	
	1 2 Valeur remplacée	
	1 2 Valeur remplacé...	
	Type de colonn...	
	Âge inséré	
	Total calculé de...	
	.00 → 0 Arrondi	
	Colonne conditi...	
	Colonne conditi...	
	1 2 3 Index ajouté	
	Colonne perm...	
	Colonne person...	
	Colonne person...	
	1 2 Valeur remplacé...	
	1 2 Valeur remplacé...	
	 Colonne suppr...	

1. Création d'une colonne âge à partir de la colonne « date de naissance ».
Le but sera par la suite de supprimer la colonne date de naissance pour renforcer l'anonymisation.



```
5  # "Âge inséré" = Table.AddColumn("#Type de colonne changé", "Âge", each DateTimeZone.LocalNow() - DateTimeZone.From([date_naissance]), type  
    nullable duration),  
6  # "Total calculé des années" = Table.TransformColumns("#Âge inséré", {{ "Âge", each Duration.TotalDays(_) / 365, type nullable number }}),  
7  Arrondi = Table.TransformColumns("#Total calculé des années", {{ "Âge", each Number.Round(_, 0), type number }}),
```

2. Création d'une colonne booléenne sur la présence d'enfants dans le foyer ou non :

```
# "Colonne conditionnelle insérée" = Table.AddColumn(Arrondi, "Enfants ?", each if [nombre_enfants] > 0 then "Oui" else if [nombre_enfants] = 0  
then "Non" else ""),
```

Cette colonne va remplacer la colonne nombre d'enfants. Plutôt que d'indiquer le nombre d'enfants des clients, on indique s'ils en ont ou pas.

3. Création d'une colonne booléenne sur la présence dans le foyer d'enfants en conduite accompagnée ou non :

```
11 # "Colonne conditionnelle insérée 1" = Table.AddColumn("#Colonne conditionnelle insérée", "enfant_conduite_accompagné ?", each if  
    [enfant_conduite_accompagne] = 0 then "Non" else if [enfant_conduite_accompagne] >= 1 then "Oui" else ""),
```

4. Création d'une colonne index qui remplacera la colonne identifiant client. Avec cette nouvelle colonne, aucun croisement ne sera possible avec d'autres base de données de l'entreprise

```
# "Index ajouté" = Table.AddIndexColumn("#Colonne conditionnelle insérée", "Index", 0, 1, Int64.Type),
```

5. Création d'une colonne catégorisant les données de revenus afin qu'elles ne soient pas disponibles de façon détaillée

```
# "Colonne personnalisée ajoutée" = Table.AddColumn("#Colonnes permutées", "Catégorie de revenu", each if [revenus] < 20000 then "0-20K"  
else if [revenus] >= 20000 and [revenus] <= 40000 then "20-40K"  
else if [revenus] > 40000 and [revenus] <= 60000 then "40-60K"  
else if [revenus] > 60000 and [revenus] <= 80000 then "60-80K"  
else if [revenus] > 80000 and [revenus] <= 100000 then "80-100K"  
else ">100K"),
```

6. Création d'une colonne (« tarif non détaillé ») catégorisant les données tarifaires en intervalles, afin qu'elles ne soient pas disponibles de façon détaillée pour les équipes de performances commerciales

```
20 # "Colonne personnalisée ajoutée 1" = Table.AddColumn("#Colonne personnalisée ajoutée", "Tarif non détaillé", each if [tarif_devis] >= 0 and  
    [tarif_devis] < 50 then "0 - 50"  
21 else if [tarif_devis] >= 50 and [tarif_devis] < 100 then "50 - 100"  
22 else if [tarif_devis] >= 100 and [tarif_devis] < 150 then "100 - 150"  
23 else if [tarif_devis] >= 150 and [tarif_devis] < 200 then "150 - 200"  
24 else if [tarif_devis] >= 200 and [tarif_devis] < 250 then "200 - 250"  
25 else if [tarif_devis] >= 250 and [tarif_devis] < 300 then "250 - 300"  
26 else if [tarif_devis] >= 300 and [tarif_devis] < 350 then "300 - 350"  
27 else if [tarif_devis] >= 350 and [tarif_devis] < 400 then "350 - 400"  
28 else "400 et plus"),
```



7. Suppression de plusieurs colonnes :

```
#"Colonnes supprimées" = Table.RemoveColumns(#"Colonne personnalisée ajoutée", {"date_naissance", "id_client", "nombre_enfants", "revenus",
"tarif_devis", "adresse"})
in
#"Colonnes supprimées"
```

Ici, on supprime les colonnes id_client, date_naissance, nombre d'enfants, revenus, adresse pour anonymiser les données et rendre impossible l'identification d'un client.

