

# Analyse multivariée de l'offre de soins en France

## Étude comparative de la structure départementale de l'offre de soins

Fadli Aaron, Moreau Matteo et Kurnaz Kubra

2026-01-08

## Contents

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Introduction</b>   | <b>2</b>  |
| 1.1      | Contexte et objectifs . . . . .                                   | 2         |
| 1.2      | Objectifs de l'analyse . . . . .                                  | 2         |
| <b>2</b> | <b>Méthodologie</b>   | <b>2</b>  |
| 2.1      | Source de données . . . . .                                       | 2         |
| 2.1.1    | Données AMELI - démographie des professionnels de santé . . . . . | 2         |
| 2.1.2    | Données FINESS – établissements de santé . . . . .                | 2         |
| 2.2      | Nettoyage et harmonisation des données . . . . .                  | 3         |
| 2.2.1    | Traitement des données AMELI . . . . .                            | 3         |
| 2.2.2    | Harmonisation des codes départementaux . . . . .                  | 3         |
| 2.3      | Construction des indicateurs pour l'ACP . . . . .                 | 3         |
| 2.3.1    | Indicateurs de structure démographique . . . . .                  | 3         |
| 2.3.2    | Indicateurs de volume . . . . .                                   | 4         |
| 2.4      | Construction des variables pour l'ACM . . . . .                   | 4         |
| 2.4.1    | Principes de discréétisation . . . . .                            | 4         |
| 2.4.2    | Variables qualitatives construites . . . . .                      | 4         |
| 2.4.3    | Équilibrage et validation . . . . .                               | 5         |
| <b>3</b> | <b>Analyse en Composantes Principales (ACP)</b>                   | <b>6</b>  |
| 3.1      | Analyse comparative 2010 vs 2024 . . . . .                        | 6         |
| 3.1.1    | Éboulis des valeurs propres . . . . .                             | 6         |
| 3.1.2    | Cercle des corrélations . . . . .                                 | 7         |
| 3.1.3    | Contributions des variables aux dimensions . . . . .              | 8         |
| 3.1.4    | Projection des individus . . . . .                                | 9         |
| 3.2      | Conclusion synthétique de l'ACP . . . . .                         | 9         |
| <b>4</b> | <b>Analyse des Correspondances Multiples (ACM)</b>                | <b>10</b> |
| 4.1      | Analyse comparative 2010 vs 2024 . . . . .                        | 10        |
| 4.1.1    | Interprétation des axes et plan factoriel . . . . .               | 11        |
| 4.1.2    | Contributions des modalités . . . . .                             | 12        |
| 4.1.3    | Projection des individus . . . . .                                | 13        |
| 4.2      | Conclusion de l'ACM . . . . .                                     | 13        |

# 1 Introduction

## 1.1 Contexte et objectifs

L'analyse de l'offre de soins constitue un enjeu majeur pour les politiques publiques. La répartition territorial des professionnels de santé ainsi que l'évolution de leur caractéristiques démographiques sont des éléments essentiels pour anticiper les besoins futurs et orienter en conséquence les politiques d'aménagement sanitaire du territoire.

Cette étude propose une analyse de l'offre de soins à l'échelle départementale, en mobilisant plusieurs approches complémentaires :

- **L'Analyse en Composantes Principales (ACP)**, qui permet d'étudier les relations entre variables quantitatives continues décrivant le volume et la structure démographique de l'offre
- **L'Analyse des Correspondances Multiples (ACM)**, qui propose une typologie qualitative des territoires à partir de variables catégorielles construites par discrétisation
- **L'analyse Factorielles des Correspondances (AFC)**, qui propose ...
- Le clustering, qui propose ...

## 1.2 Objectifs de l'analyse

L'objectif principal de cette étude est de caractériser la structure A COMPLETER

# 2 Méthodologie

## 2.1 Source de données

### 2.1.1 Données AMELI - démographie des professionnels de santé

La base AMELI, issue du portail data.ameli.fr, constitue la source principale de cette étude. Elle fournit des informations détaillées sur les effectifs de professionnels de santé selon:

- Le sexe
- La classe d'âge
- La profession
- Le département
- L'année

Ces données présentent une granularité départementale annuellement ce qui permet donc de construire des indicateurs démographiques synthétiques.

Table 1: Aperçu de 5 lignes aléatoires du jeu de données

| Année      | Département   | Sexe   | Classe d'âge   | Profession                 |
|------------|---------------|--------|----------------|----------------------------|
| 2022-01-01 | Orne          | hommes | de 40 à 44 ans | Hépato-gastro-entérologues |
| 2024-01-01 | Haute-Saône   | femmes | de 60 à 64 ans | Psychiatres                |
| 2016-01-01 | Pas-de-Calais | hommes | de 60 à 64 ans | Stomatologues              |
| 2020-01-01 | Deux-Sèvres   | femmes | de 55 à 59 ans | Médecins vasculaires       |
| 2019-01-01 | Creuse        | hommes | de 65 à 69 ans | Anesthésistes-réanimateurs |

### 2.1.2 Données FINESS – établissements de santé

La base FINESS, issue de data.gouv.fr, recense les établissements sanitaires et médico-sociaux. Après fusion des informations administratives et géographiques, elle permet d'identifier la localisation départementale de chaque établissement.

