

# Exercice1: Comprendre la dynamique des populations

Visseho Adjiwanou, PhD.

20/09/2020

---

## NOTE

1. Cet exercice est à rendre au plus grand tard le **20 septembre**.
  2. Vous le soumettez sur Moodle en deux versions, une version RMarkdown et une version pdf.
  3. Avec cet exercice, vous assimilerez la manipulation des données avec base R
  4. Je ne vous ai pas encore expliqué ce que c'est qu'un fichier RMarkdown. Mais, cet exercice fait recours à cela. Un fichier .RMD comprend deux parties principalement: la partie où vous écrivez du texte (partie blanche) et la partie où vous écrivez vos commandes (partie grise). Dans la partie où vous écrivez du texte, vous pouvez ajouter des niveaux de titres (#). Si vous ne vous sentez pas en mesure d'utiliser ce fichier, alors, utiliser le fichier "script" pour faire l'exercice. Mais rappelez vous, pour faire des commentaires dans le script, vous devez les précéder de #.
- 

## Remarque importante:

Cette solution utilise plusieurs méthodes que nous n'avons pas encore vu mais que je vais discuter demain. Aussi, vous pouvez vous limiter à lire que les parties qu'on a déjà vues en classe. Après la séance de demain, tout deviendra clair pour vous.

---

Comprendre la dynamique des populations est important pour de nombreux domaines des sciences sociales. Nous allons calculer certaines quantités démographiques à base des naissances et des décès pour deux pays et la population mondiale pour deux périodes: 1950 à 1955 et 2005 à 2010. Nous analyserons les fichiers de données CSV suivants: **Kenya.csv**, **Sweden.csv** et **World.csv**. Chaque fichier contient des données démographiques pour le Kenya, la Suède (Sweden) et le monde (World), respectivement. Le tableau ci-dessous présente les noms et les descriptions des variables dans chaque ensemble de données.

Nom	Description
<b>country</b>	Nom du pays abrégé
<b>period</b>	Période au cours de laquelle les données sont collectées
<b>age</b>	Groupe d'âge
<b>births</b>	Nombre de naissance en milliers (i.e., nombre d'enfants nés des femmes du groupe d'âge)
<b>deaths</b>	Nombre de décès en millier
<b>py.men</b>	Personne-années pour les hommes en milliers
<b>py.women</b>	Personne-années pour les femmes en milliers

Source: United Nations, Department of Economic and Social Affairs, Population Division (2013). *World Population Prospects: The 2012 Revision, DVD Edition*.

Les données sont collectées pour une période de 5 ans, où *personne-années* est une mesure de la contribution temporelle de chaque personne au cours de la période. Par exemple, une personne qui vit pendant toute la

période de cinq ans contribue pour cinq personnes-années alors que celle qui ne vit que pendant la première moitié de la période contribue à 2,5 personnes-années. Avant de commencer cet exercice, il serait judicieux d'inspecter directement chaque ensemble de données. Dans R, cela peut être fait avec la fonction `View`, qui prend comme argument le nom d'un fichier (`data.frame`) à examiner. Alternativement, dans RStudio, double-cliquer sur le nom du fichier que vous venez d'ouvrir dans l'onglet **Environment** vous permettra de visualiser les données dans une vue de type tableur. Vous pouvez directement visualiser les données dans RMarkdown avec la commande `head`.

**Remarque:** Prenez toujours soins d'enregistrer les données à utiliser dans le même fichier que votre RMarkdown pour vous simplifier la vie.

## Mise en route

Ce code vous permet de charger directement les données. Vous devez télécharger les données de Moodle et les placer avec ce fichier Rmarkdown dans le même dossier (voir remarque en haut).

```
## Télécharger les données
Sweden <- read.csv("Sweden.csv")
Kenya <- read.csv("Kenya.csv")
World <- read.csv("World.csv")
```

## QUESTION 1: 10 points

Nous commençons par calculer *le taux brut de natalité* (TBN) pour une période donnée. Le TBN est défini comme suit:

$$\text{TBN} = \frac{\text{Nombre total de naissances}}{\text{Nombre total de personnes-années vécues}}$$

Calculez le TBN pour chaque période, séparément pour le Kenya, la Suède et le Monde. Commencez par calculer le nombre total de personnes-années par groupe d'âge, enregistré sous la forme d'une nouvelle variable dans chaque fichier de données, en faisant la somme des personnes-années pour les hommes et les femmes.

## Méthode 1: Vecteurs

```
head(Kenya)
```

```
##   country  period  age  births  deaths  py.men py.women    l_x
## 1    KEN 1950-1955  0-4   0.000 398.314 2982.589 2977.828 100000.00
## 2    KEN 1950-1955  5-9   0.000  36.799 1977.935 1969.698  75157.42
## 3    KEN 1950-1955 10-14  0.000  19.520 1634.561 1633.975  71364.50
## 4    KEN 1950-1955 15-19 264.185  18.508 1588.554 1564.652  69469.47
## 5    KEN 1950-1955 20-24 485.564  21.361 1427.824 1364.061  67421.15
## 6    KEN 1950-1955 25-29 383.252  20.424 1204.917 1105.817  64854.80
```

```
# Naissance totale par période
```

```
NTK_1950_1955 <- sum(Kenya$births[1:15])
```

```
NTK_2005_2010 <- sum(Kenya$births[16:30])
```

```
# Personne année totale par période
```

```
PAK_1950_1955 <- sum(Kenya$py.men[1:15]) + sum(Kenya$py.women[1:15])
```

```
PAK_2005_2010 <- sum(Kenya$py.men[16:30]) + sum(Kenya$py.women[16:30])
```

```
# Taux brut de natalité par période
```

```
TBN_K_1950_1955 <- (NTK_1950_1955 / PAK_1950_1955) * 1000
```

```
TBN_K_2005_2010 <- (NTK_2005_2010 / PAK_2005_2010) * 1000

# Taux brut de natalité dans un tableau
TBN_K <- c("1950-1955" = TBN_K_1950_1955, "2005-2010" = TBN_K_2005_2010)
TBN_K
```

```
## 1950-1955 2005-2010
## 52.09490 38.51507
```

