

# **DICTIONNAIRE DES CODES ET DES LABELS:**

**VARIABLES SOCLES**

## Contenu

I – Listes des variables .....	2
II – Remarques sur les variables de la base socle .....	6
<b>A. Variable « statut conjugal à l’arrivée » .....</b>	<b>6</b>
<b>B. Score PHQ-4 : Score d’anxiété et de dépression (Kroenk et al. 2009) .....</b>	<b>7</b>
<b>C. Score de consommation d’alcool (AUDIT-C sous-score du score AUDIT élaboré par l’OMS) ....</b>	<b>7</b>
<b>D. Caractérisation des conditions d’arrivée .....</b>	<b>8</b>
1. Rappel des variables disponibles.....	8
2. Objectifs de l’analyse.....	9
3. Dendrogrammes pour analyser les proximités entre réponses multiples à une même question .....	10
4. Dendrogrammes croisant les différentes informations .....	12
5. Conclusion .....	14
Annexe 1. Dendrogramme croisant accompagnants (en 3 classes) et raisons de la migration (en 5 classes).....	15
Annexe 2. Dendrogramme croisant les trois variables : motifs, accompagnants et réseau .....	16
Annexe 3. Dendrogramme croisant le réseau à l’arrivée et les raisons de la venue en France ...	17

---

## **I – Listes des variables**

**Q1 (variable caractère) : Identifiant de l'enquêté**

**PRECR (variable catégorielle) : groupe consultation précaire (groupe de référence)**

Nom et valeurs des labels

PRECR:

1 Oui

2 Non

99992 Refus

99993 Ne sait pas

99994 Refus, NSP

99995 Non concerné

99996 Abandon

99997 Perte

99999 Information manquante

**PHQ4 (variable numérique) : score d'anxiété et de dépression : PHQ\_4**

Nom et valeurs des labels

PHQ4lb:

99992 Refus

99993 Ne sait pas

99994 Refus, NSP

99995 Non concerné

99996 Abandon

99997 Perte

99999 Information manquante

**ORIENT\_SEX (variable catégorielle) : Orientation sexuelle**

Nom et valeurs des labels

orient\_sex:

0 Jamais sexe

- 1 Hétéro exclusif
- 2 Homo exclusif
- 3 Bisexuel
- 99992 Refus
- 99993 Ne sait pas
- 99994 Refus, NSP
- 99995 Non concerné
- 99996 Abandon
- 99997 Perte
- 99999 Information manquante

**CONJUG\_ARRIV (variable catégorielle) : situation conjugale à l'arrivée en France**

Nom et valeurs des labels

CONJUG\_A\_FR:

- 1 seul
- 2 en couple France
- 3 en couple avec un partenaire à l'étranger
- 99992 Refus
- 99993 Ne sait pas
- 99994 Refus, NSP
- 99995 Non concerné
- 99996 Abandon
- 99997 Perte
- 99999 Information manquante

**RAISON\_ARRIV (variable catégorielle) : raison de la venue en France**

Nom et valeurs des labels

RAISON\_ARRIV:

- 1 chercher un emploi/tenter sa chance

2 rejoindre la famille (+emploi identifié)

3 menacé dans son pays

4 études

5 raisons médicales

99992 Refus

99993 Ne sait pas

99994 Refus, NSP

99995 Non concerné

99996 Abandon

99997 Perte

99999 Information manquante

**RESEAU\_FR (variable catégorielle) : les personnes que l'on connaît déjà en France à son arrivée**

Nom et valeurs des labels

RESEAU\_FR:

1 connaître son conjoint

2 connaître de la famille (sans conjoint)

3 connaître seulement amis ou connaissance

4 ne connaître personne

99992 Refus

99993 Ne sait pas

99994 Refus, NSP

99995 Non concerné

99996 Abandon

99997 Perte

99999 Information manquante

**PROFILALCOOL (variable catégorielle) : Profil de consommation d'alcool (DREES)**

Nom et valeurs des labels

Profil:

1 Non consommateur

2 Conso sans risque

3 conso a risq ponct

4 conso a risq chro

99992 Refus

99993 Ne sait pas

99994 Refus, NSP

99995 Non concerné

99996 Abandon

99997 Perte

99999 Information manquante

**AUDITC (variable numérique) : Score de consommation d'alcool AUDIT-C**

Nom et valeurs des labels

AUDITC1b:

99992 Refus

99993 Ne sait pas

99994 Refus, NSP

99995 Non concerné

99996 Abandon

99997 Perte

99999 Information manquante

**ANNEE\_ARRIV (variable numérique) : année d'arrivée en France**

Nom et valeurs des labels

ANNEE\_ARRIV1b:

99992 Refus

- 99993 Ne sait pas
- 99994 Refus, NSP
- 99995 Non concerné
- 99996 Abandon
- 99997 Perte
- 99999 Information manquante

---

## **II – Remarques sur les variables de la base socle**

### **A. Variable « statut conjugal à l'arrivée »**

La variable « statut conjugal à l'arrivée » se décline en trois modalités

- 1) Seul (e)
- 2) En couple en France
- 3) En couple avec un partenaire à l'étranger

La première modalité correspond aux personnes qui déclarent ne pas être en relation longue (définie comme une relation avec une partenaire depuis plus d'un an) à leur arrivée en France. Les modalités 2 et 3 correspondent aux personnes qui déclarent être en relation longue à leur arrivée en France. La modalité 2 correspond aux personnes dont le partenaire vit en France et la modalité 3 aux personnes dont le partenaire vit dans le pays d'origine ou dans un autre pays.

