

Valorisation d'une Production Professionnelle

Projet : Déchiffrer les Clés de la Réussite Étudiante

Rapport de synthèse démontrant l'acquisition des compétences de la SAE 4.02 :
Reporting d'une Analyse Multivariée

Auteur : Aimad HAMDAOUI

Formation : BUT Science des Données - 2ème année

Promotion : 2023-2026

Établissement : Université d'Avignon et des Pays de Vaucluse

Date de réalisation du projet : Mai 2025

Table des matières

1	Introduction et Contexte du Projet	2
2	Niveau 1 : Contextualiser et Présenter les Données	2
2.1	Contextualisation, Démarche et Data Visualisation	2
3	Niveau 2 : Restituer et Argumenter ses Résultats	3
3.1	Défense des Choix et Rigueur de l'Analyse	3
4	Conclusion	5

1 Introduction et Contexte du Projet

Ce document a pour objectif de démontrer comment le projet d'analyse de données sur la réussite étudiante, mené dans le cadre de la SAE 4.02, répond aux compétences requises pour **valoriser une production dans un contexte professionnel**.

Le projet consistait à analyser un jeu de données de 480 étudiants (système Kalboard 360) pour identifier les facteurs comportementaux et environnementaux influençant la performance académique. Pour ce faire, une série de méthodes d'analyse multivariée (ACP, ACM, Classification, AFC) a été mise en œuvre.

Ce rapport est structuré selon le référentiel de compétences fourni, illustrant point par point l'atteinte des objectifs de chaque niveau.

Valoriser une production dans un contexte professionnel

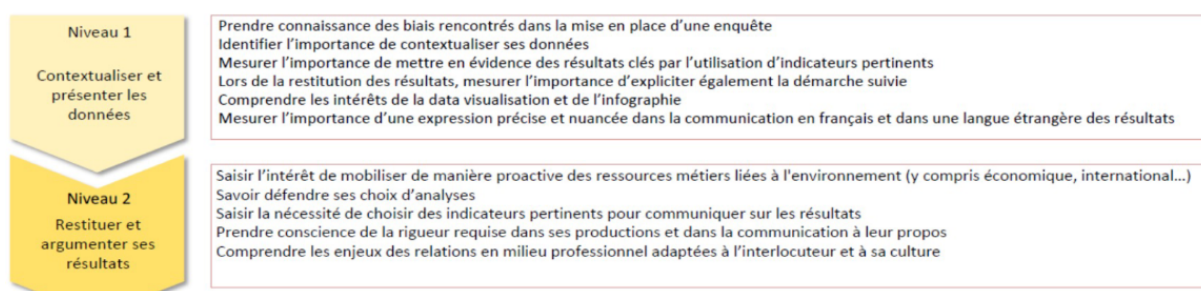


FIGURE 1 – Référentiel de compétences "Valoriser une production dans un contexte professionnel".

2 Niveau 1 : Contextualiser et Présenter les Données

Ce premier niveau de compétence a été entièrement validé à travers la phase préparatoire et la restitution initiale de l'analyse.

2.1 Contextualisation, Démarche et Data Visualisation

- **Contextualiser les données :** Le rapport original débute par une section (1.1 "Contexte et Problématique") qui pose clairement le cadre de l'étude. La source des données (Kaggle, Kalboard 360) et la description des 16 variables (Figure 2 du rapport) ont permis de donner un sens métier aux informations brutes.
- **Explicitement la démarche suivie :** La section 1.6 ("Choix et Justification des Méthodes d'Analyse") a été cruciale. Elle explique *pourquoi* chaque méthode a été choisie (ACP pour résumer l'engagement, ACM pour les liens qualitatifs, etc.), démontrant une démarche réfléchie et non une simple application technique.

- **Comprendre les intérêts de la data visualisation** : L'ensemble du rapport s'appuie sur des visualisations pour rendre les résultats complexes accessibles. Le cercle des corrélations de l'ACP (voir Figure 2 ci-dessous) a permis de visualiser instantanément la structure de l'engagement étudiant, tandis que les biplots de l'ACM et de l'AFC ont cartographié des relations multidimensionnelles.
- **Expression précise et nuancée** : Le langage utilisé dans le rapport est volontairement prudent et précis. Des termes comme "association", "tendance", "profil" ont été employés pour éviter les conclusions hâtives de causalité. La section 6.5 ("Points Importants à Garder à l'Esprit") insiste sur la distinction entre corrélation et causalité, démontrant la rigueur de l'interprétation.

