Characteristics and properties

1. **Subject-oriented –**

A data warehouse is always a subject oriented as it delivers information about a theme instead of organization’s current operations. It can be achieved on specific theme. That means the data warehousing process is proposed to handle with a specific theme which is more defined. These themes can be sales, distributions, marketing etc.  
  
A data warehouse never put emphasis only current operations. Instead, it focuses on demonstrating and analysis of data to make various decision. It also delivers an easy and precise demonstration around particular theme by eliminating data which is not required to make the decisions.

1. **Integrated –**  
   It is somewhere same as subject orientation which is made in a reliable format. Integration means founding a shared entity to scale the all similar data from the different databases. The data also required to be resided into various data warehouse in shared and generally granted manner.  
     
   A data warehouse is built by integrating data from various sources of data such that a mainframe and a relational database. In addition, it must have reliable naming conventions, format and codes. Integration of data warehouse benefits in effective analysis of data. Reliability in naming conventions, column scaling, encoding structure etc. should be confirmed. Integration of data warehouse handles various subject related warehouse.
2. **Time-Variant –**  
   In this data is maintained via different intervals of time such as weekly, monthly, or annually etc. It founds various time limit which are structured between the large datasets and are held in online transaction process (OLTP). The time limits for data warehouse is wide-ranged than that of operational systems. The data resided in data warehouse is predictable with a specific interval of time and delivers information from the historical perspective. It comprises elements of time explicitly or implicitly. Another feature of time-variance is that once data is stored in the data warehouse then it cannot be modified, alter, or updated.
3. **Non-Volatile –**  
   As the name defines the data resided in data warehouse is permanent. It also means that data is not erased or deleted when new data is inserted. It includes the mammoth quantity of data that is inserted into modification between the selected quantity on logical business. It evaluates the analysis within the technologies of warehouse.  
     
   In this, data is read-only and refreshed at particular intervals. This is beneficial in analysing historical data and in comprehension the functionality. It does not need transaction process, recapture and concurrency control mechanism. Functionalities such as delete, update, and insert that are done in an operational application are lost in data warehouse environment. Two types of data operations done in the data warehouse are:
   * Data Loading
   * Data Access

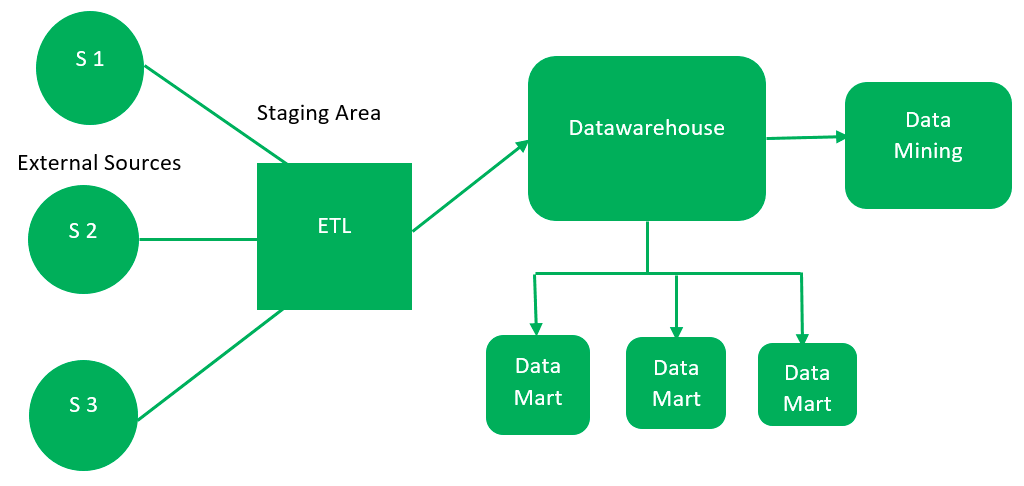
**Functions of Data warehouse:**  
It works as a collection of data and here is organized by various communities that endures the features to recover the data functions. It has stocked facts about the tables which have high transaction levels which are observed so as to define the data warehousing techniques and major functions which are involved in this are mentioned below:

1. Data consolidation
2. Data Cleaning
3. Data Integration

Development of datawarehouse approach

A **data-warehouse** is a heterogeneous collection of different data sources organised under a unified schema. There are 2 approaches for constructing data-warehouse: Top-down approach and Bottom-up approach are explained as below.

