

MAFE Methodological Note 10

Manipulation des bases biographiques

Application à la transnationalité des unions

R

Marc Thevenin (INED-SMS)

Octobre 2015



Funded under the
Socio-economic
Sciences & Humanities
Theme



TABLE DE MATIERES

INTRODUCTION	3
PARTIE I : Création de la base d'analyse en format individus/années	7
1) Base listant toutes les années de vie du répondant jusqu' à l'année de l'enquête	7
2) Base individus-années listant les pays de résidence du répondant	9
3) Base individus-années listant les pays de résidence des conjoints	13
a. Pays de résidence des partenaires	14
b. Les périodes de séparation du répondant	15
c. Les périodes d'union du répondant	16
d. Variables de statut matrimonial, correction de l'overlapping et comparaison des pays de résidence	17
PARTIE II. Exemples de création de variables d'analyse	24
1) Création d'une variable synthétique de statut matrimonial	24
2) Variable de comparaison des pays de résidence des répondants et des partenaires	25
a. Les structures d'union	25
b. La variable trans	27

INTRODUCTION

L'objectif de ce tutoriel est de donner quelques outils de manipulation des modules biographiques de l'enquête MAFE avec le logiciel R, en particulier lorsqu'il s'agit de mettre en œuvre une analyse biographique à temps discret. Concernant le langage peu de prérequis sont demandés, seule une connaissance générale sur la manipulation des données et des variables avec le logiciel est nécessaire.

Du fait de la complexité des données issues de l'enquête, il convient impérativement de consulter la documentation méthodologique présente sur le site mafeproject.com, en particulier le document décrivant toutes les bases mises à disposition (« [Introduction to the MAFE datasets](#) »).

Objectifs du tutoriel

- I. **Changement de niveau d'observation** : les modules biographiques de MAFE sont livrés en format individu-événement (ex : dans le fichier des unions, pour chaque individu, une ligne par union). Si ce type de format permet déjà de réaliser des analyses, la nécessité de combiner plusieurs modules ou l'utilisation de méthodes en temps discret demande de modifier le niveau d'observation. On va donc créer une base d'analyse qui donne, *pour chaque année*, des informations sur le pays de résidence du répondant et de son ou ses conjoints, ainsi que sa situation maritale. Le niveau d'observation passe donc de l'événement vécu à l'année vécue.

- II. **Création de « variables du moment »**: On cherchera à créer des variables synthétiques qui permettent de renseigner pour chaque instant d'observation les caractéristiques d'état d'un individu. Ces variables sont souvent appelée TVC pour « Time Varying Covariate » dans la littérature technique relative à ce type de problématique.
Par exemple, on cherchera à renseigner une variable donnant, pour chaque année, le statut matrimonial annuel du répondant et une information sur le caractère transnational des unions, en tenant compte des situations passées.

D'une base individus-événements à une base individus-périodes

En entrée : on dispose d'un fichier *individus-événements*. A l'exception de la base *general.dta*, tous les modules biographiques de MAFE sont livrés sous ce format avec une variable indiquant le rang lié à l'événement (*num_event*), l'année d'entrée dans l'événement (*qxxxd*) et l'année de sortie (*qxxxf*), ainsi que diverses caractéristiques propres à l'événement.

Dans la base *union* par exemple, le rang de l'union est donné par la variable **num_uni**, l'année de début de l'union par **q102d**, l'année de fin par **q102f**, le motif de rupture de l'union par **q102** et l'année de mariage par **q103**.

Si l'année de fin du dernier épisode est en valeur manquante (marquée NA sous R), cela indique que l'événement était toujours en cours au moment de l'enquête.

Base individu-événements

ident	num_event	qxxxd	qxxxf	VarX
SN000000	1	1970	1974	a
SN000000	2	2005	NA	b

En sortie : on souhaite disposer d'un fichier *individus-périodes* (ou plus précisément, *individus-années*). Après avoir calculé la durée de l'évènement, on va transformer la base pour à avoir une ligne par année.

Base individu-années

ident	year	num_event	Qxxxd	qxxxf	VarX
SN000000	1970	1	1970	1972	a
SN000000	1971	1	1970	1972	a
SN000000	1972	1	1970	1972	a
SN000000	2005	2	2005	2008	b
SN000000	2006	2	2005	2008	b
SN000000	2007	2	2005	2008	b
SN000000	2008	2	2005	2008	b

Ceci implique :

1. Une transformation des bases individus-événements en bases individus-périodes en tenant compte des trajectoires continues et des fin d'évènement l'année de l'enquête.
2. Une transformation, appelée souvent reshaping dans la littérature technique, des bases lorsque plusieurs événements de même type ont lieu en même temps (cas de « temporal overlapping »). Ici, le « temporal overlapping » est engendré par la présence d'unions polygames. Dans ce tutoriel, la polygamie est seulement traitée du point de vue d'un mari (le répondant) avec plusieurs épouses. Lorsque le répondant est une femme, bien qu'on dispose dans MAFE d'informations sur les épisodes de migration des autres co-épouses, on ne connaît pas leurs années d'entrée et de sortie de l'union.

Ressources mises à disposition

- **tuto_R.txt**: la totalité du programme décrit dans ce document (sauf les compléments qui figurent en annexe).
- Extraits des bases de données de MAFE au format Rdata : **general**, **migration**, **union** et **network**, pour 10 répondants et une sélection de variables. Les quelques variables catégorielles ne sont fournies que sous format numérique, les labels des codes sont explicités dans le document¹.

Documentation sur les bases

Avant toute manipulation des bases, il est fortement conseillé de consulter la documentation de l'enquête, surtout celle dédiée aux bases de données du document **presentation of the datasets**, et plus particulièrement pour ce tutoriel, la partie consacrée aux modules biographiques (à partir de la page :

http://nesstar.ined.fr/other_docs/IE0216A_PresentationDatasets.pdf

¹ Les captures des bases aux différentes étapes du programme sont issues de la version Stata.

- Base **general** : page 32-33.
- Base **house** : page 34.
- Base **migration** : page 36.
- Base **union** : pages 34.
- Base **network** pages 42-46.

Différences avec les tutoriels STATA et SAS

Par rapport aux versions Stata et Sas, la structure est sensiblement modifiée avec les variables de statut matrimonial, créées bien plus tôt dans le programme avant la correction de l'overlapping. Le programme n'utilise seulement de 2 packages externes : **epicalc** et **reshape**. Epicalc n'est plus installable directement à partir des entrepôts depuis mai 2015. Le package est néanmoins archivé à cette adresse (<https://cran.r-project.org/src/contrib/Archive/epicalc/>) ou disponible avec les données du tutoriel (répertoire zip « données d'apprentissage »). L'installation peut se faire directement dans le menu déroulant de la console R ou de R-Studio.

Capture des bases et valeurs manquantes

Les captures des bases aux différentes étapes du tutoriel sont celles réalisées sous Stata. Contrairement à Stata et Sas, les valeurs manquantes pour une variable numériques ne sont pas renseignées par un . (espace pour une variable caractère) avec R.

Avec R les valeurs manquantes sont marquées NA (Not Available), qui peut s'interpréter comme « une valeur qui n'est égale à aucune autre » (il existe également NaN « not a number »). Ne pas confondre avec NULL qui représente un objet vide.

Liste des instructions et fonctions utilisées

expand
ave
sum
cumsum
subset
paste
merge
ifelse
reshape

Les principales étapes

Informations sur le répondant

- 1) Création d'une base individu-année à partir de l'année de naissance jusqu'en 2008 (répondants)
- 2) Création d'une base individu-année pour les pays de résidence (répondants)

Informations sur le(s) conjoint(s)

- 3) Création d'une base individu-année pour les pays de résidence (conjoint(s))
- 4) Création d'une base individu-année pour les épisodes de séparation

- 5) Création d'une base individu-année pour les épisodes d'union
- 6) Appariement des épisodes d'union, de séparation et pays de résidence du(des) conjoint(s)
- 7) Création de variables dichotomique de statut matrimonial

Appariement des informations et correction de l'overlapping

- 8) Appariement des informations sur les conjoints et les pays de résidence du répondants
- 9) Correction de l'overlapping engendré par les unions simultanées
- 10) Comparaison des pays de résidence du répondant et du(des) conjoint(s)

Variables d'analyses

- 11) Création d'une variable synthétique de statut matrimonial
- 12) Création d'une variable identifiante des structures d'union
- 13) Création d'une variable synthétique de localisation de la structure d'union tenant compte du passé.

PARTIE I : Création de la base d'analyse en format individus/années

1) Base listant toutes les années de vie du répondant jusqu' à l'année de l'enquête

Objectif :

Créer une ligne d'observation par année de vie

1. Création d'une variable durée (entre l'année de naissance et 2008, année de l'enquête).
2. « Allongement » de la base pour chaque individu : les observations vont être dupliquées avec l'instruction `expand`.

On utilise ici la base *general*, dont la structure est classique, avec autant de lignes que de répondants. D'autres options sont possibles en utilisant directement les modules biographiques².

Base « general »

Variable name	Variable label
ident	Biographic individual identifier
q1a	Year of birth

Dans un premier temps une variable de durée entre 2008 (année de l'enquête) et l'année de naissance (variable **q1a**) est créée. On doit ajouter 1 à la différence entre les deux années pour obtenir le nombre exact d'années de la naissance à 2008.

