1. #remove recorded by column from data frame as it is not related

#due to same value for each instance

2. #remove scheme name due to multiple distinct values and unnecessary levels

3. #payment and payment\_type are having same values

#remove one column from them, payment removed

4. #remove feature quantity group as quantity and quantity group have same values

#keep one and remove other

5. #waterpoint\_type and wwaterpoint\_type\_group, have same values,

#remove waterpoint\_type\_group