**Comment la bibliographie 5 a traité les données ?**

**Remarque :** les données préfiltrés de l’open Damir permettent de s’intéresser seulement au régime obligatoire de l’assurance maladie. Les indicateurs préfiltrés sont nuls pour les types de prestations (PRS\_REM\_TYP) non issues de l’assurance maladie

**Nature de l’assurance :**

Seule la valeur 10 a été gardé.

Une image contenant texte, capture d’écran, Police, nombre

Le contenu généré par l’IA peut être incorrect.

**Type d’enveloppe :**

Seules les types d’enveloppes 1, 2, 3, 9, 98 ont été gardées, pour ne considérer que le régime général.

Une image contenant texte, capture d’écran, Police, nombre

Le contenu généré par l’IA peut être incorrect.

**Bénéficiaires de la CMU**

BEN\_CMU\_TOP permet de faire un filtrage sur les non bénéficiaires de CMU (On y s’intéressera pour la partie 2 su projet)

TOP\_PS5\_TRG permet de faire un filtrage sur les non bénéficiaires de CMU-C (On y s’intéressera pour la partie 2 su projet)

**Suppression des données de date :** on peut supprimer les mois et l’année

**Type de remboursement :** seuls les remboursements de référence (code 0) et les remboursements inconnus sont gardées (code 99)

Une image contenant texte, capture d’écran, Police, nombre

Le contenu généré par l’IA peut être incorrect.

Les valeurs inconnues étant nombreuses pour cette variable, nous avons dû associer à chaque code acte (PRS\_NAT), une indicatrice prenant la valeur de « 1 » si l’acte est un complément, supplément ou majoration d’acte et « 0 » si c’est un acte de référence. Cette indicatrice a ensuite permis de mettre à 0 les quantités d’actes visées. (still didn’t do it)

**Montants négatifs**

Les lignes avec des montants ou quantités négatifs sont supprimés car il s’agit d’opérations de régularisation de la sécu sociale.

**Taux de remboursement impossible**

Les lignes tels que le taux de remboursement est plus grand que 100 % sont supprimés

**Recalculation des montants remboursés**

La variable FLT\_REM\_MNT est renommée MT\_RO (pour montant RO). Il a été observé que moins de 1 % des lignes présentaient un écart de plus de 1 € entre le montant RO renseigné et celui recalculé à partir du taux de remboursement PRS\_REM\_TAU et de la base de remboursement PRS\_REM\_BSE. Excepté les cas où cela est dû à une spécificité de renseignement des compléments d’acte, **les montants ont été corrigés pour retenir la valeur calculée**. La variable du taux de remboursement PRS\_REM\_TAU est ensuite supprimée car elle

**Traitement des données manquantes et aberrantes :**

**-** Pour les lignes présentant un âge, un genre, un code acte ou une région inconnue, elles sont aussi supprimées car considérées comme données manquantes ;

- Pour les valeurs de dénombrement manquantes, est prise la valeur de la quantité d’actes.

- Si les adhérents de la ligne présentent une qualité de bénéficiaires égale à « 2 » (conjoint et assimilé) ou « 4 » (autre ayant-droit) avec un âge supérieur à 19 ans, la qualité est remplacée par « 1 » (assuré) ; si la qualité est égale à « 4 » (autre ayant-droit) avec un âge inférieur à 19 ans, la qualité est remplacée par « 3 » (enfant); si la qualité est égale à « 9 » (valeur inconnue), la ligne est supprimée

**Suppression des variables**  
Dans l’étape de tarification, le choix des variables supprimés est le suivant :   
Une image contenant texte, capture d’écran, nombre, Parallèle

Le contenu généré par l’IA peut être incorrect.

Une image contenant texte, capture d’écran, Police, nombre

Le contenu généré par l’IA peut être incorrect.