



TALEND DATA INTEGRATION

Grimaldo Oliveira



Sobre o instrutor

- Grimaldo Oliveira
 - grimaldo_lopes@hotmail.com
- Formação
 - Mestre em Tecnologias Aplicadas a Educação Universidade do Estado da Bahia.
 - Especialização em Análise de Sistemas pela Faculdade Visconde de Cairu.
 - Estatístico pela Universidade Federal da Bahia.
- Atividades
 - Mais de 10 anos atuando como Consultor de Business Intelligence.
 - Projetos Governos Maranhão, Mato Grosso e Bahia.
 - Idealizador do Blog : BI com Vatapá – bicomvatapa.blogspot.com.
 - Livro: BI COMO DEVE SER – bicomodeveser.com.br





Desenvolvimento

- Objetivo
 - Aprender de forma prática os principais comandos e funções do TALEND DATA INTEGRATION.
- Material de Consulta
 - Slides do curso.
 - Apostila com exercícios resolvidos.
 - Apostila com exercícios desafio.
 - Sites indicados.



Roteiro

- O QUE É BUSINESS INTELLIGENCE
- O QUE É DATA WAREHOUSE
- CONHECENDO O TALEND DATA INTEGRATION
- INSTALAÇÃO O TALEND DATA INTEGRATION
- ESTRUTURA DE FUNCIONAMENTO DO TALEND DATA INTEGRATION
- CONSTRUÇÃO DO DW
 - CARGA DA STAGING
 - TRANSFORMAÇÃO DOS DADOS
 - CARGA TABELA STAGING FATO
 - CARGA DA DIMENSÃO
 - CARGA DA DIMENSÃO TEMPO
 - CARGA DA FATO



Roteiro

- AUTOMATIZAÇÃO DAS CARGAS
- PRINCIPAIS COMPONENTES DO TALEND DATA INTEGRATION
 - TFILEARCHIVE
 - TFILEUNARCHIVE
 - TFILECOMPARE E TFILEOUTPUTDELIMETED
 - TFILECOPY E TFILELIST
 - TFILEDELETE E TJAVA
 - TFILEEXIST ETMSGBOX
 - TFILEROWCOUNT
 - TAGGREGATEROW, TSORTROW E TLOGROW
 - TREPLACE E TFILTERCOLUMN
 - TFILTERROW
 - TSPLITROW