On verra plus tard que cet ajout génère des doublons lorsque la transition entre deux évènements a lieu la même année (trajectoire continue). Ces doublons doivent faire l'objet d'une correction.

```
year<-general
year$fin<-2008
year$duree<-(year$fin-year$q1a+1)
```

² Se reporter à l'annexe du tutoriel Stata. La démarche est très proche: utilisation de la variable q1a dans n'importe quel module biographique en sélectionnant une seule observation par individu, ou rectification des doublons liés aux trajectoires continues si on utilise une base à trajectoires strictement continues comme *house* ou *activity* qui couvre toutes les années de vie après l'avoir allongé.

	ident	num_un~n	qla	a2008	dur
1.	E1	1	1972	2008	37
2.	E103	1	1978	2008	31
3.	E147	1	1954	2008	55
4.	F112	1	1951	2008	58
5.	F121	1	1972	2008	37
6.	F127	1	1964	2008	45
7.	I172	1	1971	2008	38
8.	S011202	1	1975	2008	34
9.	S079601	1	1936	2008	73

Pour illustrer, on a créé la variable a2008 qui est toujours égale à 2008

Pour transformer la base en individu-années (une ligne par année) on utilise la fonction `expand` du package **epicalc** appliqué sur la variable durée. La fonction `expand` va répliquer la ligne un nombre de fois égal à la valeur de **dur**, soit ici le nombre d'années d'existence du répondant jusqu'en 2008 compris.

```
library("epicalc")
```

Avant `expand`

id	duree
1	1
2	6

```
year <- expand(year, index.var = "duree")
```

Après `expand`

id	duree	
1	1	
2	6	
2	6	
2	6	
2	6	
2	6	
2	6	

Par individu (`year$ident`), la variable **age** est créée à partir de la variable `x` toujours égale à 1 avec l'instruction `ave`, à laquelle on associe la fonction `cumsum` (somme cumulée). L'instruction `ave` (pour `average`) permet d'obtenir le résultat d'une fonction pour chaque strate d'une variable (ici `ident`). On en déduit ensuite la variable `year`.

Pour initialiser les années à partir de l'année de naissance de l'individu on remonte d'un rang avec (1), soit l'âge de l'individu au 31 décembre de chaque année.

```

year$x<-1
year$age<-ave(year$x,year$ident,FUN=cumsum)-1
year$year <- year$qla+year$age
year<-subset(year,select = c("ident","age","year"))

```

Appariée aux modules transformés en format individu-années dont les trajectoires ne sont pas strictements continues, la base *year.dta* permet de récupérer toutes les années « sans évènement ».

```

+-----+
| ident  year  age |
+-----+
38. | E103  1978   0 |
39. | E103  1979   1 |
40. | E103  1980   2 |
41. | E103  1981   3 |
42. | E103  1982   4 |
43. | E103  1983   5 |
44. | E103  1984   6 |
45. | E103  1985   7 |
46. | E103  1986   8 |
47. | E103  1987   9 |
48. | E103  1988  10 |
49. | E103  1989  11 |
50. | E103  1990  12 |
51. | E103  1991  13 |
52. | E103  1992  14 |
53. | E103  1993  15 |
54. | E103  1994  16 |
55. | E103  1995  17 |
56. | E103  1996  18 |
57. | E103  1997  19 |
58. | E103  1998  20 |
59. | E103  1999  21 |
60. | E103  2000  22 |
61. | E103  2001  23 |
62. | E103  2002  24 |
63. | E103  2003  25 |
64. | E103  2004  26 |
65. | E103  2005  27 |
66. | E103  2006  28 |
67. | E103  2007  29 |
+-----+
68. | E103  2008  30 |
+-----+

```

2) Base individus-années listant les pays de résidence du répondant

La démarche présentée ici, qui utilise la base *migration*, n'est pas la plus rapide, mais elle exploite la variable *year*. Une alternative avec la base *house* est décrite en annexe des tutoriels Stata et SAS.

Objectifs :

1. Rectifier la dernière année d'un évènement qui se termine l'année de l'enquête (p1)
2. Supprimer les doublons générés par les trajectoires continues (p2)

Base « migration »

variable name	storage type	display format	value label	variable label
ident			Biographic individual identifier	
num_mig			Migration number	
q601d			Start year of the stay outside Senegal	
q601f			End year of the stay outside Senegal	
q602			Country of stay	

	ident	num_mig	q601d	q601f	q602	
1.	E1	1	1991	1992	NIGERIA	
2.	E1	2	2001	.	SPAIN	
3.	E103	1	1992	1995	SPAIN	
4.	E103	2	1997	.	SPAIN	
5.	E147	1	1994	1995	PORTUGAL	
6.	E147	2	1995	.	SPAIN	
7.	F112	1	1968	.	FRANCE	
8.	F121	1	1993	1995	FRANCE	
9.	F121	2	1998	2008	ITALY	(p2)
10.	F121	3	2008	.	FRANCE	
11.	F127	1	1989	1996	FRANCE	(p2)
12.	F127	2	1996	1997	UK	(p2)
13.	F127	3	1997	1997	PORTUGAL	(p2)
14.	F127	4	1997	1998	SERBIA & MTN	(p2)
15.	F127	5	1998	2000	HUNGARY	(p2)
16.	F127	6	2000	.	FRANCE	
17.	I172	1	1995	2004	ITALY	
18.	I172	2	2007	.	ITALY	
19.	S011202	1	2005	2008	SPAIN	(p1)
20.	S079601	1	1954	1956	INDONESIA	(p2)
21.	S079601	2	1956	1958	ALGERIA	
22.	S079601	3	1958	1961	MADAGASCAR	
23.	S079601	4	1961	1963	FRANCE	

Cas (p1) : Fin du dernier évènement l'année de l'enquête

Le temps étant discret avec comme unité l'année, le dernier évènement peut se terminer en 2008 avant la passation de l'entretien.

Si le dernier épisode de migration a eu lieu en 2008, on doit légèrement tronquer l'information pour créer la variable de durée en ramenant la fin de l'évènement à sa dernière année de survie au 31 décembre, soit l'année 2007. On génère tout d'abord la variable end_08 qui permet d'identifier ces situations et de ne pas tronquer une seconde fois ces observations lorsqu'on corrige p2.

Pour les épisodes qui n'étaient pas achevés au moment de l'entretien, on remplace la valeur manquante de **q601f** par 2008.

```
loc_resp<-migration
loc_resp$end_08[loc_resp$q601f==2008] <- 1
loc_resp$q601f[loc_resp$q601f==2008] <- 2007
loc_resp$q601f[is.na(loc_resp$q601f)] <- 2008
```

	ident	num_mig	q601d	q601f	q602	end_08
1.	E1	1	1991	1992	NIGERIA	.
2.	E1	2	2001	.	SPAIN	.
3.	E103	1	1992	1995	SPAIN	.
4.	E103	2	1997	.	SPAIN	.
5.	E147	1	1994	1995	PORTUGAL	.
6.	E147	2	1995	.	SPAIN	.
7.	F112	1	1968	.	FRANCE	.
8.	F121	1	1993	1995	FRANCE	.
9.	F121	2	1998	2007	ITALY	1
10.	F121	3	2008	.	FRANCE	.
11.	F127	1	1989	1996	FRANCE	.
12.	F127	2	1996	1997	UNITED KINGDOM	.
13.	F127	3	1997	1997	PORTUGAL	.
14.	F127	4	1997	1998	SERBIA AND MONTENEGRO	.
15.	F127	5	1998	2000	HUNGARY	.
16.	F127	6	2000	.	FRANCE	.
17.	I172	1	1995	2004	ITALY	.
18.	I172	2	2007	.	ITALY	.
19.	S011202	1	2005	2007	SPAIN	1
20.	S079601	1	1954	1956	INDONESIA	.
21.	S079601	2	1956	1958	ALGERIA	.
22.	S079601	3	1958	1961	MADAGASCAR	.
23.	S079601	4	1961	1963	FRANCE	.

On applique de nouveau la fonction `expand` après avoir créé la variable de durée. Au préalable on transforme la variable `duree` qui prend la valeur 0 si l'individu n'a jamais migré (ici il aura toujours vécu au Sénégal).

```
loc_resp$duree<-(loc_resp$q601f-loc_resp$q601d+1)
# no migration = duree=>0 #
loc_resp$duree[is.na(loc_resp$duree)] <- 0
loc_resp <- expand(loc_resp,index.var = "duree")
```

Pour la variable `year`, on doit au préalable créer un identifiant individu-épisode de migration (`id_mig`). Cet identifiant concatène avec la fonction `paste` les variables `num_mig` transformée en variable caractère et `ident`.

```
as.character(as.numeric(loc_resp$num_mig))
loc_resp$id_mig<- paste(loc_resp$ident,loc_resp$num_mig, sep=" ")
loc_resp$x<-1
loc_resp$cumx<-ave(loc_resp$x, loc_resp$id_mig,FUN=cumsum)-1
loc_resp$year <- loc_resp$q601d+loc_resp$cumx
```

Cas (p2) : Suppression des doublons générés par les trajectoires continues

Une trajectoire est considérée comme continue si le début d'un évènement a lieu la même année que la fin de l'évènement précédent³. Une même année se trouve donc inscrite sur deux lignes différentes : une fois dans la variable **q601f** en (t-1) et une fois dans la variable **q601d** en (t). L'allongement de la base avec la fonction `expand` provoque donc un doublon sur l'année.