Contrairement aux données AMELI, FINESS ne dispose pas d'une dimension temporelle homogène sur la période étudiée. Les informations issues de cette base sont donc considérées comme structurelles, c'est-à-dire constantes dans le temps.

Table 2: Aperçu de 5 établissements (données FINESS)

| N° FINESS | Raison Sociale                         | Catégorie  | Commune | Département      |
|-----------|--|--|---------|------------------|
| 280003567 | PHARMACIE BIJOT                        | Pharmacie d'Officine   | 218     | EURE ET LOIR     |
| 950780148 | POUPONNIÈRE CAR. SOC. LES POUSSINET    | Pouponnier à Caractère Social                                | 555     | VAL D'OISE       |
| 860789742 | EHPAD RESIDENCE PORTE DU MARTRAY       | Etablissement d'hébergement pour personnes âgées dépendantes | 137     | VIENNE           |
| 130040785 | LBM CERBALLIANCE PROVENCE-AZUR SITE DE | Laboratoire de Biologie Médicale                             | 086     | BOUCHES DU RHÔNE |
| 560019390 | PHARMACIE DE PENESTIN                  | Pharmacie d'Officine   | 155     | MORBIHAN         |

## 2.2 Nettoyage et harmonisation des données

### 2.2.1 Traitement des données AMELI

Afin d'éviter les doublons et les agrégats non désirés et incompatibles avec notre analyse territoriale, plusieurs traitements ont été appliqués :

- Exclusion de la modalité “tout sexe” et “sexe inconnu”
- Exclusion des agrégats territoriaux “FRANCE” et “Tout département”
- Exclusion des classes d'âge “Tout âge” et “âge inconnu”
- Exclusion des regroupements de professions de type “Ensemble des...”

Ces opérations garantissent que chaque observation corresponde à un effectif réel, rattaché à une profession, un sexe, une classe d'âge connue, un département et enfin une année.

### 2.2.2 Harmonisation des codes départementaux

La variable département qui a été utilisé pour la jointure issue de **FINESS** utilise des codes spécifiques pour certains territoires. Ces codes ont été recodés selon la nomenclature standard de l'INSEE :

- 9A -> 971 (Guadeloupe)
- 9B -> 972 (Martinique)
- 9C -> 973 (Guyane)
- 9D -> 974 (La Réunion)
- 9E -> 975 (Saint-Pierre-et-Miquelon)
- 9F -> 976 (Mayotte)

## 2.3 Construction des indicateurs pour l'ACP

L'ACP repose sur cinq variables quantitatives continues construire à partir des données, agrégées en couple département × année.

### 2.3.1 Indicateurs de structure démographique

#### Part des femmes

La part des femmes est calculée comme le rapport entre l'effectif féminin et l'effectif total :

$$\text{part\_femmes} = \frac{\text{effectif femmes}}{\text{effectif total}}$$

#### Part des professionnels de 60 ans et plus

La part des 60 ans et plus est calculée à partir des classes d'âge 60-64 ans, 65-69 ans et 70 ans et plus :

$$\text{part\_60\_plus} = \frac{\text{effectif des 60 ans et plus}}{\text{effectif total}}$$

Cet indicateur constitue une mesure centrale du vieillissement de l'offre de soins et des enjeux de renouvellement démographique.

### Âge moyen estimé

L'âge moyen est estimé par une moyenne pondérée des centres de classes d'âge, les pondérations étant les effectifs correspondants. Cette méthode est couramment utilisée lorsque l'âge individuel n'est pas disponible.

### 2.3.2 Indicateurs de volume

#### Effectif total de professionnels (transformation logarithmique)

L'effectif total correspond à la somme des effectifs de professionnels de santé, toutes professions confondues, par département et par année. Afin de limiter l'influence des départements les plus peuplés dans l'ACP, une transformation logarithmique est appliquée :

$$\text{effectif\_total\_log} = \log(1 + \text{effectif\_total})$$

Table 3: Aperçu de 5 lignes aléatoire du jeu de données

| departement | annee      | annee_f    | effectif_total_log | part_femmes | part_60_plus | age_moyen | nb_etablissements_log |
|-------------|------------|------------|--------------------|-------------|--------------|-----------|-----------------------|
| 2A          | 2011-01-01 | 2011-01-01 | 6.914731           | 0.5308151   | 0.1610338    | 47.68787  | 5.662960              |
| 34          | 2018-01-01 | 2018-01-01 | 9.149528           | 0.5857158   | 0.1795090    | 47.03592  | 7.501634              |
| 971         | 2018-01-01 | 2018-01-01 | 7.905442           | 0.6355588   | 0.2039838    | 46.63076  | 6.541030              |
| 19          | 2015-01-01 | 2015-01-01 | 7.138073           | 0.5532591   | 0.1891892    | 48.23609  | 6.196444              |
| 74          | 2021-01-01 | 2021-01-01 | 8.366138           | 0.6019079   | 0.1461145    | 44.71010  | 7.003065              |

### 2.4 Construction des variables pour l'ACM

L'ACM nécessite des variables qualitatives. Ces variables sont construites par **discrétisation en tertiles** des indicateurs quantitatifs précédents.

#### 2.4.1 Principes de discrétisation

Le choix de la discrétisation en tertiles repose sur plusieurs considérations :

- Absence de seuils clairement établis
- Volonté d'obtenir des modalités équilibrées en effectifs

**Important** : Les seuils sont calculés sur l'ensemble des observations (toutes années confondues), ce qui garantit une interprétation homogène dans le temps et permet des comparaisons inter-temporelles cohérentes.