**Remarques :** Vous ne connaissez peut-être pas le symbole \$. Mais, il permet de trouver les informations sur une variable. Donc, pour respecter entièrement la notation de base R, on écrirait: `NTK_1950_1955 <- sum(Kenya[1:15, "births"])`.

NTK = Naissance totale Kenya; PAK = Personnes années Kenya; TBN\_K = Taux brut de natalité Kenya. En utilisant les mêmes commandes sur les données du Sweden et du World, on obtient les taux bruts de natalité du Sweden et du World.

### TRES IMPORTANT: CE PROCÉDÉ EST CE QUI EST ATTENDU DE VOUS.

Mais, vous pouvez remarquer que nous avons écrit les mêmes commandes trois fois de suite, si ce n'est plus que cela, en comptant le nombre de fois que nous l'avons écrit aussi pour chaque pays. C'est dans ces cas qu'il est utile de faire recours aux fonctions.

#### Utilisation d'une fonction

```
NT <- function(data){
  NT_1 <- sum(data$births[1:15])
  NT_2 <- sum(data$births[16:30])
  # names(NT_1) <- "1950 - 1955"
  # names(NT_2) <- "2005 - 2010"
  c(NT_1, NT_2) # C() dit explicitement à fonction que nous voulons montrer les résultats
}
```

```
NT_Sweden <- NT(Sweden)
NT_Sweden
```

```
## [1] 548.069 546.474
```

```
NT_Kenya <- NT(Kenya)
NT_Kenya
```

```
## [1] 1672.493 7275.247
```

```
NT_World <- NT(World)
NT_World
```

```
## [1] 488891.5 674581.3
```

Sur la base de cet exemple, on peut créer une fonction pour calculer directement le taux brut de natalité.

```
TBN_x <- function(data){
  TBN_1 <- (sum(data$births[1:15]) / (sum(data$py.men[1:15]) + sum(data$py.women[1:15]))) * 1000
  TBN_2 <- (sum(data$births[16:30]) / (sum(data$py.men[16:30]) + sum(data$py.women[16:30]))) * 1000
  c(TBN_1, TBN_2)
}
```

```

TBN_x_Kenya <- TBN_x(Kenya)
TBN_x_Kenya

## [1] 52.09490 38.51507

TBN_x_Sweden <- TBN_x(Sweden)
TBN_x_Sweden

## [1] 15.39614 11.92554

TBN_x_World <- TBN_x(World)
TBN_x_World

## [1] 37.32863 20.21593

library(tidyverse)

## -- Attaching packages ----- tidyverse 1
## v ggplot2 3.2.1      v purrr 0.3.3
## v tibble 2.1.3       v dplyr 1.0.2
## v tidyr 1.0.2        v stringr 1.4.0
## v readr 1.3.1        v forcats 0.4.0

## Warning: package 'dplyr' was built under R version 3.6.2

## -- Conflicts ----- tidyverse_conflic
## x dplyr::filter() masks stats::filter()
## x dplyr::lag() masks stats::lag()

TBN_all <- bind_cols(Period = c("1950 - 1955", "2000 - 2005"), Kenya = TBN_x_Kenya, Sweden = TBN_x_Sweden, World = TBN_x_World)

TBN_all

## # A tibble: 2 x 4
##   Period      Kenya Sweden World
##   <chr>      <dbl> <dbl> <dbl>
## 1 1950 - 1955  52.1  15.4  37.3
## 2 2000 - 2005  38.5  11.9  20.2

```

## Méthodes 2: Base de données

En recourant à la fonction `summarise`, on peut calculer le nombre total de naissance et le nombre total de personnes-années directement.

```

# Naissance totale par période

Naissance_totale_periode <-
  Kenya %>%
  group_by(period) %>%
  summarise(nais_Kenya = sum(births))

## `summarise()` ungrouping output (override with `.groups` argument)

Naissance_totale_periode

## # A tibble: 2 x 2
##   period      nais_Kenya
##   <fct>      <dbl>