**1. Top-down approach:**

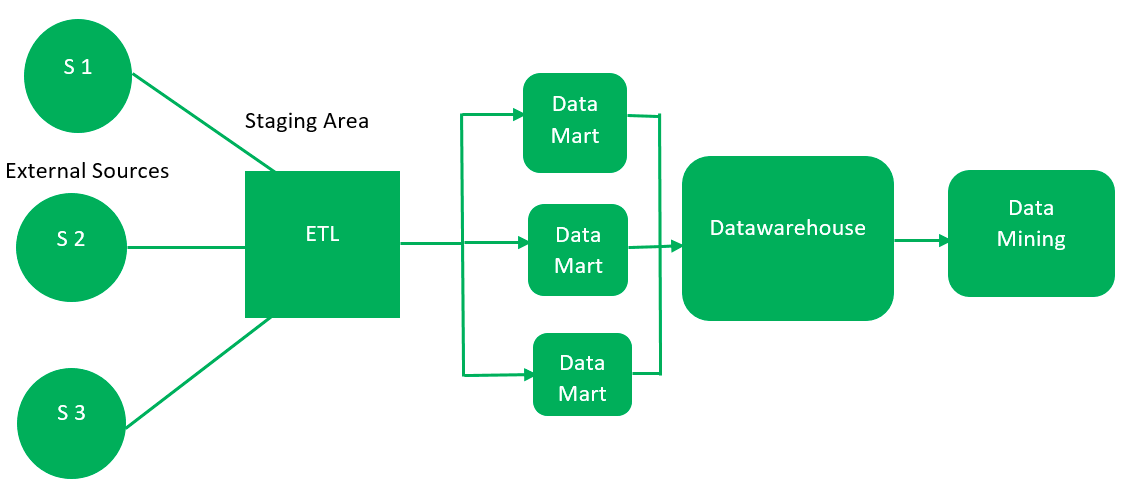


The essential components are discussed below:

1. **External Sources –**   
   External source is a source from where data is collected irrespective of the type of data. Data can be structured, semi structured and unstructured as well.
2. **Stage Area –**   
   Since the data, extracted from the external sources does not follow a particular format, so there is a need to validate this data to load into datawarehouse. For this purpose, it is recommended to use **ETL** tool.
   * **E(Extracted):** Data is extracted from External data source.
   * **T(Transform):** Data is transformed into the standard format.
   * **L(Load):** Data is loaded into datawarehouse after transforming it into the standard format.
3. **Data-warehouse –**   
   After cleansing of data, it is stored in the datawarehouse as central repository. It actually stores the meta data and the actual data gets stored in the data marts. **Note** that datawarehouse stores the data in its purest form in this top-down approach.
4. **Data Marts –**   
   Data mart is also a part of storage component. It stores the information of a particular function of an organisation which is handled by single authority. There can be as many number of data marts in an organisation depending upon the functions. We can also say that data mart contains subset of the data stored in datawarehouse.
5. **Data Mining –**   
   The practice of analysing the big data present in datawarehouse is data mining. It is used to find the hidden patterns that are present in the database or in datawarehouse with the help of algorithm of data mining.

This approach is defined by **Inmon** as – datawarehouse as a central repository for the complete organisation and data marts are created from it after the complete datawarehouse has been created. 

**2. Bottom-up approach:**



1. First, the data is extracted from external sources (same as happens in top-down approach).
2. Then, the data go through the staging area (as explained above) and loaded into data marts instead of datawarehouse. The data marts are created first and provide reporting capability. It addresses a single business area.
3. These data marts are then integrated into datawarehouse.

This approach is given by **Kinball** as – data marts are created first and provides a thin view for analyses and datawarehouse is created after complete data marts have been created.

