

## Unit-5

### What is XML?

- XML stands for EXtensible Markup Language
- XML is a markup language much like HTML
- XML was designed to carry data, not to display data
- XML tags are not predefined. You must define your own tags
- XML is designed to be self-descriptive
- XML is a W3C Recommendation
- **XML: eXtensible Markup Language**, a standard created by the W3Group for marking up data.
- **DTD: Document Type Definition**, a set of rules defining relationships within a document; DTDs can be "internal" (within a document) or "external" (links to another document).
- **XML Parser:** Software that reads XML documents and interprets or "parse" the code according to the XML standard. A parser is needed to perform actions on XML, such as comparing an XML document to a DTD.
- XML can be used to exchange, share, and store data.
- XML documents form a tree structure that starts at "the root" and branches to "the leaves".
- XML has very simple syntax rules. XML with correct syntax is "Well Formed". Valid XML also validates against a DTD.
- XSLT is used to transform XML into other formats like HTML.

### Goals of XML:-

- XML shall be straightforwardly usable over the Internet.
- XML shall support a wide variety of applications.
- XML shall be compatible with SGML.  
**SGML** or Standard Generalized Markup Language, is the ISO standard on which all XML, hence XHTML documents are based upon. If a document is not compatible with SGML, then it cannot be called XML.
- It shall be easy to write programs that process XML documents.  
XML was always intended to be easy to use and process and since XML is based on human-readable text, this makes it a lot easier for developers to figure out what is meant by the XML tags.
- XML documents shall be easy to create.

### What is HTML?

- HTML (HyperText Markup Language) is a language for describing Web pages.
- HTML is not a programming language or protocol, it is a markup language
- A markup language is a set of markup tags
- HTML uses markup tags to describe web pages
- A browser, eg Internet Explorer or Firefox, can interpret the HTML and display a Web page in the desired fashion.
- HTML markup tags are usually called HTML tags

- HTML tags are keywords surrounded by angle brackets like <html>
- HTML tags normally come in matched pairs like <b> and </b>
- The first tag in a pair is the start tag, the second tag is the end tag
- Start and end tags are also called opening tags and closing tags
- For some special HTML tags, start tag and end tag can be combined, eg <br />

## HTML Documents = Web Pages

- HTML documents describe web pages
- HTML documents contain HTML tags and plain text
- HTML documents are also called web pages

Although many people have never heard of it, XHTML is really the future of the internet. It is the newest generation of HTML (coming after HTML 4) but has many new features which mean that it is, in some ways, like XML.

## HTML vs XML

There are many differences between HTML (Hyper Text Markup Language) and XML (eXtensible Markup Language). The important differences are given below:

No.	HTML	XML
1)	HTML is used <b>to display data</b> and focuses on how data looks.	XML is a software and hardware independent tool used <b>to transport and store data</b> . It focuses on what data is.
2)	HTML is a <b>markup language</b> itself.	XML provides a <b>framework to define markup languages</b> .
3)	HTML is <b>not case sensitive</b> .	XML is <b>case sensitive</b> .
4)	HTML is a presentation language.	XML is neither a presentation language nor a programming language.
5)	HTML <b>has its own predefined tags</b> .	You <b>can define tags according to your need</b> .
6)	In HTML, it is <b>not necessary to use a closing tag</b> .	XML <b>makes it mandatory to use a closing tag</b> .
7)	HTML is <b>static</b> because it is used to display data.	XML is <b>dynamic</b> because it is used to transport data.

8)	HTML does not preserve whitespaces.	XML preserve whitespaces.
----	-------------------------------------	---------------------------

## **BUSINESS IMPORTANCE IN USING XML:-**

XML has become the basis of virtually web-connected applications. It has greatly enhanced the way businesses share information and communicate with each other. XML is also an excellent way to store data. As XML allows sharing of information between different computer platforms and different applications, businesses are opting to convert a lot of their data into XML format. For instance, suppose company A is selling something to company B. Company A would need to send a purchase order to company B. For that, both companies require a common formatting practice. XML provides the language of both the descriptions of that formatting practice or convention, and offers a suitable way to send the purchase order data. Here are the important reasons why businesses use this globally accepted web format.