#### 2.4.2 Variables qualitatives construites

##### Niveau d'offre de soins

La variable **niveau\_offre** est construite à partir de l'effectif total de professionnels par département et par année. Trois modalités sont définies :

- **Faible** : départements situés dans le premier tertile
- **Moyen** : départements situés dans le deuxième tertile
- **Élevé** : départements situés dans le troisième tertile

Cette variable caractérise le niveau global de dotation en professionnels de santé.

#### Féminisation de l'offre

La variable **feminisation** repose sur la part des femmes parmi les professionnels de santé. Trois modalités sont retenues :

- **Faible** : départements avec une féminisation inférieure au premier tertile
- **Intermédiaire** : départements situés dans le deuxième tertile
- **Forte** : départements avec une féminisation supérieure au deuxième tertile

Cette variable permet de saisir les différences territoriales et temporelles dans la composition sexuée de l'offre de soins.

#### Vieillissement de l'offre

La variable **vieillissement** est construite à partir de la part des professionnels âgés de 60 ans et plus. Trois modalités sont définies :

- **Jeune** : départements avec une faible part de professionnels de 60 ans et plus
- **Intermédiaire** : départements situés dans le deuxième tertile
- **Vieillissante** : départements avec une forte part de professionnels de 60 ans et plus

Cette variable constitue un indicateur central pour analyser les enjeux de renouvellement de l'offre de soins.

#### 2.4.3 Équilibrage et validation

La discrétisation en tertiles conduit naturellement à un équilibre parfait des modalités sur l'ensemble de la base :

- Chaque modalité représente environ un tiers des observations
- Aucune modalité rare n'est observée
- Aucune variable constante n'est introduite

Table 4: Aperçu de 5 lignes aléatoire du jeu de données

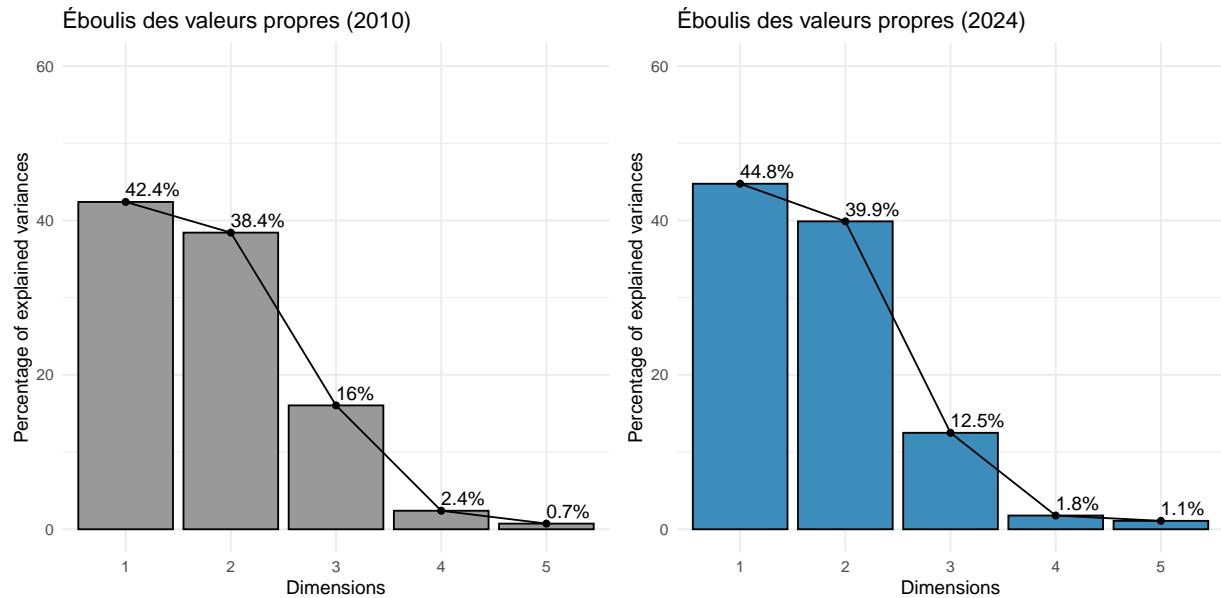
| departement | annee      | annee_f    | niveau_offre | feminisation  | vieillissement |
|-------------|------------|------------|--------------|---------------|----------------|
| 07          | 2016-01-01 | 2016-01-01 | Moyen        | Intermédiaire | Intermédiaire  |
| 93          | 2018-01-01 | 2018-01-01 | Élevé        | Faible        | Vieillissante  |
| 73          | 2022-01-01 | 2022-01-01 | Moyen        | Forte         | Jeune          |
| 67          | 2019-01-01 | 2019-01-01 | Élevé        | Intermédiaire | Vieillissante  |
| 61          | 2012-01-01 | 2012-01-01 | Faible       | Faible        | Vieillissante  |

