



## THESIS / THÈSE

### MASTER EN SCIENCES INFORMATIQUES

#### Enseigner la visualisation de données aux étudiants de promotion sociale novices en data science grâce à la modélisation de bases de données

HINANT, Philippe

*Award date:*  
2023

*Awarding institution:*  
Universite de Namur

[Link to publication](#)

#### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

#### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

UNIVERSITÉ DE NAMUR  
Faculté d'informatique  
Année académique 2022-2023

**"Enseigner la visualisation de données aux étudiants de  
promotion sociale novices en data science grâce à la  
modélisation de bases de données"**

Philippe HINANT



Promoteur : \_\_\_\_\_ (Signature pour approbation du dépôt - REE art. 40)  
Bruno DUMAS

Co-promoteur : Antoine CLARINVAL

Mémoire présenté en vue de l'obtention du grade de  
Master en Sciences Informatiques.

## Remerciements

Je tiens à exprimer ma sincère gratitude à mon promoteur, Monsieur Bruno Dumas. Je souhaite également remercier chaleureusement mon co-promoteur, Monsieur Antoine Clarinval, avec qui j'ai eu la chance de partager de nombreux échanges enrichissants, de bénéficier de sa disponibilité ainsi que pour son soutien constant et ses conseils avisés tout au long de la rédaction de mon mémoire. Ses précieuses orientations ont été cruciales pour le développement et l'aboutissement de ce travail. Sa bienveillance et son expertise ont contribué à façonner ce projet.

Je souhaite également exprimer ma reconnaissance à l'école Namur Cadet et à Madame Charlotte Coerten, Directrice adjointe de l'E AFC Namur-Cadets, pour avoir accepté la réalisation de mon mémoire au sein de l'école Namur Cadet. Je tiens à remercier Monsieur Michel Vonèche, l'enseignant avec qui j'ai collaboré tout au long de mon mémoire. Sa contribution et son expertise ont grandement enrichi mes recherches et ont été d'une valeur inestimable pour ce travail. Leurs soutiens et leurs collaborations ont été essentiels pour mener à bien ce projet de recherche.

Je voudrais aussi adresser mes plus profonds remerciements à ma compagne Sophie, qui m'a soutenu inconditionnellement au cours de ces trois années d'études. Son amour, sa patience et son encouragement ont été le moteur de ma persévérance. Sophie, ton soutien indéfectible m'a permis de relever les défis et de surmonter les obstacles que j'ai rencontrés durant ce parcours. Enfin, je tiens à exprimer toute ma gratitude et mon amour à ma fille Chiara, âgée de 7 ans. Sa présence, son sourire et sa joie de vivre ont été une source d'inspiration et de motivation quotidiennes. Chiara, tu as illuminé mes journées même les plus difficiles et m'as donné la force de poursuivre. J'espère que ce travail et l'accomplissement qu'il représente pourront, un jour, t'inspirer à réaliser tes propres rêves.

À vous tous, merci du fond du cœur.

## Résumé

Ce mémoire explore l'intégration de la visualisation de données dans l'enseignement de la modélisation de données pour les étudiants novices en data science, dans le cadre d'un bachelier en informatique de gestion enseigné en promotion social. L'objectif de cette étude est d'explorer les méthodes d'enseignement les plus efficaces pour la visualisation de données et d'évaluer l'impact d'une approche pédagogique spécifique, mise en œuvre au travers d'une unique activité de visualisation de données, sur l'apprentissage des étudiants. Les questions de recherche qui ont guidé cette étude sont : Comment enseigner la visualisation de données à un public novice de promotion sociale par le biais d'un cours sur la modélisation de base de données ? Et, comment intégrer la visualisation de données dans le cursus scolaire des étudiants de promotion sociale novices en data science ?

Pour répondre à ces questions, la méthodologie de recherche comprend une activité de visualisation de données, suivie de questionnaires pré et post observation, un focus group avec les étudiants, une entrevue semi-directive avec l'enseignant et une évaluation basée sur une grille d'observation. L'activité, centrée sur la modélisation de base de données comme porte d'entrée pour l'enseignement de la visualisation de données, s'est avérée efficace pour stimuler l'intérêt et la compréhension des étudiants pour le sujet.

Les résultats indiquent une amélioration significative de la compréhension des concepts de visualisation de données et de l'application des compétences en visualisation de données par les étudiants. De plus, l'activité a réussi à susciter un intérêt accru pour la visualisation de données parmi les étudiants, indiquant le potentiel de cette approche pour l'enseignement de la data science.

Cette recherche contribue à la littérature sur l'enseignement de la visualisation de données en proposant une approche pédagogique qui met l'accent sur la visualisation de données. Les résultats soutiennent l'intégration de la visualisation de données dans le cursus d'enseignement de la modélisation de données pour les étudiants novices, soulignant son rôle clé dans l'amélioration de l'apprentissage et de l'engagement des étudiants.

## Abstract

This dissertation explores the integration of data visualization into the teaching of data modeling for novice data science students, within the context of a bachelor's degree in business computer science taught in social promotion. The aim of this study is to explore the most effective teaching methods for data visualization and to evaluate the impact of a specific pedagogical approach, implemented through a unique data visualization activity, on student learning. The research questions guiding this study are : How to teach data visualization to a novice audience in social promotion through a course on basic data modeling ? And, how to integrate data visualization into the school curriculum of novice data science students in social promotion ?

To answer these questions, the research methodology includes a data visualization activity, followed by pre and post observation questionnaires, a focus group with students, a semi-structured interview with the teacher, and an evaluation based on an observation grid. The activity, focused on basic database modeling as a gateway to teaching data visualization, has proven effective in stimulating students' interest and understanding of the subject.

The results indicate a significant improvement in the understanding of data visualization concepts and in the application of data visualization skills by students. Moreover, the activity succeeded in arousing increased interest in data visualization among students, indicating the potential of this approach for teaching data science.

This research contributes to the literature on the teaching of data visualization by proposing a pedagogical approach that emphasizes data visualization. The findings support the integration of data visualization into the data modeling teaching curriculum for novice students, underlining its key role in enhancing student learning and engagement.

**Mots-clés** : teaching , database , modeling , novices , Infovis



# Table des matières

|     |  |           |
|-----|--|-----------|
| 1   | Introduction . . . . .   | 6         |
| 2   | Etat de l'art . . . . .  | 9         |
| 2.1 | Les fondamentaux de la visualisation de données . . . . .  | 9         |
| 2.2 | Enseignement de la visualisation de données en promotion social . . . . .  | 12        |
| 2.3 | Modélisation de bases de données et visualisation de données . . . . .   | 19        |
| 2.4 | Conclusions l'état de l'art . . . . .  | 23        |
| 2.5 | Question de recherche . . . . .  | 24        |
| 3   | Méthodologie . . . . .   | 25        |
| 3.1 | Contexte de l'étude . . . . .  | 28        |
| 3.2 | Conception de l'activité . . . . .   | 28        |
| 3.3 | Évaluation de l'activité . . . . .   | 31        |
| 3.4 | Débriefing de l'activité . . . . .   | 32        |
| 4   | Résultats . . . . .  | 35        |
| 4.1 | Présentation des résultats . . . . .   | 35        |
| 4.2 | Analyse de l'observation participante . . . . .  | 35        |
| 4.3 | Analyse des questionnaires pré- et post-test . . . . .   | 45        |
| 4.4 | Analyse du focus group . . . . .   | 55        |
| 5   | Discussions . . . . .  | 57        |
| 5.1 | Limitations de l'étude . . . . .   | 57        |
| 5.2 | Résumé des principaux résultats de l'étude . . . . .   | 57        |
| 5.3 | Réponse aux questions de recherche et aux sous-questions de recherche . . . . .  | 58        |
| 5.4 | Implications pédagogiques et perspectives pour l'intégration de la visualisation de données<br>dans les cursus scolaires des étudiants novices en data science . . . . . | 58        |
| 5.5 | Innovation et différenciation par rapport à l'état de l'art . . . . .  | 59        |
| 5.6 | Perspectives et travaux futurs . . . . .   | 60        |
| 6   | Conclusion . . . . .   | 62        |
|     | <b>Table des annexes</b>   | <b>73</b> |
| A   | Annexes A - Questionnaire pre-test et post-test . . . . .  | 75        |
| B   | Annexes B - Focus Group et Interview EAFC Cadet Namur 31/03/2023 - Document préparatoire   | 80        |

# 1 Introduction

Ce mémoire vise à contribuer à la compréhension de l'enseignement de la visualisation de données en promotion sociale, en explorant les approches pédagogiques adaptées aux étudiants novices en data science et en modélisation de bases de données. Les résultats de ce travail pourraient aider les enseignants et les chercheurs à développer des méthodes et des outils efficaces pour enseigner la visualisation de données aux étudiants en promotion sociale, en tenant compte de leurs besoins et de leurs contraintes.

La visualisation de données est devenue un domaine de plus en plus important dans de nombreuses disciplines, notamment en sciences sociales, en informatique, en économie, en biologie, en santé publique et en journalisme. La visualisation de données peut aider à comprendre des phénomènes complexes, à découvrir des modèles cachés, à communiquer des résultats de recherche et à prendre des décisions éclairées. Cependant, la visualisation de données est souvent considérée comme un domaine difficile à maîtriser, qui nécessite des compétences en mathématiques, en statistiques et en informatique.

Dans le contexte de l'enseignement en promotion sociale, la visualisation de données peut être un domaine particulièrement difficile à enseigner, en raison du manque de connaissances et de compétences préalables des étudiants. Il est donc important de développer des approches pédagogiques qui permettent aux étudiants novices d'acquérir des compétences de visualisation de données de manière efficace et efficiente. Dans notre société en constante évolution, l'importance de la compréhension et de l'analyse des données n'a jamais été aussi cruciale. La visualisation de données joue un rôle essentiel dans la communication et la compréhension de ces informations complexes. En outre, l'enseignement de la visualisation de données est devenu un enjeu important, en particulier pour les étudiants en promotion sociale, qui peuvent avoir des compétences et des connaissances variées en informatique et en data science.

Plusieurs travaux de recherche se sont intéressés à l'enseignement de la visualisation de données et ont proposé des méthodes et des outils pour enseigner la visualisation de données aux étudiants novices. Par exemple, Munzner T. [45] ou Katz A. [33] et Kerren A. [73] propose un cadre général pour la visualisation de données qui peut être utilisé pour enseigner la visualisation de données à différents niveaux d'expertise. Kelleher et Tierney [27] proposent un livre de référence sur la visualisation de données pour les débutants en data science, qui présente des concepts clés, des exemples de visualisation et des exercices pratiques.

Suite à une revue de la littérature, il apparaît que la visualisation de données est devenue un outil essentiel pour améliorer la compréhension et la communication de données complexes. Cependant, l'enseignement de la visualisation de données n'est pas encore intégré dans les cursus des étudiants en promotion sociale. C'est pourquoi, ce mémoire se propose d'explorer comment la modélisation de base de données peut être un outil pédagogique pour enseigner la visualisation de données à des étudiants en promotion sociale novices en data science.

Cependant, il y a encore peu d'études qui se sont intéressées spécifiquement à l'enseignement de la visualisation de données en promotion sociale. Le public visé par cette étude est constitué d'élèves de bachelier en informatique de gestion et techniciens codeurs en promotion sociale. Les étudiants en promotion sociale sont des adultes qui retournent à l'école pour se former ou se perfectionner dans leur domaine de travail. Ils peuvent avoir des horaires de travail à temps plein ou à temps partiel et des responsabilités familiales, ce qui peut influencer leur disponibilité et leur engagement dans l'apprentissage. De plus, les étudiants en promotion sociale peuvent avoir des antécédents scolaires limités et des besoins d'apprentissage spécifiques. L'enseignement de la visualisation de données est donc particulièrement pertinent pour les étudiants en promotion sociale de bachelier en informatique de gestion et techniciens codeurs, car cela leur permettrait d'améliorer leur compréhension et leur capacité à communiquer des informations à travers des représentations visuelles. Cette compétence est particulièrement pertinente dans le

contexte professionnel de ces étudiants, où les données sont de plus en plus importantes dans la prise de décision et la résolution de problèmes. La visualisation de données, outil essentiel dans la communication et la compréhension d'informations complexes, pose un enjeu pédagogique particulier pour les étudiants en promotion sociale. En effet, ces derniers, dotés de compétences et connaissances variées, voire lacunaires, en informatique et data science, pourraient grandement bénéficier de méthodes d'enseignement adaptées dans une société de plus en plus axée sur les données, la capacité à comprendre et à communiquer efficacement des informations complexes est devenue cruciale.

Il est important de savoir comment la visualisation de données peut être utilisée pour améliorer la représentation de données et faciliter l'apprentissage des étudiants en promotion sociale novices en data science. La problématique de ce mémoire est la suivante : "Comment utiliser les bases de données comme "proxy" porte d'entrée pour enseigner la data visualisation à des étudiants en promotion sociale qui n'ont jamais été confrontés au contenu de data science (à un public novice en data science) ?".

L'approche proposée dans ce mémoire consiste à utiliser la modélisation de base de données comme une "porte d'entrée" pour enseigner la visualisation de données aux étudiants en promotion sociale. Cette approche est basée sur l'idée que l'analyse de données et la modélisation de données sont des compétences complémentaires et que l'enseignement de l'un peut être utilisé pour renforcer l'autre. Cette approche a été explorée par plusieurs auteurs dans le domaine de l'enseignement de la data science, notamment par Peng et Matsui (2011) [66] et par Ruggles et al. (2018) [28]. Les résultats de cette recherche pourraient aider les enseignants et les chercheurs à développer des méthodes et des outils efficaces pour enseigner la visualisation de données aux étudiants en promotion sociale, en tenant compte de leurs besoins et de leurs contraintes.

Afin de mener à bien cette étude, nous avons adopté une approche mixte, combinant observation participante, questionnaires, focus group avec des étudiants et un entretien avec l'enseignant. Le chercheur a mené une observation participante, ce qui lui a permis de s'immerger dans le contexte d'enseignement, en prenant part à l'activité de visualisation de données et en observant les interactions entre les étudiants et le chercheur. Cette observation a été complétée par des questionnaires, distribués aux étudiants au début et à la fin de l'activité, pour recueillir leurs retours d'expérience et évaluer la pertinence et l'efficacité de l'approche pédagogique utilisée. Par ailleurs, un focus group avec des étudiants a été organisé pour discuter de leurs perceptions et de leurs défis concernant l'apprentissage de la visualisation de données. Enfin, un entretien individuel a été mené avec l'enseignant pour comprendre sa stratégie d'enseignement et sa perception des besoins des étudiants. L'ensemble de ces méthodes a permis une analyse riche et nuancée des défis et des opportunités liés à l'enseignement de la visualisation de données en promotion sociale.

Le mémoire est structuré en plusieurs parties, qui abordent les différents aspects de la problématique. Tout d'abord, l'état de l'art présente une revue de la littérature sur la visualisation de données et la modélisation de données ainsi que de l'enseignement de l'informatique en promotion sociale. Ensuite, la méthodologie de collecte de données et d'analyse est présentée pour répondre à la question de recherche. La première partie de ce mémoire se compose d'une revue de la littérature qui positionne le sujet dans le contexte de la Data Science et de l'enseignement. Cette partie présente également les implications, les limitations et les perspectives de recherche couvertes par la littérature. La deuxième partie de ce mémoire présente la méthodologie utilisée pour répondre à la question de recherche. Cette partie décrit les activités de cours proposées, les méthodes de collecte de données, ainsi que l'analyse des résultats. La troisième partie de ce mémoire présente les résultats obtenus. Les résultats montrent que les étudiants novices en Data Science sont capables d'améliorer leur représentation de données grâce à l'utilisation de bases de données comme porte d'entrée pour enseigner la visualisation de données. Enfin, la dernière partie de ce mémoire présente les implications de cette étude pour l'enseignement de la visualisation de données et la formation des novices en Data Science. Les méthodes pédagogiques adaptées aux besoins des étudiants en promotion sociale sont essentielles pour favoriser leur engagement dans l'apprentissage de la



visualisation de données. Les résultats de cette étude peuvent être utiles pour les enseignants et les responsables de programme qui souhaitent intégrer la visualisation de données dans les cursus scolaires des étudiants novices en Data Science.

En somme, ce mémoire a pour objectif de contribuer à l'amélioration de l'enseignement de la visualisation de données pour les étudiants en promotion sociale novices en Data Science. En utilisant des bases de données comme porte d'entrée pour l'enseignement de la visualisation de données, cette étude montre que les étudiants peuvent améliorer leur représentation de données de manière significative. Les résultats obtenus peuvent être utiles pour les enseignants et les responsables de programme qui souhaitent intégrer la visualisation de données dans les cursus scolaires des étudiants novices en Data Science.

## 2 Etat de l'art

La visualisation de données est un domaine en constante évolution qui vise à faciliter la compréhension, l'analyse et la communication des données en les présentant sous une forme visuelle accessible et attrayante. Dans le contexte de l'enseignement de la promotion sociale, l'intégration de la visualisation de données dans la modélisation de bases de données peut potentiellement aider les étudiants novices en sciences des données à mieux comprendre les concepts et les techniques associés. L'objectif de cet état de l'art est d'examiner les travaux existants dans les domaines de la visualisation de données, de l'enseignement en promotion sociale et de la modélisation de bases de données, afin de mieux comprendre comment ces éléments peuvent être combinés pour améliorer l'apprentissage des étudiants.

Pour mener à bien cet état de l'art, une recherche bibliographique a été réalisée en consultant des bases de données scientifiques, des revues spécialisées et des ouvrages de référence dans les domaines concernés. Les mots-clés utilisés pour la recherche incluent des termes tels que "teaching" , "database" , "modeling" , "novices" , "Infovis". Des critères de sélection ont été appliqués pour identifier les sources les plus pertinentes et les plus récentes, en accordant une attention particulière aux études empiriques, aux revues systématiques et aux travaux théoriques qui abordent directement les questions soulevées du mémoire.

### 2.1 Les fondamentaux de la visualisation de données

#### Définition et objectifs de la visualisation de données

La visualisation de données est une discipline qui consiste à représenter des données sous une forme graphique ou visuelle pour faciliter leur compréhension, leur analyse et leur communication. Elle vise à transformer les données brutes, souvent complexes et volumineuses, en informations visuellement accessibles et compréhensibles. Les objectifs de la visualisation de données incluent la découverte de tendances, la reconnaissance de modèles, l'identification d'anomalies et la communication efficace des résultats. Cette discipline est en constante évolution, avec l'apparition de nouveaux outils et techniques pour représenter des données de manière efficace et attrayante (Herr, 2010[68]). La visualisation de données peut être utilisée dans de nombreux domaines, notamment la science des données, la recherche scientifique, la gestion d'entreprise et la communication de l'information au grand public. Elle peut également être utilisée dans le contexte de l'enseignement pour aider les étudiants à mieux comprendre les concepts et les techniques associés aux données (Kelleher, 2018[27]).

La visualisation de données peut être utilisée pour différents types de données, qu'il s'agisse de données quantitatives, telles que les chiffres, les statistiques ou les mesures, ou de données qualitatives, telles que les textes ou les images. Les techniques de visualisation peuvent varier en fonction du type de données et des objectifs de la visualisation (Munzner, 2014[45]). Les techniques de visualisation courantes comprennent les graphiques, les tableaux, les cartes et les diagrammes (Few, 2006[80]). En outre, le pipeline de visualisation de données est un processus structuré qui comprend plusieurs étapes pour transformer les données brutes en représentations visuelles exploitables. Ce processus se compose généralement des étapes suivantes (voir figure 1) : l'acquisition des données, le prétraitement, la cartographie des données, la représentation visuelle et l'interaction (Munzner, 2014[45]; Card, Mackinlay, et Shneiderman, 1999[90]). L'interactivité joue également un rôle clé dans les visualisations, car elle permet aux utilisateurs d'interagir directement avec les représentations visuelles, offrant ainsi la possibilité d'explorer et d'analyser les données de manière flexible et personnalisée (Yi et al., 2007[79]; Heer et Shneiderman, 2012[59]). Cependant, la complexité inhérente à ce concept, couplée à la nécessité d'une certaine familiarité avec les outils informatiques et les techniques de data science, peut rendre l'interactivité difficile à appréhender pour les novices. Il est donc crucial de mettre en place un cadre pédagogique approprié pour faciliter l'acquisition de ces compétences et maximiser les bénéfices de l'interactivité dans la visualisation de données suivant Lee et al.[39].

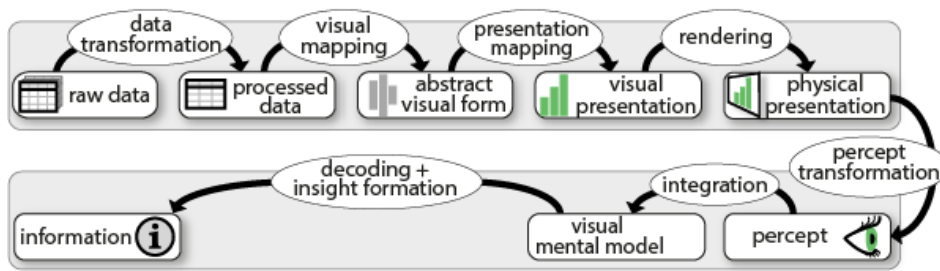


FIGURE 1 – Pipeline de visualisation de données (Jansen et al. 2013[53])

Les principes de conception de la visualisation de données sont guidés par des règles esthétiques, fonctionnelles et cognitives pour optimiser la lisibilité et l'efficacité des représentations visuelles. Selon Few [80], ces principes incluent :

**La simplicité** : privilégier des représentations visuelles épurées et sans éléments superflus.

**La clarté** : assurer une bonne lisibilité des données en utilisant des légendes, des étiquettes et des couleurs appropriées.

**L'adéquation** : choisir le type de graphique le plus adapté pour représenter les données, en fonction de leur nature et des objectifs de l'analyse.

Les techniques de visualisation comprennent une variété de représentations, telles que les graphiques linéaire, les histogramme (voir figure 2), les graphiques à secteurs, les nuages de points et les cartes géographiques (voir figure 3), entre autres. Chaque technique présente des avantages et des inconvénients, en fonction de la nature des données et des objectifs de l'analyse.

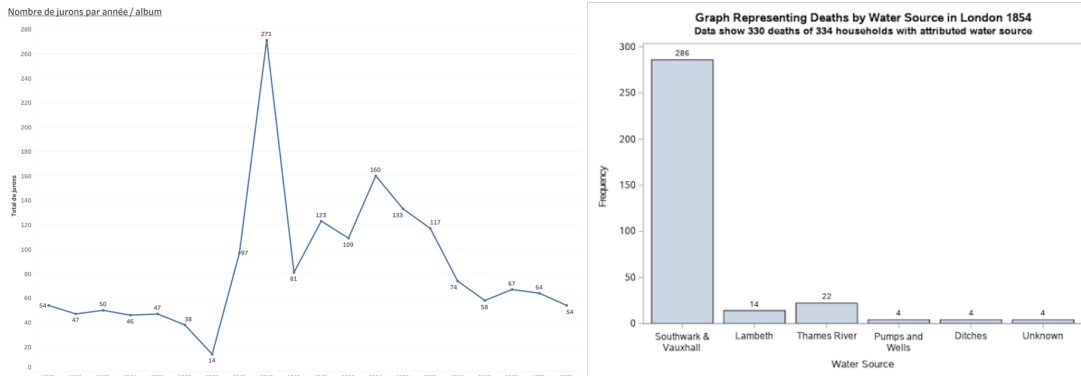


FIGURE 2 – Exemple d'un graphique linéaire et d'un graphique à barres <sup>1</sup>

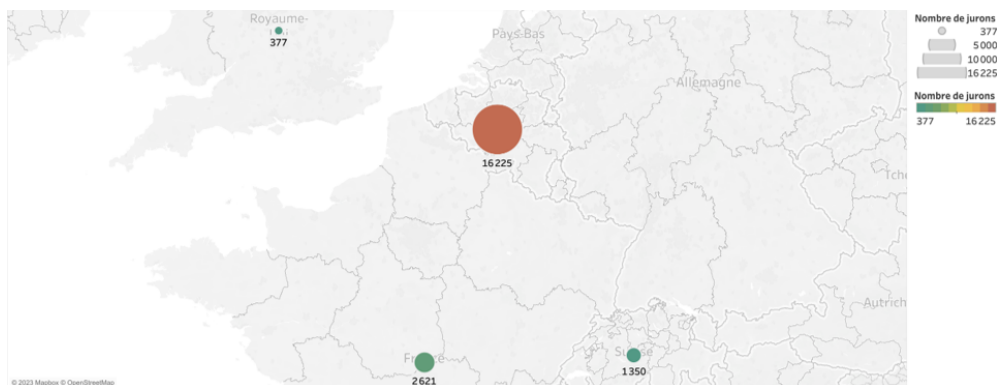


FIGURE 3 – Exemple d'un graphique géographique

Il existe de nombreux outils et logiciels de visualisation de données, allant des solutions simples et gratuites aux solutions professionnelles et payantes. Parmi les outils les plus populaires, on trouve :

**Microsoft Excel**<sup>2</sup> : un tableur largement utilisé qui propose des fonctionnalités de base pour créer des graphiques et des tableaux. Excel est essentiellement utilisé pour le traitement et l'analyse de données grâce à ses fonctions intégrées et à sa capacité à réaliser des calculs complexes.

**Tableau**<sup>3</sup> : un logiciel professionnel de visualisation de données qui offre une grande variété de graphiques et de techniques d'analyse avancées. Tableau se distingue par sa capacité à traiter des ensembles de données de grande taille et par sa flexibilité pour créer des visualisations personnalisées.

**D3.js**<sup>4</sup> : une bibliothèque JavaScript pour créer des visualisations de données interactives et personnalisées sur le web. D3.js permet d'exploiter la flexibilité du HTML, du SVG et du CSS pour créer des visualisations de données avancées, et offre une grande variété de fonctions pour manipuler les données et générer des visualisations dynamiques.

**R**<sup>5</sup> et **ggplot2**<sup>6</sup> : un langage de programmation et une bibliothèque pour créer des visualisations de données statistiques de haute qualité. R est particulièrement puissant pour l'analyse statistique et les modèles prédictifs, tandis que ggplot2 offre une interface cohérente pour créer des visualisations complexes de manière intuitive.

**PowerBI**<sup>7</sup> : une suite d'outils d'analyse commerciale de Microsoft qui permet de visualiser des données et de partager des informations à travers une organisation. Ses fonctionnalités interactives de glisser-déposer rendent la création de tableaux de bord et de rapports facile et intuitive.

**Python et Matplotlib**<sup>8</sup> : Python est un langage de programmation très populaire pour la data science, en grande partie grâce à ses nombreuses bibliothèques dédiées, comme Matplotlib. Matplotlib est une bibliothèque de visualisation de données pour Python qui permet de créer une grande variété de graphiques et de visualisations.

L'utilisation d'outils et de logiciels pour la visualisation de données varie en fonction des besoins et des compétences des utilisateurs, ainsi que des contraintes budgétaires et techniques[17] Les enseignants devraient tenir compte de ces facteurs lors du choix des outils de visualisation pour répondre aux besoins spécifiques de leurs étudiants (Germanakos, P. 2019[17] Hermans, F. 2016 citeHermans2016). Plusieurs études ont examiné l'intégration de ces outils et approches dans l'enseignement de la visualisation de données, soulignant l'importance de la sélection appropriée des outils en fonction des compétences et des besoins des étudiants (voir figure 4 (Lee, S. 2015[39],Kruk, S. 2022[1] Fekete, JD. 2007[77] Seckler, M. 2021[7] Alper,B (2017)[29]). L'intégration de ces contraintes supplémentaires, doit orienter les enseignants vers une approche pédagogique différente pour les étudiants de la promotion social.

---

1. <https://pressbooks.library.upei.ca/montelpare/chapter/john-snow-and-the-natural-experiment/>

2. <https://www.microsoft.com/fr-be/microsoft-365/excel>

3. <https://www.tableau.com/fr-fr>

4. <https://d3js.org/>

6. <https://www.r-project.org/>

6. <https://ggplot2.tidyverse.org/>

7. <https://powerbi.microsoft.com/>

8. <https://matplotlib.org/>



FIGURE 4 – Exemple d’une tablette se basant sur une application Web à destination des élèves d’une école primaire Alper,B (2017)[29]

C’est pour cela, que la visualisation de données joue un rôle crucial dans la compréhension et l’interprétation de l’information dans notre quotidien et dans l’enseignement. Les enseignants sont de plus en plus confrontés à la nécessité de former les étudiants à maîtriser les compétences en visualisation, car la quantité de données disponibles augmente rapidement et la nécessité de les comprendre est omniprésente dans divers domaines professionnels(Borkin, M. 2019[13]). En intégrant la visualisation de données dans l’enseignement, les enseignants peuvent aider les étudiants à développer des compétences en analyse de données, en pensée critique et en communication, ce qui est essentiel pour naviguer avec succès dans le monde complexe et axé sur les données d’aujourd’hui selon Brown, A.(1987)[94] et Fekete, JD. (2007)[77].

