

Naveen Kannan

RESEARCH ASSOCIATE

Cleveland, Ohio | United States Permanent Resident

+1 973 519 5397 | naveen.kannan12@gmail.com | naveenkannan.netlify.app | [naveenk2022](#) | [naveen-kannan-177120118](#)

Data Scientist experienced with automating and building scientific workflows, data analysis, visualization and Data Engineering. I have a strong passion for innovation and problem solving. I enjoy learning new skills and blogging about them on my website.

Work Experience

Research Associate

Cleveland, Ohio

POPULATION HEALTH AND QUANTITATIVE HEALTH SCIENCES, SCHOOL OF MEDICINE, CASE WESTERN RESERVE UNIVERSITY

MAY 2023 - Present

• Docker/Singularity Container Deployment

- Developed efficient Docker/Singularity containers for reproducible scientific pipelines, facilitating efficient workflows for colleagues, reducing container build times by 300 percent.
- Enabled the use of tools such as RStudio and JupyterLab (with Bioinformatics tools installed on them) through Docker containers running servers with these tools enabled, circumventing the need to have to install these tools on the server directly.
- Transferred the Docker storage cache to a physical hard drive while ensuring existing images were preserved, eliminating all future storage errors.

Cluster Configuration and Big Data Analytics

- Set up, configured and maintained both VM (4 node cluster) and Raspberry Pi (4 node cluster) bare-metal servers.
- Automated the installation and configuration of Big Data architecture such as Hadoop and Spark along with bioinformatics tools like Hail and Glow using Ansible scripts.

Genomic data Analysis

- Collaborated with the Alzheimer's Disease Sequence Project by annotating and scoring the biological effects of 362 million structural genetic variants associated with Alzheimer's Disease.
- Performed statistical analysis on the Data and presented findings in the form of a poster at the ASHG (American Society of Human Genetics) 2023 Conference in Washington DC.

Graduate Research Assistant

Cleveland, Ohio

POPULATION HEALTH AND QUANTITATIVE HEALTH SCIENCES, SCHOOL OF MEDICINE, CASE WESTERN RESERVE UNIVERSITY

JANUARY 2023 - MAY 2023

- Engineered Docker containers for streamlined deployment of specialized R packages and Python modules, optimizing analyses for single-cell RNA datasets.

Graduate Research Assistant

Cleveland, Ohio

GIS HEALTH AND HAZARDS LAB, SCHOOL OF MEDICINE, CASE WESTERN RESERVE UNIVERSITY

MAY 2022 - DECEMBER 2022

• Mapping human movement patterns after a natural disaster in the Democratic Republic of Congo.

- Developed Python scripts automating GPS coordinate extraction from GPX files of geospatial video feeds from refugee camps in the Democratic Republic of Congo post the volcanic eruption of Mt Nyiragongo.
- Processed and transformed data into a machine-readable format, enabling streamlined and automated map generation.
- Applied automated scripts to parse video feeds into frames, implementing Machine Learning Algorithms for tent detection.
- Created heat maps of tent distribution according to GPS coordinates, allowing for the mapping of the human movement pattern in the aftermath of the eruption.
- Wrote up the findings as a part of a manuscript for a peer-reviewed publication.

Junior Resident Doctor

Chennai, India

DEPARTMENT OF PSYCHIATRY, SAVEETHA MEDICAL COLLEGE

JULY 2020 - SEPTEMBER 2020

• Diagnosing and managing patients with psychiatric illnesses

- Learned to evaluate and diagnose psychiatric illnesses in stressful environments.
- Assigned and achieved a goal of management of five patients per day.

Junior Resident Doctor

Chennai, India

MADRAS MEDICAL COLLEGE

MARCH 2019 - MARCH 2020

- Rotated through the following departments: Internal Medicine, General Surgery, Obstetrics and Gynecology, Community Medicine, Psychiatry, Emergency Trauma Ward, Labor Ward.

Education

MSc in Biomedical and Health Informatics

Cleveland, Ohio

CASE WESTERN RESERVE UNIVERSITY

JANUARY 2022 - AUGUST 2023

- GPA-3.9

Skills

Operating Systems

UNIX, KUBUNTU DEBIAN LINUX, FEDORA RED HAT LINUX, RASPBAN DEBIAN LINUX, WINDOWS 11

Programming Languages

PYTHON, R, BASH, SQL

Software

HADOOP/SPARK/HIVE, JUPYTER, DOCKER, SINGULARITY, SLURM, ANSIBLE

Geospatial Analytics/Visualization

GEOPANDAS, SF (R), QGIS

Markup Languages

LATEX, MARKDOWN, RMARKDOWN

Version Control

GIT

Python-based skills

PYSARK FOR HADOOP/SPARK, TENSORFLOW, PYTORCH, PANDAS, SCIKIT-LEARN, NUMPY, SCIPY

Bioinformatics tools

SAMTOOLS, BCFTOOLS, HTSLIB, TABIX, PLINK

Teaching Experience

Teaching Assistant

A DATA DRIVEN INTRODUCTION TO GENOMICS AND HUMAN HEALTH (PQHS 451)

- Held weekly office hours, developed and graded homework.

