

# Samira MAMANE SIDI

Étudiante en Master 2 Statistique et Sciences des Données

Montpellier, France | [samiramamanesidi2050@gmail.com](mailto:samiramamanesidi2050@gmail.com) | 07 66 60 50 44  
[linkedin.com/in/samira-mamane-sidi-3508ba271](https://www.linkedin.com/in/samira-mamane-sidi-3508ba271)

## PROFIL PROFESSIONNEL

---

Étudiante en double Master 2 (Statistique & Sciences des Données + Statistique pour l'Information et l'Aide à la Décision) avec une base solide en modélisation statistique, machine learning et analyse de données. Recherche un stage de 6 mois pour appliquer mes compétences en Python, R, SQL et statistiques avancées afin de transformer des données complexes en insights business actionnables. Double cursus alliant rigueur analytique et compréhension des enjeux décisionnels en entreprise.

## FORMATION

---

<b>Master 2 - Statistique &amp; Sciences des Données (SSD)</b> Faculté des Sciences, Université de Montpellier <i>Spécialisation : Modélisation statistique, machine learning, analyse de données</i>	2025 - 2026
<b>Master 2 - Statistique pour l'Information et l'Aide à la Décision (SIAD)</b> IAE de Montpellier <i>Double diplôme axé sur l'aide à la décision et le business analytics</i>	2025 - 2026
<b>Master 1 - Statistique &amp; Sciences des Données</b> Faculté des Sciences, Université de Montpellier	2024 - 2025
<b>Licence en Mathématiques</b> Faculté des Sciences, Université de Montpellier	2021 - 2024
<b>Classes Préparatoires aux Grandes Écoles</b> Université Tunisienne Internationale de Niamey (UTI)	2020 - 2021
<b>Baccalauréat Scientifique - Mention Très Bien</b> Lycée d'Excellence de Niamey	2020

## PROJETS ACADEMIQUES ET TECHNIQUES

---

### Prédiction d'Octroi de Prêt Bancaire | IAE Montpellier

<i>Machine Learning, Classification, Gestion de Déséquilibre</i>	2025
— Analysé un jeu de données déséquilibré pour prédire l'encaissement de prêts bancaires avec des algorithmes de machine learning	
— Appliqué des techniques de rééchantillonnage (SMOTE) pour traiter le déséquilibre de classes et améliorer les performances prédictives	
— Développé et évalué plusieurs modèles de classification (Random Forest, régression logistique) avec optimisation des hyperparamètres	
— Présenté les résultats et recommandations business pour l'aide à la décision en contexte bancaire	

### Étude Comparative de Méthodes de Régression | Université de Montpellier

<i>Régression PLS, Ridge, Lasso, ACP</i>	2025
— Implémenté et comparé des méthodes de régression avancées (PLS, Ridge, Lasso) sur un jeu de données réel avec R	

- Réalisé une analyse en composantes principales (ACP) pour réduire la dimensionnalité et identifier les variables clés
- Évalué les performances des modèles via cross-validation et métriques statistiques (RMSE, R<sup>2</sup>)
- Rédigé un rapport technique détaillé documentant la méthodologie, les résultats et les conclusions statistiques

## Analyse Questionnaire Mayotte | Université de Montpellier

*Analyse Statistique, Régression Linéaire Multiple*

Février - Mai 2025

- Nettoyé et préparé un ensemble de données issues de questionnaires avec gestion des valeurs manquantes et aberrantes
- Conduit des tests statistiques (ANOVA, tests de Student) pour identifier les différences significatives entre groupes
- Développé des modèles de régression linéaire multiple pour identifier les facteurs de réussite en mathématiques
- Visualisé les résultats avec ggplot2 et présenté les insights pour orienter les décisions pédagogiques

## BikeProject - Analyse du Trafic Cycliste | Université de Montpellier

*Analyse de données, Visualisation Interactive*

Septembre - Décembre 2024

- Analysé l'impact du Tour de France 2023 sur le trafic cycliste avec des données avant/après événement
- Étude de l'impact de l'événement sur l'utilisation des vélos.
- Prédiction des trajets les plus empruntés par les cyclistes sur une semaine via modèles de Machine Learning .
- Traitement et visualisation des données avec Python, Pandas, Matplotlib pour identifier tendances et comportements.

## COMPÉTENCES TECHNIQUES

---

**Langages de Programmation :** Python (NumPy, Pandas, Scikit-learn, Matplotlib), R (dplyr, ggplot2, caret), SQL

**Machine Learning & IA :** Random Forest, Régression Logistique, SVM, Rééchantillonnage (SMOTE), Cross-validation

**Statistiques Avancées :** Régression linéaire multiple, Régression PLS/Ridge/Lasso, Tests d'hypothèses, Analyse en Composantes Principales (ACP), Séries temporelles

**Visualisation & Business Intelligence :** Power BI, Matplotlib, ggplot2,

**Outils & Méthodes :** Git/GitHub, Jupyter Notebook, RStudio, LaTeX, Méthodologie de projet Data

## COMPÉTENCES TRANSVERSALES

---

Rigueur analytique | Organisation | Autonomie | Travail en équipe | Pédagogie | Curiosité scientifique | Sens de l'analyse | Communication des résultats techniques

## LANGUES

---

**Français :** Langue maternelle | **Anglais :** Intermédiaire (B1) | **Japonais :** Débutant (A2)

## CENTRES D'INTÉRÊT

---

Lecture | Pâtisserie | Apprentissage de langues | Voyage