

OI

# MATTHIAS *Portfolio* VIET

2024 / 2025

# SOMMAIRE

01

PRESENTATION

02

POURQUOI BUT  
SD

03

LES PROJETS

04

MES OBJECTIFS

05

CE DONT JE SUIS  
LE PLUS FIÈR

02



# MATHIS VITET

présentation personnelle

## QUI SUIS-JE?

03

Je suis étudiant en première année de BUT Science des Données à l'Université Sorbonne Paris Nord. C'est une formation qui m'a attiré pour son équilibre entre mathématiques, informatique, analyse statistique et ouverture vers les enjeux du monde réel. J'ai toujours été curieux de comprendre comment on peut extraire des informations pertinentes à partir d'un grand volume de données, et cette formation m'a permis de découvrir à quel point les données peuvent être un levier d'action puissant dans des domaines variés comme la santé, l'environnement ou l'économie.

J'apprécie particulièrement les projets concrets que nous réalisons à travers les SAE, car ils nous permettent de mettre en pratique nos connaissances sur des cas réalistes. J'aime aussi manipuler des bases de données, écrire des requêtes SQL, créer des visualisations de données claires et explorer les possibilités qu'offrent des outils comme R ou Python.

Je souhaite à terme devenir data analyst ou data scientist, dans un domaine qui a du sens pour moi, comme la protection de l'environnement ou la santé publique. Je suis aussi ouvert à d'autres parcours professionnels comme l'ingénierie statistique ou le conseil en stratégie basée sur les données.

Dans cet objectif, je m'efforce de progresser en programmation, en statistiques et en visualisation. J'essaye également d'améliorer mes compétences en communication et en présentation orale, car je sais que savoir transmettre des résultats est aussi important que savoir les produire. Enfin, je commence à m'informer sur les opportunités de stages ou d'alternance dès la 2e année, pour mieux me projeter dans le monde professionnel.

## Mes compétences

Compétence technique ou transversale	Exemple concret issu de mon année
→ Programmation en R	Création de graphiques sur la qualité de l'air en Île-de-France dans le cadre d'un projet Tableau
→ SQL (interrogation de bases)	Rédaction de requêtes pour analyser des exportations bas carbone dans une SAE
→ Analyse statistique	Calcul d'estimateurs et d'intervalles de confiance en R (cours R2.08)
→ Visualisation et storytelling	Présentation d'un rapport clair sur les données de pollution aux particules fines
→ Travail d'équipe / Gestion de projet	Répartition des tâches et synthèse collective pour les SAE semestrielles
→ Veille technologique	Lecture régulière de blogs comme <i>DataCamp</i> , <i>StatQuest</i> , et <i>Towards Data Science</i>

## Mes qualités

- Curieux : J'aime comprendre en profondeur les concepts, et je vais souvent au-delà du cours pour lire ou expérimenter par moi-même (ex. : apprentissage personnel de notions de machine learning en Python).
- Autonome : Je suis capable de travailler seul sur des projets complexes, comme lors du projet sur la qualité de l'air où j'ai géré toute l'analyse des données PM2.5.
- Rigoureux : Je porte attention aux détails dans les formules statistiques et les scripts, et je relis toujours mes résultats avant de les communiquer.

## 2. POURQUOI LA SCIENCE DES DONNÉES ?

La science des données m'attire car elle se trouve à l'intersection entre mathématiques, informatique et réflexion stratégique. J'ai choisi ce domaine parce qu'il allie rigueur et créativité, et parce qu'il permet de transformer de simples chiffres en connaissances utiles à la prise de décision. Aujourd'hui, la donnée est au cœur de tous les secteurs — santé, environnement, marketing, transports, économie — et c'est précisément cette transversalité qui me motive.

J'ai découvert la data science assez récemment, en explorant des vidéos et articles sur l'intelligence artificielle et les statistiques. Très vite, j'ai voulu comprendre ce qu'il y avait "derrière les algorithmes".

C'est ce besoin de comprendre les mécanismes de prédiction, de détection de tendances et d'optimisation qui m'a amené à choisir le

BUT Science des Données.

Depuis que j'ai commencé la formation, j'apprécie tout particulièrement :

- l'analyse de données à l'aide de R et Python,
- l'organisation et l'exploration de bases de données avec SQL,
- et la visualisation des résultats pour les rendre compréhensibles et utiles, par exemple à travers l'outil Tableau.

