

ANAMIKA SHARMA

Data Scientist

CONTACT

716-867-6020

sanamica@gmail.com

linkedin.com/in/anamikas123

github.com/sanamica

SKILLS

- **Python:** data manipulation and exploratory data analysis (EDA) using pandas, numpy, regex, scikit-learn, matplotlib, seaborn, web scraping and html data manipulation with beautifulsoup and requests.
- **R:** data manipulation and EDA with dplyr, tidyverse ggplot2, web analysis, shiny app development
- **PostgreSQL:** database import and EDA, complex joins
- **Cypher :** NoSQL / graph database (using Neo4J) manipulation (incl. complex queries)
- **Microsoft Suite:** MS Excel; MS Access; MS Word, MS SQL Server
- **AWS,** SageMaker
- **SAS, SPSS**

PROJECT - Coursera

Capstone Project:

Used **SAS** to analyze the Step'n Out data and designed model to predict drug use among parolees after 9 months of being in the STEPNOOUT program.

EDUCATION

Data Science Boot Camp

Nashville Software School

Data Analysis and Interpretation Specialization

Wesleyan University on Coursera

Candidacy to Philosophy Sociology

State University of New York at Buffalo

M.A. Sociology

Jawaharlal Nehru University, New Delhi, India

TECHNICAL EXPERIENCE

Data Scientist Apprentice

Nashville Software School

Aug 2019 -present

Nine month part-time intensive data science course with emphasis in getting and organizing data, exploratory data analysis, building models and testing hypotheses, machine learning, data. visualization and storytelling using Python, R, and SQL. Projects based on real world data.

Capstone Project: Analyzed Yelp data along with income data from census to explore relation between neighborhood income and restaurant feature. Developed location based recommender and model to predict median income.

Mid-course Project: Designed **R Shiny** app to analyze how socioeconomic and environmental factors influences obesity along with county level map.

anamika.shinyapps.io/Obesity_food_environment/

- Parsed Edgar filling for Alliance Bernstein to get required fields using **NLP** in **Python**.
- Created model to predict derates on trucks for Big G Express and the prediction was able to save 46% of cost caused by derates.
- Compared referring pattern between different specialist using medicare billing data using **Python** and **Neo4J**.
- Used ETL tool to extract data into **PostgreSQL Database** and created reports using **PSQL**, **Python**, **R**.
- Analyzed data from the TN Department of Education and census to determine whether demographic and economic factors could be important factors to predict high school graduation rates in **R**.
- Analyzed Trash collection data from HUB Nashville to capture different types of complains within six month and created visualizations.
- Using **Python** and **GeoPandas** scraped web data on earthquake and analyzed secondary effects of earthquake. Further analyzed the earthquake trend in TN.
- Analyzed the computer cluster usage for Vanderbilt's ACCRE and identified best and worst group in optimizing memory usage.

WORK EXPERIENCE

Sociology Instructor

Motlow State Community College

2018-present

Columbia State Community College

2017-present

Miller Motte Technical College

2016-2017

- Instructed students various sociology courses which involved reading and explaining peer reviewed research papers
- Instructed course on research method
- Advised students on research and group projects

Language Consultant

Tennessee Language Center

2014- present

- Nepali and Hindi medical, legal, and general Interpreter
- Translate written documents from Nepali to English and vice versa
- Nepali Voice over

Research Associate

Center for Medicine, Health and Society, Vanderbilt University

2015-2016

- Conducted survey's using tablets and interviews
- Assisted in analyzing data and assisted in writing a research paper titled "Interdisciplinary Well being Construct Developed Among Resettled Refugees."