

LONG KHANH NGUYEN

ngkhlong189@gmail.com | 309-631-4516

[Linkedin](#) | [Tableau Viz](#) | [Portfolio](#)

EDUCATION

UNIVERSITY OF ROCHESTER

STEM Master of Science in Data Science

- Relevant courses: Data Mining, Comp. Statistics, Statistical Machine Learning, Database Management, Deep Learning.

Rochester, NY

Aug 2022 - Dec 2023

AUGUSTANA COLLEGE

Bachelor of Arts in Applied Mathematics

Rock Island, IL

Aug 2018 - May 2022

PROFESSIONAL EXPERIENCE

REL8ED.TO ANALYTICS

Data Science Intern

Rochester, NY

Jul 2023 - Aug 2023

- Collaborated with Coface's Data Scientists to accurately classify companies' industries using **OpenAI** embedding models for NAICS codes classification.
- Improved data collection efficiency by 80% with **Python** scrapers for concurrent validation of 1 million URLs.
- Conducted dynamic data analysis in **Excel** for visually impactful insights using Pivot Table and XLOOKUP.

UNIVERSITY OF ROCHESTER

Teaching Assistant

Rochester, NY

Aug 2023 - Dec 2023

- Mentored students in diverse statistical areas: descriptive statistics, probability and distribution, confidence intervals, hypothesis testing, inference, ANOVA, and regression analysis.
- Led multiple review sessions for exams and projects, resulting in an 80% increase in students' overall grades.
- Empowered students in understanding **R** packages and creating impactful reports with **RMarkdown**.

LEAFPRINT ENTERPRISES

Data Analyst Intern

Cedar Hills, UT

Jan 2022 - May 22

- Led global municipal research, collected 100k+ data points, and delivered actionable recommendations to enhance World Archives's municipal dictionary.
- Utilized **Python**'s pandas for data cleaning and transformation, reducing errors and retrieval time by 80%.
- Developed and maintained interactive dashboards and reports using **Tableau** to support datadriven decision-making.

BENIT PTY LTD

Data Analyst Intern

Hanoi, Vietnam

Mar 2021 - Jul 2021

- Developed an optimized **SQL** data model for TwelveAPI's Forex data, improving access time by 80%.
- Utilized **AWS S3** for efficient database management, resulting in a 70% reduction in overall latency.
- Integrated **Python** with **SQL** databases to streamline **ETL** processes, ensuring accurate financial insights via dynamic dashboards and reports.

DATA ANALYTICS PROJECTS

- Behavioral Segmentation for a hotel and resort company (ongoing)** Building a K-means clustering model on a hotel and resort dataset of more than 50k customers to segment them into different behavioral groups, and provided key strategies to enhance customer experience and maximize revenue for each segment.
- Mental health mining and prediction** Analyzed 1,259 employee observations, using association rules and SVM with Grid-CV to predict mental health treatment necessity with 74.6% accuracy in **Python** (sklearn, pandas, numpy, matplotlib, seaborn).
- Twitter Sentimental Analysis on Asian Hate During COVID-19** Conducted Twitter Sentiment Analysis on Asian Hate during COVID-19 using over 16,000 tweets, exploring sentiments with **Python** (pandas, plotly, re, nltk, textblob), and visualized most frequent hashtags using WordClouds.

TECHNICAL SKILLS

- Programming:** Python, SQL, R
- Data Visualization:** Tableau, PowerBI, Excel
- Database Management:** MySQL, PostgreSQL
- Familiar with:** Database design and development, Statistical modeling & Machine learning techniques, ChatGPT