

MISE EN CONFORMITÉ
RÈGLEMENT GÉNÉRAL PROTECTION DONNÉES
PRÉSENTATION DU 21/04/2023

SOMMAIRE

1 - CONTEXTE & ENJEUX(page3)

2 - PRÉCONISATIONS(page 4)

3 - DOCUMENTATION :

- *Requête SQL et export fichier CSV*(page 5)
- *Import CSV dans Excel*.....(page 6)
- *Excel vers Power Query*.....(page 7)
- *Éditeur Power Query*.....(page 8)
- *Anonymisation par randomisation*.....(page 9)
- *Anonymisation par généralisation*.....(page 10)
- *Tableau final anonymisé*.....(page 11)

4 - CONCLUSION(page 12)

CONTEXTE



Durée
conservation
inappropriée



Droit d'accès
client
non respecté



Sanction
de la
CNIL



Actions
correctives
à réaliser

ENJEUX



Maintenir
pilotage
de l'activité



Revoir structure
et process
internes



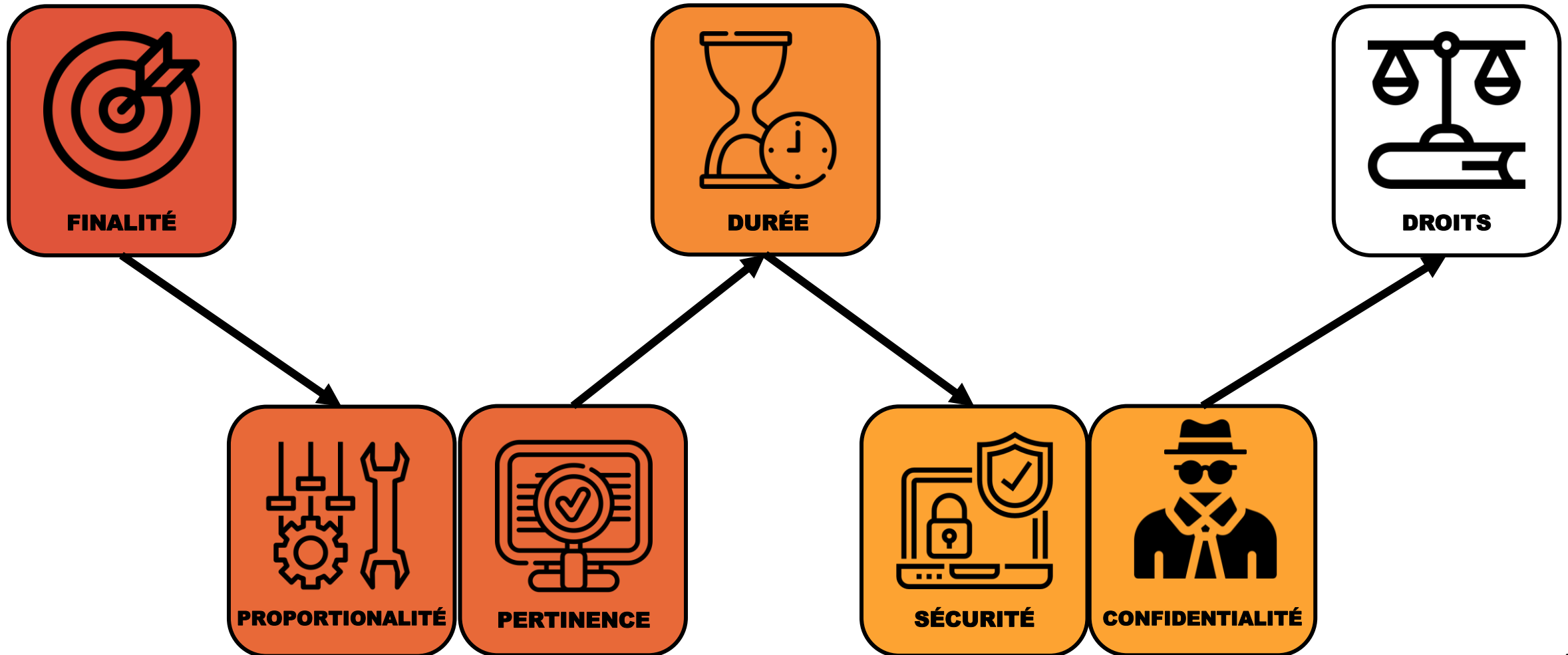
Mise en
conformité avec
règles RGPD



Éviter une
sanction plus
lourde

PRÉCONISATIONS

Respecter les 5 grands principes du RGPD :



DOCUMENTATION – Requête SQL et export fichier CSV

Requête incluant les champs strictement nécessaire au traitement

Bouton permettant d'effectuer la demande d'exportation de la requête

Fenêtre permettant de vérifier la requête à exporter et de saisir les options d'exportations (format, séparateur...)

