

Nguyen Quang Phuoc

Data Engineer Intern



phuocquang272@gmail.com



<https://github.com/NQP27>



0387592827



27/02/2003



<https://www.instagram.com/q.phuoc20>



67 Phung Khoang - Ha Dong - Ha Noi



<https://www.facebook.com/nguyenquangphuoc.27>

ABOUT ME

I am a young, passionate individual with a strong desire for learning and interested in data. I am capable of self-study and patient to solve problems. I work well both in a team and independently and want to find an intern position for data engineer job.

SKILLS

Python

SQLServer

Scraping data

Pandas

ETL

AWS

Airflow

Spark

Kafka

Data Modeling

LANGUAGES

ENGLISH

Basic English communication, good at reading and listening skill

INTERESTS

Technology Devices

EDUCATION

INFORMATION TECHNOLOGY

POSTS AND TELECOMMUNICATIONS INSTITUTE OF TECHNOLOGY

(11/2021 - 2026)

PROJECTS

SALE REPORT

- Information:** Create a Report & Analyze Sales Performance of a Business with Python, Pandas, and Matplotlib.
- Link repo:** github.com/NQP27/report-sales-pandas
- Achievement:**
 - Effective skills in data processing and cleaning, contributing to the creation of a well-structured and understandable data table.
 - Generating professional visualizations of data, facilitating easy understanding and analysis of results for the readers.

WEATHER DATA

- Information:** Extract data from <https://openweathermap.org/api>, transform and load them to a csv file using Apache Airflow, Docker.
- Link repo:** github.com/NQP27/crawl-weather-data
- Achievement:**
 - Interact with API.
 - Generating connection between Airflow and PostgreSQL, HTTP.
 - Using HttpSensor to check the readiness of the API before proceeding with other tasks.
 - Make a simple HTTP request to the API and process the response data.

DATA PIPELINE

1. **Information:** An end-to-end data engineering pipeline project utilizing Apache Airflow, Python, Apache Kafka, Apache Zookeeper, Apache Spark, and Cassandra. Everything is containerized using Docker.
2. **Link repo:** github.com/NQP27/data-engineer-project
3. **Achievement:**
 - Setting up a data pipeline with Apache Airflow.
 - Real-time data streaming with Apache Kafka.
 - Distributed synchronization with Apache Zookeeper.
 - Data processing techniques with Apache Spark.
 - Data storage solutions with Cassandra.
 - Containerizing your entire data engineering setup with Docker.