

CST8390
BUSINESS
INTELLIGENCE &
DATA ANALYTICS

Introduction to BI
BI Components & Architecture

Professor: Dr. Anu Thomas
Email: thomasa@algonquincollege.com
Office: T314

Agenda

- Assignment 3
- Final Project
- BI, Components & Architecture



Business Intelligence

- Business Intelligence systems have been around for almost 20 years now. They try to summarize large datasets and try to understand what is happening.
- Definition: Business Intelligence is the process of collecting data to put together a picture of what is going on in a business.
- These data are then analyzed to detect trends (good or bad) so that management staff can make informed decisions.



Business Intelligence

- It is meant to help management to make decisions:
 - Strategic – use information to make better decisions
 - Tactical – reaching short term goals
 - Operational – helps identify processes to optimize
- <http://www.managementstudyguide.com/strategic-decisions.htm>



Strategic vs Administrative vs Operational

Strategic Decisions	Administrative Decisions	Operational Decisions
Strategic decisions are long-term decisions.	Administrative decisions are taken daily.	Operational decisions are not frequently taken.
These are considered where The future planning is concerned.	These are short-term based Decisions.	These are medium-period based decisions.
Strategic decisions are taken in Accordance with organizational mission and vision.	These are taken according to strategic and operational Decisions.	These are taken in accordance with strategic and administrative decision.
These are related to overall Counter planning of all Organization.	These are related to working of employees in an Organization.	These are related to production.
These deal with organizational Growth.	These are in welfare of employees working in an organization.	These are related to production and factory growth.



Business Intelligence

- BI systems have 4 main components:
 - Data Warehouses
 - Business Analytics tools – manipulating and analyzing data
 - Business Performance tools – monitoring and analyzing performance
 - Visualization– portals, dashboards and scorecards.

