There are certain reasons for DWH projects failure which are as follows:

a Improper planning

b Inadequate project management

c Company not ready for a data warehouse

d Insufficient staff training

e Improper team management

f No support from top management

So, Planning for Data ware house is necessary which can be done as follows:

I. Key issues needs to be planned

      1.Value and expectation

      2.Risk assessment

      3.Top-down or bottom –up

      4.Build or Buy

      5.Single vender or best of breed

II. Business requirement ,not technology

III. Top management support

IV. Justification

The challenges which are there for Data Warehousing Project Management are as follows:

➢ DATA ACQUISITION

• Large number of sources

• Many disparate sources

• Different computing platforms

• Outside sources

• Huge initial load

• Ongoing data feeds

• Data replication considerations

➢ DATA STORAGE

- Storage of large data volumes

- Rapid growth

- Need for parallel processing

- Data storage in staging area

- Multiple index types

- Several index files

- Storage of newer data types

- Archival of old data

- Compatibility with tools

- RDBMS & MDDBMS

➤ INFO. DELIVERY

- Several user types

- Queries stretched to limits

- Multiple query types

- Web-enabled

- Multidimensional analysis

- OLAP functionality

- Metadata management

- Interfaces to DSS apps.

- Feed into Data Mining

- Multi-vendor tools

There are certain considerations in Project Management which are as follows:
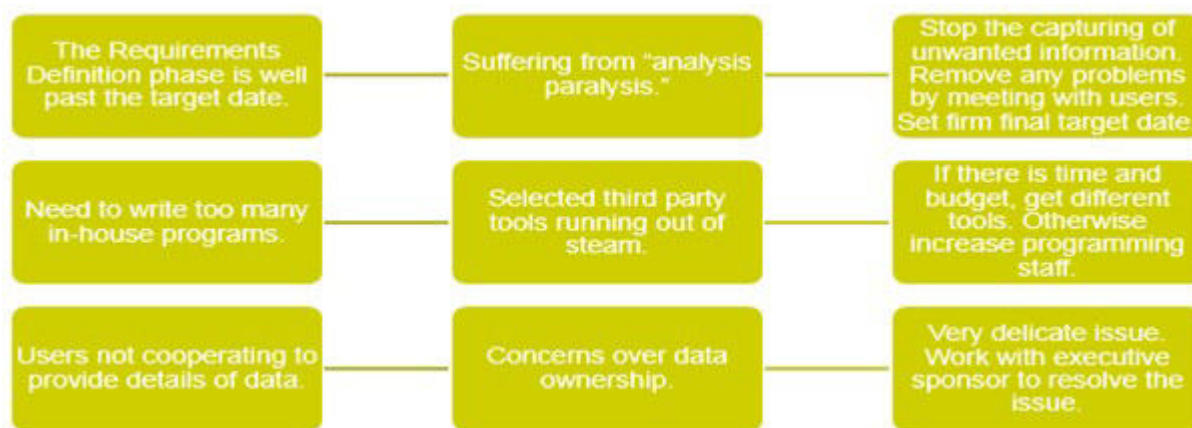
- The effort of data warehouse project has been successful if there is critical effective project management.

- Project management issues are applied to build success data warehouse projects :

o Project management principles Some of the guiding principles that pertain to data warehouse projects exclusively are as follows:

o Sponsorship , New Paradigm , Data Quality , Building for Growth , Project Politics , Dimensional Data Modeling , Project Manager, Team Roles, User Requirements, Training, Realistic Expectations, External Data

o Warning signs

The warning sign indication action is explained below:

The Requirements Definition phase is well past the target date.	Suffering from "analysis paralysis."	Stop the capturing of unwanted information. Remove any problems by meeting with users. Set firm final target date.
Need to write too many in-house programs.	Selected third party tools running out of steam.	If there is time and budget, get different tools. Otherwise increase programming staff.
Users not cooperating to provide details of data.	Concerns over data ownership.	Very delicate issue. Work with executive sponsor to resolve the issue.

o Success factors

The success factors in project management are explained below:

Queries and reports	• rapid increase in the number of queries and reports requested by the users directly from the data warehouse
Query types	• queries becoming more sophisticated
Active users	• steady increase in the number of users
Usage	• users spending more and more time in the data warehouse looking for solutions
Turnaround times	• marked decrease in the times required for obtaining strategic information

o Adopting a practical approach

Ralph Kimbal is one of the most renowned architects in the world of data warehousing technologies. His guidelines to design a data warehouse, often referred as Kimball methodology, has become one of the most followed methods by technical experts in building a decision support system across the organization.