Base de données incluant tables et champs

The screenshot displays the SQLiteStudio 3.4.3 interface. On the left, the 'Base de données' (Database) pane shows the 'CRM (SQLite 3)' database with a table 'base_client' containing 28 columns. The 'Requête' (Query) pane shows a SQL query selecting various fields from 'base_client' where 'etat_dossier' is 'complet' and 'date_demande' is greater than 2022. The 'Exporter' (Export) dialog box is open, showing the 'Format d'exportation' (Export format) set to 'CSV', the 'Sortie' (Output) set to 'Fichier' (File), and the 'Texte encodé exporté' (Exported text encoding) set to 'System'. The 'Exporter options de format' (Export format options) section shows 'Noms des colonnes dans la 1ère ligne' (Column names in the first line) checked, 'Séparateur de colonnes' (Column separator) set to '; (point virgule)' (semicolon), and 'Exporter la valeur NULL comme' (Export NULL value as) set to 'Chaine de caractères vide' (Empty string). The 'Barre d'état' (Status bar) at the bottom shows the query execution time: '[16:46:57] Requête terminée en 0.000 seconde(s)' and '[11:23:53] Requête terminée en 0.019 seconde(s)'.

Requête

```
1 select
2   , sexe
3   , date_naissance
4   , id_client
5   , enfant_conduite_accompagne
6   , nombre_enfants
7   , revenus
8   , valeur_residence_prin
9   , formation
10  , usage_vehicule
11  , type_vehicule
12  , est_rouge
13  , points_perdus
14  , age_vehicule
15  , type_conduite
16  , date_demande
17  , etat_dossier
18  , formule
19  , tarif_devis
20 from base_client
21 where etat_dossier = 'complet' and date_demande > 2022
```

Table

	sexe	date_naiss	id_client	enfant_conduite_acc
1	F	1999-09-25	871024631	0
2	F	1932-01-28	7945239	1
3	F	1978-09-15	713692930	0
4	F	2021-02-17	588888550	0
5	M	1956-06-14	69165559	0
6	F	1992-04-11	573751437	0
7	F	1941-03-18	294816298	2
8	M	1956-03-06	686791650	0

Barre d'état

[16:46:57] Requête terminée en 0.000 seconde(s).

[11:23:53] Requête terminée en 0.019 seconde(s).

DOCUMENTATION – Import CSV dans Excel

Affichage des données à charger et sélection des options (délimiteur, type de données...)

Permet de sélectionner le fichier à importer

The screenshot shows the 'Import Data' dialog box in Microsoft Excel. The dialog box is titled 'CRM 2022 complet.csv'. It has three main sections: 'Origine du fichier' (File origin) set to '1252: Europe de l'Ouest (Windows)', 'Délimiteur' (Delimiter) set to 'Point-virgule' (Comma), and 'Détection du type de données' (Data type detection) set to 'Selon les 200 premières lignes' (Based on the first 200 lines). Below these sections is a preview of the CSV data. The preview shows a table with columns: 'sexe', 'date_naissance', 'id_client', 'enfant_conduite_accompagne', 'nombre_enfants', 'revenus', 'valeur_residence_prin', 'formation', and 'usi'. The data is organized into rows, with the first row being a header. The background shows the Excel ribbon with the 'Données' tab selected. The 'Obtenir des données' (Get Data) button is highlighted in the ribbon, and a red arrow points to it from the text 'Permet de sélectionner le fichier à importer'. An orange arrow points from the text 'Affichage des données à charger et sélection des options (délimiteur, type de données...)' to the dialog box.

