Yalemzewod Gelaw

📍115 Washington Street, Victoria Park, WA 6100 ✉yalassefa@gmail.com 📱+61469738918 🌐 https://yalemzewodgelaw.com

SUMMARY

I am a Senior Data Scientist at the WA Department of Health. I use Microsoft Fabric for data deployment and data visualization, SQL Server and Snowflake for data extraction and transformation, and R and Python for data management and machine learning. With over eight years of experience as a data analyst, researcher, and statistical modeler, I have worked with various data types in academia and public services. I have a solid background in epidemiology and biostatistics and am skilled in project management, stakeholder engagement, and effective communication.

EXPERIENCE

Senior Data Scientist | HQIU, WA Department of Health | January 2021 - Present, Perth, AU

- Managed and analyzed complex, large-scale, cloud-based data to better understand the cancer patient journey related to patient safety and clinical quality.
- Performed patient-level data extraction and transformation of large hospital data from Snowflakes and SQL server.
- Created data flow, data modelling and deployment pipeline in Power BI Desktop and Service
- Developed reports and dashboard using Power BI.
- Engaged with dashboard user groups to monitor and evaluate the quality of care provided and identify areas for improvement.
- Collaborated on machine learning models to identify diabetes-related discharges from discharge summary clinical synopsis.

Postdoctoral Research Officer | Telethon Kids Institute |May 2021 – October 31, 2024, Perth, AU

- Developed four pipelines for the annual production of global malaria burden estimates.
- Managed climate and health service datasets to spatiotemporally predict malaria risk transmission using the Bayesian Framework
- Designed, built, and maintained automated systems for surveys and surveillance data gathering, transforming and integration to reduce time and effort.
- Presented and published outputs in scientific publications, reports, and educational materials.
- Collaborated with international stakeholders and provided support in cleaning and utilizing local data gathered from 3,385 health centers for decision-making.
- Conducted spatial analysis in R and QGIS, prepared training manuals, and conducted training sessions on data management and spatial analysis for over 50 data analysts and program officers.
- Led data extraction, cleaning, manipulation, and cross-validation for 42 data points in R programming language.
- Provided R-programming technical support to research assistants, improving team productivity by preparing guide scripts.

Child health Biostatistician | University of New South Wales |March 2020 - May 2021, Sydney, AU

- Developed, documented, and implemented data cleaning, analytics, and reporting for multiple projects with longitudinal data points.
- Led data cleaning, transformation, integration, and analysis of datasets, resulting in successful publications.
- Applied decision tree analysis using a machine learning algorithms to identify influential risk factors for socio-emotional trajectories of children in NSW out-of-home care services using the R *caret* package.
- Provided statistical support to clinicians and PhD students on utilizing R-programming for hospital record data management, analysis, and reporting.

Epidemiologist | Queensland Department of Health | September 2019 - February 2020, Brisbane, AU

- Analyzed routine surveillance data to provide spatial and temporal patterns of weekly reportable disease for program recommendations.
- Offered analytical support for communicable disease unit project planning and program evaluation.

- Conducted monitoring and surveillance activities, including data collection, management, and analysis.
- Utilized ETL process to extract reportable diseases, examined and visualized data to identify patterns on maps for monitoring.

Lecturer | The University of Gondar | September 2013 - July 2016, Gondar, Ethiopia
- Delivered face-to-face biostatistics courses to undergraduate Nursing, Environmental Health, and Medicine students, as well as Master of Public Health Students.
- Mentored Master of Public Health students and Bachelor of Science (Nursing and Medicine) students.
- Led and coordinated health science students in field community professional practices.

EDUCATION

PhD in Epidemiology and Biostatistics, The University of Queensland, Brisbane, Australia (2016 – 2020)
- Thesis: Towards sustainable TB control in Ethiopia–profiling high-risk geographical areas using spatial modelling
- Key skills developed: Mathematical modelling, data visualization, statistics, machine learning, spatial modelling, meta-regression, predictive analysis, and scientific writing.

MPH in Epidemiology and Biostatistics, The University of Gondar, Gondar, Ethiopia (2011 – 2013)
- Thesis: Effect of residence on mothers' health care seeking behaviour for common childhood illness in Northwest Ethiopia: a community-based comparative cross-sectional study.
- Key skills developed: Field data collection, Data entry, EpiInfo, REDCap, Data analysis, survey data management, and scientific writing.

BSc in Environmental Health Science, Debub University, Hawassa, Ethiopia (2002 – 2006)
- Key skills developed: risk assessment, surveillance, data collection and community engagement.

SKILLS

- Programming language: R and Python
- Statistics, machine learning, predictive analysis, spatiotemporal analysis, data ingestion and modelling
- Data management & analysis: RStudio, Panda, SQL
- Data warehousing (snowflake)
- Data visualization: Power BI, MS Excel, Ggplot2, R-shiny, seaborn, matplotlib
- Communication & Negotiation
- Structural Thinking & Problem-Solving
- Scientific writing
- Leadership

PUBLICATIONS

- Over 45 peer-reviewed publications, with 18 first and senior-author papers. Full list available on Google Scholar.
- For conference presentations, grants and awards please review my webpage.

PROJECT PORTFOLIO

MS Excel dashboard

R and Python programming for Data Science (GitHub)

SKILL SET TRAINING

- Google Data Analytics Certificate – Coursera (2022)
- Applied data analysis skill set – Tafe (2024)

REFERENCE

Up on request