```

```
## 1 1950-1955      1672.
## 2 2005-2010      7275.
```

```
# Généralisation pour estimer tous les indicateurs
```

```
TBN_Kenya <-
  Kenya %>%
  group_by(period) %>%
  summarise(nais = sum(births),
            pa_total = sum(py.men) + sum(py.women),
            TBN = nais/pa_total*1000)
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
TBN_Kenya
```

```
## # A tibble: 2 x 4
##   period      nais pa_total  TBN
##   <fct>      <dbl>   <dbl> <dbl>
## 1 1950-1955  1672.   32105.  52.1
## 2 2005-2010  7275.  188894.  38.5
```

```
TBN_Sweden <-
  Sweden %>%
  group_by(period) %>%
  summarise(nais = sum(births),
            pa_total = sum(py.men) + sum(py.women),
            TBN = nais/pa_total*1000)
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
TBN_Sweden
```

```
## # A tibble: 2 x 4
##   period      nais pa_total  TBN
##   <fct>      <dbl>   <dbl> <dbl>
## 1 1950-1955   548.   35598.  15.4
## 2 2005-2010   546.   45824.  11.9
```

```
TBN_World <-
  World %>%
  group_by(period) %>%
  summarise(nais = sum(births),
            pa_total = sum(py.men) + sum(py.women),
            TBN = nais/pa_total*1000)
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
TBN_World
```

```
## # A tibble: 2 x 4
##   period      nais pa_total  TBN
##   <fct>      <dbl>   <dbl> <dbl>
## 1 1950-1955 488892. 13096959.  37.3
## 2 2005-2010 674581. 33368800.  20.2
```

```
TBN <- bind_cols(period = TBN_Sweden$period, Kenya = TBN_Kenya$TBN, Sweden = TBN_Sweden$TBN, World = TBN_World$TBN)
```

```
TBN
```

```
## # A tibble: 2 x 4
##   period    Kenya Sweden World
##   <fct>      <dbl>  <dbl> <dbl>
## 1 1950-1955  52.1   15.4  37.3
## 2 2005-2010  38.5   11.9  20.2
```

On peut essayer une procédé qui fonctionne une seule fois en combinant les données des trois zones ensemble et en appliquant la fonction `summarise`.

*# Combinaison des trois bases de données. Comme ce sont les mêmes variables dans les trois bases de données*

```
KSW <- bind_rows(Kenya, Sweden, World)
head(KSW)
```

```
##   country    period    age  births  deaths  py.men py.women    l_x
## 1    KEN 1950-1955  0-4   0.000 398.314 2982.589 2977.828 100000.00
## 2    KEN 1950-1955  5-9   0.000  36.799 1977.935 1969.698  75157.42
## 3    KEN 1950-1955 10-14  0.000  19.520 1634.561 1633.975  71364.50
## 4    KEN 1950-1955 15-19 264.185  18.508 1588.554 1564.652  69469.47
## 5    KEN 1950-1955 20-24 485.564  21.361 1427.824 1364.061  67421.15
## 6    KEN 1950-1955 25-29 383.252  20.424 1204.917 1105.817  64854.80
```

*#View(KSW)*

```
TBN1 <-
  KSW %>%
  group_by(country, period) %>%
  summarise(naissance = sum(births),
            personne_annee = sum(py.men) + sum(py.women),
            TBN = (naissance / personne_annee) * 1000,
            TBN_round = round((naissance / personne_annee) * 1000, 2)
  )
```

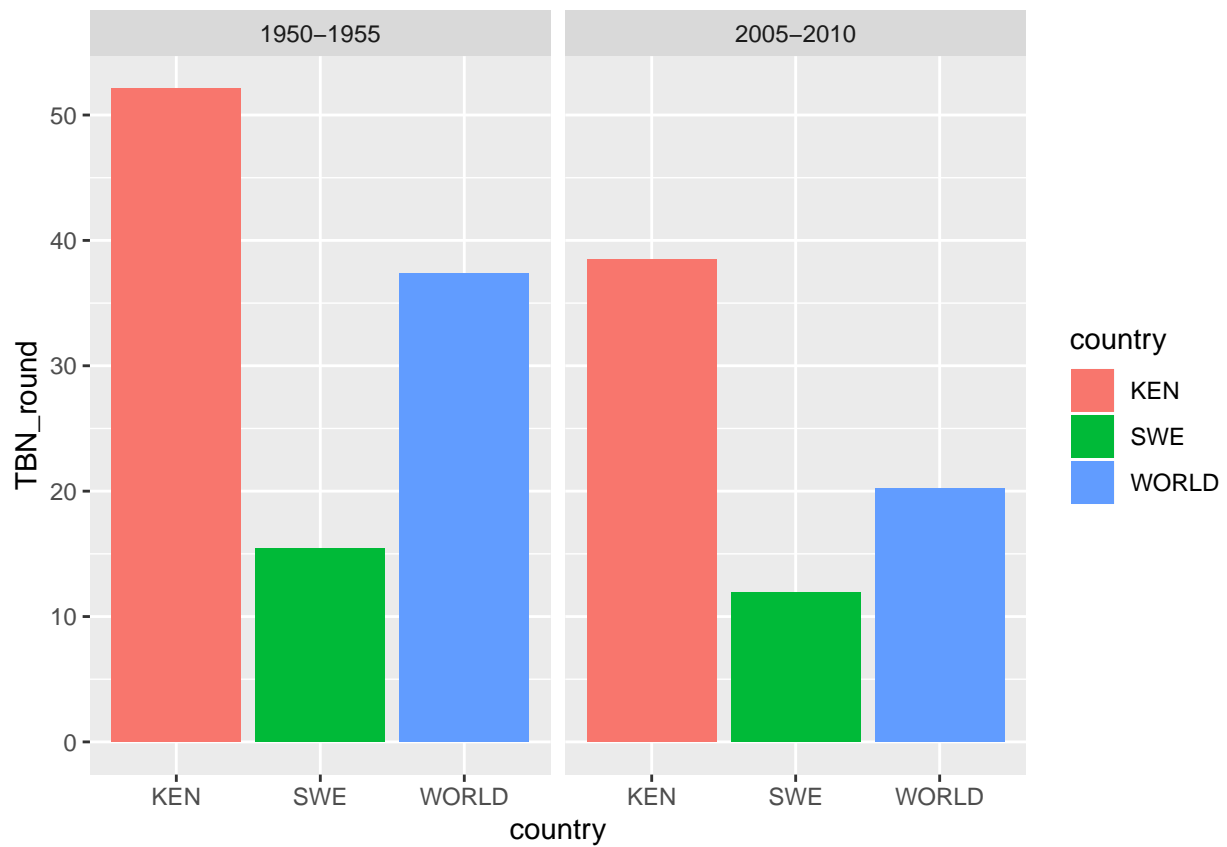
## `summarise()` regrouping output by 'country' (override with `.groups` argument)

