

# STEPHEN KANE

Data Scientist

My love of detail and observation are integral to both Data Science and Cyber Security. My commitment and patience allow me to embrace a project and see it through. I problem solve, think clearly, and take action in high-stress environments.

**Email:** [stephen.x.kane@gmail.com](mailto:stephen.x.kane@gmail.com)

**GitHub:** [github.com/QuantumKane](https://github.com/QuantumKane)

**LinkedIn:** [/in/stephenjkane](https://in/stephenjkane)

**Cell:** 210-249-1139

## SKILLS

### Technical

Beginner: Tableau

Intermediate: Python, Git, SQL, Jupyter, Applied Statistics, Pandas, Matplotlib, Numpy, Seaborn, Sklearn

### Soft

Written and verbal communication skills, teamwork, empathy, organization, analytical skills, relationship building, public speaking, and learning agility

## EXPERIENCE

### Seafood Specialist

*HEB | San Antonio | 2015-Present*

Oversee and coordinate daily Meal Simple production resulting in a 16% sales increase over the last six-month period. Supervise and train new hires with a focus on customer service.

### Book Seller

*Half Price Books | San Antonio | 2014-15*

Designed and implemented a plan to boost sales and marketing for the Business section. Increased section sales by 14% over a six-month period. Supervised and trained new hires.

### Volunteer

*Youth Code Jam | San Antonio | 2/2020*

Worked with middle and high school students at Rackspace's Cyber Security station.

## PROJECTS

### Off the Rails: Capstone Project - June 2021

Our team is accessing the Department of Transportation's database to analyze 8 years of rail accidents across the United States. We are utilizing the full data science pipeline to analyze the data and create a classification model that will predict which company is likely to be involved in a rail accident.

### NLP Web Scraping - May 2021

Using GitHub's API and python's BeautifulSoup, I scraped over 400 GitHub repositories and pulled the text off of their README files. I analyzed the text using Natural Language Processing techniques and created multiple models to predict the programming language of the repository. I experimented with oversampling using SMOTE due to an extremely imbalanced dataset. My best performing model correctly predicted the programming language at 60% accuracy, doubling the baseline.

### Network Intrusion Detection - April 2021

Using Network Intrusion data from Kaggle, I decided to find insight on anomalous activity. I went about this by predicting normal transmissions, thereby isolating suspicious activity. I created several Classification models using Scikit-learn for prediction, all with accuracies beyond that of my baseline prediction.

## EDUCATION

### Codeup

Certificate of Completion

*San Antonio, TX | December 2020 - June 2021 (expected)*

Fully-immersive, project-based 22-week Data Science career accelerator that provides students with 670+ hours of expert instruction in applied data science.

### Bachelor's Degree

Business Administration in Cyber Security

*The University of Texas at San Antonio | May 2020*

Cumulative GPA: 3.0 Major GPA: 3.20

### Relevant Coursework

#### Classes:

Network Security, Intrusion Detection, Information Assurance & Security and Operating Systems Security

#### Topics:

Basic Cryptography, Secure Networking Protocols, Traffic Analysis, Fraud Detection, Intrusion Prevention, Hacking and Viruses