John M. Snyder

Contact

Salt Lake City, UT
(316) 617-1838
snyderjo@gmail.com
Linkedin
Github
Website

Education

Master of Science
Statistics
Ohio State University
Columbus, Ohio

Bachelor of Arts
Mathematics
Grinnell College
Grinnell, Iowa

Skills

R (dplyr, ggplot2, Shiny, RMarkdown)

Python (jupyter, numpy, pandas, sklearn, TensorFlow, keras, gensim, Flask)

SQL

Git

Bash

Docker

Windows, Linux, OSX

Machine Learning

Deep Learning

Regression

Statistics

Visualization

Airflow

Employment

Biostatistician II
Stanford, Radiation Oncology / Remote / July 2022 - Feb 2023

- Prepared multiple submissions to the National Institutes of Health; creating pipelines from API database pull to finalized reports; including figures, plots, and statistics.
- Designed experiments (analysis plan, sample size estimation, randomization scheme, etc.) and implemented planned analyses for oncology clinical trials in collaboration with faculty, clinical research coordinators, and residents.
- Authored statistical methods sections of publications and protocols.
- Advised clinical research coordinators and faculty on the use R and ggplot for producing publication output.

Statistical Data Analyst, Sr.
Intermountain Healthcare, Oncology / SLC, UT / Nov '15 - June '22

- Discovered gaps in patient outcomes and $10,000+ in potential cost savings per annum by leveraging system-wide structured data to identify deviation from best practice.
- Worked with superiors to identify areas of interest; iteratively investigated the relevant data (EHR/EMR, claims, registries, SQL databases, text, images, etc.) for insights, prepared graphs, tables, reports, and presentations; and delivered them to stakeholders.
- Created a postgreSQL production environment on a redhat linux server for computationally-intensive, long-running tasks, and created ETL pipelines into it.
- Prioritized multiple objectives driven by various work-groups, physician-investigators, and executives.
- Trained a convolutional neural network utilizing cloud resources from more than 30,000+ skin lesion images to proactively identify skin cancer in underserved, rural areas.
- Evaluated the effectiveness of biometric wearable devices and the data they produced in a high-risk cancer patient population as part of a partnership with a wearables start up.
- Developed key performance indicators in collaboration with physician experts for practitioner rewards and board goals, and published them in interactive R Shiny dashboards.
- Applied regression, random forest, and loess models to investigate the association between a more affordable, institutionally developed biomarker and the industry standard for predicting recurrence in breast cancer patients.
- Examined the impact of surgeon specialty on patient survival in gastro-intestinal tumors in collaboration with research residents and physicians, and quantified the results.
- Created a report of market share and financial health from disparate data sources.
- Completed the Institute for Healthcare Delivery Research's Advanced Training Program in Healthcare Delivery.
- Led and organized presentations of statistics and machine learning as chair of the Intermountain Statistics committee, and presented topics of deep learning and survival analysis to 50+ peers.

- Developed word embeddings from millions of medical documents to identify cancer-specific nomenclature.
- Automated several recurring metrics from query to emailing finalized report combining SQL, python, R Markdown, and bash; maintaining version control using git and cloud repositories.
- De-identified data subject to federal regulation and shared it with outside collaborators.
- Co-authored more than 20 publications, aiding physician-investigators in data collection, cleaning, and analysis.
- Identified limitations in data systems and reports, and worked across teams of analysts and data architects to find solutions.
- Developed several python objects to represent pathology reports and metadata.

Quantitative Data Analyst
UW - Madison / Madison, WI / Aug '10 – Oct '15
- Produced quantitative evaluations of educational effectiveness as a consultant for more than 15,000 employees and hundreds of institutions in a labor-management environment.
- Automated the organization of multiple data sources into 600+ distinct datasets for production.
- Communicated with clients and collaborators to distinguish between technical and policy decisions, producing reports elaborating upon the consequences of each.
- Identified issue penalizing high-performing students in the conventional model, collaborated with senior investigators to develop a novel process to address it, and implemented the solution in a SAS macro.
- Clarified data limitations and shared them with clients, decreasing production time two months across four FTE's.
- Re-factored the application of business rules and policy decisions–reducing 10,000 lines of code to 2,000–making them clearer, capable of dynamic updates, and easily communicated to clients.
- Conveyed statistical results to collaborators and clients of diverse analytical background.