Smita Misra

DATA SCIENTIST

PROFILE

Accomplished Scientist with over two decades of experience in project planning, hypothesis-driven research design, implementation, and reporting. Highly skilled in identifying high-impact analytical problems and finding innovative ways to tackle problems *via* data analysis, experimentation, and statistical modeling. Skilled in various data analytical tools, advanced statistical methods, and Machine learning algorithms. An inquisitive and curious learner who enjoys data-driven storytelling.

DATA TECHNICAL TRAINING

Nashville Software School / Nashville, TN
Data Science Apprentice, Part-time SEPTEMBER 2022 - PRESENT

Intensive 9-month boot camp emphasizing the practical application of data science technical skills to real-world business questions and challenges. Utilizing data from different domains like healthcare, retail, and government

- Performed data cleanup and exploratory data analysis (EDA) using different packages and libraries from Python (*pandas, numpy, matplotlib, scikit-learn, statsmodels, SQLite*) and R (*tidyverse, dplyr)*
- Employed EDA to excavate relationships among the variables, design pertinent questions, and derive influential outcomes.
- Applied statistical methodologies to analyze trends and significant data relationships.
- Created data visuals including graphs, maps and networks (*plotly express, ggplot, Shiny, leaflet, Neo4j*) for storytelling and presentations.
- Examined data sets using advanced queries in PostgreSQL.
- Used SQLite for database creation and extracting utilizable table for advance analysis.
- Performed geospatial analysis (*geopandas*, *leaflet*)
- Performed web scraping using (*BeautifulSoup)* and API keys (*Requests*).
- Performed network analysis on graph data (*Neo4j*)
- Used Machine Learning for generating models, interpreting them, and refine methodologies to improve them.

DATA PROJECTS

- ***Medical Deserts: Exploring Access to Healthcare and its Impact on health outcomes*** - *R*, *Shiny*, *Python*

  For this project data regarding the nationwide distribution of hospitals and related information was collected from the Centre for Medicare and Medicaid database. Population-level data was gathered from the US Census Bureau, American Community Survey, and TN.gov website. Census-level data were extracted at the tract level for Davidson County, Nashville, TN. Google Maps API and US census shape files were utilized for geographical mapping and analyzing the distribution of hospitals per state at the census level and tract level for TN regions. Statistical analysis was performed to analyze the impact of predictors variables like distance of facility, PCP/100K people, and socioeconomic status, etc. on the outcome of overall population health like rate of preventable hospitalization, physically unhealthy days, and years of potential life lost.

CONTACT

smitamisra74@gmail.com
615.275.9965
Smita Misra | LinkedIn
Smita Misra | github.com

SKILLS

Python (Pandas, NumPy, matplotlib, seaborn, Plotly, geospatial models
SQL (PostgreSQL, SQLite)
R (ggplot2, tidyverse, dplyr, RShiny)
GitHub
NLP
Machine Learning
GraphPad Prism
Agile Methodology
Strategic Planning
Program Management
Public Speaking
Genomics
Transcriptomics
Proteomics
Structure-Activity Relationship
Drug Design
Gene Expression Regulation
Health Disparity Research
Genetic Manipulation

EDUCATION

Data Science Certificate
Nashville Software School
Graduation, June 2023
Doctor of Philosophy, Biochemistry, Central Drug Research Institute, Lucknow India
Master of Science, Biochemistry RMLA University India
Bachelor of Science, Biological Sciences University of Lucknow India

TRAINING & CERTIFICATION

Biostatistics Meharry Medical College
Cancer Transcriptomics Praxis AI digital learning platform.
Introduction to Python Programming, Certstaffix Training
Advanced Python Programming, Certstaffix Training