sexe	date_naissance	id_client	enfant_conduite_accompagne	nombre_enfants	revenus	valeur_residence_prin	formation	usi
F	25/09/1999	871024631	0	1	125301	0	Bachelors	Co
F	28/01/1932	7945239	1	1	50815	0	z_High School	Pri
F	15/09/1978	713692930	0	2	21204	0	z_High School	Pri
F	17/02/2021	588888550	0	0	0	106859	<High School	Pri
M	14/06/1956	69165559	0	0	null	207519	Bachelors	Co
F	11/04/1992	573751437	0	1	79606	225962	Bachelors	Pri
F	18/03/1941	294816298	2	2	130540	344195	Bachelors	Co
M	06/03/1956	686791659	0	0	51278	230340	Bachelors	Pri
M	18/02/1956	866093378	0	0	115044	311980	PhD	Pri
F	07/03/1974	401200262	0	0	null	252896	Bachelors	Co
F	13/02/1967	926073051	1	2	24060	0	Bachelors	Pri
M	14/05/2017	45830738	0	4	20822	76022	z_High School	Pri
F	20/12/1950	429992959	0	0	74426	0	PhD	Pri
M	26/07/2017	46606719	0	0	0	0	z_High School	Co
M	07/10/1945	389315422	0	1	38042	202967	z_High School	Pri
M	12/02/1938	184578161	0	0	45787	173196	Masters	Pri
F	21/03/1973	765577228	0	0	108152	329756	Masters	Pri
F	01/02/1936	106556153	0	0	53244	0	Bachelors	Co
F	04/04/1955	351243985	0	0	76954	249290	Bachelors	Pri
F	23/05/1940	243016778	0	0	43290	158771	Bachelors	Pri

DOCUMENTATION – Excel vers Power Query

Fenêtre permettant d'accéder à l'éditeur Power Query

Données
chargées
sous forme
de tableau

The screenshot displays the Microsoft Excel interface with the 'Création de tableau' (Create Table) ribbon active. A data table named 'CRM_2022_complet' is loaded, containing 29 rows and 11 columns. The columns are: sexe, date_naissance, id_client, enfant_conduite_accompagne, nombre_enfants, revenus, valeur_residence_prin, formation, usage_vehicule, type_vehicule, and est_rouge. The 'Requêtes et connexions' (Queries and Connections) pane is open on the right, showing the 'CRM_2022_complet' query with 1,158 rows. A context menu is open over the query, listing actions such as Copier, Coller, Modifier, Supprimer, Renommer, Actualiser, Charger dans..., Dupliquer, Référence, Fusionner, Ajouter, Exporter le fichier de connexion..., Déplacer vers le groupe, Déplacer vers le haut, Descendre, Afficher l'aperçu, and Propriétés....

	A	B	C	D	E	F	G	H	I	J	K
1	sexe	date_naissance	id_client	enfant_conduite_accompagne	nombre_enfants	revenus	valeur_residence_prin	formation	usage_vehicule	type_vehicule	est_rouge
2	F	25/09/1999	871024631	0	1	125301	0	Bachelors	Commercial	Sports Car	no
3	F	28/01/1932	7945239	1	1	50815	0	z_High School	Private	Minivan	no
4	F	15/09/1978	713692930	0	2	21204	0	z_High School	Private	Pickup	no
5	F	17/02/2021	588888550	0	0	0	106859	<High School	Private	z_SUV	no
6	M	14/06/1956	69165559	0	0	0	207519	Bachelors	Commercial	Panel Truck	no
7	F	11/04/1992	573751437	0	1	79606	225962	Bachelors	Private	z_SUV	no
8	F	18/03/1941	294816298	2	2	130540	344195	Bachelors	Commercial	Minivan	no
9	M	06/03/1956	686791659	0	0	51278	230340	Bachelors	Private	z_SUV	no
10	M	18/02/1956	866093378	0	0	115044	311980	PhD	Private	Minivan	no
11	F	07/03/1974	401200262	0	0	0	252896	Bachelors	Commercial	Panel Truck	yes
12	F	13/02/1967	926073051	1	2	24060	0	Bachelors	Private	Minivan	no
13	M	14/05/2017	45830738	0	4	20822	76022	z_High School	Private	z_SUV	no
14	F	20/12/1950	429992959	0	0	74426	0	PhD	Private	z_SUV	no
15	M	26/07/2017	46606719	0	0	0	0	z_High School	Commercial	Van	no
16	M	07/10/1945	389315422	0	1	38042	202967	z_High School	Private	Minivan	no
17	M	12/02/1938	184578161	0	0	45787	173196	Masters	Private	Minivan	no
18	F	21/03/1973	765577228	0	0	108152	329756	Masters	Private	z_SUV	no
19	F	01/02/1936	106556153	0	0	53244	0	Bachelors	Commercial	Van	yes
20	F	04/04/1955	351243985	0	0	76954	249290	Bachelors	Private	Minivan	yes
21	F	23/05/1940	243016778	0	0	43290	158771	Bachelors	Private	z_SUV	no
22	M	20/07/1988	411375726	0	1	67278	220556	Masters	Private	Minivan	no
23	M	02/12/1931	747049542	1	3	29431	143149	<High School	Private	Minivan	no
24	F	13/02/2000	582086109	0	0	0	147191	z_High School	Commercial	Van	yes
25	F	24/04/1941	93412915	0	3	48184	209188	z_High School	Commercial	Sports Car	no
26	M	16/12/2004	461573915	0	0	22059	123924	<High School	Commercial	Pickup	no
27	M	18/08/1957	931284804	0	0	0	0	z_High School	Private	Minivan	yes
28	F	06/12/1998	384628643	0	0	49780	0	<High School	Private	Sports Car	no
29	F	31/10/1978	922808394	0	0	76443	257889	Masters	Private	z_SUV	no

