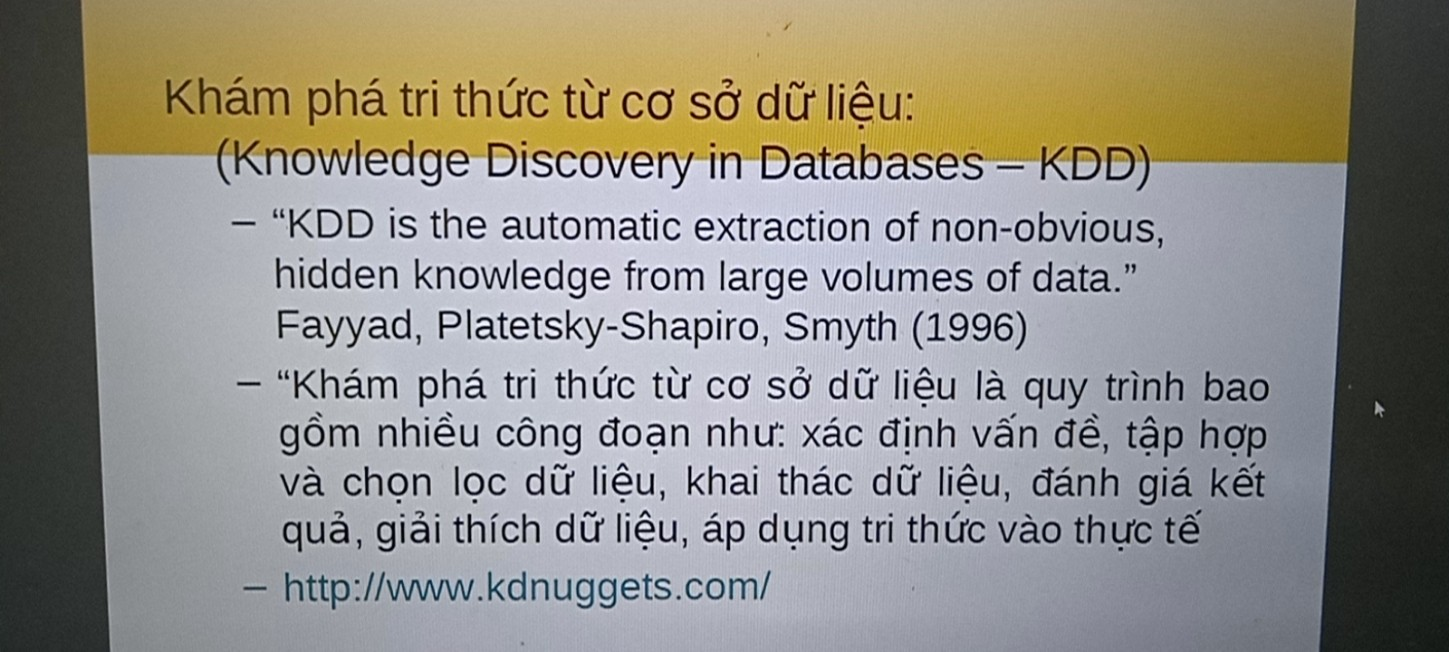