- XML is one of the most cost-effective platforms for publishing documents on the Internet
- Easy retrieval and archival of documents
- XML has flexibility, scalability and versatility
- It is an unique cross platform web publishing format
- XML can represent complex data structures
- It is used to interchange information
- XML allows automated web publishing
- Optimum security of data
- Hassle-free retrieval and archiving

XML documents can be used and shared in different ways. You can use them to send survey forms or online applications to companies, and also to send output to print on publishing systems. Large files and non-printable characters are handled well with XML. XML conversion frees up a lot of data storage space and its high versatility and good compatibility across major platforms allows data to be shared easily. As data is arranged systematically and in a more orderly way in XML format, it allows greater scalability and flexibility.

If your documents are in PDF, HTML, Text, Excel or Word format, you can convert your documents to XML format by outsourcing the task to a professional document conversion company. Through XML conversion your documents will be well formed and can be used various platforms.

## **Advantages of XML**

1. XML is platform independent and programming language independent, thus it can be used on any system and supports the technology change when that happens.
2. XML supports unicode. Unicode is an international encoding standard for use with different languages and scripts, by which each letter, digit, or symbol is assigned a unique numeric value that applies across different platforms and programs. This feature allows XML to transmit any information written in any human language.

3. The data stored and transported using XML can be changed at any point of time without affecting the data presentation. Generally other markup language such as HTML is used for data presentation, HTML gets the data from XML and display it on the GUI (graphical user interface), once data is updated in XML, it does reflect in HTML without making any change in HTML GUI.
4. XML allows validation using DTD and Schema. This validation ensures that the XML document is free from any syntax error.
5. XML simplifies data sharing between various systems because of its platform independent nature. XML data doesn't require any conversion when transferred between different systems.

### **Disadvantages of XML**

1. XML syntax is verbose and redundant compared to other text-based data transmission formats such as JSON.
2. The redundancy in syntax of XML causes higher storage and transportation cost when the volume of data is large.
3. XML document is less readable compared to other text-based data transmission formats such as JSON.
4. XML doesn't support array.
5. XML file sizes are usually very large due to its verbose nature, it is totally dependant on who is writing it.

### **Structure of XML document:-**

```
<?xml version = "1.0"?>
<contact-info>
<name>Tanmay Patil</name>
<company>TutorialsPoint</company>
<phone>(011) 123-4567</phone>
</contact-info>
```

### **What is XHTML?**

XHTML stands for eXtensible HyperText Markup Language and is a cross between HTML and XML. XHTML was created for two main reasons:

1. To create a stricter standard for making web pages, reducing incompatibilities between browsers
2. To create a standard that can be used on a variety of different devices without changes

The great thing about XHTML, though, is that it is almost the same as HTML, although it is much more important that you create your code correctly. You cannot make badly formed code to be XHTML compatible. Unlike with HTML (where simple errors (like missing out a closing tag) are ignored by the browser), XHTML code must be exactly how it is specified to be. This is due to the fact that browsers in handheld devices etc. don't have the power to show badly formatted pages so XHTML makes sure that the code is correct so that it can be used on any type of browser.

XHTML is a web standard which has been agreed by the W3C and, as it is backwards compatible, you can start using it in your web pages now. Also, even if you don't think it's really necessary to update to XHTML yet, there are three very good reasons to do so:

1. It will help you to create better formatted code on your site
2. It will make your site more accessible (both in the future and now due to the fact that it will also mean you have correct HTML and most browsers will show your page better)
3. XHTML is planned to replace HTML 4 in the future

There is really no excuse not to start writing your web pages using XHTML as it is so easy to pick up and will bring many benefits to your site.