## 2.2 Enseignement de la visualisation de données en promotion social

### Approches pédagogiques pour enseigner la visualisation de données

L’enseignement de la visualisation de données implique l’utilisation de diverses approches pédagogiques pour aider les étudiants à acquérir des compétences en matière de création, d’analyse et d’interprétation de représentations visuelles de données. Parmi les approches couramment utilisées (Kerren, 2008 [74] et Heer et al., 2012 [58]) , on trouve :

**La théorie :** expliquer les concepts fondamentaux, les principes de conception et les techniques de visualisation aux étudiants(Munzner ,2014 [45]).

**Les démonstrations :** montrer aux étudiants comment créer et analyser des visualisations de données en utilisant des outils et logiciels spécifiques (Yi, 2007 [79]).

**Les exercices pratiques :** permettre aux étudiants de créer leurs propres visualisations de données à partir de jeux de données réels ou simulés, afin de renforcer leur compréhension des concepts et de développer leurs compétences techniques (Heer, J. 2012 [59]).

**Les projets de groupe :** encourager les étudiants à travailler ensemble pour résoudre des problèmes de visualisation de données complexes et à communiquer leurs résultats (Ruggles et al., 2018 [28]).

Des recherches supplémentaires ont été menées sur l’efficacité de diverses méthodes d’enseignement. Dans leur étude, Margolis et Fisher (2002)[85] ont noté que l’acquisition de compétences en informatique est souvent un processus itératif, où la compréhension et la compétence augmentent à chaque cycle. En particulier, ils ont constaté que les étudiants qui ont été exposés à des concepts de manière répétée et ont eu l’occasion de pratiquer ces compétences ont montré une amélioration de leur compréhension et de leur performance.

Dans le même esprit, Noss et Hoyles (1996)[92] ont étudié les cultures d’apprentissage et les ordinateurs et ont découvert qu’à mesure que les étudiants devenaient plus familiers avec les concepts et les outils, ils

étaient capables de travailler de manière plus efficace et précise. Cette amélioration de l'efficacité était liée à une meilleure compréhension des concepts sous-jacents ainsi qu'à une familiarité accrue avec les outils eux-mêmes.

## Rôle de la visualisation de données dans l'enseignement des sciences des données et de l'informatique

La visualisation de données joue un rôle important dans l'enseignement des sciences des données et de l'informatique, car elle permet de :

1. **Faciliter** la compréhension des concepts abstraits et complexes liés aux données et aux algorithmes (Munzner, 2014 [45]).
2. **Améliorer** les compétences en analyse et en résolution de problèmes des étudiants (Yi, 2007 [79]).
3. **Développer** la créativité et l'esprit critique des étudiants en les encourageant à explorer différentes représentations visuelles et à évaluer leur pertinence (Isenberg, 2011 [65]).
4. **Renforcer** la capacité des étudiants à communiquer efficacement les résultats de leurs analyses à divers publics [13].

### Approches pédagogiques spécifiques pour la visualisation de données

En plus des approches pédagogiques générales mentionnées précédemment, certaines approches spécifiques ont été développées pour enseigner la visualisation de données. Par exemple, l'enseignement basé sur des projets réels peut aider les étudiants à acquérir une expérience pratique et à résoudre des problèmes concrets liés à la visualisation de données (Lee, 2015 [39]). De plus, l'utilisation de scénarios d'apprentissage collaboratifs peut encourager les étudiants à partager leurs connaissances et à apprendre les uns des autres tout en travaillant sur des projets de visualisation de données (Kruk, S. 2022 [1]). Enfin, l'intégration de l'interactivité dans les visualisations peut aider les étudiants à explorer et à comprendre les données plus efficacement, en leur permettant de manipuler les représentations visuelles et d'observer les effets de leurs actions (Heer, J. 2012 [59]).

### Études de cas et expériences d'enseignement de la visualisation de données

Plusieurs études de cas et expériences d'enseignement de la visualisation de données ont été rapportées dans la littérature, mettant en évidence les avantages et les défis de cette approche pédagogique. Parmi les exemples notables, on trouve :

1. L'intégration de la visualisation de données dans des cours de statistiques, d'analyse de données et de programmation, où les étudiants apprennent à créer des graphiques et des tableaux pour analyser et interpréter les données (Kelleher, 2005, Willett 2019 [22], [27]).
2. Des ateliers et des séminaires sur la visualisation de données, où les participants sont exposés à diverses techniques de visualisation et apprennent à les appliquer à des problèmes concrets (Heer, J. 2012 [58]).
3. Des projets interdisciplinaires qui impliquent la collaboration entre des étudiants en informatique, en design et en sciences sociales pour développer des visualisations de données innovantes et informatives.

Les résultats de ces études suggèrent que l'enseignement de la visualisation de données peut avoir des effets positifs sur la compréhension et l'application des concepts liés aux sciences des données et de l'informatique. Cependant, ils soulignent également les défis liés à l'enseignement de la visualisation de données, tels que la nécessité de former les enseignants et les étudiants aux principes de conception et aux techniques de visualisation appropriées (Ruggles, 2018 [28]). Ci-dessous le tableau de synthèse des études de cas recensées (voir table 1).

TABLE 1 – Tableau de synthèse des études de cas par article

| id | Références  | Synthèse   |
|----|---|--|
| 1  | Shmallo, Ronit et Katz, Adi (2015) [41]                   | Cette étude de cas met l'accent sur l'enseignement de la modélisation des bases de données relationnelles (BDR) en utilisant l'approche d'apprentissage à partir des erreurs. L'analyse des données recueillies à travers des enquêtes auprès des étudiants, des observations et l'analyse de données qualitatives révèle que cette approche permet d'améliorer la qualité des modèles conceptuels produits par les étudiants universitaires. En effet, l'approche d'apprentissage à partir des erreurs surpasse l'enseignement traditionnel en termes de qualité des résultats obtenus. |
| 2  | Maik Beege (2021) [2]                                     | Cette recherche se concentre sur l'apprentissage de la programmation à partir d'exemples erronés. Elle utilise une méthodologie expérimentale, impliquant à la fois un groupe de contrôle et un groupe expérimental. Les résultats montrent que les exemples avec des erreurs, qu'elles soient syntaxiques ou conceptuelles, sont bénéfiques pour l'apprentissage des étudiants en informatique et des enseignants d'informatique. En fait, ces exemples d'erreurs se sont révélés supérieurs aux exemples corrects pour améliorer l'apprentissage.                                      |
| 3  | Adi Katz et Ronit Shmallo (2016) [33]                     | Cette étude explore comment l'apprentissage par les erreurs et la visualisation de données peuvent améliorer la modélisation de données relationnelles. Les chercheurs ont analysé les données des étudiants et identifié les erreurs fréquentes pour développer des outils de visualisation. Les résultats montrent que ces outils ont aidé à réduire le nombre d'erreurs et à améliorer la compréhension des étudiants en informatique, démontrant ainsi l'efficacité de cette approche.   |
| 4  | Bachelet, R., Caron-Pargue, A., et Morlat, P. (2016) [31] | Cette étude se concentre sur les compétences informationnelles des étudiants en formation professionnelle en Communauté française de Belgique. Grâce à une enquête quantitative et qualitative, les auteurs ont mis en évidence que les étudiants manquent souvent de compétences de base telles que la recherche d'informations, la gestion de l'information et l'évaluation critique des sources d'information. Les technologies de l'information et de la communication (TIC) ont été identifiées comme des outils essentiels pour l'amélioration de ces compétences.                 |
| 5  | Kerren, A. [54]   | Cette étude se concentre sur l'enseignement de la visualisation de l'information à des étudiants en informatique. Grâce à une étude exploratoire avec des questionnaires pré et post cours, et l'observation des interactions entre les étudiants et les outils de visualisation, il a été démontré que les étudiants ont amélioré leur compréhension des principes de conception de visualisation et leur capacité à travailler avec des outils de visualisation.   |
| 6  | Lo, Yu-Ho, Yao Ming, and Huamin Qu [21]                   | Cette étude se concentre sur l'enseignement de la visualisation de données à des étudiants en informatique et en science des données. Les auteurs proposent une approche d'apprentissage interactive basée sur l'utilisation d'outils de visualisation de données et de tutoriels. Bien que la méthodologie précise ne soit pas détaillée, les résultats indiquent une amélioration de la compréhension des concepts de visualisation de données et une augmentation de l'efficacité dans la création de visualisations par les étudiants.   |

|    |   |   |
|----|---|---|
| 7  | Burch, Michael and Melby, Elisabeth (2019) [15]                 | Burch et Melby examinent les méthodes d'enseignement collaboratif et d'évaluation de groupe dans le cadre de cours de visualisation de données pour les étudiants universitaires. Utilisant une approche d'apprentissage expérientiel et pratique, ils ont conçu des projets de groupe, effectué une évaluation basée sur les pairs et encouragé la rétroaction et l'auto-réflexion. Leur étude de cas a montré que ces méthodes ont conduit à une amélioration significative de la qualité des projets de groupe.  |
| 8  | Burch et Melby (2020) [9]                                       | Burch et Melby analysent plus de 100 projets de visualisation de données pour comprendre comment les enseignants peuvent améliorer leur enseignement de la visualisation. Ils utilisent une approche d'apprentissage actif, en se concentrant sur l'expérimentation avec de nouveaux moyens d'enseigner la visualisation, y compris l'utilisation de projets. Les auteurs concluent que l'enseignement de la visualisation doit se concentrer sur la compréhension des données plutôt que sur la technologie, et que les projets collaboratifs sont un moyen efficace d'enseigner la visualisation. |
| 9  | Lanzara, G. and Palmisano, I. (2016) [35]                       | Lanzara et Palmisano proposent une méthode pour faciliter l'apprentissage de la modélisation de bases de données en utilisant la visualisation de la syntaxe et de la transition graduelle. Cela est particulièrement destiné aux étudiants en informatique. L'idée est que la visualisation de la syntaxe peut améliorer l'apprentissage de la transition graduelle dans le contexte de la modélisation de bases de données. Leur approche a été efficace pour améliorer l'apprentissage des étudiants.  |
| 10 | Adi Kat (2018) [26]   | Adi Kat a utilisé une approche appelée TSSL (Teaching with Subgoal Labeling) pour enseigner la modélisation de schémas de bases de données, en se concentrant sur la visualisation de la syntaxe de transitions graduelles. Cette approche, expérimentée avec des étudiants en informatique, a démontré une amélioration significative des compétences en modélisation de schémas de bases de données. De plus, la visualisation de la syntaxe de transitions graduelles s'est avérée bénéfique pour faciliter la compréhension des étudiants.  |
| 11 | Lars Grammel [67]   | Grammel se concentre sur la conception d'interfaces utilisateur pour aider les novices en visualisation d'information à construire leurs propres visualisations. L'étude vise à montrer comment des interfaces utilisateur adaptées peuvent faciliter la construction de visualisations d'information par les personnes qui débutent dans ce domaine, notamment les étudiants en informatique.  |
| 12 | Kucher, Kostiantyn et Martins, Rafael M. et Kerren, Andreas [5] | Cet article expose la conception d'un cours intitulé "Projet en Visualisation et Analyse de Données" destiné aux étudiants universitaires. L'accent est mis sur l'apprentissage par projet et l'analyse de données. Les résultats démontrent le succès de l'approche d'apprentissage par projet, avec une satisfaction élevée des étudiants par rapport à la qualité du cours.  |
| 13 | Byrd, V. L. (2019)[6]   | Cet article explore l'utilisation de pédagogies innovantes, comme l'apprentissage actif, l'apprentissage par projet, et le travail collaboratif, pour l'enseignement de la visualisation de données à des étudiants en ingénierie de premier cycle et des cycles supérieurs. Les résultats de cette étude montrent une amélioration de l'engagement, de la motivation et des performances des étudiants grâce à ces méthodes d'enseignement.  |



|    |   |   |
|----|---|---|
| 14 | Borkin[11]                                | Cet article propose une méthodologie simplifiée pour l'étude de conception, mettant l'accent sur la pédagogie de la visualisation et l'impact social positif. La méthodologie proposée a permis de simplifier et d'accélérer les études de conception, tout en favorisant une synergie entre la pédagogie de la visualisation et un impact social positif.  |
| 15 | Van Riper, Charles et Kirby, Robert M[72] | L'article explore des stratégies pour promouvoir l'utilisation efficace des outils de visualisation dans le domaine de l'ingénierie, avec un accent particulier sur l'apprentissage expérientiel. Les étudiants ont signalé une amélioration significative dans leur utilisation des outils de visualisation après avoir participé à un programme basé sur ces stratégies, témoignant ainsi du succès de l'approche d'apprentissage expérientiel.   |
| 16 | Tavernier, E., et Galand, B. (2015)[42]   | Cet article présente une étude sur l'analyse exploratoire de données dans les cours de sciences de l'ingénieur en promotion sociale. La recherche montre que l'analyse exploratoire de données améliore l'apprentissage en favorisant la compréhension des concepts scientifiques et l'engagement des étudiants. La méthodologie comprenait l'utilisation de techniques d'analyse de données sur des ensembles de données spécifiques. Le public cible de cette recherche était les étudiants des cours de sciences de l'ingénieur en promotion sociale.  |
| 17 | McKenna, S. (2017) [30]                   | L'article explore des stratégies pour promouvoir l'utilisation efficace des outils de visualisation dans le domaine de l'ingénierie, en mettant l'accent sur l'apprentissage expérientiel. Les résultats de l'étude montrent que les étudiants ont signalé une amélioration significative dans leur utilisation des outils de visualisation après avoir participé à un programme basé sur ces stratégies. La méthodologie de l'apprentissage expérientiel a été efficace pour aider les étudiants à développer leurs compétences en visualisation. Le public cible de la recherche est principalement constitué d'étudiants en ingénierie qui cherchent à améliorer leur utilisation des outils de visualisation. |
| 18 | Kurniawati, S. (2019) [19]                | L'article présente l'utilisation d'un outil pratique pour faciliter l'apprentissage de la normalisation des tables. L'outil comprend des exemples pratiques et permet aux étudiants en informatique d'acquérir les compétences nécessaires dans ce domaine. L'outil a été testé avec succès sur un groupe d'étudiants, démontrant ainsi son efficacité pour faciliter l'apprentissage de la normalisation des tables. Le public cible de la recherche est principalement constitué d'étudiants en informatique qui cherchent à renforcer leurs compétences en normalisation des tables.   |
| 19 | Ginestié, J. and Tricot, A. (2006) [52]   | L'article présente une étude de cas sur l'activité des élèves et des enseignants en éducation scientifique et technologique. L'étude utilise l'enseignement par projet et l'apprentissage par exploration pour promouvoir la résolution de problèmes. Les résultats montrent que les élèves sont plus impliqués et motivés lorsqu'ils sont actifs dans leur apprentissage. Les enseignants ont un rôle important à jouer en guidant et en accompagnant les élèves tout au long de leur démarche d'apprentissage. Le public cible de la recherche est principalement constitué d'élèves du secondaire et d'enseignants dans le domaine de l'éducation scientifique et technologique.                               |

|    |   |   |
|----|---|---|
| 20 | Hepworth, K. J. and Canon, C. (2018) [24]           | L'article explore l'utilisation d'une approche interdisciplinaire STEAM (Science, Technology, Engineering, Art, and Mathematics) pour améliorer l'enseignement de la visualisation de données dans le domaine de l'éducation scientifique. L'étude met en œuvre des ateliers pratiques et des outils informatiques pour permettre aux étudiants d'explorer et de créer des visualisations. Les résultats montrent que cette approche interdisciplinaire a permis aux étudiants d'améliorer la qualité de leurs visualisations. Le public cible de la recherche est principalement constitué d'étudiants en science. |
| 21 | Martins, R., Jusufi, I., and Kerren, A. (2020) [10] | Ce article présente une revue de la littérature sur l'utilisation de la visualisation pour l'interprétation des modèles d'apprentissage automatique. Les résultats des sondages montrent que les méthodes de visualisation les plus utilisées sont les graphiques, les tableaux de bord et les cartes de chaleur. Les utilisateurs expriment le besoin de méthodes de visualisation plus interactives et conviviales pour l'utilisateur final. Le public cible de cette recherche comprend des chercheurs en informatique, ainsi que des enseignants et des étudiants dans le domaine de l'informatique.            |
| 22 | L'Yi (2019) [20]                                    | Cet article présente une étude expérimentale sur différentes méthodes de recommandation de visualisations pour la construction de scatterplots (nuages de points). L'expérimentation implique des participants volontaires qui évaluent l'efficacité des méthodes de recommandation. Les résultats soulignent l'importance de la représentation des données dans l'efficacité des recommandations de visualisations. Le public cible de cette recherche n'est pas spécifié dans l'article.  |
| 23 | Pantazos, K. (2014) [46]                            | Cet article présente une évaluation des outils d'information visuelle (InfoVis) du point de vue des utilisateurs. L'évaluation a été réalisée à l'aide d'un questionnaire et d'une étude de cas. Bien que le public cible ne soit pas spécifié, les résultats ont montré que les utilisateurs préféreraient des outils InfoVis qui leur permettaient de personnaliser la présentation des données et d'interagir de manière flexible avec elles.  |
| 24 | Borkin, M. A., Yan, Z., et Horn, B. (2019) [37]     | Approche d'apprentissage par service pour enseigner la visualisation des données. Les étudiants travaillent avec des organisations à but non lucratif pour créer des visualisations de données pertinentes sur des questions sociales importantes. Développement des compétences de visualisation et de travail avec des clients.   |
| 25 | Kruk, Sergiusz et Holubowicz (2022) [1]             | Avancées dans l'enseignement des systèmes de bases de données, mettant en évidence les différentes approches et pratiques pédagogiques utilisées, ainsi que les outils et technologies disponibles pour faciliter cet enseignement.   |

### Contexte de la promotion sociale et ses objectifs

La promotion sociale est un système d'enseignement qui vise à offrir des opportunités d'apprentissage et de formation aux adultes, souvent en dehors du cadre traditionnel de l'enseignement supérieur. Les objectifs de la promotion sociale incluent l'amélioration des compétences professionnelles et personnelles des étudiants, la facilitation de l'accès à l'emploi, et le soutien à la mobilité sociale et professionnelle (Cedefop, 2014 [43]).

### Approches pédagogiques spécifiques pour l'enseignement en promotion sociale

La formation en informatique en promotion sociale se caractérise par certaines particularités et soulève des enjeux spécifiques. Ces aspects distinctifs sont les suivants :

**Un public diversifié :** les étudiants en promotion sociale peuvent avoir des niveaux de compétence, des

expériences et des attentes très variés, ce qui nécessite des approches pédagogiques adaptées et flexibles. Merrill a souligné l'importance d'adapter les méthodes d'enseignement à la diversité du public en utilisant des environnements d'apprentissage interactifs et des outils de programmation visuelle [86]. Kasworm a mis en œuvre des projets de groupe et des discussions en classe pour encourager la collaboration et le partage d'idées entre les étudiants [84]. Akcayir et Akcayir ont également étudié les avantages et les défis de la mise en œuvre de l'apprentissage mixte et des laboratoires à distance en ligne dans l'éducation professionnelle et l'enseignement de la programmation informatique [12], [23]. Leurs études ont souligné l'importance de l'engagement des étudiants, de l'adaptabilité des enseignants et des ressources technologiques dans ces approches.

**Des contraintes de temps et de ressources :** les cours en promotion sociale sont souvent dispensés en soirée ou le week-end, et les étudiants peuvent avoir des contraintes professionnelles et familiales qui limitent leur disponibilité pour l'apprentissage. Kasworm a souligné l'importance de prendre en compte ces contraintes et de proposer des modalités d'enseignement flexibles, telles que des projets de groupe et des discussions en classe, qui favorisent la collaboration et permettent aux étudiants de concilier leurs engagements personnels et professionnels avec leurs études [84]. Akcayir et Akcayir ont également abordé les défis liés à la mise en œuvre de l'apprentissage mixte et des laboratoires à distance en ligne, notamment en termes de ressources technologiques et de gestion du temps [12], [23].

**L'importance de l'employabilité et des compétences pratiques :** les formations en informatique en promotion sociale doivent être axées sur le développement de compétences pratiques et transversales, en accord avec les besoins du marché du travail et les attentes des étudiants. Selon le Cedefop, l'objectif de la promotion sociale est de permettre aux étudiants d'acquérir des compétences qui favorisent leur employabilité et leur permettent de s'adapter aux évolutions du marché du travail [43]. Les études de Akcayir et Akcayir ont également souligné l'importance de développer des compétences pratiques en programmation informatique [12].

En prenant en compte ces enjeux spécifiques, les enseignants en promotion sociale peuvent adapter leurs approches pédagogiques pour répondre aux besoins des étudiants et favoriser leur apprentissage en informatique. L'utilisation d'environnements d'apprentissage interactifs, de projets de groupe et d'une orientation pratique permet d'engager les étudiants, de les préparer aux exigences du marché du travail et de les soutenir dans leur progression professionnelle et sociale.

## **Enseignement de l'informatique et des sciences des données en promotion sociale**

La littérature regorge de nombreuses expériences et études portant sur l'enseignement de l'informatique en promotion sociale. Ces recherches mettent en lumière à la fois les défis auxquels sont confrontés les enseignants et les apprenants, ainsi que les réussites et les avantages de cette approche pédagogique.

**Étude sur l'enseignement de la programmation et de l'analyse de données en promotion sociale :** Merrill (2002) a adopté une approche pédagogique basée sur l'utilisation d'environnements d'apprentissage interactifs et d'outils de programmation visuelle [86]. Ces outils permettent aux étudiants de manipuler visuellement les concepts de programmation, ce qui facilite leur compréhension et leur apprentissage. En utilisant cette approche, Merrill a pu accroître l'engagement des étudiants et améliorer leur maîtrise des concepts de programmation. Kasworm (2003) a mis en œuvre des projets de groupe et des discussions en classe pour encourager la collaboration et le partage d'idées entre les étudiants [84]. Cette approche favorise l'interaction sociale et permet aux étudiants de construire leur compréhension de la programmation et de l'analyse de données en travaillant ensemble. En encourageant la collaboration, Kasworm a constaté une augmentation de l'engagement des étudiants et une amélioration de leur apprentissage. En combinant les approches de Merrill et de Kasworm, les enseignants en promotion sociale peuvent créer un environnement d'apprentissage stimulant où les étudiants sont activement engagés et motivés. L'utilisation d'environnements d'apprentissage interactifs et de

projets de groupe favorise l'apprentissage collaboratif et permet aux étudiants de développer des compétences en programmation et en analyse de données de manière significative et pratique.

**Étude sur l'enseignement de l'informatique en éducation professionnelle** : Akcayir et Akcayir ont étudié les avantages et les défis de la mise en œuvre de l'apprentissage mixte dans l'éducation professionnelle [23]. Ils ont souligné l'importance de l'engagement des étudiants et de l'adaptabilité des enseignants dans ce contexte.

**Étude sur l'enseignement de la programmation informatique par le biais de laboratoires à distance en ligne** : Akcayir et Akcayir ont également réalisé une étude sur l'enseignement de la programmation informatique par le biais de laboratoires à distance en ligne [12]. Ils ont examiné les avantages et les défis de cette approche et ont souligné son efficacité pour faciliter l'apprentissage des étudiants.

**Étude sur l'utilisation de l'apprentissage en ligne en mathématiques** : Dans leur étude, Hainey, Connolly et Stansfield ont examiné les perceptions des étudiants sur l'utilisation de l'apprentissage en ligne en mathématiques [64]. Ils ont exploré comment les étudiants perçoivent l'efficacité de l'apprentissage en ligne dans l'acquisition de compétences mathématiques. Leur étude met en évidence l'importance de l'engagement des étudiants dans le processus d'apprentissage en ligne pour favoriser la compréhension et l'application des concepts mathématiques.

**Étude sur la présence sociale et l'engagement des étudiants dans les environnements d'apprentissage en ligne** : Jakkaew, Wongwanich et Xie ont réalisé une étude portant sur la présence sociale, l'engagement des étudiants et les environnements d'apprentissage en ligne [25]. Ils ont examiné comment la présence sociale des étudiants dans les environnements d'apprentissage en ligne influence leur engagement et leur participation active. Leur étude souligne l'importance de la création d'un environnement en ligne propice à la collaboration, à l'interaction sociale et à l'engagement des étudiants pour favoriser un apprentissage efficace.

**Étude sur la motivation des apprenants adultes à apprendre en ligne** : Kupiainen, Hautala, Miettinen et Ruokamo ont réalisé une étude sur la motivation des apprenants adultes à apprendre en ligne [18]. Leur étude a examiné les facteurs qui influencent la motivation des adultes à poursuivre des études en ligne, en se concentrant sur les aspects tels que la flexibilité, l'interaction sociale, l'autonomie et la pertinence du contenu. Leurs résultats mettent en évidence l'importance de ces facteurs pour maintenir la motivation des apprenants adultes dans un environnement d'apprentissage en ligne.

Ces études contribuent à la compréhension de l'impact de l'apprentissage en ligne, de la présence sociale et de la motivation des apprenants sur l'efficacité de l'enseignement dans des contextes spécifiques, notamment en promotion sociale et pour les apprenants adultes. Elles soulignent également l'importance de l'engagement des étudiants, de l'adaptabilité des enseignants et de l'utilisation d'outils interactifs pour faciliter l'apprentissage de l'informatique et des sciences des données.

## 2.3 Modélisation de bases de données et visualisation de données

### Rôle de la visualisation de données dans la modélisation de bases de données

La modélisation de bases de données est un processus essentiel dans le domaine de la gestion des données. Il consiste à créer des schémas conceptuels, logiques et physiques pour représenter et structurer les données dans une bases de données. D'un autre côté, la visualisation de données est une technique puissante qui permet de représenter graphiquement les données afin de faciliter leur analyse et leur interprétation. Bien que la modélisation de bases de données et la visualisation de données soient des domaines distincts, ils sont étroitement liés et se complètent mutuellement. En effet, une bonne modélisation des données facilite l'analyse et la visualisation

ultérieure des données, car elle fournit une structure logique et cohérente pour leur représentation (Elmasri et Navathe, 2016 [32]).

### **Techniques et outils pour intégrer la visualisation de données dans la modélisation de bases de données**

L'utilisation de la modélisation de bases de données peut considérablement améliorer la visualisation des données. Voici quelques-unes des manières dont la modélisation peut contribuer à la visualisation des données :

1. **Organisation logique et structurée** : La modélisation de bases de données permet d'organiser les données de manière cohérente et structurée. Cela facilite l'accès, la manipulation et l'analyse ultérieure des données, ce qui est essentiel pour créer des visualisations de données significatives et informatives (Elmasri et Navathe, 2016 [32]).
2. **Réduction des erreurs et des incohérences** : Une modélisation de qualité contribue à réduire les erreurs et les incohérences dans les données. Lorsque les données sont précises et fiables, les visualisations qui en découlent seront également plus précises et fiables, ce qui renforce la confiance des utilisateurs dans les résultats obtenus (Elmasri et Navathe, 2016 [32]).
3. **Intégration de sources de données** : La modélisation de bases de données facilite également l'intégration de différentes sources de données. En reliant et en combinant des ensembles de données hétérogènes, il devient possible de créer des visualisations plus riches et complexes, fournissant ainsi des informations plus approfondies et pertinentes (Heer et al., 2005 [82]).

En intégrant la modélisation de bases de données à la visualisation de données, il devient possible de créer des visualisations plus puissantes, précises et exploitables, ouvrant ainsi de nouvelles perspectives pour l'analyse et la compréhension des données.

### **Études de cas et expériences d'utilisation de la visualisation de données pour faciliter la compréhension des bases de données**

Plusieurs expériences et études sur l'enseignement de l'informatique en promotion sociale ont été rapportées dans la littérature, mettant en évidence les défis et les réussites de cette approche pédagogique. Par exemple, des études sur l'enseignement de la programmation et de l'analyse de données en promotion sociale ont souligné l'importance de l'engagement des étudiants, de la motivation intrinsèque et de l'adaptabilité des enseignants pour faciliter l'apprentissage (Merrill, 2002 Kasworm, 2003 [84], [86]). Bisschop et Tondeur [70] ont étudié les compétences pour l'intégration des TIC dans l'enseignement aux adultes, se focalisant spécifiquement sur les compétences nécessaires pour intégrer les technologies de l'information et de la communication (TIC) dans l'enseignement aux adultes. Ils se sont intéressés à identifier les compétences que les enseignants doivent développer pour utiliser efficacement les TIC dans l'éducation des adultes. En revanche, l'étude de Vincent C.[83] examine plus largement les technologies de l'information et de la communication dans l'éducation des adultes. Il explore les différentes dimensions et implications des TIC dans le contexte de l'éducation des adultes, en examinant leur utilisation, leurs avantages, leurs limites, etc.

