

PROJET ANALYSE DE DONNÉES : SONDAGE FST

Module : Analyse de Données

Réalisée Par :

- KHOUDRAJI OUIAM
- CHIHAB AMAL
- EL HASSNAOUI AYMANE

Encadrée Par :

- Mr. Ouabarab Aziz





REMERCIEMENT

Au nom de notre équipe, nous tenons à exprimer notre sincère gratitude à tous ceux qui ont contribué à la réalisation de ce projet de fin de module.

Nous souhaitons remercier chaleureusement notre professeur, monsieur Ouaarab Aziz, pour son soutien et son expertise tout au long de ce semestre. Nous tenons également à remercier nos collègues et coéquipiers, pour leur collaboration, leur dévouement et leur travail acharné.

Un grand merci à la faculté des sciences et techniques pour avoir créé un environnement propice à l'apprentissage et à la recherche.

Nous sommes reconnaissantes envers toutes les personnes qui ont généreusement participé à notre projet, que ce soit en fournissant des données, en offrant leur expertise ou en apportant leur soutien. Leur contribution a été essentielle pour le succès de ce projet.

Enfin, nous souhaitons exprimer notre gratitude envers toutes les personnes qui ont contribué de près ou de loin à la réussite de ce projet. Votre aide et votre soutien ont été essentiels, et nous sommes reconnaissants de pouvoir compter sur vous.

Merci à tous pour votre implication, votre collaboration et votre dévouement. Nous sommes fiers de notre travail accompli ensemble.



SOMMAIRE

REMERCIEMENT	2
SOMMAIRE	3
INTRODUCTION	4
DESCRIPTION DU JEU DE DONNÉE	5
LA DÉMARCHE CHOISIE ET LES APPORTS DE L'ANALYSE.....	6
Analyse des fréquences :.....	6
Répartition des formations :	7
Test 2D (table de contingence) :	7
○ La relation entre la formation et le parcours après le diplôme des individus :	7
○ La relation entre le sexe et assister au cours des individus :	8
Traitement des questions à choix binaires:.....	9
Analyse de relation entre formation et la colonne personnalité :	10
Analyse des correspondances multiples (ACM) :	10
Classification :	13
K-means :	13
Analyse discriminante :	14
Anova a un facteur :	14
Analyse de l'effet :	15
CONCLUSION	15

INTRODUCTION

Ce projet a pour objectif d'analyser le comportement des étudiants de la Faculté des Sciences et Techniques (FST) à l'aide d'un questionnaire. Les données recueillies comprennent des informations sur leur profil démographique, leurs pratiques d'études, leurs aspirations professionnelles et leurs perceptions personnelles.

Le jeu de données que nous avons à notre disposition contient les réponses des étudiants à plusieurs questions. Ces questions portent sur des sujets variés tels que leur formation, leur niveau d'étude, leur nationalité, leur temps de trajet domicile-université, leur fréquence d'assistance aux cours, la composition de leur entourage, leur sentiment d'intégration, leurs pratiques d'études, leurs croyances sur l'intelligence et l'apprentissage, ainsi que leurs réactions face aux défis et aux erreurs.

Dans le cadre de ce projet, nous allons entreprendre une analyse approfondie de ces données en suivant les étapes principales de la production d'un rapport d'analyse. Nous commencerons par nettoyer et importer les données brutes, puis nous les présenterons sous forme de tableaux appropriés. Ensuite, nous procéderons à une description détaillée des variables mesurées, de la population d'individus et des conditions de collecte des données.

Nous utiliserons ensuite différentes méthodes d'analyse pour explorer les données. Nous effectuerons une analyse descriptive, en calculant des statistiques univariées et bivariées, afin de dégager des tendances, des corrélations et des différences significatives entre les groupes d'étudiants.

Pour une exploration plus approfondie, nous mettrons en œuvre des méthodes exploratoires multivariées telles que l'Analyse Factorielle des Correspondances (AFC), l'Analyse des Correspondances Multiples (ACM) et l'Analyse Factorielle Discriminante (AFD), en fonction de la nature des données et des objectifs d'analyse.

Nous procéderons également à une classification des étudiants en utilisant des méthodes de calcul de distances et en appliquant l'AFC ou l'ACM pour obtenir des groupes homogènes.

