Dao Huy Hoang

(+84) 914 209 065 | 11/01/2003 | peachhuyhoang203@gmail.com

LinkedIn   Portfolio Website   Github

District 7, Ho Chi Minh City, Vietnam

DATA ANALYTICS

A passionate and dedicated young professional focused on data analytics. Through my academic journey, I have built a strong foundation in data analysis, statistical modeling, and business intelligence tools. I am eager to apply my skills to solve real-world problems, continuously learning and adapting to new challenges. My enthusiasm for learning drives me to contribute to innovative projects in dynamic environments.

- **Short-term goal:** Enhance technical skills in data analytics tools and secure an internship to apply theoretical knowledge in practical scenarios.
- **Long-term goal:** Develop expertise in data analytics and business intelligence, eventually taking on leadership roles in analytics-driven projects.

SKILLS

Technical Skills
- Programming Languages: Python, R, C/C++
- Data Analysis Tools: Power BI, Microsoft Office (Word, Excel, PowerPoint), Google Sheets (Apps Script)
- Data Science Libraries: Numpy, Pandas, Scikit-learn, Seaborn, Matplotlib
- Operating Systems: Ubuntu
- Version Control: Git, GitHub
- ETL Tools: Databricks, AWS Glue
- Cloud Platforms:
    - Microsoft Azure (Azure Data Factory, Synapse)
    - Amazon Web Services (AWS Glue, EC2, S3, IAM, Athena)
- Technologies: Kafka, Zookeeper

Soft Skills
- Analytical Skills: Analytical Thinking, Critical Thinking
- Problem-Solving: Problem Solving, Attention to Detail
- Communication: Communication Skills, Teamwork, Presentation
- Flexibility: Adaptability, Time Management, Self-learning Skill

ACADEMY PROJECTS

Azure Data Pipeline Architecture   **December 2023**

Team size: 1

Worked with AdventureWorksLT2017 dataset.

Key Contributions:
- Data Ingestion: Configured login in SSMS and securely stored credentials using Azure Key Vault. Loaded data into Azure Synapse Workspace (Bronze layer) via Azure Data Factory.
- Data Transformation: Performed ETL on Databricks, transforming data from Bronze to Silver and Silver to Gold.
- Data Loading: Built a pipeline to move data from Bronze → Silver → Gold, triggered through Azure Synapse Analytics.
- Data Reporting: Connected Power BI to Azure Synapse for real-time data visualization from the Gold layer.

Skills: SQL server, Microsoft Azure, ETL - process, Microsoft Power BI

Stock Market Real-time Data Analysis   **August 2024**

Team size: 1
End-to-end Data Engineering Project using Kafka.
Key Contributions:

- Environment & Installation: Set up Ubuntu (WSL/VM) and configured JDK for Kafka; started Zookeeper and Kafka while adjusting memory settings and IP configuration (Public IP).
- Producers/Consumers: Created sessions for Producers and Consumers and set up Kafka topics.
- AWS Setup & CLI: Established IAM Role with S3 and Glue access, generated Access Key, and configured AWS CLI with credentials.
- Data Management: Used Jupyter Notebook to install libraries and develop code for Kafka and S3; created a Glue Crawler for S3 scanning and set up a database, followed by executing SQL queries with AWS Athena.
- Web Data Integration: Integrated APIs for real-time data and modified Producer logic accordingly.

Skills: Ubuntu, Kafka, AWS, Python (Jupyter Notebook), SQL

EDUCATION

UNIVERSITY OF SCIENCE - VNUHCM   **2021 - 2025 (expected)**

Bachelor of Science - Major: Mathematics & Computer Science
Specialization: Data Science - Current grade: 3.2 / 4.0

NGUYEN THI MINH KHAI HIGH SCHOOL FOR THE GIFTED   **2018 - 2021**

Award: Top 10 student with the highest score in class

CERTIFICATIONS

IELTS - OVERALL BAND SCORE: 6.5   **August 2024**
IDP

GOOGLE DATA ANALYTICS SPECIALIZATION   **May 2024**
COURSERA

SQL - BASIC & INTERMEDIATE & ADVANCED   **2023**
HACKERRANK

LANGUAGE

VIETNAMESE (native **)**

ENGLISH (fluent **)**