Ce que j'aime dans cette discipline, c'est qu'elle demande à la fois de la logique, de la curiosité, de la rigueur et un vrai sens critique. Il ne s'agit pas seulement de produire des chiffres, mais de les interpréter, de les mettre en contexte et d'en faire émerger du sens. Je m'efforce également de faire de la veille technologique, car c'est un domaine qui évolue très rapidement. Voici quelques

ressources que je consulte régulièrement :

-  Towards Data Science (blog sur Medium)
-  Chaîne YouTube StatQuest (excellente pour comprendre les concepts statistiques)
-  DataCamp et Kaggle pour les tutoriels interactifs
-  Articles de vulgarisation sur Sciences et Avenir ou Futura Sciences

Ces lectures et explorations m'aident à élargir mes connaissances au-delà du programme, à me familiariser avec des outils comme Pandas, Scikit-learn ou les bases du machine learning, et à mieux comprendre les enjeux éthiques liés à la donnée.

La science des données est pour moi bien plus qu'une discipline : c'est un outil puissant pour mieux comprendre le monde et agir avec pertinence. C'est cette ambition — comprendre pour agir — qui me pousse à m'investir pleinement dans cette voie.

# PROJET I



## Résumé du projet :

Dans cette SAÉ, nous avons travaillé sur une base de données contenant des informations sur les catastrophes climatiques à l'échelle mondiale. Après avoir installé PostgreSQL, Metabase et les outils nécessaires, nous avons importé les données dans une base relationnelle. L'objectif principal était de réaliser des requêtes SQL pertinentes, puis de créer des visualisations dans Metabase afin d'interpréter les tendances climatiques mondiales et régionales.

## Points de vigilance : lacunes à combler

- J'ai parfois eu des difficultés à bien modéliser les relations entre les tables, notamment dans les jointures à plusieurs niveaux (country, sub\_region, region). Cela a entraîné des erreurs de regroupement ou de doublons.
- J'ai mis du temps à bien structurer mes requêtes, notamment à cause de l'ordre des opérations (GROUP BY, ORDER BY) ou de l'utilisation des alias.
- Objectifs : revoir la logique relationnelle entre entités dans une base, m'entraîner à écrire des requêtes plus lisibles et optimisées, et m'exercer régulièrement avec des bases différentes pour améliorer ma réactivité et ma rigueur.



## Points positifs : bonnes pratiques acquises

- J'ai appris à écrire des requêtes SQL complexes avec jointures multiples (JOIN) pour regrouper et structurer des données à partir de plusieurs tables (régions, pays, types de catastrophes...).
- J'ai développé une bonne pratique de formulation progressive des requêtes, en testant étape par étape pour identifier d'éventuelles erreurs de logique ou de syntaxe.
- J'ai aussi découvert Metabase, un outil de datavisualisation que j'ai trouvé intuitif et puissant pour créer des graphiques (lignes, barres) et pour présenter visuellement les résultats de mes analyses.
- Ce projet m'a permis de lier des compétences techniques (SQL, Metabase) à des enjeux d'actualité (climat), ce qui a renforcé mon intérêt pour les applications concrètes de la science des données.

## Work in progress : progrès à accomplir

- Savoir-faire à développer : améliorer la qualité de mes requêtes SQL en utilisant des sous-requêtes ou des vues, gagner en rapidité dans la rédaction, approfondir l'utilisation de Metabase (filtres dynamiques, dashboards).
- Savoir-être à renforcer : être plus méthodique et organisé dans ma démarche (prendre des notes, schématiser la base avant d'écrire), et mieux anticiper les erreurs pour ne pas rester bloqué.
- Je souhaite encore progresser sur la lecture critique des résultats (analyser ce que montre un graphique et ce qu'il cache) et sur la création de rapports de synthèse exploitables dans un cadre professionnel.

# PROJET 2

06

## Résumé du projet :

Dans cette SAÉ, nous devions effectuer une recherche de données ouvertes sur un territoire donné afin d'en extraire les caractéristiques économiques, démographiques et culturelles les plus importantes. Ce travail a abouti à une présentation orale en anglais, appuyée par un poster numérique. L'objectif était de s'entraîner à restituer une synthèse claire et pertinente de données dans une langue étrangère, tout en respectant les contraintes de la communication professionnelle.