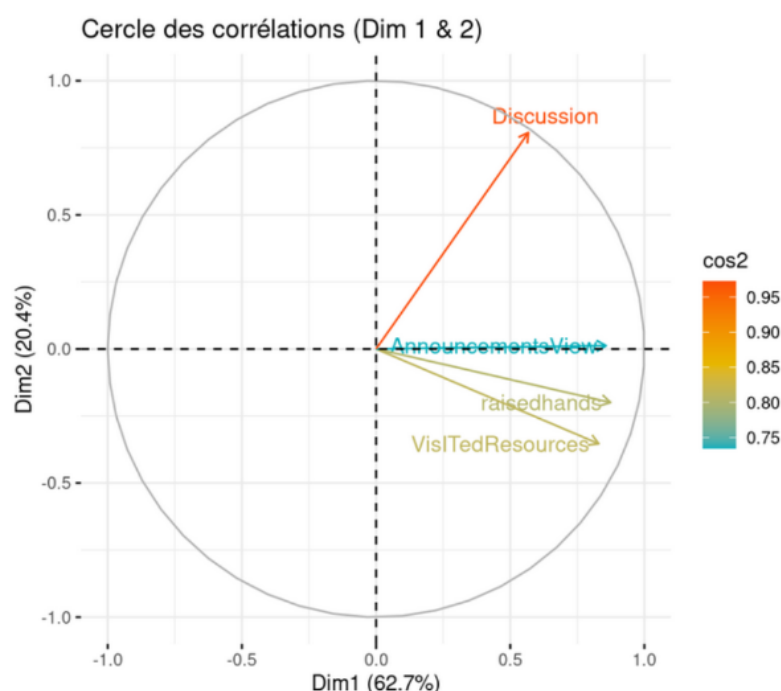


FIGURE 2 – Exemple de visualisation : le cercle des corrélations de l'ACP, montrant que toutes les mesures d'engagement contribuent à un axe principal "d'Engagement Général" (Dim 1).

3 Niveau 2 : Restituer et Argumenter ses Résultats

La seconde partie de l'analyse, la synthèse et les recommandations, ont permis de valider les compétences de ce niveau avancé.

3.1 Défense des Choix et Rigueur de l'Analyse

- **Savoir défendre ses choix d'analyses** : Comme mentionné précédemment, la justification des méthodes (section 1.6) était un pilier du rapport. Par exemple, le choix de combiner une CAH (pour déterminer le nombre de clusters) et un K-means (pour affiner les groupes) a été une décision méthodologique défendue pour garantir la robustesse de la classification des profils étudiants.
- **Choisir des indicateurs pertinents** : L'ACP a permis de construire des indicateurs synthétiques pertinents ("Engagement Général" et "Style d'Engagement") à partir des variables brutes. De même, l'ACM (voir Figure 3) a cartographié les proximités entre des dizaines de modalités de variables, offrant une vue d'ensemble intelligible des liens entre le profil scolaire, familial et la performance.

- **Prendre conscience de la rigueur requise** : La rigueur se manifeste dans la validation statistique des résultats. Par exemple, le test du Chi-deux (Figure 9 du rapport) a été utilisé avant l'AFC pour confirmer statistiquement l'existence d'un lien entre la matière et la performance ($p\text{-value} = 0.002522$), justifiant ainsi l'analyse approfondie qui a suivi (Figure 4).
- **Comprendre les enjeux des relations en milieu professionnel** : Le rapport ne s'arrête pas aux constats. La section 7 ("Conclusion Générale et Perspectives") et notamment 7.3 ("Recommandations Pratiques") traduit les résultats statistiques en conseils concrets pour les enseignants, les établissements et les parents. C'est la preuve d'une compréhension des enjeux et de la nécessité de fournir des résultats *actionnables* à un interlocuteur (ici, le milieu éducatif).

