



# 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 ENGINEER

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, Amazon Web Services (AWS)
- 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

## 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

### INTERNATIONAL ENGLISH LANGUAGE TESTING SYSTEM - IELTS 6.5

August 2024

IDP

### GOOGLE DATA ANALYTICS SPECIALIZATION

May 2024

COURSERA

### SQL - BASIC & INTERMEDIATE & ADVANCED

2023

HACKERRANK

# ACADEMY PROJECTS

## DATA ANALYST & SCIENCE

### Coffee Sales Data Analysis

August 2022

Analyzed coffee sales data, including orders, customers, and products.

Key Contributions:

- Used XLOOKUP and INDEX MATCH for data retrieval.
- Applied formulas to calculate total sales and categorize data using IF functions.
- Standardized date and number formatting for consistency.
- Checked for duplicates and converted data ranges into structured tables.
- Created Pivot Tables and Pivot Charts to summarize key metrics.
- Built an interactive dashboard with timelines and slicers for dynamic filtering.

**Skills:** Excel Functions, Pivot Chart, Pivot Table

### Sales Forecasting For a Store

Feb 2023 - May 2023

Applied machine learning algorithms to analyze whether sales are affected by relevant factors.

Key Contributions:

- Group project analyzing Walmart sales data (2010-2013) to forecast sales.
- Cleaned and visualized data to assess the impact of holidays on sales.
- Enhanced model performance by carefully selecting and eliminating specific features, improving the predictive accuracy.
- Built a Random Forest Regression model using WMAE (Weighted Mean Absolute Error) to predict sales.

**Skills:** Python, Jupyter Notebook, Random Forest, Time Series Forecasting

---

## DATA ENGINEER & CLOUD

### Azure Data Pipeline Architecture

December 2023

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

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