

ANDREAS KEFALAS

Queens, NY | 929-569-9420 | [e-mail](#) | [LinkedIn](#) | [GitHub](#) |

Authorized to work for any US employer.

DATA ENGINEER

Experienced Data Engineer with 2+ years of professional experience in data engineering, analysis, and visualization. Skilled in tackling complex challenges, promoting team synergy, and utilizing statistical methodologies. Proficient in generating impactful visualizations to reveal insights and translating technical concepts into actionable SaaS specifications. Versatile background in academia and entrepreneurship. Fluent in English, Spanish, and Greek.

SKILLS

- **Programming & Tools:** Python (pandas, numpy, SciPy), R (ggplot2, Matplotlib, Seaborn), Bash scripting, SaaS Products, Redis
- **Data Manipulation, ETL and Visualization:** Tableau, PowerBI, Apache Superset and Spark, SPSS, NetworkX, Grafana, Fivetran
- **Cloud & Databases:** AWS (EC2, S3, Lambda, Lake Formation), Snowflake, SQL, PostgreSQL, AWS Redshift, Jenkins, Docker, Cron, Kubernetes
- **Agile Environment:** Pivotal Tracker, Jira, Slack, ClickUp

PROFESSIONAL EXPERIENCE

Tru Solutions, Remote

May 2023 - Present

Data Engineer

- Analyzed energy project data, optimized operations, and developed real-time visualization dashboards using other industry-standard tools to enhance IT sector insights. Integrated internal metrics automation, transitioning manual tasks to cloud-based solutions and efficiency-oriented automations
- Automated API requests using Python scripts and Jenkins, leveraging AWS Lake Formation to extract, store, and analyze large volumes of API data, while proficiently managing both SQL and NoSQL databases for comprehensive data management solutions.
- Implemented data engineering solutions for energy projects, orchestrating the integration of multiple APIs using Fivetran to ingest data into Snowflake. Achieved a significant 30% improvement in data reporting accuracy for top management.
- Conducted data analysis for the oil rig and natural gas midstream pipeline industry, leveraging PostgreSQL expertise for efficiency.
- Utilized AWS Lake Formation for scalable data extraction, storage, and analysis, and configured AWS EC2 instances to host and run a full BI tool stack. Led the creation of custom productivity dashboards, optimizing operations at an energy facility by implementing Materialized Views and caching in Snowflake to reduce latency and warehouse costs.

Tru Solutions, Remote

Aug 2022 – May 2023

Technical Project Manager

- Managed projects and devised blueprints by collaborating with cross-functional teams, including Client Services and Product Development, to define project scope, objectives, and deliverables.
- Engineered a predictive churn model resulting in a 25% decrease in customer attrition and optimized data pipelines, reducing data processing time by 30% and contributing to quicker product updates.
- Consolidated all-in-one features such as device monitoring, ticketing, reporting, Cloud backup, and remote control. Deliver remarkable results, reducing service calls by 25% and on-site visits by 35% within the initial 2 months.
- Articulated wireframes, composed automation tool requirements, and proficiently oversaw multiple projects alongside their pipelines, encompassing ticket integration, QA testing, and proactive communication with top management.

CUNY Queens College, Queens, NY

Jan 2022 - Aug 2022

Quantitative Analysis in R Lecturer

- Developed lab assignments, evaluated homework and exams, provided feedback, offered guidance on improving code-writing skills, and taught a class of 25 students, aiding them in grasping key concepts, R package use, and executing coursework effectively.

CUNY Queens College, Queens, NY

Aug 2021 - Jan 2022

Scientific Research Assistant

- Processed and visualized large data volumes, established connections via NetworkX, automated tasks for a 40% reduction in manual work, and used Python and R for clustered environment visualizations and metadata analysis in scientific journals.

TruQC, Remote

May 2017 – Aug 2020

Junior Developer

- Designed and implemented dynamic application reports leveraging JavaScript and JSON schemas for enhanced functionality and user experience.
- Led the management of Continuous Integration builds in Jenkins, ensuring seamless environmental synchronization through timely fix commits.
- Crafted custom SQL queries to generate insightful reports, driving data-driven decision-making.
- Performed in an Agile environment, utilizing cutting-edge CI/CD tools like Jenkins and Pivotal Tracker for efficient project tracking and ticket management.

EDUCATION

- **Masters of Arts (M.A.) in Data Analytics and Applied Social Research** **Aug 2021 - May 2023**
Queens College, Queens, NY
- **Bachelor of Arts (B.S.) in Information Systems** **Aug 2013 - Dec 2017**
University of Missouri - Saint Louis, Saint Louis, MO

HONORS / AWARDS

- Queens College, **Academic Scholarship recipient** **Aug 2021 - May 2023**
- University of Missouri - Saint Louis Men's Basketball, **Scholarship recipient** **Aug 2013 - Dec 2017**

PROJECTS

- **NYC OPEN DATA - 311 Pre/Post Covid-19 Service Request Analysis** **Jan 2023 - May 2023**
Analyzed a 9 million-row dataset with R, conducting regression testing, correlation analysis, and A/B testing, while using the mapPluto library for visualizing service request patterns across NYC's five boroughs.
- **Open-Source Contributions on GitHub**
Contributed to open-source projects, including Apache Superset, by documenting existing issues and proposing solutions via sophisticated pull requests to enhance development functionality.