

# Pham Manh Hung

## Data Engineer Intern

📍 Linh Trung, Thu Duc    ✉ Hungpm267@gmail.com    ☎ 0333856974    Portfolio [🔗](#)    in Hungpm267  
🌐 Hungpm267

## Career Objective

---

As a senior student in Information Technology, I am seeking an internship opportunity to not only apply the knowledge I have gained but also learn more from a real working environment. My goal is to contribute to the success of the company while enhancing my skills in software development, problem-solving, and teamwork, ultimately preparing myself for a professional career in the tech industry.

## Education

---

**University of Information Technology - VNUHCM**, Information System

- July 2021 - July 2025
- Nearest semester GPA: 7.81

## Skills

---

**Programming Languages:** SQL, Python, MDX, C++, Java, C#.

**Programming Frameworks:** Pyspark, Numpy, Pandas, Scikit-learn, Bootstrap

**DBMS:** SQL Server, MySQL, Oracle

**Version Control Tools:** Git, Git Hub, Trello

**Technologies:** Spark, Hadoop, Microsoft SQL Server, Jupyter Notebook, VSCode, Microsoft Azure, PowerBI

**Soft Skills:** English(B2), Critical Thinking, Communication, Problem Solving, Self-Learning, Teamwork

## Projects

---

### Data Warehouse and OLAP

[Link Github](#) [🔗](#)

- Build a Data Warehouse by SSIS tool to store data of earthquake, perform analyze data with SSAS and make report With SSRS, Pivot Excel, PowerBI as well as do datamining and query data with mdx
- Team Size: 2
- Role: Responsible for SSIS, SSAS and Data mining Sections
- Tools Used: Microsoft SQL Server, PowerBI, Python, Excel, MDX

### Predict Wine Quality by Spark

[Link Github](#) [🔗](#)

- Use pyspark in Apache Spark to Exploratory Data Analysis and Preprocessing. Build model with KNN algorithm (no MLlib) to predict the quality of wine
- Team Size: 1
- Tools Used: Jupyter Notebook, Pyspark, Jupyter Notebook

### Perform ETL Process with Azure DataFactory and DataBrick

[Link Github](#) [🔗](#)

- Perform Azure Data Engineer Project. Did data collection using Azure DataFactory with data pipeline. Perform data transformation across bronze, silver, and gold layers in Azure DataBricks, and finally dump data into Azure Synapse Analytics
- Team Size: 2
- Role: perform Extract Dataset to Azure DataFactory and Store it into data lake gen

2(bronze layer, silver layer and gold layer) with azure data bricks. extract the output to Azure Synapse Analytics

- Tools Used: Azure DataFactory, Azure Data Lake gen2, Azure Data Bricks, Azure Key Vault, Pyspark

### **Visualizing the effectiveness of PPC marketing as well as GDP, population in Vietnam**

[Link Github](#) 

- In this project, I use the dataset about the level of effectiveness in PPC marketing. I also visualize the GDP and Population each province of Vietnam
- Team Size: 1
- Tools Used: PowerBI

## **Certificates**

---


**Toeic Reading, Listening:** 550

[Image](#) 

**Toeic Speaking, Writing:** 250

[Image](#) 

**IELTS mock test scores:** Listening: 5.5, Reading: 5.5, Speaking: 5.5, Writing: 5.0

[Image](#) 

**Critical Thinking Certificate**

[Image](#) 