

Partie 3 : Développement et implémentation d'un cas d'AFC, à l'aide des logiciels R ou Python.

Dans cette partie, vous allez implémenter en Python ou en R les étapes de l'AFC appliquées au tableau de contingence présenté dans la Partie 2.

L'objectif est de vérifier numériquement vos calculs manuels, de réduire la dimension, de représenter graphiquement les résultats et d'interpréter les sorties.

1. Implémentation de l'AFC. Programmez les étapes suivantes :

- Construire la matrice des effectifs et la matrice des fréquences,
- Calculer les profils-lignes et profils-colonnes, ainsi que le profil moyen,
- Mettre en œuvre le modèle d'indépendance et calculer la statistique χ^2 ,
- Réaliser la réduction dimensionnelle (SVD sur la matrice **Z**),
- Tracer la carte factorielle (nuage des profils-lignes et profils-colonnes).

2. Interprétation des résultats. En particulier, répondez aux questions suivantes :

- (a) **Association globale.** À l'aide du χ^2 et de l'inertie totale Φ^2 , la situation familiale et l'origine de l'automobile sont-elles indépendantes ? Justifiez.
- (b) **Proximité avec le profil moyen.** Quels profils-lignes se rapprochent le plus du profil moyen (marge colonnes) ? Lesquels s'en écartent le plus ?
- (c) **Qualité de représentation.** Pour chaque modalité (au moins deux lignes et deux colonnes), indiquez la qualité de représentation (\cos^2) sur le plan factoriel.
- (d) **Contributions aux axes.** Quelles modalités (lignes ou colonnes) contribuent le plus à la construction des deux premiers axes ?
- (e) **Comparaison de profils.** Quelles deux situations familiales présentent les profils-lignes les plus similaires ? (Indice : distance du khi-deux entre profils-lignes.)

Rendu attendu

Déposer sur le campus :

- Le ou les **scripts** Python/R bien commentés,
- Un court compte rendu avec les graphiques obtenus et vos réponses aux questions d'interprétation.

Outils nécessaires pour l'AFC

Pour réaliser l'Analyse Factorielle des Correspondances (AFC), il est possible d'utiliser soit le langage R, soit le langage Python. Ci-dessous, les principales bibliothèques à installer et à charger selon votre choix de logiciel.

En R : Les packages suivants sont nécessaires pour effectuer une AFC complète (calculs + graphiques) :

- **FactoMineR** — permet de réaliser directement une AFC avec la fonction **CA()**.
- **factoextra** — pour la visualisation des résultats : biplots, inerties, \cos^2 , contributions, etc.
- **ggplot2** — pour personnaliser les graphiques.
- **dplyr** et **tidyR** — pour la manipulation des tableaux de données.

Exemple de chargement :

```
install.packages(c("FactoMineR", "factoextra", "ggplot2", "dplyr", "tidyR"))
library(FactoMineR)
library(factoextra)
```

En Python : Les bibliothèques suivantes sont recommandées pour implémenter une AFC :

- **pandas** — pour la lecture et la manipulation des tableaux de données.
- **numpy** — pour les calculs matriciels.
- **matplotlib.pyplot** — pour les graphiques.
- **scipy.linalg** — pour effectuer la décomposition en valeurs singulières (SVD).
- **scikit-learn (sklearn.preprocessing)** — pour normaliser les données si nécessaire.
- (Optionnel) **prince** — une bibliothèque dédiée aux analyses factorielles (AFC, ACP, ACM).

Exemple d'installation et d'importation :

```
pip install pandas numpy matplotlib scipy scikit-learn prince
```

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from scipy.linalg import svd
from sklearn.preprocessing import StandardScaler
import prince # optionnel
```