DOCUMENTATION – Éditeur Power Query

Ruban de commandes

Édition /
visualisation
de requêtes

En-tête,
permet
modification
de type et
filtrage des
données

Historique
de
requêtes

FichierAccueilTransformerAjouter une colonneAffichage

Fermer et chargerActualiser l'aperçuRequête

PropriétésÉditeur avancéGérer

Choisir les colonnesSupprimer les colonnesGérer les colonnes

Conserv. les lignesSupprimer les lignesRéduire les lignes

Trier

Fractionner la colonneRegrouper par

Type de données : DateUtiliser la première ligne pour les en-têtesRemplacer les valeursTransformer

Fusionner des requêtesAjouter des requêtesCombiner les fichiers

Gérer les paramètresParamètres

Paramètres de la source de donnéesSources de données

Nouvelle sourceSources récentesEntrer des donnéesNouvelle requête

Requêtes

= Table.TransformColumnTypes("#En-têtes promus",{{"sexe", type text}, {"date_naissance", type date}, {"id_client", Int64.Type}, {"enfant_conduite_accompagne", Int64.Type}, {"nombre_enfants", Int64.Type}, {"revenus", Int64.Type}, {"valeur_residence_prin", Int64.Type}, {"formation", type text}, {"usage_vehicule", type text}, {"type_vehicule", type text}, {"est_rouge", type text}, {"points_perdus", Int64.Type}, {"age_vehicule", type text}, {"type_conduite", type text}, {"date_demande", type datetime}, {"etat_dossier", type text}, {"formule", type text}, {"tarif_devis", type text}})

	sexe	date_naissance	id_client	enfant_conduite_accompagne	nombre_enfants	revenus	valeur_residence_prin	formation
1	F	12/02/1938	024631	0	1	125301	0	Bachelors
2	F	21/03/1973	945239	1	1	50815	0	z_High Scho
3	F	01/02/1936	692930	0	2	21204	0	z_High Scho
4	F	04/04/1955	888550	0	0	0	106859	<High Scho
5	M	23/05/1940	165559	0	0	null	207519	Bachelors
6	F	20/07/1988	751437	0	1	79606	225962	Bachelors
7	F	02/12/1931	816298	2	2	130540	344195	Bachelors
8	M	13/02/2000	791659	0	0	51278	230340	Bachelors
9	M	24/04/1941	093378	0	0	115044	311980	PhD
10	F	16/12/2004	200262	0	0	null	252896	Bachelors
11	F	18/12/1957	073051	1	2	24060	0	Bachelors
12	M		830738	0	4	20822	76022	z_High Scho
13	F		992959	0	0	74426	0	PhD
14	M		606719	0	0	0	0	z_High Scho
15	M		315422	0	1	38042	202967	z_High Scho
16	M		184578161	0	0	45787	173196	Masters
17	F		765577228	0	0	108152	329756	Masters
18	F		106556153	0	0	53244	0	Bachelors
19	F		351243985	0	0	76954	249290	Bachelors
20	F		243016778	0	0	43290	158771	Bachelors
21	M		411375726	0	1	67278	220556	Masters
22	M		747049542	1	3	29431	143149	<High Scho
23	F		582086109	0	0	null	147191	z_High Scho
24	F		93412915	0	3	48184	209188	z_High Scho
25	M		461573915	0	0	22059	123924	<High Scho
26	M		931784804	0	0	null	null	High Scho