Kimball methodology is intended for for designing, developing, and deploying data warehouse/business intelligence systems, as described in The Data Warehouse Lifecycle Toolkit. There are other names for the Kimball approach that we will be discussion shortly.

•Bottom-up approach for data warehousing

•Kimball's dimensional modelling

•Data warehouse lifecycle model by Kimball

Summary of Kimball methodology

•Starts with one data mart (ex. sales); later on additional data marts are added (ex. collection, marketing, etc.)

•Data flows from source into data marts, then into the data warehouse

•Kimball approach is faster to implement as it is implemented in stages
Before we go ahead with details of the methodology, let us take a quick view on some essential definitions of the terms used.

Data mart

Data mart is a specific, subject-oriented repository of data that was designed to answer specific questions. Usually, multiple data marts exist to serve the needs of multiple business units (sales, marketing, operations, collections, accounting, etc.)

Data warehouse

Data warehouse is a single organizational repository of enterprise wide data across many or all subject areas. Data warehouse is an enterprise wide collection of data marts. Data warehouse is the foundation for business intelligence.

Business Intelligence refers to reporting and analysis of data stored in the warehouse.

The Kimball Lifecycle Diagram
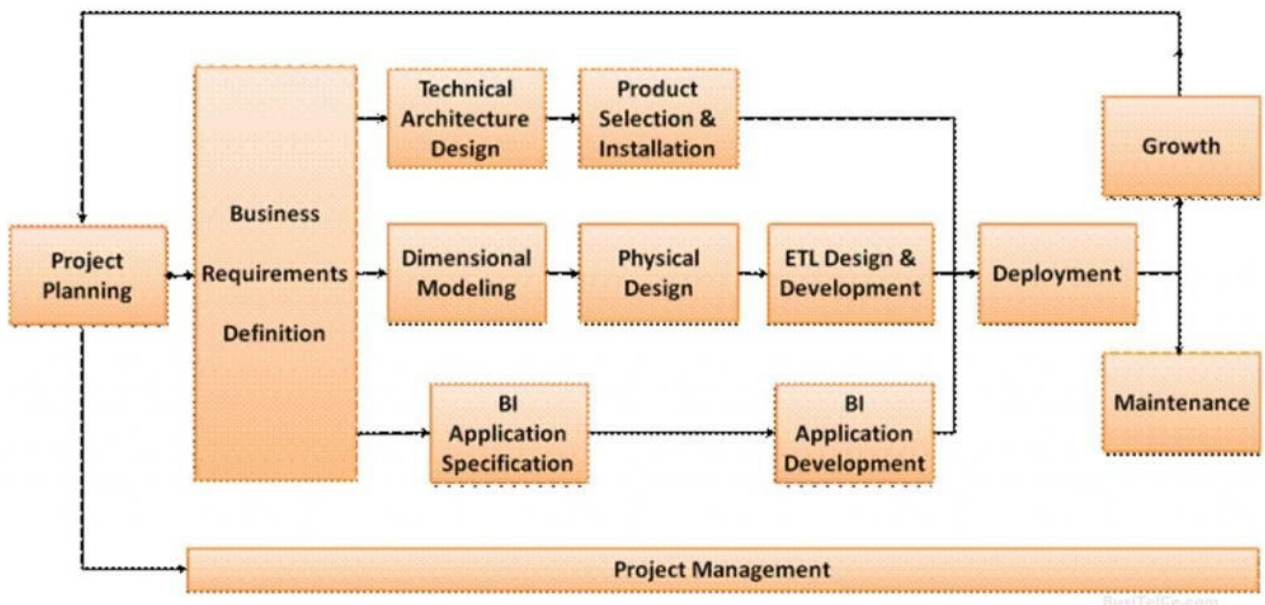
Kimball lifecycle diagram illustrates the flow of data warehouse implementation. It identifies task sequencing and highlights activities that should happen concurrently. Activities may need to be customized to address the unique needs of the organization. Also,not every detail of every lifecycle task will be required on every project – this has to be decided as per need.

