
PROFESSIONAL SUMMARY

Results driven Data Scientist with an engineering background in statistics and data analytics, specializing in experimental design and causal inference. Strong foundation in Python, SQL, and statistical modeling with expertise in hypothesis testing, observational study design, and impact evaluation. Known for turning complex data into clear, actionable insights, including large scale causal uplift modeling using Double Machine Learning IV.

CORE SKILLS & TECHNICAL EXPERTISE

- **Causal Inference & Experiments:** A/B testing, experimental design, difference-in-differences, synthetic control, uplift modeling, propensity score weighting, Double Machine Learning (DML/DMLIV), CATE estimation, impact evaluation, hypothesis testing.
- **Data Analysis & Modeling:** Python (pandas, NumPy, scikit-learn, statsmodels), SQL, machine learning, regression analysis, time-series forecasting.
- **Tools & Technologies:** BigQuery, Looker Studio, Tableau, Power BI, Excel, Git, Jupyter Notebooks.
- **Soft Skills:** Analytical thinking, problem-solving, cross-functional collaboration, Agile methodologies.
- **Languages:** English (Fluent), French (Professional), Arabic (Native).

WORK EXPERIENCE

Ozeol | Experimentation & Performance Analyst

Tunis, Tunisia | 09/2025 – 06/2026

- Analyzed key performance indicators across digital channels and product features to identify trends and opportunities, influencing strategic optimization.
- Developed and automated interactive dashboards to deliver actionable insights to marketing and product teams.
- Collaborated with stakeholders to design and analyze A/B tests and multivariate experiments, ensuring statistical rigor.
- Applied causal inference techniques (difference-in-differences, synthetic control) to quantify the impact of marketing campaigns and operational changes, informing resource allocation (e.g., measured a 10% lift in asset recovery rate).
- Provided data-driven recommendations that improved campaign efficiency and supported portfolios and processes.

Glovo | Market Analyst

Tunis, Tunisia | 01/2024 – 09/2025

- Conducted comprehensive market and user-behavior analyses using Python and SQL to identify growth opportunities.
- Built predictive models for customer segmentation and demand forecasting, optimizing pricing and marketing strategies.
- Designed and evaluated A/B tests for marketing campaigns and performed uplift modeling to target offers (e.g., a targeted promotion increased user engagement by 20%).
- Analyzed the causal impact of pricing and incentive strategies on order volume using observational techniques and statistical experiments (e.g., difference-in-differences analysis to assess expansion in a new city).
- Collaborated with sales and marketing teams to translate insights, contributing to market expansion and optimization.

United Nations Population Fund | Data Analyst Consultant

Tunis, Tunisia | 05/2023 – 07/2023

- Processed and analyzed national health survey data to derive demographic and health insights for planning.
- Created performance-monitoring dashboards (Excel, Python) for tracking key health indicators in real-time.
- Automated data cleaning scripts, reducing survey data processing time by 30% and enabling timely analysis.
- Collaborated with local and global stakeholders on health analytics projects to measure outcomes.

EDUCATION

Higher School of Statistics and Data Analysis |Engineering degree of Statistics and data analysis, Tunis Tunisia |09.2019-10.2023

- **Specialization:** Big Data, Advanced Analytics, Machine Learning.

IPEIN – Preparatory Institute for Engineering Studies | Preparatory Studies (Mathematics & Physics), Nabeul Tunisia |09.2017-06.2019

PROJECTS

- **Causal Uplift Modeling with DMLIV (Criteo Dataset) | github.com/bahakahri/causal-uplift-dmliv:** Built an end-to-end CATE estimation pipeline on 25M rows using NonParamDMLIV with LightGBM nuisance models. Validated instrument strength (F-stat 5,575) and achieved a Qini coefficient of 9,611.
- **Regression Discontinuity Design (Lending Club) | github.com/bahakahri/regression-discontinuity:** Estimated the causal effect of crossing a 675 FICO threshold on approved loan amounts using sharp RDD. Validated assumptions with McCrary density test and covariate smoothness checks.
- **Staggered Difference-in-Differences | github.com/bahakahri/staggered-difference-in-differences:** Applied the Callaway & Sant'Anna estimator to measure true feature impact under staggered treatment adoption, avoiding bias from heterogeneous treatment timing in standard DiD.