Piette [55] a abordé la pédagogie de la coopération et de l'enseignement en promotion sociale, tandis que Delvaux et al. [16] ont étudié une expérimentation de « scénario hybride » dans l'enseignement supérieur de promotion sociale. Deschryver [50] a examiné l'apprentissage autonome, la motivation et le contexte socioculturel dans la formation d'adultes en promotion sociale. Gillard et al. [57] ont étudié la formation, le travail et les technologies de l'information et de la communication du point de vue des enseignants de promotion sociale en Communauté française de Belgique. Bachelet et al. [31] ont étudié les compétences informationnelles des étudiants inscrits en promotion sociale en Communauté française de Belgique.

Plusieurs exemples et études de cas sur l'utilisation de la modélisation de bases de données pour la visualisation de données ont été rapportés dans la littérature. Par exemple, des études sur l'utilisation de modèles de

données relationnels pour faciliter la création de tableaux de bord et de rapports interactifs ont montré que la modélisation de bases de données peut contribuer à améliorer la qualité et l'utilité des visualisations de données (Few, 2006[80]; Elmasri et Navathe, 2016 [32]). Lee et al. [36] ont étudié comment les personnes comprennent des visualisations inhabituelles et ont développé un modèle basé sur le terrain pour la création de sens des visualisations d'information par les novices (voir figure 5).

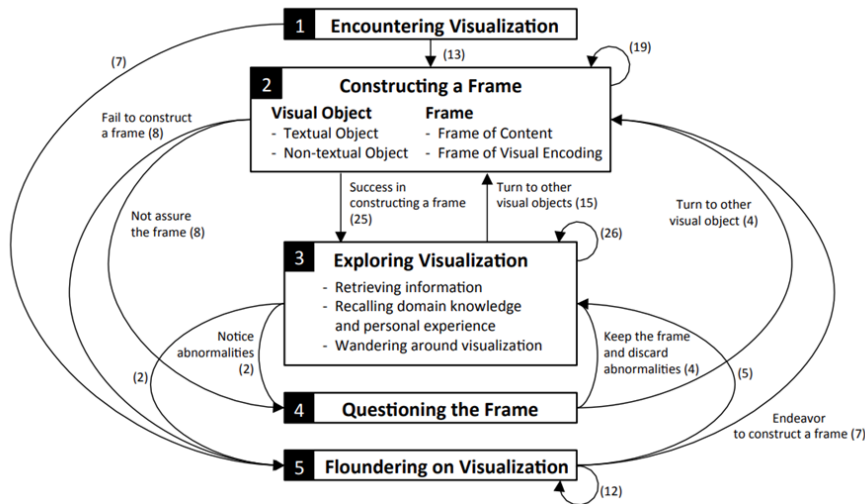


FIGURE 5 – Un modèle ancré de la visualisation de l'information pour les novices (modèle NOVIS). Les flèches indiquent les principales transitions entre les cinq activités et les nombres entre parenthèses indiquent le nombre de transitions entre les activités que nous avons observé à partir des données[36]

### Études de cas sur l'insertion d'erreur dans la modélisation de bases de données

L'insertion d'erreur dans la modélisation de bases de données a été étudiée comme outil pédagogique dans plusieurs études. Dans l'étude de Katz [33] (voir figure les auteurs ont utilisé l'apprentissage par les erreurs pour enseigner la modélisation conceptuelle de bases de données relationnelles. Les résultats ont montré que cette méthode d'enseignement était efficace pour améliorer la compréhension des concepts de modélisation et la performance des étudiants de niveau universitaire. Une autre étude, a examiné l'impact de différents types d'erreurs sur l'apprentissage de la programmation. Les résultats ont montré que l'insertion d'erreurs fréquentes était plus bénéfique pour l'apprentissage que l'insertion d'erreurs rares (Beege [3]) avec des étudiants et des apprenants dans le domaine de la programmation.

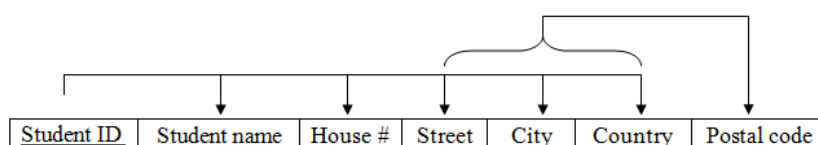


FIGURE 6 – Exemple de dépendances entre attributs qui violent la 3NF[33]

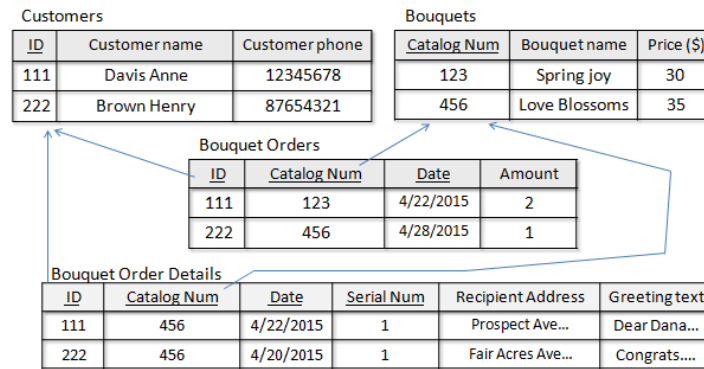


FIGURE 7 – Un schéma partiel d'un scénario de magasin de fleurs pour démontrer des exemples d'exercices erronés[33]

Enfin, dans l'étude de Maik Beege [4], les auteurs ont étudié l'utilisation de l'apprentissage par les erreurs pour améliorer la compréhension des concepts de modélisation de bases de données (voir figure 8 et 9). Les résultats ont montré que cette méthode d'enseignement était efficace pour aider les étudiants à atteindre un niveau conceptuel supérieur dans la modélisation de bases de données. Ces résultats suggèrent que l'insertion d'erreur peut être un outil pédagogique efficace pour enseigner la modélisation de bases de données et d'autres sujets liés à la data science.

Erroneous code:

```
main.py [icon] saved
1 def hallo ():
2     print ('Hallo, liebe Versuchspersonen!')
3
```

Correct code:

```
main.py [icon] saved
1 def hallo ():
2     print ('Hallo, liebe Versuchspersonen!')
3 hallo ()
```

FIGURE 8 – Exemple d'erreur sémantique : invocation manquante de la fonction hallo (la version correcte est affichée ci-dessous).[4]

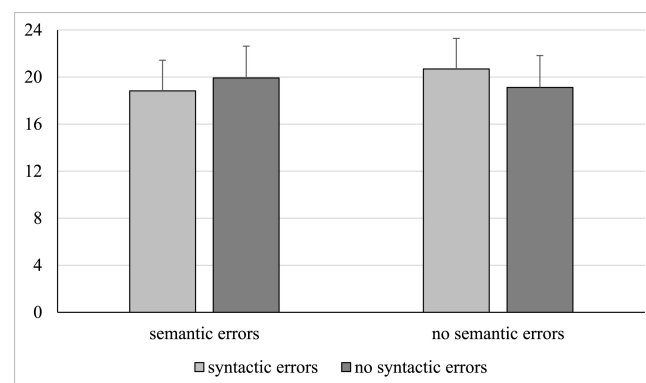


FIGURE 9 – Effet d'interaction de l'inclusion des erreurs syntaxiques et des erreurs sémantiques sur les performances des tests de connaissances sémantiques ; M et SD sont affichés.[4]

## 2.4 Conclusions l'état de l'art

### Résumé des principaux résultats de l'état de l'art

L'état de l'art met en évidence l'importance de la visualisation de données dans l'enseignement et l'apprentissage, particulièrement dans le contexte de la promotion sociale et de la formation en informatique. Il montre également que la modélisation de bases de données et la visualisation de données sont étroitement liées, et qu'une bonne modélisation des données peut faciliter l'analyse et la visualisation des données (Elmasri et Navathe, 2016 [32]). De plus, des approches pédagogiques adaptées et flexibles sont nécessaires pour enseigner la visualisation de données à un public diversifié, comme celui de la promotion sociale (Merrill, 2002 ; Kasworm, 2003 [84], [86]).

Malgré l'existence de plusieurs études examinant la visualisation de données et l'enseignement de l'informatique en promotion sociale (Bisschop et Tondeur [70], Vincent [83], Piette [55], Delvaux et al. [16], Kerren [74], Heer, J. [58]), il persiste des lacunes et des opportunités de recherche dans ce domaine. En particulier, les aspects suivants méritent une attention approfondie :

1. **L'impact de l'intégration de la visualisation de données dans les programmes d'études :** Il est important d'étudier comment l'incorporation de la visualisation de données dans les programmes d'études peut affecter les compétences et les résultats des étudiants en promotion sociale. Les recherches actuelles, y compris celles sur les approches pédagogiques (Isenberg et al. [65], Ruggles et al. [28]), n'ont pas encore pleinement exploré cet aspect (Deschryver [50], Gillard et al. [57]).
2. **L'identification des meilleures pratiques pour enseigner la visualisation de données aux étudiants novices en data science et en modélisation de bases de données :** Les recherches actuelles n'ont pas encore déterminé les approches pédagogiques les plus efficaces pour enseigner la visualisation de données aux étudiants novices, notamment en matière de data science et de modélisation de bases de données (Lee et al. [36], Few [80], Elmasri et Navathe [32]).
3. **L'évaluation des outils et des ressources pédagogiques adaptés à l'enseignement de la visualisation de données en promotion sociale :** Les études existantes, y compris celles qui examinent les outils et méthodes de visualisation (Card et al. [90], Yi et al. [79]), n'ont pas encore évalué de manière approfondie les outils et les ressources pédagogiques qui peuvent être adaptés spécifiquement à l'enseignement de la visualisation de données en promotion sociale, en tenant compte des besoins et des contraintes des étudiants dans ce contexte (Bachelet et al. [31]).

### Implications pour le mémoire et la recherche future

Les résultats de cette revue de la littérature, incluant les travaux de Bisschop et Tondeur [70], Vincent [83], Piette [55], Delvaux, E. [16], Kerren [74], Heer et al. [58], ont des implications importantes pour le mémoire et la recherche future dans ce domaine. Ils suggèrent, par exemple, de :

1. **Explorer des approches pédagogiques innovantes pour enseigner la visualisation de données en promotion sociale :** En tenant compte des besoins et des contraintes des étudiants, il est essentiel d'étudier de nouvelles méthodes pour enseigner la visualisation de données en promotion sociale, en s'appuyant sur les travaux existants sur les approches pédagogiques (Isenberg et al. [65], Ruggles et al. [28]).
2. **Évaluer l'efficacité de ces approches à travers des études empiriques :** Il est important de mesurer l'évolution des compétences et des connaissances des étudiants avant et après l'intervention en utilisant des approches pédagogiques innovantes, en s'inspirant des études existantes sur l'enseignement de l'informatique et de la visualisation de données (Lee et al. [36], Few [80], Elmasri et Navathe [32]).
3. **Contribuer à combler les lacunes identifiées dans la littérature :** En examinant l'impact de l'intégration de la visualisation de données dans les programmes d'études et en identifiant les meilleures



pratiques pour enseigner ce domaine aux étudiants novices en data science, les futurs travaux pourront aider à combler les lacunes actuelles dans la littérature (Deschryver [50], Gillard et al. [57], Card et al. [90], Yi et al. [79], Bachelet et al. [31]).

## 2.5 Question de recherche

L'élaboration de la question de recherche a été effectuée en se basant sur l'état de l'art et en ayant pour objectif de répondre à la problématique identifiée. La question de recherche a donc été élaborée en plusieurs étapes. Tout d'abord, il a été nécessaire d'identifier le contexte spécifique de l'enseignement de promotion sociale, ainsi que les compétences et les connaissances des étudiants novices en data science dans ce domaine. Ensuite, l'étude de la littérature a permis de mettre en évidence l'importance de la visualisation de données pour faciliter la compréhension et la communication des informations complexes (Lee, 2016, Card, 1999, Ware, 2014 [36], [48], [90]).

La question de recherche a ensuite été formulée en prenant en compte ces éléments, en se concentrant sur la manière dont la visualisation de données peut améliorer la représentation de données pour un public novice en data science dans le cadre de l'enseignement de promotion sociale. Plus précisément, il s'agit de comprendre comment la modélisation bases de données peuvent être utilisées comme porte d'entrée pour enseigner la visualisation de données, en prenant en compte les spécificités des étudiants en promotion sociale (Munzner, 2014 [45], Sedlmair, 2012 [60]).

La visualisation de données étant un élément clé dans l'analyse et la communication des informations, il est essentiel d'explorer les méthodes les plus efficaces pour enseigner cette compétence aux étudiants novices en promotion sociale. En utilisant les bases de données comme porte d'entrée pour enseigner la visualisation de données, les enseignants peuvent aider les étudiants à comprendre les concepts et les techniques associés à la visualisation de données tout en utilisant un contexte familier. Ainsi, la question de recherche "*Comment enseigner la visualisation de données à un public novice de promotion sociale par le biais d'un cours sur les bases de données ?*" a été élaborée pour comprendre comment cette approche peut être optimisée pour répondre aux besoins des étudiants de promotion sociale.

### 3 Méthodologie

La méthodologie utilisée pour cette étude repose sur une approche mixte qui combine plusieurs techniques de collecte de données afin d'évaluer l'efficacité de l'enseignement de la visualisation de données à l'aide de la modélisation de bases de données pour un public novice en data science [60]. Les techniques utilisées incluent des questionnaires, des entretiens semi-structurés, des focus groups et l'observation participante (voir figure 10).

L'observation participante serait utilisée comme une méthode spécifique permettant au chercheur de s'immerger dans l'activité étudiée, d'observer directement les interactions et de recueillir des données pertinentes. Dans ce contexte, l'observation participante serait utilisée pour obtenir une compréhension approfondie de l'activité d'enseignement de la visualisation de données et de la modélisation de bases de données. Elle permettrait au chercheur d'observer les interactions entre l'enseignant et les apprenants, les dynamiques de classe, et de recueillir des informations contextuelles cruciales.

En combinant ces différentes techniques de collecte de données, la méthodologie adoptée permettrait de recueillir des perspectives multiples et complémentaires sur l'enseignement de la visualisation de données. Les questionnaires fourniraient des données quantitatives sur les connaissances et les perceptions des apprenants, les entretiens semi-structurés approfondiraient les points de vue de l'enseignant, et les focus groups favoriseraient les échanges et les discussions collectives avec les étudiants.

Dans l'ensemble, cette méthodologie mixte permettrait d'obtenir une compréhension approfondie des enjeux, des besoins et des attentes des différentes parties prenantes, en combinant des données qualitatives et quantitatives. Elle permettrait également au chercheur de s'immerger dans l'activité d'enseignement grâce à l'observation participante, fournissant une perspective immersive pour une meilleure analyse et interprétation des résultats obtenus.

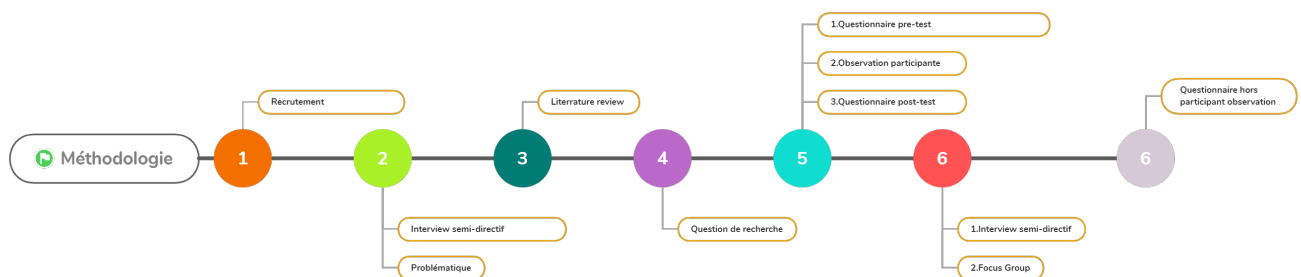


FIGURE 10 – Diagramme de flux de la méthodologie

La méthodologie combinera plusieurs techniques de collecte de données, notamment l'observation participante (Angrosino, 2007) [75], les questionnaires (Fowler et al., 2013) [51], les focus groups détaillés par Krueger (2014) [44], et les entretiens semi-directifs (Bernard, 2011) [62] (voir figure 11). L'utilisation de ces différentes techniques permettra de recueillir des données variées et complémentaires, fournissant ainsi une perspective approfondie et holistique sur l'efficacité des approches pédagogiques utilisées dans l'enseignement de la visualisation de données à ce public spécifique.

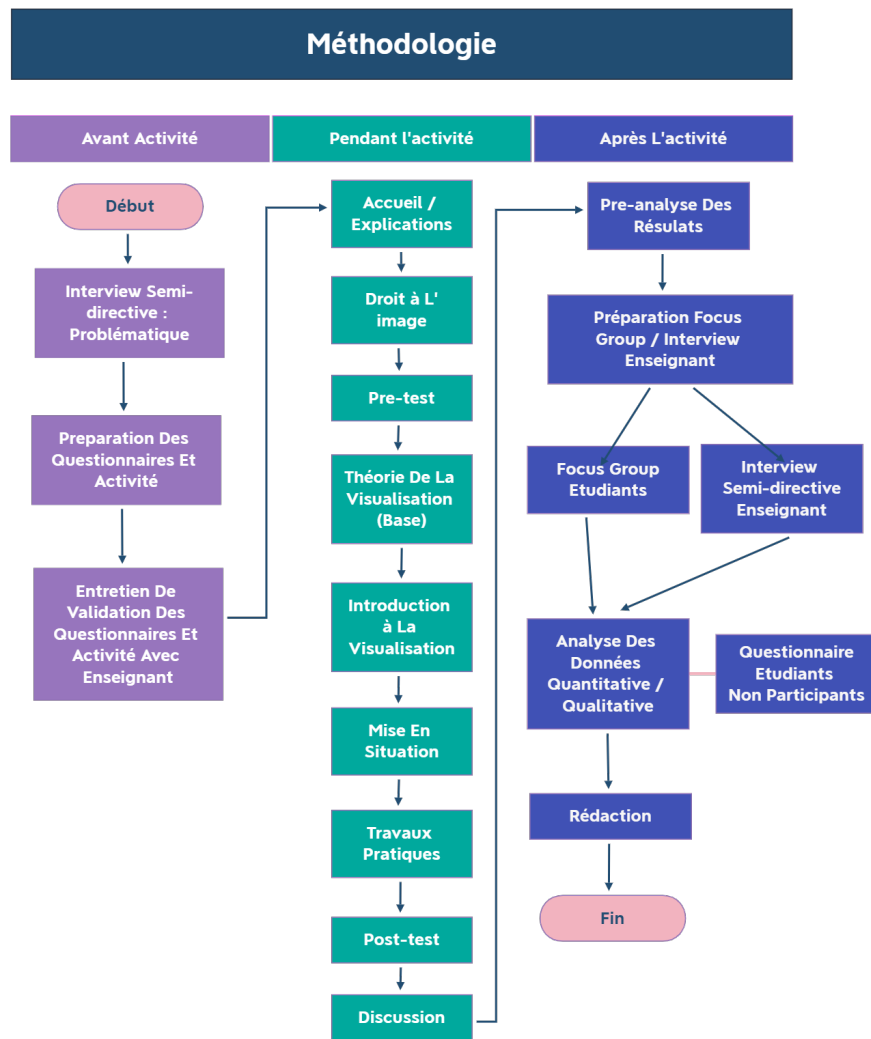


FIGURE 11 – Diagramme de flux de la méthodologie pour l'observation participante

Les résultats de cette recherche permettront d'identifier les meilleures pratiques et de formuler des recommandations pour l'intégration de la visualisation de données dans le cursus des étudiants en promotion sociale. Ainsi, plusieurs sous-questions de recherche pourront être abordées en lien avec l'état de l'art, telles que celles formulées par Ware (2014) [48] :

- *Comment enseigner la visualisation de données à un public novice de promotion sociale par le biais d'un cours sur les bases de données ?*
- *Comment intégrer la visualisation de données dans le cursus scolaire des étudiants de promotion sociale novices en data science, dans le cadre d'un bachelier en informatique de gestion ?*

Pour répondre à ces questions, la méthodologie de recherche adoptée combinera plusieurs techniques de collecte de données. L'observation participante permettra d'observer directement les interactions entre l'enseignant et les apprenants, ainsi que les dynamiques de classe, afin de comprendre comment l'enseignement de la visualisation de données est effectivement réalisé dans un cours sur les bases de données. Les questionnaires permettront de recueillir des données quantitatives sur les connaissances et les perceptions des apprenants, fournissant ainsi des informations sur l'efficacité de l'approche pédagogique utilisée. Les focus groups et les entretiens semi-directifs permettront d'approfondir les perspectives des participants, de recueillir des informations qualitatives détaillées sur leurs expériences, leurs préférences et leurs attentes.

L'analyse et l'interprétation des données collectées à l'aide de ces différentes techniques permettront de répondre aux questions de recherche de manière approfondie et holistique. Les résultats fourniront des informations sur les meilleures pratiques pour enseigner la visualisation de données à un public novice de promotion sociale, ainsi que des recommandations spécifiques pour intégrer la visualisation de données dans le cursus scolaire des étudiants de promotion sociale novices en data science. L'ensemble de ces résultats contribuera à améliorer l'enseignement de la visualisation de données dans le contexte particulier des étudiants en promotion sociale et servira de base pour formuler des recommandations pratiques adaptées à ce public spécifique.

Les outils de collecte de données sont présentés ci-dessous :

Deux interviews seront menées, la première est une interview semi-directive ayant pour objectif de définir la problématique, le contexte social et pédagogique. La deuxième, est aussi une interview semi-directive ayant pour objectif la consolidation des récoltes de données liées à l'activité d'enseignement (pre-test, post-test et activité d'enseignement) et partager la vision de l'enseignant sur l'implication de la visualisation de données dans le cadre de ses cours ou du cursus scolaire du bachelier.

Les entretiens semi-directifs sont une autre méthode de recherche qualitative utile, où le chercheur a un guide d'entretien flexible avec des questions ouvertes. Cela permet à l'entretien de suivre une certaine direction, tout en laissant la liberté aux étudiants d'explorer d'autres thèmes qu'ils jugent pertinents (Bernard, 2011 [63]). Cela peut aider à comprendre comment l'enseignant perçoit et enseigne la visualisation de données, et quelles sont les stratégies pédagogiques qu'il utilise. Cela peut également révéler les défis que l'enseignant rencontre dans l'enseignement de la visualisation de données aux étudiants novices en data science (Patton, 2014 [47]).

De plus, cela peut fournir des informations sur la manière dont l'enseignant évalue l'efficacité de l'enseignement de la visualisation de données, et comment il s'adapte aux besoins d'apprentissage de différents étudiants (Seidman, 2016 [81]). Enfin, l'interview semi-directive peut fournir un espace pour que l'enseignant partage ses réflexions et expériences personnelles, ce qui peut apporter une perspective riche et nuancée à la recherche (Bernard, 2011 [63]).

Un entretien avant l'activité sera réalisé pour permettre la présentation des questionnaires, de l'activité et de la question de recherche ainsi qu'une synthèse de l'avancé du mémoire. Il a aussi pour but, de valider les contenus avec l'enseignant.

Des questionnaires au nombre de 3 et mélangent des questions qualitatives et quantitatives sont utilisés. Elles sont rédigées sur la plate-forme Google Forms, la première et la deuxième seront encodées au format papier et la troisième de manière électronique.

Les questionnaires pré-test et post-test serviront à évaluer les connaissances initiales et finales des étudiants en matière de visualisation de données. Le questionnaire pré-test (préalables) évaluera l'expérience antérieure, la compréhension des concepts clés et la confiance en leurs compétences des étudiants. Le questionnaire post-test (post-observation) évaluera l'évolution des connaissances des étudiants, les compétences acquises, leur niveau de confiance en leurs compétences et leur expérience avec les activités de l'observation participante. Ces données permettront d'évaluer l'efficacité de l'enseignement de la visualisation de données à l'aide de bases de données pour le public cible de l'étude.

---

8. [https://www.google.com/intl/fr\\_be/forms/](https://www.google.com/intl/fr_be/forms/)

### 3.1 Contexte de l'étude

Pour recruter des étudiants de cette étude, plusieurs méthodes de recrutement couramment utilisées dans les études scientifiques ont été explorées, en tenant compte du contexte spécifique. Une prospection a été réalisée par email auprès de différentes écoles de promotion sociale offrant des bacheliers en informatique de gestion, notamment à Charleroi, La Louvière et Namur. L'EAFIC de Namur a répondu positivement à cette demande. La sous-directrice de l'établissement a exprimé l'intérêt de l'école à participer à l'étude et a mis en contact le chercheur avec M. Vonèche, le professeur en charge de l'UE "Gestion et exploitation de bases de données".

Le public cible de cette étude est constitué d'étudiants suivant des cours en soirée dans le cadre de la promotion sociale, avec un intérêt particulier pour le domaine informatique. Plus précisément, nous avons limité notre étude aux étudiants de la section "Bachelier en informatique de gestion" en première année, ainsi qu'à la classe de première année de la section "Technicien en programmation". Au total, il y a 69 étudiants inscrits dans cette section sur une durée de 4 ans, dont 5 étudiants en première année de "Bachelier en informatique de gestion", 7 étudiants en troisième année de "Bachelier en informatique de gestion" et 2 étudiants en première année de "Technicien en programmation". Parmi ces étudiants, 7 sont inscrits à les unités d'enseignement dispensées par M. Vonèche, tous en première année de "Bachelier en informatique de gestion" et en technicien informatique.

#### Contexte

Les activités de récolte de données se sont déroulées sur le site de l'EAFIC de Namur, une école offrant un enseignement pour adultes et en formation continue pour le public cible en promotion sociale. Les cours sont donnés en soirée avec un horaire de 17h45 à 21h05 du lundi au samedi suivant l'année d'étude, notamment le bachelier en informatique de gestion. Le bachelier en informatique de gestion de l'EAFIC de Namur est un programme de formation destiné aux étudiants qui souhaitent développer des compétences en informatique, en gestion et en analyse de données. Ce programme est dispensé en promotion sociale, permettant aux étudiants de suivre les cours en soirée pour concilier leurs études avec d'autres obligations, telles que le travail ou la vie familiale. Le cursus du bachelier en informatique de gestion couvre un large éventail de sujets, notamment la programmation, les bases de données, la gestion de projet, l'analyse des systèmes d'information, la sécurité informatique et la gestion des réseaux. Les étudiants apprennent également des compétences en gestion et en communication pour être en mesure de travailler efficacement au sein d'équipes interdisciplinaires et de gérer des projets informatiques.

Tout au long du programme, les étudiants ont l'occasion de mettre en pratique les compétences acquises au travers de projets, d'études de cas et de stages en entreprise. Cette approche pratique vise à préparer les étudiants à une carrière dans le domaine de l'informatique de gestion en les familiarisant avec les défis et les opportunités du secteur. À l'issue du programme, les diplômés du bachelier en informatique de gestion de l'EAFIC de Namur sont prêts à occuper des postes variés tels que développeur de logiciels, analyste de systèmes d'information, gestionnaire de bases de données, consultant en technologies de l'information ou chef de projet informatique, entre autres.

### 3.2 Conception de l'activité

#### Interview semi-directive avant question de recherche

L'interview semi-directive préalable à la définition de la question de recherche sera menée auprès de l'enseignant en charge du cours de "modélisation de données". Cette entrevue a pour objectif de récolter des informations permettant de comprendre le contexte du cours, notamment le nombre d'étudiants, la fréquentation, le type

d'étudiants, leurs expériences, leurs ressentis face à la donnée et l'organisation du module. Elle servira également à déterminer la population cible, aborder les thématiques vues dans ses cours, identifier les problèmes rencontrés par l'enseignant dans le cadre de son cours et de son activité d'enseignant dans le contexte de la promotion sociale, ainsi qu'à discuter des attentes de l'enseignant par rapport au mémoire. Enfin, cette interview permettra de communiquer avec l'enseignant sur le déroulement du mémoire et de lui faire une proposition de planning. Le guide d'entretien se trouve en annexes.

## Enquêtes

Les enquêtes (pre-test et post-test) ont été construit avec un certain nombre d'objectif pédagogique (voir Annexe A) :

**Informations générales (exclusivement au pre-test)** Cette section vise à recueillir des informations démographiques sur les étudiants. Ces informations peuvent être utilisées pour analyser les réponses en fonction de différents groupes (par exemple, par âge, par niveau d'éducation, par genre, etc.)[76].

**Compréhension de la modélisation de données** Cette section évalue la compréhension de base des étudiants en matière de modélisation de données, notamment leur capacité à identifier les composants clés d'un tableau de données (attributs, items, cellules) abordé par Kent [88] et Brehmer et al[14].

**Compréhension de la visualisation de données** Cette section évalue la compréhension des étudiants de ce qu'est la visualisation de données et de la façon dont elle se distingue d'autres formes de représentation de données (par exemple, une table de base de données ou un tableau Excel)[61].

**Utilité de la visualisation de données** Cette section vise à évaluer comment les étudiants perçoivent l'utilité de la visualisation de données, tant pour eux-mêmes que pour d'autres, et à comprendre dans quelle mesure ils pensent que la visualisation de données pourrait leur être utile dans leurs études et leur carrière future[71].