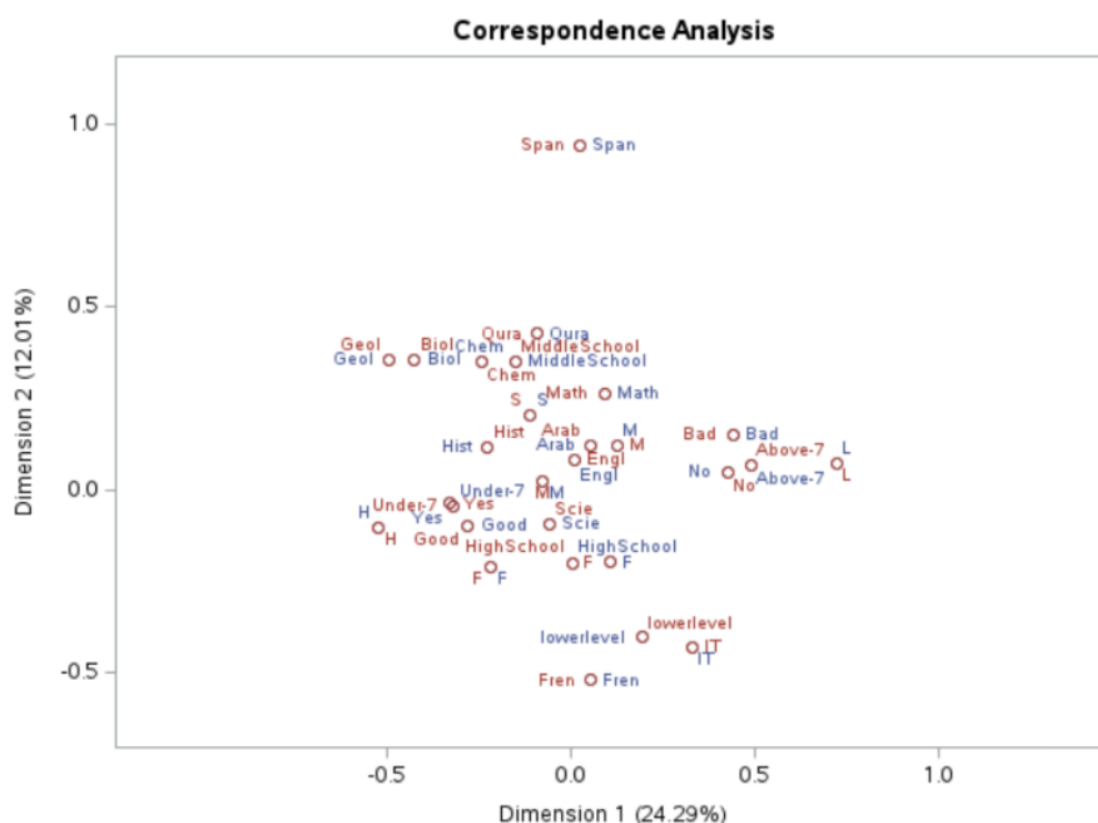


FIGURE 3 – Carte des relations (ACM) : une restitution complexe qui synthétise les liens entre toutes les variables qualitatives, opposant un profil "performant" (gauche) à un profil "en difficulté" (droite).

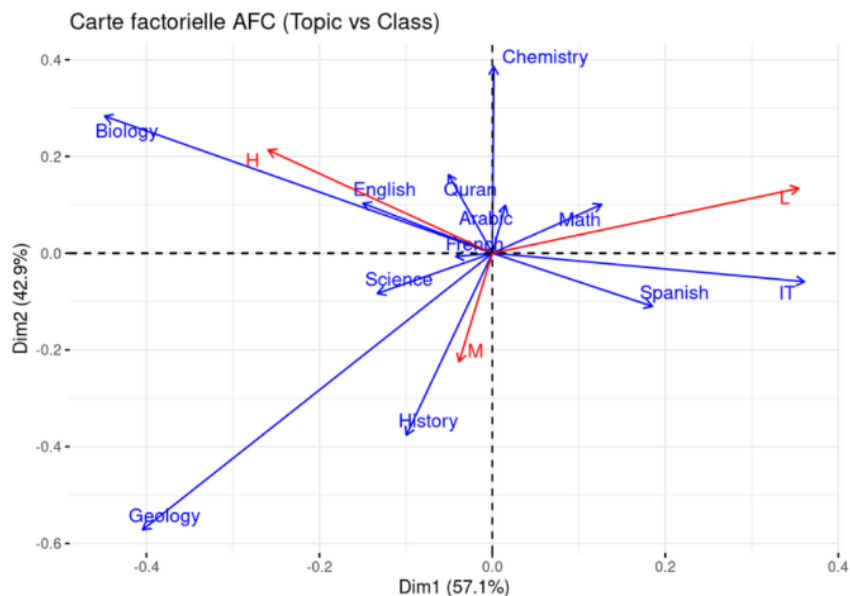


FIGURE 4 – Carte factorielle (AFC) : un zoom argumenté sur le lien entre Matière et Performance, montrant une forte association entre l'IT et la faible performance (L) et entre la Biologie et la haute performance (H).

4 Conclusion

Ce projet d'analyse de la réussite étudiante a été une application complète et rigoureuse des techniques d'analyse de données multivariées. En partant d'un jeu de données brut, une démarche structurée a été mise en place, alliant la préparation des données, le choix justifié des méthodes, l'analyse statistique et la visualisation.

Les résultats ont été non seulement présentés de manière claire, mais aussi interprétés avec prudence et traduits en recommandations pratiques. Le projet démontre ainsi une maîtrise des compétences techniques et une capacité à **valoriser une production analytique dans un contexte professionnel**, en rendant les conclusions intelligibles, argumentées et utiles pour un public cible.