Resumo Aula 4

ETL(Extract Transform Load)

O processo ETL começou a se consolidar nos anos 70 e 80, juntamente dos primeiros sistemas de Data Warehouse. Devido a necessidade de integrar os dados que se espalhavam pelos databases cresceu rapidamente e fez surgir a necessidade de ferramentas e processos para extrair dados de diferentes fontes.

1. Extração
   1. Essa fase consiste em definir as fontes de dados e selecioná-los para o DW, de forma a manter os dados no mesmo local.
   2. Serão apenas dados que estão de acordo com os objetivos do DW.
2. Transformação
   1. Esta fase consiste em limpar e alterar os dados obtidos da fonte, de forma a padronizá-los.
   2. A limpeza dos dados está ligada às inconsistências existentes na base de dados.
3. Carregamento(Load)
   1. Feitas as transformações, os dados são carregados. Esta é a fase que consiste em integrar/carregar os dados do DW.

O processo ETL é essencial para a criação das tabelas Fato e Dimensões no ambiente do DW.