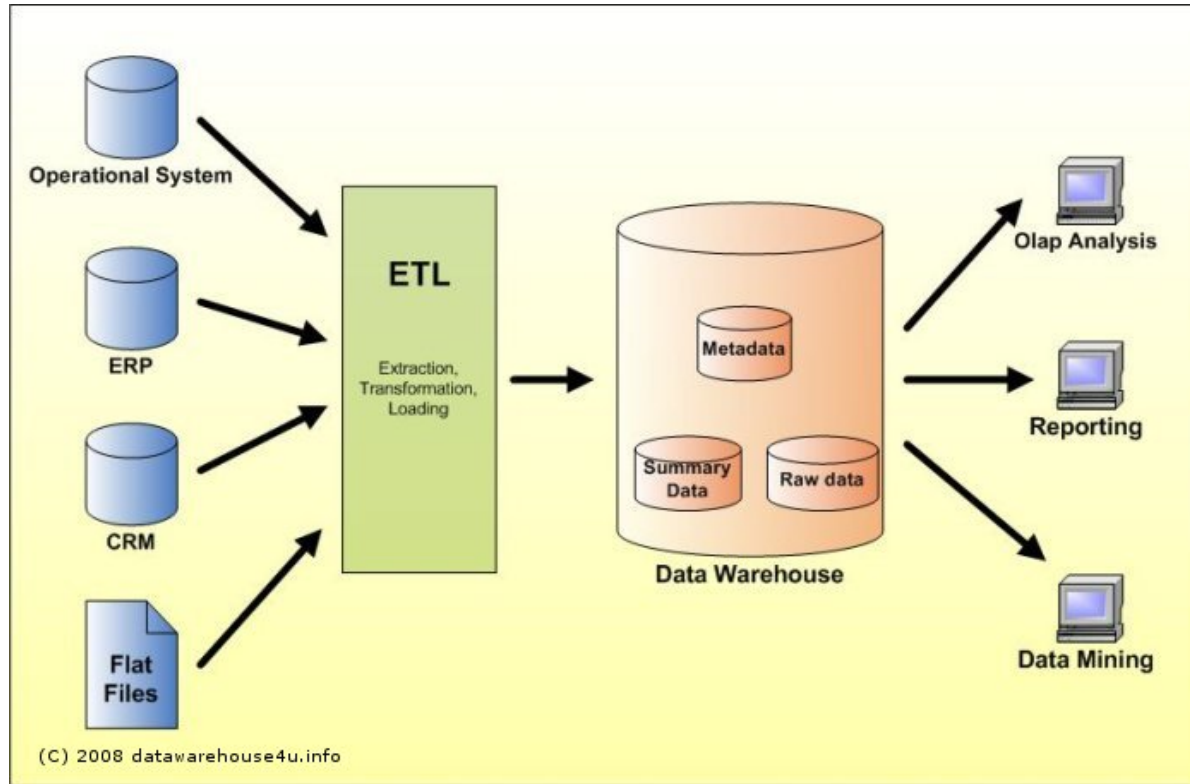


Data Warehouse

- A Data Warehouse is a large store of data that is accumulated from different sources within a company.
- The Data Warehouse pulls data from inventory, sales, etc. so that an overall picture can be created to support decision making. A Data Warehouse is typically updated on a daily basis.



Data warehouse



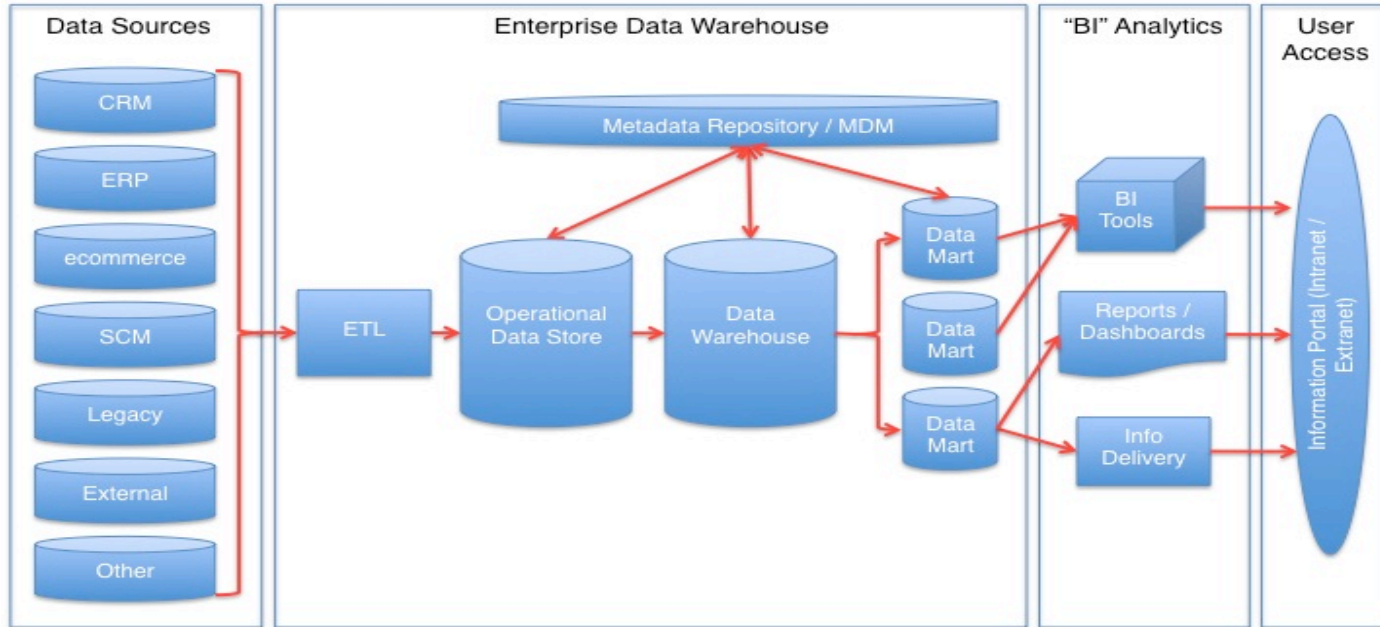
Picture taken from www.datawarehouse4u.info

ETL

- The process of gathering data is referred to as Extract, Transform, Load (ETL)
 - Extraction pulls data from its raw source.
 - Transform manipulates the data to convert it to the format you want.
 - Load stores the data in the final target (Data Warehouse or Data Mart).
- If the data won't fit into memory, then you must store it first at the target database and then transform it after. This process is referred to as ELT, where it is loaded before being transformed.