Afin d'obtenir cette distinction plusieurs étapes ont été nécessaires :

Premièrement la question Q105 : « Avez-vous toujours habité ensemble ? » a été utilisée. Les individus qui ont répondu « Oui, toujours », on considère qu'il s'agit d'une migration « de couple » et les deux partenaires sont ensemble en France soit la modalité 2 « En couple en France ». Deuxièmement, les questions suivantes : Q48 : « Pourquoi êtes-vous venu(e) en France (moment de la première année d'arrivée) ? », Q49 « Avec qui êtes-vous arrivé(e) ? » et Q50 « A ce moment la connaissiez-vous déjà des personnes qui vivaient en France ont été utilisées pour déterminer si la « migration liée au partenaire ».

On considère qu'il s'agit d'une « relation transnationale à l'arrivée » si les personnes déclarent être venues pour une autre raison que pour de se marier ou de rejoindre leur conjoint(e), si elles ne sont pas arrivées avec leur conjoint(e) en France et si elles ne connaissaient pas leur (e) conjoint(e) en France.

## B. Score PHQ-4 : Score d'anxiété et de dépression (Kroenk et al. 2009)

Le score PHQ-4 est un score variant de 0 à 12, il est construit en additionnant le score de 4 questions retrouvées dans le questionnaire CAPI : Q34\_1, Q34\_2, Q34\_3, Q34\_4.

Le PHQ-4 est catégorisé en 4 classes selon le résultat du score :

- 0-2 : Normal
- 3-5 : Léger
- 6- 8 : Modéré
- 9-12 : Sévère

## C. Score de consommation d'alcool (AUDIT-C sous-score du score AUDIT élaboré par l'OMS)

Trois questions posées dans l'enquête Parcours permettent de reconstruire le score de consommation d'alcool AUDIT-C (sous-score du test AUDIT de l'Organisation Mondiale de la Santé) qui peut varier de 0 à 12 (Bush K et al. 1998). Ces questions concernent la fréquence de consommation, la quantité de boisson consommée un jour typique, et la consommation excessive d'alcool lors d'une seule occasion : **Q44. Quelle est votre consommation d'alcool ?** (Vous ne buvez jamais d'alcool ; 1 fois par mois ou moins ; 2 à 4 fois par mois ; 2 à 3 fois par semaine ; Au moins 4 fois par semaine) ; **Q45. Combien de verres contenant de l'alcool consommez-vous un jour typique où vous buvez ?** (Non consommateurs ; 1 ou 2, 3 ou 4 ; 5 ou 6 ; 7, 8 ou 9 ; 10 ou plus) ; **Q46. Avec quelle fréquence buvez-vous six verres ou davantage lors d'une occasion particulière ?** (Non consommateurs ; Jamais ; Moins d'une fois par mois ; Une fois par mois ; Une fois par semaine ; Tous les jours ou presque).

Une autre classification proposée par la DREES (Direction de la Recherche, des Etudes, de l'Evaluation et des Statistiques) consiste à définir des « profils d'alcoolisation », selon les recommandations de l'OMS en termes de consommation d'alcool (Mouquet and Villet 2002). La construction de ces profils tient compte du volume hebdomadaire en termes de consommation d'alcool ainsi que d'une alcoolisation ponctuelle importante. Ceci permet de distinguer une consommation régulière d'une consommation occasionnelle et en même temps de dégager une pratique dite « à risque » d'une pratique dite « sans risque ».

A partir de cette méthode, une personne peut être classée dans 4 différents profils :

- Les non-consommateurs, qui ne boivent jamais de boissons alcoolisées.
- Les consommateurs sans risque, qui ont une consommation d'alcool qui ne dépasse pas les seuils et pratiques recommandés.
- Les consommateurs à risque ponctuel : la volumétrie hebdomadaire de consommation rejoint celle des consommateurs sans risque, mais ils boivent de manière excessive à certaines occasions. Cette consommation excessive ponctuelle suffit à accroître les dangers liés aux états d'ébriété et les comportements à risque de l'individu.
- Les consommateurs à risque chronique : ils ont une volumétrie hebdomadaire de consommation supérieure aux seuils recommandés ou consomment de l'alcool de manière importante et de façon récurrente. Outre les risques qu'ils ont en commun avec les consommateurs à risque ponctuel, ces individus sont également exposés à des risques physiologiques et sociaux. Caractérisation des conditions d'arrivée