### 3 Analyse en Composantes Principales (ACP)

#### 3.1 Analyse comparative 2010 vs 2024

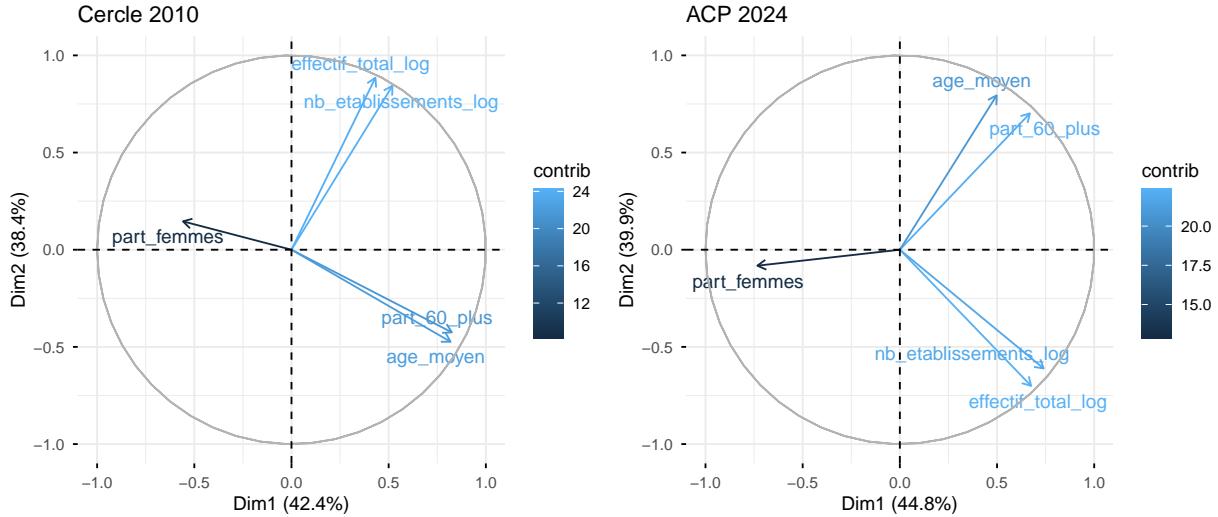
L'ACP a été réalisé sur 101 départements français pour chacune des deux années de référence (2010 et 2024) avec 5 variables quantitatives actives. Les individus sont donc chacun des départements français.

##### 3.1.1 Éboulis des valeurs propres



Les éboulis des valeurs propres montrent que les deux premières dimensions concentrent l'essentiel de l'information aussi bien en 2010 qu'en 2024. En 2010, les deux premiers axes expliquent 80,8% de la variance totale, tandis qu'en 2024 cette proportion atteint 84,6%, indiquant une structure factorielle légèrement plus concentrée dans le temps. La forte décroissance observée après la deuxième dimension suggère que les axes suivants apportent une information marginale ce qui implique que l'analyse peut être raisonnablement limitée aux deux premiers axes pour l'interprétation des structures territoriales de l'offre de soins.

### 3.1.2 Cercle des corrélations



L'analyse des cercles de corrélations met en évidence deux dimensions structurantes communes aux deux années étudiées, bien que leur positionnement sur les axes diffère.

En 2010, la première dimension est principalement associée aux variables de volume de l'offre de soins, notamment l'effectif total de professionnels de santé et le nombre d'établissements. Elle traduit un gradient opposant les départements fortement dotés à ceux disposant d'une offre plus restreinte. La seconde dimension est davantage corrélée aux variables de structure démographique, en particulier l'âge moyen et la part des professionnels âgés de 60 ans et plus.

