Type de garage

* Beaucoup de NA, 5621 sur 6614 réclamations.
* Peu de différences dans les histogrammes et le ggplot.
* 4/10

Source d’énergie

* Peut-être mieux de regarder avec la valeur du véhicule.
* Électrique seulement 2 données, pas de NA. Donc division entre diesel et régulier.
* 3/10

**Valeur du véhicule**

* Linéarité observée
* 9/10

Statut matrimonial

* Étrange d’avoir seulement autre et seul
* 4872 personnes en « autre »
* Légère tendance à la hausse pour « autre ».
* 5/10

**Catégorie sociale**

* Ggplot intéressante, recherche de plus à faire pour agglomérer les catégories et interpréter.
* 7/10

**Usage**

* Ggplot intéressant pour voyagement prof
* À voir si on agglomère certaines réponses.
* 7/10

**Âge**

* Linéarité
* 9,5/10
* Variable discrète à traiter en continu probablement.
* Modification annexe A : retirer les NA : *age\_conduct1 <- na.omit(MontantReclFr$age\_conduct)*

*mont\_recl1 <- MontantReclFr$mont\_recl[age\_conduct1]*

avant d’effectuer les calculs et graphiques à l’aide des nouvelles variables mont\_recl1 et age\_conduct1.

**Type de véhicule**

* Attention à la colinéarité
* 8/10

**Bonus malus**

* Travail à faire pour enlever les données qui ne sont pas l’intervalle : Modification annexe A : Retirer les deux valeurs de -5, qui ne font pas partie de l’intervalle de [50; 350[ mentionné pour les bonus (utiliser les variables modifiées bonus\_malus1 et mont\_recl1 dans le doc de Gabrielle) : *bonus\_malus1 <- MontantReclFr$bonus\_malus[MontantReclFr$bonus\_malus > -5]*

*mont\_recl1 <- MontantReclFr$mont\_recl[MontantReclFr$bonus\_malus > -5]*

* Rien entre 200 et 350, donc peut-être meilleur dossier de conduite.
* 9,75/10

**Âge du permis**

* Décision à prendre entre âge et âge du permis

**Âge du véhicule**

* 3 données de véhicule de plus de 150 ans.
* 6/10

**Genre**

* Plus d’homme que de femme.