Projet BE Santé & Industrie

Traitement des données manquantes :

1-Supprimer la variable ddn ( car variance = 0)

2-Supprimer les lignes des trois valeurs manquantes dans la variable diplome\_h —> 22 NA’s — 50% —> 1

3- diplome\_f ? au premier on essaye de les remplacer avec la valeur pertinente —> 3 NA’s — 45% -1 : 10% 0 : 45% 1

4- bmi\_h ? 22NA’s —> Moyenne ou median.

—> possible qu’on a des valeurs aberrantes. valeur > 50

5- age\_f ? —> Moyenne ou median , 3 NA’s —> Moyenne ou median

6- dconception ?

7- bh\_f ? 359 NA’s : 670 normal et 102 anormal

8- ct\_f ? 355 NA’s : 614 ovulation, 110 dysovulation, 52 anovulation

9- patho\_f ? 368 NA’s : 80% non