KRISTO PAPADHIMITRI

New York, NY papadhimitrik@gmail.com | GitHub: github.com/AirKristo | LinkedIn: linkedin.com/in/kristo-papadhimitri/

SUMMARY

Senior Data Analyst with 8+ years in utility analytics and 3+ years building production SQL/SAS/Python pipelines on 500M+ record datasets. MS in Computer Science (NYU Courant) with hands-on experience in causal inference, time series analysis, and NLP pipelines. Strong track record translating messy operational and market data into scalable data products, statistical analyses, and decision-ready insights for senior stakeholders.

TECHNICAL SKILLS

Data Science & Machine Learning: Python, Pandas, NumPy, Scikit-learn, SciPy, TensorFlow, Keras
Data Engineering & Databases: SQL, SAS, MapReduce, HiveQL
Visualization & Tools: Tableau, Matplotlib, Excel, Git/GitHub
Programming Languages: C++, Java

PROFESSIONAL EXPERIENCE

Senior Data Analyst | Data Analyst - Consolidated Edison, New York, NY Nov 2022 – Present

- Design and maintain SQL/SAS ETL pipelines that ingest and cleanse 500 million+ smart-meter records, ensuring data quality for demand forecasting and rate design models used by senior leadership
- Refactor legacy Python/SAS workflows to improve code maintainability and automation, reducing manual processing time by 25% (~8 analyst-hours per week)
- Conduct exploratory data analysis on customer usage patterns across residential and commercial segments, delivering insights that shaped strategic recommendations adopted by senior leadership

Program Manager - Consolidated Edison, New York, NY Mar 2020 – Nov 2022

- Built Tableau dashboards to visualize performance of 5,000+ company assets, identifying high-priority sites and enabling data-driven capital allocation decisions

Business Associate - Consolidated Edison, New York, NY Sep 2017 – Mar 2020

- Collaborated with Technology team to develop a regression model to forecast staffing needs for future storms based on historical data

TECHNICAL PROJECTS

Federal Reserve Communication Causal Inference Analysis Jan 2026

- Applied causal inference methods, specifically difference-in-differences with clustered standard errors, across 96 FOMC announcements (2014-2025) to identify a statistically significant 12 basis point volatility premium in rate-sensitive sectors vs. defensive sectors (p-value of 0.043)
- Validated causal identification by confirming parallel pre-trends across treatment and control groups in the 2-day pre-announcement window
- Applied synthetic control to the March 2020 emergency cuts, estimating financials underperformed their counterfactual by 13.6%; built automated time series panel in Python (pandas, statsmodels, scipy)

Sentiment-Driven Market Forecasting Pipeline - New York University, New York, NY Feb 2025 – May 2025

- Engineered end-to-end NLP pipeline to ingest and clean over 2 GB of Reddit posts from four finance subreddits and all 119th Congress bills, applying FinBERT to generate daily sentiment scores
- Integrated Yahoo Finance API to fetch and aggregate daily returns for all companies in the S&P 500, grouped by sector
- Identified a 2-day lead indicator: legislative sentiment hit $p < 0.05$ in every Granger-causality test (Lag 2), whereas Reddit and news never did, yielding the most consistent sector forecasts

EDUCATION

Master of Science, Computer Science, NYU COURANT INSTITUTE OF MATHEMATICAL SCIENCES May 2025

- Relevant Coursework: Machine Learning, Big Data Science, Artificial Intelligence, Introduction to Data Science

Bachelor of Science, Mathematics, NYU TANDON SCHOOL OF ENGINEERING May 2017

Minors: Economics, Finance