Exemple pour `ident= S079601`

	ident	num_mig	q601d	q601f	q602	year
113.	S079601	1	1954	1956	INDONESIA	1954
114.	S079601	1	1954	1956	INDONESIA	1955
115.	S079601	1	1954	1956	INDONESIA	1956
116.	S079601	2	1956	1958	ALGERIA	1956
117.	S079601	2	1956	1958	ALGERIA	1957
118.	S079601	2	1956	1958	ALGERIA	1958
119.	S079601	3	1958	1961	MADAGASCAR	1958
120.	S079601	3	1958	1961	MADAGASCAR	1959
121.	S079601	3	1958	1961	MADAGASCAR	1960
122.	S079601	3	1958	1961	MADAGASCAR	1961
123.	S079601	4	1961	1963	FRANCE	1961
124.	S079601	4	1961	1963	FRANCE	1962
125.	S079601	4	1961	1963	FRANCE	1963

Ici on va de nouveau légèrement tronquer l'information en privilégiant l'année du nouvel évènement. On supprime donc le doublon généré par **q601f**. On considère donc que tous les évènements se sont terminés au 31 décembre de l'année précédente. Les deux conditions (`year != 2008` et `end_08 != NA`) évitent de tronquer une seconde fois l'information lorsque le dernier évènement s'est achevé en 2008 (p1).

```
loc_resp$d[loc_resp$q601f==loc_resp$year & loc_resp$year!=2008 &
is.na(loc_resp$end_08)]<-1
loc_resp <- loc_resp[ which(is.na(loc_resp$d)),]
loc_resp<-subset(loc_resp,select = c("ident","q602","year"))
```

	ident	num_mig	q601d	q601f	q602	year
106.	S079601	1	1954	1956	INDONESIA	1954
107.	S079601	1	1954	1956	INDONESIA	1955
108.	S079601	2	1956	1958	ALGERIA	1956
109.	S079601	2	1956	1958	ALGERIA	1957
110.	S079601	3	1958	1961	MADAGASCAR	1958
111.	S079601	3	1958	1961	MADAGASCAR	1959
112.	S079601	3	1958	1961	MADAGASCAR	1960
113.	S079601	4	1961	1963	FRANCE	1961
114.	S079601	4	1961	1963	FRANCE	1962

On remarque que lorsque la trajectoire continue a lieu en 2008 (F121), le problème a été réglé en amont lorsqu'on a remplacé la valeur de `q601f` par 2007.

³ Ce type de transition est systématique dans les bases *activity*, *house* et *network*, l'individu ayant toujours un statut d'activité et un pays de résidence.

Il ne reste plus alors qu'à appairer la base *migration* avec la base *year* créée plus haut, en triant préalablement les observations par l'identifiant individuel **ident** et la variable **year**.

```
loc_resp <- merge(year, loc_resp, by=c("ident", "year"), all.x=TRUE)
```

La variable renommée **loc_resp** (précédemment **q602**) a maintenant des valeurs manquantes pour les années n'appartenant pas à un épisode migratoire, c'est à dire les années où le répondant résidait au Sénégal. Il suffit de recoder la variable en remplaçant ces valeurs manquantes par le code pays du Sénégal (q602= 99341).

```
loc_resp$q602[is.na(loc_resp$q602)] <- 99341
```

	ident	year	age	q602
291.	S079601	1936	0	SENEGAL
292.	S079601	1937	1	SENEGAL
293.	S079601	1938	2	SENEGAL
294.	S079601	1939	3	SENEGAL
295.	S079601	1940	4	SENEGAL
296.	S079601	1941	5	SENEGAL
297.	S079601	1942	6	SENEGAL
298.	S079601	1943	7	SENEGAL
299.	S079601	1944	8	SENEGAL
300.	S079601	1945	9	SENEGAL
301.	S079601	1946	10	SENEGAL
302.	S079601	1947	11	SENEGAL
303.	S079601	1948	12	SENEGAL
304.	S079601	1949	13	SENEGAL
305.	S079601	1950	14	SENEGAL
306.	S079601	1951	15	SENEGAL
307.	S079601	1952	16	SENEGAL
308.	S079601	1953	17	SENEGAL
309.	S079601	1954	18	INDONESIA
310.	S079601	1955	19	INDONESIA
311.	S079601	1956	20	ALGERIA
312.	S079601	1957	21	ALGERIA
313.	S079601	1958	22	MADAGASCAR
314.	S079601	1959	23	MADAGASCAR
315.	S079601	1960	24	MADAGASCAR
316.	S079601	1961	25	FRANCE
317.	S079601	1962	26	FRANCE
318.	S079601	1963	27	SENEGAL
319.	S079601	1964	28	SENEGAL
320.	S079601	1965	29	SENEGAL
321.	S079601	1966	30	SENEGAL
.
.
Idem
.
.
363.	S079601	2008	72	SENEGAL

3) Base individus-années listant les pays de résidence des conjoints

Objectif

Transformation d'une base de données pour traiter les situations d'overlapping (polygamie) avec la fonction reshape

a. Pays de résidence des partenaires

Dans un premier temps on va récupérer les informations sur les pays de résidence des conjoints dans la base « network ». On fait comme dans la section précédente après avoir sélectionné les partenaires dans les personnes du réseau (**q901=1**).

Base *network.dta*

```
-----  
variable      variable label  
-----  
ident         Biographic individual identifier  
num_netmig    Number of migrant's migration  
q904d         Beginning year of the migration  
q904f         Ending year of the migration  
country       Country  
q901          Link with Ego  
id_uni        Union ID  
-----
```

```
+-----+  
| ident  num_netmig  q904d  q904f  country  q901  id_uni |  
+-----+  
1. | E103      1    1979      .    SPAIN  Partner  U_E1032 |  
2. | E103      1    1989      .    SPAIN  Partner  U_E1031 |  
3. | F112      1    1965      .    FRANCE Partner  U_F1121 |  
4. | I172      1    1986    1990    GUINEA Partner  U_I1722 |  
5. | I172      2    1990    1991    SENEGAL Partner  U_I1722 |  
6. | I172      3    1991      .    ITALY  Partner  U_I1722 |  
7. | S011202  1    2000    2002    ITALY  Partner  U_S0112021 |  
8. | S011202  2    2002      .    SPAIN  Partner  U_S0112021 |  
+-----+
```

Note : dans le tableau ci-dessus on a seulement sélectionné les informations relatives aux partenaires (q901=1)

```
net_conj<-network  
net_conj<-subset(net_conj, q901==1)  
net_conj$end_08[net_conj$q904f==2008] <- 1  
net_conj$q904f[net_conj$q904f==2008] <- 2007  
net_conj$q904f[is.na(net_conj$q904f)] <- 2008  
  
## expand ##  
net_conj$duree <- (net_conj$q904f-net_conj$q904d+1)  
net_conj$duree[is.na(net_conj$duree)] <- 0  
  
net_conj <- expand(net_conj,index.var = "duree")  
  
## variable year ##  
net_conj$id_unimig <- paste(net_conj$id_uni, net_conj$num_netmig, sep="")  
net_conj$x<-1  
net_conj$cumx <- ave(net_conj$x, net_conj$id_unimig,FUN=cumsum)-1  
net_conj$year <- net_conj$q904d+net_conj$cumx  
  
## continous trajectories ##
```

```

net_conj$d[net_conj$q904f==net_conj$year & net_conj$year!=2008 &
is.na(net_conj$end_08)]<-1
net_conj <- net_conj[ which(is.na(net_conj$d)),]

net_conj<-subset(net_conj,select = c("ident","country","id_uni","year"))

```

b. Les périodes de séparation du répondant

variable name	storage type	display format	value label	variable label
ident			Biographic individual identifier	
num_union			Number of the union	
id_uni			Union ID	
q102d			Start years of the relationship	
q102f			End year of the relationship	
q103			Type of dissolution	
q105			In which year did you get married	
q107			In which country was she/he borned	

Dans les informations concernant les unions, on dispose également des causes d'une éventuelle rupture. Après l'allongement de la base, ces informations ne sont reportées que sur les années où l'union était en cours.

	ident	year	q1031	num_union
155.	F127	1984	Separation or divorce	1
156.	F127	1985	Separation or divorce	1
157.	F127	1986	Separation or divorce	1
158.	F127	1987	Separation or divorce	1
159.	F127	1988	Separation or divorce	1

Pour résoudre ce problème, on va créer une base *separation* qui donnera la causes de la rupture à partir de l'année de la séparation et jusqu'en 2008. On procèdera comme auparavant pour mettre en forme cette base et on l'appariera de nouveau à la base union elle-même mise en forme.

	ident	year	q1031	num_union
155.	F127	1984	.	1
156.	F127	1985	.	1
157.	F127	1986	.	1
158.	F127	1987	.	1
159.	F127	1988	.	1
160.	F127	1989	Separation or divorce	.
161.	F127	1990	Separation or divorce	.
162.	F127	1991	Separation or divorce	.
163.	F127	1992	Separation or divorce	.
164.	F127	1993	Separation or divorce	.
165.	F127	1994	Separation or divorce	.
166.	F127	1995	Separation or divorce	.
167.	F127	1996	Separation or divorce	.
168.	F127	1997	Separation or divorce	.
169.	F127	1998	Separation or divorce	.
170.	F127	1999	Separation or divorce	.
171.	F127	2000	Separation or divorce	.
172.	F127	2001	Separation or divorce	.
173.	F127	2002	Separation or divorce	.
174.	F127	2003	Separation or divorce	.
175.	F127	2004	Separation or divorce	.

```

176. | F127 2005 Separation or divorce . |
177. | F127 2006 Separation or divorce . |
178. | F127 2007 Separation or divorce . |
179. | F127 2008 Separation or divorce . |
-----+

