

Jimmy Nguyen

• Washington DC • jnjimmy1@gmail.com • 443-720-9048 • [linkedin.com/in/jnjimmy1](https://www.linkedin.com/in/jnjimmy1) • jnjimmy1.com

PROFESSIONAL PROFILE

Results-driven data engineer with a strong background in leading teams and implementing scalable data solutions. Proficient in ETL pipeline deployment and optimization, adept at driving efficiency in data engineering processes.

SKILLS

- Python (5 YoE) | SQL (4 YoE) | Java (8 YoE) | LUA (4 YoE) | Jupyter Notebook | PostgreSQL | MySQL | GIS | Git
- Cloud Computing | Machine Learning | NLP | ETL | CI/CD | OOP | AWS (EC2, RDS, S3) | Optimization | Agile
- Distributed Systems | Scalable Systems | MLOps | DevOps | Backend | Data Visualization | Data Modelling

PROFESSIONAL EXPERIENCE

DATA TEAM LEAD

Washington, DC

Citian Solutions

03/2022 - 06/2023

- Managed a 7 member Data Team and liaised with non-technical stakeholders to convert requests into actionable tasks via GitLab, contributing to a 20% increase in productivity and project management.
- Implemented a recommendation algorithm that accurately budgets multi-million dollar construction projects by using Excel, SQL, PL/pgSQL functions, and Python.
- Recreated the entire roadway network of the United States in extreme detail and replaced multiple QA/QC teams by leveraging GIS, web scraping, data visualization, error handling, Jupyter Notebooks, and REST APIs.
- Developed a location refinement algorithm using geocoders, predictive modeling, NLP, and distributed systems, which reduced margins of error from 1,000+ ft down to just 6 ft for millions of reported crashes.
- Optimized data transmission processes and created hundreds of visualizations and dashboards, decreasing loading times by 40% and improving user workflow speeds by 70%.

FOUNDING DATA SCIENTIST/ENGINEER

Washington, DC

Citian Solutions

08/2021 - 03/2022

- Deployed scalable ETL pipelines in Python using OOP concepts, expediting the client onboarding timeline from over 3 months down to a single week.
- Refactored and normalized PostgreSQL, MySQL, and Snowflake databases, resulting in a 60% reduction in database storage size.
- Parallelized Extract, Transform, Load (ETL) scripts using Python's multiprocessing library and automating AWS EC2 instances, accelerating intensive scripts by upwards of 996x.
- Extracted tabular and geospatial data from hundreds of APIs and data warehouses using Python and PostgreSQL, resulting in the recovery of up to 93% of previously missing data.
- Presented to stakeholders by performing Exploratory Data Analysis (EDA) on client data and generating visualizations and statistics, increasing the rate of incoming SaaS trials by 25%.

EDUCATION

UNIVERSITY OF MARYLAND

College Park, Maryland

Bachelor of Science in Computer Science; Minor in Mathematics

2017-2021

- University Honors

PROJECTS

Game Analytics and Automation

2014-Present

- Designed and implemented automation scripts using Python, utilizing data analysis and data analytics to optimize resource allocation and maximize gameplay efficiency.
- Leveraged Computer Vision techniques to navigate complex gaming environments, showcasing proficiency in image processing and pattern recognition.

Game Development

2009-Present

- Created a handful of games in LUA with tens of thousands of unique player counts.
- Employed Java and Unity Game Engine to reverse engineer popular games, showcasing proficiency in understanding and manipulating game-related data structures.

MENTORSHIP

Computer Science Tutor/Mentor

2014-2021

- Programming | Java | Python | Object Oriented Programming | Data Structures | Concurrency | Data Science

Programming Forums

2009-2021

- Actively mentored and participated on programming help forums for LUA, Java, and Python