TBN1

```
## # A tibble: 6 x 6
## # Groups:   country [3]
##   country period  naissance personne_annee  TBN TBN_round
##   <fct>    <fct>      <dbl>          <dbl> <dbl>    <dbl>
## 1 KEN      1950-1955    1672.         32105.  52.1     52.1
## 2 KEN      2005-2010    7275.        188894.  38.5     38.5
## 3 SWE      1950-1955     548.         35598.  15.4     15.4
## 4 SWE      2005-2010     546.         45824.  11.9     11.9
## 5 WORLD    1950-1955   488892.       13096959. 37.3     37.3
## 6 WORLD    2005-2010   674581.       33368800. 20.2     20.2
```

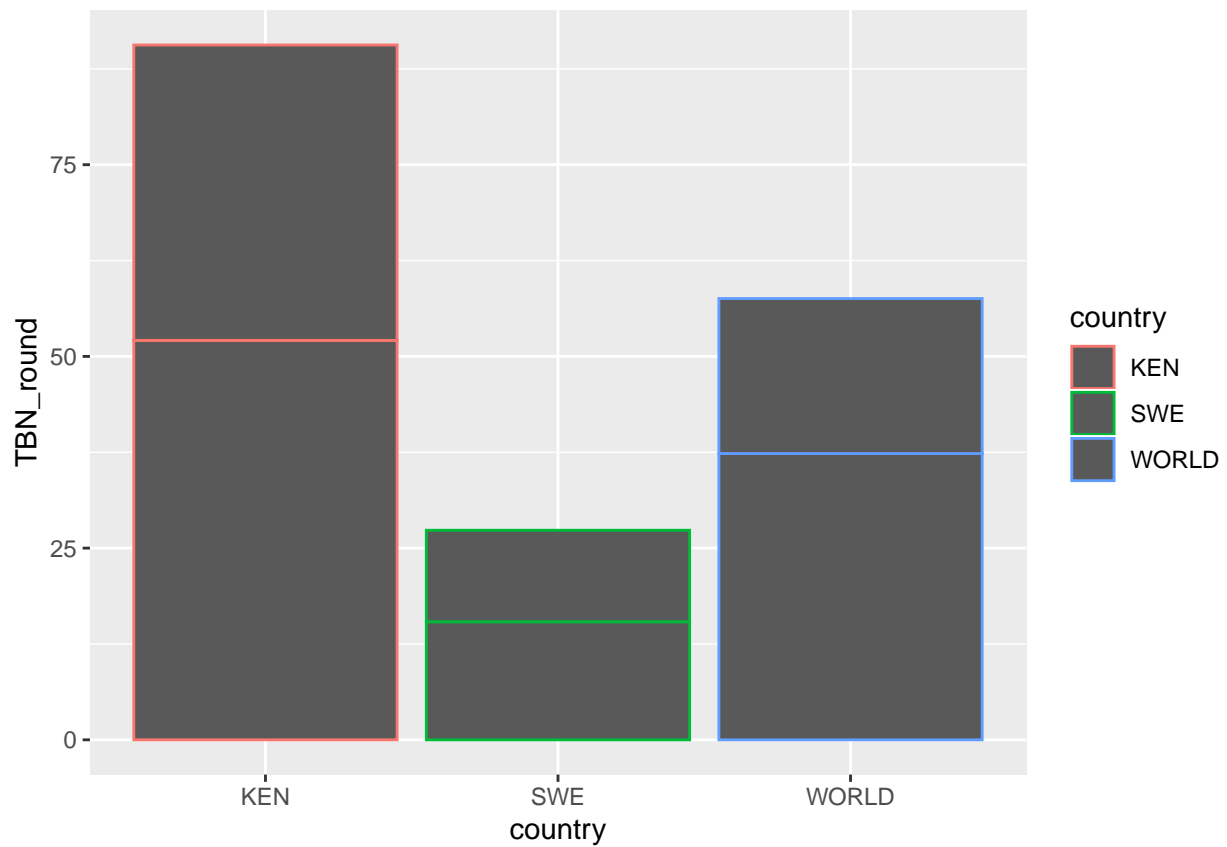
On peut représenter ces informations sous formes graphique pour une meilleure visualisation. Je ne m'attends pas qu'ils/elles me fassent ce graphique.

```
ggplot(TBN1) +
  geom_col(aes(x = country, y = TBN_round, fill = country)) +
  facet_grid(~period)
```



*# Mais ce graphique n'est pas idéal. On aimerait que les informations d'un même pays soit ensemble.*