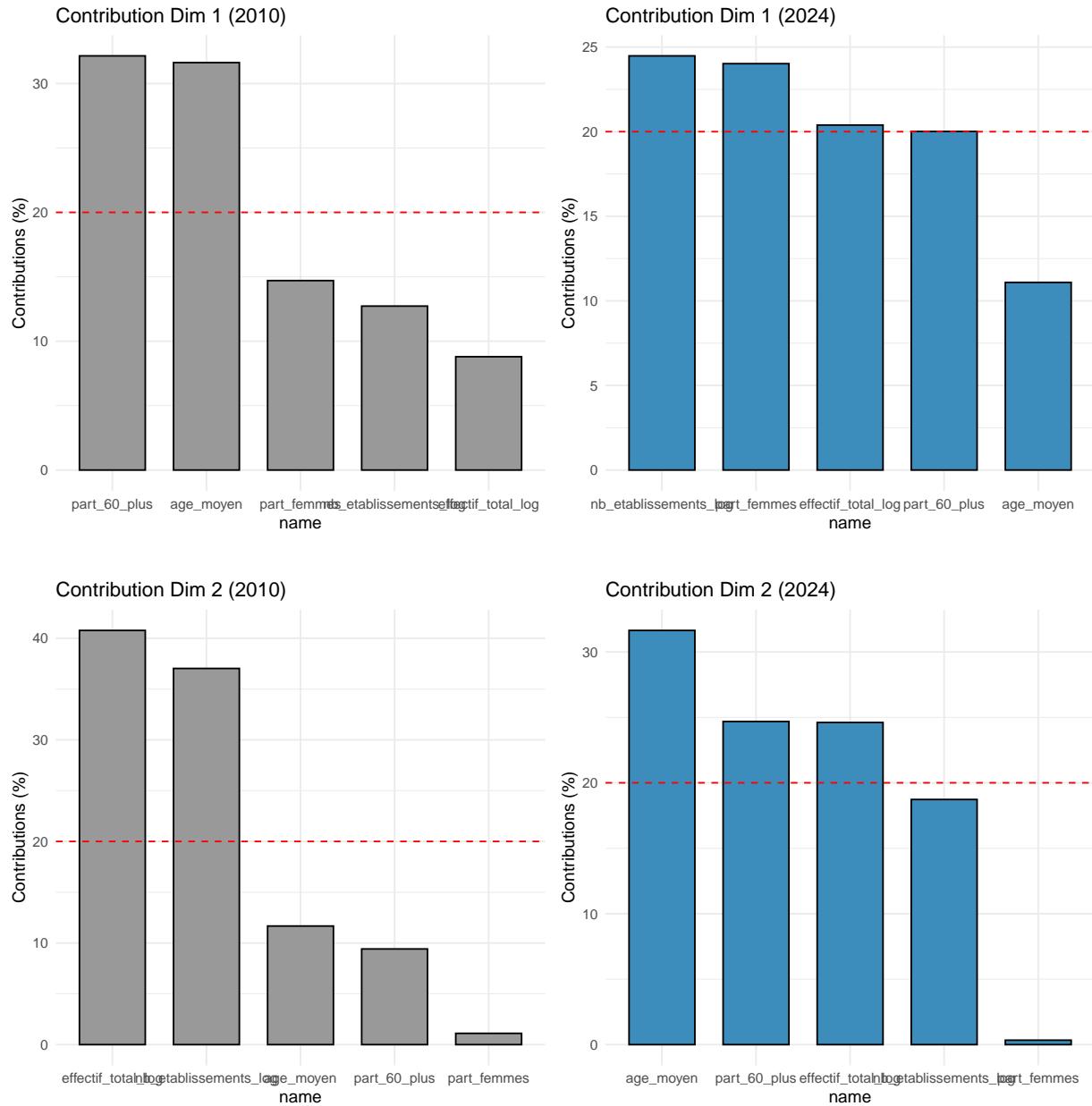
En 2024, on observe une inversion des axes : les variables qui structuraient majoritairement la première dimension en 2010 contribuent davantage à la seconde, et inversement. Cette inversion ne correspond malgré cela pas à un changement de sens des axes mais à une rotation du plan factoriel, les axes de l'ACP étant définis indépendamment pour chaque année. Ainsi, la numérotation des dimensions n'est pas directement comparable dans le temps.

Dans les deux cas, les mêmes dimensions latentes demeurent :

- une dimension liée au volume global de l'offre de soin,
- une dimension liée à la structure démographique des professionnels.

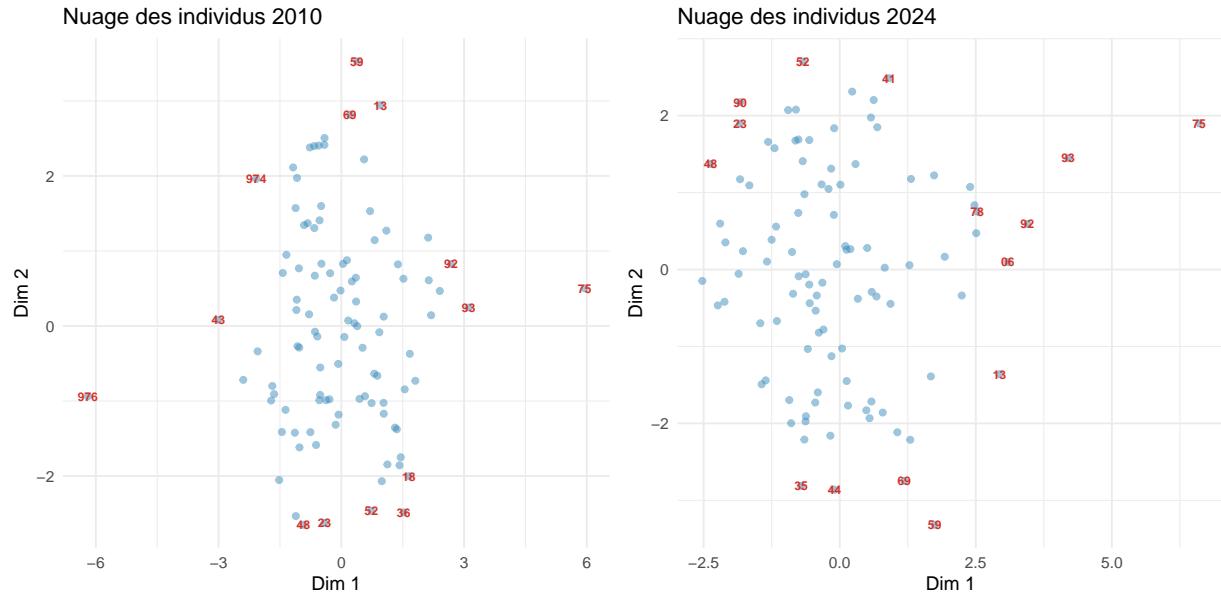
En revanche, leur poids relatif dans l'explication de la variance évolue. En 2024, les variables d'âge et de vieillissement contribuent davantage à la structuration de l'espace factoriel, suggérant que les différences territoriales sont de plus en plus liées au profil démographique des professionnels de santé, et non plus uniquement à la quantité d'offre disponible.

### 3.1.3 Contributions des variables aux dimensions



L'analyse des contributions confirme cette lecture. La dimension associée au volume est principalement portée par l'effectif total et le nombre d'établissements, tandis que la dimension démographique est dominée par l'âge moyen et la part des professionnels de 60 ans et plus. La part des femmes contribue de manière plus modérée mais régulière, traduisant une féminisation progressive et relativement homogène de l'offre de soins sur l'ensemble du territoire.

