EMMANUEL N. NARTEY

narte1en@vt.edu www.linkedin.com/in/emmarteyn

EDUCATION

- **PhD in Statistics and Analytics** May 2024
  Department of Statistics, Actuarial, and Data Sciences
  Central Michigan University, Mount Pleasant, MI
  Dissertation Title: Determining the Optimal K in Cluster Analysis and Application of Such Approaches to Feature Selection in Classification.
  Advisors: Dr. Carl Lee and Dr. Felix Famoye
- **Master of Science in Applied Statistics and Analytics** August 2020
  Department of Statistics, Actuarial, and Data Sciences
  Central Michigan University, Mount Pleasant, MI
- **Bachelor of Arts in Mathematics with Statistics** May 2016
  Department of Mathematics
  University of Ghana, Accra, Ghana

PROFESSIONAL EXPERIENCE

Center for Biostatistics and Health Data Science (VT), Roanoke, VA ***May 2024 - Present***

Research Scientist

- Collaborate with researchers at Virginia Tech and Virginia Tech Carilion School of Medicine (VTCSOM), focusing on study design, data wrangling, data analysis, writing statistical methods sections, performing sample size and power computations for grant proposals.
- Conduct appropriate statistical analyses using statistical languages and provide complex statistical programming and analysis support for large datasets.
- Mentor medical students, graduate students and student interns in their quantitative research studies.

Statistical Consulting Center (CMU), Mount Pleasant, MI ***August 2022 – May 2024***

Graduate Assistant

- Executed data cleaning and data integration for faculty and graduate students across diverse disciplines in the University especially clients in the Health and Medical Sciences, and the Education fields using SAS and R.
- Conducted statistical hypothesis testing and built inferential statistical models like GLM and Repeated Measures ANOVA with SAS, R, and SPSS to answer quantitative research questions.
- Summarized results from statistical analyses and document such evidence in manuscripts for publication.
- Promoted the reproducibility of data management and analysis by writing re-useable computer codes like SAS Macros, SPSS Syntax, and R functions.

Amazon.com, Bellevue, WA ***May 2023 – August 2023***

Data Engineer Intern

- Developed an automated anomaly detection mechanism to alert the data engineering team when outlier data points are observed in the time series of business metrics to mitigate against publishing questionable metrics in enterprise analytics systems.
- Built a robust and extensible time series-based anomaly detector using serverless cloud computing solutions that can process more than a million rows of data across multiple dimensions per day.
- Built a robust dashboard to visualize customers' traffic to the Re-Commerce platform, enabling the business subject matter experts to delve into data at intricate levels.
- Implemented natural language-driven analytics on a vital data repository, facilitating spontaneous visual generation for business users without creating additional dashboards.

Amazon.com, Bellevue, WA ***May 2022 – August 2022***

Data Engineer Intern

- Spearheaded a project to automate data quality checks using Deequ and AWS microservices, enhancing data integrity and operational efficiency.
- Developed and tested proof of concept for data unit testing on a Spark big data platform using Scala, demonstrating proficiency in both language and system.
- Collaborated with business partners to understand and develop metrics for data quality checks, ensuring alignment with business objectives and data accuracy.
- Implemented data quality anomaly detection on production-level datasets, identifying and addressing major limitations for optimized performance.
- Played a key role in developing and refining a new use case for anomaly detection, demonstrating adaptability and problem-solving skills in a dynamic environment.

Amway Corporation, Ada, MI ***May 2021 – August 2021***

Data Science Intern

- Conducted extensive exploratory data analysis, feature engineering, and developed segmentation models using BigQuery, GCP AI Platform, and Google Cloud Storage.
- Authored a comprehensive Technical Design Document, outlining approaches for machine learning model development, showcasing proficiency in technical documentation and planning.
- Developed a parametrized scoring algorithm for RFM segmentation in Python, contributing novel and reusable sets of code to the project's codebase.
- Built and optimized data pipelines for scalable and configurable model deployment, enhancing the efficiency and effectiveness of data processes.
- Collaborated effectively with data analysts to implement dashboard visualizations, providing new and relevant business insights on customer segments.