Enfin, nous analyserons la dimension de la personnalité des étudiants en utilisant des scores qui varient de manière croissante ou fixe. Nous interpréterons ces scores pour identifier les caractéristiques personnelles des étudiants.

L'ensemble de ces analyses nous permettra de fournir une description complète des données et de dégager des implications pertinentes pour mieux comprendre le comportement des étudiants de la FST.

DESCRIPTION DU JEU DE DONNÉE

Le jeu de données comprend les variables suivantes :

- **Timestamp** : Horodatage de la réponse.
- **Sexe** : Genre de l'étudiant.
Modalités : Masculin/Féminin.
- **Votre formation** : Formation suivie par l'étudiant.
Modalités : Tronc commun, LST, Ingénieur, Master, Formation continue, Doctorat.
- **Niveau d'étude** : Niveau d'études de l'étudiant.
Modalités : Bac+1, Bac+2, Bac+3, Bac+4, Bac +5, Plus.
- **Avez-vous la nationalité marocaine ?** : Indication de la nationalité marocaine de l'étudiant.
Modalités : Oui, seulement celle-ci ; Oui et une autre ou d'autre nationalité(s) ; Non.
- **Que comptez-vous faire après l'obtention de votre diplôme** : Plan de carrière envisagé par l'étudiant.
Modalités : Continuer dans un autre parcours de formation, Travailler dans le public, Travailler dans le privé, Créer mon entreprise, Je ne sais pas.
- **Quelle est la durée de trajet de votre domicile à la faculté en minutes ?** : Temps de trajet domicile-faculté.
Modalités : Moins de 15, Entre 15 et 30, Entre 30 et 45, Entre 45 et 60, plus de 60.
- **Vous assistez aux cours...** : Fréquence d'assistance aux cours.
Modalités : Systématiquement, Assez régulièrement, Rarement, Jamais.
- **Votre entourage que vous fréquentez, est-il composé d'étudiants ayant globalement le même âge que vous ?** : Composition de l'entourage fréquenté par l'étudiant en termes d'âge.
Modalités : Ayant globalement le même âge que vous ; Plus jeunes que vous ; Plus âgé(e)s que vous, D'âges différents.
- **Votre entourage que vous fréquentez, est-il composé d'étudiants ayant le même niveau d'études que vous ?** : Composition de l'entourage fréquenté par l'étudiant en termes de niveau d'études.
Modalités : Ayant le même niveau d'étude que vous ; De niveau inférieur ; de niveau supérieur ; Redoublants ; Issus de différents niveaux d'études.
- **Vous vous sentez pleinement intégré(e) à la population étudiante de votre formation et/ou de votre université.** : Sentiment d'intégration à la population étudiante.
Modalités : Pas du tout d'accord ; Pas vraiment d'accord ; Ni d'accord, ni pas d'accord ; Plutôt d'accord ; Tout à fait d'accord.
- **Quelle(s) pratique(s) mettez-vous en œuvre pour vos études ?** : Pratiques mises en œuvre par l'étudiant pour ses études.
Modalités : Noter tout ce que l'enseignant dit en cours ; Compléter les cours par des informations complémentaires ; Réviser et relire les cours ; Retravailler les cours en faisant des fiches, des résumés ; Faire les exercices proposés par les enseignant(e)s ; Poser des questions aux enseignant(e)s ; Travailler les annales des examens ; Avoir des rendez-vous avec les enseignant(e)s ; Faire d'autres exercices que ceux proposés par les enseignant(e)s, avec des modalités : jamais, Presque tous les jours, Environ une fois par semaine, Plusieurs fois par semaine, Une fois ou deux.