## Points positifs : bonnes pratiques acquises

- Ce projet m'a permis de progresser en expression orale en anglais, en apprenant à présenter des données chiffrées et des indicateurs de manière claire, fluide et structurée.
- J'ai appris à sélectionner l'information pertinente dans un grand volume de données territoriales, ce qui est très utile dans un contexte de data science.
- J'ai aussi développé des compétences en travail d'équipe, notamment pour répartir les tâches, harmoniser notre discours et concevoir un visuel cohérent (poster).
- Ce projet m'a sensibilisé à l'importance de la vulgarisation : savoir adapter son langage à son public, même dans une langue étrangère.

**Provence-Alpes-Côte d'Azur (THE FRENCH RIVIERA)**

**Economy and Key Industries**

**Perfume industry**

**Technology and innovation : Sophia Antipolis**

**Housing**

**Agriculture and local products**

**Challenges and future outlooks**

**Sustainable development & tourism management**

**Urbanization & environmental preservation**

**Attractionness for investors and residents**

**Répartition des projets**

Région	Nombre de projets en cours	%
Île-de-France	207	44%
Provence-Alpes-Côte d'Azur	75	16%
Centre	60	13%
Grand Est	58	13%
Hauts-de-France	57	13%
Bretagne	10	2%
Occitanie	10	2%
Normandie	10	2%
Grande-Côte d'Or	4	1%
Pays de la Loire	4	1%
Autres	1	0.2%
Total	466	100%

## Points de vigilance : lacunes à combler

- J'ai rencontré des difficultés sur certains termes techniques en anglais, notamment pour traduire correctement des indicateurs économiques ou sociaux.
- Lors de la soutenance, j'ai eu du mal à gérer le stress et à répondre de manière fluide aux questions en anglais.
- Objectifs : renforcer mon vocabulaire spécialisé en anglais (via Quizlet, glossaires, articles spécialisés) et m'entraîner davantage à l'oral (exposés courts, répétitions, auto-enregistrement pour travailler l'intonation).

## Work in progress : progrès à accomplir

- Savoir-faire : améliorer mes compétences en présentation visuelle (design de poster, mise en page), progresser en prononciation et en fluidité lors d'interventions orales.
- Savoir-être : apprendre à parler avec plus d'assurance face à un public, même dans une langue étrangère, et mieux gérer le travail en groupe sous contrainte de temps.
- Il me reste aussi à développer une routine de veille anglophone sur des sujets liés à la data science, pour enrichir mon vocabulaire et m'habituer à entendre/parler de données en anglais.

# PROJET 3

## Résumé du projet :

Dans cette SAÉ en lien avec l'économie, nous avons étudié la Coopérative Bio d'Île-de-France. Le but était d'analyser son modèle économique et ses performances à partir de données réelles, pour évaluer si elle répond aux enjeux sociaux, environnementaux et économiques à travers les circuits courts. Nous avons mené une étude complète : contexte, marché, performances, limites, et propositions d'amélioration.

### Points positifs : bonnes pratiques acquises

- J'ai appris à analyser une organisation réelle à partir de données issues de sources diverses (rapports officiels, sites spécialisés, données chiffrées), ce qui m'a permis de faire le lien entre théorie et réalité économique.
- J'ai renforcé mes compétences en recherche documentaire et en synthèse d'information, deux atouts essentiels pour les futurs projets de data science.
- J'ai mieux compris le fonctionnement d'un modèle coopératif et son rôle dans les circuits courts, ce qui m'a donné une vision plus concrète des enjeux sociaux de la donnée.
- Ce projet m'a aussi permis de travailler ma capacité d'analyse critique, en ne me contentant pas des chiffres positifs mais en intégrant aussi les limites et les perspectives d'amélioration.



# PROJET 3

09

## Points de vigilance : lacunes à combler

- J'ai parfois eu des difficultés à structurer mes idées clairement dans les parties les plus longues du rapport, notamment pour organiser logiquement les informations économiques, sociales et environnementales.
- L'analyse des indicateurs financiers et sociétaux (comme le chiffre d'affaires ou les créations d'emplois) aurait pu être plus rigoureuse si j'avais mieux maîtrisé certains concepts.
- Objectifs : m'entraîner à structurer des rapports analytiques plus efficacement, approfondir mes connaissances en lecture de données économiques, et apprendre à utiliser Excel ou R pour représenter ces indicateurs plus visuellement.