**Choix de la meilleure visualisation pour un scénario donné** Cette section évalue la capacité des étudiants à choisir la visualisation de données la plus appropriée pour un scénario donné et à expliquer les raisons de leur choix. Cela permet d'évaluer leur compréhension de l'utilisation pratique de la visualisation de données[56].

**Compétences en visualisation de données** Cette section évalue les compétences pratiques des étudiants en matière de visualisation de données, en leur demandant s'ils seraient capables de créer différents types de visualisations à partir d'un tableau de données. Elle évalue également leur compréhension des principes de base de la visualisation de données, tels que la nécessité de ne pas toujours afficher toutes les données disponibles et le fait que la création d'une visualisation est un processus à plusieurs étapes[49].

**Compréhension de l'activité (exclusivement au Post-test)** Cette section vise à recueillir des commentaires sur les difficultés rencontrées par les étudiants lors de l'activité et à évaluer leur confiance dans leur compréhension de l'ensemble de données. Cela peut aider à identifier les domaines qui peuvent nécessiter des explications ou des instructions supplémentaires à l'avenir[34].

Le questionnaire pré-test et post-test est administré aux étudiants en format papier au début et à la fin de l'activité (voir tableau 2 et tableau 4). Les étudiants remplissent les questionnaires de manière anonyme pour assurer l'intégrité des résultats. Les discussions pendant l'activité sont enregistrées à l'aide d'un dispositif d'enregistrement audio placé au centre de la salle de classe pour assurer une couverture maximale. Les enregistrements sont ensuite transcrits pour analyse.

D'un point de vue des considérations éthiques, tous les étudiants à l'activité d'enseignement ont donné leur consentement éclairé avant de participer à l'étude. Ils ont été informés de leurs droits, y compris le droit de se retirer de l'étude à tout moment. Toutes les données recueillies lors de l'activité sont traitées de manière

anonyme pour protéger la confidentialité des étudiants.

| Question   | Questionnaire        | Justification  |
|--|----------------------|--|
| Q0_1 : Situation professionnelle   | Pré-test             | Cette question aide à comprendre le contexte professionnel des étudiants, ce qui peut influencer leur capacité à comprendre et utiliser la visualisation de données. C'est pertinent pour la question de recherche car il peut y avoir des différences dans la compréhension de la visualisation de données en fonction du contexte professionnel. |
| Q1_1 : Parmi les propositions suivantes, sélectionner le tableau où un attribut est entouré  | Pré-test - Post-test | Cette question évalue la compréhension de base des étudiants de la visualisation de données. Cela nous aide à comprendre leur niveau de compétence initial, ce qui est crucial pour mesurer l'efficacité de l'intervention.  |
| Q21 : Est-il possible pour moi de réaliser une visualisation similaire à celle-ci en disposant du tableau de données? (il s'agit d'un graphique linéaire, axe des y les revenus, axe des x les mois de l'année sur une période de 2 ans) | Pré-test - Post-test | Cette question évalue la capacité des étudiants à créer une visualisation de données après l'intervention. C'est crucial pour répondre à la question de recherche, car cela nous indique si l'intervention a amélioré leur capacité à utiliser la visualisation de données.  |
| Q28 : Quelles sont les difficultés que j'ai rencontrées pendant l'activité?  | Post-test            | Cette question donne un aperçu des défis que les étudiants ont rencontrés pendant l'intervention, ce qui peut aider à identifier les domaines qui nécessitent une amélioration ou un soutien supplémentaire. C'est pertinent pour la question de recherche car cela peut aider à expliquer les variations dans les résultats entre les étudiants.  |

TABLE 2 – Tableau d'exemple des questions sélectionnées du questionnaire

Le troisième questionnaire a été mis en place dans le but d'offrir une perspective complémentaire à notre enquête. Il a été administré à un panel d'étudiants suivant le même cursus, mais n'ayant pas participé à l'activité de visualisation de données, et par conséquent, n'ayant pas été soumis aux questionnaires pré- et post-activité. Cette démarche s'inscrit dans la lignée des recherches en pédagogie où il est souvent recommandé d'établir une comparaison entre les résultats des étudiants ayant bénéficié d'une intervention spécifique et un groupe de contrôle n'ayant pas été exposé à cette même intervention [38].

L'idée de ce troisième questionnaire était de mesurer la compréhension et la compétence en visualisation de données parmi un groupe d'étudiants n'ayant pas été exposés à l'activité proposée. Cette démarche permet d'établir un point de référence en termes de compétences en visualisation de données parmi les étudiants de ce cursus. Ces données, couplées aux résultats des questionnaires pré- et post-activité, offrent une vue d'ensemble sur l'efficacité de l'activité en termes d'amélioration des compétences en visualisation de données [69], [78].

### Activité d'enseignement

Le lieu de l'activité d'enseignement est une salle de classe de l'E AFC située sur la Place de l'Ecole des Cadets à 5000 Namur, au troisième étage dans le local C307, pouvant accueillir jusqu'à 30 participants. L'activité est

structurée en trois grandes phases et une présentation PowerPoint sert de fil conducteur tout au long de l'activité. L'activité d'enseignement est basée sur une observation participante qualitative. L'observation participante est une forme particulière d'observation naturaliste dans laquelle le chercheur s'engage activement dans la situation qu'il ou elle observe. Cela signifie généralement que le chercheur devient un membre du groupe ou de la communauté qu'il ou elle étudie, participant à ses activités et interagissant avec ses membres (Angrosino, Michael V, 2007[75], Noss, Richard et Hoyles, Celia, 1996 [92], Margolis, Jane, 2002 [85]). L'observation participante a été choisie comme méthode de recherche en raison de sa capacité à fournir un aperçu approfondi du processus d'apprentissage des étudiants. Contrairement aux méthodes de recherche plus traditionnelles, l'observation participante permet au chercheur de s'immerger dans la situation d'apprentissage, offrant une perspective plus nuancée et détaillée de la manière dont les étudiants interagissent avec le matériel d'apprentissage et entre eux.

Les objectifs d'apprentissage spécifiques de cette activité d'enseignement sont triples : (1) les étudiants seront capables de comprendre les principes fondamentaux de la visualisation de données, y compris les différents types de visualisations et leurs usages ; (2) ils seront capables d'appliquer ces principes pour créer leurs propres visualisations de données à partir de jeux de données présentés ; et (3) ils seront capables de critiquer et d'évaluer efficacement la qualité des visualisations de données (Bonnemains, Chaste, et Berger, 2020 [8]).

La réalisation de l'activité se fait en trois phases : avant, pendant, et après. La phase "Avant" consiste à préparer l'activité et à recueillir des données de base. Cette phase comprend plusieurs étapes, comme le remplissage d'une grille d'évaluation de l'activité par le chercheur, la préparation du matériel d'enregistrement vidéo, l'accueil et l'explication de l'activité aux étudiants, et la distribution et la collecte des questionnaires pré-activité. Ces étapes sont conformes aux recommandations de Bonnemains, Chaste, et Berger (2020)[8] sur la mise en place d'une structure d'activité claire et précise.

La phase "Pendant" est la partie centrale de l'activité d'enseignement. Elle comprend une introduction à la théorie de base de la visualisation de données, l'introduction à la visualisation de données en utilisant le livre de Munzner Tamara comme référence, une mise en situation basée sur une modélisation de données fournie par l'enseignant, et enfin, un travail pratique où les étudiants sont amenés à représenter le cycle de vie d'une visualisation de données et à proposer une visualisation à partir d'un énoncé contenant des erreurs. Ces activités sont alignées avec l'idée que l'apprentissage se produit mieux dans un contexte authentique (Noss et Hoyles, 1996[92]).

La phase "Après" se divise en quatre parties : le remplissage du questionnaire post-test par les étudiants, une discussion ouverte sur l'activité, une explication des objectifs de l'étude et des démarches entreprises pour recueillir et analyser les données (comme recommandé par Angrosino, 2007[75]), et enfin, une explication de la manière dont les données seront utilisées et des mesures prises pour garantir la confidentialité des participants (conformément à Margolis, 2002[85]).

### 3.3 Évaluation de l'activité

#### Grille d'évaluation

L'approche méthodologique pour évaluer l'activité de visualisation de données se concentre sur une évaluation holistique des performances et des comportements des étudiants. Cette approche utilise une grille d'évaluation qui permet d'évaluer de manière systématique et cohérente plusieurs aspects de l'activité (Leung, 2015) [40] et est décomposé comme suit.

**Participation des étudiants :** Comment les étudiants se sont-ils engagés dans l'activité ? Ont-ils été actifs et impliqués tout au long de l'activité ? Ont-ils posé des questions, discuté entre eux, utilisé les outils de



visualisation de données de manière appropriée ?

**Compréhension des concepts de visualisation de données :** Les étudiants ont-ils démontré une compréhension des concepts clés de la visualisation de données lors de l'activité ? Ont-ils été capables d'interpréter les visualisations de données de manière appropriée ?

**Application des compétences de visualisation de données :** Les étudiants ont-ils été capables d'appliquer leurs compétences en visualisation de données pour analyser et interpréter les données ? Ont-ils utilisé les outils de visualisation de données de manière efficace ?

**Réactions des étudiants à l'activité :** Comment les étudiants ont-ils réagi à l'activité ? Ont-ils semblé apprécier l'activité, la trouver intéressante ou utile ? Y a-t-il eu des moments de confusion ou de frustration ?

**Progrès des étudiants :** Les étudiants ont-ils montré des progrès dans leurs compétences en visualisation de données au cours de l'activité ? Y a-t-il des preuves d'apprentissage ou de développement ?

**Effets sur l'intérêt pour la visualisation de données :** L'activité a-t-elle influencé l'intérêt des étudiants pour la visualisation de données ? Cela peut être difficile à évaluer uniquement à partir d'un enregistrement vidéo, mais il pourrait y avoir des indices dans les commentaires ou les réactions des étudiants.

**Le succès de l'activité d'enseignement :** Elle sera évalué en fonction de plusieurs critères, premièrement, une amélioration significative des scores aux tests de pré- à post-test indiquera un apprentissage réussi. Deuxièmement, les commentaires positifs des étudiants lors de la discussion du focus group indiqueront une appréciation de l'activité. Enfin, une application réussie des compétences de visualisation de données lors de l'activité, telle qu'évaluée par la grille d'évaluation, sera également un indicateur de succès.

### 3.4 Débriefing de l'activité

#### Interview semi-directive après observation participante

L'interview semi-directive après l'observation participante sera menée deux semaines après l'activité d'enseignement. Elle abordera plusieurs thèmes et questions ouvertes. Tout d'abord, elle explorera l'appréciation de l'enseignant avant l'activité et le contact pour le mémoire, en se demandant s'il avait un intérêt initial pour l'introduction de la visualisation de données dans l'enseignement de la modélisation de données, et s'il avait constaté des retours sur la visualisation de données de la part de ses étudiants avant l'activité. Ensuite, l'interview abordera l'appréciation de l'enseignant pendant l'activité, en évaluant comment il a ressenti l'activité, ce qu'il a apprécié, ce qu'il n'a pas apprécié et ce qui doit être changé. On lui demandera également s'il a constaté des difficultés particulières pendant l'activité, que ce soit pour lui-même ou pour les étudiants.

L'interview s'intéressera également à l'appréciation de l'enseignant après l'activité sur le retour des étudiants, en se demandant s'il a constaté une évolution ou des retours de ses étudiants dans les premiers cours suivis après l'activité. De plus, l'interview évaluera l'appréciation de l'enseignant après l'activité de son point de vue, en se demandant si l'intérêt pour la visualisation de données est présent en tant que professeur, et si c'est le cas, dans quelle proportion. On lui demandera également s'il est intéressant d'intégrer les concepts de visualisation de données dans le cursus scolaire des étudiants en bachelier et technicien en informatique, et si oui, sous quel format et quelle fréquence. Enfin, l'interview permettra à l'enseignant de faire des remarques complémentaires (voir Annexe B).

#### Focus group

Un focus group (groupe de discussion) est une méthode de recherche qualitative qui rassemble un petit groupe de personnes pour discuter d'un sujet spécifique sous la direction d'un chercheur. Cette méthode est particulièrement utile pour explorer les attitudes, les perceptions, les opinions et les idées des étudiants sur le sujet à l'étude [89][44]. Cela pourra permettre d'obtenir des informations détaillées et nuancées sur la manière dont les étudiants novices en data science perçoivent et interagissent avec la visualisation de données [91]. Les

discussions en groupe pourraient révéler des perspectives uniques et fournir des aperçus inattendus sur la manière dont les étudiants abordent l'apprentissage de la visualisation de données. De plus, un groupe de discussion peut encourager les étudiants à interagir et à discuter entre eux, ce qui peut générer de nouvelles idées et perspectives. Ces interactions peuvent révéler des différences d'opinion, des malentendus communs, ou des points de consensus parmi les étudiants, ce qui peut être précieux pour comprendre les défis auxquels ils sont confrontés lors de l'apprentissage de la visualisation de données [89]. Enfin, un groupe de discussion peut également fournir un environnement plus détendu et moins formel que d'autres méthodes de recherche, ce qui peut encourager les étudiants à partager librement leurs pensées et leurs expériences [93].

Le lieu, l'horaire et les interlocuteurs sont clairement définis. Le fait de tenir le focus group dans un environnement familier pour les étudiants peut aider à faciliter la conversation et à obtenir des données de meilleure qualité [89]. L'objectif du focus group est clairement défini : comprendre les attentes des étudiants concernant l'introduction de la visualisation de données dans leur formation et recueillir des commentaires sur l'activité d'enseignement. Cela correspond bien à l'objectif de votre mémoire, qui est d'identifier les meilleures pratiques pour enseigner la visualisation de données [87].

Les thèmes et les questions du focus group sont pensés et alignés sur l'objectif de la recherche. Les questions cherchent à comprendre comment les étudiants ont vécu l'activité, quels étaient leurs points forts et leurs points faibles, et comment ils aimeraient voir la visualisation de données intégrée dans leur cursus en dehors de l'activité proposée. Ce type de questions ouvertes peut aider à susciter des discussions approfondies et à recueillir des données riches et nuancées [89]. Enfin, le focus group sera enregistré, ce qui permettra de revoir les discussions et de s'assurer que toutes les données importantes sont capturées [87](voir Annexe B).

## Analyse des données

Pour l'analyse des données, une approche mixte sera adoptée. Les réponses aux questionnaires pré-test et post-test seront analysées pour déterminer si des améliorations significatives des connaissances et des compétences en visualisation de données ont eu lieu. Les transcriptions des discussions seront analysées qualitativement à l'aide de l'analyse de contenu pour identifier les thèmes et les motifs communs. L'analyse des données de cette étude sera réalisée en plusieurs étapes (voir figure 12 :

**1. Nettoyage des données :** Après la collecte des données, la première étape sera de nettoyer et de préparer les données pour l'analyse. Cela impliquera d'examiner les questionnaires pour détecter d'éventuelles erreurs ou omissions, de transcrire les enregistrements audio des discussions, et de préparer les données de la grille d'évaluation pour l'analyse.

**2. Analyse quantitative :** L'analyse quantitative portera sur les données recueillies à partir des questionnaires pré-test et post-test. Les scores des questionnaires seront analysés pour déterminer si des améliorations significatives des connaissances et des compétences en visualisation de données ont eu lieu. Des analyses de corrélation pourraient également être effectuées pour examiner les relations entre différents facteurs (par exemple, le niveau d'engagement des étudiants et l'amélioration des scores).

**3. Analyse qualitative :** L'analyse qualitative portera sur les données recueillies à partir des discussions et de la grille d'évaluation et des enregistrements vidéos. L'analyse de contenu sera utilisée pour examiner les transcriptions des discussions, avec un focus sur les thèmes récurrents relatifs à l'expérience d'apprentissage des étudiants, leurs perceptions de l'activité, et leurs commentaires sur les difficultés ou les réussites qu'ils ont rencontrées. Les données de la grille d'évaluation seront analysées pour évaluer les performances des étudiants en termes de participation, de compréhension des concepts de visualisation de données, d'application des compé-

tences de visualisation de données, de réactions à l'activité et de progrès.

**4. Intégration des résultats :** Enfin, les résultats des analyses quantitative et qualitative seront intégrés pour fournir une image complète de l'efficacité de l'activité d'enseignement. Cette approche mixte permettra de tirer parti des forces des deux types d'analyse - la capacité de l'analyse quantitative à fournir des mesures objectives et généralisables, et la capacité de l'analyse qualitative à fournir un aperçu détaillé et nuancé des expériences d'apprentissage des étudiants.

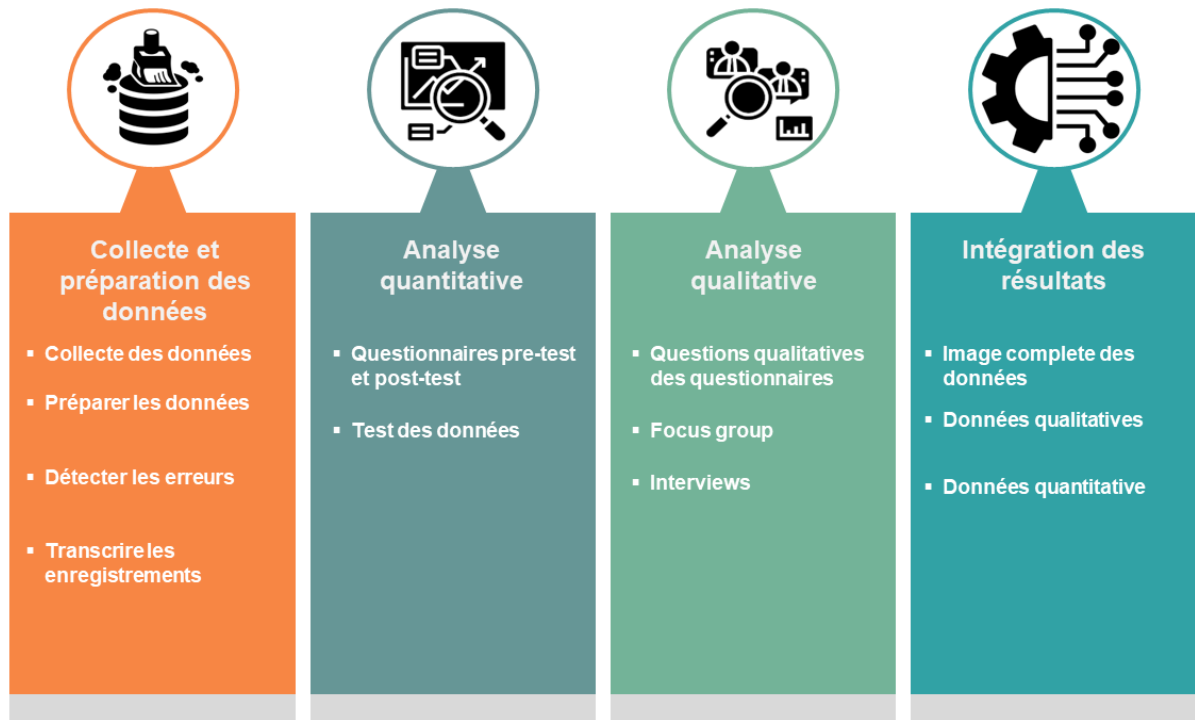


FIGURE 12 – Schéma de l'analyse des données

## 4 Résultats

### 4.1 Présentation des résultats

Au cours de cette étude, nous avons examiné les résultats d'une activité d'enseignement sur la visualisation de données réalisée auprès d'un échantillon de 14 étudiants en promotion sociale. Cet échantillon était composé de 5 étudiants de première année, 2 techniciens en informatique et 7 étudiants de troisième année inscrits dans les cursus de "Bachelier en informatique de gestion" et "Technicien en informatique". L'activité a été conçue dans le but d'introduire les étudiants aux concepts de base de la visualisation de données et de leur fournir une expérience pratique dans l'utilisation d'outils de visualisation.

L'activité d'enseignement a été évaluée à l'aide d'une variété de méthodes, y compris l'observation participante, les questionnaires de pré- et post-test, une discussion de focus group et des interviews. Nous avons utilisé une grille d'évaluation pour évaluer plusieurs aspects de l'activité, y compris la participation des étudiants, leur compréhension des concepts de visualisation de données, leur application des compétences en visualisation de données, et leurs réactions à l'activité. Dans la section suivante, nous présentons une analyse détaillée des résultats de l'activité d'enseignement. Nous discutons de la manière dont les étudiants ont réagi à l'activité, de leur progrès dans l'apprentissage des compétences en visualisation de données, et de l'effet de l'activité sur leur intérêt pour la visualisation de données.

Dans le cadre de notre évaluation globale de l'enseignement en visualisation de données, un troisième questionnaire a été administré à un groupe d'étudiants non participants, mais appartenant au même cursus que ceux qui ont participé à l'activité. L'objectif de ce questionnaire était d'évaluer les connaissances et compétences de base en visualisation de données de ces étudiants, ainsi que leur perception de l'utilité de ces compétences, sans l'influence de l'activité pédagogique spécifique que nous avons menée.

Ce troisième questionnaire a permis de constituer un groupe de comparaison comprenant un nombre limité de répondants (7), ce qui restreint la portée de notre analyse comparative. Cependant, malgré la taille réduite de l'échantillon, il nous permet néanmoins d'obtenir des informations complémentaires pour évaluer l'impact de notre intervention. En comparant les résultats de ce questionnaire avec ceux des étudiants ayant participé à l'activité, nous pouvons obtenir une indication plus précise des améliorations apportées par notre enseignement. Les réponses fournies par ce groupe de non-participants peuvent également nous éclairer sur l'état général des connaissances et des attitudes à l'égard de la visualisation de données parmi les étudiants du même cursus. Bien que les résultats obtenus à partir de cet échantillon réduit doivent être interprétés avec prudence, ils enrichissent néanmoins notre compréhension de l'efficacité de l'activité d'enseignement en visualisation de données.

### 4.2 Analyse de l'observation participante

L'observation participante fournit des informations précieuses sur la manière dont les étudiants interagissent avec les concepts de visualisation de données et les activités pratiques proposées. Il est possible de constater que les étudiants sont généralement engagés et intéressés par les activités, et qu'ils sont capables de mettre en pratique les concepts appris pendant la partie théorique. Les étudiants semblent également apprécier l'approche pédagogique adoptée, qui mêle théorie et pratique (voir figure 13).



FIGURE 13 – Les étudiants pendant l'observation participante

Dans cette section, nous présentons les résultats de notre étude, fondée sur une méthodologie d'observation participante structurée en trois phases distinctes : Avant, Pendant et Après l'activité. Cette approche a été conçue pour offrir une vue holistique de l'activité de visualisation de données et de son impact sur les étudiants. En tout, 14 étudiants en promotion sociale ont participé à cette activité, permettant de recueillir une riche collection de données à analyser.

La première phase, "Avant", a permis de préparer et d'introduire l'activité, tout en recueillant des données de base sur les participants à travers un questionnaire pré-test. La deuxième phase, "Pendant", était le cœur de l'activité d'enseignement, avec une présentation théorique, une introduction à la visualisation de données, une mise en situation avec la base de données de Tintin, et un exercice pratique à réaliser par les étudiants. Enfin, la phase "Après" a consisté en un questionnaire post-test, une discussion ouverte sur l'activité, une explication des objectifs de l'étude et des remerciements.

Nous avons analysé les résultats obtenus à chaque étape, en portant une attention particulière à l'évolution de la compréhension et de l'application des concepts de visualisation de données par les étudiants. Nous avons également pris en compte les réactions et les commentaires des étudiants concernant l'activité.

Dans les sections suivantes, nous détaillerons les résultats obtenus pour chaque phase, en discutant des implications pour l'enseignement de la visualisation de données et en identifiant les points forts et les points à améliorer pour les futures activités d'enseignement.

### Phase 1 - Avant)

La première phase, intitulée "Avant", est destinée à introduire l'activité et à recueillir des données de base. Cette phase comprend plusieurs étapes, notamment :

1. Préparation de l'enregistrement vidéo qui servira pour la complétion de la grille d'évaluation de l'activité.
2. La préparation de l'activité, comprenant le matériel servant à l'enregistrement vidéo de l'activité (trépied et caméra), l'utilisation d'un PowerPoint comme fil conducteur<sup>9</sup>, la vérification du matériel de projection et la préparation d'une farde plastifiée par participant comprenant les questionnaires pre et post test, un stylo bille et l'exercice pour le travail pratique avec des exemples de visualisation.
3. Le chercheur répartit les étudiants dans la salle de classe afin de faciliter la gestion de cette dernière (voir figure 14). Un plan est réalisé avec les personnes présentes.

9. IHDCM050-Philippe-HINANT-Observation.pptx

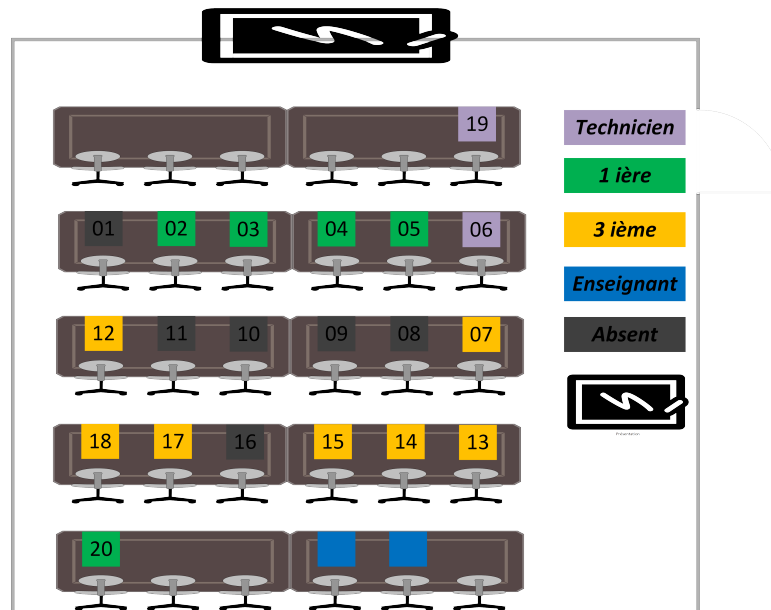


FIGURE 14 – Plan de la classe pendant l'observation

4. L'accueil des étudiants et l'explication de l'activité aux étudiants, avec une présentation du chercheur, une présentation succincte de l'objet de l'étude, l'explication du déroulement de l'activité et le contenu de la farde plastifiée.
5. Les étudiants sont invités à compléter l'autorisation aux droits à l'image.
6. Le pre-test, avec la distribution et la collecte des questionnaires pré-activité, destiné à mesurer les connaissances des étudiants avant l'activité.

### Pre-test

La plupart des étudiants (85,7%) déclarent posséder une expérience préalable en visualisation de données. En ce qui concerne la représentation de la distribution des notes obtenues par les étudiants, le graphique en aires s'avère être le choix prédominant (35,7%), suivi par le graphique linéaire (21,4%). Le graphique en anneau a été sélectionné par 14,3% des participants, tandis que le nuage de points et le graphique à barres ont recueilli 7,1% des préférences chacun. Un segment de 14,3% des étudiants n'a pas opté pour une proposition spécifique.

La pertinence de la visualisation pour répondre à la question a été le principal facteur motivant le choix de la majorité des étudiants (81,25%). Cependant, d'autres considérations comme la clarté et la simplicité de la visualisation (12,5%) ainsi que son attrait esthétique (6,25%) ont également joué un rôle. En terme de visualisation optimale pour les besoins du service des ressources humaines, le graphique en aires arrive une nouvelle fois en tête (35,7%), suivi par le graphique linéaire (21,4%). Le nuage de points et le graphique à barres ont été sélectionnés par 14,3% des étudiants respectivement, tandis que 14,3% n'ont sélectionné aucune option.

Face à la question de la réalisation autonome de certaines visualisations, la majorité des étudiants s'estime capable de les produire sans assistance. Cependant, certains étudiants ont exprimé des difficultés de compréhension de certaines visualisations ou un manque de connaissance sur les outils pour les réaliser. En réponse à la question sur la possibilité de représenter l'intégralité des données dans une visualisation, une majorité des étudiants (64,3%) a répondu de manière erronée, suggérant une méconnaissance des principes de la visualisation de données.

Finalement, en ce qui concerne la compréhension des étapes de construction d'une visualisation de données, les réponses ont été variées. Une moitié des étudiants a fourni une réponse correcte ou partiellement correcte (50%), tandis que d'autres ont répondu de manière incorrecte (30%) ou vague (20%). Ces résultats suggèrent une variabilité dans la compréhension du processus de construction de la visualisation et indiquent le besoin potentiel d'une formation plus approfondie.

