La base de données BASEAUTO contient des informations relatives à des montants annuels de remboursement de sinistres suite à un ou plusieurs accidents automobiles.

Les variables qui caractérisent les assurés et leur véhicule sont les suivantes :

- PERMIS : ancienneté de permis (en mois)

- ACV : ancienneté du véhicule (en années) (10 : au moins 10 ans d’ancienneté)

- SEX : sexe du conducteur (M : Masculin, F : Féminin)

- STATUT : statut matrimonial (C : Célibataire, A : Autre)

- CSP : catégorie socioprofessionnelle (1 à 100 : codes de CSP)

- USAGE : usage du véhicule (1 : Privé, 2 : Privé-Trajet, 3 : Professionnel, 4 : Tournées)

- AGECOND : âge du conducteur (en années)

- K8000 : option kilométrage limité (N : Non, O : Oui)

- RM : bonus/malus (de 50 à 350)

- CAR : type de véhicule (ber : berline, brk: break, bus: bus, cab : cabriolet, cpe : coupe, cte : camionnette, mspf : monospace formule, msphf : monospace hors formule, tt : tout-terrain)

- CLA : classe SRA de prix du véhicule (de A à Z1 : du prix le plus bas au prix le plus élevé)

- ALI : type d'alimentation du véhicule (car : carburateur, elc : électrique, gpl : gpl, ids : injection directe suralimentée, inj : injection, ins : injection suralimentée)

- ENE : type d'énergie du véhicule (EL : électrique, ES : essence, GO : gasoil)

- VIT : vitesse maximum du véhicule (tranches en km/h, S220 : supérieur à 220 km/h)

- SEGM : segment du véhicule (0 : non renseigné, A : segment le plus bas, B, M1, M2, H : segment le plus élevé)

- GARAGE : garage (1 : Individuel clos, 2 : Clos collectif, 3 : Autre cas)

- CHARGE : charge de sinistres (en euros)