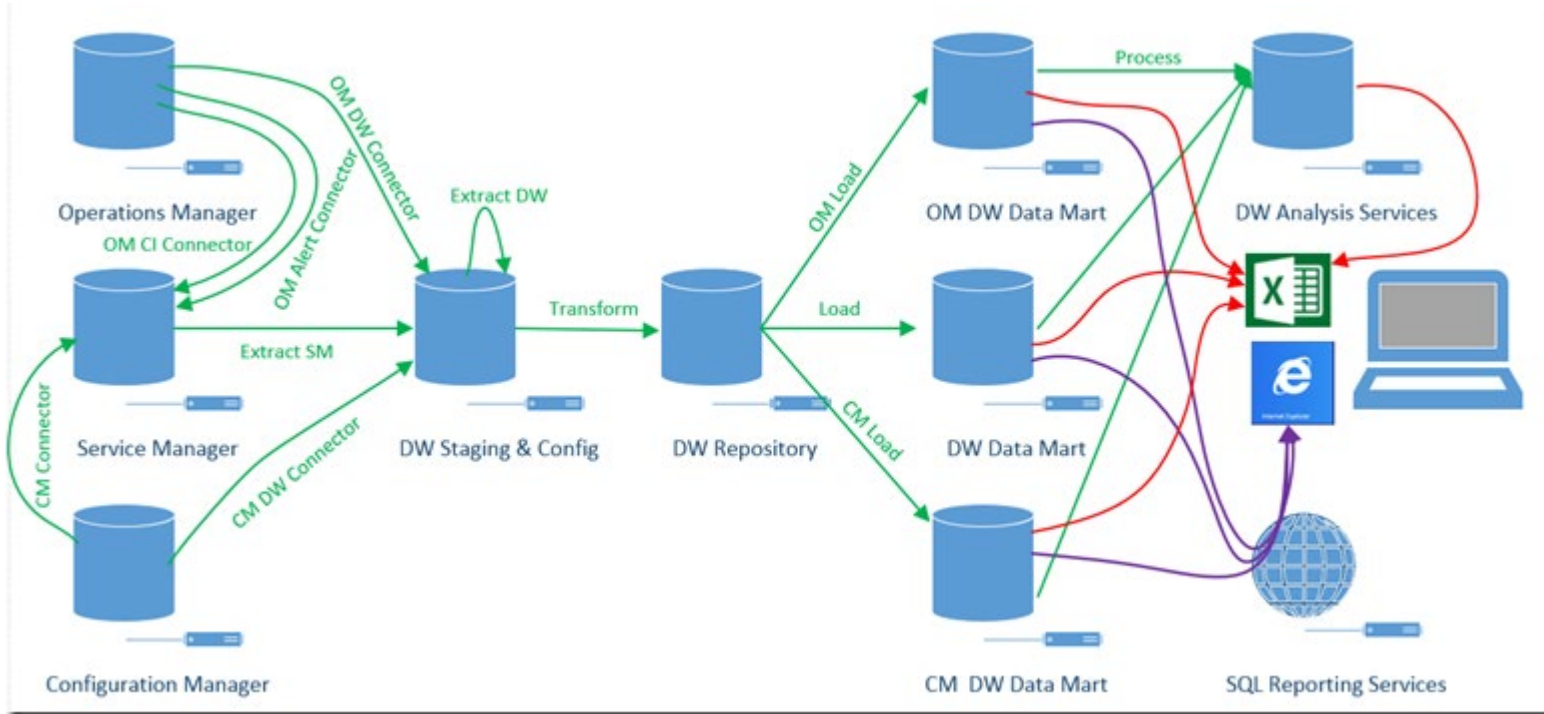


Data Warehouse



Picture taken from https://en.wikipedia.org/wiki/Data_warehouse

Data warehouse (Contd.)



Picture taken from <https://blogs.technet.microsoft.com/servicemanager/2013/02/16/new-data-warehouse-architecture-diagram/>

Data Warehouse vs Database

	Data Warehouse	Database
Definition	Pulls together data from different sources for reporting and analysis	Normalized data organized in columns, rows and tables
Purpose	To store large quantities of historical data and enable fast, complex queries across all data	To store current transactions and enable fast access to specific transactions for ongoing business processes
Data structure	Use fewer tables Doesn't exclude data redundancies	Related data separated into multiple tables. Data is organized to make sure no redundant data
Data	Denormalized data Offers better performance when reading data for analytics purposes	Normalized data More complex queries required to read the data as a single data combines data from many tables