## The Main Changes

There are several main changes in XHTML from HTML:

- All tags must be in lower case
- All documents must have a doctype
- All documents must be properly formed
- All tags must be closed
- All attributes must be added properly
- The name attribute has changed
- Attributes cannot be shortened
- All tags must be properly nested

At a glance, this seems like a huge amount of changes but once you start checking though the list you will find that very little on your site actually needs to be changed. In this tutorial I will go through each of these changes explaining exactly what is different.

## Benefits of XHTML:-

- **Compatibility** XHTML is both forward compatible with the structured world of XML and backward compatible with older HTML-based browsers.
- **Interoperability** XHTML's logical structure allows easier repurposing of content for different platforms.
- **Extensibility** XHTML documents can be extended to add new tags with namespaces. Or, because XHTML is XML, authors can add new tags without namespaces. These tags will be ignored by the browser, but they can be used by XML tools for exchanging data.
- **XML Conforming** XHTML documents have a reliable structure and conform to the rules of XML. XML processors can reliably manipulate and validate these conforming cargos of content for easier document and data exchange.

## What is a Data Warehouse?

A Data Warehouse (DW) is a relational database that is designed for query and analysis rather than transaction processing. It includes historical data derived from transaction data from single and multiple sources.

A Data Warehouse provides integrated, enterprise-wide, historical data and focuses on providing support for decision-makers for data modeling and analysis.

A Data Warehouse is a group of data specific to the entire organization, not only to a particular group of users.

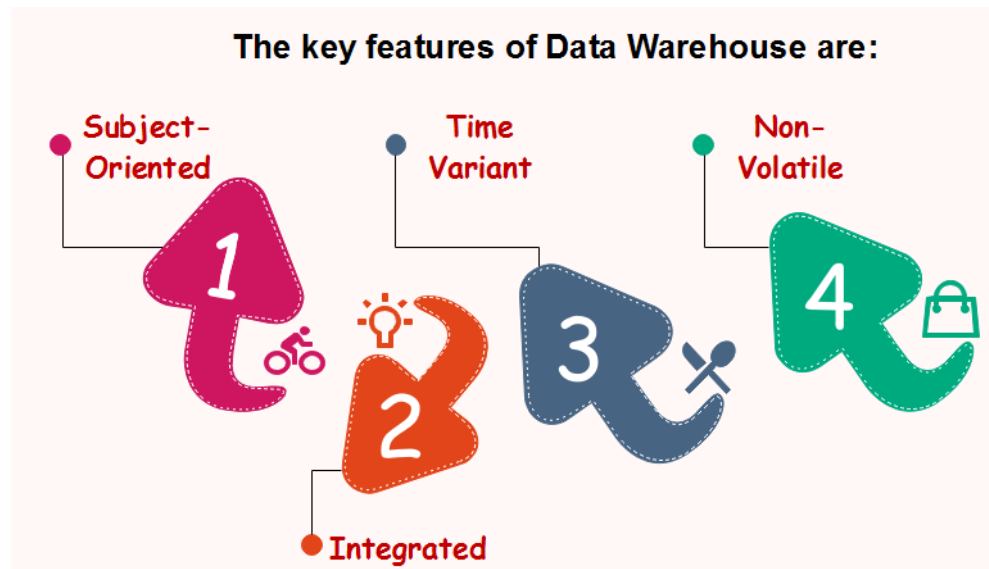
It is not used for daily operations and transaction processing but used for making decisions.

A Data Warehouse can be viewed as a data system with the following attributes:

- It is a database designed for investigative tasks, using data from various applications.
- It supports a relatively small number of clients with relatively long interactions.
- It includes current and historical data to provide a historical perspective of information.
- Its usage is read-intensive.
- It contains a few large tables.

"Data Warehouse is a subject-oriented, integrated, and time-variant store of information in support of management's decisions."