LIST OF BIOMEDICAL PUBLICATIONS

- ***Hop Teaming Analysis:*** *Python, SQLite, Neo4j*
  For this project, we used the Hop Teaming dataset which aims to capture referrals between healthcare providers based on Medicare claims. We used the NPPES data to identify the provider using their NPI (National Provider Index). We used these data to decode the referral patterns among different specialties within Nashville/Davidson County and draw network graphs.
- ***Impact of Affordable Housing Developments on single-family home sales in Davidson County*** *– R, tidyverse, simple features, leaflet, Shiny*
  We used the single home sale data from the Metro Nashville Planning Department, the Low-income housing tax credit information from the Department of Housing and Urban Development (HUD) database. We used the county-level shape files to find the proximity of the development to the single-family homes and analyzed its impact on home prices.
- ***Vanderbilt ACCRE User Analysis*** *- Python, plotly express*
  Vanderbilt ACCRE provided data on jobs performed using its high-performance computing resources in order to complete an analysis of user activity generally and inefficient user activity specifically.
- ***LahmanBaseball*** *- SQL*
  Completed a thorough analysis of a baseball database consisting of 20 tables and over 100 years' worth of data.

PROFESSIONAL EXPERIENCE:

- **Associate Professor** (Tenure Track, February **2022** - Present**)**, School of Graduate Studies and Research, Center for Women's Health Research, **Meharry Medical College,** Nashville, TN, USA. **Research Focus:** Stress-mediated regulation of RNA turnover and development of resistance, breast cancer research, non-coding RNA biology, Cancer transcriptomics, TCGA, Health disparity in cancer onset and outcome, Obesity, and Metabolic disorder**. Full-Time**
- **Assistant Professor** (Tenure Track, July **2012** - January **2022)**, School of Graduate Studies and Research, Center for Women's Health Research, **Meharry Medical College,** Nashville, TN, USA. **Research Focus:** Cancer transcriptomics, Stress mediated regulation of RNA turnover, breast cancer research, Redox regulation, non-coding RNA biology**. Full-Time**
- **Instructor** (July **2008** - June **2012**), Center for AIDS Health Disparities Research (CAHDR), **Meharry Medical College** School of Medicine**,** Nashville, TN, USA. **Research Focus:** DNA damage response, Gene expression regulation, breast cancer**. Full-Time**
- **Postdoctoral Fellow** (July **2003** - June **2008**), Department of Microbiology and Immunology School of Graduate Studies and Research Medicine, **Meharry Medical College** Nashville, TN, USA. **Research Focus:** Host-Pathogen interaction, manipulation of host macrophages by protozoan parasite *Leishmania*, antisense therapy-RNaseH. **Full-Time**
- **Research Associate**, Department of Microbiology and Immunology, School of Medicine, Meharry Medical College, Nashville, TN, **Full-Time** (July **2002** – June **2003**)

CURRENT JOB RESPONSIBILITIES: Associate Professor, Meharry Medical College(MMC)

- Provide scientific leadership and expertise in project design, development, management, and execution, allocation of appropriate resources, managing budget, supervising the research team to ensure scientific soundness, and technical feasibility of the project, interpreting results, and preparing reports and manuscripts
- Reviewed and combined multiple sources of information, research articles, reports, and other publications, to gain a broader and deeper understanding of a particular subject.
- Analyze findings during the research by examining, cleaning, transforming, and modeling data with the goal of discovering useful information, drawing conclusions, and supporting decision-making.
- Working on clinical data. Collecting multiple pieces of data from various sources to form a coherent and comprehensive understanding. Combing various data sets, transforming, cleaning the data, and then using statistical and machine learning methods to identify patterns, relationships, and insights.
- Member of the Institutional overall surveillance committee & Admissions Committee
- Chairperson and member of multiple Graduate Student Research thesis advisory committees.
- Communicate progress, plans, evaluations, and execution of research with the leadership.
- Conducted experiments and examined medical data to identify acute and chronic diseases.
- Conduct hypothesis-driven reach to understand the molecular contributions of disparity in breast cancer progression and response to chemotherapeutic agents (e.g., DNA damaging agents).
- Performed Genomic/RNA Sequencing activity and planned Next Generation Sequencing (NGS) to identify genomics/transcriptomics alteration/dysregulation in tropical diseases and cancer research.
- Provide advice on work methods, practices, and procedures.
- Conduct active research, performed experiments, generated data, and performed data analyses).
- Develop review procedures & training to ensure scientific research protocol standards are maintained.
- Published & presented scientific findings in peer-reviewed scientific journals and conferences.