Big Data

* Big data is a broad term for data sets so large or complex that traditional data processing applications are inadequate
* Often refers simply to the use of predictive analytics
* Analysis of data sets can find new correlations to "spot business trends, prevent diseases, combat crime
* Characteristics:
* Volume

The quantity of generated and stored data. The size of the data determines the value and potential insight- and whether it can actually be considered big data or not.

* Variety

The type and nature of the data. This helps people who analyze it to effectively use the resulting insight.

* Velocity

In this context, the speed at which the data is generated and processed to meet the demands and challenges that lie in the path of growth and development.

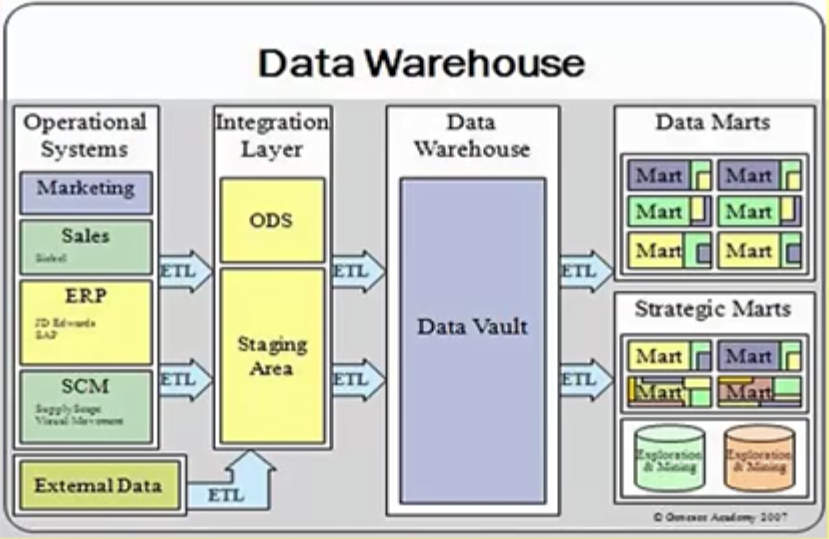
* Variability

Inconsistency of the data set can hamper processes to handle and manage it.

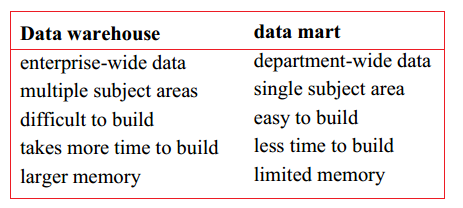
* Veracity

The quality of captured data can vary greatly, affecting accurate analysis.

Data warehousing



* DWs are central repositories of integrated data from one or more disparate sources.
* A data mart is a simple form of a data warehouse that is focused on a single subject



* This definition of the data warehouse focuses on data storage. The main source of the data is cleaned, transformed, cataloged and made available for use by managers and other business professionals for data mining

GIS geographical information system

* A geographic information system (GIS) lets us visualize, question, analyze, and interpret data to understand relationships, patterns, and trends.
* Advantages:
* Cost Savings from Greater Efficiency
* Better Decision Making
* Improved Communication
* Better Record Keeping



SAP

* SAP HANA is a static tool which in an in-money , column oriented , relational database management system developed by SAP company
* It is an enterprise-wide information system designed to coordinate all the resources, information, and activities needed to complete business processes such as order fulfillment, billing, human resource management, and production planning
* It’s also used as enterprise resource planning
* In the case of an SAP component or product like ECC, the programs and data reside together in the same database.
* SAP typically focused on best practice methodologies for driving its software processes, but more recently expanded into vertical markets
* SAP based on a three-tier client/server structure:

1. Presentation Layer (GUI)

2. Application Layer

3. Database Layer

* Server-to-server communications can be encrypted with the SAP cryptographic library.