**1. Nhập dữ liệu (Input)**

**Mục đích:** Đưa dữ liệu thô vào hệ thống để hệ thống có thể tiến hành xử lý.

**Nội dung chi tiết:**

* Nguồn dữ liệu có thể từ con người (nhập tay), cảm biến, file (CSV, Excel), hệ thống khác (API), hoặc thiết bị đo.
* Dữ liệu bao gồm các trường (fields) cụ thể — ví dụ: mã sinh viên, họ tên, mã môn, điểm, ngày tháng.
* Các bước cần thực hiện khi nhập:
  + **Xác thực (validation):** kiểm tra dạng dữ liệu, kiểm tra ràng buộc (ví dụ điểm phải trong khoảng 0–10), kiểm tra bắt buộc (trường không được để trống).
  + **Tiền xử lý ban đầu:** loại bỏ khoảng trắng thừa, chuẩn hóa định dạng ngày/tháng, chuyển mã ký tự nếu cần.
  + **Ghi nhật ký (logging):** lưu lại ai đã nhập, thời gian, nguồn dữ liệu.
* **Thiết bị/ giao diện:** biểu mẫu web/mobile, form Excel, cổng API.

**Kết quả:** Dữ liệu thô đã được đưa vào hệ thống và sẵn sàng cho bước xử lý.

**2. Xử lý dữ liệu (Processing)**

**Mục đích:** Biến đổi dữ liệu thô thành thông tin có ý nghĩa dựa trên thuật toán hoặc quy tắc nghiệp vụ.

**Nội dung chi tiết:**

* Các phép toán cơ bản: tính toán (cộng, trung bình, tổng), lọc, sắp xếp, nhóm, tính chỉ số (ví dụ: trung bình, phần trăm).
* Các bước xử lý thường gặp:
  + **Tính toán:** ví dụ tính điểm trung bình, tổng tiền, tỷ lệ phần trăm.
  + **Áp dụng quy tắc nghiệp vụ:** ví dụ xếp loại học lực theo ngưỡng (>=8 Giỏi, 6.5–<8 Khá, ...).
  + **Kiểm tra bất thường (anomaly detection):** phát hiện điểm nhập sai (ví dụ 150 điểm).
  + **Gộp và kết hợp dữ liệu:** kết hợp thông tin học sinh với thông tin lớp, môn học.
  + **Cập nhật dữ liệu:** lưu kết quả tính toán vào bảng/tệp kết quả.
* **Hiệu suất & độ chính xác:** chú ý đến chi phí tính toán với dữ liệu lớn (tối ưu hóa truy vấn, xử lý theo lô — batch hoặc theo luồng — stream).

**Kết quả:** Tập thông tin đã qua xử lý, sẵn sàng lưu trữ hoặc xuất ra.

**3. Lưu trữ (Storage)**

**Mục đích:** Giữ lại dữ liệu và kết quả xử lý để truy xuất, báo cáo hoặc phân tích sau này.

**Nội dung chi tiết:**

* **Loại lưu trữ:** cơ sở dữ liệu quan hệ (MySQL, PostgreSQL), NoSQL (MongoDB), file (CSV, Excel), lưu trữ đám mây (S3), hoặc hệ thống dữ liệu lớn (HDFS).
* **Cấu trúc lưu trữ:** thiết kế bảng (tables), định nghĩa khóa chính/khóa ngoại, chỉ mục (index) để truy vấn nhanh.
* **Bảo mật & phân quyền:** phân quyền đọc/ghi, mã hóa dữ liệu nhạy cảm, quản lý truy cập.
* **Sao lưu & phục hồi (backup & recovery):** lịch sao lưu định kỳ, kịch bản khôi phục khi mất dữ liệu.
* **Quản lý phiên bản (versioning):** ghi lại lịch sử sửa đổi nếu cần (audit trail).

**Kết quả:** Dữ liệu được lưu an toàn, có thể truy xuất theo yêu cầu.

**4. Trình bày kết quả (Output)**

**Mục đích:** Cung cấp thông tin đã xử lý cho người dùng hoặc hệ thống khác dưới dạng dễ hiểu và hữu ích.

**Nội dung chi tiết:**

* **Hình thức xuất:** giao diện web, báo cáo PDF/Word, biểu đồ/ dashboard, file CSV/Excel, thông báo email hoặc in ấn.
* **Tùy biến theo đối tượng:** báo cáo cho giáo viên có chi tiết, cho phụ huynh tóm tắt, cho ban giám hiệu có thêm phân tích lớp/ trường.
* **Tính trực quan:** dùng bảng, biểu đồ cột/đường, highlight điểm cao/thấp, sắp xếp theo thứ tự cần thiết.
* **Tự động hóa:** lên lịch gửi báo cáo, xuất file hàng loạt.

**Kết quả:** Người dùng nhận được thông tin đúng lúc, rõ ràng, và có thể hành động dựa trên đó.

**Ví dụ thực tế: Hệ thống quản lý điểm học sinh**

Dưới đây là mô tả chi tiết từng bước theo ví dụ cụ thể.

**A. Nhập dữ liệu**

* **Nguồn:** giáo viên nhập điểm kiểm tra, điểm thi qua form web hoặc upload file Excel.
* **Trường dữ liệu mẫu:** student\_id, student\_name, class, subject, score\_type (quy ước: KT, Thi), score\_value, date, teacher\_id.
* **Kiểm tra:** điểm phải nằm trong phạm vi 0–10; student\_id tồn tại trong danh sách học sinh; tránh trùng bản ghi.
* **Ghi log:** lưu ai đã nhập, thời gian, nguồn dữ liệu (tay/Excel/API).

**B. Xử lý dữ liệu**

* **Tính điểm trung bình:** ví dụ công thức: average = (average\_of\_quizzes \* 0.3) + (exam\_score \* 0.7) — tùy theo quy định nhà trường.
* **Xếp loại:** áp dụng ngưỡng: >= 8.0 → Giỏi, 6.5–<8.0 → Khá, 5.0–<6.5 → Trung bình, <5.0 → Yếu.
* **Rounding:** làm tròn đến 2 chữ số thập phân nếu cần.
* **Phát hiện ngoại lệ:** nếu score\_value > 10 hoặc < 0 thì báo lỗi cho người nhập.
* **Tính thêm:** xếp hạng lớp, phân tích điểm theo môn, thống kê điểm yếu/phổ điểm.

**C. Lưu trữ**

* **Mô tả bảng đơn giản:**
  + Students(student\_id, name, class, dob)
  + Scores(score\_id, student\_id, subject, score\_type, score\_value, date, teacher\_id)
  + Averages(student\_id, subject, average, grade, term)
* **Sao lưu:** backup hàng ngày, lưu trữ bản sao trên cloud.
* **Bảo mật:** chỉ giáo viên và admin có quyền ghi; phụ huynh chỉ có quyền đọc báo cáo con mình.

**D. Xuất kết quả (Output)**

* **Báo cáo bảng điểm:** hiển thị danh sách điểm theo học sinh/ theo lớp, có cột điểm trung bình và xếp loại.
* **Biểu đồ:** histogram phân bố điểm, biểu đồ tiến bộ theo thời gian.
* **In ấn / PDF:** xuất phiếu báo điểm cá nhân để in hoặc gửi dưới dạng PDF.
* **Gửi thông báo:** tự động gửi email/SMS khi điểm mới được cập nhật (nếu hệ thống hỗ trợ).