As per Kimball Lifecycle, we start building a data warehouse with understanding business requirements and determining how best to add value to the organization. The organization must agree on what the value of this data is before deciding to build a data warehouse to hold it. Once the requirements are gathered, implementation phase begins with design steps across three different tracks – technology, data, and BI applications. Once we are done with this implementation, the Lifecycle comes back together to deploy the query tools, reports, and applications to the user community.

The incremental approach of the Lifecycle helps to deliver business value in a short span of time and at the same time helps to build a enterprise wide information resource in a long term.

Lets get little deeper into the Kimball guideline as per the Lifecycle diagram.

Business Requirements Definition

Success of the project depends on a solid understanding of the business requirements. Understanding the key factors driving the business is crucial for successful translation of the business requirements into design considerations. There are many methods and tools to ensure requirement gathering is best done – we will not dwell more into it now.

What follows business requirement definition is three concurrent tracks focusing on

•Technology

•Data

•Business intelligence applications

Technical Architecture Design

Objective here is to finalize overall architectural framework and vision. We do this based on consider business requirements, current technical environment, and planned strategic technical directions of the organization. Based on the technical architecture, we do:

•Evaluation and selection of
Products that will deliver needed capabilities
Hardware platform
Database management system
Extract-transformation-load (ETL) tools
Data access query tools
Reporting tools must be evaluated

•Installation of selected products/components/tools

•Testing of installed products to ensure appropriate end-to-end integration within the data warehouse environment.

Data track – Dimensional modeling

Data track primaraly deals with design of the dimensional model. Dimensional modeling is a vast subject area comprising of many methods, suggestions, and best practices. Here, a detailed data analysis of a single business process is performed to identify the fact table granularity, associated dimensions and attributes, and numeric facts. Primary constructs of a

dimensional model are fact tables and dimension tables. Two important design menthods in dimension modeling are – star schema and snowflake schema.

Star schema contains a central fac table directly connected to dimension table. The snowflake schema is a variation of the star schema used in a data warehouse. The snowflake schema is a more complex schema than the star schema because the tables which describe the dimensions are normalized

Physical design

We start with a logical data model definition and then proceed to physical design of the model. In physical design, we define the data structures. Some key activities include:

•setting up the database environment

•setting up appropriate security

•preliminary performance tuning strategies, from indexing to partitioning and aggregations.

•if required, OLAP databases are also designed during this process.

ETL design and development

This is by far the most important stage of the lifecycle and it take 70% of the effort to achieve. In this stage, we perform extraction, transformation, and loading (ETL) of source data into the target models. Raw data is extracted from the operational source systems and is being transformed into meaningful information for the business. Data quality conditions are continuously monitored.

Kimball calls ETL a "data warehouse back room". ETL system strives to deliver high throughput, as well as high quality output.

BI Application track

BI applications deliver business value from the DW/BI solution, rather than just delivering the data. The goal is to deliver capabilities that are accepted by the business to support and enhance their decision making.

First, we start to identify the candidate BI applications and appropriate navigation interfaces to address the users' needs and needed capabilities. This will give desired specification of the tool required. Later, we configure the business metadata and tool infrastructure. This is

followed by construction and validation of the specified analytic and operational BI applications and the navigational portal.

Deployment

Deployment should be deferred until all the pieces such as training, documentation, and validated data are not ready for production release. Also, it is critical that deployment be well orchestrated and adequately planned.

Deployment should ensure the results of technology, data, and BI application tracks are tested and fit together properly. While deploying the solution, appropriate education and support infrastructure has to be in place.

Maintenance

Maintenance begins once the system is deployed into production. Maintenance work ensures ongoing support, education, and communication with business users. Also, technical operational tasks that are necessary to keep the system performing optimally are conducted as needed. Usage monitoring, performance tuning, index maintenance, and system backup are done periodically by technical experts.

Growth process

Organization has to reason to be happy if the data warehouse system tends to grow. DW growth is considered as a symbol of success. There may be new requests which requires attention. The key here is to build upon the foundation that has already been established.