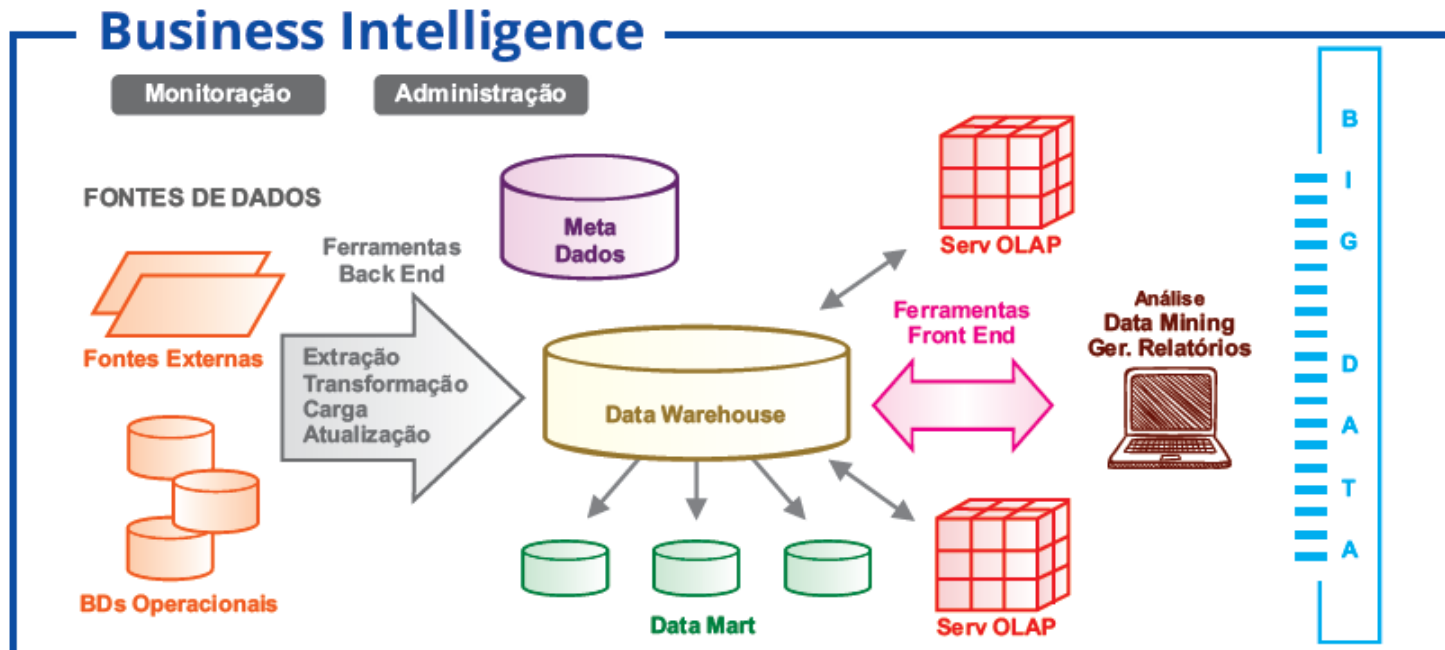
Roteiro

- PRINCIPAIS COMPONENTES DO TALEND DATA INTEGRATION
 - TFILEINPUTXML
 - TLOGCATCHER , TROWGENERATION,TDIE E TLOGROW
 - TLOGCATCHER , TROWGENERATION,TWARN E TLOGROW
 - TSENDMAIL
- VARIÁVEL DE CONTEXTO
- EXPORTAR JOB
- EXPORTAR PROJETO
- IMPORTAR PROJETO



Sobre

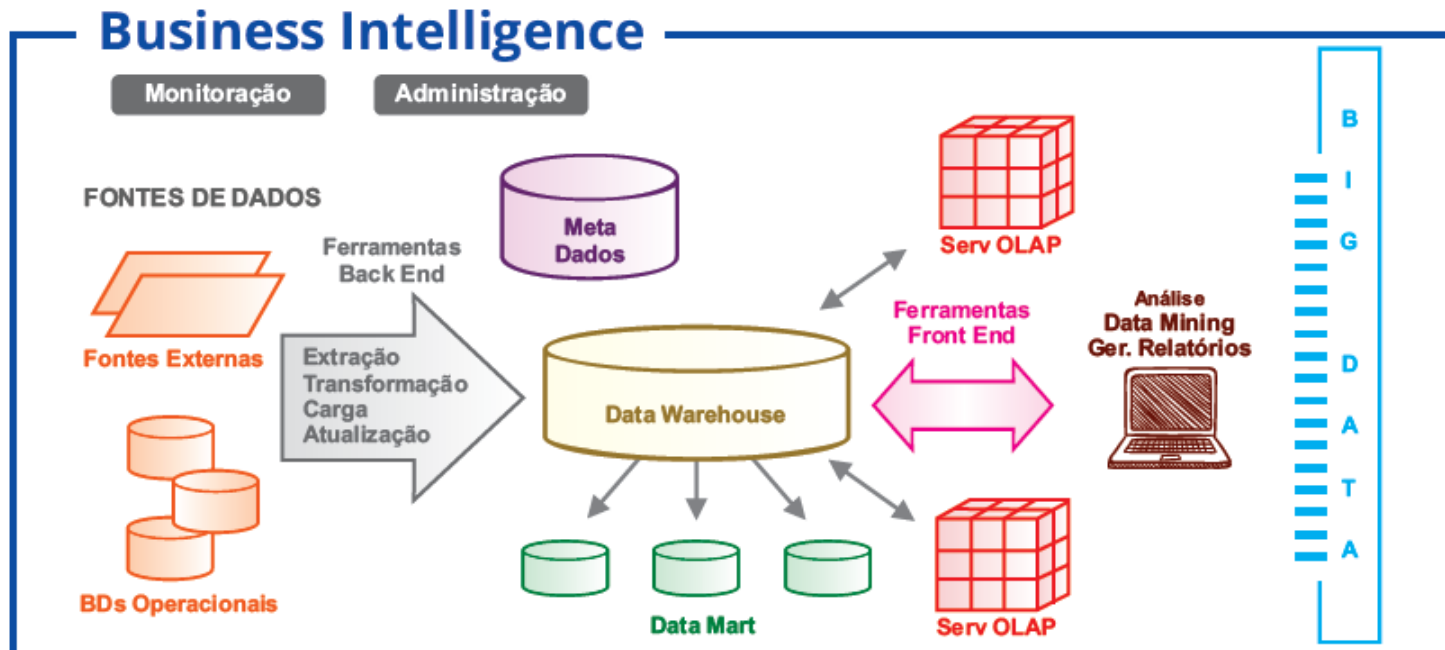
- O QUE É BUSINESS INTELLIGENCE
 - Na tradução livre, significa Inteligência de negócios ou simplesmente “BI”. É um conjunto de técnicas e ferramentas que, reunidas, possibilitam ao gestor a tomada de uma decisão com base em métricas ou valores que são extraídos de seus diversos sistemas e consolidados.





