David DeStephano

davedestep@gmail.com   (904)-303-2059   https://github.com/davedestep

Education:

Columbia University, Mailman School of Public Health, New York, NY   May 2020
  Master of Public Health in Epidemiology, Certificate in Applied Biostatistics
Florida State University, Tallahassee, FL   May 2017
  Bachelor of Science International Affairs, Concentration in Economics GPA 3.86/4.00.
  Minor: Statistics, Certificate: SAS Programming & Data Analysis
  Graduated Phi Beta Kappa/Florida State Honors Program

Skills:

R, Python, Spark, SQL (Server, MySQL, Postgre), SAS, Stata, QGIS, Tableau, Qlikview, Mplus, Microsoft Office (Word, Excel, Access, PowerPoint), machine learning (NumPy, pandas, scikit-learn, caret, Keras, XGBoost), Linux, Spanish

Experience:

Columbia Herbert Irving Cancer Center, *Biostatistician*, New York, New York   January 2022 – Present
- Write code in R and Python for data analytics and processing of EHR and real-world data including SEER-Medicare, Flatiron, Epic, MarketScan, DRG, etc. Spearheading AWS EMR infrastructure using pyspark and sparklyr for big data.

Milliman, *Healthcare Data Analyst*, New York, New York   August 2020 – January 2022
- Apply innovative techniques to large claims datasets (MarketScan, Medicare (5%, 100% IR), Medicaid, CMS MSSP ACO claims files, etc.) and provide outcomes, utilization, and cost insights for providers and pharmaceutical clients.
- Perform and advise statistical analyses, including matching, linear, logistic, GLM, ML, and survival regression in R, Python, SAS. Write/consult on abstracts, white papers, statistical analysis plans, and manuscripts for HEOR projects.
- Develop Qlikview dashboards used by ACOs and write data pipeline/ETL code using pyspark, git, Luigi, and Jenkins.

Columbia Department of Anesthesiology, *Data Analyst*, New York, New York   May 2019 – July 2020
- Conduct analysis of large data sets for studies evaluating complications and outcomes in children going for surgery
- Utilized propensity scores and other comorbidity adjustment methods, multiple imputation, linear, logistic, longitudinal, and survival data analysis, and generated results and figures by writing R, SAS, and SQL code to analyze and manipulate EHR, CMS (Medicaid), and research datasets. Collaborated with faculty in Biostatistics department on in progress and published manuscripts: "Prenatal Exposure to General Anesthesia and Childhood Behavioral Deficit".
- Work extensively with ICD 9 and ICD 10 codes using ICD specific R packages and HCUP resources for SAS.

NYC Department of Health and Mental Hygiene, *Survey Reporting Analyst*, New York, New York June 2019 – August 2019
- As part of HRTP, write SAS Macro and SQL code to automate reporting tables in the Bureau of HIV/AIDS Research and Evaluation unit, reducing the turnaround time of key analysis and dissemination projects by up to 90 percent.
- Generated graphics using data visualization software (Tableau, R) which were distributed internally to senior leaders.

Columbia Department of Neurology, *Research Assistant*, New York, New York   October 2018 – May 2019
- Created regression models in SAS and R analyzing cost trends for heart disease patients. Developed principal component analysis and hierarchical models to compare CVD categories in the medical literature with classifications.
- Use SAS to clean, manipulate, and combine data from 15 large administrative data sets including SPARCS and HCUP, resulting in 13 years of data with over 40 million observations used for analysis.

Fundación En Vía, *Monitoring and Evaluation Volunteer*, Oaxaca de Juárez, Mexico   May 2018 – July 2018
- Carried out a pilot evaluation of the En Vía microfinance program, which serves 250 women, from start to finish.
- Conducted a literature review to design evaluation methodology using resources published by international researchers.
- Developed logic model and the performance indicators/metrics to be used to guide the study.
- Led a team of five to design and administer a survey to measure business outcomes, SES, and other information.
- Created a database to house the collected information, developed data dictionary, wrote a research guide for future evaluations, created indexes used to quantify performance indicators, and analyzed the gathered quantitative data.

SS&C Advent, *Associate Data Analyst*, Jacksonville, Florida   June 2017 – May 2018
- Reconcile investment portfolios, prepared SSRS financial and investment performance reports for clients including portfolio managers and investment companies. Communicated with clients remotely over the phone and in Salesforce.
- Extracted data using SQL Server for billing inquiries, ad hoc troubleshooting, securities research and data exploration.
- Conducted data manipulation and spreadsheet creation using Excel including VLOOKUP and pivot tables, using transaction data from multimillion-dollar investment portfolios, to address client inquiries & data related issues.

Florida Department of Health, *Intern,* Tallahassee, Florida   August 2016 – December 2016
- Conducted data analysis using data from HCUP and patient surveys, billing in HMS (including private insurance and Medicaid/Medicare), used Electronic Health Records (EHR) to input and pull patient data, created Excel reports.