On supprime également la colonne tarif_devis pour éviter que les équipes commerciales n'aient accès aux données tarifaires détaillées conformément à la demande de l'entreprise.

Un aperçu de l'extract finale :

	A	B	C	D	E	F	G	H	I	J	K	L	M
	index	annee	sexe	est_rouge	usage_vehicule	type_vehicule	age_vehicule	formule	Âge	enfant_conduite_accompagné ?	Enfants ?	Catégorie de revenu	Tarif non détaillé
2	0	2022	F	Non	Commercial	Sports Car		7 dev_integral	25	Non	Oui	>100K	300 - 350
3	1	2022	F	Non	Private	Minivan		1 dev_express	93	Oui	Oui	40-60K	250 - 300
4	2	2022	F	Non	Private	Pickup		4 dev_express	46	Non	Oui	20-40K	250 - 300
5	3	2022	F	Non	Private	z_SUV		5 dev_integral	4	Non	Non	0-20K	300 - 350
6	4	2022	M	Non	Commercial	Panel Truck		12 dev_vip	68	Non	Non		400 et plus
7	5	2022	F	Non	Private	z_SUV		15 dev_premium	33	Non	Oui	60-80K	400 et plus
8	6	2022	F	Non	Commercial	Minivan		9 dev_express	84	Oui	Oui	>100K	250 - 300
9	7	2022	M	Non	Private	z_SUV		10 dev_premium	69	Non	Non	40-60K	300 - 350
10	8	2022	M	Non	Private	Minivan		14 dev_premium	69	Non	Non	>100K	350 - 400
11	9	2022	F	Oui	Commercial	Panel Truck		10 dev_integral	51	Non	Non		300 - 350
12	10	2022	F	Non	Private	Minivan		10 dev_express	58	Oui	Oui	20-40K	250 - 300
13	11	2022	M	Non	Private	z_SUV		1 dev_premium	8	Non	Oui	20-40K	350 - 400
14	12	2022	F	Non	Private	z_SUV		18 dev_express	74	Non	Non	60-80K	250 - 300
15	13	2022	M	Non	Commercial	Van		6 dev_integral	7	Non	Non	0-20K	300 - 350
16	14	2022	M	Non	Private	Minivan		1 dev_integral	79	Non	Oui	20-40K	300 - 350
17	15	2022	M	Non	Private	Minivan		13 dev_express	87	Non	Non	40-60K	250 - 300
18	16	2022	F	Non	Private	z_SUV		15 dev_integral	52	Non	Non	>100K	300 - 350
19	17	2022	F	Oui	Commercial	Van		8 dev_express	89	Non	Non	40-60K	250 - 300
20	18	2022	F	Oui	Private	Minivan		11 dev_express	70	Non	Non	60-80K	250 - 300
21	19	2022	F	Non	Private	z_SUV		10 dev_integral	85	Non	Non	40-60K	250 - 300
22	20	2022	M	Non	Private	Minivan		16 dev_express	36	Non	Oui	60-80K	250 - 300
23	21	2022	M	Non	Private	Minivan		dev_premium	93	Oui	Oui	20-40K	350 - 400
24	22	2022	F	Oui	Commercial	Van		7 dev_premium	25	Non	Non		350 - 400
25	23	2022	F	Non	Commercial	Sports Car		1 dev_premium	84	Non	Oui	40-60K	400 et plus
26	24	2022	M	Non	Commercial	Pickup		1 dev_express	20	Non	Non	20-40K	250 - 300
27	25	2022	M	Oui	Private	Minivan		dev_express	67	Non	Non		250 - 300
28	26	2022	F	Non	Private	Sports Car		1 dev_integral	26	Non	Non	40-60K	250 - 300
29	27	2022	F	Non	Private	z_SUV		dev_integral	46	Non	Non	60-80K	250 - 300
30	28	2022	F	Non	Private	z_SUV		1 dev_vip	56	Non	Non	40-60K	400 et plus
31	29	2022	F	Non	Private	z_SUV		17 dev_premium	18	Non	Non	>100K	350 - 400
32	30	2022	F	Non	Private	z_SUV		1 dev_express	50	Non	Oui	20-40K	250 - 300
33	31	2022	M	Oui	Commercial	Minivan		5 dev_integral	40	Non	Oui	20-40K	300 - 350
34	32	2022	F	Non	Private	z_SUV		9 dev_express	67	Non	Non	>100K	250 - 300
35	33	2022	M	Non	Private	z_SUV		6 dev_premium	7	Non	Non	20-40K	400 et plus
36	34	2022	M	Non	Private	Sports Car		12 dev_express	47	Oui	Oui	60-80K	250 - 300