

## Intelligence Artificielle et Données : feuille 3

### Analyse de données Analyse en Composantes Principales : ACP

#### Exercice 1.

On étudie les consommations annuelles en 1972, exprimées en francs, de 8 denrées alimentaires (les variables), les individus étant 8 catégories socio-professionnelles. Les données sont des moyennes par CSP :

|      | PAO | PAA | VIO | VIA | POT | LEC | RAI | PLP |
|------|-----|-----|-----|-----|-----|-----|-----|-----|
| AGRI | 167 | 1   | 163 | 23  | 41  | 8   | 6   | 6   |
| SAAG | 162 | 2   | 141 | 12  | 40  | 12  | 4   | 15  |
| PRIN | 119 | 6   | 69  | 56  | 39  | 5   | 13  | 41  |
| CSUP | 87  | 11  | 63  | 111 | 27  | 3   | 18  | 29  |
| CMOY | 103 | 5   | 68  | 77  | 32  | 4   | 11  | 30  |
| EMPL | 111 | 4   | 72  | 66  | 34  | 6   | 10  | 28  |
| OUVR | 130 | 3   | 76  | 52  | 43  | 7   | 7   | 16  |
| INAC | 138 | 7   | 117 | 74  | 53  | 8   | 12  | 20  |

|                                  |                      |
|----------------------------------|----------------------|
| AGRI = Exploitants agricoles     | PAO = Pain ordinaire |
| SAAG= Salariés agricoles         | PAA = Autre pain     |
| PRIN = Professions indépendantes | VIO = Vin ordinaire  |
| CSUP = Cadres supérieurs         | VIA=Autre vin        |
| CMOY= Cadres moyens              | POT= Pommes de terre |
| EMPL= Employés                   | LEC=Légumes secs     |
| OUVR = Ouvriers                  | RAI=Raisin de tables |
| INAC = Inactifs                  | PLP= Plats préparés  |

En utilisant python et les bibliothèques nécessaires , générer un rapport avec les éléments suivants :

- La matrice de corrélation
- Le tableau des valeurs propres et le tracé de ces valeurs
- Le tableau des vecteurs propres
- Les coordonnées et qualités des individus
- Les coordonnées des variables sur les axes

Répondez ensuite aux questions suivantes :

1. Quel est le nombre d'axes que vous retenir ? Justifiez votre réponse.
2. Interprétez les axes que vous avez retenus.

## Exercice 2. (A rendre)

Le tableau suivant fournit la structure du bilan d'un groupe pétrolier de 1969 à 1984 :

| Année | NET   | INT  | SUB  | LMT   | DCT   | IMM   | EXP   | VRD   |
|-------|-------|------|------|-------|-------|-------|-------|-------|
| 1969  | 17.93 | 3.96 | 0.88 | 7.38  | 19.86 | 25.45 | 5.34  | 19.21 |
| 1970  | 16.21 | 3.93 | 0.94 | 9.82  | 19.11 | 26.58 | 5.01  | 18.40 |
| 1971  | 19.01 | 3.56 | 1.91 | 9.43  | 17.87 | 25.94 | 5.40  | 16.88 |
| 1972  | 18.05 | 3.33 | 1.73 | 9.72  | 18.83 | 26.05 | 5.08  | 17.21 |
| 1973  | 16.56 | 3.10 | 2.14 | 9.39  | 20.36 | 23.95 | 6.19  | 18.31 |
| 1974  | 13.09 | 2.64 | 2.44 | 8.10  | 25.05 | 19.48 | 11.61 | 17.59 |
| 1975  | 13.43 | 2.42 | 2.45 | 10.83 | 22.07 | 22.13 | 11.17 | 15.49 |
| 1976  | 9.83  | 2.46 | 1.79 | 11.81 | 24.10 | 22.39 | 11.31 | 16.30 |
| 1977  | 9.46  | 2.33 | 2.30 | 11.46 | 24.45 | 23.07 | 11.16 | 15.77 |
| 1978  | 10.93 | 2.95 | 2.25 | 10.72 | 23.16 | 24.17 | 9.64  | 16.20 |
| 1979  | 13.02 | 3.74 | 2.21 | 7.99  | 23.04 | 19.53 | 12.60 | 17.87 |
| 1980  | 13.43 | 3.60 | 2.29 | 7.09  | 23.59 | 17.61 | 16.67 | 15.72 |
| 1981  | 13.37 | 3.35 | 2.58 | 6.76  | 23.94 | 18.04 | 15.42 | 16.54 |
| 1982  | 11.75 | 2.74 | 3.11 | 7.37  | 25.04 | 18.11 | 14.71 | 17.18 |
| 1983  | 12.59 | 3.05 | 3.85 | 7.12  | 23.40 | 19.17 | 11.86 | 18.97 |
| 1984  | 13.00 | 3.00 | 4.00 | 7.00  | 24.00 | 20.00 | 12.00 | 17.00 |

Les postes de bilan sont les suivants :

NET : Situation nette ; représente l'ensemble des capitaux propres de l'entreprise.

INT : Intérêts ; représente l'ensemble des frais financiers supportés par l'entreprise.

SUB : Subventions ; représente le montant total des subventions accordées par l'Etat.

LMT : Dettes à long et moyen terme.

DCT : Dettes à court terme.

IMM : Immobilisations ; représente l'ensemble des terrains et du matériel de l'entreprise.

EXP : Valeurs d'exploitation.

VRD : Valeurs réalisables et disponibles ; ensemble des créances à court terme de l'entreprise.

Les données ont été ventilées en pourcentage par année, la somme des éléments d'une même ligne vaut 100, de manière à éviter les effets dus à l'inflation.

Générez le rapport correspondant à l'ACP de ce cas et analysez les différents axes.