## Work in progress : progrès à accomplir

- Savoir-faire à développer : mieux utiliser les données chiffrées pour construire des arguments solides, créer des tableaux de bord simples (Excel/Tableau) pour rendre mes analyses plus lisibles, et automatiser certaines tâches d'analyse.
- Savoir-être à renforcer : apprendre à m'adapter rapidement à un nouveau domaine professionnel (ici, l'économie sociale et solidaire), développer ma rigueur dans la formulation d'analyses, et améliorer mes réflexes de travail collaboratif dans des projets à long terme.
- Je souhaite aussi progresser dans la capacité à proposer des pistes d'amélioration concrètes à partir des limites observées, en liant données et solutions possibles.

<https://youtu.be/jW6kqLzzwiQU>

# PROJET 4

## Résumé du projet :

Dans cette SAÉ, nous devions construire un tableau de bord interactif dans Excel à partir d'un fichier d'étudiants anonymisé, afin d'aider le service communication de l'IUT à mieux comprendre le profil des étudiants par lycée, sexe, baccalauréat, bourse, etc. Cela impliquait la structuration des données, la création de TCD, segments, filtres dynamiques, et l'utilisation de Power Query pour nettoyer et enrichir les données.



**Power Query**



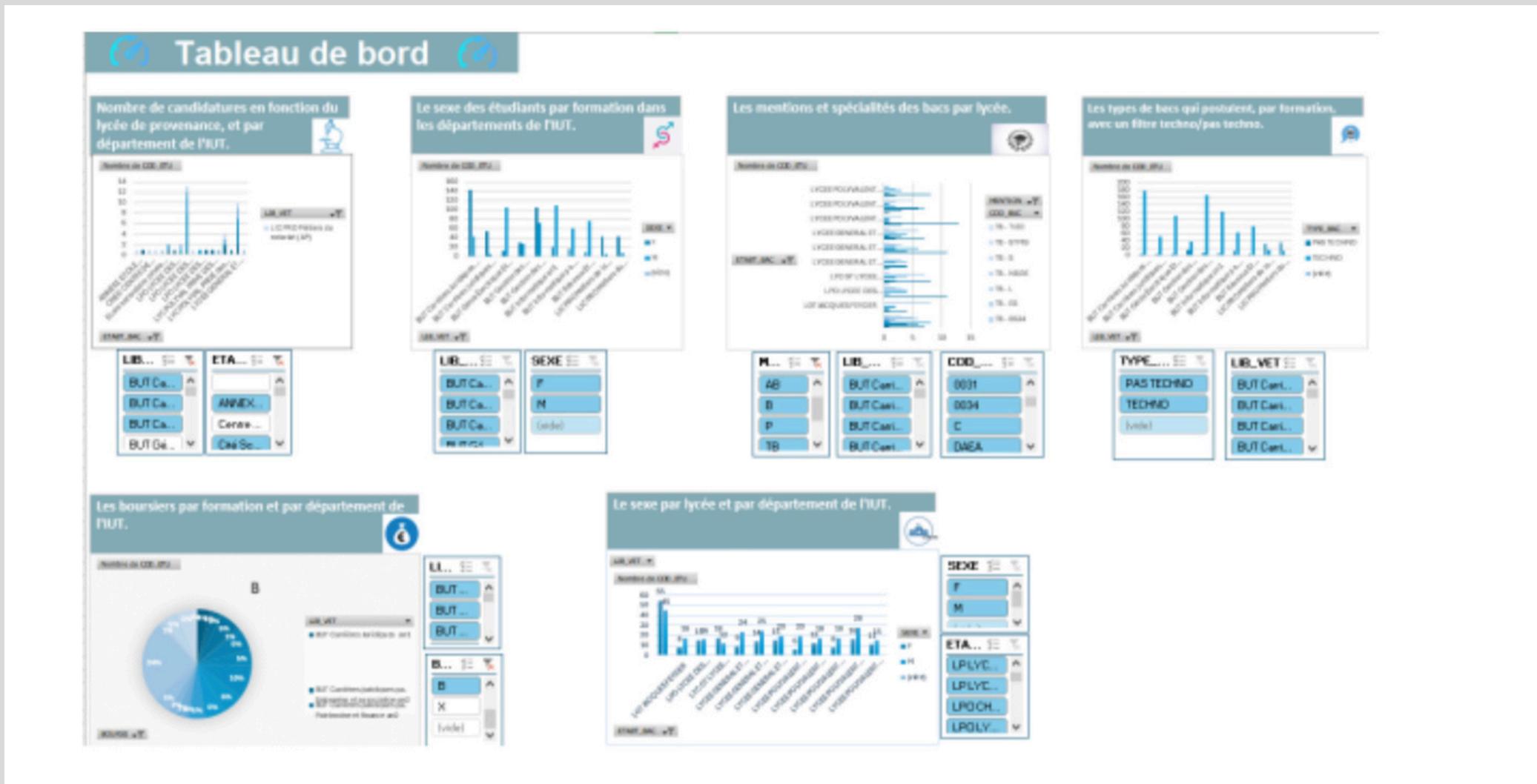
**Excel**

Points positifs : bonnes pratiques acquises

- J'ai appris à utiliser Power Query pour nettoyer, transformer et structurer des données automatiquement.
- J'ai su construire des tableaux croisés dynamiques clairs et exploitables, avec des libellés bien formatés et des filtres pertinents.
- J'ai aussi pris l'habitude de travailler proprement, en renommant mes feuilles, mes champs, et en soignant l'esthétique du tableau final.

## Points de vigilance : lacunes à combler

- Je dois encore progresser sur la logique des relations entre tables et la création de modèles de données dans Excel.
- J'ai parfois eu du mal avec des erreurs techniques (problèmes de types de données, formules non actualisées).
- Objectif : approfondir l'utilisation de Power Query et Power Pivot, et créer une fiche de bonnes pratiques Excel pour éviter de répéter certaines erreurs.



