ANNIE (NGHI) NGUYEN

(513) 690-1854| nghi.npv@gmail.com | LinkedIn | Github

EDUCATION

Miami University | Oxford, OH
B.Sc. Data Science and Statistics | GPA: 3.87 *Expected May 2025*
Minor: Finance, Computer Science

Coursework: Regression Analysis, Advanced Data Visualization, Probability, Inferential Statistics, Database Systems, Data Structure and Algorithms, Statistical Programming, Statistical Learning

SKILLS

- Programming languages: Python, R, SQL, Java, SAS, Processing
- Tools: Excel, PLINK, Tableau, Snowflake

EXPERIENCE

MIAMI UNIVERSITY Oxford, OH
Computer Science Department Teaching Assistant *Jan 2024 - Present*

- Collaborate with the professor to evaluate and grade labs and assignments for 40 students via Canvas.
- Assist 40 students during lab sessions and lectures, facilitate learning of Processing and Java, and aid in understanding course material and completing assignments.

OXFORD UNIVERSITY CLINICAL RESEARCH UNIT Remote
Research Assistant *May 2023 - Aug 2023*
The effectiveness of tafenoquine for treating P. vivax malaria using individual patient data.

- Optimized 727K+ malaria clinical records, incorporating PCA dimensionality reduction method.
- Incorporated quality control and population stratification analysis on data from a cohort of 900 patients.
- Deduced *tafenoquine*'s effective duration from Bayesian mixed-effects model to support previous findings.

LEE LAB Oxford, OH
Research Assistant *Mar 2022 - Feb 2024*
Prioritizing disease candidate genes in GWAS utilizing knockout mouse phenotype data

- Conducted meta-analysis on 14 million most frequent genetic variations in humans and gene-edited mice.
- Summarized test statistics from 115 human physiology traits for 9 million genes using Simes method.
- Developed biostatistical website **PRIDIC** to promote efficient data visualization for public accessibility from integrated data.

PROJECTS

Health Vitality App | Github

- Developed a stroke liability prediction model based on 5,000 observations from a stroke prediction data set
- Created a user-friendly body fat calculation model with health data inputs, referencing 250 body fat data
- Visualized worldwide global health trends from 6,500 data using advanced interactive plots with RShiny

Improving Legal Services Customer Experience: ABA Raw Data for Predictive Analysis | Github

- Visualized key insights from customer demographics, question categories, and sentiment analysis
- Identified Random Forests as the optimal model for resource optimization and enhanced service quality at the ABA using Elastic Nets, Random Forests, and Neural Networks on 400,000+ data inputs.

DataViz: World's Cancer Occurrences 1990-2019 | Report

- Illustrated insights into global cancer trends and regional disparities from 7,000+ observations with R
- Explored cancer mortality rates in world regions, identifying variations and potential correlations with prosperity, and elderly population for better understanding of cancer epidemiology

ASK ME ABOUT

- Solitaire card games, Badminton, Building Lego models