COMPUTING EXPERIENCE

- **Programming Languages**: Python, R, SAS, SQL, Scala
- **Software**: R Markdown, SAS 9.4, GitHub, Minitab, SPSS, MS Office, AWS, Google Cloud, Tableau, Looker Studio, Amazon Quicksight

RESEARCH INTERESTS

Supervised and Unsupervised Machine Learning: Feature Selection, Cluster Analysis, Interpretable ML; Statistical Computing; Generalized Linear Models; Complex Sample Design and Analysis; Survival Analysis

PUBLICATIONS, ABSTRACTS, PRESENTATIONS, WORKSHOPS AND GUEST LECTURES

Journal Article

- Palit, S., Sufyani, T., Inungu, J. N., Cheng, C. I., & **Nartey, E.** (2024). Behavioral Determinants of Childhood Obesity in the United States: An Exploratory Study. *Journal of Obesity*, *2024*(1), 9224425. https://doi.org/10.1155/2024/9224425
- Ameh, G., **Nartey, E.**, Inungu, J., & Shayestah, J. (2023). Racial Disparities in Oral Health; Analysis of 2020 Behavioral Risk Factor Surveillance System. *Acta Scientific Dental Sciences*, Vol. 7, pp. 29-39. DOI: 10.31080/ASDS.2023.07.1551

Works in Progress

- Kolb, R., **Nartey, E.**, Lozano, A., Hanlon, A., Ramirez, V., & Parma, V. Smell and taste dysfunction with GLP-1 RAs in the FDA adverse event reporting system: a pharmacovigilance assessment, (manuscript in preparation)
- Gearhardt, A., Hutelin, Z., **Nartey, E.**, Ahrens, M., Baugh, M., Fazzino, T., LaFata, E., Sonneville, K., & DiFeliceantonio, A. Nutritional characteristics of foods with addictive potential: a machine learning approach, (manuscript in preparation)
- Vaughn-Cooke, M., & **Nartey, E.** Beyond Diagnoses: Identifying Functional Subgroups Among Adults with Chronic Conditions Using NHANES Data, (manuscript in preparation)
- Dell, G., Landgraf, K., Ngounly, C., Choi, H., Waymack, S., **Nartey, E.**, Hanlon, A., Patterson, A., & Faltin, F. An Evaluation of FDA's FAERS Database's Ability to Identify Safety Signals Over Time Using Montelukast's Neurotoxicity Issues, (manuscript in preparation)
- **Nartey, E.**, Lee, C., & Famoye, F. A comparative review of some internal validation techniques for determining the number of clusters in numeric data, (manuscript in preparation)
- **Nartey, E.**, Lee, C., & Famoye, F. Numeric feature selection via the average silhouette score towards optimal classification in machine learning, (manuscript in preparation)

Abstracts

- Gainey, M., Stettinius, A., Holmes, H., **Nartey, E.**, Stayrook, S., Maxwell, A., Vlaisavljevich, E. & Rao, J. (2025). Evaluation and Characterization of a Novel DNA Lysis Technique for Extraction of Nontuberculous Mycobacteria DNA. *IDWeek 2025 Annual Meeting*.
- Kolb, R., **Nartey, E.**, Lozano, A., Hanlon, A., Ramirez, V., & Parma, V. (2025) Smell and taste dysfunction with GLP-1 RAs in the FDA adverse event reporting system: a pharmacovigilance assessment. *16th Pangborn Sensory Science Symposium.*
- Gearhardt, A., Hutelin, Z., **Nartey, E.**, Ahrens, M., Baugh, M., Fazzino, T., LaFata, E., Sonneville, K., & DiFeliceantonio, A. (2025) The Addictive Potential of Ultra-Processed Foods: Identifying Key Nutritional Predictors. *SSIB 2025 Annual Meeting.*
- Kolb, R., **Nartey, E.**, Lozano, A., Hanlon, A., Ramirez, V., & Parma, V. (2025, March) Reports of Smell and Taste Adverse Events from GLP-1 RAs. *AChemS 2025 Annual Meeting.*