```

Pour créer la variable de durée de la base *separation*, on va utiliser *q102f* comme variable de début et 2008 pour marquer la fin.

```

separation<-union
separation <- subset(separation,select =
c("ident", "num_union", "id_uni", "q102f", "q103"))
separation <- na.omit(separation)

separation$duree <- (2008-separation$q102f+1)
library("epicalc")
separation <- expand(separation,index.var = "duree")
separation$x<-1

separation$cumx <- ave(separation$x, separation$id_uni ,FUN=cumsum)-1
separation$year <- separation$q102f+separation$cumx

separation<-subset(separation,select =
c("ident", "id_uni", "num_union", "year", "q103"))

```

c. Les périodes d'union du répondant

Avant de traiter le problème de l'overlapping, on transforme la base *union* en format individus-années, puis on l'apparie aux bases *separation* et *network*. Pendant l'opération, on doit créer une variable qui servira à contrôler la situation d'union et corriger l'information sur les années de mariages.

```

union<-subset(union,select = c("ident", "q102d", "q102f", "q107", "q105",
" id_uni", "num_union"))

union$end_08[union$q102f==2008] <- 1
union$q102f[union$q102f==2008] <- 2007
union$q102f[is.na(union$q102f)] <- 2008

```

La variable suivante *uni* (=1) permettra de contrôler si une année donnée, la personne est ou non en union.

```
union$uni <-1
```

```

# expand #
union$duree<-(union$q102f-union$q102d+1)
library("epicalc")
union <- expand(union,index.var = "duree")
union$x<-1
union$cumx<-ave(union$x, union$id_uni ,FUN=cumsum)-1
union$year <- union$q102d+union$cumx
union$x <- NULL

#correction trajectoires continues#
union$d[union$q102f==union$year & union$year!=2008 & is.na(union$end_08)]<-1
union <- union[ which(is.na(union$d)),]

```

Comme pour les séparations, l'année du mariage est reportée sur toutes les lignes même lorsque l'union n'était qu'informelle c'est-à-dire lorsque $q105 < \text{year}$. Après cette correction, la variable $q105$ est positionnée à 0 si la personne n'est pas mariée, 1 sinon.

```
union$M <- 0
union$M[union$q105<=union$year] <- 1
union$q105 <- NULL
union$q105 <- union$M
union$M <- NULL

union<-subset(union,select = c("ident", "year", "id_uni", "num_union", "q107",
"q105", "id_uni", "uni"))
```

On apparie enfin les bases *union*, *separation*, et *loc_conj* avec l'instruction *merge*. L'option *by* donne les clés d'identification de l'appariement (ici *ident* et *num_union*), *all=TRUE* permet d'apparier toutes les observations des 2 bases et pas seulement les communes. Enfin l'option *sort=TRUE* trie les bases d'origine en suivant la hiérarchie donnée par *by*.

```
loc_conj <- merge(union,separation, by=c("ident","num_union",
"id_uni","year"), all=TRUE, sort=TRUE)
loc_conj <- merge(loc_conj,net_conj, by=c("ident","id_uni","year"), all.x=TRUE,
all.y=FALSE, sort=TRUE)
```

A ce stade, le pays de résidence du partenaire (variable *country*) avant la première migration se trouve en valeur manquante si le conjoint n'est pas né dans un des pays européens où l'enquête a été réalisée⁴. On utilise la variable **q107** (pays de naissance du conjoint) de la base *union* pour récupérer cette information.

```
loc_conj$country[is.na(loc_conj$country)]<-0
loc_conj$q107[is.na(loc_conj$q107)]<-0

loc_conj$country[loc_conj$country==0 ] <- loc_conj$q107[loc_conj$country==0]

loc_conj <- subset(loc_conj, select = c("ident", "year", "num_union", "q103",
"q105", "uni", "q107", "country"))
```

d. Variables de statut matrimonial, comparaison de l'overlapping et comparaison des pays de résidence

L'**overlapping** s'observe lorsqu'au moins deux évènements de même type se déroulent simultanément. Une situation classique d'overlapping est, par exemple, lorsqu'un individu a plusieurs enfants vivants en même temps. Ici, il s'agit des unions polygames.

⁴ Par exemple dans MAFE Sénégal ceux qui ne sont pas nés en Espagne, France ou Italie (se reporter à la description du module *network* dans le document « *presentation of the dataset* »).

	ident	num_un~n	id_uni	q102d	q102f	q103	q105	q107	
1.	E1	1	U_E11	1990	.	.	1993	SENEGAL	
2.	E103	1	U_E1031	1992	1996	Separation	1992	SENEGAL	
3.	E103	2	U_E1032	2001	2008	Separation	2001	SPAIN	
4.	E147	1	U_E1471	1979	.	.	1979	SENEGAL	
5.	E147	2	U_E1472	1990	.	.	1990	SENEGAL	Overlapping à partir
6.	E147	3	U_E1473	2004	.	.	.	SENEGAL	de 1990
7.	F112	1	U_F1121	1967	.	.	1968	SENEGAL	
8.	F121	1	U_F1211	1998	.	.	1998	SENEGAL	
9.	I172	1	U_I1721	1985	1993	Deceased	1987	SENEGAL	
10.	I172	2	U_I1722	2002	.	.	2004	SENEGAL	
11.	S011202	1	U_S0112021	1999	.	.	1999	SENEGAL	
12.	S079601	1	U_S0796011	1965	.	.	1965	SENEGAL	
13.	S079601	2	U_S0796012	1977	.	.	1977	SENEGAL	Overlapping à partir
14.	S079601	3	U_S0796013	1984	1988	Deceased	1984	SENEGAL	de 1977

Dans l'exemple ci-dessous, après l'allongement de la base, les années renseignées par la variable **year** sont uniques jusqu'en 1989 (une union), doublées de 1990 à 2003 (deux unions simultanées), puis triplées de 2004 à l'année de l'enquête (trois unions simultanées).

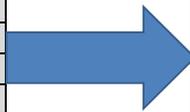
	ident	num_union	q103	q105	year	loc_conj
30.	E147	1	.	1979	1979	SENEGAL
31.	E147	1	.	1979	1980	SENEGAL
32.	E147	1	.	1979	1981	SENEGAL
33.	E147	1	.	1979	1982	SENEGAL
34.	E147	1	.	1979	1983	SENEGAL
35.	E147	1	.	1979	1984	SENEGAL
36.	E147	1	.	1979	1985	SENEGAL
37.	E147	1	.	1979	1986	SENEGAL
38.	E147	1	.	1979	1987	SENEGAL
39.	E147	1	.	1979	1988	SENEGAL
40.	E147	1	.	1979	1989	SENEGAL
41.	E147	1	.	1979	1990	SENEGAL
42.	E147	2	.	1990	1990	SENEGAL
43.	E147	1	.	1979	1991	SENEGAL
44.	E147	2	.	1990	1991	SENEGAL
45.	E147	1	.	1979	1992	SENEGAL
46.	E147	2	.	1990	1992	SENEGAL
47.	E147	1	.	1979	1993	SENEGAL
48.	E147	2	.	1990	1993	SENEGAL
49.	E147	1	.	1979	1994	SENEGAL
50.	E147	2	.	1990	1994	SENEGAL
51.	E147	1	.	1979	1995	SENEGAL
52.	E147	2	.	1990	1995	SENEGAL
53.	E147	1	.	1979	1996	SENEGAL
54.	E147	2	.	1990	1996	SENEGAL
55.	E147	1	.	1979	1997	SENEGAL
56.	E147	2	.	1990	1997	SENEGAL
57.	E147	1	.	1979	1998	SENEGAL
58.	E147	2	.	1990	1998	SENEGAL
.
.	.	Idem de 1999 à 2003
.
69.	E147	1	.	1979	2004	SENEGAL
70.	E147	2	.	1990	2004	SENEGAL
71.	E147	3	.	.	2004	SENEGAL
72.	E147	1	.	1979	2005	SENEGAL
73.	E147	2	.	1990	2005	SENEGAL
74.	E147	3	.	.	2005	SENEGAL
75.	E147	1	.	1979	2006	SENEGAL

76.	E147	2	.	1990	2006	SENEGAL
77.	E147	3	.	.	2006	SENEGAL
78.	E147	1	.	1979	2007	SENEGAL
79.	E147	2	.	1990	2007	SENEGAL
80.	E147	3	.	.	2007	SENEGAL
81.	E147	1	.	1979	2008	SENEGAL
82.	E147	2	.	1990	2008	SENEGAL
83.	E147	3	.	.	2008	SENEGAL

Pour résoudre ce problème, on va changer le format de la base en l'élargissant avec la fonction **reshape**. Les tableaux ci-dessous illustrent la démarche avec une variable propre à chaque évènement (X) qui ne varie pas au cours du temps (exemple : le sexe d'un individu). De (t1 à (t4) l'évènement est unique ; de (t5) à (t9), il s'observe 2 fois simultanément ; et en (t10) suite à la fin de l'évènement 1, il est à nouveau unique.

Avant reshape

t	num_event	X
t1	1	v1
t2	1	v1
t3	1	v1
t3	1	v1
t4	1	v1
t5	1	v1
t5	2	v2
t6	1	v1
t6	2	v2
t7	1	v1
t7	2	v2
t8	1	v1
t8	2	v2
t9	1	v1
t9	2	v2
t10	2	v2



Après reshape

t	num_event_1	X_1	num_event_2	X_2
t1	1	v1	NA	NA
t2	1	v1	NA	NA
t3	1	v1	NA	NA
t3	1	v1	NA	NA
t4	1	v1	NA	NA
t5	1	v1	2	v2
t6	1	v1	2	v2
t7	1	v1	2	v2
t8	1	v1	2	v2
t9	1	v1	2	v2
t10	NA	NA	2	v2

Avant cette transformation du format de la base va créer des variables de statut matrimonial.