### 3.1.4 Projection des individus



Les projections des départements sur le plan factoriel révèlent une forte dispersion selon la dimension de volume, mettant en évidence d'importantes inégalités territoriales en matière d'offre de soins. Certains départements se distinguent par des positions extrêmes, reflétant des situations atypiques de forte concentration ou de sous-dotation.

Entre 2010 et 2024, la structure selon la dimension démographique devient plus visible, avec une différenciation accrue de certains territoires en fonction du vieillissement de leur offre de soins. Cette évolution confirme le rôle croissant des enjeux démographiques dans l'analyse territoriale de l'offre de soins.

## 3.2 Conclusion synthétique de l'ACP

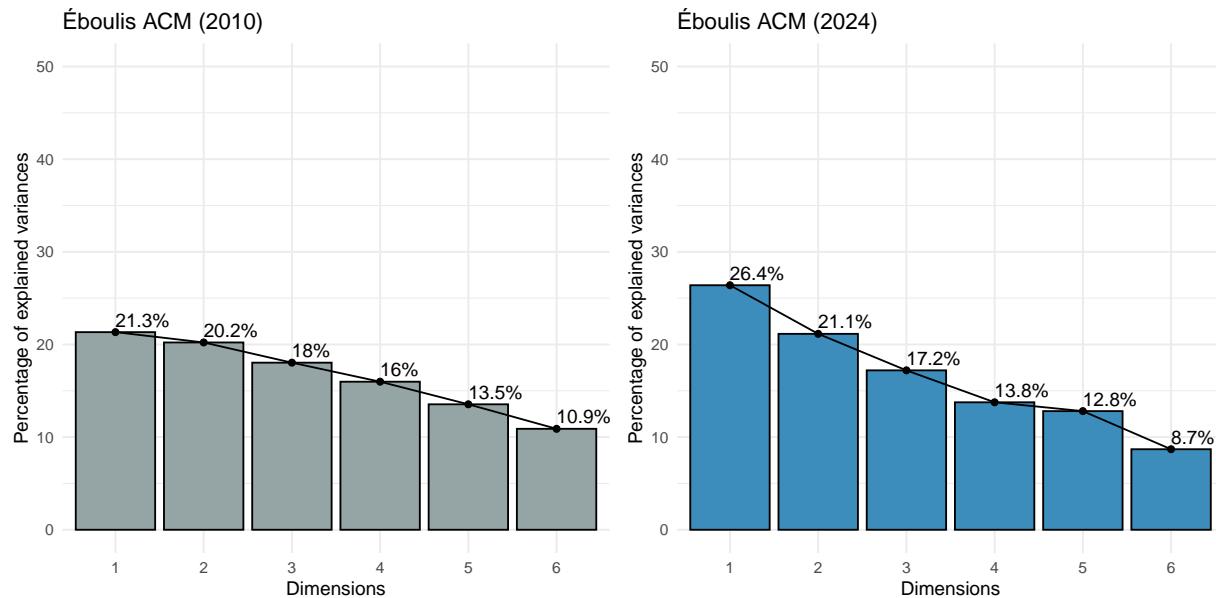
L'analyse en composantes principales met ainsi en évidence une structuration stable de l'offre de soins autour de deux dimensions fondamentales : **le volume et la structure démographique**, tout en soulignant une évolution notable de leur importance relative. Entre 2010 et 2024, le vieillissement des professionnels de santé apparaît comme un facteur de plus en plus discriminant entre les départements, au même titre que le niveau global de dotation.

## 4 Analyse des Correspondances Multiples (ACM)

### 4.1 Analyse comparative 2010 vs 2024

Comme précédemment, l'ACM a été réalisée sur 101 départements français pour chacune des deux années de références (2010 et 2024), avec 3 variables qualitatives actives (9 modalités au total).