- Si je dois travailler dur à quelque chose, cela veut dire que je ne suis pas intelligent(e). : Opinion sur la corrélation entre le travail dur et l'intelligence.
Modalités : Vrai ou faux.
- J'aime essayer des choses difficiles. : Attitude envers les défis difficiles.
Modalités : Vrai ou faux.
- Lorsque je fais une erreur, j'ai honte ou je suis gêné(e). : Réaction face aux erreurs.
Modalités : Vrai ou faux.
- J'aime qu'on me dise que je suis intelligent(e). Réaction face aux compliments sur l'intelligence.
Modalités : Vrai ou faux.
- J'abandonne généralement quand quelque chose devient difficile ou frustrant. : Tendance à abandonner face à la difficulté.
Modalités : Vrai ou faux.
- Ça ne me dérange pas de faire des erreurs car elles m'aident à apprendre. : Perception des erreurs comme un moyen d'apprentissage.
Modalités : Vrai ou faux.
- Il y a des choses pour lesquelles je ne serai jamais doué(e): Perception des capacités et aptitudes n'importe qui peut apprendre certaines choses s'il y travaille dur. : Croyance en l'apprentissage par le travail dur.
Modalités : Vrai ou faux.
- Les gens naissent soit stupides, moyens ou intelligents et ils ne peuvent changer cela par la suite. : Croyance en la fixité des capacités intellectuelles.
Modalités : Vrai ou faux.
- Je suis fier(e) quand je fais de mon mieux, même si ce n'est pas parfait : Réaction face aux efforts personnels.
Modalités : Vrai ou faux.

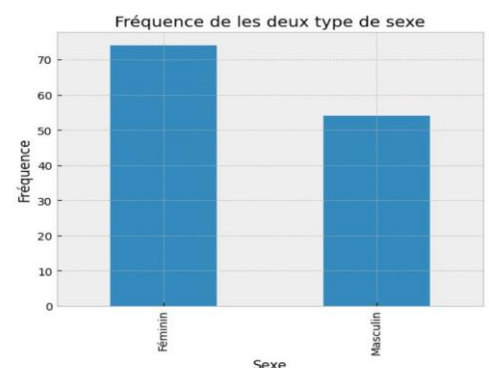
LA DÉMARCHE CHOISIE ET LES APPORTS DE L'ANALYSE

Analyse des fréquences :

Dans cette étape, nous avons calculé les fréquences relatives des hommes et des femmes dans notre dataset, cela nous a permis de déterminer la répartition proportionnelle des deux sexe , le graphe suivant represente la repartition de sexe :

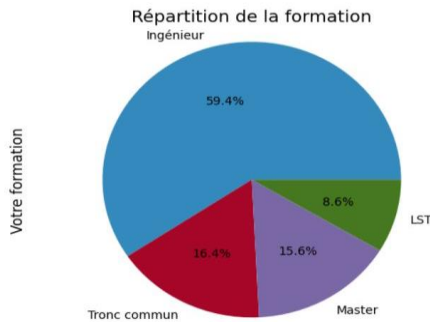
Cette analyse initiale nous a donné une idée de la représentativité de chaque groupe dans notre échantillon.

Dans notre dataset , on remarque que les femmes sont plus nombreuses que les hommes .



Répartition des formations :

Nous avons ensuite examiné la répartition des formations au sein de notre échantillon. Pour cela, nous avons identifié les différentes catégories de formations présentes dans notre dataset, le graphe suivant représente la répartition de la formation :



Nous avons calculé les fréquences relatives de chaque catégorie de formation afin de comprendre comment elles étaient réparties parmi les individus de notre échantillon.

Dans notre dataset, on remarque que la formation ingénieur domine, suivi par tron commun.

Test 2D (table de contingence) :

Le test 2D, également connu sous le nom de table de contingence, est une étape permettant d'évaluer les relations entre les variables explicatives, nous avons construit 2 tables de contingence croisant les catégories de deux variables

La relation entre la formation et le parcours après le diplôme des individus :

Votre formation	Ingénieur	LST	Master
Que comptez-vous faire après l'obtention de vot...			
Continuer dans un autre parcours de formation	13	6	5
Créer mon entreprise	12	1	2
Je ne sais pas	6	2	3
Travailler dans le privé	35	1	8
Travailler dans le public	10	1	2

Cette table représente la répartition des individus de chaque formation selon leur choix après le diplôme.

Interprétations :

- Ces chiffres indiquent que les diplômés en ingénierie ont une forte proportion d'individus intéressés par la poursuite de leur formation ou par l'entrepreneuriat, ainsi qu'une préférence

Votre formation	Tronc commun
Que comptez-vous faire après l'obtention de vot...	
Continuer dans un autre parcours de formation	14
Créer mon entreprise	0
Je ne sais pas	6
Travailler dans le privé	0
Travailler dans le public	1

marquée pour travailler dans le secteur privé.