Paramètres d'une requ...

PROPRIÉTÉS

NomCRM 2022 complet

Toutes les propriétés

ÉTAPES APPLIQUÉES

Source

En-têtes promus

Type modifié

18 COLONNES, 999+ LIGNES

Profilage de la colonne en fonction des 1000 premières lignes

APERÇU TÉLÉCHARGÉ À 100%

8

DOCUMENTATION – Anonymisation par randomisation*

Données brutes format nombre entier

Modification type de données (Vrai / Faux)

The screenshot shows a data transformation tool interface. On the left, a list of data types is displayed, including 'Nombre décimal', 'Devise', 'Nombre entier', 'Pourcentage', 'Date/Heure', 'Date', 'Heure', 'Date/Heure/Fuseau horaire', 'Durée', 'Texte', 'Vrai/Faux', and 'Binaire'. The 'Vrai/Faux' option is highlighted. On the right, a panel titled 'PROPRIÉTÉS' shows the 'Nom' field set to 'CRM 2022 complet'. Below this, a panel titled 'ÉTAPES APPLIQUÉES' lists the steps taken: 'Source', 'En-têtes promus', 'Type modifié', 'Colonnes permutées', 'Duplication de la colonne "enfant_conduite_accompagn...', and 'Colonne renommée "enfant_conduite_accompagne_Vra...'. The 'Type modifié' step is highlighted, and a sub-step 'Type modifié (Nbr entier en Vrai/Faux)' is shown below it.

enfant_conduite_accompagne	enfant_conduite_accompagne_Vrai/Faux
0	1.2
0	\$
0	1.2
0	%
0	Date/Heure
0	Date
0	Heure
0	Date/Heure/Fuseau horaire
0	Durée
1	A ^B C
0	Texte
0	Vrai/Faux
0	Binaire
0	Utilisation des paramètres régionaux...
0	FALSE
2	TRUE
0	FALSE
0	FALSE
0	FALSE
1	TRUE

* **RANDOMISATION** : modifier les attributs dans un jeu de données de telle sorte qu'elles soient moins précises, tout en conservant la répartition globale.
Exemple : transformer le nombre d'enfants en conduite accompagnée par une simple réponse Vrai / Faux

DOCUMENTATION – Anonymisation par généralisation*

Modification type de données (nombre entier)

Création colonne définissant modèle de regroupement désiré

The screenshot shows the CRM 2022 Power Query Editor interface. The main data table has columns: **tarif_devis**, **date_naissance**, **Transformer-Date-Age**, **Transformer-Durée-Total Années**, **Modification Type - Nbr Entier**, and **tranche_age**. The **tranche_age** column contains age ranges like "18-36", "90-108", "36-54", etc.

The formula bar at the top displays the following M code:

```
Table.AddColumn("#Type modifié1", "tranche_age", each let rangeSize = 18, offset = 0, inclusive = false, rangeIndex = Number.RoundDown((
```

On the right, the **Paramètres d'une requête** (Query Parameters) pane is visible, showing the query name "CRM 2022 complet" and a list of applied steps under **ÉTAPES APPLIQUÉES**:

- Source
- En-têtes promus
- Type modifié
- Colonnes permuées
- Duplication de la colonne "date_naissance"
- Colonne renommée "Transformer-Date-Age"
- Âge calculé
- Duplication de la colonne "Transformer-Date-Age"
- Colonne renommée "Transformer-Durée-Total Années"
- Total calculé des années
- Duplication de la colonne "Transformer-Durée-Total An..."
- Colonne renommée "Modification Type - Nbr Entier"
- Type modifié1
- Ajouter colonne à partir d'exemples - colonne renommée "tranche_age" - définir modèle tranche

Données transformées format âge total années (nombre décimal)

* **GÉNÉRALISATION** : modifier l'échelle des attributs des jeux de données, ou leur ordre de grandeur, afin de s'assurer qu'ils soient communs à un ensemble de personnes.
Exemple : transformer les dates de naissance des utilisateurs pour les regrouper par tranche d'âge

DOCUMENTATION – *Tableau final anonymisé*

Rappel des principes RGPD :

1 – Finalité :

Collecte des données dans un but bien précis, légal et légitime.

