Rapport de traitement des données

|  |  |
| --- | --- |
| **Titre du projet** | *Mise en conformité RGPD* |

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Auteur** | **Description** | **Date** |
| *V1* | *MINE Marine* | *Rapport* | *08.01.2024* |

|  |
| --- |
| **Introduction** |
| *Ce projet est né suite à une plainte d’un client à la CNIL pour non-respect des règles du RGPD, notamment la conservation trop longue de données le concernant.*  *Suite à la sanction de la CNIL, consistant à une limitation du traitement des données, Dev’Immediat a besoin de se mettre en règles vis à vis du RGPD ainsi que de pouvoir travailler sur des données, pseudonymisées.* |

|  |
| --- |
| **Contenu du rapport :** |
| * **Préconisations** en lien avec les recommandations émises pour la mise en conformité au RGPD :   + Informations aux clients :     - Mise à disposition, via un lien hypertexte présent sur la le site, d’une politique RGPD. Elle doit être mise à disposition aux clients ainsi qu’aux visiteurs qui pourraient être invités à enregistrer des données personnelles.   + Sélection des données :     - Collecte des données strictement nécessaire au traitement de l’assurance automobile (création d’un devis, suivi du dossier)     - Interdiction stricte de traiter des données relatives à la santé (numéro de sécurité sociale, groupe sanguin). Suppression obligatoire.   + Fin de vie des données :     - Si le client n’est plus client, stocker les données dans une base de données séparée. Ne conserver que les données des clients en vigueur.     - S’il s’agit d’un prospect, stocker les données dans une base de données séparée passé un délai de 3 ans à compter du dernier contact.   + Sécurité des données :     - Sectoriser les droits d’accès à la base de données et réviser régulièrement les droits d’accès utilisateurs     - Relance cyclique à destination des clients afin de vérifier que les données stockées sont toujours en vigueur     - Mise en avant des données absurdes. Par exemple, certains clients ont une date de naissance erronée, les considérants comme mineurs.   + Anonymisation :     - En cas d’export de données à destination d’une personne n’ayant pas les mêmes droits d’accès, anonymiser et/ou pseudonymiser les données afin d’éviter toute identification directe ou indirecte.     - En cas d’archivage définitif, mettre en place des règles d’anonymisation et/ou de pseudonymisation * **Documentation** des traitements effectués :   + Rédaction de la requête SQL.     - Sélection des colonnes utiles uniquement, filtrage sur les dossiers complets et les demandes effectuées en 2022.   **SELECT** sexe, date\_naissance, ID\_client, enfant\_conduite\_accompagne, nombre\_enfants, usage\_vehicule, type\_vehicule, points\_perdus, age\_vehicule, type\_conduite, tarif\_devis  **FROM** base\_client  **WHERE** etat\_dossier = 'complet'  **AND** date\_demande LIKE '%2022%'   * + Accès à Power Query via Excel     - Ouverture du fichier CSV sur Excel via le bouton « A partir d’un fichier CSV/texte » présent dans le menu « Données »      * + - Vérifier le contenu de l’aperçu et charger les données      * + - Dans le menu « Données », le bouton « Requêtes et connexion » ouvre le volet à droite      * + - Power Query s’ouvre en double cliquant sur la requête présente sur le panneau de droite      * + - Faire attention à **ne pas supprimer ou déplacer le fichier CSV d’origine**, à défaut le lien serait cassé. Il faudra refaire le lien.   + Déterminer les colonnes à pseudonymiser.     - Anonymiser consiste à supprimer les données tandis que la pseudonymisation consiste à modifier les données. Le but est le même : ne pas permettre l’identification directe ou indirecte à une personne physique.     - **Sexe :** donnée binaire ne permettant pas d’identifier une personne     - **Date\_naissance** : données précise d’identification directe : à modifier en calculant l’âge et en faisant des regroupements par intervalle.     - **ID\_client** : Bien que ce soit un ID, par un croisement de données, il est possible en interne de retrouver la personne. Recréer un ID distinct sans table de correspondance     - **enfant\_conduite\_accompagne** : donnée personnelle de type familiale. Créer un résultat en oui/non.     - **nombre\_enfants** : donnée personnelle de type familiale. Créer un résultat en oui/non.     - **usage\_vehicule :** il ne s’agit pas d’une donnée personnelle et la donnée peut être utile pour le traitement des dossiers en assurance automobile.     - **type\_vehicule** : il ne s’agit pas d’une donnée personnelle et la donnée peut être utile pour le traitement des dossiers en assurance automobile.     - **points\_perdus :** il ne s’agit pas d’une donnée personnelle et la donnée peut être utile pour le traitement des dossiers en assurance automobile.     - **age\_vehicule :** Par l’âge du véhicule, il est possible de trouver, en croisant les données avec la carte grise, de trouver son propriétaire. On privilégiera des intervalles d’âge     - **type\_conduite :** il ne s’agit pas d’une donnée personnelle et la donnée peut être utile pour le traitement des dossiers en assurance automobile.     - **tarif\_devis :** il ne s’agit pas d’une donnée personnelle mais consigne de ne pas donner accès aux données financières. Il faudra regrouper les tarifs par intervalle de prix.   + Modification des données via Power Query     - **« >point en virgule »** : Conversion des textes en nombre. Excel et Power Query ne considèrent pas les valeurs chiffrées avec un point comme un nombre décimal, mais comme du texte. Il faut dans un premier temps remplacer les points par des virgules     - **« >texte en nombre »** : Les virgules en place, il faut donc expliquer à Power Query que ce sont dorénavant des nombres.     - **« +nouvel\_ID »** : création d’un nouvel ID pour remplacer l’ID enregistré dans le CRM     - **« +conduite accompagnee »** : afin de remplacer la valeur chiffrée du nombre d’enfant faisant la conduite accompagnée, mise en place d’une réponse en « oui/non »     - **« +enfants »** : afin de remplacer la valeur chiffrée du nombre d’enfant total, mise en place d’une réponse en « oui/non »     - **« +intervalle\_age\_vehicule »** : création d’intervalle d’âge pour les véhicules     - **« +age\_client »**: calcul de l’âge du client à partir de sa date de naissance     - **« +intervalle\_age\_client »** : création d’intervalles pour l’âge des clients   + Méthodologie     - **Remplacement de caractères**   Sélectionner les deux colonnes concernées : « age\_vehicule » et « tarif\_devis » puis cliquer sur « Remplacer les valeurs » dans le menu « Transformer ». Indiquer que l’on veut remplacer les points par des virgules    La formule se résume ainsi :     * + - **Conversion des textes en nombre**   Bien qu’étant composés de virgule, les valeurs restent néanmoins sous format de texte.  Il faut pour chaque colonne cliquer sur le petit symbole de format à gauche du nom de la colonne et choisir le nouveau format. Pour l’âge de la voiture, « nombre entier », pour le tarif, « nombre décimal ».    La formule se résume ainsi :     * + - **Création d’un nouvel ID**   Cliquer sur « Colonne d’index » dans le menu « Ajouter une colonne »    Choisir le type d’index. De préférence, commencer par 1 afin que le 0 n’ait pas d’impact lors d’un exercice de logique.  La formule se résume ainsi :     * + - **Colonne « oui/non »**   Cette méthodologie doit être appliquées pour les colonnes « enfant\_conduite\_accompagne » et « nombre\_enfants ».  Cliquer sur « Colonne conditionnelle » dans le menu « Ajouter une colonne » puis compléter selon les conditions souhaitées.    La formule se résume ainsi :       * + - **Calcul de l’âge**   Il n’y a pas de bouton spécifique à ce calcul, il faut l’écrire manuellement.  Pour ce faire, il faut dans un premier temps créer une nouvelle colonne « Colonne personnalisée » dans le menu « Ajouter une colonne » puis compléter la zone de formule    La formule se compose de plusieurs variables, caractérisé par la présence du « Let » afin de lister les variables et « In » afin de demander le calcul final   * « Dureejour » : nombre de jours entre la date de naissance et « aujourd’hui » * « Dureean » : nombre de jours converti en années   Le résultat est donc le nombre d’années, arrondi à l’entier.   * + - **Création d’intervalles**   Les intervalles sont à mettre en place pour l’âge des voitures ainsi que l’âge des clients que nous venons de calculer.  Il faut utiliser la colonne conditionnelle, comme vu pour les colonnes en « oui/non »  Comme il n’y a pas de condition « Entre » il faut veiller à ce que les conditions se faussent au fur et à mesure. Il faut commencer par les valeurs hautes puis finir par les petites valeurs.  IL faut également veiller à mettre des crochets ou d’autres caractères textes car un tiret entre les valeurs donnera un résultat calculé dans l’export CSV.    La formule se résume ainsi :       * + Export CSV     - Une fois toutes les manipulations de pseudonymisation réalisées, il faut maintenant exporter ces données.     - Pour ce faire, il suffit de cliquer sur « Enregistrer sous » et de choisir le format « CSV (séparateur : point-virgule) »   + CSV nettoyé et prêt à être communiqué     - Bien que l’exercice de pseudonymisation ait été réalisé, les colonnes d’origine sont toujours présentes     - Il faut donc ouvrir le fichier et supprimer toutes les colonnes contenant les données personnelles, à savoir :     - Date de naissance     - ID\_client     - Enfant\_conduite\_accompagne     - Nombre\_enfants     - Age\_vehicule     - Tarif\_devis     - Age\_client     - Puis enregistrer sous un autre nom.   Le fichier final est sous ce format peut être communiqué en l’état. |

|  |
| --- |
| **Conclusion** |
| *Une fois :*   * *La suppression des données sensibles et des données non pertinentes* * *L’archivage des données trop anciennes* * *La mise à jour des informations clients avec correction des aberrations* * *La gestion des droits d’accès* * *L’automatisation de l’anonymisation et la pseudonymisation des données avant communication à un service n’ayant pas les mêmes droits d’accès*   *La base de données du CRM sera conforme au RGPD et pourra faire l’objet d’un nouveau contrôle de la CNIL, qui, en cas d’avis favorable, lèvera sa sanction de limitation du traitement des données.* |