## Work in progress : progrès à accomplir

- Savoir-faire à développer : maîtriser l'utilisation des relations entre tables, l'ajout de colonnes calculées via DAX, et la création de cartes interactives dans Excel.
- Savoir-être : être plus rigoureux dans la vérification des résultats (ne pas se fier uniquement à la première impression visuelle), et apprendre à tester toutes les combinaisons de filtres possibles.
- Je souhaite encore apprendre à concevoir des tableaux de bord plus évolués, avec des graphiques combinés, des menus déroulants et des indicateurs synthétiques.

# PROJET 5

## Résumé du projet :

Dans cette SAÉ, nous avons participé au Challenge national DataViz du 23 mai 2024.

En groupe, nous avons travaillé sur un jeu de données fourni par l'INSEE sur la scolarisation des jeunes de 15 à 29 ans selon les départements. L'objectif était de produire une infographie visuelle synthétique, en analysant les évolutions spatiales, temporelles et thématiques. Le projet s'est conclu par une présentation orale de 5 minutes devant un jury, après une journée de challenge.

Points positifs : bonnes pratiques acquises

- J'ai appris à structurer un projet de datavisualisation de A à Z : choix des données, indicateurs pertinents, graphiques adaptés, mise en page finale.
- J'ai acquis des réflexes de clarté visuelle : choix des couleurs, hiérarchie des éléments, légendes efficaces.
- J'ai amélioré ma capacité à travailler sous pression, à communiquer dans un temps limité et à écouter les conseils d'un jury pour m'améliorer.



## Work in progress : progrès à accomplir

- Savoir-faire à développer : créer des visualisations interactives, apprendre les bonnes pratiques UX en visualisation, mieux exploiter les typologies (flux, temporalité, disparité...).
- Savoir-être : être plus rapide dans la prise de décisions collectives, gagner en confiance pour proposer des idées visuelles originales, et accepter la critique constructive pour progresser.
- Je veux aussi apprendre à adapter mes visualisations à différents publics : experts, grand public, décideurs — en ajustant le niveau de complexité et le vocabulaire utilisé.