2 – Proportionnalité et pertinence :

Informations collectées doivent être pertinentes et strictement nécessaires au regard de la finalité.

3 – Durée de conservation :

Doit être fixée en fonction du type d'informations collectées et de la finalité du fichier, ne peut pas rester indéfinie.

4 – Sécurité et confidentialité :

Responsable doit garantir sécurité et confidentialité des informations collectées et s'assurer que seules les personnes autorisées aient accès à ces informations

Liste des actions réalisées pour la création du tableau final (ci-dessous) :

- . Les champs *métier / employeur / num_ss / groupe_sanguin / id_site_web / nom / email / adresse / lat / lon* n'ont pas été sélectionnés lors de la requête SQL afin d'empêcher l'identification directe des personnes présentes dans le fichier **(1-2)**
- . Un index, créé avec des entiers, remplace les valeurs du champ *id_client* (**colonne A**) afin d'éviter l'identification pour quelqu'un ayant accès aux données et connaissant les *id_client*. **(1-2-4)**
- . Le champ *date_de_naissance* a été modifié en âge puis en *tranche_âge* (**C**), bien plus anonyme **(2-4)**
- . Les champs *enfant_conduite_accompagne* (**D**) / *vehicule_est_rouge* (**K**) ont été modifiés en type booléen, permettant de conserver la répartition globale mais avec moins de précision **(2-4)**
- . Les champs *revenus* (**F**) / *valeur_residen_prin* (**G**) / *points_perdus* (**L**) / *age_vehicule* (**M**) ont été modifiés en tranches, pertinentes pour réaliser des analyses mais rendant le traçage, sur quelques valeurs, impossible **(1-2-4)**
- . Le champ *date_demande* (**O**) a été modifié au format date (sans heure) puis mis au premier de chaque mois, permettant de conserver la répartition globale mais avec moins de précision **(2-4)**
- . Le champ *tarif_devis* (**R**) a été modifié en entier, puis en tranche et enfin transformée en classe afin d'éviter de révéler la politique tarifaire de l'entreprise **(1-2-4)**

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	id_client	sexe	tranch	enfant_co	nor	tranche_rev	tranche_vale	formation	usage_vel	type_veh	vehic	tranch	tranche	type_conduite	date_demande	etat_de	formule	tarif_devis
2	1	F	18-36	FAUX	1	125000-150000	0-50000	Bachelors	Commercial	Sports Car	FAUX	0-3 pts	5-10 ans	Highly Urban/ Urban	01/02/2022	complet	dev_integral	A
3	2	F	90-108	VRAI	1	50000-75000	0-50000	High School	Private	Minivan	FAUX	0-3 pts	0-5 ans	Highly Urban/ Urban	01/05/2022	complet	dev_express	A
4	3	F	36-54	FAUX	2	0-25000	0-50000	High School	Private	Pickup	FAUX	0-3 pts	0-5 ans	Highly Rural/ Rural	01/01/2022	complet	dev_express	A
5	4	F	0-18	FAUX	0	0-25000	100000-150000	<High School	Private	SUV	FAUX	3-6 pts	5-10 ans	Highly Urban/ Urban	01/04/2022	complet	dev_integral	A
6	5	M	54-72	FAUX	0	N/C	200000-250000	Bachelors	Commercial	Panel Truck	FAUX	0-3 pts	10-15 ans	Highly Urban/ Urban	01/03/2022	complet	dev_vip	B

CONCLUSION



**Importance de la
collecte et du
traitement données**



**Exigence process
nécessaire même si
contraignante**



**Sécurité données
et confidentialité
indispensables**



**Permettra d'éviter
nouvelles sanctions
et/ou esclandres**