DEVOIR_1_R_FACTORS

Leslye Patricia NKWA

CONTEXTE

Nous étudions l'impact des facteurs socio-économiques et de santé sur le niveau de stress des individus à Douala.

Variables choisies

```
ID (facteur - identifiant unique)
```

Âge (quantitative discrète - en années)

Sexe (qualitative nominale - "Homme" ou "Femme")

Revenu mensuel (quantitative continue - en FCFA)

Niveau d'éducation (qualitative ordinale - "Primaire", "Secondaire", "Supérieur")

Nombre d'enfants (quantitative discrète)

Heures de sommeil par nuit (quantitative continue)

Pratique d'un sport (qualitative binaire - "Oui" ou "Non")

Maladies chroniques (qualitative nominale - "Aucune", "Diabète", "Hypertension", etc.)

Niveau de stress (score de 1 à 10) (quantitative ordinale)

Code

```
# Fixons une graine pour rendre l'exemple reproductible
set.seed(123)
# (1) Facteur : Identifiant unique
ID <- factor(1:100) # Chaque individu a un ID unique
# (2) Quantitative discrète : Âge en années
Age <- sample(18:70, 100, replace = TRUE)
# (3) Qualitative nominale : Sexe (sans hiérarchie)
Sexe_qualitative <- sample(c("Homme", "Femme"), 100, replace = TRUE)
Sexe_qualitative
       [1] "Homme" "Homme" "Femme" "Femme" "Homme" "Femme" "Homme" "Femme"
    [10] "Homme" "Homme" "Femme" "Femme" "Homme" "Homme" "Homme" "Femme"
    [19] "Homme" "Femme" "
    [28] "Femme" "Femme" "Femme" "Femme" "Femme" "Femme" "Femme" "Femme" "Femme" "Femme"
    [37] "Femme" "Homme" "Femme" "Homme" "Homme" "Homme" "Homme" "Homme"
    [46] "Homme" "Homme" "Femme" "Femme" "Femme" "Homme" "Femme" "Homme"
    [55] "Homme" "Femme" "Homme" "Femme" "Femme" "Femme" "Homme" "Homme"
    [64] "Homme" "Femme" "Femme" "Femme" "Homme" "Homme" "Homme" "Homme"
    [73] "Femme" "Femme" "Homme" "Femme" "Femme" "Homme" "Homme" "Femme"
    [82] "Femme" "Homme" "Femme" "Femme" "Femme" "Femme" "Femme" "Femme"
    [91] "Homme" "Femme" "
 [100] "Homme"
Sexe <- factor(Sexe_qualitative,levels = c("Homme", "Femme"), ordered = TRUE )</pre>
# (4) Quantitative continue : Revenu (avec décimales)
Revenu \leftarrow round(rnorm(100, mean = 2500, sd = 800), 2)
# (5) Qualitative ordinale : Niveau d'éducation (Primaire < Secondaire < Supérieur)
NiveauEducation <- factor(</pre>
              sample(c("Primaire", "Secondaire", "Supérieur"), 100, replace = TRUE),
```

```
levels = c("Primaire", "Secondaire", "Supérieur"),
    ordered = TRUE
)
# (6) Quantitative discrète : Nombre d'enfants (valeurs entières)
NbEnfants <- sample(0:5, 100, replace = TRUE)</pre>
# (7) Quantitative continue : Heures de sommeil (mesurable en décimales)
Sommeil <- round(runif(100, 4, 9), 1)
# (8) Qualitative binaire ou nominale : Pratique du sport (Oui/Non)
Sport <- factor(sample(c("Oui", "Non"), 100, replace = TRUE))</pre>
# (9) Qualitative nominale : Maladies chroniques
MaladiesChroniques <- factor(</pre>
    sample(c("Aucune", "Diabète", "Hypertension", "Asthme"), 100, replace = TRUE))
# (10) Quantitative ordinale : Niveau de stress (note de 1 à 10)
Stress <- sample(1:10, 100, replace = TRUE)
# Création du DataFrame
df <- data.frame(ID, Age, Sexe_qualitative, Sexe, Revenu, NiveauEducation, NbEnfants, Sommei
head(df)
  ID Age Sexe_qualitative Sexe Revenu NiveauEducation NbEnfants Sommeil Sport
1 1 48
                    Homme Homme 2861.62
                                                                 5
                                                                       4.7
                                             Secondaire
                                                                             Non
2 2 32
                                                                       5.4
                    Homme Homme 2921.48
                                                Primaire
                                                                 3
                                                                             Oui
                                                                       6.9
3 3 68
                    Femme Femme 2315.79
                                               Primaire
                                                                 3
                                                                             Oui
                                                                 2
                                                                       4.8
                                                                             Oui
4 4 31
                    Femme Femme 3617.94
                                               Primaire
                    Femme Femme 3910.92
                                               Supérieur
                                                                 5
                                                                       4.5
                                                                             Non
                    Homme Homme 2888.48
                                               Supérieur
                                                                       6.1
                                                                             Non
                                                                 5
  MaladiesChroniques Stress
              Aucune
1
2
                          3
              Aucune
```