## D. Caractérisation des conditions d'arrivée

### 1. Rappel des variables disponibles

<b>Motifs de la venue en France</b>		
<b>Nom et label</b>	<b>Modalités</b>	<b>Effectifs % pondérés (N=2468)</b>
<b>Q48_1</b> Raison de la venue en France : rejoindre conjoint(e) ou fiancé	1. Oui 2. Non	16% 84%
<b>Q48_2</b> Raison de la venue en France : rejoindre enfants	1. Oui 2. Non	2% 98%
<b>Q48_3</b> Raison de la venue en France : rejoindre autre membre de la famille	1. Oui 2. Non	12% 88%
<b>Q48_4</b> Raison de la venue en France : pour vous marier	1. Oui 2. Non	1% 99%
<b>Q48_5</b> Raison de la venue en France : étudier	1. Oui 2. Non	15% 85%
<b>Q48_6</b> Raison de la venue en France : chercher un travail	1. Oui 2. Non	20% 80%
<b>Q48_7</b> Raison de la venue en France : prendre un emploi déjà identifié	1. Oui 2. Non	2% 98%
<b>Q48_8</b> Raison de la venue en France : tenter sa chance	1. Oui 2. Non	24% 76%
<b>Q48_9</b> Raison de la venue en France : vous étiez menacé dans votre pays	1. Oui 2. Non	18% 82%
<b>Q48_10</b> Raison de la venue en France : pour raisons médicales	1. Oui 2. Non	5% 95%
<b>Q48_11</b> Raison de la venue en France : venue avec ses parents	1. Oui 2. Non	2% 98%
<b>Q48_12</b> Raison de la venue en France : ne sait pas	1. Oui 2. Non	0% 100%
<b>Q48_13</b> Raison de la venue en France : pour une autre raison	1. Oui 2. Non	0% 100%
<b>Q48_14</b> Raison de la venue en France : non réponse	1. Oui 2. Non	0% 100%
<b>Q48_15</b> Raison de la venue en France : séjour court	1. Oui 2. Non	2% 98%
<b>Avec qui est on arrivé en France</b>		
<b>Nom et label</b>	<b>Modalités</b>	<b>Effectifs % pondérés (N=2468)</b>
<b>Q49_1</b> Avec qui êtes-vous arrivé(e) : conjoint(e)	1. Oui 2. Non	4% 96%
<b>Q49_2</b> Avec qui êtes-vous arrivé(e): enfant(s)	1. Oui 2. Non	5% 95%
<b>Q49_3</b> Avec qui êtes-vous arrivé(e) : parents	1. Oui 2. Non	7% 93%
<b>Q49_4</b>	1. Oui	77%

Avec qui êtes-vous arrivé(e) : seul	2. Non	23%
<b>Q49_5</b>	1. Oui	7%
Avec qui êtes-vous arrivé(e): autre personne	2. Non	83%
<b>Réseau en France</b>		
<b>Nom et label</b>	<b>Modalités</b>	<b>Effectifs % pondérés (N=2468)</b>
<b>Q50_1</b>	1. Oui	62%
Vous connaissez déjà en France une ou des personnes de votre famille	2. Non	38%
<b>Q50_2</b>	1. Oui	16%
Vous connaissez déjà en France conjoint(e) ou fiancé(e)	2. Non	84%
<b>Q50_3</b>	1. Oui	44%
Vous connaissez déjà en France un ou des amis ou connaissances	2. Non	56%
<b>Q50_4</b>	1. Oui	15%
Vous ne connaissiez personne en France	2. Non	85%
<b>Q50_5</b>	1. Oui	0%
Non réponse sur réseau en France	2. Non	100%

## 2. Objectifs de l'analyse

On dispose de trois types des questions :

- Avec qui êtes-vous arrivés ?
- Raisons de l'arrivée en France
- Qui connaissiez vous en France à votre arrivée ?

On veut arriver à :

- Soit quelques variables clés qui caractérisent au mieux ces conditions de l'arrivée, avec un nombre restreint de catégories pour chaque variable
- Soit une typologie des conditions d'arrivée : recréer une nouvelle variable qui combine les différentes variables citées.
- Soit les deux.

On utilise des dendrogrammes pour :

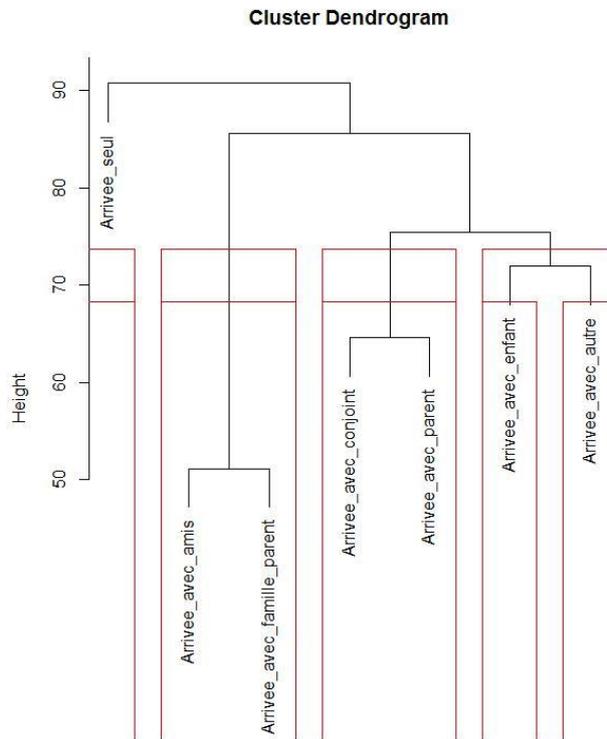
- analyser les proximités entre différentes réponses pour une même variable (puisque plusieurs réponses étaient possibles)
- analyser les proximités entre ces différentes variables : cela permettra de décider quelles sont les variables qui structurent le plus la situation d'arrivée en France, et quelle pourrait être une combinaison de variables pour créer une typologie

On complète cette analyse en cluster par :

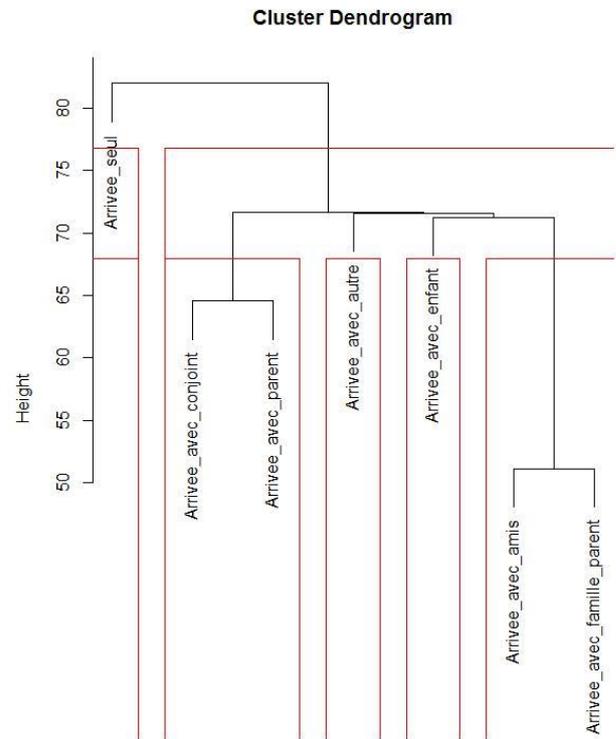
- Une réflexion sur le sens des catégories utilisées
- L'examen des effectifs considérés, pour éviter des catégories à très faible effectif

