



Lino Galiana

<mailto:lino.galiana@insee.fr> | github.com/linogaliana

ETUDES

ENSAE	2016-2017
Cursus ingénieur voie Data Science, Parcours statistique et apprentissage: <i>Machine Learning - Monte Carlo Markov Chain - Computational Statistics - Dynamic statistical models with hidden variables - Online Learning - Apprentissage statistique en génomique</i>	
Université Pierre et Marie Curie, Paris VI	2016-2018
<i>Master 1 Mathématiques</i>	
Paris School of Economics (PSE)	2015-2016
<i>Master 2 Analyse et Politiques Economiques (APE)</i>	
Ecole Normale Supérieure (ENS) Lyon	2013-2018
<i>Normalien au département Economie. L3 Econométrie, M1 Economie, L3 Maths Appliquées</i>	
Classe préparatoire B/L	2011-2013
<i>Lycée Henri IV</i>	

EXPÉRIENCES

Data scientist	Depuis septembre 2018
<i>Insee, Département des Etudes Economiques</i>	
<ul style="list-style-type: none">• Travaux académiques innovants sur les inégalités à partir de données <i>big data</i>: données de téléphonie mobile, données de tickets de caisse. Exploitation des données via <code>pyspark</code> et <code>ElasticSearch</code> ou avec des API <code>R</code> ou <code>Python</code> pour des données <code>PostgreSQL</code> ou <code>S3</code>• Production des statistiques de déplacements de population à la suite de l'annonce du confinement de Mars 2020 puis lors de la première phase de déconfinement en Juin 2020. Travail avec <code>R</code> et <code>Python</code> pour harmoniser les sources et les valoriser avec représentations cartographiques. Données stockées sur un système <code>S3</code>.• Contribution au passage à <code>R</code> et <code>Python</code> à l'Insee (packages, création de l'environnement <code>Anaconda</code> repris pour l'ensemble de l'Insee, images <code>Docker R</code> et <code>Python</code>). Lancement du projet <code>utilitR</code> pour proposer une documentation collaborative présentant des recommandations sur les manières d'assurer reproductibilité et pérennité d'un projet statistique mené sur langage <i>open source</i> (<code>R</code>).• Contribution à l'adoption de <code>Git</code> en interne. Création d'une formation, reprise nationalement, sur le Travail collaboratif avec <code>R</code> et <code>Git</code>. Contributions multiples sur <code>GitHub</code> (https://github.com/linogaliana/ et https://github.com/InseeFrLab/)• Cours de Macroéconomie dans la préparation concours Administrateur interne• Suivi d'une formation en <i>Big Data: Architecture et Données</i> (IGPDE)	
Stagiaire	Dec. 2017 - Mai 2018
<i>Insee, DMCSI - SSP-lab</i>	
<ul style="list-style-type: none">• Exploration des apports des données de téléphonie mobile en complément des données fiscales (bases Filosofi) pour mesurer les phénomènes de ségrégation urbaine	
Stagiaire	Mai 2017 - Nov. 2017
<i>EDF - Direction Economie, Tarifs et Prix</i>	
<ul style="list-style-type: none">• Intégration des remontées rapides des données Linky au modèle de prévision de consommation électrique J+1 d'EDF	

ENSEIGNEMENTS

Enseignant: Python pour le data-scientist ou économiste

Depuis septembre 2020

ENSAE ParisTech

- Présentation de l'éco-système Python pour la *data science* avec applications sur données open-data. Structuration du cours en 4 parties principales: traitement des données, visualisation, modélisation, *natural language processing* (NLP) avec un complément sur 'Git'
- Intégration continue et déploiement en continu (CI/CD) du site <https://linogaliana-teaching.netlify.app/> à partir du dépôt Github (<https://github.com/linogaliana/python-datascientist>)

Enseignant d'économie et de mathématiques

2016-2020

Sciences Po

- Cours d'Economie de la ville (M1 Stratégies Territoriales et Urbaines), de Microéconomie (1A, Programme Europe-Afrique) et Mathématiques (1A, Programme Europe-Amérique)

COMPÉTENCES

Programmation: R; Python; Git; Pyspark ; Elasticsearch ; Shell; PostgreSQL; SQL; Docker; Kubernetes; C++
Langues: Anglais, Espagnol, Italien (débutant)

PROJETS

Publications

<https://github.com/linogaliana/label-openfood>

Effet des inégalités socioéconomiques sur les risques de développer une forme grave de la Covid-19

Dossiers de la Drees, A venir

Inégalités socio-spatiales de consommation d'aliments gras, sucrés et salés *Document de Travail Insee, A venir*

Financial Wealth over the Life-Cycle: A Microsimulation Approach *Document de Travail Insee, A venir*

Residential segregation, daytime segregation and spatial frictions: an analysis from mobile phone data

Document de Travail Insee, 2020

La mixité sociale est plus forte en journée sur les lieux d'activité que pendant la nuit dans les quartiers de résidence *Insee Analyses, 2020*

Retour partiel des mouvements de population avec le déconfinement *Insee Analyses, 2020*

Les taux d'épargne dans la zone euro: comment les comparer ? comment les expliquer ? *L'Economie Française, 2017*

Projets open-source

pynsee: une API pour accéder aux données de l'Insee depuis Python <https://github.com/InseeFrLab/Py-Insee-Data>

utilitR: une documentation utile pour R. <https://github.com/InseeFrLab/utilitR>

btbpy: lissage spatial en python. <https://github.com/InseeFrLab/btbpy>

Posts de blog pour **spyrales**, communauté d'agents de l'Etat intéressés par R et python

Site web indexé par **R-bloggers**

Hackathon Insee, *Les champs de Sirène* ; Lauréat challenge Epidemium data *Fight4Cancer*

CONFERENCES

JMA Annecy 2021 ; UEA 2021 Virtual Conference ; 2019 Paris CEPR Workshop on social mobility ; 2019 Workshop on Pensions and Ageing ; Oxford Netmob 2019 ; UEA Amsterdam 2019 ; BigSurv Barcelona 2018 ; JMS Paris 2018

INTÉRÊTS PERSONNELS

Référent pour la gestion de deux sites de compostage à Malakoff

Association jardin partagé Les Nouveaux

Football, Tennis, vélo, escalade, randonnée, plongée sous-marine (niveau I), cinéma, littérature

Animateur à TrENSistor, radio de l'ENS Lyon