

# Data Engineer Trainee

## Entrance Assessment

---

### **Vietnam Office 01**

294 – 296 Truong Sa, Ward 02,  
Phu Nhuan District, HCMC

### **Vietnam Office 02**

09 Hoa Cau, Ward 07,  
Phu Nhuan District, HCMC

### **Singapore Office**

7 Straits View, Marina One East Tower,  
#05-01, Singapore 018936

### **Contact**

Web: <https://kyanon.digital>

Email: [contact@kyanon.digital](mailto:contact@kyanon.digital)

Phone: (+84) 283 5171 080

## Phần 1: Bài tập

### Câu 1: Tìm hiểu Data Pipeline

- Giải thích khái niệm **Data Pipeline** và các thành phần chính (Ingest, Transform, Storage, Orchestration, Monitoring).
- So sánh ưu nhược điểm khi triển khai pipeline bằng **Airflow** vs **n8n**
- Trình bày ví dụ một pipeline điển hình trong công ty thương mại điện tử (ví dụ: dữ liệu đơn hàng → báo cáo doanh thu).

### Câu 2: Bài tập kỹ thuật (Data Aggregation)

- Dữ liệu đầu vào: **orders.csv** với cấu trúc:

```
order_id,order_date,customer_id,amount,status
1,2025-09-01,1001,250,completed
2,2025-09-01,1002,300,pending
3,2025-09-02,1001,100,completed
.....
```

- Sử dụng **Python** hoặc **SQL** thực hiện:
  - Đọc hoặc import dữ liệu từ **orders.csv**.
  - Lọc các đơn hàng có **status = completed**.
  - Tính tổng amount theo ngày và xuất ra file **report.csv**

```
[
  {"date": "2025-09-01", "amount": "250"},
  {"date": "2025-09-02": "amount": "100"}
]
```

### Câu 3: Viết 1 doc/pdf gồm các phần sau (mở quyền truy cập)

- Giới thiệu bản thân: Nêu ngắn gọn về quá trình học tập (Ngành, Trường, GPA), trình độ Tiếng Anh và một số thông tin khác về bản thân.
- Bạn hiểu thế nào là **Data Engineer** và tại sao công việc này cần thiết.
- Trình bày suy nghĩ về khả năng yêu thích công việc Data Engineer, mong muốn trở thành **Data Engineer**.

- Nêu cảm nhận về những gì học được khi làm **Câu 1**.
- Đính kèm mã nguồn bài Câu 2 (link Git hoặc file nén).
- Chia sẻ lộ trình phát triển của bạn theo từng giai đoạn: 3 tháng, 6 tháng, 1 năm, 3 năm.

## **Phần 2: Yêu cầu**

- **Ngôn ngữ trình bày:** Tiếng Việt
- **Thời gian làm bài:** 36h từ ngày nhận đề bài
- **Cách thức nộp bài:** Nộp trực tiếp [TAI ĐÂY](#)