- Ces chiffres suggèrent que les diplômés en LST ont une proportion relativement faible d'individus envisageant la poursuite de leur formation ou l'entrepreneuriat. La majorité envisage de continuer dans un autre parcours de formation.
- Ces chiffres indiquent que les diplômés de master ont une proportion significative d'individus intéressés par la poursuite de leur formation, ainsi qu'une préférence pour travailler dans le secteur privé.

- Ces chiffres suggèrent que les individus ayant suivi un tronc commun ont une proportion importante d'individus intéressés par la poursuite de leur formation, mais moins d'intérêt pour l'entrepreneuriat ou le travail dans le secteur privé.
- | Votre formation | Ingénieur | LST |
|--|-----------|----------|
| Que comptez-vous faire après l'obtention de vot... | | |
| Continuer dans un autre parcours de formation | 0.342105 | 0.157895 |
| Créer mon entreprise | 0.800000 | 0.066667 |
| Je ne sais pas | 0.352941 | 0.117647 |
| Travailler dans le privé | 0.795455 | 0.022727 |
| Travailler dans le public | 0.714286 | 0.071429 |
-
- | Votre formation | Master | Tronc commun |
|--|----------|--------------|
| Que comptez-vous faire après l'obtention de vot... | | |
| Continuer dans un autre parcours de formation | 0.131579 | 0.368421 |
| Créer mon entreprise | 0.133333 | 0.000000 |
| Je ne sais pas | 0.176471 | 0.352941 |
| Travailler dans le privé | 0.181818 | 0.000000 |
| Travailler dans le public | 0.142857 | 0.071429 |

(pourcentages) :

Cette table va nous aider d'examiner les choix dominants dans l'esprit des différentes formations concernant le parcours après le diplôme .

Interprétations :

- Les individus ayant suivi une formation en ingénierie ont une forte association avec le choix de travailler dans le secteur privé (coefficient de contingence de 0.795455). Cela suggère que les étudiants en ingénierie sont plus enclins à chercher des opportunités d'emploi dans le secteur privé après l'obtention de leur diplôme. Cette tendance peut être due à la demande élevée de compétences techniques dans le secteur privé.
- Les individus ayant suivi une formation en LST ont une probabilité relativement faible de créer leur propre entreprise (coefficient de contingence de 0.066667). Cela indique que les diplômés en LST sont moins enclins à entreprendre et à se lancer dans des activités entrepreneuriales après l'obtention de leur diplôme. Ils peuvent préférer des carrières plus axées sur l'emploi salarié ou la poursuite de leurs études.

○ La relation entre le sexe et assister au cours des individus :

Interprétations :

Pour les personnes de sexe féminin, 11 d'entre elles suivent assez régulièrement des cours, 6 le font rarement et 57 le font systématiquement.

Pour les personnes de sexe masculin, 17 d'entre elles suivent assez régulièrement des cours, 2 le font rarement et 35 le font systématiquement.

cours	Assez régulièrement	Rarement	Systematiquement
Sexe			
Féminin	11	6	57
Masculin	17	2	35
Coefficient de contingence :			
cours	Assez régulièrement	Rarement	Systematiquement
Sexe			
Féminin	0.148649	0.081081	0.770270
Masculin	0.314815	0.037037	0.648148

Le coefficient de contingence mesure la relation entre les deux variables (sexe et fréquence des cours). Les valeurs entre 0 et 1 indiquent le degré de dépendance entre les variables. Plus la valeur est proche de 1, plus il y a une forte relation entre les variables.

En utilisant le coefficient de contingence, on peut observer que la fréquence des cours est plus fortement liée au sexe chez les hommes que chez les femmes. Cela est suggéré par les coefficients de contingence plus élevés pour les hommes dans les catégories "Assez régulièrement" (0.314815) et "Systématiquement" (0.648148). Pour les femmes, les coefficients de contingence sont plus bas dans ces catégories, ce qui indique une relation moins forte entre le sexe et la fréquence des cours.

Traitement des questions à choix binaires:

Dans notre étude, nous avons utilisé un ensemble de questions fermées pour évaluer si la personnalité est perçue comme fixe ou croissante. Les participants ont répondu à un total de 10 questions, chacune avec deux choix possibles : "vrai" ou "faux". Sur la base de ces réponses, nous avons défini des critères pour déterminer si la personnalité est considérée comme fixe ou croissante.

