#### Vì sao phải khám phá tri thức và khai phá dữ liệu?

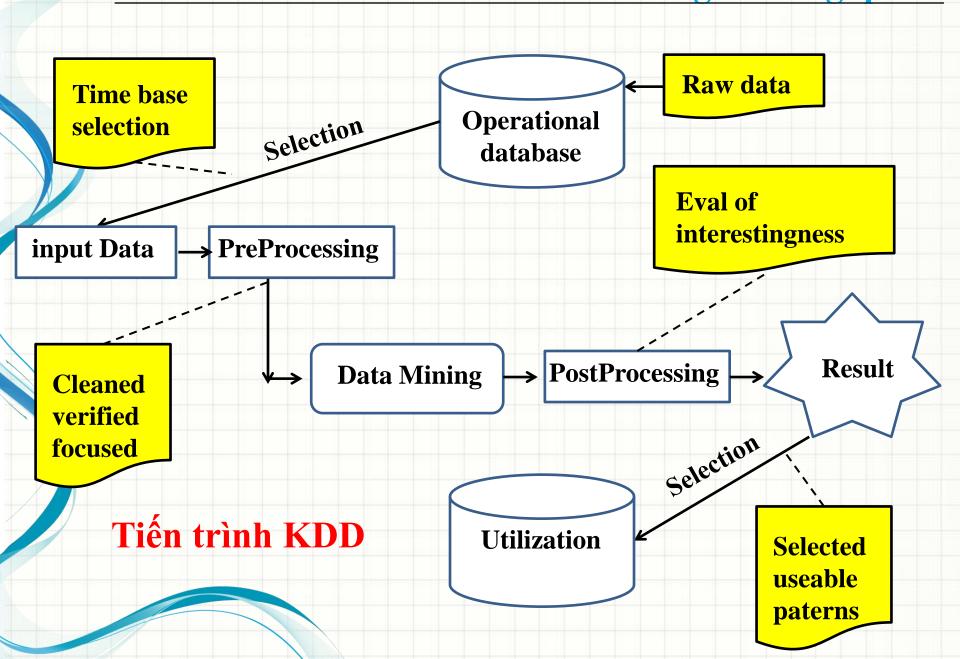
- Khối lượng dữ liệu lớn
- Nhu cầu trợ giúp nhanh ra quyết định
- Tìm kiếm tri thức tiềm ẩn
- Tìm kiếm các mô hình, dự báo ẩn sâu trong cơ sở dữ liệu

#### Khám phá tri thức là gì?

Khám phá kiến thức là quá trình xác định giá trị, tính hữu ích, và các mô hình tối ưu trong hệ cơ sở dữ liệu.

## Khai phá dữ liệu là gì?

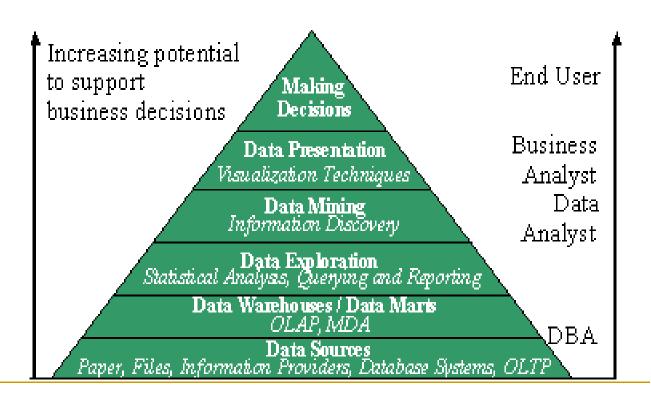
Khai phá dữ liệu là một bước trong quá trình phát hiện kiến thức bao gồm các thuật toán khai thác hệ dữ liệu theo các phương pháp cụ thể để tìm ra các mẫu hoặc các mô hình trong dữ liệu.



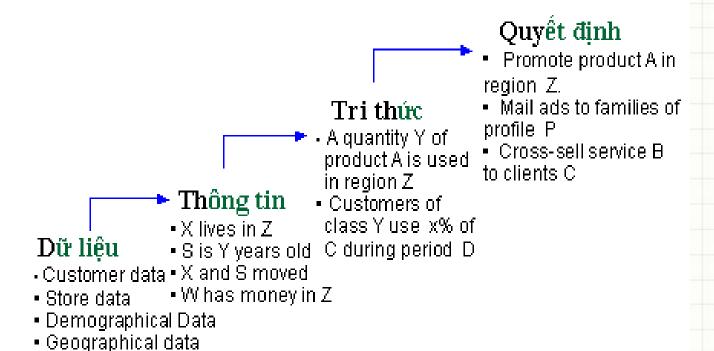
#### Lợi ích của khai phá dữ liệu?

- Cung cấp tri thức hỗ trợ ra quyết định
- Dự báo
- Khái quát dữ liệu

#### Khai phá dữ liệu







## Các lĩnh vực có liên quan đến khám phá tri thức?

- Học máy
- Thống kê, phân tích dữ liệu
- Kho dữ liệu
- Hệ chuyên gia

#### Các ứng dụng trong khai phá dữ liệu?

- Phân tích dữ liệu, hỗ trợ ra quyết định:
  - \* Phân tích quản lý thị trường
  - \* Quản lý và phân tích rủi ro
  - \* Quản lý và phân tích các sai hỏng
- Nghiên cứu khoa học và hệ thống thông tin:
  - \* Khai thác dữ liệu web, văn bản text,..
  - \* Quản lý mạng, internet
  - \* Khảo sát, cơ sở dữ liệu,...

## Tiến trình khai phá dữ liệu?

Nghiên cứu lĩnh vực và đặt bài toán

Tạo và thu thập dữ liệu đầu vào

Tiền xử lý dữ liệu: làm sạch, mã hóa,..

Rút gọn/chiều

Lựa chọn tác vụ khai phá dữ liệu

# Tiến trình khai phá dữ liệu? (tt)

Chọn phương pháp khai phá dữ liệu

Khai phá dữ liệu: Tìm kiếm tri thức

Đánh giá mẫu tìm được

Biểu diễn tri thức

Sử dụng các tri thức vừa khám phá