Kaleb Coberly

📍 Bellingham, WA, USA

✉️ kaleb.coberly@gmail.com  
🔗 linkedin.com/in/kaleb-coberly

💻 github.com/KalebCoberly  
🌐 kalebcoberly.com

# Summary

Backend software engineer, data engineer with a background in research and analysis. Builds and deploys tools, analytics, and pipelines that help others reliably use data. Supports users and developers with documentation, dev tools, issue triage, hands-on training, and debugging. Improves systems with a focus on clarity and maintainability. Prioritizes data quality, reproducibility, validity, and meaningful results.

# Skills

* *Programming & Scripting:* Python, R, Bash, SQL, NoSQL.
* *DevOps & Tooling:* Git, Make, Docker, GitHub Actions, Jenkins, CI/CD, Python packaging, debugging, QC and testing frameworks, pytest, pytype, black, flake8, jake.
* *Data Engineering:* ETL pipelines, RDBMS, database design, data modeling, data cleaning, data validation, methodological implementation, data versioning, MySQL, MariaDB, MongoDB, ElasticSearch, XML, JSON, CSV.
* *ML Engineering:* ML pipelines, feature engineering, performance evaluation, model validation, scikit-learn, caret.
* *Analysis & Visualization:* Pandas, NumPy, matplotlib, seaborn, ggplot2, datatable, dplyr, knitr, Jupyter Notebook, RStudio, R Markdown, mermaid, HTML, Excel.
* *Documentation & Support:* Sphinx, reStructuredText, user training, contributor onboarding, issue triage.
* *Platforms & Environments:* Linux, MacOS, Windows, HPC (Slurm, UGE), Kubernetes (k8s), Bitbucket/Stash, VS Code, RStudio, JIRA.

# Experience

## Volunteer Software Engineer, Bellingham Food Bank, Nov 2024–Present

* Built and deployed a tool to automate delivery planning, reducing staff workload by several hours with each weekly use. See <https://cricketsandcomb.org/#projects>.
* Designed modular, CLI-based system integrating into staff workflows; led incremental rollout and training.
* Wrote complete documentation for non-technical users; maintain ongoing support based on user feedback.
* Implemented CI/CD, local + Docker testing, and reusable GitHub Actions workflows with high test coverage.
* Released as an open-source project; maintain infrastructure and lead collaborative development at all skill levels.

## Research Engineer, Central Computation (GBD), IHME, Nov 2022–Oct 2024

* Developed and maintained internal Python and dev tools for data modeling, versioning, access, and dev workflows.
* Contributed to a large in-house Python ecosystem: features, bugfixes, major test coverage increases, and documentation.
* Maintained Jenkins builds across ~100 packages; debugged issues across DBs, Python APIs/CLIs, Docker, k8s, and HPC.
* Directly supported research teams: Investigated failed model runs, data access issues, and tooling gaps.
* Analyzed system usage and tool performance via ElasticSearch; ran ad-hoc reports and built automated report.

## Data Analyst, IHME – Pandemics Team, Nov 2021–Dec 2022

* Owned multiple data pipelines for COVID-19 production and reporting; reduced key runtime by 65%.
* Refactored vaccine and booster models for maintainability and extensibility.
* Built diagnostic tools and plots, with automated reporting for data QC; inspected and cleaned data.
* Led early UGE-to-Slurm transition for production jobs.
* Built ETL pipeline in R of national data into standardized, versioned output.
* Produced executive reports for world leaders, running ETL into R Markdown for PDFs; adapted to changing model outputs and feature requests.

# Education

B.S., Data Management and Data Analytics – Western Governors University, 2021

B.A., Critical Studies and Pedagogy – The Evergreen State College, 2008

# Certifications

* Quantum Computing (Qiskit & Python) – Udemy, Apr 2025
* Mathematics for ML & Data Science – DeepLearning.AI, Jan 2025
* Oracle Database SQL Certified Associate – Oracle, Jul 2020
* CompTIA A+, Network+, Project+ – CompTIA, 2019–2020