Nicholas **Anthony**

DATA SCIENTIST · SOFTWARE ENGINEER

570-449-6615 | nanthony007@gmail.com | nanthony007.github.io | nanthony007 | @nanthony95

Skills

Languages	Python, Go, Rust, React/Vue, SQL, R
Skill Sets	Data Pipelines, Machine Learning, NLP, Web Scraping, Web Backend, CLI, REST, Automation, Library Development
Tools	PostgreSQL, MySQL/SQL-Server, MongoDB, CI/CD (Actions), Docker, REDCap, GitHub
Deployment	AWS, Heroku, Docker, Digital Ocean, Netlify/Vercel
Frameworks	Pandas, Scikit-Learn, TensorFlow, Keras, FastAPI, Django, Echo, Fiber, Cobra, Click, Clap, Next/Nuxt, Streamlit, TailwindCSS, Bootstrap

Experience

University of Kentucky
DATA SCIENTIST — *Lexington, KY* · *Dec. 2021 - Present*

- Built CLI application to extract drug mentions from text records using Rust and string similarity metrics
- Developed Tableau Web Data Connectors to directly connect web-based APIs to Tableau for team utilization
- Built automated, end-to-end, open data pipeline for enhancing and publishing open medical examiner data
- Built procalcitonin web-application for healthcare providers regarding PCT levels
- Created FL-Health-Charts website scraper to automate retrieval of online data
- Performed network analysis of Google Scholar publications. Transformed analysis into interactive web application
- Utilized NLP machine learning techniques medical concepts from 50GB of patient care narratives and link concepts to the SNOMEDCT (and UMLS) clinical health databases
- Built predictive machine learning model to estimate geographic location based on demographic data
- Migrated team to GitHub for open-source work

Global Fortification Data Exchange (GFDx)
SOFTWARE ENGINEER — *Remote* · *July. 2022 - Present*

- Migrated architecture from existing vendored services to in-house solutions
  item Implemented data modeling, validation, and standardization across organization
- Hired web developer and designer and supervised development of frontend dashboards
- Created automated DB synchronization pipeline for Data Steward database and Production
- Created data pipeline for country survey submissions

MatBoss
SOFTWARE ENGINEER, ANALYTICS PROJECT LEAD — *Minneapolis, MN* · *Jan. 2020 - Dec. 2022*

- Deployed full-stack Analytics platform using Docker with FastAPI backend and Vue frontend
- Developed Universal Data model to standardize/validate data and reduce errors throughout organization
- Built user-based feedback system to assist with data error reduction/elimination
- Released open-source wrestling package on PyPI for standard wrestling data models/containers
- Performed statistical analysis on match data to identify key factors for success. Integrated analytical results into customer dashboard
- Created various charts and infographics for marketing on social media

Raise Leaders
SOFTWARE ENGINEER — *West Palm Beach, FL* · *March. 2020 - Dec. 2020*

- Built full-stack web application using django backend, bootstrap frontend, and Postgres database
- Developed attendance tracking system for virtual zoom classes
- Built avatar customization system using rank-based badges to encourage youth participation

Veritas Wrestling Systems

<div align="right">Clearwater, FL</div>

FOUNDER

<div align="right">Sept. 2017 - Dec. 2019</div>

- Deployed first publicly accessible database for wrestling using django, heroku, and postgresql
- Engineered CLI and Desktop data collection software using tkinter to collect information in real-time
- Built predictive models to determine likely match outcomes
- Created user-dashboard for fantasy wrestling to increase marketability and fan base
- Acquired by MatBoss LLC

Research

geoPIPE: Geospatial Pipeline for Enhancing Open Data for Substance Use Disorders Research

<div align="right">University of Kentucky</div>

ACCEPTED FOR PRESENTATION AT AMIA (AMERICAN MEDICAL INFORMATICS ASSOCIATION)

<div align="right">2022</div>

- A paper on enhancing open Medical Examiner datasets using geospatial analysis and an end-to-end data pipeline
- Provides a framework for conducting geospatial analysis on Social Determinants of Health (SDOH) within a targeted spatial region
- Built an end-to-end data pipeline that consumes data from multiple sources, performs geocoding and spatial joining, extracts drugs mentions, automates descriptive analysis and reporting, and publishes the enhanced data for consumption by internal and external researchers

Opioid overdoses increase at home during the COVID-19 stay-at-home order period in Cook County, Illinois

<div align="right">University of Kentucky</div>

PUBLISHED

<div align="right">2022</div>

- A paper on observational analysis of stimulant and opioid-involved deaths during COVID-19 lockdown in Cook County.
- Examines time (pre, during, post-lockdown) and spatial elements (residential housing, hotel/motel, etc) of overdose deaths throughout 2020
- Odds of opioid-involved deaths at home were 37% higher during stay-at-home orders
- Repeat locations of fatal overdoses could be suitable for naloxone kit placement
- Available at: https://doi.org/10.1016/j.focus.2022.100007

Social Determinants of Health from Electronic Medical Records of Patients with Substance Use Disorders during the "Triple Wave"

<div align="right">University of Kentucky</div>

ACCEPTED WITH REVISIONS

<div align="right">2022</div>

- A descriptive paper on Social Determinants of Health (SDOH) in Electronic Medical Records (EMRs) for patients with Substance Use Disorder (SUD)
- This study uses electronic medical records (EMRs) from a safety net hospital system from 14,032 patients in Kentucky from 2017 to 2019 in order to 1) define three types of SUD cohorts with shared/unique risk factors, 2) identify patients with unstable housing using novel methods for EMRs and 3) link patients to their residential neighborhood to obtain quantitative perspective on social vulnerability
- Found statistically significant risk for patients experiencing unstable housing in all three cohorts

Education

Learning and Cognition

<div align="right">Florida State University</div>

MASTER OF SCIENCE

<div align="right">May. 2018 - Dec. 2019</div>

Data Science

<div align="right">The Data Incubator</div>

FELLOW

<div align="right">May. 2020 - August. 2020</div>

Computer Science

<div align="right">Rasmussen Uniersity</div>

BACHELOR OF SCIENCE

<div align="right">Feb. 2019 - May. 2020</div>

Psychology

<div align="right">University of Florida</div>

BACHELOR OF SCIENCE

<div align="right">July. 2013 - May. 2017</div>