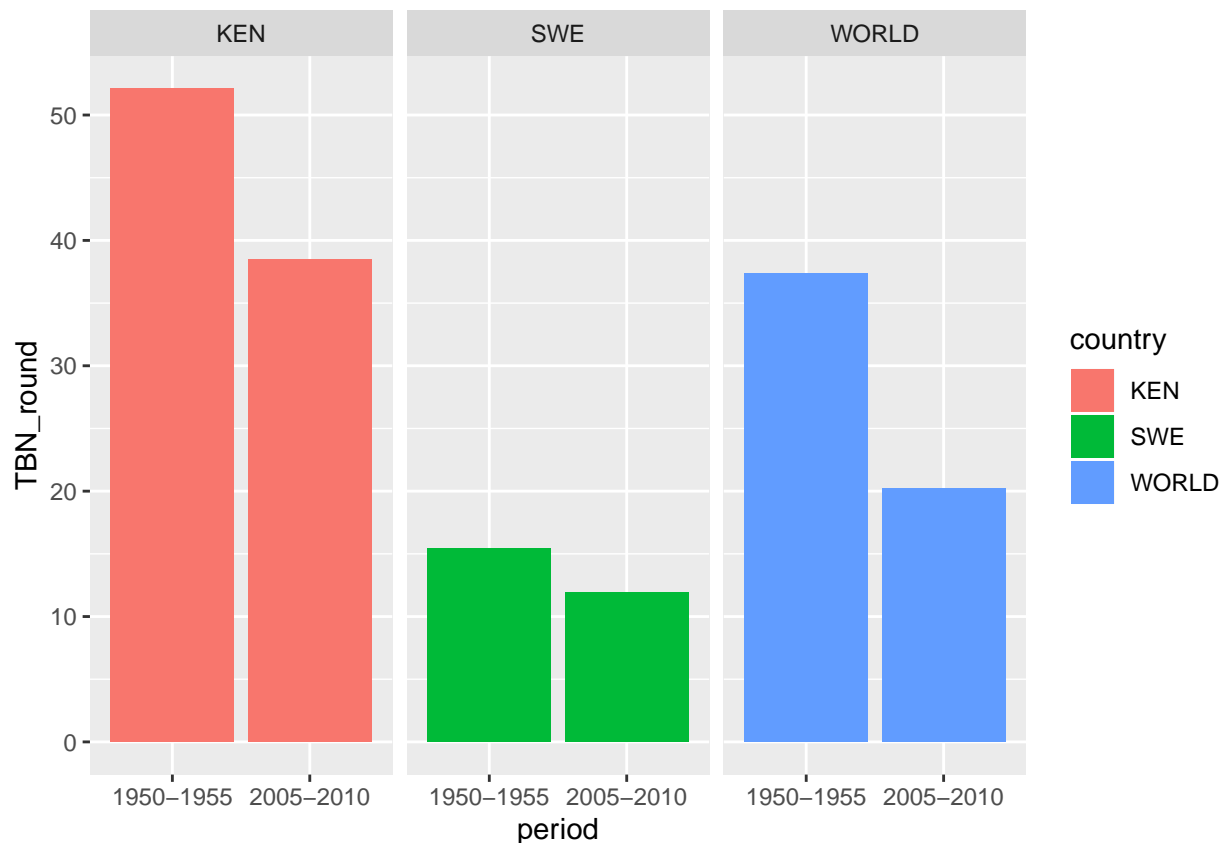
```
ggplot(TBN1) +  
  geom_col(aes(x = country, y = TBN_round, color = country))
```



*# Toujours pas idéal.*

```
ggplot(TBN1) +  
  geom_col(aes(x = period, y = TBN_round, fill = country)) +  
  facet_grid(~country)
```





## QUESTION 2: 5 points

Prenez un moment pour réfléchir à la manière d'organiser les résultats précédents.

Stockez les résultats du TBN sous forme de matrice de dimension (2, 3), 2 étant le nombre de période et 3 le nombre de pays/région. Décrivez brièvement les tendances observées dans les TBN calculés.

### Réponse 2

J'ai déjà répondu à ces questions aux lignes 148 et 202. Je les ramène juste ici pour montrer comment les résultats sont organisés.

TBN\_all *# Ligne 148*

```
## # A tibble: 2 x 4
##   Period      Kenya Sweden World
##   <chr>      <dbl>  <dbl> <dbl>
## 1 1950 - 1955  52.1   15.4  37.3
## 2 2000 - 2005  38.5   11.9  20.2
```

TBN *# Voir ligne 202*

```
## # A tibble: 2 x 4
##   period      Kenya Sweden World
##   <fct>      <dbl>  <dbl> <dbl>
## 1 1950-1955  52.1   15.4  37.3
## 2 2005-2010  38.5   11.9  20.2
```

## Votre interprétation

Nous observons qu'au cours de la période de 1950 à 1955, le Kenya avait un TBN bien supérieur à la moyenne mondiale, alors que la Suède était bien inférieure. En examinant les données de 2005 à 2010, nous constatons que les TBN des deux pays ont diminué par rapport à la première période. Cela semble être une tendance mondiale.