### 3. Dendrogrammes pour analyser les proximités entre réponses multiples à une même question

#### a. avec qui êtes-vous arrivé

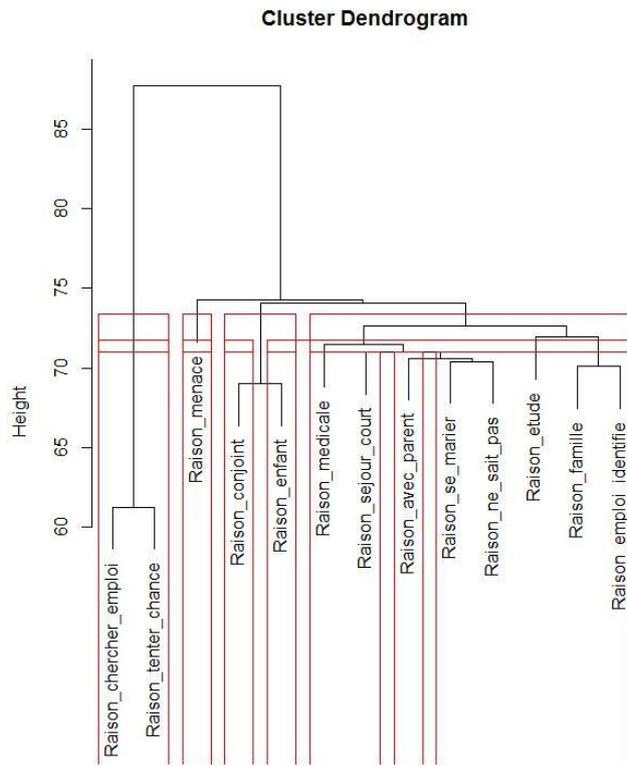


```
dist(t(scale(var_cah1)))  
hclust(*, "ward")
```

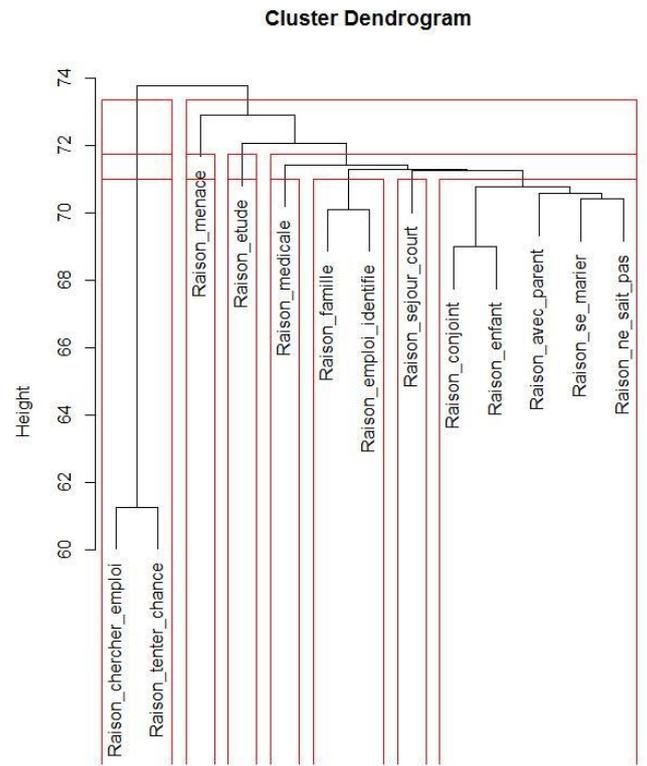


```
dist(t(scale(var_cah1)))  
hclust(*, "average")
```

