Oluwaseun Daniel Fowotade

+1 (541) 250-2647 | mailto:fowotadeoluwaseun15@gmail.com | LinkedIn |Website

SUMMARY

Versatile data scientist and statistician with 5+ years of experience in data analysis and 2+ years in statistical programming, specializing in R, Python, SAS, and SQL. Skilled in machine learning, feature selection, and experimental design with expertise in clinical trial data (SDTM, ADaM, TFLs), FDA/CDISC compliance, and bioinformatics (DESeq2, scRNA-seq, GSEA). Strong problem-solver and team player, with a track record of developing predictive models, automating workflows, and optimizing data-driven solutions.

SKILLS & CERTIFICATIONS

Programming: Python, R, SAS, SQL
Machine Learning & Statistical Modeling: Generalized Linear Regression, Feature Selection, Time Series Analysis, Statistical Methods, Multi-variate Analysis, Linear Models
Bioinformatics & Computational Biology: Git, Linux, Dimension Reduction, Enrichment and Cluster Analysis, scRNA-seq, GSEA
Data Visualization: Microsoft Excel, AWS QuickSight, Power BI, Spotfire
Cloud Platforms: GCP (experience with PySpark, BigQuery, DataProc, and DataPrep)
Clinical Data Standards: CDISC, SDTM, ADaM
Professional Certifications: SAS Programming 1: Essentials (April 2022), Medidata Clinical Programmer Certificates (April 2023), Google Data Analytics (June 2022)

STATISTICS & DATA ANALYSIS EXPERIENCE

Statistical Consultant | **Oregon State University, OR** September 2024 – June 2025

- Used **ANOVA**, **Poisson**, and **negative binomial regression** to analyze the effects of fertilizer treatments on fly egg-laying and attraction, helping identify the most effective pest control strategy.
- Applied **logistic regression (GLM)** to predict pollen fitness from genomic data, exploring feature selection methods specifically **LASSO** to improve model AUROC by 70% and interpretability.
- Utilized **statistical modeling,** including **multiple linear regression** and **Monte Carlo simulations**, to analyze the effects of ocean acidification on Olympia oysters, evaluating growth, survival, and shell dissolution rates, and identifying resilience factors contributing to a 20% higher survival rate compared to Pacific oysters.
- Implemented **experimental design analysis**, including **factorial analysis, multiple linear regression**, and **ANOVA**, to assess the effects of age, location, and fraction on biomass composition, identifying significant variations in lignin and sugar across locations.

Graduate Teaching Assistant | Oregon State University, OR **September 2023 – June 2025**

- Led recitation sessions for 30+ students on statistical methods using R, facilitating workshops and discussions on advanced techniques to enhance problem-solving skills and promote collaboration.
- Ensured confidentiality, evaluated and gave **constructive feedback**, fostered critical thinking, and promoted collaboration.
- Collaborated with 5+ other teaching assistants to ensure accuracy of grading practices across different course sections ensuring consistency in grading and clear communication across sections.
- Assisted in designing course activities and instructional materials aligned with learning objectives, supported applied learning through portfolio-based assignments, and provided student support via Canvas and Zoom office hours.

Statistical Programmer (Remote) | ClinFocus INC, Atlanta, GA April 2021 - August 2023

- Developed and validated analysis programs for **Phase I–IV tri**als, ensuring compliance with **regulatory standards** and alignment with **statistical plans and biostatistician guidance**.
- Transformed raw data listings into the **Study Data Tabulation Model (SDTM)** format using SAS programming.

- Programmed **ADaM** datasets and generated **TFL** outputs using SAS, ensuring accuracy through quality control reviews and thorough documentation.
- Ensured compliance with Regulatory Standards by providing final deliveries in accordance with **CDISC** requirements for **FDA** submissions.
- Designed and **implemented automation scripts** and tools in SAS to streamline repetitive statistical programming tasks, improving workflow efficiency and reducing manual effort.
- Reviewed and improved data management documentation, including **eCRF** designs**, vendor data specs**, and **edit checks**, ensuring accurate data collection for statistical programming.

Actuarial Analyst | Sanlam Insurance, Lagos State, Nigeria November 2019 – March 2021

- Cleaned and analyzed datasets using statistical techniques, generating reports on profitability, expenses, experience analysis, and premium collection rates.
- Conducted in-depth analysis of pricing, claims, and financial performance to ascertain competitiveness and provided valuable pricing recommendations based on the analysis.
- Communicated analytical insights by creating charts and graphs in Power BI, facilitating informed and strategic decision-making for the management team.
- Utilized various models such as quote models, Run-off Triangles, and Pricing Models to derive valuable insights and enhance decision-making processes.

PROJECTS

Developing an ADMM Algorithm for Microbiome Multi-Omic Networks (Masters Project) | R

- Developed an ADMM algorithm for gut microbiome data analysis, implementing structured optimization techniques to reduce dimensionality, improve predictive modeling, and enhance computational efficiency for biomarker discovery in multi-omic networks.

Netflix Data Analysis and Visualization | R

- Developed an interactive Shiny dashboard to analyze Netflix data (2008–2021), using trend detection and categorical data analysis to visualize content growth, geographic diversity, and rating distributions.

Survival Analysis of Treatment Efficacy in Primary Biliary Cirrhosis | SAS

- Developed SAS programs for survival analysis, including Kaplan-Meier estimation, Cox modeling, and Weibull AFT models, to assess treatment effectiveness for liver cirrhosis patients.

Feature Selection Using Genetic Algorithms for Regression | R

- Used a genetic algorithm for feature selection in regression modeling, reducing feature dimensionality from 60 to 27 predictors—a 55% reduction—while improving predictive accuracy and model interpretability.

Time Series Analysis of Indian Climate Data | Python

- Developed SARIMA models using Python to analyze and forecast climate trends in India.

Yelp Dataset Wrangling & Analysis | Python, SQL

- Processed large-scale JSON data into structured CSV format, optimized ETL workflows, and analyzed business reviews and user trends.

EDUCATION

Master of Science in Statistics Expected June 2025
Oregon State University, OR

Bachelor of Science in Mathematics 2018
University of Ilorin, Nigeria