Pour la perspective de la personnalité fixe :

Réponses "vrai" pour les questions : 1, 3, 4,5, 7, 8,9

Réponses "faux" pour les questions : 2, 6, 10

Pour la perspective de la personnalité croissante :

Réponses "vrai" pour les questions : 2, 6, 10

Réponses "faux" pour les questions : 1, 3, 4,5, 7, 8,9

En analysant les réponses des participants, nous avons calculé un score basé sur leur tendance vers une perspective fixe ou croissante de la personnalité. Si le score était supérieur à 5, cela indiquait une prédominance de la perspective de la personnalité fixe, tandis qu'un score inférieur ou égal à 5 révélait une tendance vers la perspective de la personnalité croissante.

Cette approche nous permet de catégoriser les participants en fonction de leur perception de la personnalité en tant que fixe ou croissante, en utilisant des questions spécifiques pour chaque perspective. Cela nous permettra d'explorer les corrélations entre ces croyances et d'autres variables d'intérêt dans notre étude, et d'approfondir notre compréhension de la manière dont les individus perçoivent le développement et le changement de la personnalité.

Analyse de relation entre formation et la colonne personnalité :

Votre formation	Ingénieur	LST	Master	Tronc commun
personality				
croissant	70	10	20	20
fixe	6	1	0	1

Chi-square statistic: 1.9150907571960203
P-value: 0.5902155658041921
Degrees of freedom: 3
Expected frequencies: [[71.25 10.3125 18.75 19.6875]
[4.75 0.6875 1.25 1.3125]]

Interprétations :

L'interprétation des données de la table de contingence montre qu'il y a une concentration plus élevée de personnes ayant la formation d'Ingénieur et ayant une personnalité croissante. Avec 70 personnes dans cette catégorie, elle représente le groupe le plus important parmi tous les groupes de formation et de personnalité.

En revanche, il y a seulement 6 personnes ayant la formation d'Ingénieur et ayant une personnalité fixe. Ce chiffre est beaucoup plus bas en comparaison aux autres combinaisons

de formation et de personnalité.

L'analyse du test du Chi-carré et la p-value associée indiquent qu'il n'y a pas suffisamment de preuves pour établir une relation significative entre la formation et la personnalité dans l'échantillon étudié. Cela signifie que la répartition observée des formations et des personnalités pourrait être due au hasard plutôt qu'à une relation véritable entre ces variables.

Analyse des correspondances multiples (ACM) :

L'Analyse des Correspondances Multiples (ACM) est une technique d'analyse factorielle adaptée aux variables qualitatives. Elle permet d'étudier les relations entre les différentes variables explicatives et de visualiser ces relations dans un espace multidimensionnel. L'ACM nous a aidés à comprendre les regroupements, les tendances et les structures sous-jacentes dans nos données qualitatives.

Remarque:

Étant donné que l'interprétation de la totalité des variables de notre jeu de données s'est avérée difficile en raison de leur nature illisible, nous avons pris la décision de nous concentrer sur deux études d'Analyse en Composantes Principales (ACP) afin d'obtenir une vision plus claire de la dataset. Cette approche nous a permis d'obtenir des résultats significatifs et d'effectuer une interprétation plus approfondie des données.