Case Western Reserve University

AUGUST 2023 - DECEMBER 2023

Leadership

Vice President

STEINER HOUSE INTERNATIONAL CO-OP

- **Community Leadership**

- Facilitated seamless integration of new residents, fostering community through monthly social events.
- Collaborated with the executive board to ensure effective communication and daily management of the house.
- Organized monthly events, contributing to a vibrant and inclusive living environment.

Admissions Manager

STEINER HOUSE INTERNATIONAL CO-OP

- **Residential Admissions and Selections**

- Screened and interviewed prospective residents, facilitating transparent communication of results within the house.
- Successfully admitted diverse graduate students globally, including those from India, Bangladesh, China, Vietnam, Thailand, Nigeria, Ireland, and Iran.
- Oversaw the onboarding process for 8 graduate students, ensuring a smooth transition to life in the United States at Steiner House.

Ordering Manager

STEINER HOUSE INTERNATIONAL CO-OP

- **Inventory Management**

- Managed inventory for Steiner House, with a capacity of 21 students.
- Responsible for ensuring a steady supply of consumables and non-consumables for the smooth functioning of the House.

Projects

- https://naveen-kannan.shinyapps.io/final_east_palestine_dashboard_2
 - Coded scripts to scrape publicly available air and water quality data from USGS(United States Geological Survey) and AirNow AQI for counties within a 30-mile radius of the Norfolk Southern train derailment in February 2023.
 - Implemented an interactive Shiny dashboard in R, offering users the ability to explore and analyze daily trends in Ohio, West Virginia, and Pennsylvania counties throughout January and February.
 - The dashboard includes dynamic time series plots for comprehensive visualization of air and water quality metrics during the specified timeframe.

Data Science Blog

Cleveland, Ohio

- https://naveenkannan.netlify.app/blog_main.html
 - Authored tutorials on topics ranging from Docker/Singularity container creation to data engineering tutorials on setting up Spark and Hadoop.

Publications

- Jayakrishnan Ajayakumar ^a, Andrew J. Curtis ^a, Felicien Maisha ^b, Sandra Bempah ^c, Afsar Ali ^{b,d}, J. Glenn Morris Jr. ^{b,e}, **Naveen Kannan** ^a, and Grace Armstrong ^a, ^aDepartment of Population and Quantitative Health Sciences, School of Medicine, Case Western Reserve University, Cleveland, Ohio, USA; ^bEmerging Pathogens Institute, University of Florida, Gainesville, Florida, USA; ^cDepartment of Geography, Kent State University, Kent, Ohio, USA; ^dDepartment of Environmental & Global Health, College of Public Health and Health Professions, University of Florida, Gainesville, Florida, USA; ^eCollege of Medicine, University of Florida, Gainesville, Florida, USA, Using Spatial Video and Deep Learning for Automated Mapping of Relief Camps. [Manuscript under review]
- **N. Kannan**¹, N. Wheeler¹, Genome Center for Alzheimer's Disease, L-S. Wang², Y. Leung², W. S. Bush¹; ¹Cleveland Inst. for Computational Biology, Dept. for Population and Quantitative Hlth.Sci., Case Western Reserve Univ., Cleveland, OH, ²Dept. of Pathology and Lab. Med., Penn Neurodegeneration Genomics Ctr., Perelman Sch. of Med., Univ. of Pennsylvania, Philadelphia, PA., Annotation and scoring of the deleteriousness of individual genetic variants in the 4th release of the Alzheimer's Disease Sequencing Project;(PB4451). Presented at the Annual Meeting of The American Society of Human Genetics, November 3, 2023, in Washington DC.
- Suebsarn R., W. T., **Naveen K.**, Ellis W., Atsadaporn N., Chitchanok B., Chantira C., Christine S., Heba A. Using regression analysis and machine learning approach to understand the association between comorbidities, medication nonadherence, activity of daily living, and heart condition status among older adults in the United States. [Manuscript under review]
- Segamalai, D., Abdul Jameel, A. R., **Kannan, N.**, Anbalagan, A., Duraisamy, B., Raju, P., & Devy Gounder, K. (2017). Mediastinal pseudocyst: varied presentations and management—experience from a tertiary referral care centre in India. *HPB Surgery*, 2017.
- Vellaisamy, R., **Kannan, N.**, Anbalagan, A., Raju, P., Duraisamy, B., Murugesan, C. S., & Gounder, K. D. (2016). Endoscopic access to hepatic duct through duodenum during follow up—after primary surgery for hepatolithiasis. *HPB*, 18, e530.
- Ramasamy, V., Vellaisamy, R., **Kannan, N.**, & Gounder, K. D. (2016). Refined technique of access loop in hepatobiliary surgery. *HPB*, 18, e593-e594.
- **Kannan, N.**, Vellaisamy, R., Govindarajan, M., & Gounder, K. D. (2016). Pellagra following pancreaticoduodenectomy for malignant pancreatic carcinoid with pluripotent hormonal potential. *HPB*, 18, e381-e382.