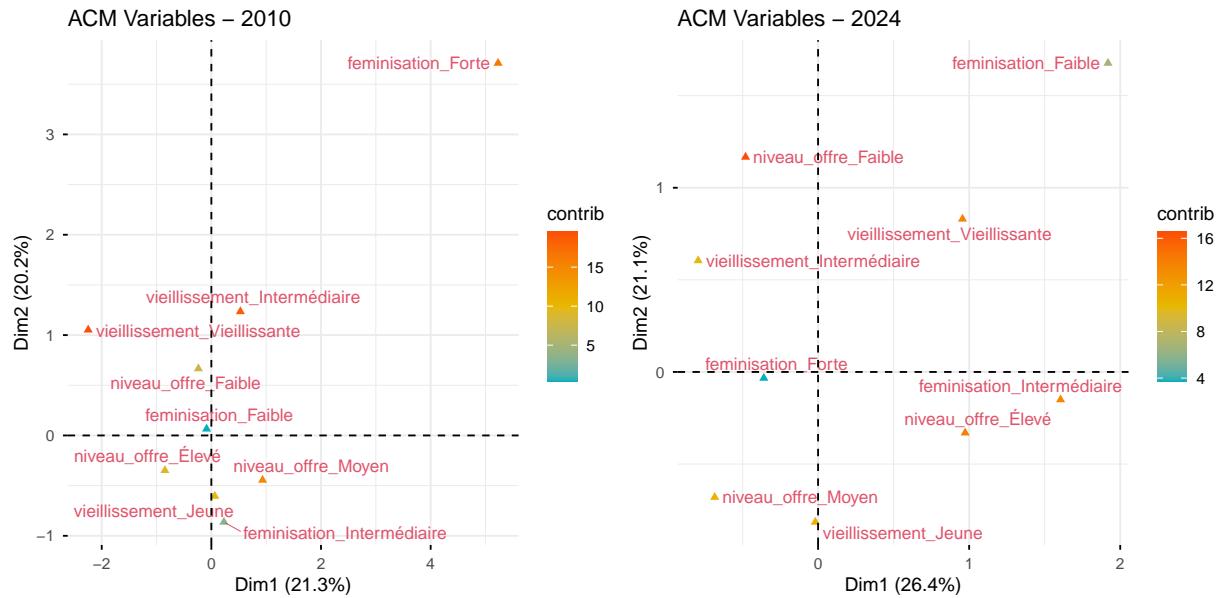
##Éboulis des valeurs propres



Les éboulis des valeurs propres montrent que, pour les deux années étudiées, les premières dimensions expliquent une part substantielle de l'inertie totale. En 2010, les deux premiers axes cumulent environ 41,5% de l'intertie, tandis qu'en 2024 cette proportion atteint 47,5%, traduisant une structuration légèrement plus marquée des profils départementaux dans le temps.

Comme c'est fréquemment le cas en ACM, l'intertie est plus diffuse que dans une ACP, ce qui justifie une interprétation prudente et principalement sur les deux premières dimensions.

#### 4.1.1 Interprétation des axes et plan factoriel



L’analyse des modalités projetées sur le plan factoriel met en évidence des structures comparables entre 2010 et 2024, bien que leur positionnement sur les axes diffère.

En 2010, la première dimension oppose principalement les départements caractérisés par une offre de soins faible et une structure vieillissante à ceux disposant d’une offre plus élevée et d’un profil plus favorable. La seconde dimension est davantage structurée par la féminisation de l’offre, opposant des territoires faiblement féminisés à ceux où la part des femmes est plus importante.

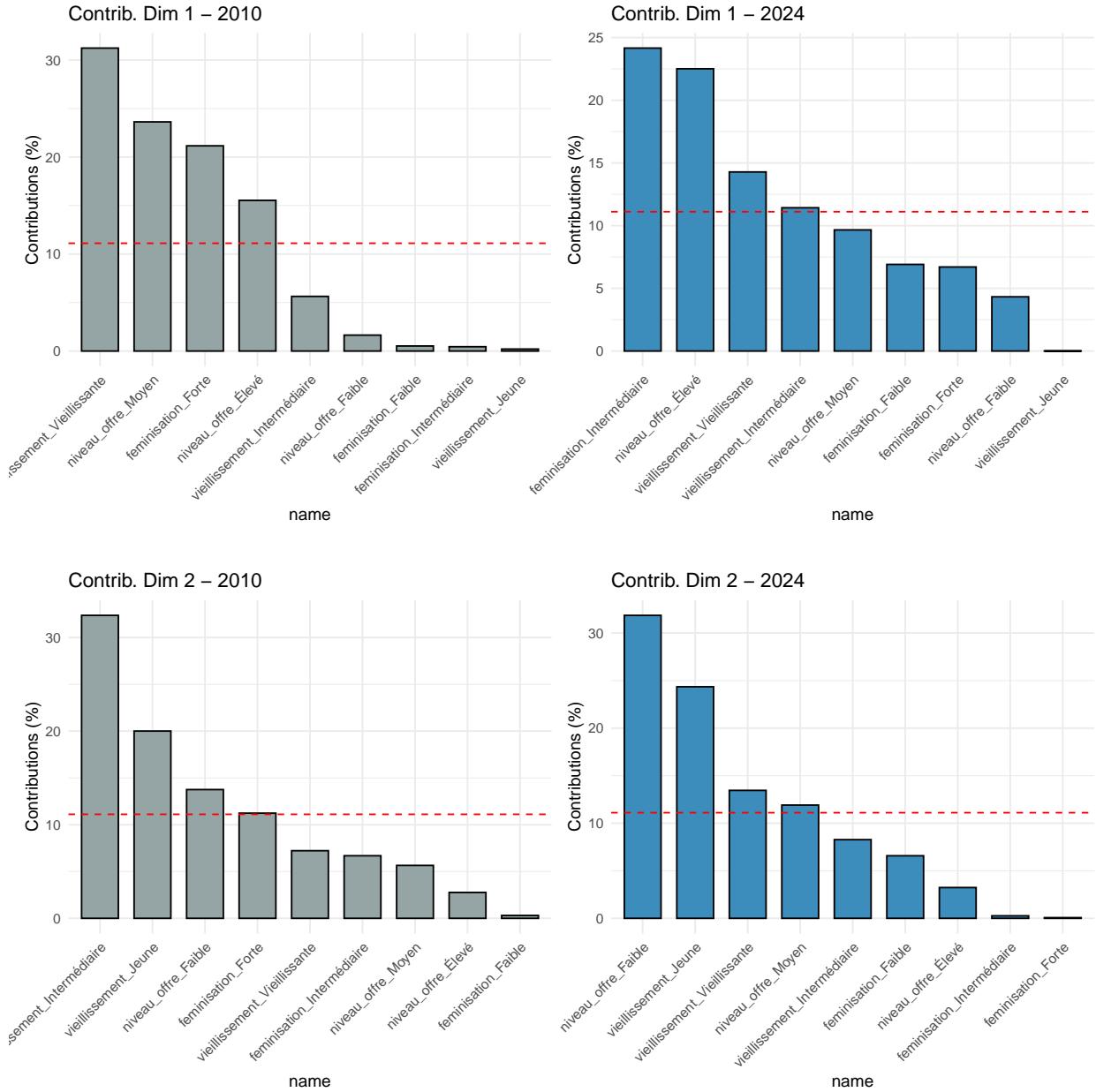
En 2024, on observe une réorganisation du plan factoriel : certaines modalités qui structuraient principalement la première dimension en 2010 contribuent davantage à la seconde, et inversement. Cette évolution ne doit pas être interprétée comme une inversion de sens des axes, mais comme précédemment par une rotation du repère factoriel, les axes étant définis indépendamment pour chaque année.