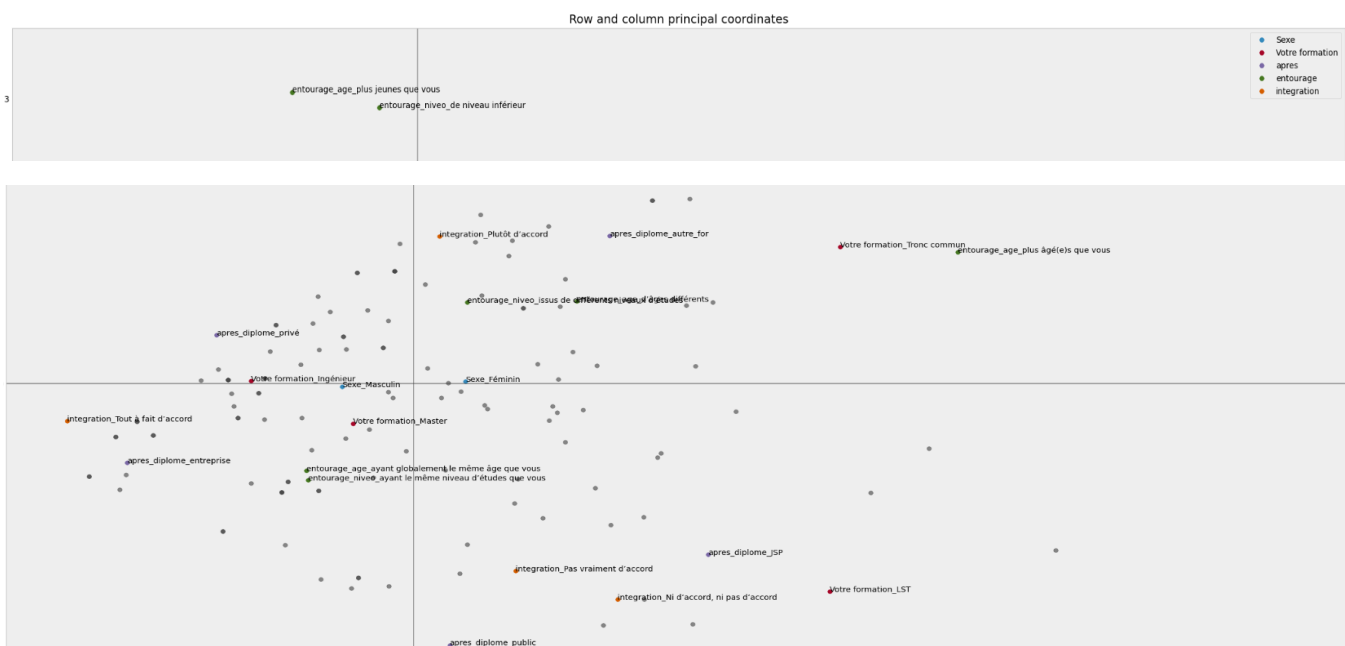
Les graphes ci-dessous représente les individus et les variables dans un plan factoriel afin d'observer les associations et les différences entre les catégories de variables explicative :



- Les étudiants ont tendance à ne jamais prendre de rendez-vous avec les enseignants.
- Le sexe féminin est majoritaire dans notre jeu de données.
- Les étudiants posent fréquemment des questions une ou deux fois.
- Les étudiants qui révisent le cours une fois par semaine font également des exercices une ou deux fois.
- Les étudiants en ingénierie sont principalement des femmes qui prennent des notes sur le cours au moins une fois par semaine.

- Les étudiants qui résument le cours une fois par semaine prennent des rendez-vous avec l'enseignant une ou deux fois.
- Les étudiants qui prennent des notes sur le cours une ou deux fois par semaine révisent également le cours une ou deux fois par semaine.
- Les étudiants qui prennent des notes sur le cours une fois par semaine sont généralement ceux qui font des recherches supplémentaires une ou deux fois et qui sont en tronc commun.
- Les étudiants qui ne prennent jamais de notes sur le cours font des exercices une ou deux fois par semaine.
- Les étudiants qui résument toujours le cours prennent des rendez-vous avec l'enseignant plus d'une fois par semaine et travaillent toujours sur les examens.
- Les étudiants qui posent toujours des questions font toujours des recherches supplémentaires et révisent toujours le cours.
- Les étudiants qui font des exercices plus d'une fois par semaine révisent également le cours plus d'une fois par semaine.
- Les étudiants qui posent des questions une fois par semaine font des recherches supplémentaires une fois par semaine, révisent le cours une fois par semaine et font des exercices une fois par semaine.
- Les modalités "faire_exercices_jamais", "informations_complementaires_jamais", "resumer_cours_jamais", "travailler_examens_jamais", "poser_questions_jamais", "faire_autres_exercices_jamais", "noter_cours_jamais", "faire_exercices_1_ou_2" sont des modalités rares, c'est-à-dire que ces réponses sont les moins choisies par les étudiants.
- De plus, les modalités "poser_questions_toujours", "informations_complementaires_toujours", "resumer_cours_toujours", "faire_autres_exercices_toujours", "travailler_examens_toujours" et "rendez-vous_enseignant_plusieurs_fois_par_semaine" sont également des modalités rares.
- Les modalités "rendez-vous_jamais", "votre_formation_1st", "votre_formation_ingénieur", "poser_questions_1_ou_2" et "sexe_feminin" sont des modalités fréquentes dans notre jeu de données.