**b. les raisons de l'arrivée**



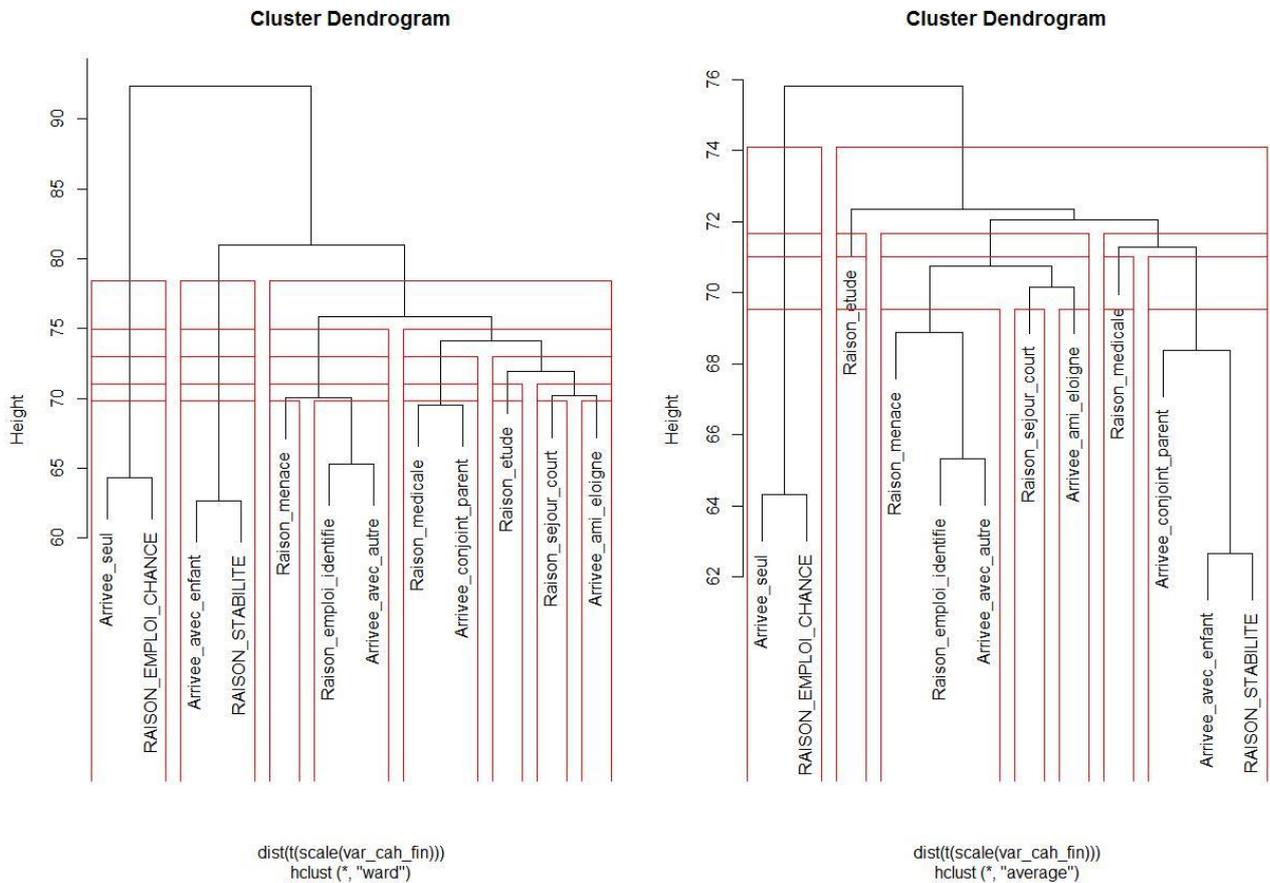
dist(t(scale(var\_cah2)))  
hclust (\*, "ward")



dist(t(scale(var\_cah2)))  
hclust (\*, "average")

#### 4. Dendrogrammes croisant les différentes informations

##### a. Dendrogramme croisant accompagnants et raisons de la migration avec un nombre maximum de classes

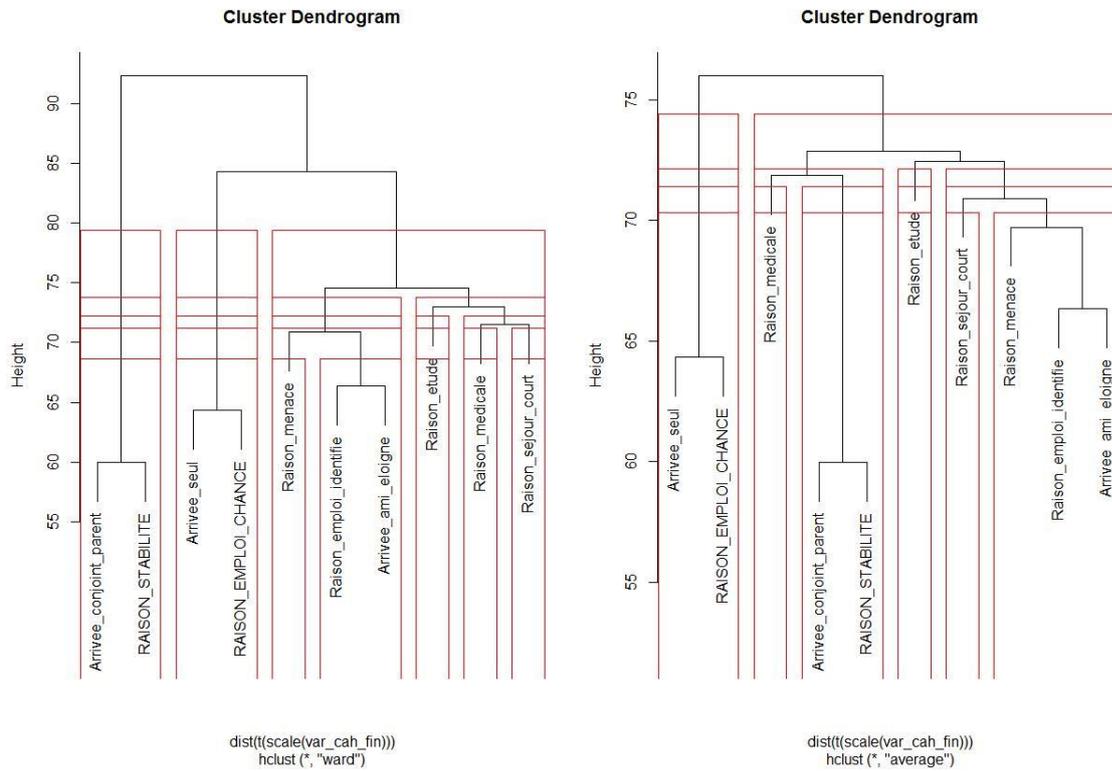


⇒ On voit sur le dendrogramme que l'on peut regrouper les accompagnants en trois catégories

1. arrivé seul
2. arrivé avec conjoint, parent, enfant
3. arrivé avec amis, autre membres de la famille, et autre

En revanche, ce n'est pas très concluant sur les raisons qui se distribuent vraiment différemment entre les deux dendrogrammes, avec des regroupements qui n'ont pas beaucoup de sens. On refait les dendrogrammes avec les accompagnants en trois catégories et les raisons en 7 catégories, pour voir si c'est plus clair.

