

Omer Haluk Demirhan

Statistics · Computational Statistics · Statistical Machine Learning

www.halukdemirhan.com
London, E14 9YD
(+44) 07778 847275
halukdemirhan@hotmail.com

EDUCATION

University of Greenwich, London – *MSc Management of Business Information Technology*

JAN 2024 – JAN 2025 | MERIT

Dissertation: *Assessing Data Quality Management Issues in Open-Sourced Data Sets with Statistical Methodology*. Developed and validated a statistical framework for evaluating the quality of openly published datasets, combining quantitative diagnostics with qualitative assessment of documentation and provenance. Implemented as an R package providing anomaly detection, visual diagnostics and automated reporting.

halukdemirhan.com/assets/msc-thesis-haluk-demirhan-public.pdf

Marmara University, Istanbul – *BSc Statistics*

2019 – 2023

Coursework in probability theory, statistical inference, stochastic processes, regression and multivariate analysis, time series, experimental design and machine learning. Marmara University is ranked 8th nationally in Turkey, with its Department of Statistics among the top five in the country.

PUBLICATIONS AND PREPRINTS

Demirhan, O. H. datagrade: Computable ISO/IEC 25012 Data Quality Measures for R. *Submitted to the Journal of Open Source Software*, August 2026. Archived version: doi.org/10.5281/zenodo.22064234

Demirhan, O. H. Measurable and Unmeasurable: Operationalising the ISO/IEC 25012 Data Quality Model with Explicit Applicability Semantics. *Manuscript in preparation for arXiv (cs.SE)*, 2026.

RESEARCH AND SOFTWARE

datagrade – *R package for statistical assessment of data quality*

2025 – PRESENT | AUTHOR AND MAINTAINER | MIT LICENCE

Open-source implementation of the measurement framework developed in my MSc dissertation, treating data quality as a governance and assurance problem rather than a preprocessing step. `dg_assess()` scores a dataset across the five ISO/IEC 25012 quality characteristics, accuracy, completeness, consistency, credibility and currentness, via fifteen measured properties, and returns a structured diagnostic report covering missing values, duplicates, outliers and redundant columns, designed for audit trails and model risk documentation rather than interactive inspection. Requires no configuration, has three dependencies, and assesses 10 million rows in approximately 38 seconds. Submitted to CRAN.

Each property is computed as a ratio A/B in which the denominator counts the items for which the property is defined at all, so properties that cannot be evaluated return NA rather than zero and are dropped from aggregation instead of being scored as de-

RESEARCH INTERESTS

Statistical methodology for data quality

Computational statistics

Statistical machine learning

Empirical evaluation of software and AI systems

Statistical software development

STATISTICAL METHODS

Statistical modelling and regression

Missing data and anomaly detection

Experimental design and A/B testing

Multivariate and high-dimensional analysis

Mixed quantitative and qualitative analysis

TECHNICAL SKILLS

R (package development, CRAN tooling, ggplot2), Python (NumPy, pandas, scikit-learn), FastAPI, React Native, Flutter, Dart, TypeScript, Docker, Git, Redis, Firebase, REST APIs, GA4.

LANGUAGES

English (Advanced)

Turkish (Native)

Russian (Basic)

LINKS

halukdemirhan.com

github.com/ohalukdemirhan

fects. The taxonomy follows ISO/IEC 25012, but the measurement functions are my own implementations of named properties and are not conformant with ISO/IEC 25024. Establishing specification-based, statistically defensible measurement for these dimensions is the planned second version and an open methodological problem.

github.com/ohalukdemirhan/datagrade

doi.org/10.5281/zenodo.22064234

agent-estateMCP – *specification and reference implementation for an agent operations registry*

2025 – PRESENT | AUTHOR AND MAINTAINER | MIT LICENCE

An append-only registry that autonomous software agents write to, providing a persistent audit trail for agent-operated software estates. Agents record observations, findings and attempted work; a single operation, plan approval, is reserved for a human, making the governance boundary explicit and enforceable rather than procedural. Implemented as an MCP server over HTTP exposing thirty-one tools across five record types, in Python with SQLite storage, and accompanied by a written specification, threat model and safe-publication guidance.

github.com/ohalukdemirhan/agent-estateMCP

MSc Dissertation Research – *Data quality in open-sourced data sets*

2024 – 2025 | SUPERVISOR: DR SOLOMON EBENUWA

Quantitative and qualitative analysis of publicly available datasets to characterise recurring data quality failures and to quantify their effect on downstream statistical analysis. Combined distributional and missingness diagnostics with structured assessment of metadata and provenance, and validated the resulting scoring framework against independently annotated datasets.

TEACHING

Art Project Lab, Istanbul – *Design Technology Lecturer*

MAY 2020 – AUG 2022

Taught digital design, layout principles and visual storytelling to secondary and pre-university students. Led creative projects for academic events and exhibitions, delivered print and digital material to fixed academic deadlines, and mentored students preparing portfolios for AP Art and Design programmes. Held concurrently with undergraduate study.

PROFESSIONAL EXPERIENCE

Commercium Limited, London – *Founder*

APR 2026 – PRESENT

AI software company developing subscription mobile applications, automation systems and generative content pipelines. Responsible for the full technical stack, including model integration, backend services, data infrastructure and analytics instrumentation.

Stealth consumer app startup, London – *Co-Founder*

JUL 2025 – JUN 2026

Designed and built consumer mobile applications in React Native and FastAPI, scaling to over 40,000 downloads and more than \$20K in monthly recurring revenue. Built the experimentation and analytics pipeline used to evaluate product changes, including cohort retention analysis and controlled A/B testing of growth interventions.

AGS Global, Istanbul – *Quantitative Market Researcher*

AUG 2022 – JUL 2023

Conducted quantitative market analysis for more than 100 exporting firms with combined annual revenue above \$3M, within the Turkish Government's Turquality programme. Produced data-driven market entry recommendations from trade, pricing and demand data.

REFEREES

Referees available on request.