

Innocent BIGIRIMANA

DATA ANALYST/SCIENTIST EN ALTERNANCE

🏠 Toulouse
✉ bigirimainnocent596@gmail.com
☎ 07 58 13 31 42
📍 Cholet

COMPÉTENCES

Analyse des données

- Collecter, nettoyer et traiter les données
- Identifier des tendances et des opportunités à partir des données

Compétences SQL avancées et gestion de bases de données

API Rest

Modélisation statistique avancée

- Analyses statistiques avancées
- Création de modèles prédictifs pour optimiser les processus métier.
- Machine Learning

Création de visualisations

- Développer des dashboards interactifs (Power BI).
- Présenter des insights clairs et exploitables aux équipes métier.

INFORMATIQUES

- RStudio
- SQL
- Power BI
- Pack Office
- Python: Numpy, Pandas, Scikit-Learn, Scipy, statsmodels, Matplotlib,...

CERTIFICATS UDEMY

- Data Science et Machine Learning | MasterClass Python: 45h
- Devenir un expert de Power BI: 47h

QUALITÉS

- Esprit d'analyse et de synthèse
- Motivé
- Organisé
- Autonome et Force de proposition

LANGUES

Français

Swahili

Anglais

CENTRES D'INTÉRÊT

- Bénévolat
- Musique
- Sport
- IA

RÉSEAUX SOCIAUX

in @bigirimana-innocent
🐦 @bigirimainnocent12

Passionné par la transformation des données en insights stratégiques. Fort d'une expertise approfondie en analyse de données, visualisation et Machine Learning, je maîtrise l'art de révéler les tendances cachées et de traduire les données complexes en recommandations concrètes pour optimiser la prise de décision business.

EXPÉRIENCES PROFESSIONNELLES

Data Analyst (Stage)

De février 2025 à août 2025 Lallemand SAS Blagnac

- Développer une architecture de données centralisée avec API REST.
- Migrer et structurer 12 000 fichiers Excel dans un data warehouse SQL Server optimisé, réduisant les temps de requête de 90%.
- Développer une application Python/Streamlit d'analyse automatisée avec dashboard interactif pour le traitement des données de fermentation.
- Créer un tableau de bord Power BI avec KPIs métier, améliorant la prise de décision.
- Réaliser des analyses statistiques avancées (ANOVA, T-tests, corrélations) pour identifier les facteurs critiques de performance fermentaire.
- Développer et déployer 2 modèles de Machine Learning prédictifs avec 85% de précision pour anticiper les cinétiques de fermentation.
- Modéliser l'impact des variables environnementales et des souches de levures sur les rendements, optimisant les processus de production.

Statisticien (Stage)

De mai 2022 à août 2022 Institut National de la Statistique du Burundi (INSBU)

- Manipuler les données issues des enquêtes pour en assurer la qualité.
- Visualiser les tendances et relations à travers des représentations graphiques.
- Analyser statistiquement les informations pour en extraire des informations pertinentes.
- Développer un modèle prédictif afin d'anticiper les valeurs futures de l'indice des prix à la consommation du Burundi.

DIPLOMES ET FORMATIONS

Mastère 2 Data Science

De septembre 2025 à septembre 2026 Kaischool Paris

Master Mathématiques et applications parcours Statistique

De septembre 2023 à août 2025 Université de Strasbourg

Cours principaux: GLM, Régression linéaire, Série temporelle, Tests Statistiques, Analyse de données avancées, base de données, Statistique bayésienne, Machine Learning, Réseaux de neurones...

Licence en probabilité et statistique

D'octobre 2019 à octobre 2022 Université du Burundi

Cours principaux: Mathématiques actuarielles vie et IARD, Actuariat, Econométrie I et II, Analyse de la qualité de données, Traitement des données d'enquêtes, R, logiciel statistique (Stata), Modélisation statistique, Modélisation et prévision...

Diplôme d'État (Bac S)

De septembre 2017 à juillet 2018 Lycée Saint Marc

RÉALISATIONS

Depuis septembre 2023

- Classification des prix des téléphones (Régression logistique):**
Réalisation d'un modèle de régression logistique pour prédire la gamme de prix de téléphones portables (4 classes) à partir de 21 caractéristiques techniques (RAM, batterie, écran, connectivité...). Le projet a inclus la préparation et l'analyse exploratoire des données, la séparation des variables qualitatives et quantitatives, l'évaluation via précision et F1-score, et l'identification des variables les plus influentes sur la classification.
- Prédiction des Dépenses Médicales en Assurance Maladie:**
Modélisation avec RandomForestRegressor pour estimer les frais médicaux à partir de données socio-démographiques et de santé (27,000 lignes, 7 variables). Analyse exploratoire, détection des facteurs clés (tabagisme, IMC), entraînement et optimisation du modèle avec validation croisée, puis évaluation (R2: 0,87; MAPE: 12%).