

Anas Zakroum

DATA SCIENTIST · STATISTICIAN · DATA ANALYST

+33 663 384 180 | anaszkr@gmail.com | lvl0-statistician

About Me

Process oriented data analyst/scientist with 2 years of experience. Experienced in analyzing and interpreting data. Provide insights, analytics, and business intelligence needed to guide decisions.

Professional Experience

Data Scientist at CIRAD

2023 - Present · Montpellier - FR

- Supported quantitative data-related projects
- Data Science Counseling for collaborators with different backgrounds
- Extracted data from different sources, transformed data into actionable format
- Conducted data analysis in collaboration with interdisciplinary teams and stakeholders
- Constructed visualization dashboards and developed ML models according to the needs of the team
- Technical Skills :** Network Science, Clustering, Regression, Time-series Analysis and Forecasting, Random Forests, Parametric Estimation, PCA, MCA, etc.
- Tools :** Bash, R, Python, NetworkX, GraphML, GIS, Git.

Data Scientist at Biotope

2022 - 2023 · Mèze - FR

- Automated detection and identification of animals appearing in camera trap images allowing for reducing image annotation time.
 - Design of an API pipeline meant for end users
 - Design, test and implementation of a Deep Learning model architecture for detection and classification of species
 - 50 % time reduction for image processing
- Query database to process metadata
- Analysis and interpretation of KPIs with regards to species endangerment status
- Technical Skills :** A/B Testing, Object Oriented Programming, CNN, RNN, NN fine-tuning, Statistical Modeling, Data Analysis, etc .
- Tools :** SQL, Bash, docker, Python (TensorFlow, Keras), R Shiny, Microsoft PowerBI, Git.

Statistical Analyst · Intern at INSEE

2021 - 3 months · Marseille - FR

- Created an API to explore air quality drivers
- Statistical analysis and modeling of indoor and outdoor air quality
- Harmonization and cleaning of data
- Explainable Prediction of air quality using advanced regularization techniques to handle colinearities in high dimensional data
- Technical Skills :** A/B Testing, ANOVA, ML, THEME, PCA, PLS, Dimensionality Reduction.
- Tools :** SQL, R, R Shiny, Python, Git.

Skills

Programming	Python, R, Shell Scripting, Java, Flask
Operating Systems	GNU/Linux (Debian and derivatives, Red Hat Enterprise Linux, etc.)
Database Management	SQL, PostgreSQL, NoSQL, etc.
Tools	git, docker, AWS, Microsoft PowerBI, PySpark, Hadoop, \LaTeX
Other / Soft	Time Management, Teamwork, Communication, Technical writing, Mentoring, Documentation

Education

Msc in Statistics and Data Science at Montpellier University

2020 - 2022 · Montpellier - FR

- Mentoring bachelor students, Scientific event organization
- Courses:** STOCHASTIC PROCESSES · STATISTICAL INFERENCE · MACHINE LEARNING · DEEP LEARNING · SURVIVAL ANALYSIS · CLUSTERING REGRESSION · TIME SERIES FORECASTING · GENERALIZED LINEAR MODELS · BAYESIAN STATISTICS · STOCHASTIC CONTROL · MARKOV CHAINS LATENT VARIABLES MODELING · MULTIVARIATE DATA ANALYSIS · CONVEX OPTIMIZATION · INFORMATION THEORY · DECISION THEORY

Bachelor in Mathematics and applications at Montpellier University

2016 - 2020 · Montpellier - FR

Languages

- English** Fluent, Advanced writing skills, TOEIC : 885/990
- French** Fluent, Advanced writing skills

A few projects examples

Below you can find examples of projects I have worked ranging from software development to statistical data analysis and predictive machine learning.

Good wine, bad wine ? Why ? at Montpellier University

2022 • Montpellier-FR

- Exploration the driving factors of wine quality
- Build a classification model of wine qualities and overcome multi-collinearities known to affect model performance
- Construction of KPIs and interpretation of metrics according to needs
- Writing technical report
- **Technical Skills:** ANOVA, Discriminant Analysis, PLS-DA, Classification, Clustering
- **Tools:** R, Git, \LaTeX .

Worldwide statistics visualizer at Montpellier University

2022 • Montpellier-FR

- Construction of our own data visualization API for countries
- Package development in Python
- Managing input/output dependencies
- Writing technical documentation
- Unit Testing and deployment
- **Technical Skills:** Object Oriented Programming, Dependencies management, Unit Testing, CI/CD.
- **Tools:** Git, Flask, Python (io, pandas, numpy, matplotlib, sys, pytest, os)

Understanding polygamous unions, a statistical perspective at Montpellier University

2022 • Montpellier-FR

- Exploring the factors driving of polygamous union formation
- Life trajectory data surveyed
- Statistical analysis of categorical data
- Data Visualization
- Design of a methodology suited for the study
- Construction of a Time-To-Event model handling high correlation among data
- **Technical Skills:** Statistical Inference, Survival Analysis, Cox-Regression, Proportional hazards, MCA.
- **Tools:** R, \LaTeX

Bicycle traffic forecasting at Montpellier University

2021 • Montpellier-FR

- Build ETL pipeline to read and preprocess data
- Develop a methodology suited for the task
- Build and train a machine learning model to predict bicycle traffic
- Write technical report
- **Technical Skills:** Input/Output, Time Series processing and forecasting, Technical Writing, Autoregressive modeling.
- **Tools:** Python (pandas, scikit-learn), \LaTeX

References available upon request.