**b. Dendrogramme croisant accompagnants (en 3 classes) et raisons de la migration (en 7 classes)**

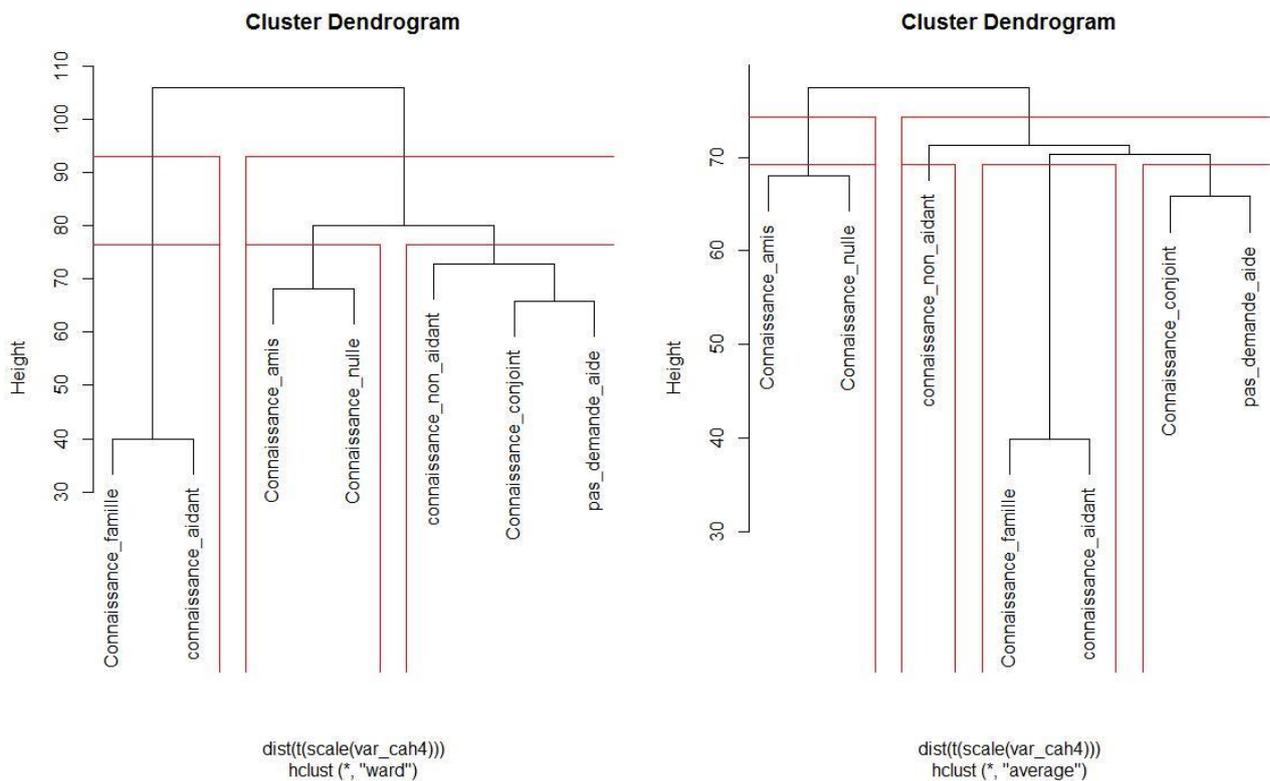


↳ A partir de ces deux dendrogrammes, nous décidons de partir sur une catégorisation des motifs en cinq classes où l'on rebascule le séjour court (2%) dans emploi/chance en tenant compte de la nature imprévue du séjour prolongé, et l'emploi identifié dans « stabilité » vu que cette catégorie ne peut pas rester seule (2%) et ne se rapproche pas statistiquement d'une autre catégorie en particulier

On arrive à la catégorisation suivante :

1. chercher un emploi/tenter sa chance
2. stabilité
3. menacé dans son pays
4. études
5. raisons médicales

### c. Dendrogramme croisant le réseau à l'arrivée et le fait d'avoir bénéficié d'aide de son entourage



Ce dendrogramme 3 nous renseigne sur un type de relation entre ces deux informations : en général

- les personnes qui connaissaient de la famille en arrivant ont pu bénéficier d'aide
- celles qui avaient un conjoint en arrivant n'ont pas eu besoin d'aide
- les choses sont plus mitigées pour les personnes qui connaissaient juste des amis en arrivant

NB : le fait d'avoir bénéficié de l'aide des personnes qu'on connaissait à l'arrivée est renseignée dans la question Q51 du questionnaire principal « CAPI » (base CAPI\_QUET.dta)

L'information importante sur les connaissances à l'arrivée est donc donnée par la catégorisation suivante :

- Connaître son conjoint en France
- Connaître de la famille (sans conjoint)
- Connaître seulement amis ou connaissances (ni famille ni conjoint)
- Ne connaître personne

## 5. Conclusion

Les jeux de variables qui seront utilisés sont les suivants :