Variables binaires de statut matrimonial

Les variables **V, D, U, M** indiquent si l'individu est –pour chaque année donnée, veuf, séparé ou divorcé, en union informelle, marié. La variable célibataire (C) est créée plus tard.

Du fait de l'overlapping, on a besoin d'un identifiant individu/année (id).

```
as.character(as.numeric(loc_conj$year))
loc_conj$id <- paste(loc_conj$ident,loc_conj$year, sep=" ")
```

On transforme les valeurs manquantes des variables uni, q103 et q105 en 0.

```
loc_conj$uni[is.na(loc_conj$uni)] <- 0
loc_conj$q103[is.na(loc_conj$q103)] <- 0
loc_conj$q105[is.na(loc_conj$q105)] <- 0
```

Les variables V et D sont créées à partir de la variable q103 et les variables U et M à partir de la variable q105. Pour chaque individu/année (id) on cherche à savoir si l'individu est dans l'une des situations une année donnée.

Rappel pour la syntaxe ifelse :

loc_conj\$x <- ifelse(loc_conj\$q103==2,1,0) se traduit par si q103=2 alors x=1 sinon 0.

Le code pour veuvage est q103=2

Le code pour séparation/divorce est q103=1

```
# V=widowed #
loc_conj$x      <- ifelse(loc_conj$q103==2,1,0)
loc_conj$cumx  <- ave(loc_conj$x, loc_conj$id ,FUN=sum)
loc_conj$V     <- ifelse(loc_conj$cumx>0,1,0)
#loc_conj$x    <- NULL
#loc_conj$cumx <- NULL

# D=divorced, separated #
loc_conj$x      <- ifelse(loc_conj$q103==1,1,0)
loc_conj$cumx  <- ave(loc_conj$x, loc_conj$id ,FUN=sum)
loc_conj$D     <- ifelse(loc_conj$cumx>0,1,0)
loc_conj$x     <- NULL
loc_conj$cumx  <- NULL

# U= free union#
loc_conj$x      <- ifelse(loc_conj$q105==0 & loc_conj$uni==1,1,0)
loc_conj$cumx  <- ave(loc_conj$x, loc_conj$id ,FUN=sum)
loc_conj$U     <- ifelse(loc_conj$cumx>0,1,0)

loc_conj$x     <- NULL
loc_conj$cumx  <- NULL

# M= marriage#
loc_conj$x      <- ifelse(loc_conj$q105==1 & loc_conj$uni==1,1,0)
loc_conj$cumx  <- ave(loc_conj$x, loc_conj$id ,FUN=sum)
loc_conj$M     <- ifelse(loc_conj$cumx>0,1,0)
loc_conj$x     <- NULL
loc_conj$cumx  <- NULL
```

Comparaison des pays de résidence et élargissement de la base avec la fonction reshape

Pour corriger ces situations d'overlapping on va privilégier ici la solution la plus simple, soit l'élargissement de la base sur chaque séquence d'union (**num_union**).

Cela aura pour conséquence de créer également des blocs de variables pour les individus dont les unions se sont succédées sans se chevaucher (période union1=> rupture union1 +période d'au moins un an sans union => période union2.....). C'est avec cette option que la suite du tutoriel sera développée.

Variables nbu et num_uni

Au préalable, deux variables vont être créées: la variable **nbu** qui renvoie pour chaque année le nombre total d'unions en cours, et **num_uni** identique à num_union (rang de l'union), mais seulement pour les années où elle n'a pas été rompue. nbu est générée en sommant les valeurs de uni.

```

# num_uni #
loc_conj$num_uni <- loc_conj$num_union
loc_conj$num_uni[loc_conj$uni==0] <- NA

#variable nbu#
loc_conj$nbu<-ave(loc_conj$uni,loc_conj$id, FUN=sum)

```

Appariement des bases loc_resp et loc_conj

On récupère les informations sur le pays de résidence du répondant, on crée une de nouveau une variable id dans la base loc_resp, puis on merge la base avec loc_conj (id comme clé d'appariement). On ne conserve que les âges du répondant supérieurs à 14.

La variable nbu est recodée 0 si valeur manquante, ce qui correspond à la période de célibat avant la première union.

```

## loc_rc: merge loc_conj & loc_resp ##

# variable id #
loc_resp$id<- paste(loc_resp$ident,loc_resp$year, sep=" ")

# merge #
loc_rc <- merge(loc_resp, loc_conj, by=c("id", "ident", "year"), all.x=TRUE,
all.y=FALSE, sort=TRUE)
loc_rc <- subset(loc_rc, age > 13)

loc_rc$nbu[is.na(loc_rc$nbu)]<-0

```

Comparaison des pays de résidence du répondant et du/des conjoints

La variable loc compare une année donnée le pays de résidence du répondant et du conjoint. Elle est égale à 1 s'ils vivent dans le même pays, 0 sinon.

Comme un répondant peu avoir plusieurs conjoint une même année, la variable tloc somme les valeurs de loc.

La variable uni_loc compare si le tloc, nombre d'union unifié une année donnée, est égale au nombre d'union en cours. S'il y a plusieurs unions simultanées, si tloc<nbu alors au moins un conjoint vi dans un pays différent.

```

loc_rc$loc <- ifelse(loc_rc$q602==loc_rc$country,1,0)
loc_rc$loc[is.na(loc_rc$num_uni)] <- 0
loc_rc$tloc <- ave(loc_rc$loc, loc_rc$id ,FUN=sum)
loc_rc$uni_loc <- ifelse(loc_rc$tloc==loc_rc$nbu,1,0)
loc_conj$loc <- NULL
loc_conj$tloc <- NULL

```

Elargissement de la base

En reprenant l'illustration précédente, avec le `reshape` on cherche le résultat suivant :

ident	Year	id	num_uni_1	var_1	num_uni_2	var_2	nbu
1	t1	1t1	1	v1	NA	NA	1
1	t2	1t2	1	v1	NA	NA	1
1	t3	1t3	1	v1	NA	NA	1
1	t3	1t4	1	v1	NA	NA	1
1	t4	1t4	1	v1	NA	NA	1
1	t5	1t5	1	v1	2	v2	2
1	t6	1t6	1	v1	2	v2	2
1	t7	1t7	1	v1	2	v2	2
1	t8	1t8	1	v1	2	v2	2
1	t9	1t9	1	v1	2	v2	2
1	t10	1t10	NA	.	2	v2	1

Les arguments de l'instruction `reshape` sont la base, la liste des variables dont la valeur n'est pas unique pour un identifiant donné (`v.name=`), l'identifiant (`idvar=`), la variable qui est à l'origine de l'overlapping (`timevar=`) et la direction du reshape (`direction=`).

```
library(reshape)

loc_rc <- subset(loc_rc,select = c("ident", "year", "age", "q602", "num_union",
"num_uni", "id", "country", "nbu", "V", "D", "U", "M", "uni_loc"))

analys <- reshape(loc_rc, v.names=c("country", "num_uni"), idvar=c("id"),
timevar="num_union", direction="wide")
analys$id <- NULL
```

Le résultat du `reshape` avec les variables de localisation des partenaires donne le résultat suivant:

	ident	year	age	q602	loc_conj1	num_uni1	loc_conj2	num_uni2
41.	E147	1968	14	SENEGAL
42.	E147	1969	15	SENEGAL
43.	E147	1970	16	SENEGAL
44.	E147	1971	17	SENEGAL
45.	E147	1972	18	SENEGAL
46.	E147	1973	19	SENEGAL
47.	E147	1974	20	SENEGAL
48.	E147	1975	21	SENEGAL
49.	E147	1976	22	SENEGAL
50.	E147	1977	23	SENEGAL
51.	E147	1978	24	SENEGAL
52.	E147	1979	25	SENEGAL	SENEGAL	1	.	.
53.	E147	1980	26	SENEGAL	SENEGAL	1	.	.
54.	E147	1981	27	SENEGAL	SENEGAL	1	.	.
55.	E147	1982	28	SENEGAL	SENEGAL	1	.	.
56.	E147	1983	29	SENEGAL	SENEGAL	1	.	.
57.	E147	1984	30	SENEGAL	SENEGAL	1	.	.
58.	E147	1985	31	SENEGAL	SENEGAL	1	.	.
59.	E147	1986	32	SENEGAL	SENEGAL	1	.	.
60.	E147	1987	33	SENEGAL	SENEGAL	1	.	.
61.	E147	1988	34	SENEGAL	SENEGAL	1	.	.
62.	E147	1989	35	SENEGAL	SENEGAL	1	.	.
63.	E147	1990	36	SENEGAL	SENEGAL	1	SENEGAL	2
64.	E147	1991	37	SENEGAL	SENEGAL	1	SENEGAL	2
65.	E147	1992	38	SENEGAL	SENEGAL	1	SENEGAL	2

66.	E147	1993	39	SENEGAL	SENEGAL	1	SENEGAL	2
67.	E147	1994	40	PORTUGAL	SENEGAL	1	SENEGAL	2
68.	E147	1995	41	SPAIN	SENEGAL	1	SENEGAL	2
69.	E147	1996	42	SPAIN	SENEGAL	1	SENEGAL	2
70.	E147	1997	43	SPAIN	SENEGAL	1	SENEGAL	2
71.	E147	1998	44	SPAIN	SENEGAL	1	SENEGAL	2
72.	E147	1999	45	SPAIN	SENEGAL	1	SENEGAL	2
73.	E147	2000	46	SPAIN	SENEGAL	1	SENEGAL	2
74.	E147	2001	47	SPAIN	SENEGAL	1	SENEGAL	2
75.	E147	2002	48	SPAIN	SENEGAL	1	SENEGAL	2
76.	E147	2003	49	SPAIN	SENEGAL	1	SENEGAL	2
77.	E147	2004	50	SPAIN	SENEGAL	1	SENEGAL	2
78.	E147	2005	51	SPAIN	SENEGAL	1	SENEGAL	2
79.	E147	2006	52	SPAIN	SENEGAL	1	SENEGAL	2
80.	E147	2007	53	SPAIN	SENEGAL	1	SENEGAL	2
81.	E147	2008	54	SPAIN	SENEGAL	1	SENEGAL	2

Note : seulement pour les deux premières unions du répondant. Les variable q1031, q1032, q1051 et q1052.

