

Daniel RAKOTOMALALA, PhD

*Thèse en économie appliquée, Data scientist (niveau R avancé), Chef de projet
Data en administration publique*

rakdanielh@gmail.com

+ 33 782 50 10 67

<https://github.com/danielrak>



Expériences professionnelles

Data scientist, chef de projet Innovations, Données et Expérimentations en Education (IDEE) – depuis 2024

*Direction de l'évaluation, de la prospective et de la performance (DEPP) - Ministère
chargé de l'Education nationale*

- Etude statistique, en collaboration avec le Ministère de la Justice
 - Objet : analyse des parcours scolaires d'un échantillon particulier de jeunes
 - Opérations massives de construction de données, notamment en panel
 - Calcul d'indicateurs socio-démographiques et de parcours éducatif
 - Restitutions régulières d'avancement (notes et réunions à l'intention des partenaires)
- Traitements statistiques à grande échelle : problèmes complexes, expérimentations de solutions, solutions reproductibles et documentées
 - *Pipeline* de mise à disposition des données aux équipes de recherche conventionnées. Développement d'un processus automatique permettant à une petite équipe de gérer techniquement des demandes régulières sur un centaine de projets de recherche
 - Contributions à la préparation de l'arrêt du logiciel SAS et au passage à R de toute l'institution. Ex. conversion de format (SAS → parquet) de ~4500 fichiers. Fiabilisation systématique des intrants et sorties
 - Cadrage technique du processus du [Code Statistique Non Signifiant](#) (CSNS) pour l'institution
- Outillage statistique pour la DEPP (packages, fonctions, programmes de routine, etc.) : maintenance, amélioration et innovation
 - Traitements de données structurées (le design minimise le lien entre taille des inputs et lourdeur des opérations) – cf. package CRAN [scrutr](#)
 - Package de chiffrement en masse des données administratives
 - *Matching* probabiliste du nom des classes entre données d'enquêtes chercheurs et données administratives
- Formateur technique pour la DEPP : modélisation statistique
- Chef de projet IDEE
 - Suivi continu du [partenariat](#) entre la DEPP et IDEE
 - Collaboration avec des équipes métiers transversales et des personnels de direction
 - [Catalogue public des données](#)
 - Plateforme sécurisée d'accès à distance aux données individuelles (pour les travaux de recherche)

Data manager / Ingénieur statisticien pour le programme IDEE – CDD 2022 à 2024

J-PAL Europe + DEPP

- Responsable du volet « données administratives » du programme
 - ~ 75% à la DEPP – interventions techniques et suivi de près du programme
 - ~ 25% au J-PAL Europe – capitalisation et transfert des connaissances (ex. documentation de projet), utiles aux collègues du J-PAL travaillant sur des projets similaires
- Contribution à la confection et aux orientations des livrables numériques susmentionnés
 - Pour le catalogue : recrutement et supervision de personnels de saisie, construction d'un système de saisie, suivis précis des avancements
 - Pour la plateforme : ex. animations d'ateliers et de comités, diverses restitutions précises des besoins pour les futurs développements
- Implication dans la gouvernance du programme
- Gestion des données en éducation en lien avec les programmes de recherche
- Automatisation des processus d'exploitation des données *via* des outils statistiques structurants (codes paramétrables, packages R, entre autres)
- Pour en savoir plus : [blog](#) et [vidéo](#)

Chargé d'enseignements en statistiques et en mathématiques – 2021 à 2022

Université de La Réunion

- Contrat d'Attaché temporaire d'enseignement et de recherche (ATER)
- Etudiants en L1 AES (Administration Economique et Sociale) et L1 Economie-Gestion
- Confection de programmes générant automatiquement des exercices variés de statistiques

Chargé d'études statistiques et d'évaluation – 2020 à 2021

Direction régionale Pôle emploi Réunion

- Mesure quantitative de l'impact des services proposés par Pôle emploi Réunion sur le retour à l'emploi
 - Construction d'un processus de suivi longitudinal de 15 cohortes mensuelles (~ 100 000 obs.) de demandeurs d'emploi : mobilisation d'une base de gestion + appariement avec d'autres sources
 - Production d'indicateurs de retour à l'emploi à 6, 12 et 18 mois
 - Estimation économétrique des effets des services de Pôle emploi (ex. entretiens téléphoniques, entretiens en présentiel, relances, ateliers, etc.) sur la probabilité de retour à l'emploi
 - Présentation des résultats (confidentiels) à la direction et autres acteurs de haut niveau
- Demandes ponctuelles de chiffrages statistiques et promotion de R (sortie de SAS)

Compétences

Compétences en économie appliquée – niveau doctorat

- Economie de l'éducation
- Evaluation des politiques publiques / évaluation d'impact : essais randomisés contrôlés, appariement, doubles différences, variables instrumentales, régressions sur une discontinuité
- Statistiques et économétrie – avancé

Niveau avancé sur R – plus de 8 ans d'utilisation soutenue du langage et outils associés pour la résolution de problèmes complexes

- Traitement de données
- Programmation, notamment packaging : [Certificat de compétences consultable](#)
- Restitution (documents dynamiques – Rmd et Quarto)

Exploitation des données – excellente culture et capacités opérationnelles

- Gestion des bases de données volumineuses
- Aspects réglementaires (typiquement RGPD)
- Data engineering
- Autres langages de programmation : Python, SAS, SQL, VBA
- Ops : Git (GitHub et GitLab – y compris en collaboration avec des collègues), CI/CD

Conduite de projet

- Cadrage et spécifications (fonctionnelles / techniques), recueil des besoins (notamment via des ateliers), comitologie, relations avec les parties prenantes internes et externes, etc.
- Travail en mode Agile
- Collaboration avec un personnel spécifique au projet (vacataires, par exemple) : recrutement, aspects RH et logistiques, maintien de la motivation, supervision du travail, restitution à la hiérarchie, etc.

Langues

- Français (natif) : lu, écrit et parlé
- Anglais (professionnel) : lu, écrit et parlé

Compétences transversales

- Tout outil de bureautique (Word, Excel, PPT, etc.)
- Travail en équipe
- Adaptation, rigueur, organisation, fiabilité

Formations

Doctorat en sciences économiques – déterminants des performances éducatives à La Réunion, dirigé par Yves CROISSANT – 2018 à 2022
([consulter la thèse](#))

Université de La Réunion

Repo Git dédié à la thèse (sans données) : <https://github.com/danielrak/phd>.

Master 2 en Economie appliquée – option méthodes quantitatives et modélisation pour l’entreprise (MQME)

Université de La Réunion

- Major de promotion, mention bien
- Connaissance pratique des méthodes de Machine Learning

Activités de recherche

Conférences scientifiques nationales et internationales

- Ecole thématique du CNRS sur l’évaluation des politiques publiques (ETEPP) – 2022
- Conférence annuelle Théorie et évaluation des politiques publiques (TEPP) – 2021
- French econometrics conference (FEC) - 2019
- Journée de microéconomie appliquée (JMA) - 2019