Sobre

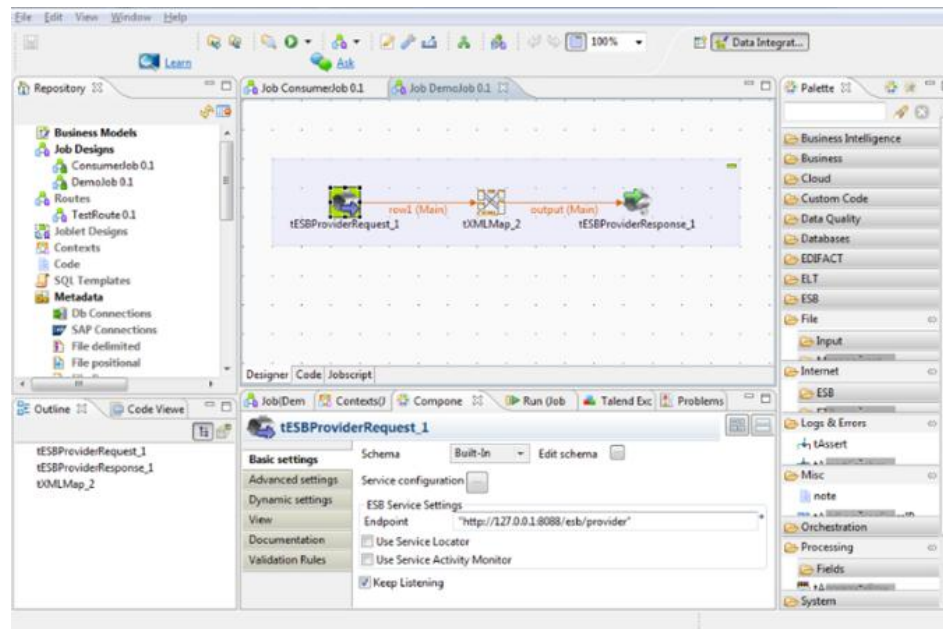
- O QUE É DATA WAREHOUSE
 - Também conhecido como “armazém de dados”, é um banco de dados que integra e consolida os diversos sistemas e fontes de dados (arquivos de texto, banco de dados, mensagens, dentre outras) de uma organização para apoio na tomada de decisão.