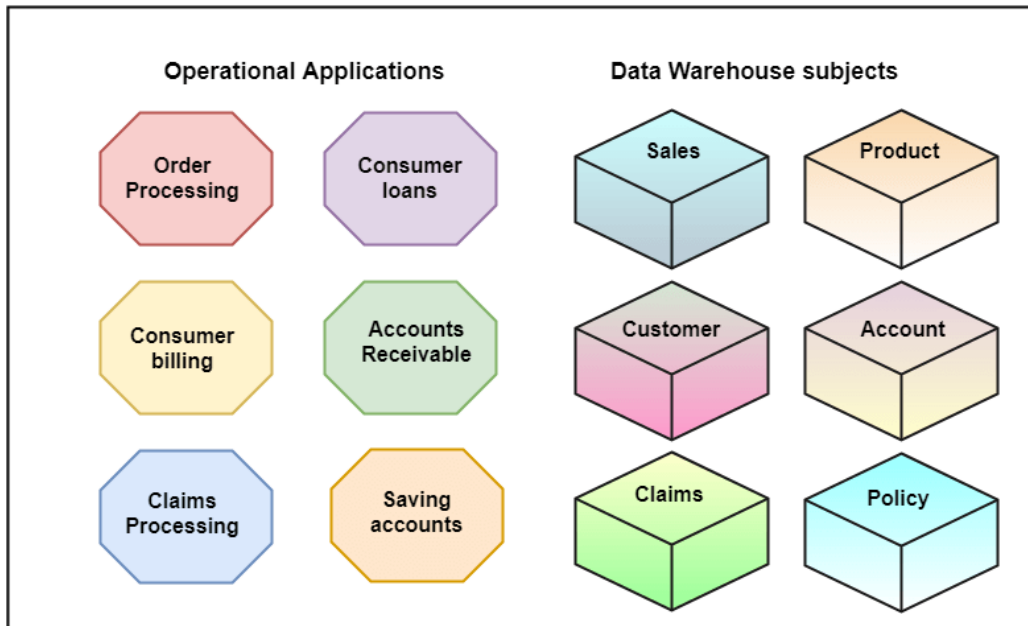
Characteristics of Data Warehouse



## Subject-Oriented

A data warehouse target on the modeling and analysis of data for decision-makers. Therefore, data warehouses typically provide a concise and straightforward view around a particular subject, such as customer, product, or sales, instead of the global organization's ongoing operations. This is done by excluding data that are not useful concerning the subject and including all data needed by the users to understand the subject.

