Mohammad Haghjoo



Data Science & Analytics

Working Student (Werkstudent) · Available from Dec 2026 · up to 20 h/week · Nürnberg, Erlangen, Munich or hybrid

Nürnberg, Germany · mohammad.haghjoo@fau.de · mhmmdhaghjoo.github.io · github.com/mhmmdhaghjoo · linkedin.com/in/mohammad-haghjoo-b241823b0

PROFILE

Data Science & Analytics professional with five years of full-time (plus six years of part-time) industry experience owning production PostgreSQL databases, automated ETL reporting pipelines, internal Python systems, and applied ML models. Focused on data-quality auditing, statistical analysis, and clear reporting across SQL warehouses, Tableau Public, Streamlit, and R Shiny. Currently pursuing an M.Sc. in International Information Systems at FAU Erlangen-Nürnberg (B.Sc. Computer Engineering); available as Working Student from Dec 2026.

TECHNICAL SKILLS

Programming & SQL	Python (pandas, NumPy, SciPy), SQL, PostgreSQL, R (tidyverse, ggplot2), Kimball Star Schema
Statistics & Analytics	EDA, Hypothesis Testing, Bayesian Inference, A/B Experimentation, RFM & Cohort Analytics
BI & Data Apps	Tableau Public, Executive Dashboard Design, Data Storytelling, Streamlit, R Shiny, Plotly
ML & Engineering	scikit-learn, Calibration, Error & Leakage Auditing, FastAPI, pytest, mypy, ruff, Git, GitHub Actions

PROFESSIONAL EXPERIENCE

08/2020 – 10/2025

Information Systems Specialist (Databases, ETL, Python Systems & ML)

Shiraz Petrochemical Co. · Tehran, Iran · Full-time (5 years)

- **Production Database Reliability:** Owned PostgreSQL design and administration across five years, delivering normalized schemas, indexing/query tuning, role-based access control, and scheduled backup routines for multi-department systems.
- **Automated ETL & Recurring Reporting:** Built SQL data-integration pipelines and cross-system interfaces consolidating operational tables into recurring analytical reports for internal business teams.
- **Internal Python Systems & Workflow Automation:** Engineered and maintained Python enterprise tools (pandas, NumPy) and scheduled automation scripts replacing manual data imports, recurring report assembly, and backup checks.
- **Applied ML on Production Data:** Developed machine-learning models in Python (pandas, scikit-learn) on petrochemical production datasets to support operational analysis and troubleshooting.

01/2014 – 06/2020

Information Technology Specialist (Databases, SQL Reporting & Python Tools)

Tankerranan Petroleum Transport Company · Rafsanjan, Iran · Part-time (6 years)

- **Database Operations & Continuity:** Maintained and tuned PostgreSQL operational databases across six years, handling routine updates, indexing, and scheduled backups to keep company records reliable and accessible.
- **Recurring Operational SQL Reporting:** Delivered structured SQL queries and recurring operational reports giving management regular visibility into day-to-day transport operations.
- **Internal Python Automation Tools:** Built internal Python utilities and data-processing scripts (pandas, NumPy) automating repetitive reporting tasks alongside undergraduate engineering studies.

SELECTED OPEN-SOURCE PROJECTS

Retail Analytics Warehouse & Executive BI SQL, Kimball Star Schema & Tableau Public · v0.1.0
github.com/mhmmdhaghjoo/retail-analytics-warehouse

Kimball star-schema SQL warehouse and 10-check financial reconciliation audit on **541,909 real European retail transactions** (UCI Online Retail). Removes 5,268 duplicates, isolates 9,251 credit returns and 132,186 guest checkouts, reconciles **£10.64M gross / £9.75M net revenue** to the penny (**10/10 DQ checks PASS**), and publishes an interactive **Tableau Public executive dashboard** (RFM, geographic map, cohorts).

Marketing Campaign Decision & Calibration Lab Tabular ML, FastAPI & Streamlit · v0.1.0
github.com/mhmmdhaghjoo/marketing-campaign-decision-lab

Leakage-audited tabular ML, probability calibration, FastAPI scoring endpoint, and interactive **Streamlit** simulator on **45,211 real European marketing calls** (UCI Bank Marketing). Exposes post-call `duration` leakage (0.91 inflated → **0.81 honest ROC-AUC, 0.008 ECE**), audits chronological prior shift, and turns a **−€29.8k blanket loss** into **+€40.2k net profit (+164.5% ROI)**.

Experimentation & Bayesian Inference Workbench (R + Shiny) Statistics, R & Shiny · v0.1.0
github.com/mhmmdhaghjoo/ab-experimentation-workbench-r

Pre-readout A/B auditing and decision workbench in **R and Shiny** on a reproducible 20,000-session checkout benchmark. Automates Sample Ratio Mismatch (χ^2) checks, Frequentist vs. exact **Bayesian Beta-Binomial** posteriors (`P(B>A)=90.7%`), power planning (14,768/arm), and **Holm-Bonferroni** subgroup correction catching false-positive noise (p=0.026 → 0.102).

EDUCATION

10/2025 – 03/2028 expected

M.Sc. International Information Systems
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)

Relevant coursework: Data Science: Foundations, Tools, Applications in Socio-Economics & Marketing; People Analytics; Experimental Behavioral Research in Data Science; Knowledge Discovery in Databases; Pattern Recognition; AI Assistance in the Data Science Value Chain (Seminar); Business & Global Operations Strategy; Introduction to Software Engineering; Free & Open Source Software.

10/2013 – 07/2020

B.Sc. Computer Engineering (Software Engineering)
Islamic Azad University, Central Tehran Branch

Relevant coursework: Principles of Database Design; Database Lab; Engineering Statistics & Probability; Artificial Intelligence; Data Structures; Algorithm Design; Software Engineering I & II.

LANGUAGES

English C1 (IELTS Academic 7.5) · **German** B1 · **Persian** Native · **Turkish** Intermediate