PARTIE II. Exemples de création de variables d'analyse

L'objectif est de générer des variables de type TVC (Time Varying Covariate) sur le statut matrimonial du répondant et sur les localisations (pays de résidence) des différents partenaires en distinguant les unions « transnationale », « réunifiée » et « toujours unifiée ». L'intérêt des deux dernières modalités est que les valeurs dépendent de celles prises dans le passé.

1) Création d'une variable synthétique de statut matrimonial

On peut créer la variable C correspondant à une situation de célibat. On en profite pour recoder les variables V,S,U et M.

```

analys$D[is.na(analys$D)] <- 0
analys$V[is.na(analys$V)] <- 0
analys$U[is.na(analys$U)] <- 0
analys$M[is.na(analys$M)] <- 0

analys$C <- ifelse(analys$nbu==0 & analys$D==0 & analys$V==0 ,1,0)

```

A partir des 5 indicatrices, on peut facilement synthétiser l'information avec la variable catégorielle **marital**, qui chaque année donne une information sur le statut matrimonial du répondant. Le choix des modalités est une possibilité parmi d'autres.

```

analys$mar[analys$C==1] <- "Alone, never in union"
analys$mar[analys$D==1 & analys$nbu==0] <- "Alone, separated"
analys$mar[analys$V==1 & analys$nbu==0] <- "Alone, widowed"
analys$mar[analys$U==1 & analys$nbu==1] <- "Monogamy, informal union"
analys$mar[analys$M==1 & analys$nbu==1] <- "Monogamy, married"
analys$mar[analys$U==1 & analys$nbu>1 & analys$M==0] <- "Polygamy, only informal unions"
analys$mar[analys$M==1 & analys$nbu>1] <- "Polygamy, at least one marriage"
table(analys$mar)

table(analys$marital)

```

	ident	year	num_unil	q1031	q1051	nbu	C	V	D	U	M	marital
82.	F112	1965	1	0	0	0	0	Alone, never in union
83.	F112	1966	1	0	0	0	0	Alone, never in union
84.	F112	1967	1	.	.	1	0	0	0	1	0	Monogamy, informal union
85.	F112	1968	1	.	Marriage	1	0	0	0	0	1	Monogamy, married
86.	F112	1969	1	.	Marriage	1	0	0	0	0	1	Monogamy, married
87.	F112	1970	1	.	Marriage	1	0	0	0	0	1	Monogamy, married
88.	F112	1971	1	.	Marriage	1	0	0	0	0	1	Monogamy, married
89.	F112	1972	1	.	Marriage	1	0	0	0	0	1	Monogamy, married
90.	F112	1973	1	.	Marriage	1	0	0	0	0	1	Monogamy, married
91.	F112	1974	1	.	Marriage	1	0	0	0	0	1	Monogamy, married
92.	F112	1975	1	.	Marriage	1	0	0	0	0	1	Monogamy, married
93.	F112	1976	1	.	Marriage	1	0	0	0	0	1	Monogamy, married
94.	F112	1977	1	.	Marriage	1	0	0	0	0	1	Monogamy, married
95.	F112	1978	1	.	Marriage	1	0	0	0	0	1	Monogamy, married
96.	F112	1979	1	.	Marriage	1	0	0	0	0	1	Monogamy, married
	.											
	etc											
	.											
125.	F112	2008	1	.	Marriage	1	0	0	0	0	1	Monogamy, married

marital	Fréquence	Pourcentage	Fréquence cumulée	Pct cumulé
Alone, never in union	75	21.74	75	21.74
Alone, separated	6	1.74	81	23.48
Alone, widowed	9	2.61	90	26.09
Monogamy, informal union	24	6.96	114	33.04
Monogamy, married	132	38.26	246	71.30
Polygamy, only informal unions	9	2.61	255	73.91
Polygamy, at least one marriage	90	26.09	345	100.00

2) Variable de comparaison des pays de résidence des répondants et des partenaires

Objectifs

- Dépendance de la valeur d'une variable aux valeurs passées : « unions réunifiées » ou « toujours unifiées ».
- Stratification des comparaisons par structure d'union.

a. Les structures d'union

La présence d'unions multiples oblige à établir des règles de priorité dans la codification. Par exemple, si un homme est en même temps dans une union transnationale avec une épouse et réunifiée avec une autre épouse, on considérera que son statut est « transnational ». D'autres règles de priorités pourraient être établies selon la recherche menée.

On va raisonner ici en termes de structure d'unions en créant la variable **suni_id** qui identifie – pour chaque répondant et chaque année – la combinaison des unions en cours (par exemple, si l'union 1 et l'union 2 sont en cours en même temps, **suni_id**=12 ; si, à un autre moment, les unions 2 et 3 sont simultanées, alors **suni_id**=23). Les comparaisons des pays de résidence se feront alors sur chaque structure. De nouveau cette règle n'est pas absolue, d'autres choix peuvent être opérés.

Sans union simultanée

Si un individu n'a jamais connu de situation d'overlapping, la variable **suni_id** est toujours égale à **num_uni_X**.

ident	t	num_uni_1	num_uni_2	num_uni_3	suni_id
2	t1	NA	NA	NA	NA
2	t2	1	NA	NA	1
2	t3	1	NA	NA	1
2	t3	1	NA	NA	1
2	t4	NA	2	NA	2
2	t5	NA	2	NA	2
2	t6	NA	2	NA	2
2	t7	NA	2	NA	2
2	t8	NA	2	NA	2
2	t9	NA	2	NA	2
2	t10	NA	NA	NA	NA
2	t11	NA	NA	3	3
2	t12	NA	NA	3	3
2	t13	NA	NA	3	3

Avec unions simultanées

Dans l'exemple illustré par le tableau ci-dessous on observe de l'overlapping. Les 3 unions permettent d'identifier 4 structures d'unions différentes: la structure [1] entre t2 et t3, la structure [12] entre t4 et t9, la structure [23] en t11 et t12 et la structure [2] en t10 et en t13.

ident	t	num_uni_1	num_uni_2	num_uni_3	suni_id
1	t1	NA	NA	NA	NA
1	t2	1	NA	NA	1
1	t3	1	NA	NA	1
1	t3	1	NA	NA	1
1	t4	1	2	NA	12
1	t5	1	2	NA	12
1	t6	1	2	NA	12
1	t7	1	2	NA	12
1	t8	1	2	NA	12
1	t9	1	2	NA	12
1	t10	NA	2	NA	2
1	t11	NA	2	3	23
1	t12	NA	2	3	23
1	t13	NA	2	NA	2

Le format de la variable d'identification de la structure d'union doit être de type caractère sans quoi l'on risquerait d'avoir la même valeur de **suni_id** pour deux structures différentes. Par exemple, avec les unions simultanées 1 et 2, on obtiendrait le même id que l'union unique 3 avec un format numérique. Pour créer cette variable, on concatène les variables num_uni_X dont le nombre maximal d'occurrences est égale à 5, après les avoir transformé en format caractère.

```

analys$num_uni.1[is.na(analys$num_uni.1)]<- "."
analys$num_uni.2[is.na(analys$num_uni.2)]<- "."
analys$num_uni.3[is.na(analys$num_uni.3)]<- "."
analys$num_uni.4[is.na(analys$num_uni.4)]<- "."
analys$num_uni.5[is.na(analys$num_uni.5)]<- "."

analys$suni_id<-
paste(analys$ident,analys$num_uni.1,analys$num_uni.2,analys$num_uni.3,
analys$num_uni.4, analys$num_uni.5, sep=" ")
analys$suni_id[analys$nbu==0]<-NA

```

La dépendance au passé consiste à créer des modalités « toujours unifié » et « réunifié » lorsque la structure d'union a été au moins une année en situation « transnationale ».

