On veut étudier la santé perçue en fonction de l’activité professionnelle.

On ajuste sur

* Age
* Sexe
* Alcool
* Tabac
* Epice
* Csp
* Niveau d’éducation
* Sport/activité physique régulière
* Pénibilité
* Dépression / Stress
* Pathologies lourdes : cvd/cancer/TS

On commence par recoder les variables qui nous intéressent :

* On code la variable activpro en se basant sur Adm12a, Adm11, Adm12 et Adm10 (par ordre d’importance dans la prise de décision). La variable activpro aura 5 catégories :
  + T pour Travailleur
  + R pour Retraité
  + I pour Inactif
  + C pour Chômeurs
  + AR pour les individus indécidables
* On renomme sexe en H et F car plus lisible
* On normalise l’âge en passant au log