## Points de vigilance : lacunes à combler

- Je dois encore progresser sur l'analyse exploratoire rapide des données, pour gagner du temps dans l'identification des indicateurs utiles.
- Je n'ai pas toujours su évaluer la pertinence de certains graphiques (ex. : graphiques trop chargés ou mal équilibrés).
- Objectifs : m'exercer à construire des dataviz claires en moins d'une heure, approfondir les outils comme Tableau, Flourish, ou Power BI, et étudier des exemples professionnels.



# 4. Mes objectifs

14

01

## Objectifs à court terme (stage en BUT 2 / alternance)

Dès maintenant, je me fixe pour objectifs :

- Trouver un stage pour le BUT 2 dans une structure qui utilise les données pour des finalités concrètes : collectivité locale, entreprise du secteur bio, santé ou transport.
- Me constituer un portfolio solide (comme celui-ci) avec mes projets les plus parlants, accompagnés de visualisations, extraits de code et réflexions personnelles.
- Approfondir mes compétences en Python, SQL, dataviz (Tableau, Power BI), et commencer à me former à des outils avancés (notamment des bases en machine learning ou des API).
- M'habituer à répondre à des offres, préparer des lettres de motivation, des entretiens, avec l'aide des ateliers ou du service orientation de l'IUT. Je sais que la clarté des objectifs est un moteur de motivation, et je suis prêt à ajuster mes choix en fonction des expériences concrètes que je vivrai dans les mois à venir.

02

## Objectifs à moyen terme (formation post-BUT, premier emploi)

À l'issue du BUT Science des Données, plusieurs voies me semblent possibles :

- Poursuivre en master, notamment un Master Data Science, Statistique publique, ou un master spécialisé en environnement, santé ou économie des données.
- Ou bien intégrer une école d'ingénieur en admission parallèle (comme ENSAI, INSA, IMT Atlantique), pour renforcer mes compétences mathématiques et en machine learning.
- Dans tous les cas, mon objectif est de renforcer mon profil technique tout en me spécialisant progressivement, en fonction des secteurs qui m'attirent le plus.

En parallèle, je serai attentif aux opportunités d'emploi de type chargé d'études junior ou assistant data analyst pour démarrer dans le monde professionnel dès la sortie du BUT si nécessaire.

# 03

## Objectifs à long terme (dans 10 à 15 ans)

Dans une dizaine d'années, je me vois travaillant dans un domaine qui utilise les données pour répondre à des enjeux concrets et utiles pour la société, comme la santé publique, la transition écologique ou encore la mobilité urbaine.

Je souhaite exercer un métier tel que data analyst, data scientist ou chargé d'études statistiques, au sein d'une entreprise, d'une collectivité ou d'une organisation internationale.

Au-delà de la technique, j'aimerais occuper un rôle qui combine l'analyse, la visualisation, la communication et l'aide à la décision, afin de produire de l'impact réel grâce aux données.



# Ce dont je suis le plus fier...

Ce dont je suis le plus fier cette année, c'est ma capacité à progresser en partant de zéro sur des compétences techniques complexes, et à livrer des projets concrets, aboutis, et utiles.

En particulier, je suis fier de deux expériences marquantes :

## Le Challenge DataViz (SAÉ 2.04)

Participer au Challenge national DataViz a été un véritable défi, car nous avions très peu de temps pour comprendre les données, construire un message visuel, puis le défendre à l'oral devant un jury.

Ce jour-là, j'ai réussi à :

- collaborer efficacement en équipe sous pression,
- produire une infographie claire et cohérente sur un sujet complexe (la scolarisation des jeunes), et surtout, à présenter ce travail à l'oral avec confiance, malgré le stress.
- C'est une vraie victoire personnelle et une expérience que je n'aurais pas cru possible il y a quelques mois.

Mon autonomie dans les projets techniques (R, SQL, Excel)

Je suis aussi fier d'avoir su :

- écrire mes premières vraies requêtes SQL complexes avec jointures,
- automatiser le traitement de fichiers clients en R,
- et concevoir un tableau de bord complet dans Excel avec Power Query, sans copier-coller, en reliant les bonnes sources de données.

Ces projets m'ont prouvé que je pouvais m'autoformer, chercher des solutions, persévérer, et trouver du plaisir dans le processus d'apprentissage.

E-MAIL

[mathis.vitet95@gmail.com](mailto:mathis.vitet95@gmail.com)

MOBILE

07 69 18 26 41