Pour repérer ces trajectoires passées on va compter le nombre d'années de la structure d'union (cuni) et le nombre d'années où la structure d'union était unifiée avec la variable tloc créée plus haut (cuni_loc).

```
# cuni
analys$uni      <- ifelse(analys$nbu>0,1,NA)
analys$cuni     <-ave(analys$uni, analys$suni_id ,FUN=cumsum)

# cuni_loc
analys$cuni_loc<-ave(analys$uni_loc, analys$suni_id ,FUN=cumsum)
```

b. La variable trans

Pour une année donnée, les différentes situations sont donc les suivantes:

- Si **uni = 0** : Le répondant n'est pas en couple.
- Si **uni_loc = 0** : les partenaires de l'union ne vivent pas dans le même pays. Il y a transnationalité de la structure.
- Si **uni_loc =1** et **cuni>cuni_loc** : Les membres de la structure vivent tous dans le même pays l'année t, mais cela n'a pas été toujours le cas (nombre d'années de l'union>nombre d'années où les deux personnes vivaient dans le même pays). Il y a réunification du couple.
- Si **uni_loc =1** et **cuni=cuni_loc** : Les membres de la structure ont toujours vécu dans le même pays.
-

```
analys$trans1[is.na(analys$uni)]<-"No partner/spouse"
analys$trans1[analys$uni_loc==0]<-"Transnational"
analys$trans1[analys$uni_loc==1 & analys$cuni_loc<analys$cuni]<-"Reunified"
analys$trans1[analys$uni_loc==1 & analys$cuni_loc==analys$cuni]<-"Always unified"
table(analys$trans1)
```

trans1	Fréquence	Pourcentage	Fréquence cumulée	Pctage cumulé
No partner/spouse	90	26.09	90	26.09
Transnational	54	15.65	144	41.74
Reunified	57	16.52	201	58.26
Always unified	144	41.74	345	100.00

Pour avoir plus de détails au niveau du statut matrimonial, on peut utiliser la variable **marital**. Ici on a juste distingué les épisodes « jamais en union », des « séparations et veuvage ».

```

analys$trans2[analys$trans1=="No partner/spouse" & analys$C==1 ]<-"Never in union"
analys$trans2[analys$trans1=="No partner/spouse" & analys$C==0 ]<-"Separated or widowed"
analys$trans2[analys$uni_loc==0]<-"Transnational"
analys$trans2[analys$uni_loc==1 & analys$cuni_loc<analys$cuni]<-"Reunified"
analys$trans2[analys$uni_loc==1 & analys$cuni_loc==analys$cuni]<-"Always unified"

table(analys$trans2)

```

trans2	Fréquence	Pourcentage	Fréquence cumulée	Pct cumulé
Never in union	75	21.74	75	21.74
Separated or widowed	15	4.35	90	26.09
Transnational	54	15.65	144	41.74
Reunified	57	16.52	201	58.26
Always unified	144	41.74	345	100.00

Extrait de la base trans2 entre 1950 et 2008 (tous les 5 ans)

	ident	year	age	suni_id	uni	uni_loc	cuni	cuni_loc	trans2
5.	E1	1990	18	1....	1	1	1	1	Always unified
10.	E1	1995	23	1....	1	1	6	5	Reunified
15.	E1	2000	28	1....	1	1	11	10	Reunified
20.	E1	2005	33	1....	1	0	16	10	Transnational
23.	E1	2008	36	1....	1	0	19	10	Transnational
27.	E103	1995	17	1....	1	0	4	3	Transnational
32.	E103	2000	22	Separated or widowed
37.	E103	2005	27	.2...	1	0	5	0	Transnational
40.	E103	2008	30	Separated or widowed
43.	E147	1970	16	Never in union
48.	E147	1975	21	Never in union
53.	E147	1980	26	1....	1	1	2	2	Always unified
58.	E147	1985	31	1....	1	1	7	7	Always unified
63.	E147	1990	36	12...	1	1	1	1	Always unified
68.	E147	1995	41	12...	1	0	6	4	Transnational
73.	E147	2000	46	12...	1	0	11	4	Transnational
78.	E147	2005	51	123..	1	0	2	0	Transnational
81.	E147	2008	54	123..	1	0	5	0	Transnational
82.	F112	1965	14	Never in union
87.	F112	1970	19	1....	1	1	4	3	Reunified
92.	F112	1975	24	1....	1	1	9	8	Reunified
97.	F112	1980	29	1....	1	1	14	13	Reunified
102.	F112	1985	34	1....	1	1	19	18	Reunified
107.	F112	1990	39	1....	1	1	24	23	Reunified
112.	F112	1995	44	1....	1	1	29	28	Reunified
117.	F112	2000	49	1....	1	1	34	33	Reunified
122.	F112	2005	54	1....	1	1	39	38	Reunified
125.	F112	2008	57	1....	1	1	42	41	Reunified
130.	F121	1990	18	Never in union
135.	F121	1995	23	Never in union
140.	F121	2000	28	1....	1	0	3	0	Transnational
145.	F121	2005	33	1....	1	0	8	0	Transnational
148.	F121	2008	36	1....	1	0	11	1	Transnational
151.	F127	1980	16	Never in union
156.	F127	1985	21	1....	1	1	2	2	Always unified
161.	F127	1990	26	.2...	1	0	2	0	Transnational
166.	F127	1995	31	..3..	1	0	2	0	Transnational
171.	F127	2000	36	..3..	1	0	7	0	Transnational
176.	F127	2005	41	..3.5	1	1	4	4	Always unified
179.	F127	2008	44	..3.5	1	1	7	7	Always unified
180.	I172	1985	14	1....	1	1	1	1	Always unified
185.	I172	1990	19	1....	1	1	6	6	Always unified
190.	I172	1995	24	Separated or widowed
195.	I172	2000	29	Separated or widowed
200.	I172	2005	34	.2...	1	0	4	0	Transnational
203.	I172	2008	37	.2...	1	0	7	0	Transnational
205.	S011202	1990	15	Never in union
210.	S011202	1995	20	Never in union
215.	S011202	2000	25	1....	1	0	2	1	Transnational
220.	S011202	2005	30	1....	1	1	7	2	Reunified
223.	S011202	2008	33	1....	1	0	10	3	Transnational
224.	S079601	1950	14	Never in union
229.	S079601	1955	19	Never in union
234.	S079601	1960	24	Never in union
239.	S079601	1965	29	1....	1	1	1	1	Always unified
244.	S079601	1970	34	1....	1	1	6	6	Always unified
249.	S079601	1975	39	1....	1	1	11	11	Always unified
254.	S079601	1980	44	12...	1	1	4	4	Always unified
259.	S079601	1985	49	123..	1	1	2	2	Always unified
264.	S079601	1990	54	12...	1	1	10	10	Always unified
269.	S079601	1995	59	12...	1	1	15	15	Always unified
274.	S079601	2000	64	12...	1	1	20	20	Always unified
279.	S079601	2005	69	12...	1	1	25	25	Always unified
282.	S079601	2008	72	12...	1	1	28	28	Always unified
287.	S093701	1950	18	Never in union
292.	S093701	1955	23	Never in union
297.	S093701	1960	28	Never in union
302.	S093701	1965	33	1....	1	1	5	5	Always unified
307.	S093701	1970	38	.2...	1	0	1	0	Transnational
312.	S093701	1975	43	.234.	1	1	1	1	Always unified
317.	S093701	1980	48	.2345	1	1	3	3	Always unified

322.		S093701	1985	53	.2345	1	1	8	8	Always unified	
327.		S093701	1990	58	.2345	1	1	13	13	Always unified	
332.		S093701	1995	63	.2345	1	1	18	18	Always unified	
337.		S093701	2000	68	.2345	1	1	23	23	Always unified	
342.		S093701	2005	73	.2345	1	0	28	27	Transnational	
345.		S093701	2008	76	.2345	1	0	31	27	Transnational	

-----+

Extrait de la base trans2 entre 1985 et 1995

	ident	year	age	suni_id	uni	uni_loc	cuni	cuni_loc	trans2
1.	E1	1986	14	Never in union
2.	E1	1987	15	Never in union
3.	E1	1988	16	Never in union
4.	E1	1989	17	Never in union
5.	E1	1990	18	1....	1	1	1	1	Always unified
6.	E1	1991	19	1....	1	0	2	1	Transnational
7.	E1	1992	20	1....	1	1	3	2	Reunified
8.	E1	1993	21	1....	1	1	4	3	Reunified
9.	E1	1994	22	1....	1	1	5	4	Reunified
10.	E1	1995	23	1....	1	1	6	5	Reunified
24.	E103	1992	14	1....	1	1	1	1	Always unified
25.	E103	1993	15	1....	1	1	2	2	Always unified
26.	E103	1994	16	1....	1	1	3	3	Always unified
27.	E103	1995	17	1....	1	0	4	3	Transnational
58.	E147	1985	31	1....	1	1	7	7	Always unified
59.	E147	1986	32	1....	1	1	8	8	Always unified
60.	E147	1987	33	1....	1	1	9	9	Always unified
61.	E147	1988	34	1....	1	1	10	10	Always unified
62.	E147	1989	35	1....	1	1	11	11	Always unified
63.	E147	1990	36	12...	1	1	1	1	Always unified
64.	E147	1991	37	12...	1	1	2	2	Always unified
65.	E147	1992	38	12...	1	1	3	3	Always unified
66.	E147	1993	39	12...	1	1	4	4	Always unified
67.	E147	1994	40	12...	1	0	5	4	Transnational
68.	E147	1995	41	12...	1	0	6	4	Transnational
102.	F112	1985	34	1....	1	1	19	18	Reunified
103.	F112	1986	35	1....	1	1	20	19	Reunified
104.	F112	1987	36	1....	1	1	21	20	Reunified
105.	F112	1988	37	1....	1	1	22	21	Reunified
106.	F112	1989	38	1....	1	1	23	22	Reunified
107.	F112	1990	39	1....	1	1	24	23	Reunified
108.	F112	1991	40	1....	1	1	25	24	Reunified
109.	F112	1992	41	1....	1	1	26	25	Reunified
110.	F112	1993	42	1....	1	1	27	26	Reunified
111.	F112	1994	43	1....	1	1	28	27	Reunified
112.	F112	1995	44	1....	1	1	29	28	Reunified
126.	F121	1986	14	Never in union
127.	F121	1987	15	Never in union
128.	F121	1988	16	Never in union
129.	F121	1989	17	Never in union
130.	F121	1990	18	Never in union
131.	F121	1991	19	Never in union
132.	F121	1992	20	Never in union
133.	F121	1993	21	Never in union
134.	F121	1994	22	Never in union
135.	F121	1995	23	Never in union
156.	F127	1985	21	1....	1	1	2	2	Always unified
157.	F127	1986	22	1....	1	1	3	3	Always unified
158.	F127	1987	23	1....	1	1	4	4	Always unified
159.	F127	1988	24	1....	1	1	5	5	Always unified
160.	F127	1989	25	.2...	1	0	1	0	Transnational
161.	F127	1990	26	.2...	1	0	2	0	Transnational
162.	F127	1991	27	.2...	1	0	3	0	Transnational
163.	F127	1992	28	..34.	1	1	1	1	Always unified
164.	F127	1993	29	..34.	1	1	2	2	Always unified
165.	F127	1994	30	..3..	1	0	1	0	Transnational
166.	F127	1995	31	..3..	1	0	2	0	Transnational
180.	I172	1985	14	1....	1	1	1	1	Always unified
181.	I172	1986	15	1....	1	1	2	2	Always unified
182.	I172	1987	16	1....	1	1	3	3	Always unified
183.	I172	1988	17	1....	1	1	4	4	Always unified
184.	I172	1989	18	1....	1	1	5	5	Always unified
185.	I172	1990	19	1....	1	1	6	6	Always unified
186.	I172	1991	20	1....	1	1	7	7	Always unified
187.	I172	1992	21	1....	1	1	8	8	Always unified
188.	I172	1993	22	Separated or widowed
189.	I172	1994	23	Separated or widowed
190.	I172	1995	24	Separated or widowed
204.	S011202	1989	14	Never in union
205.	S011202	1990	15	Never in union

