

SAMEER RAU

sameer.rau2@gmail.com | github: /sameerrau | 301.956.5964

Multidisciplinary technologist with 4 years of data science + data engineering experience, coding for the public good.

EDUCATION

George Washington University, Washington, DC
B.S. – Computer Science (*Minors: Statistics, Music*) (2013 – 2017)

WORK & RESEARCH EXPERIENCE

Technologist (2018 – present)

Customer Value Partners (CVP), Fairfax, VA
Health Care Quality Analytics and Reporting – Centers for Medicare & Medicaid Services (CMS)

- Architected several end-to-end data pipelines built on AWS EMR and managed by Apache Airflow to ingest large claims data sources (terabytes in size), optimize customer performance, and improve healthcare quality outcomes.
- Spearheaded creation of company-wide internal data library using the CKAN API.
- Applied robotic process automation (RPA) to automate the generation of health quality metrics from a large claims dataset.
- Led machine learning efforts to predict hospital readmissions among discharged patients.

Pro Bono Technical Advisor (2021 – present)
AKALAKA, Durham, NC

- Automated and designed Medicaid waiver registration process for disability caregivers accessing a custom virtual training and support program offered by AKALAKA, an early-stage startup.

Research Assistant, Department of Computer Science (2017)
George Washington University, Washington, DC

- Utilized **Vicon motion capture technology** to identify biomechanical similarities between duetting dance performers; conducted correlational analysis in R.

Analytics Intern Associate (2016)
ZoomRx, Cambridge, MA

- Conducted statistical analyses of survey data for a healthcare research and analytics firm.
- Created in-depth statistical analysis tutorials for R packages (caret, ggplot, dplyr, tidyverse).

Demographics Intern (2014 – 2015)
World Bank Group, Washington, DC

- Leveraged Python and SQL to create a database profiling centenarians and supercentenarians and presented findings as part of The World Bank's **World Population Project**.
-

HONORS & AWARDS

Best Insight, American Statistical Association DataFest (2016)

Received "Best Insight" award in annual competition co-sponsored by Google and Pew. Utilized text mining techniques in R to optimize click-to-revenue conversion. Developed optimal advertising recommendations for machine-generated ads based on >1,000,000 rows of data.

CVP High Performer Award (*Three-time winner*) (2019, 2020, 2021)

Awarded to an individual deemed to have gone above and beyond the call of duty to support government agencies with innovative technical solutions.

SKILLS

Completed Coursework: Big Data and Analytics, Regression Analysis, Database Systems, Algorithms & Data Structures, Computer Graphics, Software Engineering, Artificial Intelligence

Languages and Algorithms: Amazon Web Services (S3, EMR, EC2) • Python (Pandas, NumPy, Matplotlib) Spark • R • Cloudera Hadoop • Java • SQL • SAS • C/++ • MAYA • Machine learning algorithms (random forest, k-nearest neighbors, regression, LSTM modeling)

Certifications: SAS Certified Base Programmer for SAS 9