

TP 3 – Analyse des Correspondances Multiples (ACM)

🎯 Objectif

Découvrir comment appliquer l'ACM sur plusieurs variables qualitatives, et apprendre à interpréter les résultats graphiques :

- Scree plot (valeurs propres)
- Carte factorielle des individus
- Carte factorielle des modalités
- Biplot (individus + modalités)

1. Jeu de données

On considère un échantillon fictif d'étudiants, décrit par plusieurs **variables qualitatives** :

- **Sexe** : Homme / Femme
- **Régime alimentaire** : Végétarien / Omnivore / Végan
- **Boisson préférée** : Café / Thé / Jus
- **Sport préféré** : Foot / Basket / Natation

Chaque ligne = un étudiant.

```
import pandas as pd

# Création d'un petit dataset fictif
data = {
    "Sexe": ["H", "F", "H", "F", "H", "F", "H", "F", "H", "F"],
    "Régime": ["Omnivore", "Végétarien", "Omnivore", "Végan", "Végétarien",
               "Omnivore", "Végan", "Végétarien", "Omnivore", "Omnivore"],
    "Boisson": ["Café", "Thé", "Jus", "Café", "Thé",
                "Café", "Jus", "Thé", "Jus", "Café"],
    "Sport": ["Foot", "Basket", "Foot", "Natation", "Basket",
              "Foot", "Natation", "Basket", "Foot", "Natation"]
}
df = pd.DataFrame(data)
```

2. Application de l'ACM

```
import prince # si besoin : pip install prince
import matplotlib.pyplot as plt
import seaborn as sns

# Lancer l'ACM
mca = prince.MCA(n_components=4, random_state=42)
mca = mca.fit(df)
```

3. Valeurs propres et inertie

```
eigvals = mca.eigenvalues_
inertia = [val/sum(eigvals)*100 for val in eigvals]

print("Valeurs propres :", eigvals)
print("Inertie par axe (%) :", inertia)
print("Inertie cumulée Dim1+Dim2 :", sum(inertia[:2]))
```

Question 1 :

- Quelle est la part d'inertie expliquée par Dim1 et Dim2 ?
- Combien d'axes semble-t-il utile de garder ?

4. Scree plot (courbe des valeurs propres)

```
plt.figure(figsize=(6, 4))
sns.barplot(x=list(range(1, len(inertia)+1)), y=inertia, color="skyblue")
plt.plot(range(1, len(inertia)+1), inertia, marker="o", color="red")
plt.title("Scree plot (ACM)")
plt.xlabel("Axes")
plt.ylabel("% Inertie expliquée")
plt.show()
```

Question 2 :

- Identifiez le **coude** : à partir de quel axe les valeurs propres deviennent faibles ?

5. Carte factorielle des individus

```
fig, ax = plt.subplots(figsize=(6, 6))
mca.plot_coordinates(df, ax=ax, show_row_points=True, show_column_points=False)
plt.title("ACM - Carte des individus")
plt.show()
```

Question 3 :

- Les individus semblent-ils se regrouper ?
- Que cela signifie-t-il sur leurs préférences ?

6. Carte factorielle des modalités

```
fig, ax = plt.subplots(figsize=(6, 6))
mca.plot_coordinates(df, ax=ax, show_row_points=False, show_column_points=True)
plt.title("ACM - Carte des modalités")
plt.show()
```

Question 4 :

- Quelles modalités semblent proches (ex. Café + Omnivore) ?
- Quelles modalités s'opposent clairement ?

7. Biplot (individus + modalités)

```
fig, ax = plt.subplots(figsize=(6, 6))
mca.plot_coordinates(df, ax=ax, show_row_points=True, show_column_points=True)
plt.title("ACM - Biplot (Individus + Modalités)")
plt.show()
```

Question 5 :

- Que signifie la proximité entre un individu et une modalité ?
- Donnez un exemple de lecture (ex. un groupe d'étudiants "Végétarien + Thé + Basket").