Note: La plupart des pays ont connu une baisse de la fécondité aux 20ème et 21ème siècles, mais à des rythmes différents. Ce phénomène s'appelle la transition démographique qui s'est développée selon un processus en quatre étapes. \* **Étape 1:** Mortalité (infantile) et fécondité très élevées \* **Étape 2:** Baisse de la mortalité mais fécondité toujours élevée. Période de forte croissance démographique. \* **Étape 3:** Baisse continue de la mortalité et baisse de la fécondité \* **Étape 4:** Nouvel équilibre avec des niveaux très faibles de mortalité et de fécondité.

Alors que les pays occidentaux sont à la dernière phase de la transition, la plupart des pays en développement sont encore en phase 2 ou trois.

## QUESTION 3: 5 points

Le TBN est facile à comprendre mais pose des problèmes de comparaisons. Pourquoi le taux brut de natalité n'est pas un bon indicateur de comparaison du niveau de fécondité entre pays?

### Réponse 3

- Mets au dénominateur les hommes et les femmes.
- Mets au dénominateur les femmes qui ne sont pas en âge de procréer.
- Ne tient pas compte de la structure par âge de la population. Dans une population vieillissante, on s'attend à trouver peu de naissance.

## QUESTION 4: 10 points

Une solution intermédiaire au TBN est le calcul du *taux général de fécondité (TGF)*. Celui-ci se calcule par:

$$\text{TGF} = \frac{\text{Nombre total de naissances}}{\text{Nombre total de personnes-années vécues par les femmes de 15-49 ans}}$$

Calculer ce taux pour le Kenya, la Suède et le Monde. Présenter les résultats aussi sous forme de matrice. Pourquoi cela vous paraît déjà meilleur que le TBN? Interpréter les résultats. Quel est le problème du taux général de fécondité.

### Réponse 4

je ne m'attends pas forcément qu'ils/elles créent une fonction pour faire le travail. Ils/elles peuvent procéder de la même manière qu'au 1.

```
TGF_x <- function(data){  
  TGF_1 <- (sum(data$births[4:10]) / sum(data$py.women[4:10])) * 1000  
  TGF_2 <- (sum(data$births[19:25]) / sum(data$py.women[19:25])) * 1000  
  c(TGF_1, TGF_2)  
}  
  
TGF_x_Kenya <- TGF_x(Kenya)  
TGF_x_Kenya
```

```
## [1] 234.5367 158.3425
TGF_x_Sweden <- TGF_x(Sweden)
TGF_x_Sweden

## [1] 62.8168 53.0123
TGF_x_World <- TGF_x(World)
TGF_x_World

## [1] 151.7275 77.7116
library(tidyverse)
TGF_all <- bind_cols(Period = c("1950 - 1955", "2000 - 2005"), Kenya = TGF_x_Kenya, Sweden = TGF_x_Sweden, World = TGF_x_World)
TGF_all

## # A tibble: 2 x 4
##   Period      Kenya Sweden World
##   <chr>      <dbl>  <dbl> <dbl>
## 1 1950 - 1955   235.    62.8 152.
## 2 2000 - 2005   158.    53.0 77.7
```

Votre interprétation

## QUESTION 5: 10 points

Pour remédier à ces problèmes, nous calculons l'*indice synthétique de fécondité* (ISF). Contrairement au TBN, l'ISF s'ajuste en fonction de la composition par âge de la population féminine en âge de procréer. Pour ce faire, nous devons d'abord calculer le *taux de fécondité par âge* (TFA), qui représente le taux de fécondité chez les femmes en âge de procréer [15, 50]. Le TFA pour la tranche d'âge  $[x, x + \delta)$ , où  $x$  est l'âge de départ et  $\delta$  la largeur de la tranche d'âge (mesurée en années), est défini comme suit:

$$\text{TFA}_{[x, x+\delta)} = \frac{\text{Nombre de naissances des femmes âgées de } [x, x + \delta)}{\text{Nombre de personnes-années vécues par les femmes âgées } [x, x + \delta)}$$

Notez que les crochets,  $[ , ]$ , incluent la limite alors que les parenthèses,  $( , )$ , l'excluent. Par exemple,  $[20, 25)$  représente la tranche d'âge supérieure ou égale à 20 ans et inférieure à 25 ans. Dans les données démographiques typiques, la tranche d'âge  $\delta$  est fixée à 5 ans.

Calculez le taux de fécondité par âge pour la Suède et le Kenya ainsi que pour le monde entier pour chacune des deux périodes. Stockez les TFA résultants séparément pour chaque région. Interpréter ces résultats. Comment devrez-vous présenter ces résultats pour une interprétation optimale?