- Examinant la relation entre formation , l'entourage , sexe , parcours après diplôme et intégration .



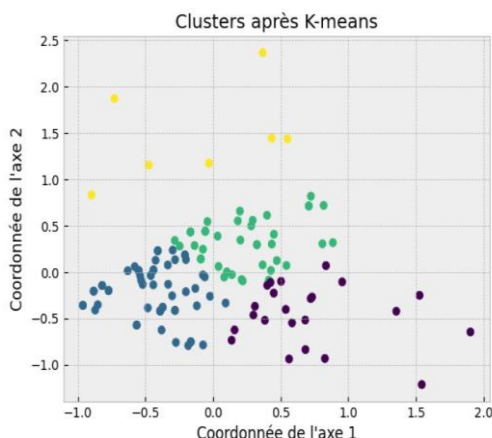
Interprétations :

- Les étudiants qui ont un entourage d'âge globalement du même âge ont également un entourage ayant globalement le même niveau.
- Généralement, les ingénieurs veulent continuer à travailler dans le secteur privé.
- Les étudiants qui sont totalement intégrés ont tendance à choisir de créer une entreprise après le diplôme.
- Les étudiants qui ne savent pas quoi faire après le diplôme ne sont pas vraiment intégrés.
- Les étudiants du tronc commun ont un entourage d'âge plus élevé.
- Les étudiants qui ont un entourage d'âge plus jeune ont un entourage de niveau inférieur.

Classification :

La classification nous a permis de regrouper les individus ayant des caractéristiques similaires, ce qui a facilité l'interprétation de nos résultats.

K-means :



K-means est une méthode de partitionnement qui divise l'échantillon en K groupes distincts en minimisant la distance intra-cluster et maximisant la distance inter-cluster.

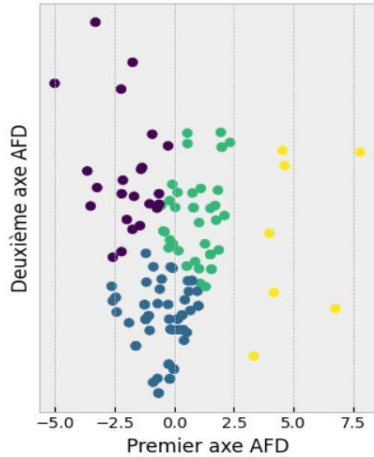
Nous avons effectué une classification des données en utilisant l'algorithme K-means. Il attribue chaque individu à l'un des 4 clusters en fonction de leurs coordonnées dans l'espace MCA. Les résultats sont ensuite affichés pour montrer le cluster auquel chaque individu est attribué.

Note : Les données de la classification sont les coordonnées des individus dans l'espace factoriel .

L'utilisation des coordonnées des individus dans l'espace factoriel MCA dans la préparation pour la classification K-means permet de réduire la dimensionnalité, faciliter la visualisation et identifier des regroupements d'individus similaires. Cela peut aider à améliorer la performance et l'interprétation des résultats de la méthode K-means dans le contexte des données catégorielles.

Analyse discriminante :

Analyse Factorielle Discriminante



L'analyse discriminante est l'étape finale de notre démarche. Elle consiste à construire une fonction discriminante qui maximise la séparation entre les différentes classes de notre variable cible. Cette fonction est basée sur les variables explicatives sélectionnées précédemment.

Les valeurs propres représentent la variance expliquée par chaque composante discriminante dans l'analyse discriminante linéaire (AFD). Dans notre cas, nous avons 4 valeurs propres : [0.46434055, 0.23458622, 0.15155463, 0.1495186]. Cela signifie que la première composante discriminante explique environ 46,43% de la variance, la deuxième composante discriminante explique environ 23,46% de la variance, la troisième composante discriminante explique environ 15,16% de la variance, et la quatrième composante discriminante explique environ 14,95% de la variance.

