

exercice 1:

1. D. Toutes les réponses
2. C. Les réponses A et B sont correctes.
3. A. La progression du cours est corrélée positivement avec la première composante.
4. D. Le cours "apprenez à programmer en Python" est court.
5. D. les. Les réponses A, B et C.

exercice 2:

L'Analyse en Composantes Principales (ACP) est une technique d'analyse multivariée largement utilisée en informatique décisionnelle pour réduire la dimensionnalité des données, identifier des relations entre variables, et explorer la structure sous-jacente des données. Voici des réponses aux questions posées :

1. Objectifs de l'Analyse en Composantes Principales (ACP) en Informatique décisionnelle :

L'ACP vise principalement à atteindre les objectifs suivants en informatique décisionnelle :

- Réduire la dimensionnalité des données : L'ACP permet de condenser l'information contenue dans un grand nombre de variables en un nombre plus restreint de composantes principales, tout en conservant une grande partie de la variance des données.
- Identifier des relations entre les variables : Elle permet de mettre en évidence les corrélations ou les associations entre les variables, ce qui peut aider à comprendre comment ces variables sont liées les unes aux autres.
- Explorer la structure sous-jacente des données : L'ACP peut révéler des tendances, des patterns ou des groupes de variables qui ne sont pas immédiatement perceptibles dans les données brutes.
- Visualisation : En réduisant la dimensionnalité, l'ACP facilite la visualisation des données dans un espace à deux ou trois dimensions, ce qui peut aider à interpréter les résultats.

2. Étude du tableau des valeurs propres :

2.1. La somme des valeurs propres correspond à la variance totale des données. Elle est égale au nombre de variables (dans ce cas, 8).

2.2. Le choix de n'étudier que les deux premières composantes principales est justifié par le fait que ces deux composantes expliquent une proportion significative de la variance totale des données. Le tableau des valeurs propres (Eigenvalues) indique que la première composante principale explique 55.88% de la variance, et la deuxième en explique 26.44%. Ensemble, elles expliquent environ 82.32% de la variance totale. C'est une proportion suffisamment élevée pour réduire la dimensionnalité tout en conservant une grande partie de l'information.

3. Analyse des résultats de l'ACP :

3.1. Pour sélectionner les individus (les années) bien représentés, on se base sur le \cos^2 (carré de la corrélation entre les individus et les composantes principales). Les années avec des valeurs de \cos^2 élevées sont bien représentées. Dans le tableau 2, les années 1969, 1970, 1971, 1972, 1973, 1978, 1979, 1980, 1981, 1982, 1983, et 1984 ont des \cos^2 élevés, ce qui signifie qu'elles sont bien représentées sur le plan factoriel.

3.2. Pour sélectionner les variables corrélées avec les premières composantes principales, on se base sur les coefficients de corrélation de Pearson dans le tableau 3. Les variables ayant des coefficients élevés en valeur absolue sont fortement corrélées avec les

composantes principales. Dans ce cas, NET, INT, SUB, LMT, DCT, IMM, EXP, et VRD ont des corrélations significatives avec les composantes principales.

3.3. En analysant la position des années bien représentées par rapport aux variables corrélées, on peut observer comment ces années se distinguent les unes des autres en fonction des variables. Par exemple, on peut comparer les années 1969 et 1970, qui sont bien représentées, et observer comment elles diffèrent en termes de variables corrélées pour mieux comprendre l'évolution de la structure de bilan.

L'ACP permet donc de synthétiser l'information contenue dans un grand nombre de variables en quelques composantes principales, ce qui facilite l'interprétation des données et l'identification des tendances et des relations entre les années et les variables du bilan du groupe pétrolier sur la période de 1969 à 1984.

Examinons les questions une par une :

****I. Etude du tableau des valeurs propres : ****

1. La somme des valeurs propres correspond à la somme cumulée des pourcentages de variance expliquée, soit 100%.

2. On choisit de n'étudier que les deux premières composantes principales car elles expliquent ensemble 84,92% de la variance totale, ce qui représente une proportion significative de la variance totale.

****II. Etude des qualités de représentation dans le premier plan principal : ****

3. La qualité de représentation des individus par le plan factoriel est quantifiée par la somme des carrés des cosinus des coordonnées des individus sur les axes factoriels.

4. L'individu le moins bien représenté par le premier plan principal est I15, et l'individu le mieux représenté est I9.

****III. Etude du nuage des individus : ****

5. Les individus dont la contribution à la formation de la première composante principale est supérieure à la moyenne sont : I1, I4, I7, I9, et I10. Pour chacun d'eux, le signe de la coordonnée correspondante est négatif. Cela caractérise l'axe en termes d'opposition entre ces individus.

6. Pour la deuxième composante principale, les individus dont la contribution est supérieure à la moyenne sont : I8, I11, I3, et I14. Pour chacun d'eux, le signe de la coordonnée correspondante est positif. Cela caractérise l'axe en termes d'opposition entre ces individus.

****IV. Etude du nuage des variables : ****

7. La représentation graphique des variables montre qu'elles sont toutes très bien représentées dans le plan (CP1, CP2) car elles ont des \cos^2 élevés, ce qui indique une proximité avec le plan factoriel.

8. Les variables qui sont corrélées positivement avec le premier facteur principal sont CAL, CUB, et PUZ, tandis que les variables qui sont corrélées négativement sont COM, MEM, et VOC. Cet axe peut être caractérisé comme représentant des capacités cognitives.

9. Les variables qui ont joué un rôle dominant dans la formation du deuxième axe sont VOC, COM, et MEM.

****Exercice 3 : ****

5. Le principe de l'analyse en composantes principales est de réduire la dimensionnalité des données tout en conservant autant d'informations que possible.

6. Le choix de n'étudier que les deux premières composantes principales est justifié par le fait qu'elles expliquent ensemble une part importante de la variance totale.

7. Le pourcentage d'information quantifié par les deux premières composantes principales sélectionnées est de 85,77%.

8.1. Les variables corrélées avec les deux premières composantes principales peuvent être identifiées à partir de la matrice des composantes après rotation.

8.2. Les conclusions adéquates pour cette étude peuvent être tirées en analysant les deux graphiques résultants de l'ACP, en observant la distribution des points et la contribution des variables.