## Réponse 5

```
## calcul du TFA

Sweden.TFA <- Sweden$births / Sweden$py.women
Kenya.TFA <- Kenya$births / Kenya$py.women
World.TFA <- World$births / World$py.women
Sweden.TFA

## [1] 0.0000000000 0.0000000000 0.0000000000 0.0389089519 0.1277108826
## [6] 0.1252436647 0.0873641591 0.0486037714 0.0162101857 0.0013418290
```

```
## [11] 0.0000000000 0.0000000000 0.0000000000 0.0000000000 0.0000000000
## [16] 0.0000000000 0.0000000000 0.0000000000 0.0059709097 0.0507320271
## [21] 0.1162085625 0.1322744621 0.0625923991 0.0121600765 0.0006143942
## [26] 0.0000000000 0.0000000000 0.0000000000 0.0000000000 0.0000000000

## Ajout des labels
names(Sweden.TFA) <- names(Kenya.TFA) <- names(World.TFA) <-
  Sweden$age

## resultats pour les femmes d'âges reproductifs [15, 50)
Sweden.TFA[4:10] # 1950-1955

##      15-19      20-24      25-29      30-34      35-39      40-44
## 0.038908952 0.127710883 0.125243665 0.087364159 0.048603771 0.016210186
##      45-49
## 0.001341829

Sweden.TFA[19:25] # 2005-2010

##      15-19      20-24      25-29      30-34      35-39      40-44
## 0.0059709097 0.0507320271 0.1162085625 0.1322744621 0.0625923991 0.0121600765
##      45-49
## 0.0006143942

Kenya.TFA[4:10]

##      15-19      20-24      25-29      30-34      35-39      40-44      45-49
## 0.16884585 0.35596942 0.34657814 0.28946367 0.20644016 0.11193267 0.03905205

Kenya.TFA[19:25]

##      15-19      20-24      25-29      30-34      35-39      40-44      45-49
## 0.10057087 0.23583536 0.23294721 0.18087964 0.13126805 0.05626214 0.03815044

World.TFA[4:10]

##      15-19      20-24      25-29      30-34      35-39      40-44      45-49
## 0.09029521 0.23763370 0.25245229 0.20416410 0.13810534 0.06360832 0.01519064

World.TFA[19:25]

##      15-19      20-24      25-29      30-34      35-39      40-44
## 0.048489719 0.151971307 0.146980966 0.093813813 0.046689639 0.016268995
##      45-49
## 0.004510245
```

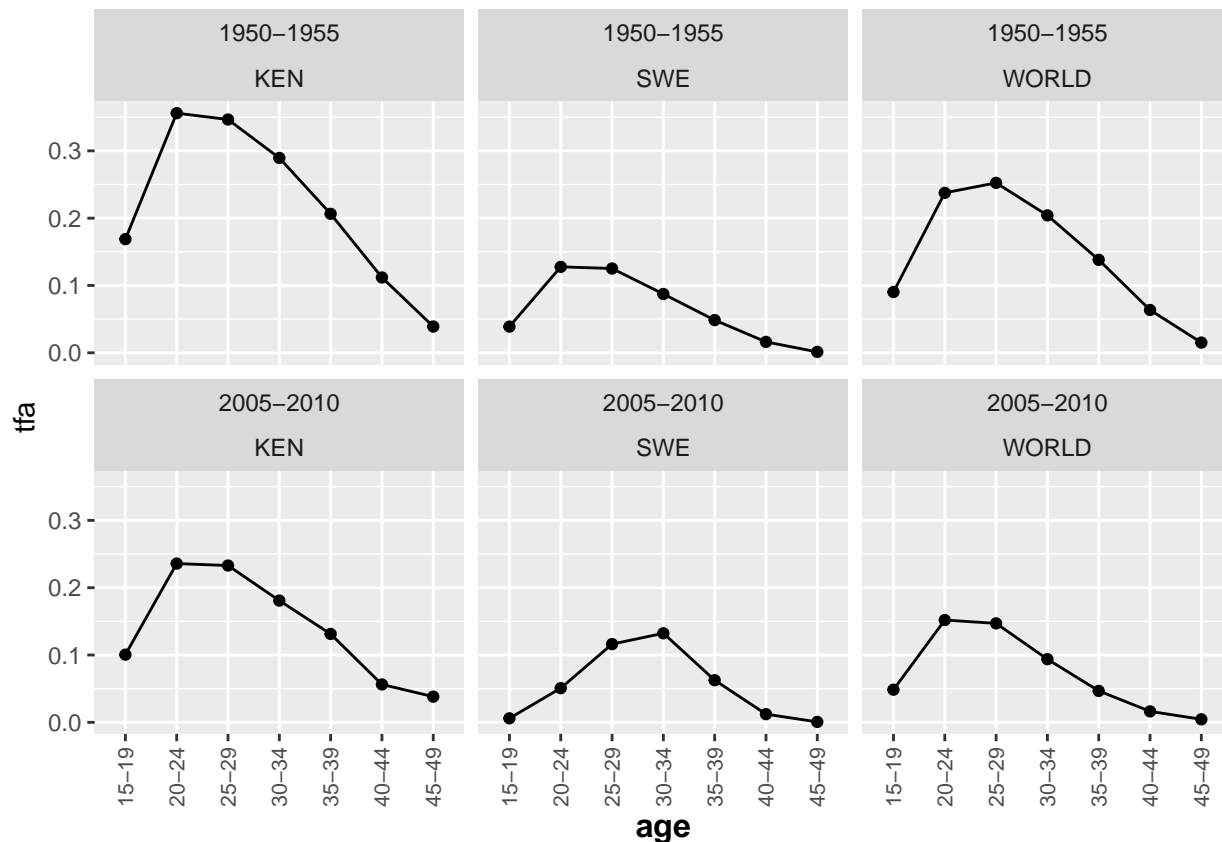
## Méthode 2

Il est plus simple de travailler avec la base de données KSW qui combine les données du Kenya, de la Suède et du Monde. Il s'agit de créer en premier lieu la variable TFA pour chaque groupe d'âges.

```
KSW <-
  KSW %>%
  filter(age != "0-4" & age != "5-9" & age != "10-14" & age != "50-54" & age != "55-59" & age != "60-69" & age != "65-69")
  mutate(tfa = births / py.women)

ggplot(KSW, aes(x = age, y = tfa)) +
  geom_line(aes(group = "period")) +
```

```
geom_point() +
facet_wrap(period~country) +
theme(axis.title.x = element_text(face="bold", size= 12),
      axis.text.x = element_text(angle=90, vjust=0.5, size = 8))
```