Les variables recatégorisées

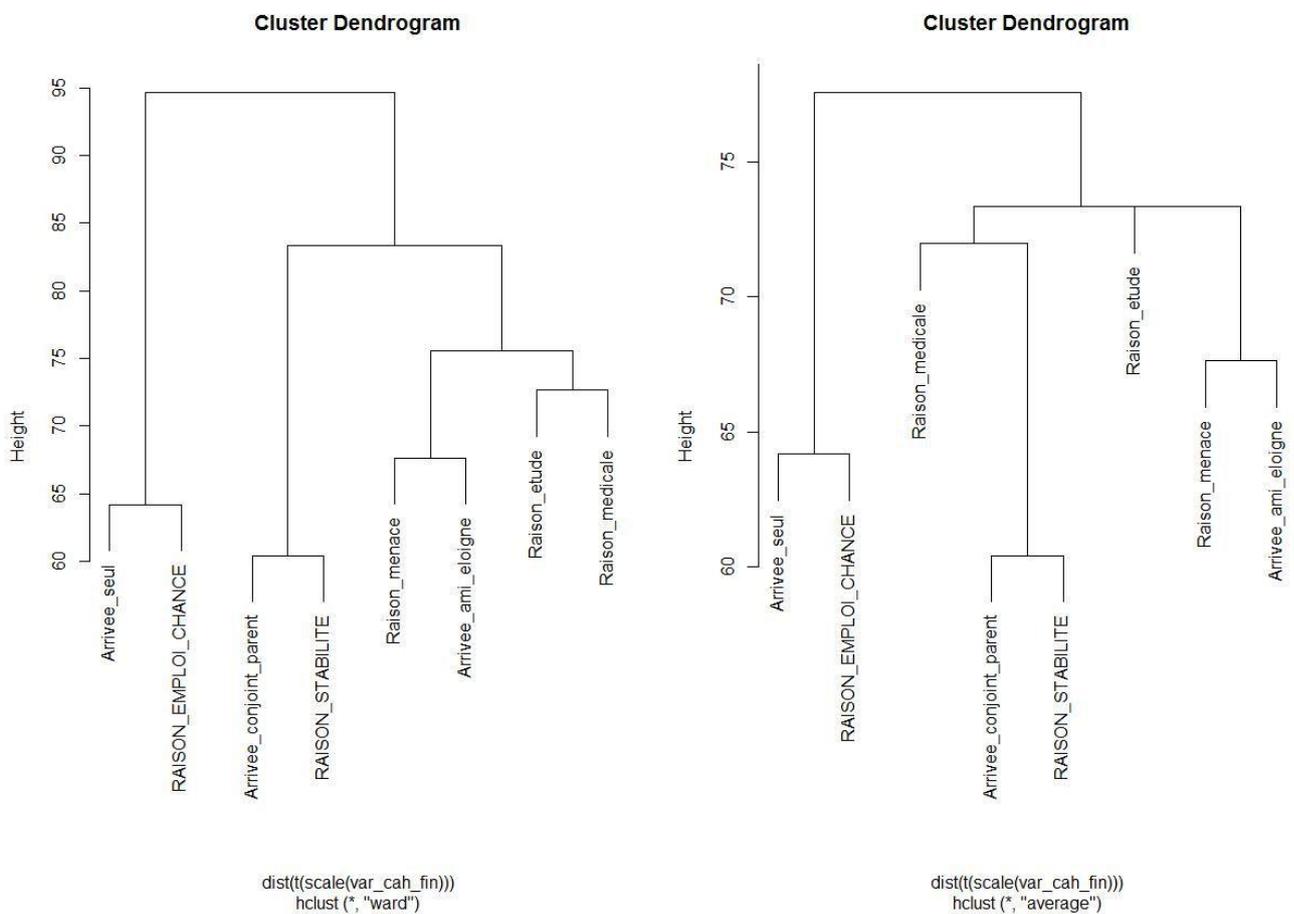
### Raison de la venue en France

1. chercher un emploi/tenter sa chance
2. raisons familiales/emploi identifié
3. menacé dans son pays
4. études
5. raisons médicales

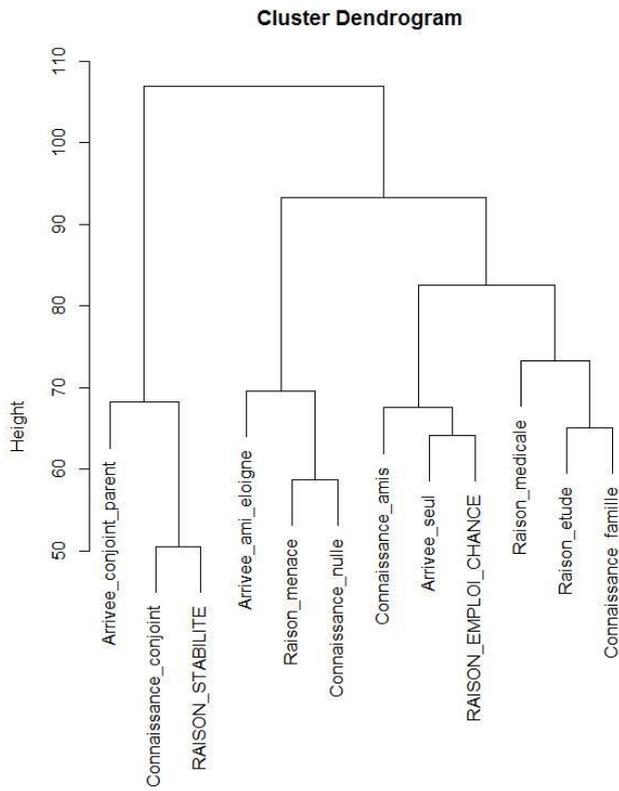
### Réseau en France

1. Connaître son conjoint en France
2. Connaître de la famille (sans conjoint)
3. Connaître seulement amis ou connaissances (ni famille ni conjoint)
4. Ne connaître personne

