

CAVAN DONOHOE

Looking to improve your company's data workflow

Senior Data Scientist / Statistical Programmer with 6+ years of experience building production-grade R packages, automated clinical trial pipelines, and reproducible TFL workflows in regulated environments. Deep expertise in R, CDISC standards (SDTM/ADaM), statistical programming, and end-to-end automation across large-scale clinical datasets. Proven track record of saving thousands of analyst hours through tooling, package development, and workflow optimization.



PROFESSIONAL EXPERIENCE

February 2021
|
August 2024

Senior Data Scientist

GRAIL

📍 Menlo Park, CA

- Authored and maintained multiple production R packages used company-wide for clinical trial analysis, validation, and reporting.
- Built automated R Markdown pipelines generating hundreds of Tables, Figures, and Listings (TFLs) to support trial accuracy assessments and regulatory reporting.
- Developed and documented derived variables under CDISC standards (SDTM, ADaM, CDASH) with Git/GitLab collaboration and AWS S3 data management.
- Created an R Shiny app to visualize cancer screening timelines and key events leading to diagnosis.
- Provided statistical programming support for data cleaning, database locks, SAP-driven analyses, and exploratory investigations.
- Partnered with biostatisticians to translate SAPs into analysis specifications and reusable code.
- Built reproducible pipelines with unit tests and a shared code library to ensure audit-ready, backward-compatible outputs.

Reporting Analyst

24 Hour Home Care

📍 El Segundo, CA

- Automated weekly accounts receivable aging report in R with HTML/CSS email delivery, saving ~5 hours per week.
- Built ad hoc dashboards and data visualizations to support leadership requests.
- Served as analytics liaison across teams to align reporting needs and delivery.
- Automated batch invoice emails and Azure Web App exports, reducing manual effort across teams.

Actuarial Analyst

HMSA

📍 Honolulu, HI

- Automated CRG band renewal in R, reducing runtime from 8 hours to under 1 minute.
- Automated demographic and plan risk factor workflows and forecasting models with Excel/R integration.
- Automated CRG proposal and mental health parity processes; documented methods and led stakeholder walkthroughs.
- Fit empirical distributions using KS tests and MLE for actuarial modeling.



EDUCATION

2018

Bachelor of Science

University of California, Irvine

📍 Irvine, CA

- Major in Mathematics
- Minor in Statistics

[Back to Home Page](#)

[PDF version](#)

[Word version](#)

CONTACT INFO



cavandonohoe@gmail.com



github.com/cavandonohoe



linkedin.com/in/cavan-donohoe/



(805) 404-3321

SKILLS

R: tidyverse, data.table, ggplot2, Shiny, R Markdown, package development, unit testing, automation, resampling (bootstrap/jackknife)

Clinical & Statistics: CDISC (SDTM, ADaM, CDASH), survival analysis, model validation

Databases: SQL (PostgreSQL)

Developer Tooling: Git/GitLab, CI workflows, AWS S3, AI-assisted development (Cursor, GitHub Copilot, Gemini, Codex) with full manual review and validation

Excel: Advanced formulas, automation, VBA

This resume was made with the R package [pagedown](#).

Last updated on 2025-12-28.

 AWARDS

- BAS UCLA Case Competition Finalist ([Project Presentation](#) and [Project Memorandum](#))

 CREDENTIALS

- Passed 3 Actuarial Exams (Exam P, FM, SRM)