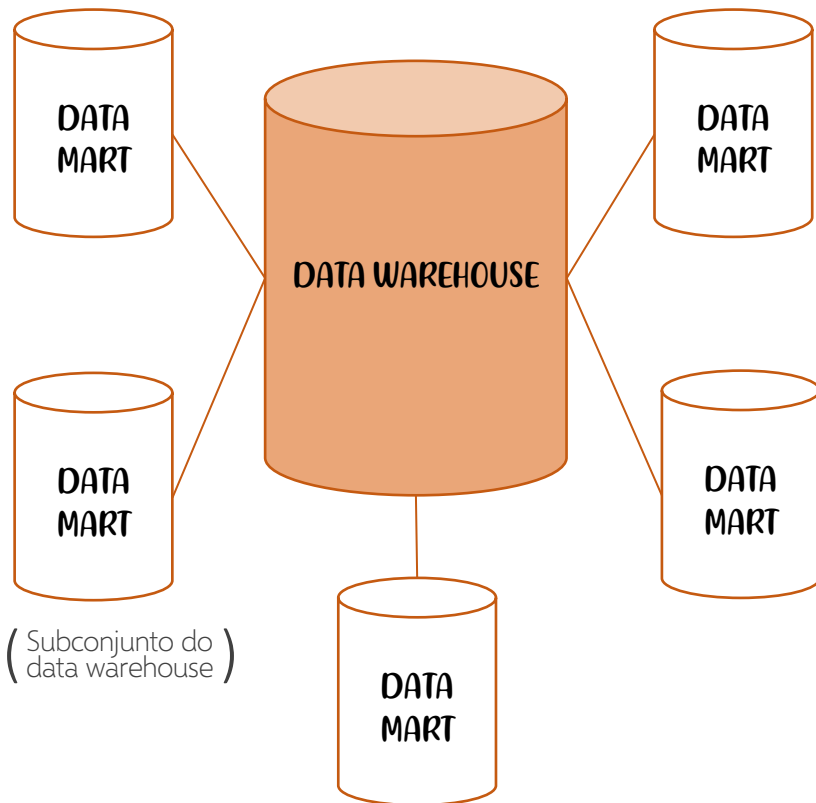


data warehouse

ASPECTOS GERAIS

- = Armazém de **dados** (Maior que um banco de dados)
- Para dar suporte ao processo de tomada de **decisão**
- Especializado em **gerar relatórios**
É eficiente para esse tipo de processamento



CARACTERÍSTICAS ESSENCIAIS

CAI MUITO!

ORIENTADO POR ASSUNTO

Dados organizados conforme temas específicos

INTEGRADO

Dados de múltiplas fontes são armazenados de forma consistente, padronizada e uniforme.

Inconsistências são removidas e conflitos de informação consertados

NÃO VOLÁTIL

Os dados, uma vez inseridos no D.W., não mudam (mas podem ser excluídos).

Ficam disponíveis para que os usuários façam consultas e gerem relatórios

VARIÁVEL NO TEMPO (Histórico)

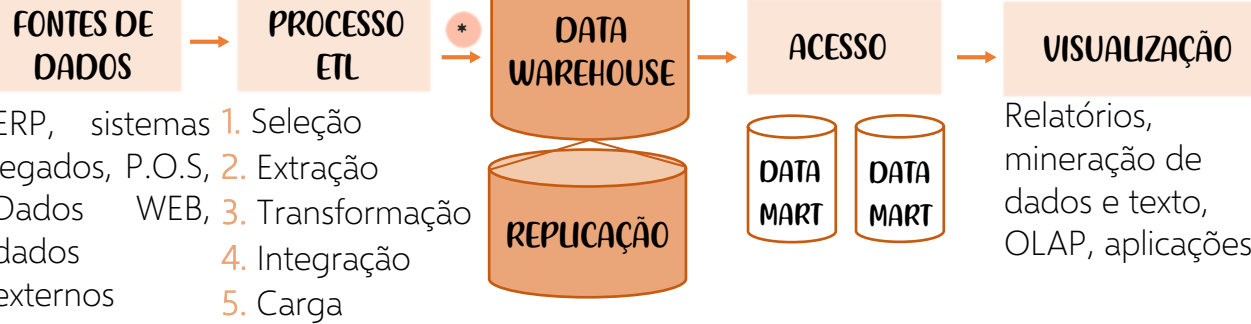
Dados históricos são mantidos no D.W. (antigos)

Diferente do banco de dados transacional (Geralmente mantém só os dados mais recentes)

TIPOS

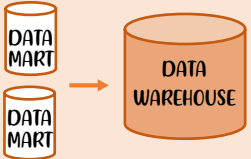
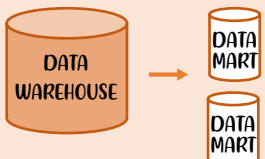
- **EDW:** (Enterprise Data Warehouse)
 - Repositório central
 - Tático/ Estratégico
- **ODS:** (Operational Data Store)
 - Repositório intermediário
 - Utilizado para relatórios operacionais + fonte de dados para o EDW
- **DM:** (Data Mart)
 - Subconjunto de dados do D.W. (Normalmente referentes a um assunto específico)
 - Dá suporte à decisão para uma específica unidade do negócio

PROCESSO



* **Staging área:** localização temporária para onde os dados dos sistemas de origem são copiados

TEÓRICOS

CRITÉRIO	KIMBALL	INMON
ARQUITETURA	<i>Dimensional Data Warehouse</i>	<i>Enterprise Data Warehouse</i>
ABORDAGEM	<i>Bottom Up</i> De baixo para cima 	<i>Top Down</i> De cima para baixo 
INVESTIMENTO INICIAL	Baixo	Alto
NORMALIZAÇÃO	Desnormalizado	Normalizado (3ª forma normal)
DIRECIONAMENTO	Usuários Finais	Profissionais de T.I.
CONSULTA	<i>No data Warehouse</i>	<i>No data mart</i>

data warehouse

ETL (*Extract, transform, load*)



- É o processo **mais crítico** e demorado da construção do D.W.
- = Extração de dados de fontes homogêneas ou heterogêneas
 - + Transformação/ limpeza de dados
 - + Carga de dados no D.W.
- Os dados devem ser **corretamente trabalhados**, se não, sua informações podem prejudicar a tomada de decisão