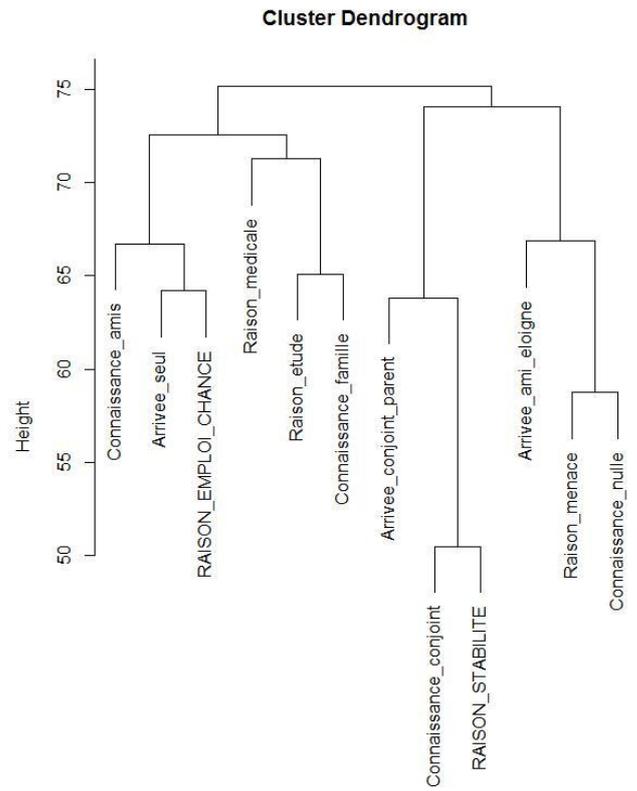
## Annexe 1. Dendrogramme croisant accompagnants (en 3 classes) et raisons de la migration (en 5 classes)



## Annexe 2. Dendrogramme croisant les trois variables : motifs, accompagnants et réseau

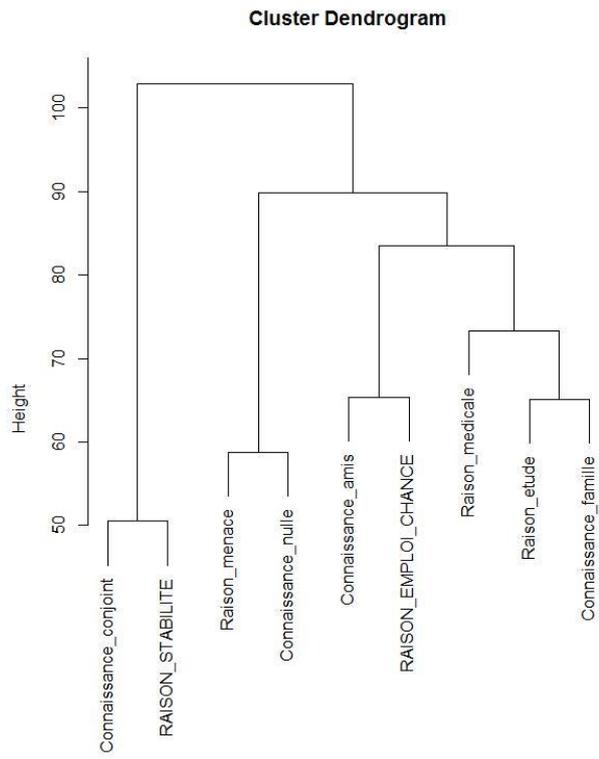


dist(t(scale(var\_cah\_fin2)))  
hclust(\*, "ward")

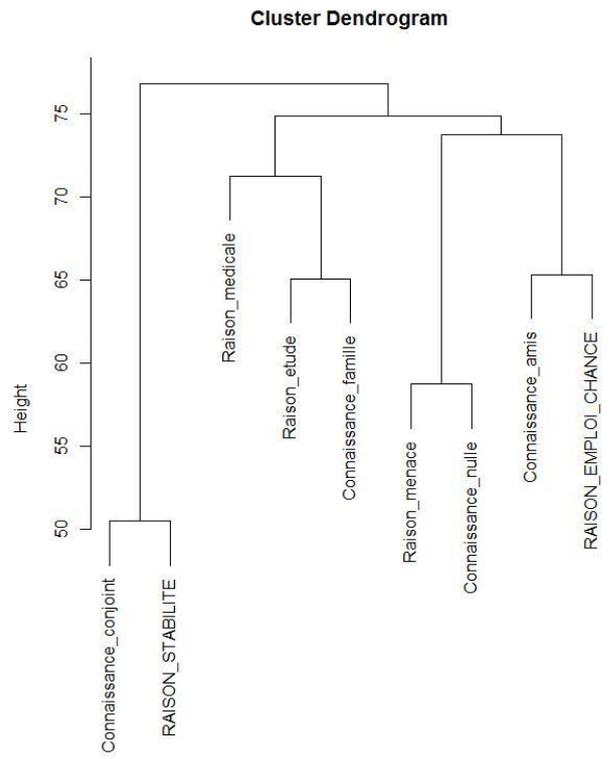


dist(t(scale(var\_cah\_fin2)))  
hclust(\*, "average")

### Annexe 3. Dendrogramme croisant le réseau à l'arrivée et les raisons de la venue en France



```
dist(t(scale(var_cah_fin2[, c(1:6, 8, 11, 12)])))
hclust(*, "ward")
```



```
dist(t(scale(var_cah_fin2[, c(1:6, 8, 11, 12)])))
hclust(*, "average")
```