

Datascientist

Conrad THIOUNN

à la recherche d'un stage non rémunéré en datascience de 3-4 mois à p. de juin 2022

Administrateur stagiaire INSEE – ENSAE – Institut Polytechnique de Paris, titulaire d'un master en statistique - ENSAI

conrad.thiounn@ensae.fr

<https://www.linkedin.com/in/conrad-thiounn/>

<https://github.com/cthiounn>

Expériences professionnelles

2019-09 -	Chef de projet informatique pour le recensement de la population	<i>Insee - Montrouge</i>
2021-09	Projet de remplacement de la taxe d'habitation par un répertoire des logements (37 millions, 100 variables) <ul style="list-style-type: none">Conception de la solution (co-construction avec le métier - rédaction - prototypage)Mise en œuvre de la solution (développement en organisation agile Scrum - JAVA Spring)	
2016-09 -	Développeur informatique pour le recensement de la population	<i>Insee – Montrouge</i>
2019-09	Responsable des redressements des données sur les individus et les logements (8 millions et 5 millions par an, 300 variables) <ul style="list-style-type: none">Programmation Orientée Objet en JAVA, alimentant le travail de reprise manuelle de 150 questionnaires.Migration à PostgreSQL, GIT	
2014-09 -	Statisticien au sein de la direction des statistiques d'entreprises	<i>Insee – Malakoff</i>
2016-09	Responsable des redressements des données annuelles d'entreprises (8 millions, 100 variables, 100 questionnaires sectoriels) <ul style="list-style-type: none">Algorithmie en pseudo-langage (compilé en JAVA)Travaux quotidiens avec SQL et SAS, alimentant des expertises <i>ad hoc</i> de 100 questionnaires.	
2011-04 -	Administrateur système et réseau	<i>Insee – Malakoff</i>
2014-09	<ul style="list-style-type: none">Administration de la partie frontale de l'infrastructure de production (messagerie, reverse-proxy, DNS, Antivirus)Mise en conformité et application des règles de sécurité au parc informatique sous Linux et Windows ServerMise en place d'un système de supervision avec Nagios et programmation de scripts en python et shell	

Travaux statistiques

2022	Statistiques et maladies infectieuses <ul style="list-style-type: none">Modélisation SIR/SEIR, inférence bayésienne, Metropolis-Hastings (R, R shiny)Dockerisation et déploiement (docker, kubernetes, helm) https://github.com/cthiounn/Seminaire-Stat-ENSAE-2A https://modelisation-sir-ensae-2a.lab.sspcloud.fr/
2021 - 2022	Analyse Statistique et Empirique des modèles de génération Text-to-Image (Dall-E) <ul style="list-style-type: none">Datasets (MSCoco, Flickr30k)Réseau de neurones (pytorch, transformers) https://github.com/cthiounn/images-generated-by-dalle-mini / https://github.com/cthiounn/dalle-tiny
2021	Que peut-on dire des statistiques de délits et des crimes à partir des établissements de boucheries-charcutiers ? <ul style="list-style-type: none">Datavisualisation (geopandas, plotly, matplotlib, seaborn)Analyse multivariée [PCA et clustering Kmeans] (sklearn, scipy, prince)Régression linéaire (statsmodels, sklearn) https://github.com/cthiounn/python-datascience-ENSAE-2A
2013	Refonte de l'anti-spam de l'Insee <ul style="list-style-type: none">Analyse textuelle (utilisation de R.Temis, AFC, CAH)
2009	Identification de sous-groupes cliniques et évolutifs homogènes au sein des vascularites associées aux ANCA <ul style="list-style-type: none">Statistique exploratoire univariée et bivariée (SAS et Rcmdr)Analyse factorielle (ACM) / Méthode de partitionnement (CAH) / Bootstrap

Formations

-2021	ENSAE (admis en deuxième année directe)
2011 - 2019	Qualifications ministérielles de chef de projet, PSE-trad, PSE-CRA, PSE-ER, analyste
2012 - 2013	Master Droit, Économie, Gestion, mention Statistique Économétrie, spécialité Statistique Publique (ENSAI – Rennes I)
2008 - 2011	ENSAI (spécialisation en 3ème année en statistiques et traitements des données, ex-Data Science & Ingénierie des données)