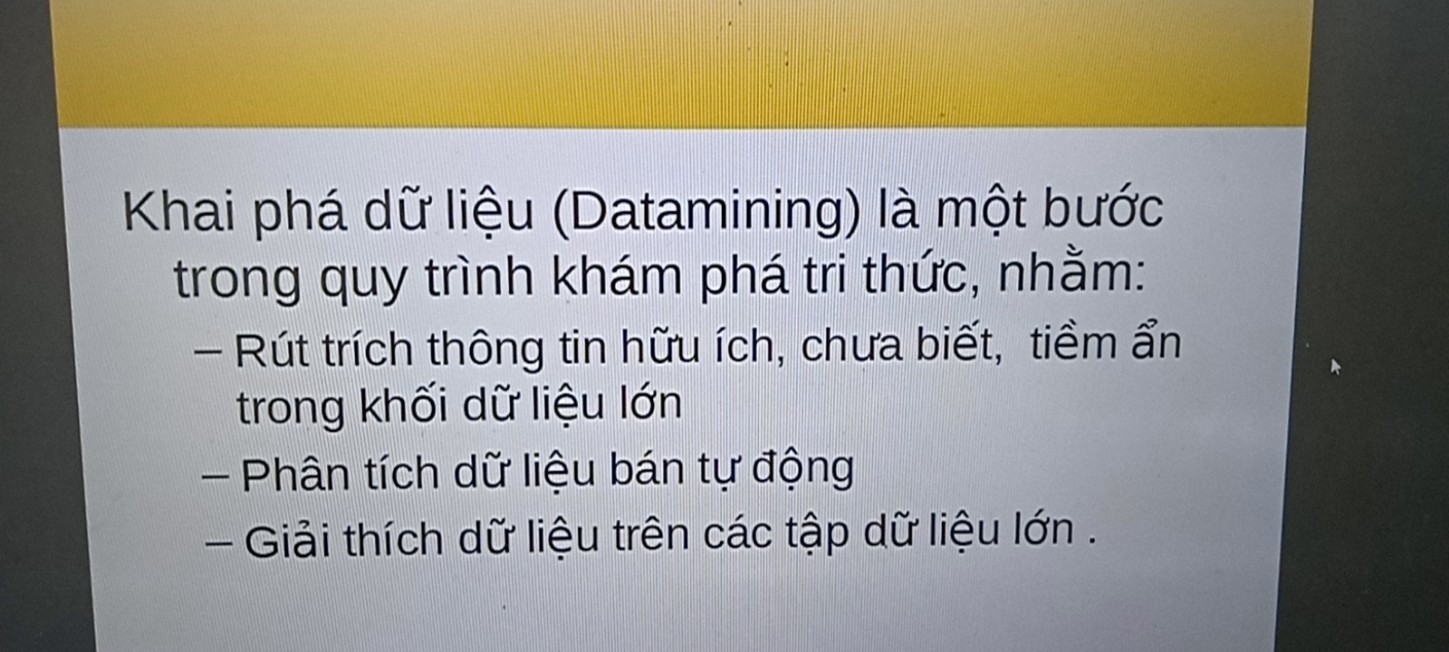
Công nghệ dữ liệu lớn :