OnLine Transaction Processing System - OLTP

- These are systems that process any type of transaction (Banking, Ticket Reservation, Point of Sale, etc.)
- Online means that it is always running, waiting for input.
- They ensure speed of transactions and maintain data integrity.
- Mainly INSERT, UPDATE, DELETE operations



OnLine Analytical Processing - OLAP

- OLAP systems operate long-running queries through the data looking for patterns. These are the systems that generate business reports and mines the data.
- They work with data from the Data Warehouse to provide summaries.
- <http://www.solveglobal.com/blog/2014/04/data-warehouse-vs-olap-cube/>



OLTP Query Example

```
SELECT PASSWORD FROM ACCOUNTS  
WHERE USERID = 1131234567
```

OLAP Query Example

```
SELECT dealer, year, SUM(price)  
FROM (Sales NATURAL JOIN Autos) JOIN Days ON date = day  
WHERE model = 'Gobi' AND  
       color = 'red' AND  
       (year = 2001 OR year = 2002)  
GROUP BY year, dealer;
```



OLTP vs. OLAP

OLAP	OLTP
Works on historical data	Works on current data
Helps analyze the business	Helps run the business
Provides summarized data	Provides raw data
Use long-running queries to refresh summaries	Focus on fast & secure queries
Used by specialized users for decision support	Used by normal company staff



Data Marts

- A Data Mart is a mini warehouse, that typically addresses one portion of the warehouse: sales, marketing, finance, etc. Each part of the organization would have their own specialized data mart.
- It is extracted from a Data Warehouse and can store summarized data instead of raw data.



Decision Making

- In Decision Making, it is important to understand what is happening, so you can do something about it:
 - Define the problem.
 - Constructing a Model for computer simulations.
 - Identify and Evaluating possible solutions
 - Recommending potential solutions.



Decision Support Systems

- There are several kinds of decisions that can be faced:
 - Structured decisions – choosing whether to buy a new product or service, what inventory levels are needed, etc.
 - Semi-structured decisions – setting budgets, pricing
 - Unstructured decisions – no clear solutions: advertising, designs
- Read: <http://www.inc.com/encyclopedia/decision-support-systems.html>



Types of DSS

- Data Driven – Process numerical data from data warehouse. Produces dashboards and scorecards.
- Document Driven – Works on documents, videos, transcripts, media (Wikipedia is an example).
- Knowledge Driven – Uses knowledge and past experiences to support decision making. (Neural networks)
- Model Driven – Uses models and simulation to support decision making.
- Communication Driven – Group support systems for collaboration (Facebook, Skype)



Dashboards

- Once the data has been gathered, loaded and transformed, the results can be presented.
- Dashboards are a graphical presentation showing whatever data that a decision maker needs to know.
- This process is automated on a regular basis. For reports that are not needed on a regular basis, there are ad-hoc reports which are executed when needed.



Dashboards

- Dashboards are graphical interfaces that display Key Performance Indicators (KPIs). It gives managers a quick overview of the KPIs.
- A dashboard should draw your attention to items that are not within expected ranges (sales are down, accidents are high etc.)
- Dashboards should be hierarchical, meaning that a user should be able to “drill down” to view more of the data.