3

Aucune

3

```
4 Asthme 2
5 Hypertension 10
6 Hypertension 8
```

tail(df)

```
ID Age Sexe_qualitative Sexe Revenu NiveauEducation NbEnfants Sommeil
95
     95 40
                       Homme Homme 1625.13
                                                  Supérieur
                                                                     5
     96 52
                                                                     2
                                                                           5.4
96
                       Femme Femme 2653.95
                                                  Supérieur
97
     97 57
                       Femme Femme 2399.07
                                                 Secondaire
                                                                     1
                                                                           8.6
98
     98
       65
                       Femme Femme 1389.25
                                                  Supérieur
                                                                     2
                                                                           8.7
99
     99 47
                       Homme Homme 2875.92
                                                                     2
                                                  Supérieur
                                                                           6.7
                       Homme Homme 3268.31
100 100 29
                                                  Supérieur
                                                                     0
                                                                           5.1
    Sport MaladiesChroniques Stress
95
      Non
                      Asthme
96
      Non
                                   9
                     Diabète
97
      Oui
                     Diabète
                                   1
98
      Oui
                                  6
                      Aucune
99
      Non
                     Diabète
                                  9
100
                                  10
     Non
                      Aucune
```

Vérification des types de données str(df)

```
'data.frame': 100 obs. of 11 variables:
$ ID
                     : Factor w/ 100 levels "1","2","3","4",...: 1 2 3 4 5 6 7 8 9 10 ...
                     : int 48 32 68 31 20 59 67 60 54 69 ...
$ Age
                           "Homme" "Femme" "Femme" ...
$ Sexe_qualitative : chr
                    : Ord.factor w/ 2 levels "Homme"<"Femme": 1 1 2 2 2 1 2 1 2 1 ...
$ Revenu
                           2862 2921 2316 3618 3911 ...
$ NiveauEducation : Ord.factor w/ 3 levels "Primaire"<"Secondaire"<..: 2 1 1 1 3 3 1 1 3</pre>
                    : int 5 3 3 2 5 5 3 1 2 5 ...
$ NbEnfants
                    : num 4.7 5.4 6.9 4.8 4.5 6.1 5.8 8.2 5.3 4.1 ...
$ Sommeil
                    : Factor w/ 2 levels "Non", "Oui": 1 2 2 2 1 1 1 2 2 1 ...
$ Sport
$ MaladiesChroniques: Factor w/ 4 levels "Asthme", "Aucune", ..: 2 2 2 1 4 4 1 1 1 2 ...
                     : int 2 3 3 2 10 8 8 2 6 7 ...
$ Stress
```

Understand sample with replace = false by default with || sample(x= 13:15, size= 2, replace)