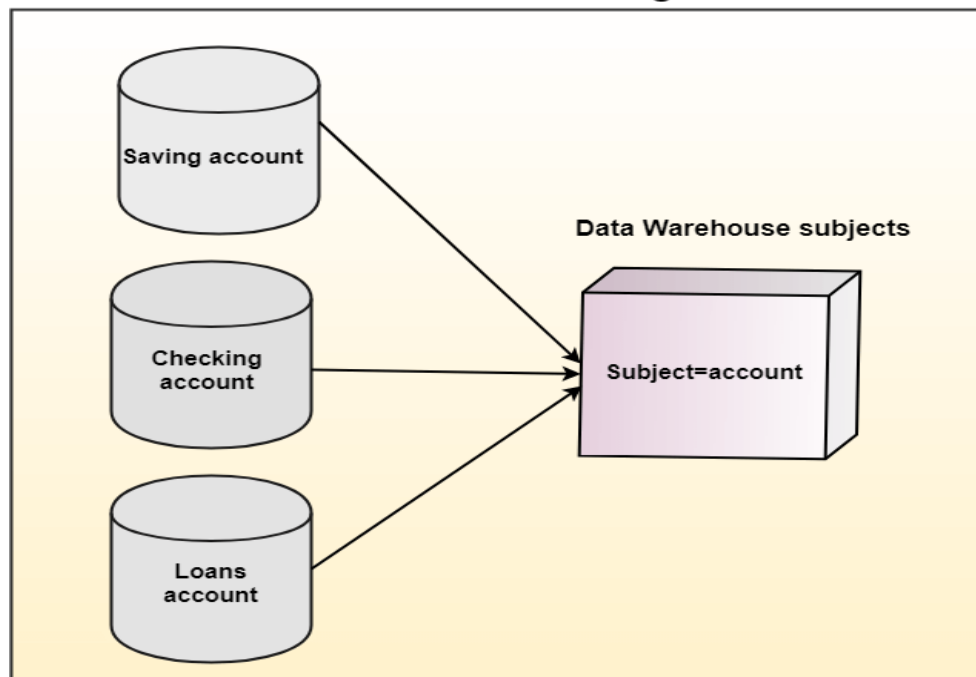
## Data Warehouse is Subject-Oriented



## Integrated

A data warehouse integrates various heterogeneous data sources like RDBMS, flat files, and online transaction records. It requires performing data cleaning and integration during data warehousing to ensure consistency in naming conventions, attributes types, etc., among different data sources.

## Data Warehouse is Integrated



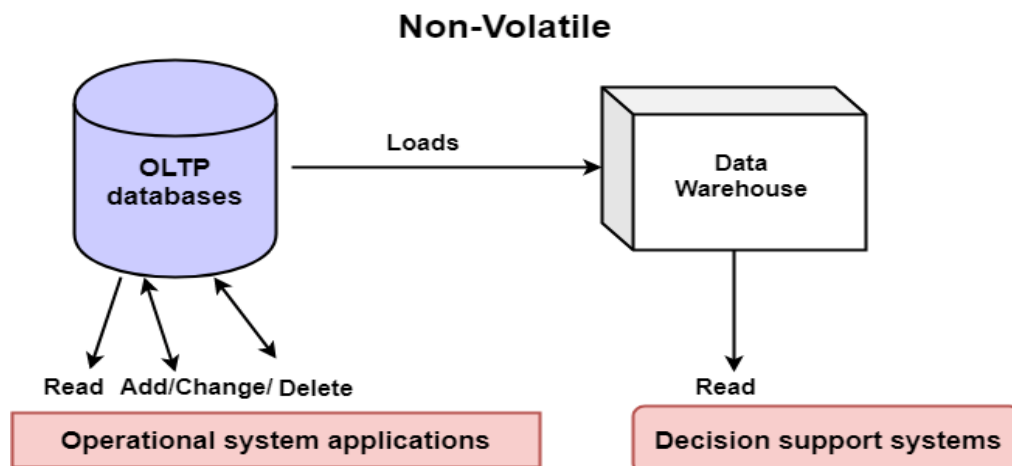
## Time-Variant

Historical information is kept in a data warehouse. For example, one can retrieve files from 3 months, 6 months, 12 months, or even previous data from a data warehouse. These variations with a transactions system, where often only the most current file is kept.