## Phase 2 - Pendant

La deuxième phase, intitulée "Pendant", est la phase principale de l'activité d'enseignement et se divise en 4 parties :

1. La théorie de base de la visualisation de données, qui se concentre sur les principes de bases, les différentes façons de nommer la visualisation de données, la définition de la visualisation de données, une définition officielle tirée du livre de Tamara Munzner "«Les systèmes de visualisation informatisés fournissent des représentations visuelles d'ensembles de données conçues pour aider les gens à effectuer des tâches plus efficacement. La visualisation convient lorsqu'il est nécessaire d'augmenter les capacités humaines plutôt que de remplacer les personnes par des méthodes informatiques de prise de décision (voir figure 15) <sup>10</sup>." [45] et de Charles Miglietti, un exemple concret de visualisation, l'âge de cette science, l'exemple de la carte réalisée par John Snow en 1854 (voir figure 16) <sup>11</sup>, la représentation de ces mêmes données avec les techniques actuelles via des graphiques à barres et géographiques, la visualisation d'une interface graphique d'une application web et le cycle de vie d'une visualisation de données.



FIGURE 15 – Exemple simple de l'utilité d'une visualisation de données

2. L'introduction à la visualisation de données, qui se base sur le livre de Munzner Tamara [45] (voir figure 17) et aborde les données, la compréhension des données, la signification des données, les différents types de données, les types d'ensemble de données et les types d'attributs de données.

10. <https://www.toucantoco.com/>

11. <https://www.matter.vc/data-re-imagined/>

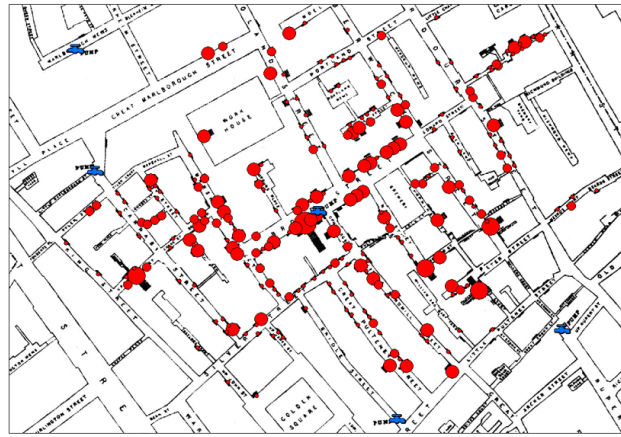


FIGURE 16 – Carte réalisée par John Snow en 1854 relevant la propagation du cholera dans le quartier de Soho à Londres en 1854

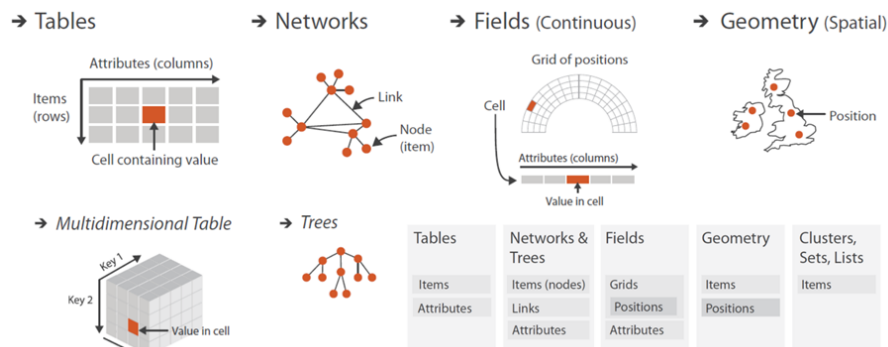


FIGURE 17 – Abstraction de données - Type d'ensemble de données

3. La mise en situation prend comme départ une base de données fournie par l'enseignant. Cette base de données contient les jurons utilisés par différents personnages dans les albums de Tintin, avec un descriptif des personnages (nom, prénom, fonction, etc.) ainsi que la liste des jurons par album et par personnage.

Pour mettre en parallèle les points abordés dans la partie précédente, l'insertion d'erreurs commence par le schéma de la base de données, et ce afin de susciter chez les étudiants une participation proactive durant tout le restant de l'activité. Le contexte des mises en situation reprend à chaque fois le même canevas, à savoir :

- Le demandeur, représenté par le terme "SIJA" (La société internationale des jurons anonymes), qui souhaite savoir un certain nombre de réponses, notamment sur l'utilisation des jurons dans les albums de Tintin tel que : "Bachi-bouzouk, Moule à gaufres, Tonnerre de Brest ou Hurluberlu" (voir figure 18)



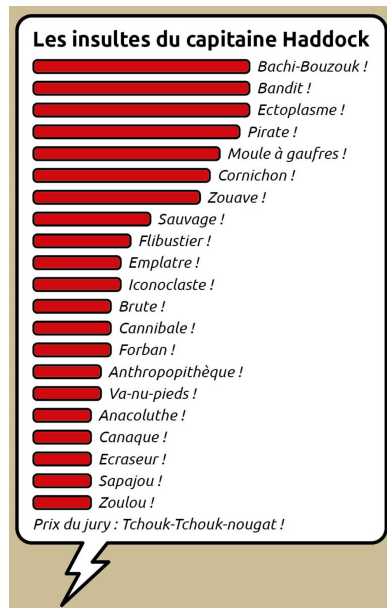


FIGURE 18 – Jurons du capitaine Haddock recensés par la SIJA

- Une représentation du cycle de vie d'une demande (What ?, Why, How ?) (voir figure 19).

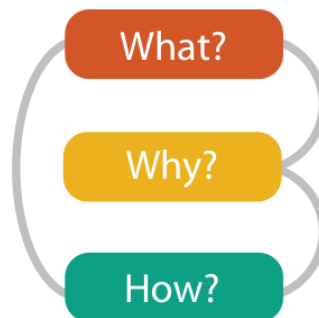


FIGURE 19 – Cycle de vie d'une visualisation de données

- Un énoncé, suivi d'une discussion avec les étudiants.
- Une proposition de requête SQL (avec ou sans insertion d'erreur), suivi d'une discussion avec les étudiants.
- Une proposition de jeux de données, suivi d'une discussion avec les étudiants.
- Une discussion avec les étudiants sur le How ?, suivi d'une proposition de visualisation (avec ou sans insertion d'erreur, voir figure 20).

## Combien y a-t-il de personnages gentils et de méchants?

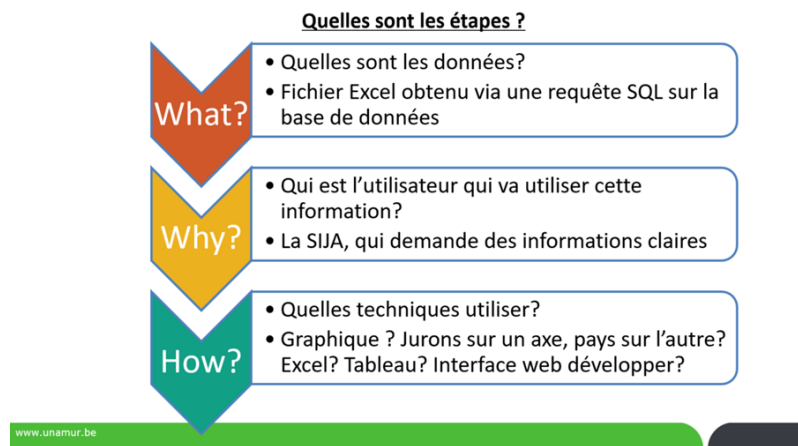


FIGURE 20 – Exemple d'une mise en situation

4. La quatrième partie, "Travaux pratiques", consiste en un exercice pratique où les étudiants doivent représenter le cycle de vie d'une visualisation de données et proposer une visualisation à partir d'un énoncé contenant des erreurs. Des exemples sont fournis aux étudiants pour les aider dans cette tâche (voir tableau 3 et figure ??).

| Titre de l'album              | ID Album | Nombre de jurons | Date de parution |
|-------------------------------|----------|------------------|------------------|
| LE CRABE AUX PINCES D'OR      | A08      | 53               | 1940             |
| L'ETOILE MYSTERIEUSE          | A09      | 57               | 1947             |
| LE SECRET DE LA LICORNE       | A10      | 50               | 1947             |
| LE TRESOR DE RACKHAM LE ROUGE | A11      | 53               | 1947             |
| LES 7 BOULES DE CRISTAL       | A12      | 61               | 1948             |
| LE TEMPLE DU SOLEIL           | A13      | 99               | 1949             |
| AU PAYS DE L'OR NOIR          | A14      | 16               | 1971             |
| OBJECTIF LUNE                 | A15      | 78               | 1953             |
| ON A MARCHE SUR LA LUNE       | A16      | 90               | 1954             |
| L'AFFAIRE TOURNESOL           | A17      | 101              | 1956             |
| COKE EN STOCK                 | A18      | 91               | 1958             |
| TINTIN AU TIBET               | A19      | 61               | 1960             |
| LES BIJOUX DE LA CASTAFIORE   | A20      | 32               | 1963             |
| VOL 714 POUR SIDNEY           | A21      | 34               | 1968             |
| TINTIN ET LES PICAROS         | A22      | 34               | 1976             |

TABLE 3 – Table de données utilisée pour la partie pratique de l'activité

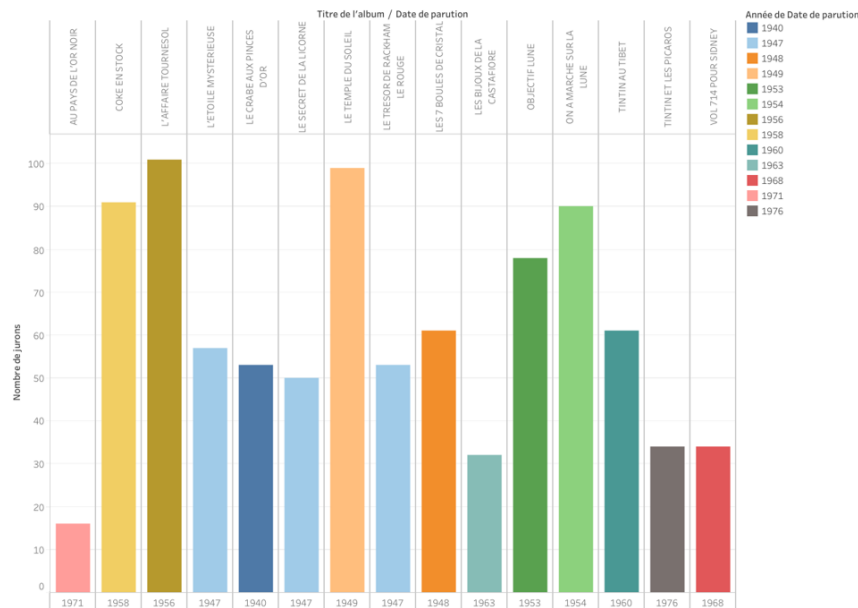


FIGURE 21 – Proposition de solution pour la partie pratique de l'activité

### Phase 3 - Après (post-test)

La phase post-test est divisée en 5 parties :

1. Remplissage du questionnaire post-test
2. Discussion sur l'activité
3. Explication des objectifs de l'étude et des démarches entreprises
4. Remerciements
5. Grille d'évaluation de l'activité est remplie par le chercheur après l'activité sur base des enregistrements vidéos pour évaluer certains critères sur des éléments observables et suivant une échelle d'évaluation.

Après la fin de l'activité, les étudiants remplissent le questionnaire post-test. Le questionnaire contient des questions similaires à celles du questionnaire pré-test, mais cette fois-ci les questions portent sur la compréhension de la visualisation de données, l'utilisation de différentes techniques de visualisation de données, etc. Une discussion ouverte est ensuite menée avec les étudiants pour recueillir leurs opinions sur l'activité et leurs expériences personnelles. Les étudiants sont encouragés à fournir des commentaires constructifs pour aider à améliorer les futures activités d'enseignement. Les discussions portent sur les points forts et les faiblesses de l'activité, les parties les plus utiles, les parties les plus difficiles, etc.

Le chercheur explique les objectifs de l'étude et les démarches entreprises pour recueillir des données. Il explique que l'étude vise à évaluer l'efficacité de l'activité d'enseignement en visualisation de données et que les données recueillies seront utilisées à des fins de recherche. Il explique également comment les données ont été collectées, comment elles seront traitées et comment les résultats seront analysés. Le chercheur remercie les étudiants de leur participation et leur implication à l'activité et leur rappelle que leurs commentaires et leurs contributions seront précieux pour améliorer les futures activités d'enseignement.

Il est intéressant de noter que plusieurs étudiants ont mentionné qu'ils avaient besoin de plus de temps pour bien comprendre les données et choisir la meilleure visualisation, ce qui indique que ces compétences nécessitent peut-être davantage de pratique et de formation. En général, les réponses montrent que les étudiants ont fait face à des difficultés spécifiques lors de l'activité de visualisation de données, ce qui souligne l'importance de fournir une formation adéquate pour améliorer les compétences en visualisation de données. Ces résultats peuvent être

utiles pour l'amélioration de l'activité et la mise en place de mesures pour aider les étudiants à mieux comprendre et utiliser les visualisations de données.

### **Activité d'enseignement**

L'analyse de l'activité d'enseignement de la visualisation de données a été réalisée en utilisant la grille d'évaluation et les données recueillies lors de l'observation, des questionnaires post-test, du focus group et de l'interview. Les résultats de l'analyse sont présentés ci-dessous :

**Participation des étudiants :** Les étudiants ont activement participé à l'activité. Ils ont posé des questions, discuté entre eux et utilisé les outils de visualisation de données. La participation active des étudiants indique un haut niveau d'engagement et d'implication dans l'activité.

**Compréhension des concepts de visualisation de données :** Les étudiants ont démontré une bonne compréhension des concepts clés de la visualisation de données. Ils ont été capables d'interpréter les visualisations de données et de comprendre leur utilité et leur application dans le domaine de la data science.

**Application des compétences de visualisation de données :** Les étudiants ont été capables d'appliquer leurs compétences en visualisation de données pour analyser et interpréter les données. Ils ont utilisé efficacement les outils de visualisation de données pour représenter les données.

**Réactions des étudiants à l'activité :** Les étudiants ont réagi positivement à l'activité. Ils ont apprécié l'approche d'apprentissage par les erreurs et ont trouvé l'activité intéressante et utile. Il y a eu quelques moments de confusion, mais ces moments ont été résolus grâce à des discussions et des clarifications.

**Progrès des étudiants :** Les étudiants ont montré des progrès significatifs dans leurs compétences en visualisation de données. Les scores aux tests de pré- à post-test ont montré une amélioration significative, indiquant un apprentissage et un développement réussis.

**Effets sur l'intérêt pour la visualisation de données :** L'activité a suscité un intérêt accru pour la visualisation de données parmi les étudiants. Les commentaires et les réactions des étudiants indiquent une appréciation de l'activité et un intérêt pour des activités d'apprentissage similaires à l'avenir.

**Évaluation globale de l'activité d'enseignement :** Dans l'ensemble, l'activité d'enseignement a été un succès. Les étudiants ont démontré une amélioration des compétences en visualisation de données, une compréhension accrue de son importance et un intérêt accru pour le sujet. Les commentaires des étudiants ont été positifs, et ils ont exprimé leur appréciation pour l'activité et leur intérêt pour des activités similaires à l'avenir.

En conclusion, l'activité d'enseignement a atteint ses objectifs en améliorant les compétences en visualisation de données des étudiants, en augmentant leur compréhension de l'importance de la visualisation de données et en suscitant un intérêt pour le sujet.

### **Post-test**

Les étudiants ont montré une meilleure compréhension des principes fondamentaux de la visualisation de données. En effet, 92,9% des étudiants ont répondu correctement à la question "Une visualisation doit toujours afficher l'entièreté des données disponibles dans les tables de données?" et 85,7% ont répondu correctement à la question "Une visualisation de données est un processus nécessitant plusieurs étapes de construction?".

Les étudiants ont également montré une amélioration de leur capacité à réaliser des visualisations de données. Les pourcentages d'étudiants qui ont répondu "Oui, je sais la réaliser seul" ont augmenté pour la plupart des visualisations présentées lors du post-test, tandis que les pourcentages de ceux qui ont répondu "Non, je ne connais pas d'outils permettant de réaliser une telle visualisation" ont diminué pour la plupart des visualisations. Les étudiants ont identifié des difficultés rencontrées pendant l'activité, notamment la compréhension des données à partir d'un énoncé et la construction de la visualisation elle-même.

En conclusion, les résultats du post-test suggèrent que la formation a été efficace pour transmettre les connaissances sur les principes fondamentaux de la visualisation de données et pour améliorer les compétences des étudiants dans la réalisation de visualisations de données. Cependant, il reste encore des difficultés à surmonter pour les étudiants, notamment en ce qui concerne la compréhension des données à partir d'un énoncé et la construction de la visualisation elle-même.

### **Analyse de l'interview semi-directive avec l'enseignant**

L'interview semi-directive avec l'enseignant a révélé plusieurs points importants sur l'intégration de la visualisation de données dans son cours de modélisation de données.

#### **Avant l'activité :**

1. L'enseignant avait déjà utilisé la visualisation de données dans un contexte professionnel, mais ne percevait pas sa pertinence dans le cadre de son cours de bases de données.
2. Les étudiants manifestaient peu d'intérêt pour la visualisation de données, probablement due à une exposition limitée à ce concept dans leur cursus.

#### **Pendant l'activité :**

1. L'enseignant a apprécié l'engagement actif des étudiants et la structure de l'activité, notamment la possibilité d'évaluer l'évolution des étudiants grâce aux questionnaires avant et après l'activité.
2. Aucune difficulté majeure n'a été observée, ce qui montre la faisabilité de l'intégration de la visualisation de données dans l'enseignement.

#### **Après l'activité :**

1. Les étudiants ont manifesté un nouvel intérêt pour la visualisation de données, discutant des liens entre les données brutes et leur intégration dans des interfaces, ainsi que de la nécessité d'une réflexion approfondie pour une intégration efficace.
2. L'enseignant a vu une réelle valeur ajoutée à la visualisation de données dans son enseignement. Il envisage de l'intégrer dans son cours l'année suivante en utilisant des exemples concrets à partir de fichiers Excel ou CSV et en organisant une session dédiée à l'enseignement de la visualisation de données.
3. L'enseignant a également souligné l'importance d'intégrer les concepts de visualisation de données plus tôt dans le cursus, et pas seulement dans les cours de troisième année. Il suggère de cibler les objectifs de cours du module d'enseignement cible et d'utiliser des outils facilitant la conception, tels que HTML, JavaScript ou un logiciel tiers.

L'enthousiasme exprimé par l'enseignant suite à l'activité est très révélateur de l'impact positif que la visualisation de données peut avoir sur l'enseignement. Cette réponse enthousiaste ne se limite pas à une appréciation subjective de l'activité, mais se traduit aussi par une demande concrète de la part de l'enseignant : obtenir les outils utilisés pendant l'activité pour pouvoir les intégrer en tant que support à part entière de son cours dans les années à venir. Cette requête met en évidence une réelle volonté d'approfondir l'utilisation

de la visualisation de données dans son enseignement et d'en faire un pilier de son cours de modélisation de données.

De plus, l'intérêt manifesté par l'école pour poursuivre cette étude scientifique témoigne d'une reconnaissance institutionnelle de la valeur de la visualisation de données dans l'enseignement. Cette volonté de continuer la recherche dans ce domaine pourrait ouvrir la voie à des développements plus larges et plus systématiques de l'utilisation de la visualisation de données dans le curriculum, avec des bénéfices potentiels significatifs pour l'apprentissage des étudiants. Cette perspective est enthousiasmante et constitue un fort encouragement à continuer à explorer et à développer l'utilisation de la visualisation de données dans l'enseignement.

En conclusion, cette interview a démontré que l'enseignant a pris conscience de l'importance de la visualisation de données dans son enseignement et a identifié des moyens concrets de l'intégrer dans son cours. De plus, les étudiants ont montré un intérêt accru pour la visualisation de données suite à l'activité. Ces résultats suggèrent que l'intégration de la visualisation de données dans l'enseignement peut avoir des effets bénéfiques sur l'engagement des étudiants et sur leur compréhension des concepts abordés.

### 4.3 Analyse des questionnaires pré- et post-test

L'analyse des réponses aux questionnaires pré- et post-test permet d'évaluer l'évolution des connaissances et des compétences des étudiants en visualisation de données. Les résultats montrent une amélioration significative des connaissances des étudiants en matière de visualisation de données et de modélisation de bases de données après l'intervention. Cela suggère que l'introduction de la visualisation de données dans le cadre de l'enseignement en promotion sociale est bénéfique pour les étudiants novices en data science. Après avoir consulté les résultats du pre-test et du post-test, il semble y avoir une amélioration significative de la compréhension des concepts de visualisation de données chez les étudiants. En effet, le score moyen au post-test (17,8/20) est nettement supérieur au score moyen au pre-test (8,2/20).

De plus, la majorité des étudiants ont obtenu une note supérieure ou égale à 15/20 au post-test, ce qui montre une compréhension satisfaisante des concepts présentés. Il est important de noter que même les étudiants ayant obtenu des scores inférieurs à 15/20 ont obtenu des scores supérieurs au pre-test, montrant ainsi une amélioration de leur compréhension. En conclusion, l'activité semble avoir été efficace pour améliorer la compréhension des concepts de visualisation de données chez les étudiants.

Dans la phase de recherche où les questionnaires pré- et post-tests ont été administrés, il a été constaté que les étudiants ont mis environ 35 minutes pour compléter le pré-test, tandis que le post-test a été complété en environ 15 minutes. Cette réduction substantielle du temps nécessaire pour achever le test indique non seulement une familiarité accrue avec le matériel et le format du test, mais aussi une augmentation potentielle de la compétence et de la compréhension des sujets couverts.

Comme Margolis (2002)[85] l'a noté, l'acquisition de compétences en informatique, y compris la visualisation de données, est souvent un processus itératif, où la compréhension et la compétence augmentent à chaque cycle. Il est donc raisonnable de s'attendre à ce que les étudiants aient besoin de moins de temps pour répondre aux questions après avoir été exposés au matériel de cours.

Il est également intéressant de noter que cette tendance à la réduction du temps nécessaire pour répondre aux questions de test est cohérente avec les travaux de Noss et Hoyles (1996)[92]. Dans leur étude sur les cultures d'apprentissage et les ordinateurs, ils ont observé que, à mesure que les étudiants devenaient plus familiers avec les concepts et les outils, ils étaient capables de travailler de manière plus efficace.

Ces résultats sont en accord avec le constat dans notre étude que les étudiants étaient capables de compléter le post-test plus rapidement que le pré-test, ce qui suggère une familiarité accrue et une meilleure compréhension des concepts. Ces résultats renforcent l'idée que la pratique répétée et l'exposition à des concepts dans différents contextes peuvent être des éléments clés dans l'enseignement des compétences en informatique et en visualisation de données.

## Démographie

La synthèse des données démographiques du pré-test montre un échantillon diversifié d'étudiants, répartis entre différentes catégories d'âge, de niveaux d'éducation et de situations professionnelles, ce qui est typique dans un contexte d'enseignement de promotion sociale.

Premièrement, concernant la répartition par âge, fait de la majorité des étudiants (42,9%) ont entre 25 et 30 ans. Cela indique que la formation est principalement suivie par des adultes relativement jeunes qui cherchent probablement à améliorer leurs compétences ou à se reconvertir professionnellement. Il est également intéressant de noter que la formation attire aussi des personnes plus âgées, avec des étudiants dans la tranche d'âge 40-50 ans.

Deuxièmement, la situation professionnelle des étudiants est majoritairement marquée par le chômage (57,1%), ce qui suggère que la formation pourrait être un moyen pour eux d'améliorer leur employabilité. Cependant, il y a aussi des employés et des fonctionnaires parmi les étudiants, ce qui montre que la formation peut également intéresser ceux qui cherchent à acquérir de nouvelles compétences tout en étant déjà en emploi.

Troisièmement, en ce qui concerne le niveau d'éducation, la majorité des étudiants ont un niveau d'éducation secondaire supérieur. Cela montre que la formation est accessible à ceux qui n'ont pas nécessairement suivi un parcours d'éducation supérieure traditionnel, ce qui est l'un des objectifs de l'enseignement de promotion sociale. Enfin, il y a une prédominance masculine parmi les étudiants (85,7%), ce qui peut refléter une tendance générale dans le domaine de l'informatique. Cependant, la présence de participantes (14,3%) montre également que la formation est ouverte et accessible aux femmes.

Dans l'ensemble, ces résultats suggèrent que la formation en informatique dans le contexte de l'enseignement de promotion sociale attire un large éventail de personnes, avec des profils variés en termes d'âge, de situation professionnelle et de niveau d'éducation (voir figure 22).

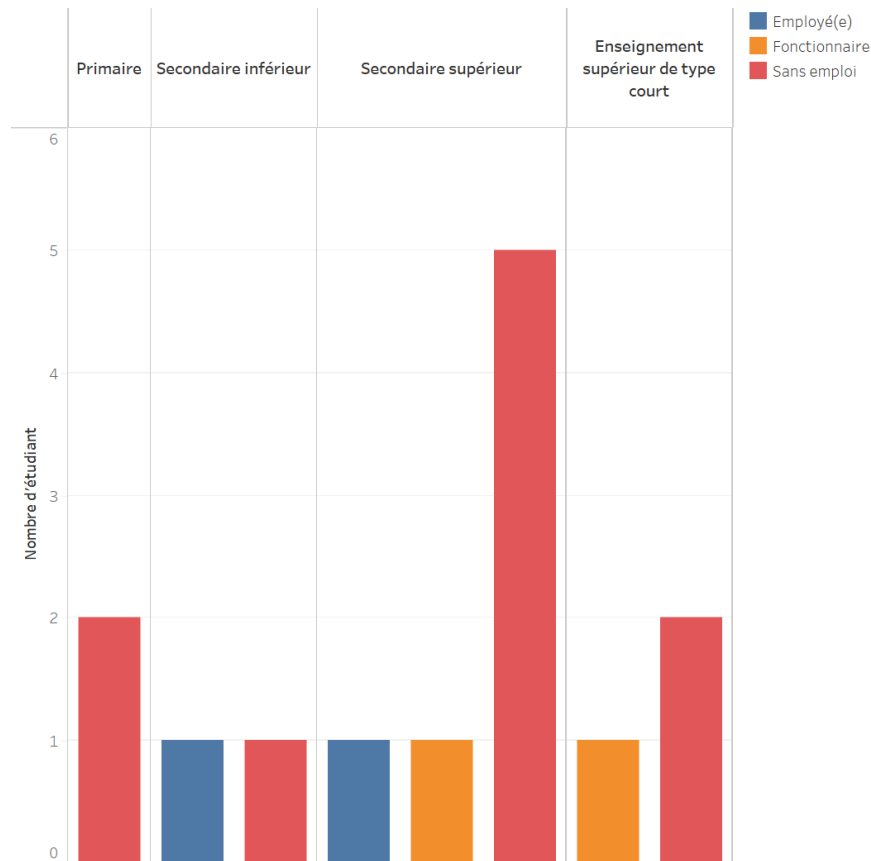


FIGURE 22 – Graphique représentant la démographie

### Compréhension de la modélisation de données

L'activité d'enseignement semble avoir eu un impact positif sur la compréhension de la modélisation de données par les étudiants. Cette conclusion est appuyée par une amélioration notable des scores dans le post-test par rapport au pré-test.

Dans le pré-test, les réponses correctes étaient assez dispersées. Cependant, après la formation, les étudiants ont montré une amélioration significative de leurs compétences en matière de modélisation de données. En effet, 13 sur 14 ont réussi à répondre correctement à toutes les questions du post-test (voir figure 23). Plus précisément, les étudiants de la 1<sup>ère</sup> année et ceux du programme de Technicien en informatique ont montré une progression importante, passant de 50% à 100% de réponses correctes. Les étudiants de la 3<sup>e</sup> année ont également progressé, passant de 66,7% à 100% de réponses correctes.

Il est important de noter qu'un participant (appartenant à la catégorie 1<sup>ère</sup> année) n'a pas montré d'amélioration dans le post-test. Cela pourrait suggérer la nécessité de méthodes pédagogiques supplémentaires ou différentes pour certains étudiants. Dans l'ensemble, ces résultats suggèrent que la formation a été efficace pour améliorer la compréhension de la modélisation de données par les étudiants. Cependant, il convient d'examiner plus en détail les besoins individuels des étudiants pour s'assurer que tous profitent de la formation.

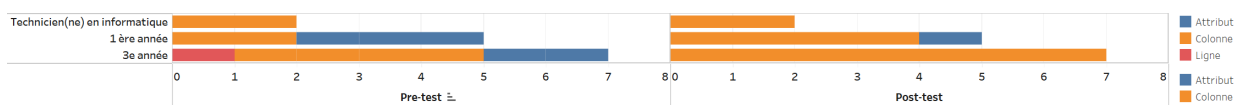


FIGURE 23 – Graphique compréhension de donnée : "Parmi les propositions suivantes, sélectionner le tableau où un attribut est entouré" (Q1\_1) pré-test / post-test (colonne est la bonne réponse)



## Compréhension de la visualisation de données

Concernant la question Q02 "La visualisation de données c'est (expliquer)" :

Dans le pré-test, il est apparent que la plupart des étudiants ont offert une définition assez générale de la visualisation de données. La visualisation de données était souvent envisagée comme une méthode pour faciliter la lecture et la compréhension des informations (5 occurrences de "compréhension"), présenter les données de manière compréhensible (11 occurrences de "données"), ou simplement afficher les données à l'écran (7 occurrences de "affichage"). On peut également noter l'importance de la représentation et de l'analyse des données avec respectivement 2 et 2 occurrences. Deux étudiants n'ont pas fourni de réponse.