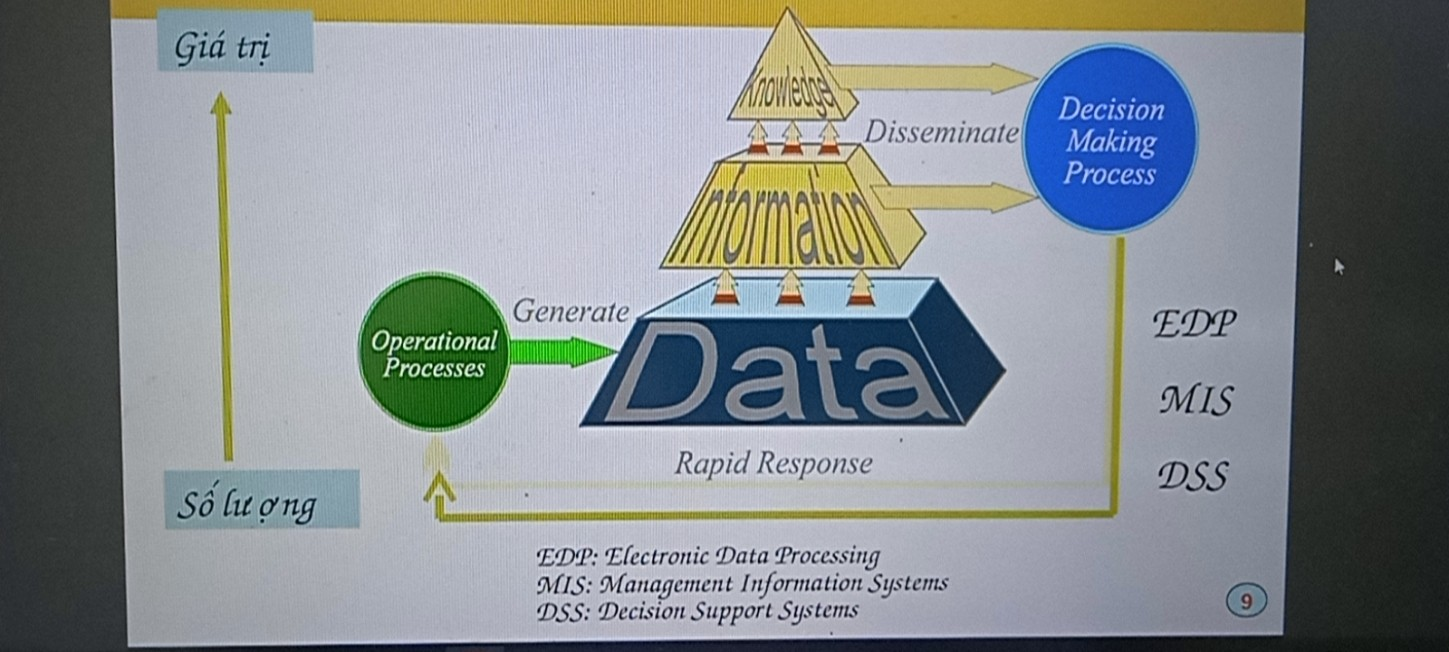
* Node js
* Hadoop Apache
* Mapreduce
* Anaconda
* OOP
* Pandas ( Để làm quen với Python và Pandas thì mình gửi các bạn 5 bài thực hành Pandas nhé. [Các bạn về cài đặt Anaconda rồi sử dụng Jupyter hoặc spider hoặc cùng lắm là trình Idle để làm bài tập; nếu sử dụng Google Colab thì vẫn sử dụng được Pandas nhưng phải tra cứu cách tải dữ liệu lên]. Đây là bài tập nhóm <=6 người cùng làm để.) . Theo dõi trên nhóm có dữ liệu .
* Hadoop Tutorial . ( Học trên trang Java T Point ) .
* Numpy ( The following are elements that make up a 3D object Visualization - - Renderer - Scene - Camera - Mesh - Lighting - Animating Nghĩa là làm 3D là làm 1 trong những cái trên á. (làm hết là siu nhơn )
* Xem thêm về power bi
* Về khai phá dữ liệu có một điểm chú ý:



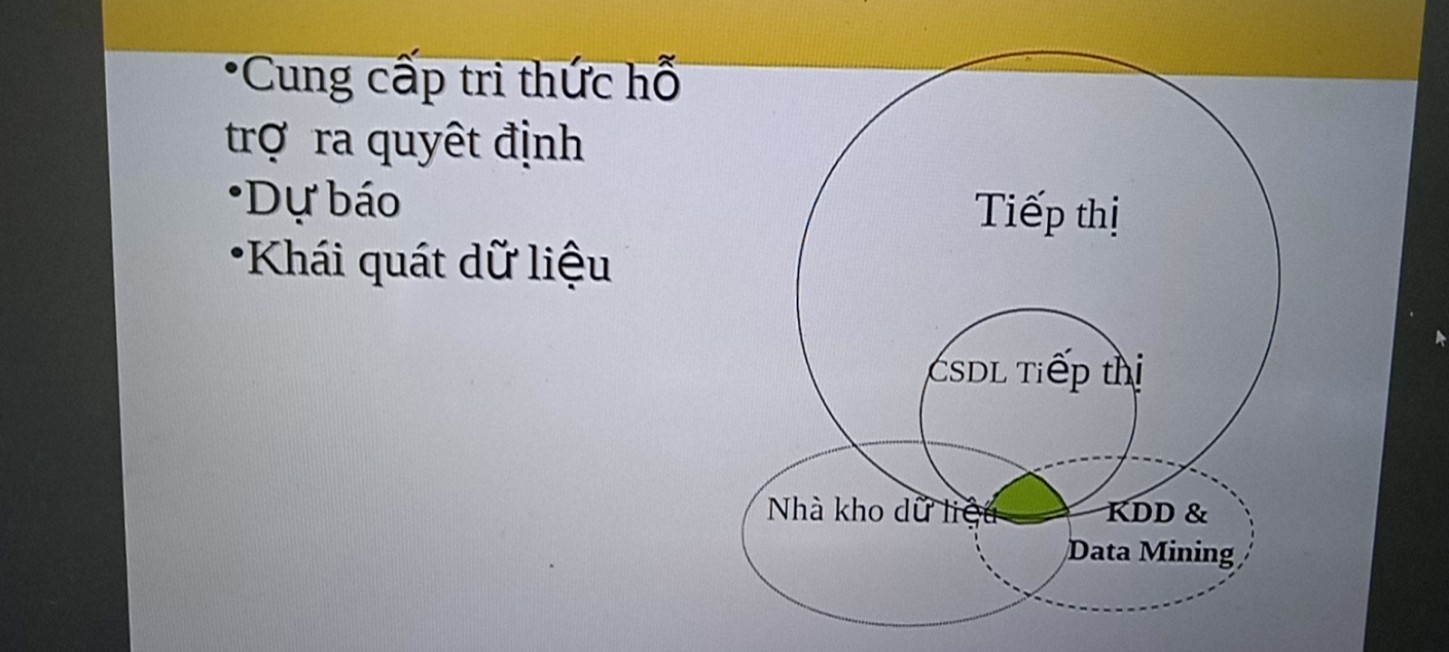


Bị ngập tràn dữ liệu mà thiếu tri thức

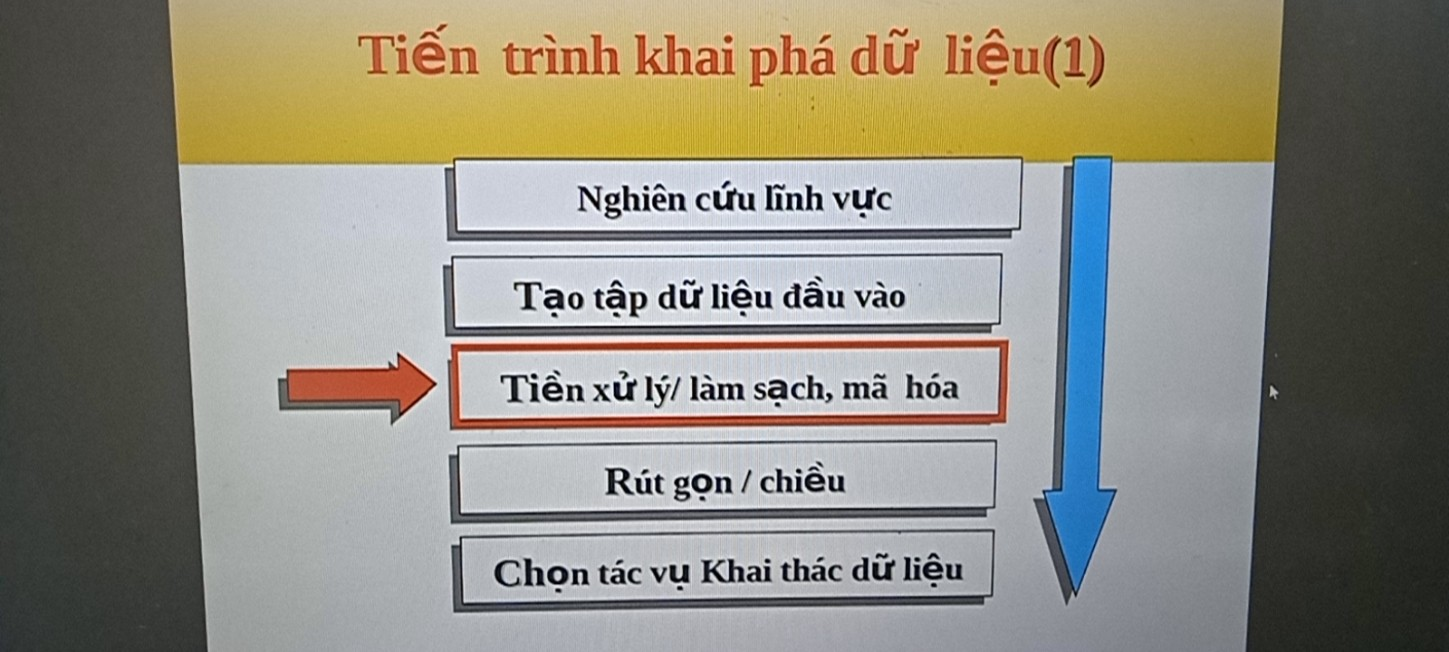




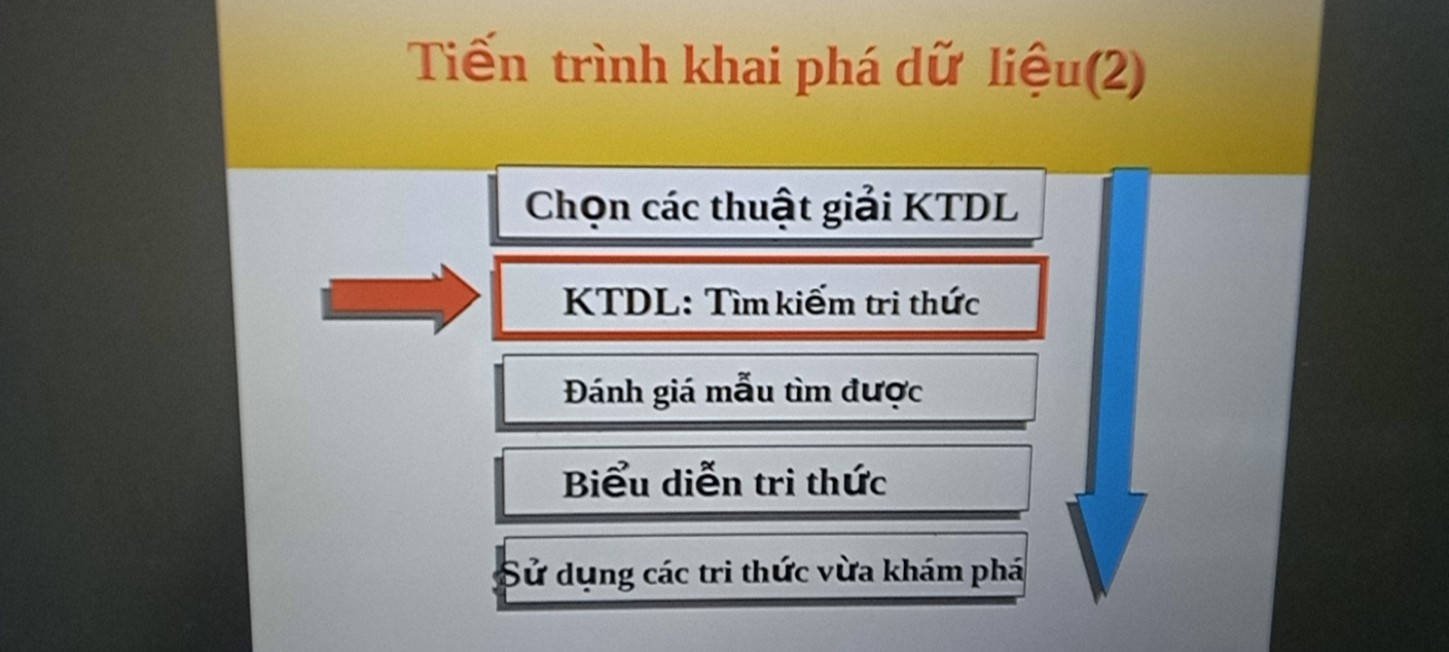
* Hệ MIS là hệ thông tin quản trị thường cho doanh nghiệp ( <https://eller.arizona.edu/departments-research/schools-departments/mis/what-is-mis#:~:text=MIS%20(Management%20Information%20Systems)> Đường link .
* Ví dụ về khai phá dữ liệu trong kinh doanh



Mô hình truyền thống



Bài toán về rút gọn chiều là một trong những bài toán quan trọng



* Rút gọn chiều cho dữ liệu ( <https://en.m.wikipedia.org/wiki/Dimensionality_reduction> )
* Thay data warehouse = data pool , data lake .



* Nhắc lại về quy hoạch tuyến tính .
* Chuyển sang Firebase của Google để có khái niệm CSDL trên cloud .