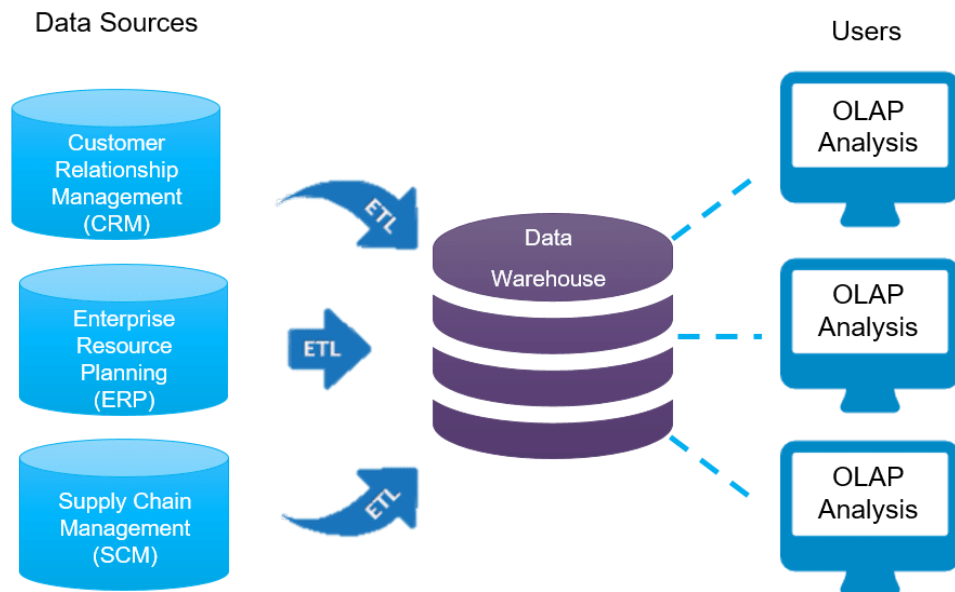
## Non-Volatile

The data warehouse is a physically separate data storage, which is transformed from the source operational RDBMS. The operational updates of data do not occur in the data warehouse, i.e., update, insert, and delete operations are not performed. It usually requires only two procedures in data accessing: Initial loading of data and access to data. Therefore, the DW does not require transaction processing, recovery, and concurrency capabilities, which allows for substantial speedup of data retrieval. Non-Volatile defines that once entered into the warehouse, and data should not change.





- Functions of Data Warehouse Tools and Utilities:-
- **Data Extraction** – Involves gathering data from multiple heterogeneous sources.
- **Data Cleaning** – Involves finding and correcting the errors in data.
- **Data Transformation** – Involves converting the data from legacy format to warehouse format.
- **Data Loading** – Involves sorting, summarizing, consolidating, checking integrity, and building indices and partitions.
- **Refreshing** – Involves updating from data sources to warehouse.



## Data mart

A **Data Mart** is focused on a single functional area of an organization and contains a subset of data stored in a Data Warehouse. A Data Mart is a condensed version of Data Warehouse and is designed for use by a specific department, unit or set of users in an organization. E.g., Marketing, Sales, HR or finance. It is often controlled by a single department in an organization.

Data Mart usually draws data from only a few sources compared to a Data warehouse. Data marts are small in size and are more flexible compared to a Data warehouse.

### Why do we need Data Mart?

- Data Mart helps to enhance user's response time due to reduction in volume of data
- It provides easy access to frequently requested data.
- Data marts are simpler to implement when compared to corporate Datawarehouse. At the same time, the cost of implementing Data Mart is certainly lower compared with implementing a full data warehouse.

- Compared to Data Warehouse, a datamart is agile. In case of change in model, datamart can be built quicker due to a smaller size.
- A Datamart is defined by a single Subject Matter Expert. On the contrary data warehouse is defined by interdisciplinary SME from a variety of domains. Hence, Data mart is more open to change compared to Datawarehouse.
- Data is partitioned and allows very granular access control privileges.
- Data can be segmented and stored on different hardware/software platforms.

## Types of Data Mart

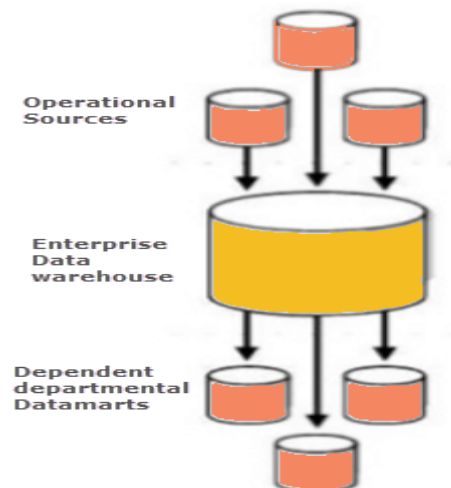
There are three main types of data mart:

1. **Dependent:** Dependent data marts are created by drawing data directly from operational, external or both sources.
2. **Independent:** Independent data mart is created without the use of a central data warehouse.
3. **Hybrid:** This type of data marts can take data from data warehouses or operational systems.

## Dependent Data Mart

A dependent data mart allows sourcing organization's data from a single Data Warehouse. It is one of the data mart example which offers the benefit of centralization. If you need to develop one or more physical data marts, then you need to configure them as dependent data marts.