206.	S011202	1991	16	Never in union	
207.	S011202	1992	17	Never in union	
208.	S011202	1993	18	Never in union	
209.	S011202	1994	19	Never in union	
210.	S011202	1995	20	Never in union	
259.	S079601	1985	49	123..	1	1	2	2	Always unified
260.	S079601	1986	50	123..	1	1	3	3	Always unified
261.	S079601	1987	51	123..	1	1	4	4	Always unified
262.	S079601	1988	52	12...	1	1	8	8	Always unified
263.	S079601	1989	53	12...	1	1	9	9	Always unified
264.	S079601	1990	54	12...	1	1	10	10	Always unified
265.	S079601	1991	55	12...	1	1	11	11	Always unified
266.	S079601	1992	56	12...	1	1	12	12	Always unified
267.	S079601	1993	57	12...	1	1	13	13	Always unified
268.	S079601	1994	58	12...	1	1	14	14	Always unified
269.	S079601	1995	59	12...	1	1	15	15	Always unified
322.	S093701	1985	53	.2345	1	1	8	8	Always unified
323.	S093701	1986	54	.2345	1	1	9	9	Always unified
324.	S093701	1987	55	.2345	1	1	10	10	Always unified
325.	S093701	1988	56	.2345	1	1	11	11	Always unified

326.	S093701	1989	57	.2345	1	1	12	12	Always unified
327.	S093701	1990	58	.2345	1	1	13	13	Always unified
328.	S093701	1991	59	.2345	1	1	14	14	Always unified
329.	S093701	1992	60	.2345	1	1	15	15	Always unified
330.	S093701	1993	61	.2345	1	1	16	16	Always unified

331.	S093701	1994	62	.2345	1	1	17	17	Always unified
332.	S093701	1995	63	.2345	1	1	18	18	Always unified

Extrait de la base trans2 entre 2000 et 2008

	ident	year	sunid_id	uni	uni_loc	cuni	cuni_loc	trans2
15.	E1	2000	1....	1	1	11	10	Reunified
16.	E1	2001	1....	1	0	12	10	Transnational
17.	E1	2002	1....	1	0	13	10	Transnational
18.	E1	2003	1....	1	0	14	10	Transnational
19.	E1	2004	1....	1	0	15	10	Transnational
20.	E1	2005	1....	1	0	16	10	Transnational
21.	E1	2006	1....	1	0	17	10	Transnational
22.	E1	2007	1....	1	0	18	10	Transnational
23.	E1	2008	1....	1	0	19	10	Transnational
32.	E103	2000	Separated or widowed
33.	E103	2001	.2...	1	0	1	0	Transnational
34.	E103	2002	.2...	1	0	2	0	Transnational
35.	E103	2003	.2...	1	0	3	0	Transnational
36.	E103	2004	.2...	1	0	4	0	Transnational
37.	E103	2005	.2...	1	0	5	0	Transnational
38.	E103	2006	.2...	1	0	6	0	Transnational
39.	E103	2007	Separated or widowed
40.	E103	2008	Separated or widowed
73.	E147	2000	12...	1	0	11	4	Transnational
74.	E147	2001	12...	1	0	12	4	Transnational
75.	E147	2002	12...	1	0	13	4	Transnational
76.	E147	2003	12...	1	0	14	4	Transnational
77.	E147	2004	123..	1	0	1	0	Transnational
78.	E147	2005	123..	1	0	2	0	Transnational
79.	E147	2006	123..	1	0	3	0	Transnational
80.	E147	2007	123..	1	0	4	0	Transnational
81.	E147	2008	123..	1	0	5	0	Transnational
117.	F112	2000	1....	1	1	34	33	Reunified
118.	F112	2001	1....	1	1	35	34	Reunified
119.	F112	2002	1....	1	1	36	35	Reunified
120.	F112	2003	1....	1	1	37	36	Reunified
121.	F112	2004	1....	1	1	38	37	Reunified
122.	F112	2005	1....	1	1	39	38	Reunified
123.	F112	2006	1....	1	1	40	39	Reunified
124.	F112	2007	1....	1	1	41	40	Reunified
125.	F112	2008	1....	1	1	42	41	Reunified
140.	F121	2000	1....	1	0	3	0	Transnational
141.	F121	2001	1....	1	0	4	0	Transnational
142.	F121	2002	1....	1	0	5	0	Transnational
143.	F121	2003	1....	1	0	6	0	Transnational
144.	F121	2004	1....	1	0	7	0	Transnational
145.	F121	2005	1....	1	0	8	0	Transnational
146.	F121	2006	1....	1	0	9	0	Transnational
147.	F121	2007	1....	1	1	10	1	Reunified
148.	F121	2008	1....	1	0	11	1	Transnational
171.	F127	2000	..3..	1	0	7	0	Transnational
172.	F127	2001	..3..	1	0	8	0	Transnational
173.	F127	2002	..3.5	1	1	1	1	Always unified
174.	F127	2003	..3.5	1	1	2	2	Always unified
175.	F127	2004	..3.5	1	1	3	3	Always unified
176.	F127	2005	..3.5	1	1	4	4	Always unified
177.	F127	2006	..3.5	1	1	5	5	Always unified
178.	F127	2007	..3.5	1	1	6	6	Always unified
179.	F127	2008	..3.5	1	1	7	7	Always unified
195.	I172	2000	Separated or widowed
196.	I172	2001	Separated or widowed
197.	I172	2002	.2...	1	0	1	0	Transnational
198.	I172	2003	.2...	1	0	2	0	Transnational
199.	I172	2004	.2...	1	0	3	0	Transnational
201.	I172	2006	.2...	1	0	5	0	Transnational
202.	I172	2007	.2...	1	0	6	0	Transnational
203.	I172	2008	.2...	1	0	7	0	Transnational
215.	S011202	2000	1....	1	0	2	1	Transnational
216.	S011202	2001	1....	1	0	3	1	Transnational
217.	S011202	2002	1....	1	0	4	1	Transnational
218.	S011202	2003	1....	1	0	5	1	Transnational
219.	S011202	2004	1....	1	0	6	1	Transnational
220.	S011202	2005	1....	1	1	7	2	Reunified
221.	S011202	2006	1....	1	1	8	3	Reunified
222.	S011202	2007	1....	1	0	9	3	Transnational
223.	S011202	2008	1....	1	0	10	3	Transnational

274.	S079601	2000	12...	1	1	20	20	Always unified
275.	S079601	2001	12...	1	1	21	21	Always unified
276.	S079601	2002	12...	1	1	22	22	Always unified
277.	S079601	2003	12...	1	1	23	23	Always unified
278.	S079601	2004	12...	1	1	24	24	Always unified
279.	S079601	2005	12...	1	1	25	25	Always unified
280.	S079601	2006	12...	1	1	26	26	Always unified
281.	S079601	2007	12...	1	1	27	27	Always unified
282.	S079601	2008	12...	1	1	28	28	Always unified
337.	S093701	2000	.2345	1	1	23	23	Always unified
338.	S093701	2001	.2345	1	1	24	24	Always unified
339.	S093701	2002	.2345	1	1	25	25	Always unified
340.	S093701	2003	.2345	1	1	26	26	Always unified
341.	S093701	2004	.2345	1	1	27	27	Always unified
342.	S093701	2005	.2345	1	0	28	27	Transnational
343.	S093701	2006	.2345	1	0	29	27	Transnational
344.	S093701	2007	.2345	1	0	30	27	Transnational
345.	S093701	2008	.2345	1	0	31	27	Transnational

-----+-----