Components of datawarehouse:

Load manager

Query manager

Warehouse manager

Enduser access tool

## 16 Steps on Your Business Intelligence Roadmap

### 1 Go into the process with eyes wide open

### 2. Determine stakeholder objectives

### 3. Choose a sponsor

### 4. BI is not just a technology initiative

### 5. Employ a Chief Data Officer (CDO)

### 6. Assess the current situation

### 7. Define a budget

### 8. Think of security, privacy, and compliance

### 9. Clean the data

### 10. Develop a “Data Dictionary”

### 11. Ensure data literacy

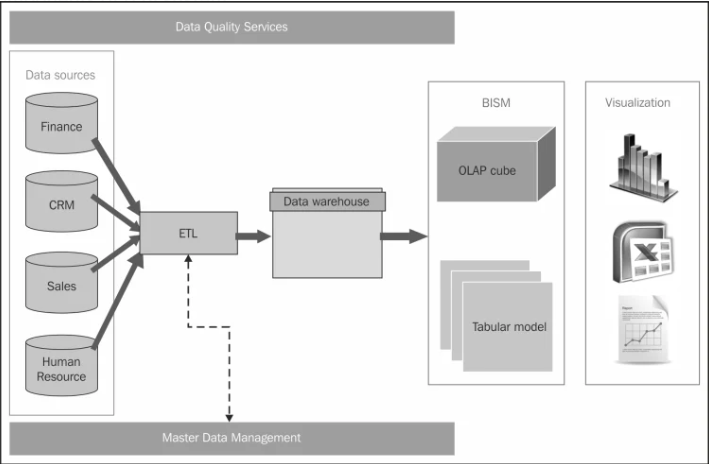
### 13. Choose the right tool/partner for your business

### 14. Rely on interactive data visualizations

### 15. Think about possible roadblocks and find solutions

### 16. Pursue a phased approach

The key elements of a business intelligence architecture are:



* Source systems
* ETL process
* Data modelling
* Data warehouse
* Enterprise information management (EIM)
* Appliance systems
* Tools and technologies

****Source Systems – Transaction Processing Systems****

This is the starting point for any BI initiative. Organization data is first created in these databases. Point to note: if you do not capture the data in the operational system, you can’t analyse it.

Operational systems (OLTP) form the bulk of the data needed for the data warehousing. In addition to that, source systems may also include data from secondary sources such as market data, benchmarking data etc. Business Intelligence architecture should address all these various data sources which are of different formats and standards.

****ETL Process****

In an ETL process data is extracted from the operational systems and loaded into a data warehouse. ETL, which stands for Extract Transform Load, is usually done using custom solutions available in the market. IBM Websphere Data Stage, Oracle Data Integrator, Ab Initio, and Microsoft Integration Services are examples of such tools.

****Data Modeling****

Data modeling will help to address what exactly is needed from data sources, the format of the data, and how it will be related to other data elements. It is not feasible to extract everything from a source system as that comes with cost issues. Data modeling will help to organize the data and therefore will minimize cost of storage replication, and effort needed to build a data warehouse.

****Data Warehouse****

Warehouse will have data extracted from various operational systems, transformed to make the data consistent, and loaded for analysis. A data warehouse will help in achieving cross-functional analysis, summarized data, and maintaining one version of the truth across the enterprise.

****Enterprise Information Management (EIM)****

EMI is another BI jargon which may stump some beginners. The term usually refers to ETL tools, data modeling tools, data quality, data profiling, metadata management, and master data management.

****BI Hardware****

It is important to make decisions on the hardware requirements to maintain a high performance and scalable BI system. Apart from server configurations, we have data warehouse appliances to combine the server, the database, and the data storage into one system. Netezza and DATAllegro are some well-known appliances in the market.

****Tools and Technologies****

Another important component of business intelligence architecture is what tools and technologies to implement. It is not just the front-end UI tools, but the tools used for EIM as well. There are cloud solutions, SaaS model, many full-fledged BI solutions (such as MSBI, Oracle BI suites, Microstrategy and more) to choose from. BI framework should have guidelines to make decisions on what is required for the organization.