- Hanlon, A., **Nartey, E.**, Grubb, C., Lyons, G., Ratcliffe, S. & Lozano, A. (2025, September) CUBE: A Collaborative Biostatistics Experience to Bring Diversity and Awareness to the Field of Collaborative Biostatistics. *StatFest 2024*.
- Salski, S., Yahia, N., Cheng, C., & **Nartey, E.** (2023, April). Will physical activity act as a coping mechanism to reduce stress eating and improve body image concerns among college students during the COVID-19 pandemic? *Health Professions Annual Research Symposium, Central Michigan University*.
- **Nartey, E.**, & Lee, C. (2021, April). Sales Volume Forecasting with Statistical and Machine Learning Models. *Student Creative and Research Endeavors Exhibition (SCREE), Central Michigan University*.

Presentations

- **Nartey, E.** (2025, June). "Interview Panel: Navigating How to Get a Job." Delivered at Quantitative Staff Network Meeting.
- **Nartey, E.** (2025, April). "Strategic Internships: Securing the Right Opportunity & Maximizing Your Experience as a Statistics and Science Student." Delivered at University of Ghana Actuarial & Statistics Society Career Fair Series.
- Hanlon, A., **Nartey, E.**, Lozano, A., Pomann, G.M., Sebro., R., & Coles, A. (2025, March). "Advancing Mid-Level Collaborative Biostatisticians in Historically Marginalized Groups: Challenges, Recommendations, and Strategies." Chaired Panel at ENAR 2025 Spring Meeting.
- Oaks, A., Patippe, C., **Nartey, E.**, Lee, J.M., & To, M.A. (2024, September). "The Graduate Student Experience." Contributed to Panel at StatFest 2024.
- **Nartey, E.** (2021, November). "CRISP-DM: A Framework for Executing Data Science Projects." Delivered at Statistics and Data Science Club Meeting, Central Michigan University.
- **Nartey, E.** (2020, October). "Interactive Data Visualization – A Tableau Use Case." Delivered at Statistics and Data Science Club Meeting, Central Michigan University.

Conducted Workshops and Guest Lectures

- **Nartey, E.** (2025, February). "CUBE: Bringing Awareness to Biostatistics and Health Data Science." Workshop conducted at the Uplifting Black Men Conference 2025, Virginia Tech.
- **Nartey, E.** (2024, November). "Working With a Collaborative Biostatistician and Ethical Data Analyses" Guest lecture for TBM 5105 Professional Development and Ethics Course, Virginia Tech.
- **Nartey, E.** (2023, November). "SQL Technical Workshop" Workshop delivered at Statistics and Data Science Club Meetings, Central Michigan University.
- **Nartey, E.** (2022, April). "MPH Statistical Software Workshop: A Two-Day Hands-On Session Introducing Public Health Graduate Students to Data Management and Statistical Analysis in R." Workshop conducted at College of Health Professions, Central Michigan University.

Peer Review

- Journal of Clinical and Translational Science – 2025
- Journal Reviewer for BMC Health Services Research – 2023 – 2024

TEACHING EXPERIENCE

Virginia Tech, Roanoke, VA ***June 2024 – July 2024***

R Programming Mentor

- Revised the content for the introductory R programming course for the Collaborative Undergraduate Biostatistics Experience (CUBE) program.
- Taught and mentored nine students to use the R language for their collaborative research projects during CUBE 2024.

Central Michigan University, Mount Pleasant, MI ***August 2018 – May 2022***

Teaching Assistant (Instructor of Record)

- Devised and implemented lesson plans for Introductory Statistics, consistently teaching and engaging sixty students each academic year, resulting in a 90% average pass rate under my instruction.
- Earned a 3.96/5.0 average on student opinion surveys, reflecting a teaching quality consistently above the department mean.
- Conducted weekly office hours and tutoring sessions at the university Math Assistant Center, complemented by completing advanced instructor training modules.
- Assessed and graded a wide range of student assessments, including quizzes, project reports, and exams, ensuring timely and constructive feedback.
- Mentored new teaching assistants, providing guidance and support to uphold the department's educational standards.

