

Mustafa Sezer Başar

Data Scientist / Machine Learning Engineer

Ankara, Türkiye · +90 533 024 11 76 · mustafasezerbasar@gmail.com · github.com/mustafa-sezer-basar · linkedin.com/in/mustafa-sezer-basar

PROFESSIONAL SUMMARY

Database analyst at a retail group, leading a three-person technology team. Works on end-to-end ML pipelines, RAG architectures and measurement design, tying a model's output to an actual business decision. Treats measuring how wrong a method is as part of the job: control groups, placebos and hypotheses written before the run.

SKILLS

- Data Science & AI:** Python (pandas, scikit-learn, NumPy), RAG architecture, LLM API integration, structured output, Computer Vision (face_recognition, Tesseract OCR)
- Databases:** SQL (joins, window functions, indexing), PostgreSQL (pgvector), MySQL, SQLite
- Web & Full-Stack:** JavaScript / TypeScript, React, Next.js, Express.js, REST API

PROJECTS

- Tax Audit System with Rule Extraction from Legislation** — The rule engine of the field audit system: reads legislation with RAG and turns it into audit rules. A collector he wrote feeds the corpus; rules are held in JSON with article references and validity dates, pass human review, and are then applied deterministically by a Go service. The LLM is off the hot path, so every audit run stays reproducible. *(Python (LLM), Go, Next.js, MySQL, Redis)*
- Data Quality Audit & Reporting Tool** — A tool that cleans dirty tabular data into a multi-sheet Excel report — the real subject being the cleaner's own error rate. It is scored against a synthetic file whose defects are known: 10 silent errors against 361 for a competent twenty-line pandas cleaner. The measurement found three defects in the tool itself, none visible from its output. 134 tests. *(Python, pandas, openpyxl)*
- Schedule Risk and Critical Path Analysis** — A Monte Carlo study of how likely a schedule is to meet its committed date: on a twenty-activity ERP migration the critical-path date holds in 34% of runs. Adding parallel paths pushes the finish 12.6 days out with no task lengthened. The pre-registered hypothesis was refuted and reported. *(Python (NumPy, SciPy, Monte Carlo))*

EXPERIENCE

Paulmark — Database Analyst — Business Development & Strategy

June 2026 – Present · Ankara

- Leads a three-person technology team, moving a Nebim V3 and Power BI retail stack off Excel-based processes.
- Closed the gap between the teams' own Excel files and the master data; inconsistencies used to surface only once they had grown.

Myrco Spa Center (Barceló Cappadocia) — Co-Founder / Operations Lead

September 2024 – February 2025 · Cappadocia

- Ran the spa centre inside the Barceló Cappadocia hotel as a subcontractor; built the 15-person team from scratch.
- Cut team-caused errors 95% in two months with a Spa Management app he built; error = an unauthorised entry counted on camera, or a session off the sold duration by $\pm 5\%$.

EDUCATION

Hacettepe University — BA in Business Administration

August 2022 – July 2026 · GPA 3.12/4.00 · English-medium · Graduated

Middle East Technical University (METU) — Civil Engineering

September 2018 – August 2022 · English-medium · Transferred to Business Administration at Hacettepe University

LANGUAGES

Turkish (Native) · English (Advanced — English-medium undergraduate education)