## Votre interprétation

De manière générale, l'ASFR a diminué au fil du temps pour les deux pays et dans le monde entier. L'ASFR pour la Suède et le Kenya était le plus élevé pour la tranche d'âge [20, 25 ans] en 1950-1955. Cependant, en 2005-2010, l'ASFR de la Suède était le plus élevé pour [30, 35), tandis que celui du Kenya était le plus élevé pour [20, 25). Ceci suggère que les femmes suédoises retardent la procréation jusqu'à un âge plus avancé par rapport aux femmes au Kenya. Dans le monde entier, l'ASFR était le plus élevé pour la tranche d'âge [20, 30 ans] au cours des deux périodes.

En conclusion, on observe des traits caractéristiques de différents niveaux de la transition démographique.

- Au Kenya, on observe une baisse généralisée de la fécondité qui touche tous les âges, en témoigne un affaissement des courbes de fécondité par âge au cours des deux périodes.
- Par contre en Suède, on observe un déplacement de la courbe vers la droite, témoignant d'un avancement de l'âge au premier mariage. Si l'âge à la première naissance augmente, il est presque certain que les femmes ne peuvent pas **retrouver** toutes les naissances reportées. D'où, une faible baisse de la fécondité que dans le cas du Kenya. Mais, encore une fois, sachez que la Suède est passé par le stage du Kenya et que le Kenya va se retrouver au stade de la Suède.

## QUESTION 6: 10 points

En utilisant le TFA, nous pouvons définir l'ISF comme le nombre moyen d'enfants que les femmes mettent au monde si elles vivent tout au long de leur vie reproductive.

$$\text{ISF} = \text{TFA}_{[15, 20)} \times 5 + \text{TFA}_{[20, 25)} \times 5 + \dots + \text{TFA}_{[45, 50)} \times 5$$

Nous multiplions chaque taux de fécondité par âge par 5 car la tranche d'âge est de 5 ans. Calculez l'ISF pour la Suède et le Kenya ainsi que le monde entier pour chacune des deux périodes. Comme dans la question précédente, continuer à supposer que la tranche d'âge en âge de procréer des femmes est de [15, 50). Stockez les deux ISF résultants pour chaque pays et le monde en tant que matrice de dimension (2, 3). En général, comment est-ce que l'indice synthétique de fécondité a-t-il évolué au cours de la période de 1950 à 2000 dans le monde? Commentez les résultats de la Suède et du Kenya.

### Réponse 6

```
## calcul du l'indice synthétique pour les femmes d'âges reproductifs [15, 50)

Sweden.ISF <- c(sum(Sweden.TFA[4:10] * 5), sum(Sweden.TFA[19:25] * 5))
Kenya.ISF <- c(sum(Kenya.TFA[4:10] * 5), sum(Kenya.TFA[19:25] * 5))
World.ISF <- c(sum(World.TFA[4:10] * 5), sum(World.TFA[19:25] * 5))

## Ajout des labels
names(Sweden.ISF) <- names(Kenya.ISF) <- names(World.ISF) <-
  c("1950-1955", "2005-2010")

## Résultats
Sweden.ISF

## 1950-1955 2005-2010
## 2.226917 1.902764

Kenya.ISF

## 1950-1955 2005-2010
## 7.591410 4.879568

World.ISF

## 1950-1955 2005-2010
## 5.007248 2.543623
```

### Méthode 2

```
isf <-
  KSW %>%
  group_by(country, period) %>%
  summarise(isf = round(sum(tfa)*5, 2))

## `summarise()` regrouping output by 'country' (override with `.groups` argument)
isf

## # A tibble: 6 x 3
## # Groups:   country [3]
```

##	country	period	isf
##	<fct>	<fct>	<dbl>
## 1	KEN	1950-1955	7.59
## 2	KEN	2005-2010	4.88
## 3	SWE	1950-1955	2.23
## 4	SWE	2005-2010	1.9
## 5	WORLD	1950-1955	5.01
## 6	WORLD	2005-2010	2.54

## Votre interprétation

Les ISF qui en résultent constituent une preuve supplémentaire de la réduction mondiale de la fécondité. Au Kenya, les femmes avaient en moyenne 7,6 enfants dans les années 50, puis ramenées à 4,9 en 2010. Il s'agit d'une baisse substantielle, mais le taux de fécondité du Kenya reste élevé par rapport au reste du monde. En Suède, nous trouvons que la fécondité a diminué sous le plafond de 2 enfants par femme au cours de la deuxième période. Cela suggère que sans l'immigration, la population suédoise pourrait commencer à diminuer.

Remarque: dans de nombreux pays développés, les femmes ont des enfants à un âge de plus en plus âgé et en ont moins. La croissance démographique mondiale est principalement due aux taux de fécondité relativement élevés dans les pays en développement.

---

BON DEVOIR

---

## ANNEXES

### 1. Glossaire

Fécondité: C'est l'étude des naissances

Fécondité = Fertility

Fécondité # Fertilité

Taux brut de natalité = Crude birth rate (CBR)

Taux global de fécondité générale

Indice synthétique de fécondité (ISF) = Total fertility rate (TFR)

Taux de fécondité par âge = Age specific fertility rate (ASFR)