University of Ghana, Accra, Ghana ***September 2016 – July 2017***

Teaching Assistant

- Facilitated tutorial sessions in Algebra, Trigonometry, and Calculus I for undergraduate students, enhancing the understanding and performance of in these core mathematical areas.
- Assessed and graded a variety of assignments and exams, providing detailed feedback to support continuous student improvement.
- Collaborated with faculty members to prepare and organize educational materials, contributing to the streamlined delivery of the mathematics curriculum.

MENTORING

VTCSOM, Roanoke, VA **June *2024 – Present***

Biostatistics Mentor

- Provide statistical mentorship to medical students on research and inquiry projects, supporting study design, power and sample size estimation, data analysis, and results dissemination.
- Current mentees include Blaire Barton, Hiam Baidas, Tristan Colaizzi, Kestrel Green, Mia Iqbal, Colby Mallett, Ailene Edwards, Emma Frye, Monique Gainey, Carter Gottschalk, Lucy Honeycutt, Varun Mishra, Jacqueline Urdang, and Isaiah Yim.
- Former mentees include Kristina Gueco, Brian Hansen, Paula Lewis, Tyler Shick, and Raymond Uymatiao.

Virginia Tech, Roanoke, VA ***January 2025 – May 2025***

Undergraduate Research Award Mentor

- Mentored Gabriel Dell on an undergraduate research project focused on advanced statistical signal detection in FAERS data, integrating parallel computing to improve dashboard performance.
- Guided manuscript development and supported a dashboard prototype to enhance the identification of emerging drug safety signals.

Virginia Tech, Roanoke, VA ***August 2024 – May 2025***

CMDA Capstone Sponsor

- Sponsored two undergraduate Capstone teams, providing curated datasets and mentorship for the following projects: "Patterns of Psychiatric Complaints: A Comparative Analysis of Singulair and Other Asthma and Allergic Rhinitis Drugs", and "An Evaluation of FAERS using Disproportionality and Time-Series on Multiple Sclerosis drugs".
- Guided students on statistical methods and analytical strategies throughout the project lifecycle and offered feedback on preliminary and final results.
- Team 1: Hyunjoon Choi, Gabriel Dell, Katelyn Landgraf, Coby Nguonly; Team 2: Edan McDonald, Shweatha Rameshkumar, Yimmi Tran, and Jackson Walsh.

AWARDS AND CERTIFICATIONS

- AWS Certified Cloud Practitioner (*October 2023*)
- The Carpentries Instructor Training Program Sponsorship (*September 2023*)
- Outstanding Doctoral TA (*2021/2022 Academic Year*)
- SAS Certified Predictive Modeler Using SAS Enterprise Miner 14 (*December 2019*)
- SAS Certified Base Programmer for SAS 9 (*December 2018*)

LEADERSHIP AND SERVICE

Statistics and Data Science Club, Central Michigan University (CMU), Mount Pleasant, MI

- **Treasurer, Former President, Former Vice President** (*August 2021 – May 2024*)
    - Propelled the club's strategic direction by overseeing financial resources, planning, and managing club activities, including executive board meetings and industry events.
    - Championed the club's growth by organizing special events technical workshops and securing industry experts to provide enriching learning experiences for members.

CMU Baseball Analytics Club

- **Member** (*August – December 2021*)
    - Conducted thorough analysis of Rapsodo Pitching and Hitting data, providing weekly reports that assisted in the optimization of team performance.

2019 Special Olympics, Mount Pleasant, MI

- **Event Volunteer** (*May 2019*)
    - Coordinated logistics and officiated at Bocce games, ensuring smooth operation and a positive experience for participants.

iVolunteer Network Ghana, Accra, Ghana

- **Co-Founder** (*February 2016 – July 2018*)
    - Spearheaded a volunteer initiative to bridge educational gaps, leading teaching programs in Mathematics, English Language, and Integrated Science for underprivileged middle schools.

PROFESSIONAL MEMBERSHIPS

- East North American Region (ENAR), International Biometric Society
- American Statistical Association (ASA)
- Whole Health Consortium
- Cochrane Collaboration
- The Carpentries
- Institute for Operations Research and the Management Sciences (INFORMS)