Cependant, dans le post-test, il y a eu une amélioration notable en termes de précision et de pertinence des réponses. Les étudiants ont davantage souligné la transformation des données brutes (4 occurrences de "brut") en informations compréhensibles (8 occurrences de "comprendre"), l'importance de l'audience cible (4 occurrences de "public" et 4 occurrences de "cible"), et le but de rendre les données plus lisibles (6 occurrences de "lisible") et utilisables (5 occurrences de "utilisable"). Notons qu'un étudiant n'a pas fourni de réponse.

Pour la question Q03 "Est-ce qu'une table (d'une base de données) est considérée comme une visualisation de données?" (voir figure 24) :

Dans le pré-test, la majorité des étudiants ont répondu par l'affirmative, tandis que deux ont répondu par la négative et un participant n'était pas sûr. Dans le post-test, tous les étudiants ont répondu "Non", ce qui montre une amélioration significative de la compréhension de ce qu'est une visualisation de données.

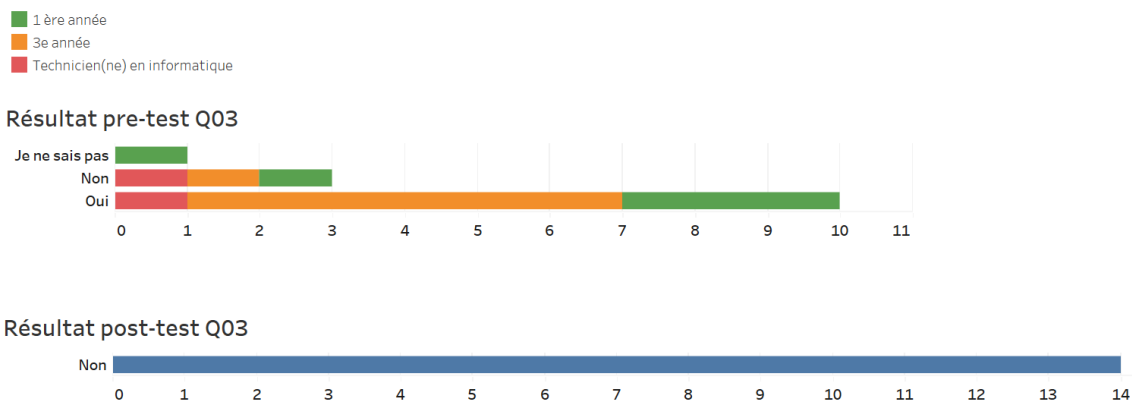


FIGURE 24 – Est-ce qu'une table (d'une base de données) est considérée comme une visualisation de données?(non est la bonne réponse)

Pour la question Q04 "Est-ce qu'un tableau Excel est considéré comme une visualisation de données?" :

Dans le pré-test, la majorité des étudiants ont répondu "Oui" (85,7%), en supposant que le tableau Excel était formaté pour la visualisation. Un participant a répondu "Non" (7,1%), en argumentant que si les données n'étaient pas transformées, alors le tableau Excel ne pouvait pas être considéré comme une visualisation de données. Un autre participant n'était pas sûr.

Dans le post-test, la majorité des étudiants ont toujours répondu "Oui" (71,34%), en maintenant l'hypothèse que le tableau Excel était formaté pour la visualisation. Toutefois, le nombre de réponses "Non" a augmenté(de 7,1% à 28,6%), avec des étudiants précisant que sans transformation des données, un tableau Excel ne peut être considéré comme une visualisation de données.

Cette nuance dans les réponses indique une compréhension plus profonde des étudiants sur ce qu'est une visualisation de données. Ils reconnaissent que la simple présentation de données dans un tableau Excel ne suffit pas à en faire une visualisation de données, sauf si ces données ont été transformées ou formatées dans un but de visualisation. Cela démontre une évolution significative dans la compréhension des étudiants entre le pré-test et le post-test.

### **Utilité de la visualisation de données**

Dans le cadre du pre-test, les étudiants à travers toutes les années ont perçu l'utilité de la visualisation de données principalement pour améliorer la lisibilité et la compréhension des données. Ils ont généralement convenu que tout le monde, indépendamment de leur domaine, pourrait bénéficier de la visualisation de données. Pour leur propre utilisation, ils ont envisagé son application dans la gestion de projet, l'informatique et la communication des informations. Dans le contexte de leur carrière future, ils ont vu la visualisation de données comme un outil potentiel pour améliorer leur efficacité, en particulier en termes de communication et de présentation des informations.

Après la formation, les étudiants ont continué à valoriser la lisibilité et la compréhension offertes par la visualisation de données. Ils ont renforcé l'idée que la visualisation de données est utile pour quiconque a besoin de comprendre et d'extraire des informations pertinentes à partir de grands ensembles de données. L'importance de la visualisation de données pour l'analyse rapide des données, la gestion de projet et la communication a également été mise en avant. En termes de carrière, les étudiants ont souligné l'utilité de la visualisation de données pour la prise de décisions éclairées, la présentation claire des informations et l'amélioration de la communication avec les clients et les collègues.

Globalement, les étudiants des différentes années ont montré une compréhension croissante de l'utilité de la visualisation de données du pré-test au post-test. Alors qu'ils ont initialement vu la visualisation de données principalement comme un outil pour améliorer la lisibilité et la compréhension des données, ils ont progressivement reconnu son rôle dans l'analyse rapide des données, la gestion efficace des projets et la communication claire et efficace des informations. Cela montre que la formation a renforcé leur appréciation de la valeur de la visualisation de données pour une variété de tâches et de contextes.

### **Choix de la meilleure visualisation pour un scénario donné**

Au pre-test, pour la Q09, les étudiants ont, pour la majorité, choisi le diagramme géographique, qui est la meilleure option pour visualiser une répartition géographique. Ils ont principalement justifié leur choix par le fait qu'il permet de voir directement où se trouvent les clients. Pour la Q11 (Historique des engagements du personnel) la majorité des répondants ont choisi le graphique à barres et les histogrammes. Les répondants ont souligné l'importance de voir facilement combien de personnes ont été engagées chaque année. A la question Q13 (Répartition du budget), le graphique circulaire a été la réponse la plus fréquente (voir figure 25). Les répondants ont mentionné qu'il permet de voir clairement et directement les répartitions du budget.

Résultat pre-test Q13

Résultat post-test Q13

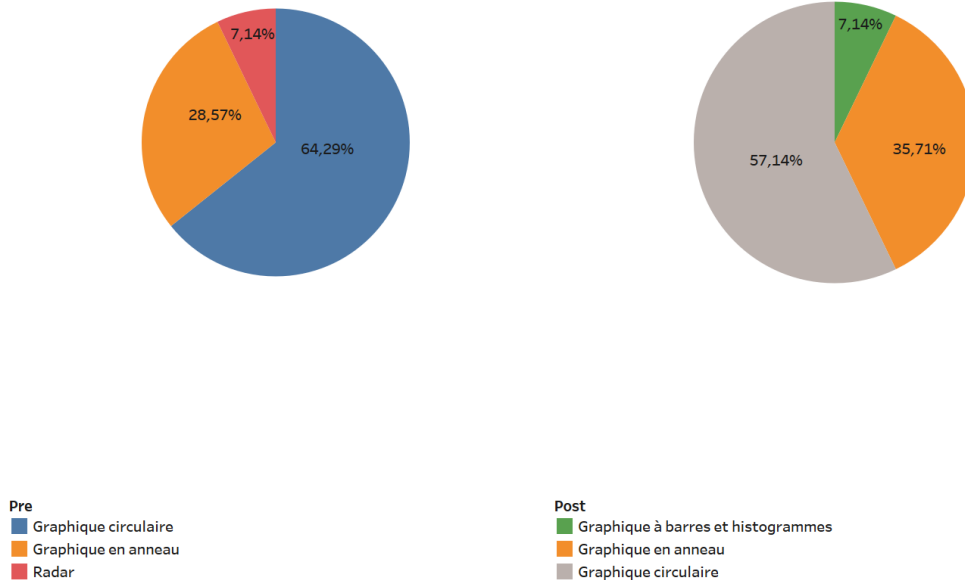


FIGURE 25 – Choix de la meilleur visualisation pour un scénario donné (préférence des visualisations) pre-test et post-test (Q13)

Pour répondre à la Q15 (Compétences des candidats), il y a une grande variété de réponses pour cette question, avec une légère préférence pour le graphique radar et le graphique linéaire. Les répondants ont justifié leur choix par le fait qu'ils permettent de visualiser facilement les différences de compétences entre les candidats. Pour la Q17 (Planning du projet), le diagramme de Gantt a été le choix majoritaire, mentionné comme étant le plus logique pour visualiser un planning de projet. Ensuite pour la Q19 (Projections RH), les réponses sont assez variées, mais le graphique en aires et le graphique en anneau semblent être les plus populaires. Les répondants ont mentionné qu'ils permettent de voir directement les nombres pour chaque catégorie d'employés.

Après l'activité, les réponses ont évolué. Lorsqu'il s'agit de représenter des données géographiques, le diagramme géographique est le type de graphique préféré par les répondants. Ils apprécient sa capacité à visualiser facilement et directement les données demandées. Il facilite la représentation du territoire et la répartition, et offre une visualisation claire du pays et de ses régions. En ce qui concerne l'affichage du nombre d'embauches pour chacune des huit dernières années ou de l'évolution des années, le graphique à barres et histogrammes est le choix privilégié. Les répondants le trouvent logique pour ce type de données, étant donné sa simplicité et son expressivité (voir figure 26).



FIGURE 26 – Choix de la meilleur visualisation pour un scénario donné (préférence des visualisations) pre-test et post-test

Le graphique circulaire, quant à lui, est souvent choisi pour représenter des données de budget ou de répartition. Les répondants ont souligné que ce type de graphique rend les chiffres plus facilement représentables sous formes circulaires, et qu'il affiche clairement les différentes proportions. En matière de comparaison de compétences entre candidats, le graphique radar est jugé comme le plus adapté. Il permet de voir si un candidat se démarque ou a une meilleure performance dans un critère, ce qui est crucial pour l'évaluation.

Le diagramme de Gantt est privilégié pour la gestion de projets. Il permet de visualiser les différentes tâches à accomplir dans le temps, une caractéristique utile pour la planification et le suivi de projets. Il a également été mentionné que ce type de diagramme est enseigné dans les cours d'analyse. Enfin, le graphique linéaire est privilégié pour représenter l'évolution de données dans le temps ou pour afficher trois lignes de données différentes sur le même graphique. Les répondants l'ont également cité pour sa capacité à permettre une représentation prévisionnelle.

Le pré-test a permis de recueillir les réponses initiales des étudiants sur leurs préférences et leurs perceptions concernant l'utilisation de divers types de graphiques pour représenter des données. Les réponses étaient variées, montrant une compréhension diverse des différents types de graphiques. Par exemple, certains étudiants ont préféré les graphiques à barres pour leur simplicité, tandis que d'autres ont favorisé les graphiques linéaires pour leur capacité à montrer l'évolution des données dans le temps.

Le post-test, en revanche, a révélé une meilleure compréhension et une plus grande appréciation des différents types de graphiques. On a constaté une tendance générale à préférer des types de graphiques plus complexes et plus informatifs, comme le diagramme de Gantt et le graphique radar. Cela indique que les étudiants ont acquis une meilleure compréhension de ces types de graphiques et de leurs utilisations potentielles au cours de la

période entre les deux tests. De plus, dans le post-test, les étudiants ont montré une plus grande conscience des avantages et des inconvénients de chaque type de graphique, et étaient plus capables d'expliquer pourquoi ils préféraient certains types de graphiques à d'autres. Par exemple, ils ont noté que le diagramme de Gantt est particulièrement utile pour visualiser la progression d'un projet, tandis que le graphique radar est utile pour comparer plusieurs éléments sur plusieurs critères. En somme, le post-test révèle une meilleure compréhension, une plus grande appréciation et une utilisation plus judicieuse des différents types de graphiques parmi les étudiants, par rapport au pré-test.

### **Compétence en visualisation de données**

Dans le pre-test on constate que dans le domaine de la visualisation des données, les compétences des étudiants étaient assez diverses. En ce qui concerne la réalisation d'un graphique linéaire (Q21), la majorité des étudiants se sentait à l'aise et capable de le faire seuls. Néanmoins, quelques-uns ont exprimé le besoin d'un expert ou l'absence de connaissance des outils appropriés. Pour le diagramme de Gantt (Q22), les réponses étaient plus hétérogènes. Bien que certains étudiants se sentaient confiants pour le créer eux-mêmes, d'autres ont mentionné qu'ils auraient besoin de l'aide d'un expert ou qu'ils ne connaissaient pas les outils nécessaires.

Lorsqu'il s'agissait de créer un graphique à bulles (Q23), beaucoup d'étudiants ont exprimé des difficultés. Certains ont admis ne pas comprendre le sens de ce type de visualisation, tandis que d'autres ont indiqué ne pas connaître les outils pour le réaliser. Seuls quelques étudiants ont affirmé pouvoir le créer seuls.

En ce qui concerne le graphique en forme de donut (Q24), plusieurs étudiants ont déclaré être capables de le faire eux-mêmes. Cependant, un certain nombre d'entre eux ont exprimé le besoin d'un expert ou l'absence de connaissance des outils nécessaires.

Pour la visualisation géographique (Q25), la majorité des étudiants ont exprimé des difficultés. Beaucoup ont indiqué ne pas connaître les outils nécessaires pour réaliser ce type de visualisation, et quelques-uns ont mentionné qu'ils auraient besoin de l'aide d'un expert. Seuls quelques étudiants se sont déclarés capables de le faire eux-mêmes.

Sur la question de savoir si une visualisation doit toujours afficher toutes les données disponibles (Q26), les avis étaient partagés. Certains étudiants ont affirmé que c'était le cas, tandis que d'autres ont déclaré le contraire. Quelques étudiants ont avoué ne pas savoir. Enfin, sur la question du processus de construction d'une visualisation de données (Q27), la plupart des étudiants ont reconnu que cela nécessitait plusieurs étapes, allant de la collecte des données à leur mise en forme pour la visualisation. Néanmoins, certains étudiants n'étaient pas sûrs de ce processus.

Dans le post-test, on observe une augmentation de la confiance des étudiants quant à leurs compétences en visualisation de données. La majorité d'entre eux se sentait à l'aise pour réaliser un graphique linéaire (Q21) et un diagramme de Gantt (Q22) seuls. Cependant, certains ont toujours indiqué avoir besoin d'un expert ou ne pas connaître les outils nécessaires pour la réalisation de ces graphiques. Concernant le graphique à bulles (Q23), malgré le fait que certains étudiants ont maintenant indiqué être capables de le réaliser seuls, de nombreux étudiants ont continué à exprimer des difficultés. Un nombre encore plus important d'étudiants a exprimé le besoin d'un expert ou a admis ne pas connaître les outils nécessaires pour réaliser ce type de visualisation.

En ce qui concerne le graphique en forme de donut (Q24), la plupart des étudiants ont déclaré être capables de le réaliser seuls. Cependant, une portion non négligeable d'entre eux a toujours exprimé le besoin d'un expert ou l'absence de connaissance des outils nécessaires. En ce qui concerne la visualisation géographique (Q25),

malgré le fait que certains étudiants ont maintenant indiqué être capables de la réaliser seuls, la majorité des étudiants ont continué à exprimer des difficultés. Beaucoup ont indiqué ne pas connaître les outils nécessaires pour réaliser ce type de visualisation, et quelques-uns ont mentionné qu'ils auraient besoin de l'aide d'un expert.

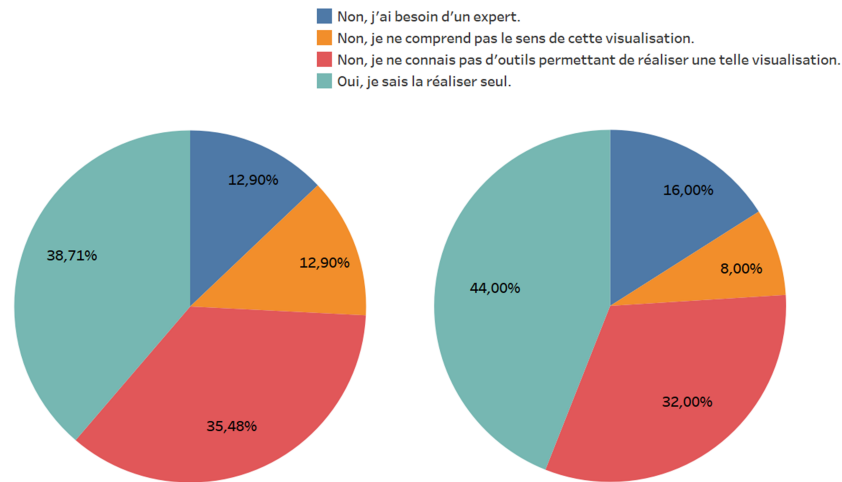


FIGURE 27 – Compétences en visualisation de données (Q21 - Q25) pre-test / post-test

Sur la question de savoir si une visualisation doit toujours afficher toutes les données disponibles (Q26), la majorité des étudiants a répondu par la négative, ce qui indique une meilleure compréhension de la nécessité de sélectionner et de filtrer les données pour la visualisation.

Enfin, sur la question du processus de construction d'une visualisation de données (Q27), la plupart des étudiants ont démontré une bonne compréhension du processus, en mettant en avant les étapes de la collecte des données, de leur mise en forme et de la définition de la visualisation en fonction du public cible et de l'objectif de la visualisation.

### Compréhension de l'activité

Il semble que la construction d'une visualisation à partir d'un énoncé a été le défi le plus couramment cité par les étudiants. La partie théorique et la définition d'une visualisation sont également des domaines dans lesquels plusieurs étudiants ont rencontré des difficultés. Cependant, malgré ces difficultés, la majorité des étudiants se sont sentis assez confiants dans leur compréhension de l'ensemble de données, avec un score moyen de 4,14 sur une échelle de 1 à 5. La majorité des étudiants ont choisi le score 4, indiquant un niveau de confiance assez élevé (voir figure 28).

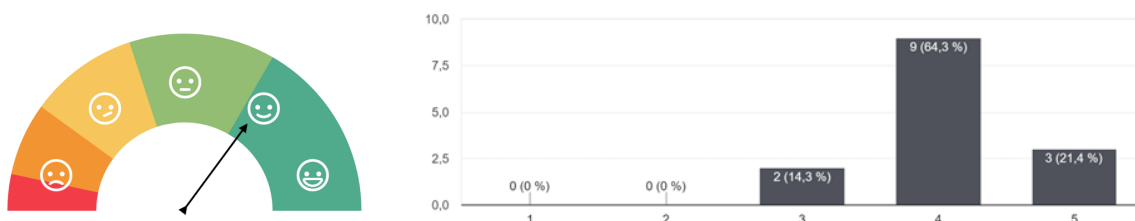


FIGURE 28 – Représentation de la confiance des étudiants en la compréhension de la visualisation

## Résultats du troisième questionnaire

Il est important de noter que la taille de l'échantillon de cette enquête est relativement petite, avec seulement sept répondants au total. Cela limite la généralisation des résultats à une population plus large. Les futures études dans ce domaine pourraient chercher à recueillir des données auprès d'un échantillon plus large pour obtenir des résultats plus robustes et représentatifs.

**Informations générales :** Les informations démographiques recueillies par le troisième questionnaire sont cohérentes avec les précédentes, démontrant une participation continue des mêmes groupes démographiques. Il n'y a pas eu de changements significatifs dans la composition des répondants, ce qui renforce la fiabilité des données. Parmi les sept répondants à ce questionnaire, la majorité étaient des hommes, représentant cinq sur sept participants. L'âge des répondants était principalement concentré dans la tranche d'âge des 25-30 ans, qui comprenait quatre personnes. En ce qui concerne le niveau d'éducation, la plupart des répondants ont déclaré avoir un niveau d'éducation secondaire supérieur, avec deux personnes ayant reçu un enseignement supérieur de type court. La situation professionnelle était assez équilibrée avec trois personnes employées et trois personnes sans emploi. Quant au domaine d'étude ou d'emploi, trois participants travaillaient ou étudiaient dans le domaine de l'informatique, tandis que les autres n'ont pas spécifié leur domaine ou étaient en deuxième ou quatrième année d'études, ou étaient sans emploi.

**Compréhension de la modélisation de données :** Les réponses au troisième questionnaire montrent que les étudiants non participants ont une compréhension variable de la modélisation des données, avec une distinction claire entre les concepts d'attribut, d'item et de cellule. Cependant, cette compréhension semble être moins approfondie que celle observée dans les résultats du pré-test.

**Compréhension de la visualisation de données :** Le troisième questionnaire a révélé que les étudiants non participants ont des points de vue divers sur ce que constitue une visualisation de données, certains reconnaissant qu'une table ou un tableau Excel peut être considérée comme une forme de visualisation. Cette diversité d'opinions reflète une compréhension moins cohérente que celle observée chez les étudiants ayant participé à l'activité.

**Utilité de la visualisation de données :** Les étudiants non participants reconnaissent largement l'utilité de la visualisation de données, même s'ils ne l'ont pas expérimentée dans le cadre de l'activité. Ils voient la visualisation de données comme un outil potentiel pour améliorer leur compréhension des données et leur prise de décision basée sur les données, tant dans leur futur parcours scolaire que professionnel. Ces résultats sont en accord avec ceux observés dans le pré-test.

**Choix de la meilleure visualisation pour un scénario donné :** Les réponses au troisième questionnaire montrent que les étudiants non participants sont moins confiants dans leur capacité à choisir la meilleure visualisation pour un scénario donné. Les réponses révèlent une variété de critères utilisés pour prendre cette décision, suggérant une compréhension moins uniforme de ce que constitue une "meilleure" visualisation.

**Compétences en visualisation de données :** Les étudiants non participants ont montré une variété de niveaux de confiance dans leurs capacités à créer des visualisations de données, allant de la confiance à la réalisation de graphiques simples à l'incertitude face à des visualisations plus complexes. Les réponses à ces questions suggèrent que ces étudiants peuvent bénéficier d'une formation supplémentaire en visualisation de données.

**Comparaison avec le premier pre-test :** Le niveau de compréhension et de maîtrise de la visualisation des données parmi les participants du troisième questionnaire semble similaire à celui du groupe précédent, avec une légère tendance à une meilleure compréhension et à une utilisation plus répandue de la visualisation des données. Cela peut être attribué à leur expérience personnelle ou professionnelle. Toutefois, tout comme dans le premier pre-test, il existe une variété d'opinions sur le type de graphique le mieux adapté à chaque situation, ce qui indique la nécessité d'une formation plus approfondie sur le choix des types de visualisation en fonction des données et des objectifs à atteindre.

Les réponses ont montré que la majorité des étudiants étaient confiants dans leur compréhension de l'ensemble de données après avoir réalisé l'activité de visualisation de données. Cela indique que l'activité a été utile pour renforcer la compréhension des étudiants des données et de leur visualisation.

## 4.4 Analyse du focus group

### Interview post-test

L'interview post-test menée auprès de l'enseignant a permis de recueillir ses impressions et observations sur l'activité de formation. L'enseignant a noté que la formation était bien structurée et qu'elle a permis de fournir une compréhension claire des concepts clés en matière de visualisation de données. Il a également souligné que les activités pratiques étaient bien conçues et ont aidé les étudiants à appliquer les concepts théoriques appris en classe.

L'enseignant a noté que les étudiants avaient posé des questions pertinentes et avaient montré un intérêt actif dans la formation. Il a également indiqué qu'il était satisfait du niveau de participation et de l'interaction des étudiants avec les différents éléments de la formation. L'enseignant a recommandé de renforcer davantage la partie pratique de la formation, notamment en fournissant des exemples supplémentaires pour aider les étudiants à appliquer les concepts de visualisation de données à des scénarios du monde réel. Il a également suggéré que des ateliers de formation plus avancés soient proposés pour aider les étudiants à se perfectionner davantage dans l'utilisation des outils de visualisation de données.

En conclusion, l'enseignant a été satisfait de la qualité globale de la formation et a recommandé d'inclure plus d'exemples pratiques pour aider les étudiants à consolider leurs connaissances en visualisation de données.

### Focus Group

Le focus group a permis de recueillir des commentaires qualitatifs des étudiants sur l'activité. Globalement, les étudiants ont exprimé leur satisfaction quant à la qualité de l'activité d'enseignement, en particulier sur la clarté et la pertinence des exemples présentés. Les étudiants ont également apprécié la partie pratique de l'activité, qui leur a permis de mettre en pratique les connaissances théoriques acquises.

Certains étudiants ont suggéré que la partie théorique aurait pu être plus interactive, en incluant plus de discussions de groupe et des activités en sous-groupes. D'autres ont exprimé le souhait d'avoir davantage de temps pour la partie pratique, en particulier pour construire des visualisations plus complexes. Certains ont également mentionné que la quantité de données fournies dans l'activité était suffisante (dû à l'insertion d'erreur volontaire de la part du chercheur).

La majorité des étudiants ont exprimé le besoin d'intégrer les concepts de visualisation de données dans leur cursus scolaire, tant théorique que pratique, en privilégiant des introductions spécifiques dans certains cours du bachelier ("Modélisation de données", "Projet Web", "Analyse", "Statistique", etc) et en corrélant le tout avec des outils spécifiques aux besoins de ces mêmes cours.

En ce qui concerne les aspects à améliorer, certains étudiants ont suggéré d'inclure davantage d'exemples de visualisations pour différents types de données et contextes, ainsi que d'intégrer des exemples de visualisations interactives. Certains ont également suggéré d'explorer des outils de visualisation de données plus avancés, tels que des outils de visualisation de données via des travaux pratiques.

En somme, le focus group a permis de recueillir des commentaires utiles pour améliorer la qualité de l'activité



de formation, en particulier en ce qui concerne la partie théorique et les exemples présentés. Les étudiants ont exprimé leur satisfaction quant à la qualité de la formation et ont suggéré des améliorations pour les futures sessions. Le focus group permet d'obtenir des retours d'expérience des étudiants sur l'activité et leur compréhension des concepts de visualisation de données. Les discussions montrent que les étudiants trouvent la visualisation de données utile pour comprendre et analyser les informations complexes, et qu'ils seraient intéressés par un module d'enseignement spécifique à la visualisation dans leur cursus scolaire. Le focus group permet également d'identifier des pistes d'amélioration pour l'enseignement de la visualisation de données en promotion sociale, telles que l'utilisation d'outils et de ressources pédagogiques adaptés, ou l'adaptation des activités pratiques aux besoins et aux compétences des étudiants.

## 5 Discussions

Cette section présente une discussion des résultats obtenus et une réflexion sur les implications pour l'enseignement de la visualisation de données et la formation des novices en data science. Elle aborde également les réponses aux questions de recherche et aux sous-questions de recherche. Cette étude est en accord avec l'état de l'art sur l'importance de l'adaptation des méthodes pédagogiques aux besoins des étudiants en promotion sociale. Cedefop (2014)[43] a souligné l'importance de l'enseignement adapté pour les étudiants adultes, tandis que Merrill (2002)[86] a souligné l'importance des méthodes pédagogiques adaptées aux styles d'apprentissage des étudiants. D'autres auteurs, tels que Kasworm (2003)[84], ont souligné l'importance de la prise en compte des expériences et des besoins des étudiants pour faciliter leur engagement dans l'apprentissage.

### 5.1 Limitations de l'étude

Il est important de noter que les résultats de cette étude doivent être interprétés avec prudence, étant donné la taille limitée de l'échantillon. Des recherches supplémentaires sont nécessaires pour confirmer et approfondir ces résultats, notamment en élargissant l'échantillon d'étudiants et en explorant d'autres contextes d'enseignement. La réalisation d'une étude longitudinale à l'EAFIC et dans d'autres établissements de promotion social pourrait permettre une analyse plus complète.

### 5.2 Résumé des principaux résultats de l'étude

Les résultats de cette étude apportent une contribution significative à l'état de l'art de l'enseignement de la visualisation de données et de la modélisation de données en data science, en mettant en évidence plusieurs aspects importants.

Tout d'abord, les résultats démontrent l'efficacité de l'approche pédagogique utilisée, basée sur l'utilisation de la modélisation de bases de données comme point d'entrée pour enseigner la visualisation de données aux étudiants novices en data science en promotion sociale. Cette approche a permis une amélioration significative des connaissances et de la compréhension des concepts clés de la visualisation de données chez les étudiants, telle qu'observée dans les questionnaires post-test. Cela confirme l'importance de proposer des activités pratiques et des mises en situation pour faciliter l'apprentissage et favoriser l'engagement des étudiants.

