LUONG NGUYEN YEN NHI
Data Engineer

Ho Chi Minh City • (+84) 378 282 298 • lnyn.netlify.app • ynnhi1508@gmail.com

CAREER OBJECTIVE

Aspiring Data Engineer seeking a fresher or internship role to grow professionally and contribute to the success of the team and company through dedicated efforts and impactful data-driven solutions.

- **Short-term Goal:** Secure a Data Engineer internship to gain practical experience and enhance my technical skills.
- **Long-term Goal:** Become a Senior Data Engineer within the next five years.

TECHNICAL SKILLS

Python, PostgreSQL, MySQL, Spark, Kafka, Airflow, Docker, Azure

PROJECTS

Real-Time NYC Trip Analytics | Python, Postgres, Kafka, Spark, PowerBI, Docker   *Source: Github*
August 2024

- Implemented Docker to run Kafka and Kafka UI.
- Employed Spark to transform and compute streaming data from Kafka, performing real-time analytics.
- Calculated average hourly revenue, identified trip discrepancies between boroughs, and detected trips with unusual durations (e.g., under 2 minutes) or abnormal fare amounts (e.g., total amount not matching the sum of individual fees).
- Transferred data and computations to PostgreSQL for monitoring.
- Visualized and analyzed trip trend insights more easily with a PowerBI dashboard.

Automated Job Data Pipeline | Python, Playwright, Postgres, Airflow   *Source: Github*
October 2024

- Used Requests and Playwright to scrape data from TOPCV.
- Cleaned the data and stored it in Postgres.
- Generated a "due date" field, representing the application deadline based on the remaining time to apply from the scraped data.
- Developed a procedure to update the "remaining time to apply" daily, e.g., from "24 days left to apply" today to "23 days left to apply" tomorrow.
- Utilized Airflow to orchestrate the data scraping, cleaning, and storage processes.

EDUCATION

Sai Gon University	Ho Chi Minh City
Bachelor of English Linguistic	August 2020