Dans les deux cas, les dimensions interprétables restent similaires :

- une dimension liée à la dotation globale en offre de soins,
- une dimension liée à la structure démographique, en particulier au vieillissement et à la féminisation.

La différence observée entre 2010 et 2024 traduit une évolution du poids relatif de ces dimensions dans la structuration des territoires.

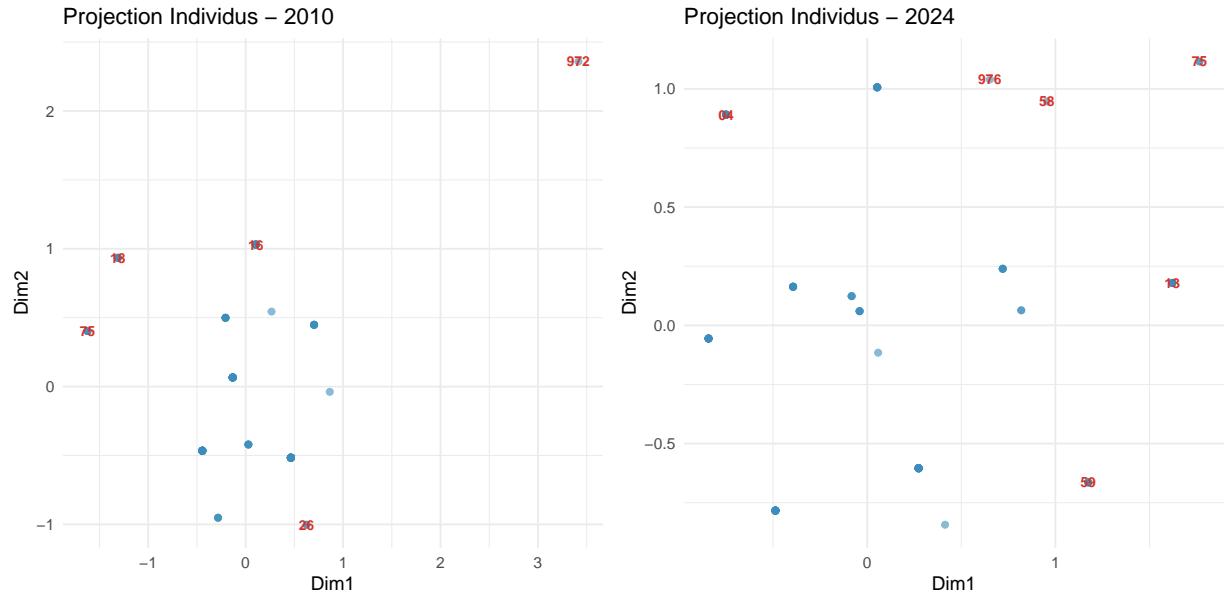
#### 4.1.2 Contributions des modalités



L'analyse des contributions confirme cette lecture. Les modalités associées à une offre élevée et à un vieillissement marqué contribuent fortement aux axes principaux, traduisant leur rôle structurant dans la différenciation des départements.

En 2024, les modalités liées au vieillissement de l'offre apparaissent plus contributives qu'en 2010, ce qui suggère une montée en importance de cet enjeu dans la typologie territoriale. La féminisation, bien que présente sur les deux dimensions, joue un rôle plus transversal et moins discriminant.

#### 4.1.3 Projection des individus



La projection des départements sur le plan factoriel met en évidence une structuration nette selon les profils qualitatifs définis. Certains se distinguent par des positions extrêmes, reflétant des configurations spécifiques combinant faible niveau d'offre, vieillissement prononcé ou, à l'inverse forte dotation et structure plus favorable.

Entre 2010 et 2024, la dispersion des individus selon les dimensions liées au vieillissement devient plus marquée, suggérant une différenciation accrue des territoires face aux enjeux de renouvellement de l'offre de soins.

## 4.2 Conclusion de l'ACM

L'ACM met ainsi en évidence une typologie territoriale relativement stable de l'offre de soins, structurée autour du niveau de dotation et de la structure démographique des professionnels. Toutefois, la comparaison temporelle souligne une montée en importance des modalités liées au vieillissement entre 2010 et 2024, confirmant les résultats obtenus par l'ACP et renforçant l'idée d'un enjeu démographique central dans l'évolution de l'offre de soins en France.