Dependent Data Mart in data warehouse can be built in two different ways. Either where a user can access both the data mart and data warehouse, depending on need, or where access is limited only to the data mart. The second approach is not optimal as it produces sometimes referred to as a data junkyard. In the data junkyard, all data begins with a common source, but they are scrapped, and mostly junked.

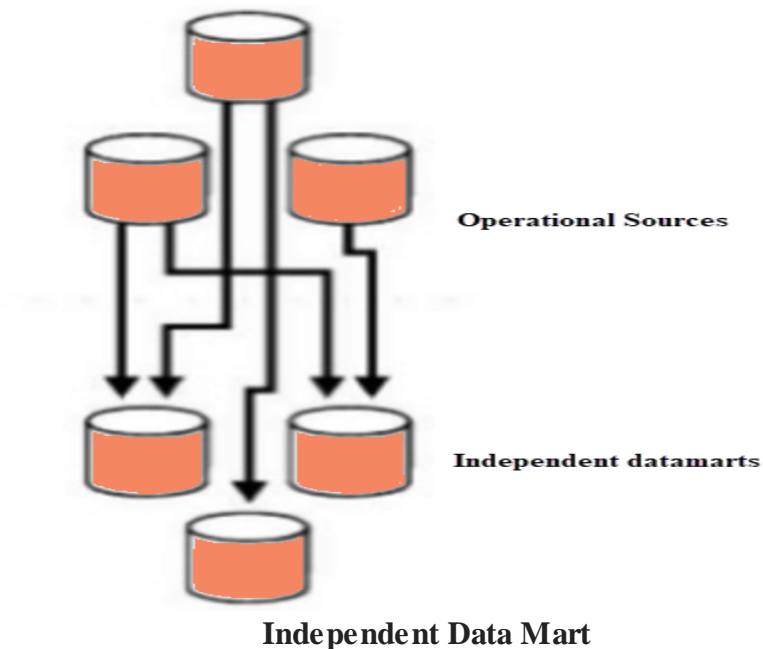


## Independent Data Mart

An independent data mart is created without the use of central Data warehouse. This kind of Data Mart is an ideal option for smaller groups within an organization.

An independent data mart has neither a relationship with the enterprise data warehouse nor with any other data mart. In Independent data mart, the data is input separately, and its analyses are also performed autonomously.

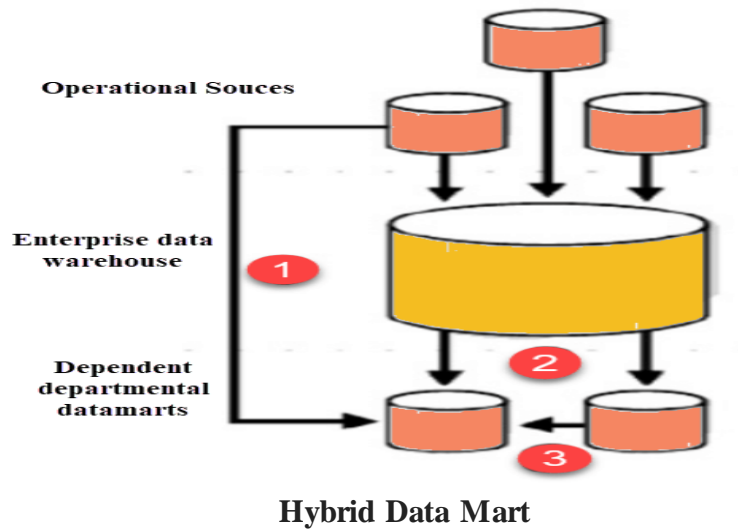
Implementation of independent data marts is antithetical to the motivation for building a data warehouse. First of all, you need a consistent, centralized store of enterprise data which can be analyzed by multiple users with different interests who want widely varying information.



## Hybrid Data Mart

A hybrid data mart combines input from sources apart from Data warehouse. This could be helpful when you want ad-hoc integration, like after a new group or product is added to the organization.