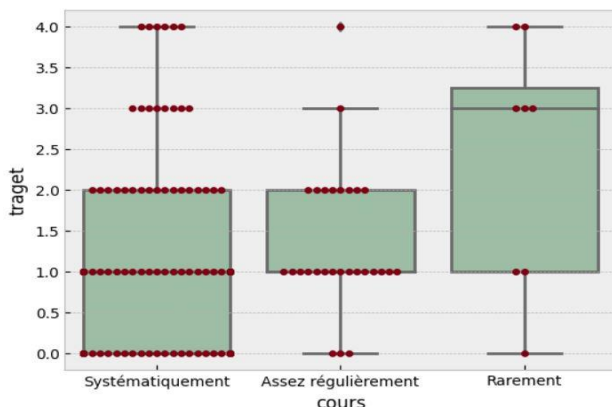
Les coefficients discriminants correspondent aux coefficients attribués à chaque variable dans chaque composante discriminante. Ils déterminent l'importance de chaque variable pour séparer les classes dans l'AFD. Dans notre cas, nous avons une matrice de coefficients discriminants de taille (4, 27). Chaque ligne représente une composante discriminante et chaque colonne représente une variable. Par exemple, la première ligne de coefficients discriminants indique l'importance des variables dans la première composante discriminante, la deuxième ligne indique l'importance des variables dans la deuxième composante discriminante, et ainsi de suite.

Il est important de noter que les coefficients discriminants peuvent être difficiles à interpréter directement en raison de la transformation linéaire effectuée lors de l'AFD. Pour une interprétation plus approfondie des résultats, il peut être utile de visualiser les données à l'aide de graphiques ou de réaliser des analyses supplémentaires.

Anova a un facteur :

L'analyse de variance à un facteur, également appelée ANOVA à un facteur, est une méthode statistique utilisée pour évaluer si la moyenne d'une variable dépend d'une variable catégorielle à un seul niveau, c'est-à-dire d'un seul facteur.

Affichage de boîte à moustache :



La boîte à moustaches (boxplot) est un outil visuel utilisé dans l'analyse de variance (ANOVA) pour représenter graphiquement la distribution des données pour chaque groupe comparé.

Dans le contexte de l'ANOVA, la boîte à moustaches permet de comparer les distributions des différentes catégories du facteur étudié. Chaque boîte représente la répartition des valeurs à l'intérieur d'un groupe, tandis que la ligne médiane de la boîte représente la médiane de chaque groupe.

Analyse de l'effet :

	df	sum_sq	mean_sq	F	PR(>F)
C(cours)	2.0	9.511986	4.755993	3.646279	0.028901
Residual	125.0	163.042702	1.304342	NaN	NaN

l'effet du facteur
"présence(cours)" sur la
variable "trajet" est significatif.
La p-value associée à cet effet
est de 0.028901, ce qui est

inférieur au seuil de significativité couramment utilisé de 0,05. Par conséquent, on peut conclure qu'il y a un effet significatif du facteur "présence(cours)" sur la variable "target". Cela indique que les moyennes de la variable "target" diffèrent de manière statistiquement significative entre les différents niveaux du facteur "présence(cours)".

CONCLUSION

En résumé, ce projet d'analyse du comportement des étudiants de la Faculté des Sciences et Techniques (FST) à l'aide d'un questionnaire a permis d'explorer de manière approfondie divers aspects de leur vie académique, de leurs aspirations et de leurs perceptions personnelles. Les données recueillies ont fourni une vision détaillée de la population étudiante et ont révélé des informations précieuses sur leurs pratiques d'études, leur intégration, leurs croyances et leurs réactions face aux défis.

En conclusion, ce projet a permis d'obtenir une meilleure compréhension du comportement des étudiants de la FST et a fourni des indications précieuses pour améliorer leur expérience universitaire. Ces informations peuvent servir de base solide pour des interventions et des politiques visant à favoriser leur épanouissement académique et personnel. De plus, les résultats de cette étude peuvent également constituer une référence importante pour de futures recherches dans le domaine de l'éducation et de la psychologie des étudiants.

Les résultats obtenus grâce à cette étude approfondie des données fournissent une description complète du comportement des étudiants de la FST. Ces résultats ont des implications importantes pour l'amélioration des pratiques pédagogiques, l'optimisation de l'environnement d'apprentissage et le soutien aux aspirations professionnelles des étudiants.