

# Benson Cyril Nana Boakye

🌐 [bensoncyril123.github.io/portfolio](https://bensoncyril123.github.io/portfolio)    ☎ 541-286-0930    ✉ [nanaboab@gmail.com](mailto:nanaboab@gmail.com)    in [linkedin.com/in/cyrilbenson](https://linkedin.com/in/cyrilbenson)    📁 [github.com/bensoncyril123](https://github.com/bensoncyril123)

## EDUCATION

### Oregon State University

September 2023 - May 2025

*Master of Science degree in Statistics*

*Corvallis, Oregon*

### Kwame Nkrumah University of Science and Technology (KNUST)

August 2018 – September 2022

*Bachelor of Science in Actuarial Science*

*Kumasi, Ghana*

## SKILLS

**Programming Languages:** Python, R-studio, PostgreSQL, SAS, SAS Macros, SPSS.

**Machine Learning & Technologies:** Data Science pipeline(eg. Cleansing, wrangling, visualization, Predictive modelling, interpretation), Experimental Design, Clinical Data Standards (CDISC, CDASH, ADaM), DNAnexus, Time Series, Tableau, Tensorflow, Excel, AWS, Neural Networks, Hadoop, MongoDB, Data Structures, Power BI, Azure, Survival Analysis.

## EXPERIENCE

### Statistical Consultant

November 2024 – May 2025

*Oregon State University*

*Corvallis, Oregon*

- Optimized statistical designs for 10+ biomedical studies, including RCTs and longitudinal studies, reducing required sample sizes by 20% while preserving statistical power and study integrity.
- Analyzed genomic data (RNA-seq) using R and Python, identifying 5 biomarkers and contributing to publications.
- Advised clients on analyzing immune responses to TB antigens and their links to malnutrition, using logistic and GLM's.
- Evaluated the correlation between breath condensate pH, reflux events, and tracheal aspirate pepsin levels using SAS.

### Graduate Teaching Assistant

September 2023 – May 2025

*Oregon State University*

*Corvallis, Oregon*

- Introduced experimental design principles (randomization, blocking, factorial designs) to optimize data collection and analysis (ETL processes), improving students' ability to evaluate and apply research methodologies by 20%.
- Integrated SVMs, Decision Trees, and Python libraries to streamline practical exercises, cutting preparation time by 40%.
- Led recitation sessions for 500+ students on statistical methods using R, Python, & SAS, boosting performance by 15%.
- Ensured confidentiality, evaluated and gave constructive feedback, fostered critical thinking, and promoted collaboration.

### Data Science, ML & Actuarial Summer Analyst - Intern

September 2021 – November 2021

*Social Security and National Insurance Trust (SSNIT)*

*Accra, Ghana*

- Applied unsupervised machine learning in python for anomaly detection on 2TB of data, reducing process time by 20%.
- Built predictive models in Python(Tensorflow) to improve customer behavior analysis, increasing relevance by 24%.
- Wrangled 11TB of trading data in Hadoop with Scala, visualizing 16 datasets for 500+ clients to track liquidity impact.
- Communicated with executives to finalize required statistical analyses, summary tables, listings and figures (TLF's).

## PROJECTS

### Survival analysis of treatment efficacy in Primary Biliary Cirrhosis (PBC) | SAS

December 2024

- Analyzed Cyclosporin A's effect on PBC survival using Kaplan-Meier, improving prediction accuracy by 15%.
- Implemented AFT and Cox models to assess survival, identifying high bilirubin and low albumin as key risk factors.

### Predictive modeling and data analysis of breast cancer | Python

November 2024

- Designed and assessed 5 classification models (Logistic Regression, KNN, Random Forest, SVM with RBF kernel, and Gradient Boosting), achieving up to 94% precision in breast cancer through diagnosis using FNA image features.
- Achieved 98.25% accuracy in classifying breast cancer as malignant or benign using an SVM with an RBF kernel.

### Modeling Female Sex Worker distribution in Sub-Saharan Africa for HIV prevention | R-studio

October 2024

- Deployed mixed-effects models (negative binomial, zero-inflated, poisson, etc.) to analyze data, with an AIC of 2863.208.
- Applied REML estimation to confirm the inclusion of country, region, and dataYear as random effects ( $p < 0.0001$ ).

### Air quality analysis with R-shiny app | R-studio

January 2024

- Developed an interactive app to explore PM2.5 levels in cities across the USA and India, allowing users to analyze trends.
- Added event annotations, like COVID-19, to highlight air quality changes and guide public health decisions..

## LEADERSHIP AND EXTRACURRICULAR

- **Administrative Office Specialist** | *Kidspirit-USA* | Streamlined office operations, ensuring efficient workflow. **2024**
- **West Regional Welfare Head** | *GRASAG-USA* | Managed 700+ student resources, healthcare, and financial aid. **2023**
- **Research Assistant** | *Dept. of Statistics-KNUST* | Led a team of 4, assigning tasks and ensuring timeline adherence **2022**
- **Clubs and Societies President** | *KNUST* | Co-founded 7 clubs with 2000+ members and partnered with NGOs. **2022**
- **Vice President** | *Actuarial Science Association* | Led 500+ members, and initiated a "Food Bank" for 30 students. **2021**
- **Certifications** | Applied Data Science - World Quant Univ. | SAS Programming 1: Essentials | Data Science Foundation