It is the best data mart example suited for multiple database environments and fast implementation turnaround for any organization. It also requires least data cleansing effort. Hybrid Data mart also supports large storage structures, and it is best suited for flexible for smaller data-centric applications.



- **Operational Data stores:-**An operational data store (ODS) is a central database that provides a snapshot of the latest data from multiple transactional systems for operational reporting. It enables organizations to combine data in its original format from various sources into a single destination to make it available for business reporting.
- **Operational data store benefits:-**
  - The ODS provides a consolidated repository into which previously isolated or inefficiently communicating IT systems can feed.
  - ODS reporting, which is focused on a snapshot of operational data, can be more sophisticated than reports from individual underlying systems. The ODS is architected to provide a consolidated view of data integrated from multiple systems, so reports can provide a holistic perspective on operational processes.
  - The up-to-date view into operational status also makes it easier for users to diagnose problems before digging into component systems. For example, an ODS enables service representatives to immediately find a customer order, its status, and any troubleshooting information that might be helpful.
  - An ODS contains critical, time-sensitive business rules, such as those automatically notifying a financial institution when a customer has overdrawn an account. These rules, in aggregate, are a kind of process automation that greatly improves efficiency, which would be impossible without current and integrated operational data.

## E-Marketing

E-Marketing (electronic marketing) refers to the marketing conducted over the Internet. Two synonyms of E-Marketing are Internet Marketing and online marketing which are frequently interchanged. E-Marketing is the process of marketing a brand (company, product, or service) using the Internet through computers and mobile devices mediums. By such a definition, eMarketing encompasses all the activities a business conducts via the worldwide web with the aim of attracting new business, retaining current business and developing its brand identity.

## **Advantages of e-Marketing**

- Ability to target your customers faster and cheaper
- Reduction of marketing costs through automation of electronic media
- Near real-time interaction between the marketer and the end user
- Ability to quantify and collect user data
- One-to-one marketing experience
- Increased interactivity
- Ability to implement marketing strategies in a short time-frame
- Ability to scale with the market
- Appeal to specific interests
- Geo-targeting

## **Types of E-Marketing**

- Search engine optimization (SEO)
- Paid advertising (PPC)
- Email marketing
- Social media marketing
- Mobile marketing
- Online Marketing:- Online marketing is a set of tools and methodologies used for promoting products and services through the internet. Online marketing includes a wider range of marketing elements than traditional business marketing due to the extra channels and marketing mechanisms available on the internet.
- Online marketing can deliver benefits such as:
  - Growth in potential
  - Reduced expenses
  - Elegant communications
  - Better control
  - Improved customer service
  - Competitive advantage
- Online marketing is also known as internet marketing, web marketing, or digital marketing. It includes several branches such as social media marketing (SMM), search engine optimization (SEO), pay-per-click advertising (PPC), and search engine marketing (SEM).

## **Importance of E-Marketing**

The return on investment (ROI) from E-Marketing can far exceed that of traditional marketing strategies. Also, the transparency of the internet allows the marketer to have access to analytics and data in a near real-time fashion which will allow the marketer to make changes to align with the market's reaction thus making E-Marketing a preferred solution for Marketing Professionals.

- Meeting the Needs of Website Visitors :-
- **Successful Web businesses:**  
Realize every visitor is a potential customer (partner)
- **Crafting Web presence** is an important concern  
Know visitor characteristic variations
- Visitor at site for a reason
- Web site visitor motivations

Learning about company products or services

- Buying products or services
- Obtaining warranty, service, repair policy information
- Obtaining general company information
- Obtaining financial information
- Identifying people
- Obtaining contact information

**Website maintenance**:-is the act of regularly checking your **website** for issues and mistakes and keeping it updated and relevant. This should be done on a consistent basis in order to keep your **website** healthy, encourage continued traffic growth, and strengthen your SEO and Google rankings.

- Choose the right website builder
- Update Regularly
- Review your content
- Review your content
- Use Social Media

**Advantages of Digital Marketing:-**

- Global Reaching
- Lower Cost
- Measurable Results
- Personalization