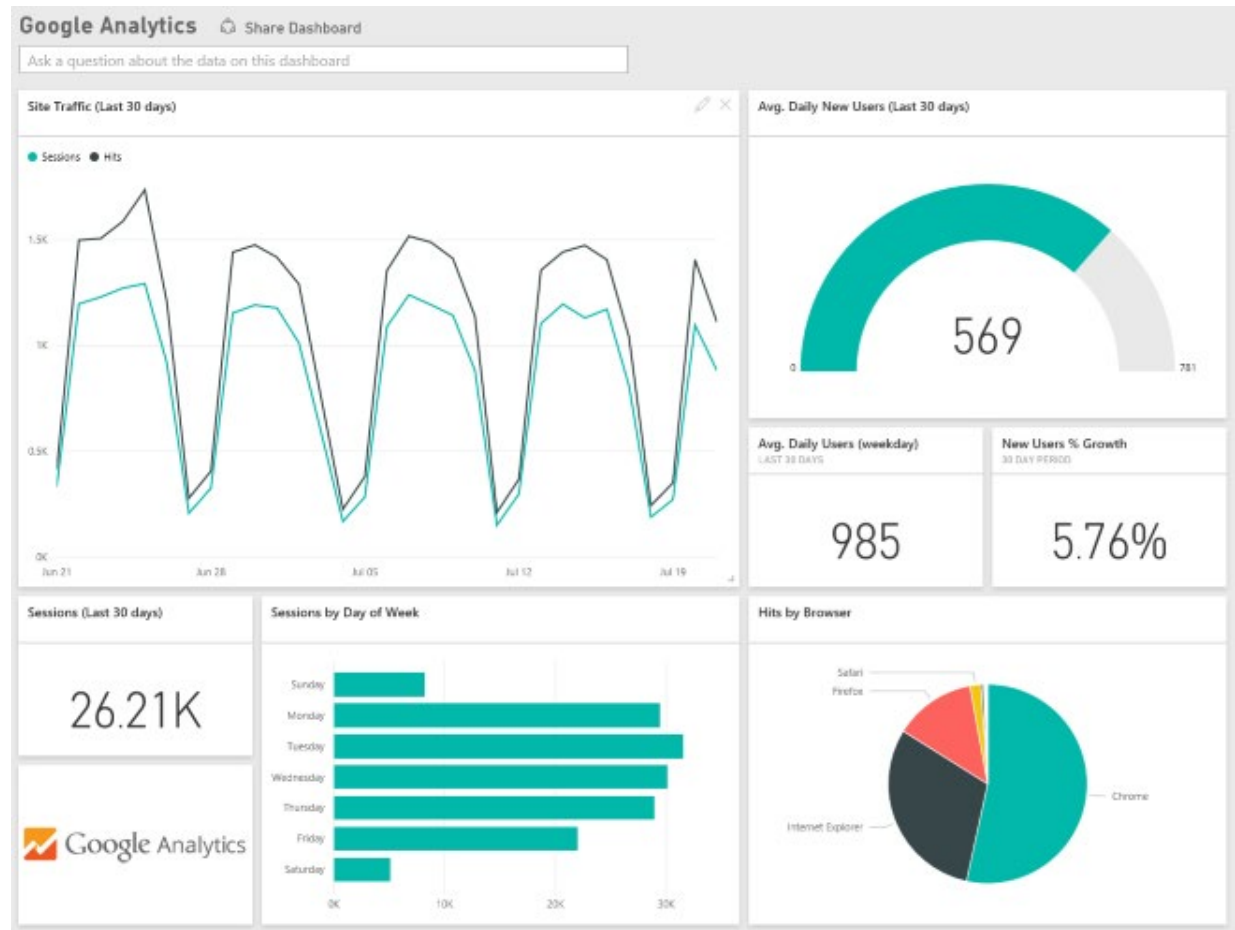


Dashboard Example



<http://enterprise-dashboard.com/sas-dashboards-to-get-slick-kpi-gauges-out-of-the-box-in-next-release/>

Dashboard Example



Drilling down

- Clicking on a dashboard item (for example: total sales) should show another screen showing the data that was used in the calculation.
- Drilling down on total sales should show a detailed screen showing sales in all provinces, or states. Clicking on one of the provinces should then show sales by city in the province. Clicking on a city should then show sales per store within that city.



Key Performance Indicators KPI

- Key performance indicators are a set of indicators to measure data against a goal, or success metric.
- Did the company meet their sales goal for the month?
- Were products built and delivered on time?
- Was quality (defects per 1000 units) maintained?
- <https://www.klipfolio.com/resources/kpi-examples>
- <http://www.pnmsoft.com/resources/bpm-tutorial/key-performance-indicators/>
- <https://www.collegesontario.org/en/resources/2019-20-key-performance-indicators>



Review

- Data is retrieved from the organization (Extraction)
- It is collected, cleaned converted and processed (Transformed)
- The results are stored in databases (Loaded)
- Decision makers can then view the data in graphical form using Dashboards.
- There are different data types (Structured, Unstructured, Semi-structured).



Review

- Watch videos on Business Intelligence:
 - <https://www.youtube.com/watch?v=LFnewuBsYiY>
 - https://www.youtube.com/watch?v=KGHbY_Sales
- Read Chapter 1 of textbook (skip case study)
- Some additional references:
 - <https://www.promptcloud.com/blog/business-intelligence-Vs-data-analytics/>
 - <https://www.inetsoft.com/evaluate/demo/flashdemo.jsp>
 - https://www.youtube.com/watch?v=eiY8GEF_0-U