De plus, cette étude met en évidence l'intérêt des étudiants pour l'intégration de la visualisation de données dans leur cursus scolaire en promotion sociale. Les entretiens avec les enseignants et le focus group avec les étudiants ont souligné l'importance de cette compétence pour leur formation professionnelle et leur employabilité. Ainsi, cette étude contribue à la prise de conscience de l'importance de la visualisation de données dans l'enseignement de la data science en promotion sociale.

Par ailleurs, cette recherche apporte une perspective novatrice en soulignant l'insertion d'erreurs comme une approche pédagogique efficace pour l'enseignement de la visualisation de données et de la modélisation de données. Cette approche permet aux étudiants de développer des compétences de résolution de problèmes et de mieux comprendre les concepts clés. Les résultats de cette étude sont en cohérence avec l'état de l'art qui met en avant l'importance de l'apprentissage à partir d'erreurs dans les domaines de l'informatique et de la data science.

En conclusion, cette étude apporte une contribution significative en proposant une approche pédagogique novatrice pour l'enseignement de la visualisation de données et de la modélisation de données en promotion sociale. Les résultats démontrent l'efficacité de cette approche pour améliorer les connaissances des étudiants et favoriser leur engagement dans l'apprentissage. De plus, cette recherche met en évidence l'intérêt des étudiants

pour l'acquisition de compétences en visualisation de données dans leur parcours scolaire. Enfin, l'insertion d'erreurs comme approche pédagogique offre de nouvelles perspectives pour renforcer la compréhension des concepts clés et développer des compétences de résolution de problèmes chez les étudiants.

### **5.3 Réponse aux questions de recherche et aux sous-questions de recherche**

#### **Comment les novices en data science améliorent/ construisent-ils une représentation de données par l'intermédiaire de la visualisation de données ?**

Les résultats de l'étude démontrent que les novices en data science améliorent et construisent leur représentation de données grâce à une approche pédagogique axée sur l'apprentissage par la pratique et les mises en situation. Les activités pratiques et les exercices concrets leur permettent d'explorer et d'appliquer les concepts clés de la visualisation de données, favorisant ainsi leur compréhension et leur capacité à construire des représentations de données pertinentes et efficaces. L'utilisation de techniques d'insertion d'erreurs dans les exercices pratiques s'est avérée particulièrement bénéfique, permettant aux étudiants d'identifier et de corriger les erreurs courantes dans leurs visualisations de données.

#### **Comment intégrer la visualisation de données dans le cursus scolaire des étudiants en promotion sociale novices en data science dans le cadre d'un bachelier en informatique de gestion ?**

L'intégration de la visualisation de données dans le cursus scolaire des étudiants en promotion sociale novices en data science peut être réalisée de plusieurs manières. Une approche consiste à inclure un module d'enseignement spécifique à la visualisation de données, dans lequel les concepts clés et les techniques de visualisation sont enseignés de manière approfondie. Une autre approche consiste à intégrer les concepts de visualisation de données dans les cours existants, tels que la modélisation de données, afin de montrer leur pertinence et leur application concrète. Les enseignants doivent également adapter les méthodes pédagogiques aux besoins spécifiques des étudiants en promotion sociale, en mettant l'accent sur l'apprentissage par la pratique, les mises en situation et l'utilisation d'outils adaptés à leurs besoins.

### **5.4 Implications pédagogiques et perspectives pour l'intégration de la visualisation de données dans les cursus scolaires des étudiants novices en data science**

Les résultats de cette étude ont plusieurs implications pour l'enseignement de la visualisation de données et la formation des étudiants en promotion sociale novices en data science.

Avant tout, les résultats de cette étude ont plusieurs implications significatives pour l'enseignement de la visualisation de données et la formation des novices en data science. L'une des contributions les plus importantes de cette recherche réside dans le développement d'une activité d'enseignement pratique, adaptable et facilement utilisable par les enseignants. Cette activité a été conçue en tenant compte des besoins spécifiques des étudiants en promotion sociale et peut être intégrée de manière flexible dans les cours existants, tels que la modélisation de données. Elle offre aux enseignants une ressource concrète pour introduire les concepts clés de la visualisation de données et faciliter l'apprentissage des étudiants. L'activité peut être adaptée en fonction des niveaux de compétence et des besoins des étudiants, offrant ainsi une approche pédagogique personnalisée et efficace. En fournissant aux enseignants un outil concret et adaptable, cette recherche contribue à combler un écart dans la littérature et offre une perspective pratique pour l'intégration de la visualisation de données dans les cursus scolaires des novices en data science en promotion sociale.

Tout d'abord, il est essentiel de proposer des activités pratiques et des mises en situation pour faciliter l'apprentissage de la visualisation de données chez les étudiants novices. Les enseignants doivent concevoir des activités variées et engageantes qui permettent aux étudiants d'expérimenter les concepts de visualisation de

données et de construire leur propre représentation de données. Cela peut inclure des exercices pratiques avec des jeux de données réels, des études de cas concrets et des projets collaboratifs.

Ensuite, les méthodes pédagogiques doivent être adaptées aux besoins spécifiques des étudiants en promotion sociale. Ces étudiants peuvent avoir des contraintes de temps et des expériences variées, il est donc important de proposer des approches flexibles et personnalisées. Les enseignants doivent être conscients des défis auxquels sont confrontés ces étudiants et les soutenir tout au long de leur apprentissage, en fournissant des ressources supplémentaires, des tutoriels et des séances de soutien individuelles si nécessaire.

Enfin, il est recommandé d'envisager l'inclusion d'un module d'enseignement spécifique à la visualisation de données dans le cursus scolaire des étudiants novices en data science et ce en complément de module spécifique dans les différents cours du cursus. Ce module peut couvrir les principes fondamentaux de la visualisation de données, les techniques de visualisation courantes et les meilleures pratiques pour la création de représentations visuelles efficaces. L'inclusion de ce module permettra aux étudiants d'acquérir des compétences essentielles en visualisation de données et de les appliquer dans leur futur travail en data science.

En résumé, l'intégration de la visualisation de données dans le cursus scolaire des étudiants en promotion sociale novices en data science présente des implications pédagogiques importantes. En adoptant une approche pratique, en adaptant les méthodes pédagogiques aux besoins des étudiants et en proposant un module d'enseignement spécifique, les établissements d'enseignement peuvent fournir aux étudiants les compétences et les connaissances nécessaires pour comprendre et utiliser efficacement la visualisation de données dans le domaine de la data science.

## 5.5 Innovation et différenciation par rapport à l'état de l'art

Notre étude se démarque de l'état de l'art grâce à une approche innovante et différenciée et réside dans plusieurs aspects de notre étude.

Tout d'abord, nous avons appliqué la méthode d'apprentissage par les erreurs à l'enseignement de la visualisation de données pour des étudiants en promotion sociale novices en data science. Bien que cette approche soit couramment utilisée dans des domaines tels que la programmation et la modélisation de données, son application à la visualisation de données est encore relativement peu explorée. En adoptant cette approche, nous avons pu aider les étudiants à mieux comprendre les concepts de visualisation de données en leur permettant d'expérimenter et de corriger activement leurs erreurs, ce qui a conduit à une amélioration significative de leurs compétences et de leur compréhension.

De plus, notre étude a été réalisée dans le contexte spécifique de la promotion sociale, ce qui constitue une différenciation importante par rapport aux recherches existantes. La plupart des études sur l'enseignement de la visualisation de données se concentrent principalement sur des contextes d'enseignement traditionnels, en négligeant souvent les spécificités et les besoins particuliers des étudiants en promotion sociale. En nous intéressant spécifiquement à cette population d'étudiants, nous avons pu identifier les défis et les opportunités propres à leur apprentissage de la visualisation de données, et proposer des approches pédagogiques adaptées à leurs besoins.

Enfin, notre approche d'intégration de la visualisation de données dans les cours existants, tels que la modélisation de données, se distingue également de la plupart des études antérieures. Plutôt que de proposer un nouveau cours dédié à la visualisation de données, nous avons préconisé l'intégration de ces concepts dans des cours déjà existants, ce qui peut être plus réaliste pour de nombreux établissements d'enseignement. Cette approche permet aux étudiants d'acquérir des compétences en visualisation de données tout en suivant leur

curus régulier, ce qui favorise une meilleure intégration des connaissances et une pertinence directe dans leurs futurs domaines d'application.

En résumé, notre étude se distingue par son approche novatrice d'apprentissage par les erreurs dans l'enseignement de la visualisation de données, son exploration du contexte spécifique de la promotion sociale et son approche d'intégration dans les cours existants. Ces différenciations par rapport à l'état de l'art contribuent à une meilleure compréhension de l'enseignement de la visualisation de données et offrent des perspectives intéressantes pour l'amélioration de la formation des novices en data science dans des contextes de promotion sociale.

## 5.6 Perspectives et travaux futurs

L'enseignement de la visualisation de données aux étudiants en promotion sociale présente des défis uniques mais aussi de nouvelles opportunités pour la recherche future. Plusieurs directions de travaux futurs se dessinent à partir de cette étude. Premièrement, une expansion de cette recherche à un échantillon d'étudiants plus large et plus diversifié en promotion sociale est nécessaire pour valider davantage nos conclusions. Il serait intéressant de voir comment les méthodes que nous avons explorées s'appliquent à des étudiants avec différents niveaux de compétence en informatique, avec des expériences de travail variées, ou des étudiants venant de différentes régions.

Deuxièmement, nous avons constaté que l'apprentissage par les erreurs peut améliorer l'efficacité de l'enseignement de la visualisation de données. Cependant, la manière optimale d'intégrer cette approche dans le processus d'enseignement reste à être explorée. Par exemple, déterminer quels types d'erreurs sont les plus instructifs pour les étudiants, et comment ces erreurs peuvent être le mieux intégrées dans le processus d'enseignement.

Troisièmement, l'extension de nos approches à d'autres domaines de la data science, tels que le Machine Learning ou l'analyse de données, pourrait être une direction de recherche intéressante. Il serait utile d'étudier comment améliorer l'apprentissage des étudiants en promotion sociale dans ces domaines clés, qui sont aussi essentiels pour réussir en data science.

De plus, cette étude a suscité un vif intérêt de la part de l'institution d'enseignement en promotion sociale. Les responsables pédagogiques sont conscients de la nécessité d'intégrer les compétences en visualisation de données dans leur programme, surtout compte tenu de l'importance croissante de la data science dans le monde professionnel. Ils sont également préoccupés par la complexité de la tâche pour des étudiants qui peuvent avoir une expérience et des compétences informatiques limitées.

Par conséquent, il a été exprimé le souhait de poursuivre cette recherche, avec pour objectif de développer des méthodes d'enseignement innovantes et adaptées à leur public. Cette perspective comprend la mise en place d'un suivi longitudinal des étudiants pour évaluer l'évolution de leurs compétences en visualisation de données, et la recherche de partenariats avec des entreprises du domaine pour garantir la pertinence et l'actualité de l'enseignement dispensé. Étant donné que les étudiants en promotion sociale peuvent avoir des engagements personnels et professionnels variés, une étude sur comment développer des approches d'enseignement flexibles et intégrées pour ce public serait précieuse.

La collaboration avec les enseignants sera également un élément clé de cette continuation de l'étude. Leur expérience et leur connaissance des étudiants en promotion sociale seront des ressources précieuses pour adapter les méthodes d'enseignement et intégrer efficacement l'apprentissage de la visualisation de données dans le curriculum. Cette étude marque donc le début d'un engagement à long terme pour l'amélioration de l'enseignement de la visualisation de données en promotion sociale.

En conclusion, bien que cette étude ait apporté des informations utiles sur l'enseignement de la visualisation de données en promotion sociale, il reste encore beaucoup à découvrir. Nous espérons que nos résultats inciteront d'autres chercheurs à explorer ce domaine d'étude crucial.

## 6 Conclusion

Dans cette étude, nous avons examiné l'importance de l'intégration de la visualisation de données dans l'enseignement de la promotion sociale et son impact sur l'apprentissage des étudiants novices en data science. Nous avons identifié plusieurs contributions et implications significatives qui peuvent guider les enseignants et les institutions d'enseignement dans l'intégration de la visualisation de données dans le curriculum.

Tout d'abord, l'utilisation de la visualisation de données offre aux étudiants un moyen concret et attrayant de comprendre et d'appliquer les concepts et techniques liés aux données. En proposant des activités pratiques, des projets de groupe et des tutoriels interactifs, les enseignants peuvent favoriser l'apprentissage actif et la construction de connaissances chez les étudiants en promotion sociale.

De plus, cette étude a proposé une activité d'enseignement spécifique à la visualisation de données qui est adaptable et utilisable facilement par les enseignants. Cette ressource concrète offre aux enseignants un outil pratique pour introduire les concepts clés de la visualisation de données et encourager les étudiants à construire leur propre représentation de données.

En mettant l'accent sur l'engagement des étudiants dans l'apprentissage, cette étude a souligné l'importance de l'apprentissage collaboratif, de l'utilisation de projets concrets et de l'application des connaissances dans des situations réelles. Ces approches pédagogiques actives favorisent non seulement l'apprentissage des étudiants, mais également leur intérêt et leur motivation pour la visualisation de données.

En conclusion, l'intégration de la visualisation de données dans l'enseignement de la promotion sociale présente des avantages significatifs pour les étudiants novices en data science. En adoptant des approches pédagogiques adaptées, en fournissant des activités pratiques et engageantes, et en offrant des outils et des ressources adaptés, les enseignants peuvent faciliter l'apprentissage et la compréhension des concepts et techniques de la visualisation de données chez les étudiants en promotion sociale.

Il est donc essentiel que les institutions d'enseignement reconnaissent l'importance de la visualisation de données et envisagent son intégration systématique dans leur curriculum. En investissant dans l'amélioration de l'apprentissage des étudiants en promotion sociale, ces institutions contribuent à les équiper avec des compétences cruciales pour réussir dans un monde axé sur les données.

# Bibliographie

- [1] S. KRUK et HOLUBOWICZ, « Advances in database systems education : Methods, tools, curricula, and way forward, » Education and Information Technologies, t. 27, n° 1, p. 1-39, 2022.
- [2] M. BEEGE, S. SCHNEIDER, S. NEBEL, J. ZIMM, S. WINDISCH et G. D. REY, « Learning programming from erroneous worked-examples. Which type of error is beneficial for learning ? » Learning and Instruction, t. 75, p. 101 497, 2021, ISSN : 0959-4752.
- [3] M. BEEGE, S. SCHNEIDER, S. NEBEL, J. ZIMM, S. WINDISCH et G. D. REY, « Learning programming from erroneous worked-examples. Which type of error is beneficial for learning ? » Learning and Instruction, t. 75, 2021, Cited by : 6.
- [4] M. BEEGE, S. SCHNEIDER, S. NEBEL, J. ZIMM, S. WINDISCH et G. D. REY, « Learning programming from erroneous worked-examples. Which type of error is beneficial for learning ? » Learning and Instruction, t. 71, p. 101 497, 2021.
- [5] K. KUCHER, R. M. MARTINS et A. KERREN, « Project in Visualization and Data Analysis : Experiences in Designing and Coordinating the Course, » in Eurographics 2021 - Education Papers, B. SOUSA SANTOS et G. DOMIK, éd., The Eurographics Association, 2021, ISBN : 978-3-03868-132-8.
- [6] V. L. B. PH.D., « Innovative Pedagogy for Teaching and Learning Data Visualization, » in 2021 ASEE Virtual Annual Conference Content Access, Virtual Conference : ASEE Conferences, juill. 2021.
- [7] M. SECKLER, J. HEER et Y. JANSEN, « What is Data ? - Exploring the Meaning of Data in Data Physicalisation Teaching, » Scientific Programming, t. 2021, p. 6 629 086, 2021.
- [8] V. BONNEMAINS, P. CHASTE et D. BERGER, « Teaching data visualisation : a practice-based model with R, » Journal of Statistics Education, t. 28, n° 1, p. 55-65, 2020.
- [9] M. BURCH et E. MELBY, « What More than a Hundred Project Groups Reveal about Teaching Visualization, » J. Vis., t. 23, n° 5, p. 895-911, oct. 2020, ISSN : 1343-8875.
- [10] A. CHATZIMPARMPAS, R. M. MARTINS, I. JUSUFI et A. KERREN, « A survey of surveys on the use of visualization for interpreting machine learning models, » Information Visualization, t. 19, n° 3, p. 207-233, 2020.
- [11] U. H. SYEDA, P. MURALI, L. ROE, B. BERKEY et M. A. BORKIN, « Design Study "Lite" Methodology : Expediting Design Studies and Enabling the Synergy of Visualization Pedagogy and Social Good, » in Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, sér. CHI '20, Honolulu, HI, USA : Association for Computing Machinery, 2020, p. 1-13, ISBN : 9781450367080. adresse : <https://doi.org/10.1145/3313831.3376829>.
- [12] M. AKCAYIR et G. AKCAYIR, « Teaching computer programming through online remote laboratories : A systematic review, » Computers & Education, t. 128, p. 412-426, 2019.
- [13] M. BORKIN, Z. BYLINSKII, P. DECAMP, G. FEINBERG, S. GHOSH, C. GOUVEIA et R. UNGER, « Visualization Literacy : The Role of Visualization in the K-16 Education System, » IEEE Computer Graphics and Applications, t. 39, n° 6, p. 20-34, 2019.



- [14] M. BREHMER et T. MUNZNER, « What more than a hundred project groups reveal about teaching visualization, » IEEE Transactions on Visualization and Computer Graphics, n° 1, p. 246-256, 2019.
- [15] M. BURCH et E. MELBY, « Teaching and Evaluating Collaborative Group Work in Large Visualization Courses, » in Proceedings of the 12th International Symposium on Visual Information Communication and Interaction, sér. VINCI '19, Shanghai, China : Association for Computing Machinery, 2019, ISBN : 9781450376266.
- [16] E. DELVAUX, A. DOPPAGNE et C. ELSEN, « Une expérimentation de «scénario hybride» dans l'enseignement supérieur de promotion sociale, » Revue internationale des technologies en pédagogie universitaire, t. 16, p. 17-30, 1 2019.
- [17] P. GERMANAKOS, M. CONSTANTINIDES, M. KASINIDOU et G. SAMARAS, « A metacognitive perspective of infovis in education, » cited By 0, 2019, p. 331-336.
- [18] R. KUPIAINEN, J. HAUTALA, M. MIETTINEN et H. RUOKAMO, « A systematic review of the factors influencing adult learners' motivation to learn in online settings, » Computers & Education, t. 138, p. 83-97, 2019.
- [19] S. KURNIAWATI, T. SUTABRI et Y. WIBISONO, « Facilitating Learning by Practice and Examples : a Tool for Learning Table Normalization, » ACM Transactions on Computing Education (TOCE), t. 19, n° 4, p. 1-20, 2019.
- [20] S. L'YI, Y. CHANG, D. SHIN et J. SEO, « Toward Understanding Representation Methods in Visualization Recommendations through Scatterplot Construction Tasks, » Computer Graphics Forum, t. 38, n° 3, p. 201-211, 2019.
- [21] L. Y.-H. LO, Y. MING et H. QU, « Learning Vis Tools : Teaching Data Visualization Tutorials, » 2019 IEEE Visualization Conference (VIS), p. 11-15, 2019.
- [22] W. WILLETT, « Designing Visualizations for Novices : A User-Centered Design Approach, » IEEE Transactions on Visualization and Computer Graphics, t. 25, n° 1, p. 919-929, 2019.
- [23] M. AKCAYIR et G. AKCAYIR, « Advantages and challenges associated with blended learning implementation in vocational education : A case study, » The Internet and Higher Education, t. 37, p. 33-40, 2018.
- [24] K. J. HEPWORTH et C. CANON, « Improving Science Students' Data Visualizations : A STEAM-Based Approach, » Dialectic, 2018.
- [25] W. JAKKAEW, S. WONGWANICH et H. XIE, « Social presence, student engagement, and online learning environments : A study using Learning Management Systems, » Computers & Education, t. 116, p. 1-12, 2018.
- [26] A. KATZ, « Applying TSSL in database schema modeling : Visualizing the syntax of gradual transitions, » Journal of Technology and Science Education, t. 8, n° 4, 2018.
- [27] J. D. KELLEHER et B. TIERNEY, Data science fundamentals. CRC Press, 2018.
- [28] C. RUGGLES, A. KACHKAEV et M. A. BORKIN, « Training novices to create valid visualizations, » IEEE Computer Graphics and Applications, t. 38, n° 2, p. 7-13, 2018.
- [29] B. ALPER, N. H. RICHE, F. CHEVALIER, J. BOY et M. SEZGIN, « Visualization Literacy at Elementary School, » in Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, sér. CHI '17, Denver, Colorado, USA : Association for Computing Machinery, 2017, p. 5485-5497, ISBN : 9781450346559.
- [30] S. MCKENNA, A. LEX et M. MEYER, Worksheets for Guiding Novices through the Visualization Design Process, 2017. arXiv : 1709.05723 [cs.HC].

- [31] R. BACHELET, A. CARON-PARGUE et P. MORLAT, « Étude des compétences informationnelles des apprenants inscrits en promotion sociale en Communauté française de Belgique, » Les Cahiers pédagogiques, t. 527, p. 50-53, 2016.
- [32] R. ELMASRI et S. NAVATHE, « Fundamentals of database systems, » 2016.
- [33] A. KATZ et R. SHMALLO, « Learning from errors as a pedagogic approach for reaching a higher conceptual level in database modeling, » Lecture Notes in Business Information Processing, t. 249, p. 93-102, 2016, Cited by : 1.
- [34] A. KIRK, Data visualisation : A handbook for data driven design. 2016.
- [35] G. LANZARA et I. PALMISANO, « Visualizing the syntax of gradual transitions in teaching database modeling, » International Journal of Information and Education Technology, n° 8, p. 547-552, 2016.
- [36] S. LEE, S.-H. KIM, Y.-H. HUNG, H. LAM, Y.-a. KANG et J. S. YI, « How do People Make Sense of Unfamiliar Visualizations ? : A Grounded Model of Novice's Information Visualization Sensemaking, » IEEE Transactions on Visualization and Computer Graphics, t. 22, n° 1, p. 499-508, 2016.
- [37] M. A. BORKIN, Z. YAN et B. HORN, « Visualization education through social impact : a service-learning approach for visualization pedagogy, » IEEE Computer Graphics and Applications, t. 35, n° 4, p. 10-15, 2015.
- [38] D. T. CAMPBELL, J. C. STANLEY et N. L. GAGE, Experimental and quasi-experimental designs for generalized causal inference. Ravenio Books, 2015.
- [39] S. LEE, S.-H. KIM et Y.-H. HUNG, « How do people make sense of unfamiliar visualizations ? : A grounded model of novice's information visualization sensemaking, » IEEE Transactions on Visualization and Computer Graphics, t. 22, n° 1, p. 499-508, 2015.
- [40] Y. LEUNG, « Validity, reliability, and generalizability in qualitative research, » Journal of Family Medicine and Primary Care, t. 4, n° 3, p. 324-327, 2015.
- [41] R. SHMALLO et A. KATZ, « Educating conceptual modeling of relational database using learning from errors approach, » juin 2015.
- [42] E. TAVERNIER et B. GALAND, « Analyse exploratoire de données pour des cours de sciences de l'ingénieur en promotion sociale, » in EMOI-TICE 2015 : Environnement MOTivant Informatique et TICE, Presses Universitaires de Louvain, 2015, p. 109-124.
- [43] CEDEFOP, Skills, qualifications and jobs in the EU : The making of a perfect match ? Publications Office of the European Union, 2014.
- [44] R. A. KRUEGER et M. A. CASEY, Focus Groups : A Practical Guide for Applied Research, 5<sup>e</sup> éd. SAGE Publications, Inc, 2014.
- [45] T. MUNZNER, Visualization analysis and design. CRC press, 2014.
- [46] K. PANTAZOS, « Constructing Visualizations with InfoVis Tools – An Evaluation from a User Perspective, » Computer Graphics Forum, t. 33, n° 3, p. 21-30, 2014.
- [47] M. Q. PATTON, Qualitative research & evaluation methods : integrating theory and practice. Sage Publications, 2014.
- [48] C. WARE, « Information Visualization : Perception for Design : Second Edition, » in avr. 2014.
- [49] A. CAIRO, The functional art : An introduction to information graphics and visualization. 2013.
- [50] N. DESCHRYVER, Apprentissage autonome, motivation et contexte socioculturel : le cas de la formation d'adultes en promotion sociale. Presses universitaires de Louvain, 2013.
- [51] F. J. FOWLER, Survey Research Methods, 5<sup>e</sup> éd. SAGE Publications, Inc, 2013.

- [52] J. GINESTIÉ et A. TRICOT, Activité d'élèves, activité d'enseignants en éducation scientifique et technologique. 2013.
- [53] Y. JANSEN et P. DRAGICEVIC, « An interaction model for visualizations beyond the desktop, » IEEE Transactions on Visualization and Computer Graphics, t. 19, n° 12, p. 2396-2405, 2013.
- [54] A. KERREN, « Information visualization courses for students with a computer science background, » IEEE Computer Graphics and Applications, t. 33, n° 2, p. 12-15, 2013, cited By 1.
- [55] J. PIETTE, « Pédagogie de la coopération et enseignement en promotion sociale, » Recherche en éducation, t. 18, p. 119-126, 2013.
- [56] S. FEW, Show me the numbers : Designing tables and graphs to enlighten. 2012.
- [57] N. GILLARD, A. PÉRILLEUX et B. ROLAND, « Formation, travail et technologies de l'information et de la communication : Le point de vue d'enseignants de promotion sociale en Communauté française de Belgique, » Les Cahiers pédagogiques, t. 506, p. 58-60, 2012.
- [58] J. HEER et M. BOSTOCK, « Interactive data visualization : Foundations, techniques, and applications, » Foundations and Trends® in Human-Computer Interaction, t. 2, n° 1, p. 1-104, 2012.
- [59] J. HEER et B. SHNEIDERMAN, « Interactive Dynamics for Visual Analysis, » Queue, t. 5, n° 1, fév. 2012, ISSN : 1077-2626.
- [60] M. SEDLMAIR, M. MEYER et T. MUNZNER, « Design Study Methodology : Reflections from the Trenches and the Stacks, » IEEE Transactions on Visualization and Computer Graphics, t. 18, n° 12, p. 2431-2440, 2012.
- [61] C. WARE, Information visualization : perception for design. 2012.
- [62] H. R. BERNARD, Research Methods in Anthropology : Qualitative and Quantitative Approaches, 5<sup>e</sup> éd. Rowman & Littlefield Publishers, 2011.
- [63] H. R. BERNARD, Research methods in anthropology : qualitative and quantitative approaches. AltaMira Press, 2011.
- [64] T. HAINEY, T. M. CONNOLLY et M. STANSFIELD, « Students' perceptions of the use of e-learning in mathematics, » Computers & Education, t. 57, n° 3, p. 2025-2035, 2011.
- [65] P. ISENBERG, N. ELMQVIST, J. SCHOLTZ, D. CERNEA, K.-L. MA et H. HAGEN, « Collaborative Visualization : Definition, Challenges, and Research Agenda, » IEEE Comput. Graph. Appl., t. 31, n° 6, p. 62-69, nov. 2011, ISSN : 0272-1716.
- [66] R. D. PENG et E. MATSUI, « Visualizing incomplete and partial data, » Journal of statistical software, t. 46, n° 5, p. 1-29, 2011.
- [67] L. GRAMMEL, « User interfaces supporting information visualization novices in visualization construction, » IEEE Transactions on Visualization and Computer Graphics, t. 16, n° 6, p. 1017-1026, 2010.
- [68] B. HERR, « Visual strategies for developing a new pedagogy for teaching statistics, » Statistics Education Research Journal, t. 9, n° 2, p. 22-38, 2010.
- [69] H. L. ROEDIGER, A. C. BUTLER et E. J. MARSH, « Optimizing learning and instruction : Theory, research, and application, » Contemporary educational psychology, t. 35, n° 3, p. 181-192, 2010.
- [70] P. BISSCHOP et J. TONDEUR, « Competenties voor ICT integratie in het volwassenenonderwijs, » Tijdschrift voor Didactiek der -Wetenschappen, t. 26, p. 5-22, 1 2009.
- [71] S. FEW, Now you see it : simple visualization techniques for quantitative analysis. 2009.
- [72] C. VAN RIPER et R. M. KIRBY, « Promoting efficient use of visualization tools through education, » Journal of Computing in Civil Engineering, t. 23, n° 6, p. 428-436, 2009.