# [The Business Intelligence Cycle](https://intellidsi.com/the-business-intelligence-cycle/)

## [IMG_256](https://intellidsi.com/wp-content/uploads/2014/09/The-Business-Intelligence-Cycle.jpg)**Phase 1: Setting Clear Objectives**

The effective use of BI begins with an awareness that BI is fundamentally different from reporting and has strategic implications if it is used to its full potential.

At the outset of a BI project (and on a periodic basis thereafter), business leaders should define the playing field in which they mean to operate. In many organizations, the initial focus for BI is on understanding the company’s existing customers, especially with respect to their buying behavior. Others may choose to zero in on sales performance and profitability by product line, division, or territory.

In most cases, BI initiatives will address multiple domains, but defining those areas of interest is a critical first step, largely because it determines how you will allocate resources, especially with respect to the second phase of the business intelligence cycle, gathering and organizing data. This is particularly important if your organization intends to enrich and extend your internal data with additional information from external sources.

## **Phase 2: Gathering and Organizing Data**

Whatever your company’s top priorities may be, it’s helpful to brainstorm a list of potential questions around those topics. Ultimately, this will shed light on the kind of information you need and what you need to do to consolidate and harmonize that information.

If you’re analyzing customer buying behavior, for example, you may need to aggregate information from your enterprise resource planning (ERP) system (which contains sales order transactions) and combine it with sales pipeline information from your customer relationship management (CRM) system, including sales quotes and deals that never came to fruition.

For many projects, internal data sources are likely to be sufficient. By brainstorming a list of questions in advance, you can begin to determine whether you should include data enrichment as part of your overall BI strategy. If the focus is on knowing your customer, for example, you may want to consider extending your corporate datasets with demographic details available from third-party sources.

As you consider your overall approach toward gathering and organizing the data for your BI initiative, you’ll need to determine how you can best harmonize and consolidate the information and make it available to the users who will rely on your business intelligence systems for meaningful insights.

BI tools that have pre-built integration to ERP are a distinct advantage–especially if they present information in a context front-line employees can easily understand and use. Systems that require advanced database skills or custom programming will cost more to operate and ultimately generate less value because of the technical barriers that stand in the way of widespread user adoption.

## **Phase 3: Analyzing the Information**

Analysis is where the magic happens; it’s where BI distinguishes itself from business reporting in general. Business reporting is more operational in nature than BI. Operational reports often have a short-term focus, and they are used to drive the daily decisions business leaders must make. Business reports may work with real-time transactional data connected directly to the source system.

On the other hand, BI typically has a long-term focus and concerns itself with trends and patterns. BI usually involves, not real-time data, but aggregated or summarized data that may have been loaded into a data warehouse and transformed for analysis. This distinction means that the data used in BI does not necessarily have a direct connection to source systems because it doesn’t need one.

Business intelligence facilitates exploration. BI makes it easy to perform ad hoc inquiries, which often prompt users to ask new questions they had never before considered.

BI makes it easy to take that next step by offering quick answers to those follow-up inquiries. It’s fast, flexible, and open-ended. It helps users to explore new territory and get rapid answers to questions that challenge conventional wisdom. When leaders face “what if” scenarios, BI helps them understand their options and narrow down their choices using a data-driven approach.

This is especially valuable when BI tools are available across a broad base of users within the organization. This is referred to as “data democratization,” and it can have a transformative effect because it empowers workers to better understand the forces that should impact their decisions, regardless of where they may sit within the organizational hierarchy.

## **Phase 4: Measuring and Monitoring KPIs**

Alongside this ad hoc analysis, you’ll want to leverage your BI systems to measure and monitor the key performance indicators (KPIs) that align to your organization’s strategic objectives. As discussed, business intelligence differs from traditional approaches to financial reporting in its focus. Rather than focusing on short-term operational matters, BI focuses on the data at a higher-level, moving from operational thinking to a more managerial approach. Because it can aggregate high volumes of data from disparate sources, BI is naturally suited for use in tracking both financial and operational KPIs. In other words, BI has the capacity for comprehensive