ANDREAS KEFALAS

Queens, NY | 929-569-9420 | andrekef@gmail.com | LinkedIn | Github
Authorized to work for any US employer.

DATA ENGINEER

Experienced Data Engineer with over 2 years of professional experience adept at tackling complex challenges and promoting team synergy. Skilled in data engineering, analysis, and visualization, with a solid grounding in statistical methodologies. Proficient in generating impactful visualizations to unveil vital insights. Capable of translating technical concepts into actionable SaaS specifications. Possesses versatile experience spanning academia and entrepreneurship. Fluent in English, Spanish, and Greek languages.

SKILLS

- **Programming & Tools:** Python (pandas, numpy, SciPy), R (ggplot2, Matplotlib, Seaborn), Bash scripting, SaaS Products
- **Data Manipulation, ETL and Visualization:** Tableau, PowerBI, Apache Superset and Spark, SPSS, NetworkX, Grafana, Fivetran
- **Cloud & Databases:** AWS (EC2, S3, Lamda, Lake Formation), Snowflake, SQL, PostgreSQL, AWS Redshift, Jenkins, Docker Containers, 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 processes to cloud-based solutions and automating tasks to improve efficiency.
- Proficient in handling both SQL and NoSQL databases for comprehensive data management solutions.
- 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.

EDUCATION

Masters of Arts (M.A.) in Data Analytics and Applied Social Research - GPA: 3.5/4.0   **Aug 2021 - May 2023**
Queens College, Queens, NY

Bachelor of Arts (B.S.) in Information Systems - GPA: 3.8/4.0   **Aug 2013- Dec 2017**
University of Missouri - Saint Louis, Saint Louis, MO

HONORS / AWARDS

Queens College, **Scholarship recipient** 2021-2023
University of Missouri - Saint Louis Men's Basketball, **Scholarship recipient** 2013-2017
Dean's List for 4 semesters

PROJECTS

NYC OPEN DATA - 311 Pre/Post Covid-19 Service Requesst Analysis   **Jan 2023 - May 2023**

- Extracted and modeled data from a 9 million-row, 13-year-old dataset, conducting comprehensive analysis with R, including regression testing, correlation analysis, and A/B testing methodologies, and created visualizations using the mapPluto R library to understand service request patterns across New York City'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.