- [73] A. KERREN, J. STASKO et J. DYKES, « Teaching information visualization, » t. 4950 LNCS, p. 65-91, 2008, cited By 16.
- [74] A. KERREN, J. T. STASKO et J. DYKES, « Teaching Information Visualization, » in Information Visualization, 2008.
- [75] M. V. ANGROSINO, Naturalistic Observation. Walnut Creek, CA : Left Coast Press, 2007.
- [76] D. A. DILLMAN, J. D. SMYTH et L. M. CHRISTIAN, Internet, Mail, and Mixed-Mode Surveys : The Tailored Design Method. 2007.
- [77] J.-D. FEKETE, « Teaching information visualization, » in Information Visualization, Berlin, Heidelberg : Springer, 2007, p. 57-77.
- [78] H. PASHLER, D. ROHRER, N. J. CEPEDA et S. K. CARPENTER, « Effective learning and the allocation of study time, » Transfer of learning from a modern multidisciplinary perspective, t. 12, p. 325-349, 2007.
- [79] J. S. YI, Y. a. KANG, J. T. STASKO et J. A. JACKO, « Toward a Deeper Understanding of the Role of Interaction in Information Visualization, » IEEE Trans. Vis. Comput. Graph., t. 13, n° 6, p. 1224-1231, déc. 2007, ISSN : 1077-2626.
- [80] S. FEW, Information dashboard design : The effective visual communication of data. O'Reilly Media, Inc., 2006.
- [81] I. SEIDMAN, Interviewing as qualitative research : A guide for researchers in education and the social sciences. Teachers college press, 2006.
- [82] J. HEER, F. B. VI'EGAS et M. WATTENBERG, « Voyagers and voyeurs : Supporting asynchronous collaborative information visualization, » in CHI'05 extended abstracts on Human factors in computing systems, ACM, 2005, p. 1212-1215.
- [83] C. VINCENT, Les technologies de l'information et de la communication dans l'éducation des adultes. Presses universitaires de France, 2004.
- [84] C. KASWORM, « Adult learners in higher education : Barriers to success and strategies to improve results, » ERIC Digest, t. 2003, p. 1-5, 2003.
- [85] J. MARGOLIS et A. FISHER, Unlocking the clubhouse : Women in computing. MIT press, 2002.
- [86] M. D. MERRILL, First principles of instruction. Wadsworth Belmont, CA, 2002.
- [87] M. BLOOR, J. FRANKLAND, M. THOMAS et K. ROBSON, Focus Groups in Social Research. mars 2001, ISBN : 9780761957423.
- [88] W. KENT, Data and reality : a timeless perspective on perceiving and managing information in our imprecise world. 2000.
- [89] R. A. KRUEGER et M. A. CASEY, Focus Groups : A Practical Guide for Applied Research. Sage Publications, 2000.
- [90] S. K. CARD, J. D. MACKINLAY et B. SHNEIDERMAN, « Readings in information visualization : using vision to think, » Morgan Kaufmann, 1999.
- [91] D. L. MORGAN, Focus Groups as Qualitative Research. Sage Publications, 1997.
- [92] R. NOSS et C. HOYLES, Windows on Mathematical Meanings : Learning Cultures and Computers. Springer, 1996.
- [93] D. W. STEWART et P. N. SHAMDASANI, Focus Groups : Theory and Practice. Sage Publications, 1990.
- [94] A. BROWN, « Metacognition, executive control, self-regulation, and other more mysterious mechanisms, » Metacognition, Motivation, and Understanding, p. 65-116, 1987.



# Table des figures

|    |  |    |
|----|--|----|
| 1  | Pipeline de visualisation de données (Jansen et al. 2013[53]) . . . . .  | 10 |
| 2  | Exemple d'un graphique linéaire et d'un graphique à barres <sup>12</sup> . . . . .   | 10 |
| 3  | Exemple d'un graphique géographique . . . . .  | 10 |
| 4  | Exemple d'une tablette se basant sur une application Web à destination des élèves d'une école primaire Alper,B (2017)[29] . . . . .  | 12 |
| 5  | Un modèle ancré de la visualisation de l'information pour les novices (modèle NOVIS). Les flèches indiquent les principales transitions entre les cinq activités et les nombres entre parenthèses indiquent le nombre de transitions entre les activités que nous avons observé à partir des données[36] | 21 |
| 6  | Exemple de dépendances entre attributs qui violent la 3NF[33] . . . . .  | 21 |
| 7  | Un schéma partiel d'un scénario de magasin de fleurs pour démontrer des exemples d'exercices erronés[33] . . . . .   | 22 |
| 8  | Exemple d'erreur sémantique : invocation manquante de la fonction hallo (la version correcte est affichée ci-dessous).[4] . . . . .  | 22 |
| 9  | Effet d'interaction de l'inclusion des erreurs syntaxiques et des erreurs sémantiques sur les performances des tests de connaissances sémantiques ; M et SD sont affichés.[4] . . . . .  | 22 |
| 10 | Diagramme de flux de la méthodologie . . . . .   | 25 |
| 11 | Diagramme de flux de la méthodologie pour l'observation participante . . . . .   | 26 |
| 12 | Schéma de l'analyse des données . . . . .  | 34 |
| 13 | Les étudiants pendant l'observation participante . . . . .   | 36 |
| 14 | Plan de la classe pendant l'observation . . . . .  | 37 |
| 15 | Exemple simple de l'utilité d'une visualisation de données . . . . .   | 38 |
| 16 | Carte réalisée par John Snow en 1854 relevant la propagation du cholera dans le quartier de Soho à Londres en 1854 . . . . .   | 39 |
| 17 | Abstraction de données - Type d'ensemble de données . . . . .  | 39 |
| 18 | Jurons du capitaine Haddock recensés par la SIJA . . . . .   | 40 |
| 19 | Cycle de vie d'une visualisation de données . . . . .  | 40 |
| 20 | Exemple d'une mise en situation . . . . .  | 41 |
| 21 | Proposition de solution pour la partie pratique de l'activité . . . . .  | 42 |
| 22 | Graphique représentant la démographie . . . . .  | 47 |
| 23 | Graphique compréhension de donnée : "Parmi les propositions suivantes, sélectionner le tableau où un attribut est entouré" (Q1_1) pre-test / post-test (colonne est la bonne réponse) . . . . .  | 47 |
| 24 | Est-ce qu'une table (d'une base de données) est considérée comme une visualisation de données?(non est la bonne réponse) . . . . .   | 48 |
| 25 | Choix de la meilleur visualisation pour un scénario donné (préférence des visualisations) pre-test et post-test (Q13) . . . . .  | 50 |
| 26 | Choix de la meilleur visualisation pour un scénario donné (préférence des visualisations) pre-test et post-test . . . . .  | 51 |
| 27 | Compétences en visualisation de données (Q21 - Q25) pre-test / post-test . . . . .   | 53 |

28 Représentation de la confiance des étudiants en la compréhension de la visualisation . . . . . 53

# Liste des tableaux

|   |   |    |
|---|---|----|
| 1 | Tableau de synthèse des études de cas par article . . . . .               | 14 |
| 2 | Tableau d'exemple des questions sélectionnées du questionnaire . . . . .  | 30 |
| 3 | Table de données utilisée pour la partie pratique de l'activité . . . . . | 41 |
| 4 | Questionnaire Pre et Post-test . . . . .                                  | 80 |





# Annexes



## A Annexes A - Questionnaire pre-test et post-test

| Section / Question                        | Options / Notes   |
|---|---|
| Q0_1 : Situation professionnelle          | Sans emploi, Employé(e), Indépendant(e), Retraité(e), Fonctionnaire   |
| Q0_2 : Âge                                |   |
| Q0_3 : Dernier diplôme obtenu             | Primaire, Secondaire inférieur, Secondaire supérieur, Enseignement supérieur de type court, Enseignement supérieur de type long |
| Q0_4 : Genre                              | Masculin, Féminin, Ne souhaite pas le préciser  |
| Q0_5 : En quelle année scolaire vous êtes | 1ère année, 2nde année, 3e année, 4e année, Technicien(ne) en informatique  |

| NCLI | NOM       | ADRESSE              | LOCALITE  | CATI | COMPTE |
|------|-----------|----------------------|-----------|------|--------|
| 0002 | GOFFIN    | 72, r. de la Gare    | Namen     | B2   | -3200  |
| B112 | HANDELNE  | 23, r. Dumont        | Pothens   | C1   | 1250   |
| 0332 | MONTI     | 112, r. Neuse        | Genève    | B2   | 0      |
| 0512 | GILLET    | 14, r. de l'Éto      | Toulouse  | B1   | -6700  |
| C003 | AVIRON    | 8, r. de la Cure     | Toulouse  | B1   | -1700  |
| C123 | MERCIER   | 25, r. Leneche       | Namen     | C1   | -2300  |
| C400 | FERRARD   | 65, r. du Tertre     | Pothens   | B2   | 350    |
| D003 | MERCIER   | 201, boulevard Nord  | Toulouse  |      | -2200  |
| F010 | TOUSSAINT | 5, r. Godebold       | Pothens   | C1   | 0      |
| F011 | PONCELET  | 17, Cité des Étables | Toulouse  | B2   | 0      |
| F400 | JACOB     | 79, ch. de Moulins   | Rhexelles | C2   | 0      |
| H111 | VANBIST   | 180, r. Flammont     | Lille     | B1   | 720    |
| H720 | NEUMAN    | 40, r. Bonnat        | Toulouse  | 0    | 0      |
| L422 | FRANCK    | 90, r. de Wippen     | Namen     | C1   | 0      |
| L423 | FRANCK    | 3, av. des Rives     | Namen     | C1   | -4800  |
| S127 | VANDERVA  | 3, av. des Rives     | Namen     | C1   | -4800  |
| S128 | GUILLAUME | 14, ch. des Rives    | Pays      | B1   | 0      |

Colonne

| NCLI | NOM       | ADRESSE              | LOCALITE  | CATI | COMPTE |
|------|-----------|----------------------|-----------|------|--------|
| 0002 | GOFFIN    | 72, r. de la Gare    | Namen     | B2   | -3200  |
| B112 | HANDELNE  | 23, r. Dumont        | Pothens   | C1   | 1250   |
| 0332 | MONTI     | 112, r. Neuse        | Genève    | B2   | 0      |
| 0512 | GILLET    | 14, r. de l'Éto      | Toulouse  | B1   | -6700  |
| C003 | AVIRON    | 8, r. de la Cure     | Toulouse  | B1   | -1700  |
| C123 | MERCIER   | 25, r. Leneche       | Namen     | C1   | -2300  |
| C400 | FERRARD   | 65, r. du Tertre     | Pothens   | B2   | 350    |
| D003 | MERCIER   | 201, boulevard Nord  | Toulouse  |      | -2200  |
| F010 | TOUSSAINT | 5, r. Godebold       | Pothens   | C1   | 0      |
| F011 | PONCELET  | 17, Cité des Étables | Toulouse  | B2   | 0      |
| F400 | JACOB     | 79, ch. de Moulins   | Rhexelles | C2   | 0      |
| H111 | VANBIST   | 180, r. Flammont     | Lille     | B1   | 720    |
| H720 | NEUMAN    | 40, r. Bonnat        | Toulouse  | 0    | 0      |
| L422 | FRANCK    | 90, r. de Wippen     | Namen     | C1   | 0      |
| L423 | FRANCK    | 3, av. des Rives     | Namen     | C1   | -4800  |
| S127 | VANDERVA  | 3, av. des Rives     | Namen     | C1   | -4800  |
| S128 | GUILLAUME | 14, ch. des Rives    | Pays      | B1   | 0      |

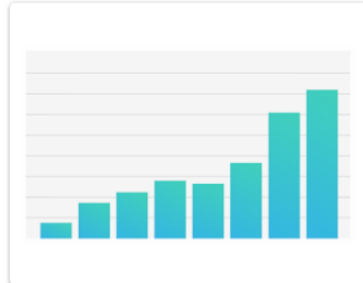
Ligne

| NCLI | NOM       | ADRESSE              | LOCALITE  | CATI | COMPTE |
|------|-----------|----------------------|-----------|------|--------|
| 0002 | GOFFIN    | 72, r. de la Gare    | Namen     | B2   | -3200  |
| B112 | HANDELNE  | 23, r. Dumont        | Pothens   | C1   | 1250   |
| 0332 | MONTI     | 112, r. Neuse        | Genève    | B2   | 0      |
| 0512 | GILLET    | 14, r. de l'Éto      | Toulouse  | B1   | -6700  |
| C003 | AVIRON    | 8, r. de la Cure     | Toulouse  | B1   | -1700  |
| C123 | MERCIER   | 25, r. Leneche       | Namen     | C1   | -2300  |
| C400 | FERRARD   | 65, r. du Tertre     | Pothens   | B2   | 350    |
| D003 | MERCIER   | 201, boulevard Nord  | Toulouse  |      | -2200  |
| F010 | TOUSSAINT | 5, r. Godebold       | Pothens   | C1   | 0      |
| F011 | PONCELET  | 17, Cité des Étables | Toulouse  | B2   | 0      |
| F400 | JACOB     | 79, ch. de Moulins   | Rhexelles | C2   | 0      |
| H111 | VANBIST   | 180, r. Flammont     | Lille     | B1   | 720    |
| H720 | NEUMAN    | 40, r. Bonnat        | Toulouse  | 0    | 0      |
| L422 | FRANCK    | 90, r. de Wippen     | Namen     | C1   | 0      |
| L423 | FRANCK    | 3, av. des Rives     | Namen     | C1   | -4800  |
| S127 | VANDERVA  | 3, av. des Rives     | Namen     | C1   | -4800  |
| S128 | GUILLAUME | 14, ch. des Rives    | Pays      | B1   | 0      |

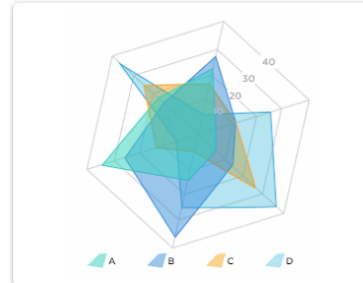
Attribut

|   |                                 |
|---|---------------------------------|
| Q1_1 : Parmi les propositions suivantes, sélectionner le tableau où un attribut est entouré           | La bonne réponse est "colonne"  |
| Q1_2 : Parmi les propositions suivantes, sélectionner le tableau où un item est entouré               | La bonne réponse est "ligne"    |
| Q1_3 : Parmi les propositions suivantes, sélectionner le tableau où une cellule est entourée          | La bonne réponse est "attribut" |
| Q02 : La visualisation de données c'est   |                                 |
| Q03 : Est-ce qu'une table (d'une base de données) est considérée comme une visualisation de données ? | Oui/Non/Je ne sais pas          |
| Q04 : Est-ce qu'un tableau Excel est considéré comme une visualisation de données ?                   | Oui/Non/Je ne sais pas          |
| Q05 : Quelle utilité voyez vous à l'utilisation de la visualisation de données ?                      |                                 |

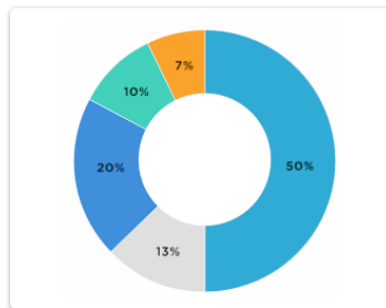
| Section / Question   | Options / Notes |
|--|-----------------|
| Q06 : À qui une visualisation de données peut être utile ?   |                 |
| Q07 : Je pense que la visualisation de données pourrait m'être utile dans la suite de mes études             |                 |
| Q08 : Je pense que la visualisation de données pourrait m'être utile dans ma future carrière professionnelle |                 |



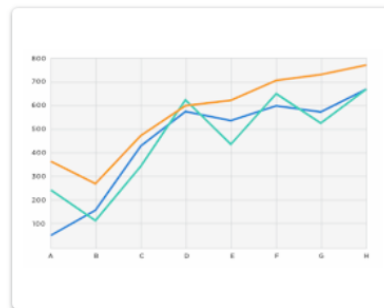
Graphique à barres et histogrammes



Radar



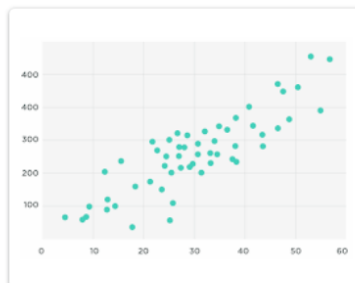
Graphique en anneau



Graphique linéaire



Diagramme géographique



Nuage de points / diagramme de dispersion

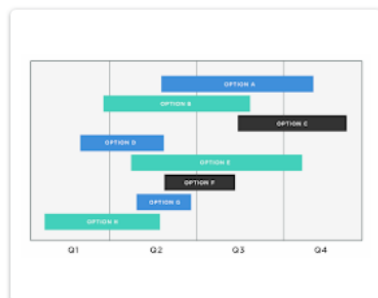
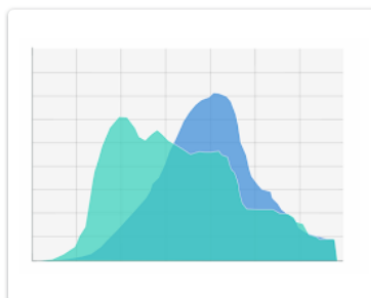


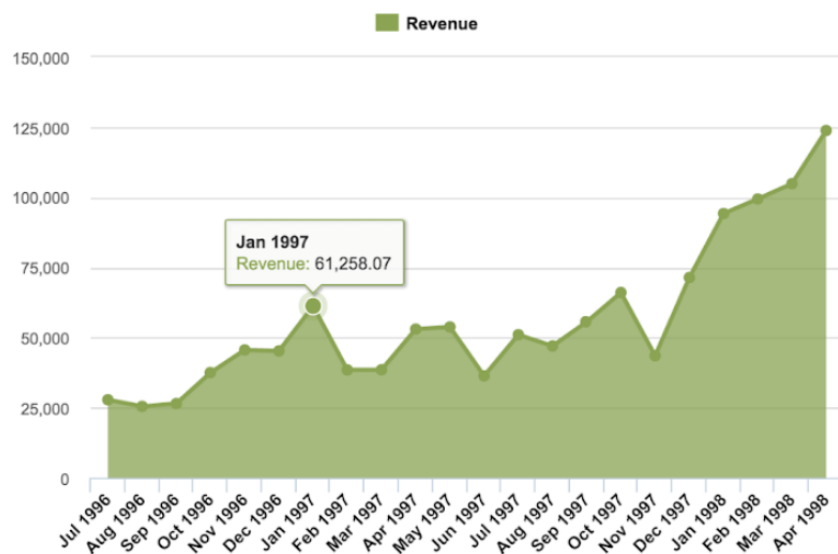
Diagramme de Gantt



Graphique en aires

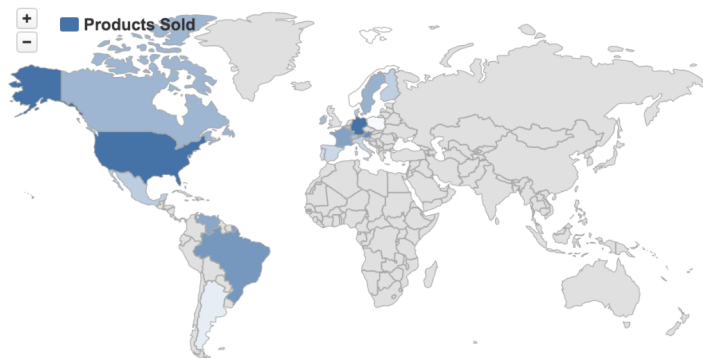
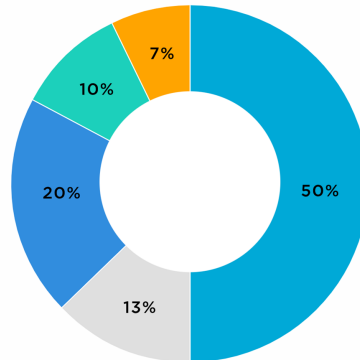
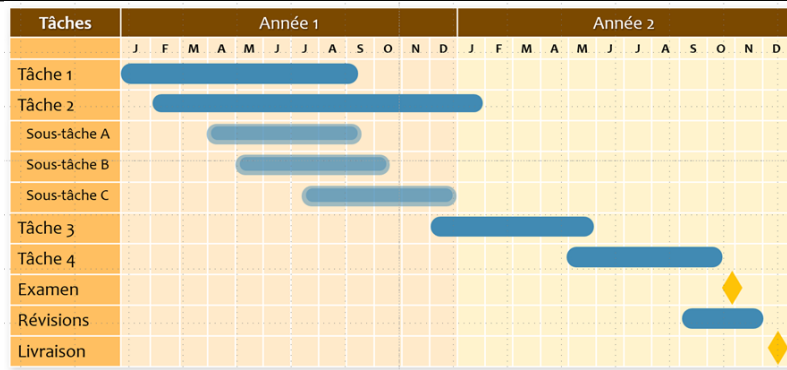
| Section / Question                                | Options / Notes   |
|---|---|
| <p><input type="radio"/> Graphique circulaire</p> | <p><input type="radio"/> Aucune de ces propositions</p> |

|   |  |
|---|--|
| Q09 : Quelle est la meilleure visualisation pour                | Scénario : Service marketing souhaite connaître la répartition de ses clients                              |
| Q10 : Quels sont les critères qui vous ont motivés à répondre ? |  |
| Q11 : Quelle est la meilleure visualisation pour                | Scénario : Service RH souhaite obtenir l'historique des engagements du personnel                           |
| Q12 : Quels sont les critères qui vous ont motivés à répondre ? |  |
| Q13 : Quelle est la meilleure visualisation pour                | Scénario : Service comptabilité souhaite visualiser la répartition du budget                               |
| Q14 : Quels sont les critères qui vous ont motivés à répondre ? |  |
| Q15 : Quelle est la meilleure visualisation pour                | Scénario : Visualiser les compétences des 5 candidats au poste de développeur                              |
| Q16 : Quels sont les critères qui vous ont motivés à répondre ? |  |
| Q17 : Quelle est la meilleure visualisation pour                | Scénario : Créer le planning du projet d'une nouvelle application de gestion du courrier                   |
| Q18 : Quels sont les critères qui vous ont motivés à répondre ? |  |
| Q19 : Quelle est la meilleure visualisation pour                | Scénario : Service RH a besoin d'une projection des engagements, des départs à la retraite et des absences |
| Q20 : Quels sont les critères qui vous ont motivés à répondre ? |  |



## Section / Question

## Options / Notes



Q21 : Est-il possible pour moi de réaliser la visualisation suivante? (graphique linéaire)

- Oui, je sais la réaliser seul.
- Non, j'ai besoin d'un expert.
- Non, je ne comprend pas le sens de cette visualisation
- Non, je ne connais pas d'outils permettant de réaliser une telle visualisation.

| Section / Question  | Options / Notes  |
|---|--|
| Q22 : Est-il possible pour moi de réaliser la visualisation suivante ? (diagramme de Gantt)                   | <ul style="list-style-type: none"> <li>• Oui, je sais la réaliser seul.</li> <li>• Non, j'ai besoin d'un expert.</li> <li>• Non, je ne comprend pas le sens de cette visualisation</li> <li>• Non, je ne connais pas d'outils permettant de réaliser une telle visualisation.</li> </ul> |
| Q23 : Est-il possible pour moi de réaliser la visualisation suivante ? (graphique bubble)                     | <ul style="list-style-type: none"> <li>• Oui, je sais la réaliser seul.</li> <li>• Non, j'ai besoin d'un expert.</li> <li>• Non, je ne comprend pas le sens de cette visualisation</li> <li>• Non, je ne connais pas d'outils permettant de réaliser une telle visualisation.</li> </ul> |
| Q24 : Est-il possible pour moi de réaliser la visualisation suivante ? (graphique donut)                      | <ul style="list-style-type: none"> <li>• Oui, je sais la réaliser seul.</li> <li>• Non, j'ai besoin d'un expert.</li> <li>• Non, je ne comprend pas le sens de cette visualisation</li> <li>• Non, je ne connais pas d'outils permettant de réaliser une telle visualisation.</li> </ul> |
| Q25 : Est-il possible pour moi de réaliser la visualisation suivante ? (graphique géographique)               | <ul style="list-style-type: none"> <li>• Oui, je sais la réaliser seul.</li> <li>• Non, j'ai besoin d'un expert.</li> <li>• Non, je ne comprend pas le sens de cette visualisation</li> <li>• Non, je ne connais pas d'outils permettant de réaliser une telle visualisation.</li> </ul> |
| Q26 : Une visualisation doit toujours afficher l'entièreté des données disponibles dans les tables de données |  |
| Q27 : Une visualisation de données est un processus nécessitant plusieurs étapes de construction              |  |



| Section / Question  | Options / Notes  |
|---|--|
| Q28 : Quelles sont les difficultés que j'ai rencontré pendant l'activité? | <ul style="list-style-type: none"> <li>• La partie théorique</li> <li>• La définition d'une visualisation</li> <li>• Les types de dataset</li> <li>• Les types de données</li> <li>• Les types d'attributs</li> <li>• La construction d'une visualisation à partir d'un énoncé</li> <li>• Le diagramme UML de la base de données</li> <li>• Le nettoyage de la table de données (nom des colonnes, spécification des attributs, etc.)</li> <li>• Aucune</li> </ul> |
| Q29 : Confiance dans la compréhension de l'ensemble de données            | Echelle Likert, de 1 à 5, 1 pas du tout d'accord, 5 tout à fait d'accord   |

TABLE 4 – Questionnaire Pre et Post-test

## B Annexes B - Focus Group et Interview EAFC Cadet Namur 31/03/2023 - Document préparatoire

### Focus Group et Interview EAFC Cadet Namur 31/03/2023

Lieu : EAFC Namur Cadets, Place de l'Ecole des Cadets,6, 5000 NAMUR, 3ième étage, local C313

Date : 31/03/2023

### Focus Group

Type : Échantillon homogène et par questionnement

Horaire : De 18h10 à 19h10

Interlocuteur(s) : Échantillon ayant participé à l'activité d'enseignement

Objectif : S'orienté sur les exigences des étudiants sur comment est qu'ils veulent être introduit à la visualisation de données, qu'est ce qu'ils veulent comme cours de visualisation.

Autre : Le focus group sera enregistrer via une vidéo smartphone

## Thèmes et Questions

|   |  |
|---|--|
| <p>Apprendre avec leurs connaissances actuelles →<br/>Pour consolider</p> | <ul style="list-style-type: none"> <li>• Comment ont-ils ressenti l'activité ?</li> <li>• Ce qu'ils ont apprécié</li> <li>• Ce qu'ils n'ont pas apprécié ?</li> <li>• Ce qu'il faut changer ?</li> <li>• Est-ce que c'est possible de leurs apprendre la visualisation de données avec leurs backgrounds en prenant l'angle de la base de données ?</li> <li>• Est-ce qu'ils voient un intérêt à la visualisation de données ?</li> <li>• Est-ce que cet intérêt est plus fort après l'activité ?</li> </ul> |
| <p>Exigences étudiants</p>  | <ul style="list-style-type: none"> <li>• Veulent-ils l'apprendre ?</li> <li>• Uniquement de la théorie ? Pratique ? Mixte des deux ?</li> <li>• Comment voudraient-ils que le cours soit construit ?</li> <li>• Dans quel cadre (dans le cours de DB, un autre ou plusieurs, ou dans un nouveau cours) ?</li> <li>• Quelle fréquence ? Pendant un semestre, une fois semaine, une heure semaine ?</li> </ul>   |

## Interview

Type : semi-directif à réponses libres

Horaire : De 17h00 à 17h40

Interlocuteur(s) : Michel Voneche, enseignant du cours de modélisation de données pour les bacheliers et des cours de programmation niveau 1 et 2 pour les techniciens en programmation

Objectif : Recueillir des informations complétant la récolte de données et la vision de l'enseignant sur cette dernière

Autre : L'interview sera enregistrée (audio) sur smartphone

## Thèmes et Questions

|                               |  |
|-------------------------------|--|
| Avant l'activité              | <ul style="list-style-type: none"> <li>• A-t-il eu un intérêt de base pour l'introduction de la visualisation de donnée afin de l'intégrer à l'enseignement de la modélisation de données ?</li> <li>• A-t-il constaté des retours sur la visualisation de données de la part de ses étudiants avant l'activité ?</li> </ul>   |
| Pendant l'activité            | <ul style="list-style-type: none"> <li>• Comment a-t-il ressenti l'activité ?</li> <li>• Ce qu'il a apprécié</li> <li>• Ce qu'il n'a pas apprécié ?</li> <li>• Ce qu'il faut changer ?</li> <li>• A-t-il constaté des difficultés particulières pendant l'activité ? Pour lui ou pour les étudiants</li> </ul>   |
| Après l'activité (étudiants)  | <ul style="list-style-type: none"> <li>• A-t-il constaté une évolution ou des retours de ses étudiants dans les premiers cours suivis après l'activité</li> </ul>  |
| Après l'activité (professeur) | <ul style="list-style-type: none"> <li>• Est-ce que, en tant que professeur, l'intérêt de la visualisation données est-il présent, si oui, dans quelle proportion ?</li> <li>• Est-il intéressant d'intégrer les concepts de visualisation de données dans le cursus scolaire des étudiants en bachelier et technicien en informatique ? Si oui, quel format ? Quelle fréquence ?</li> </ul> |