Sobre

- CONHECENDO O TALEND
 - É uma ferramenta de ETL, que oferece um ambiente gráfico para construir, gerenciar e manter processos de integração de dados em sistemas de Business Intelligence.

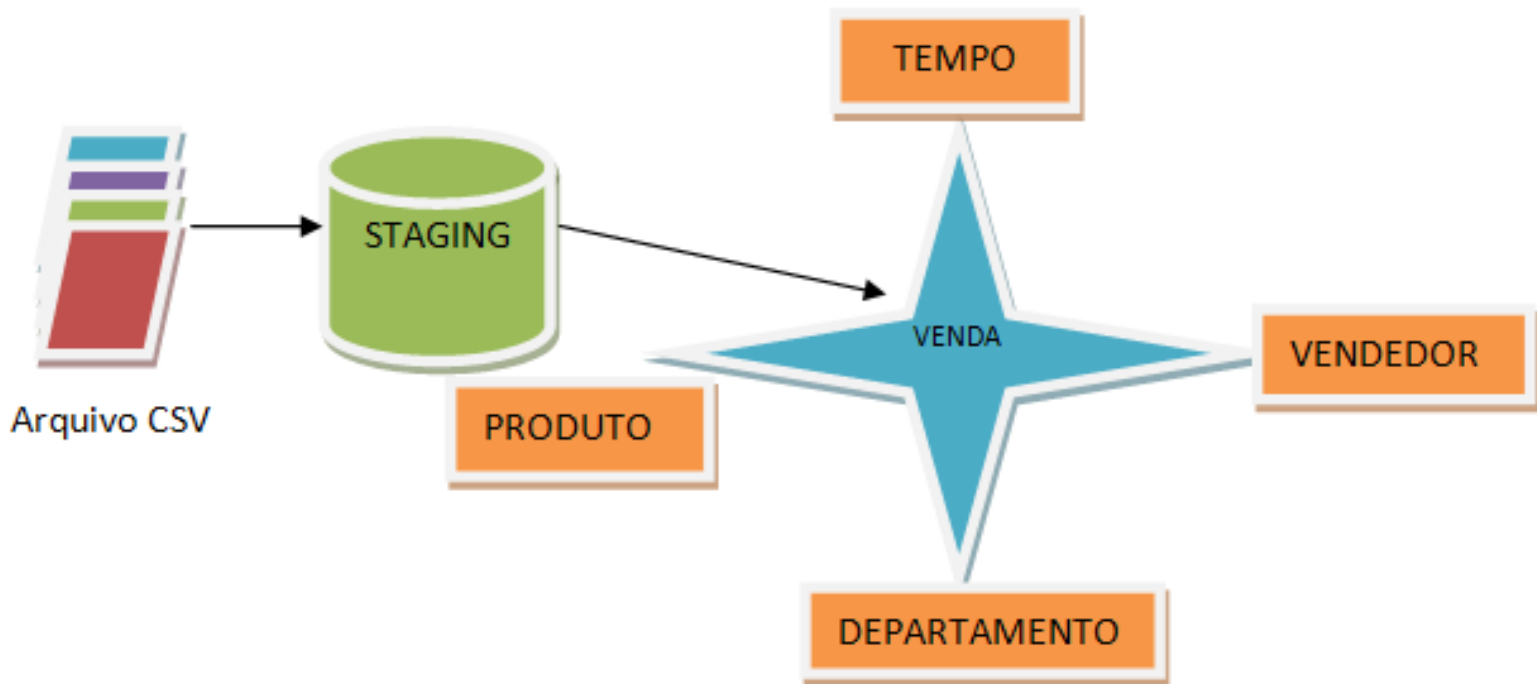


Aplicação Prática



Aplicação Prática

- DATA WAREHOUSE DE VENDAS





Sobre

- STAGING AREA
 - Staging Area ou área auxiliar serve como ponto único para a carga efetiva no data warehouse.
 - A cada carga seu conteúdo é limpo.
 - Evita acesso à produção em caso de recarga durante o dia.

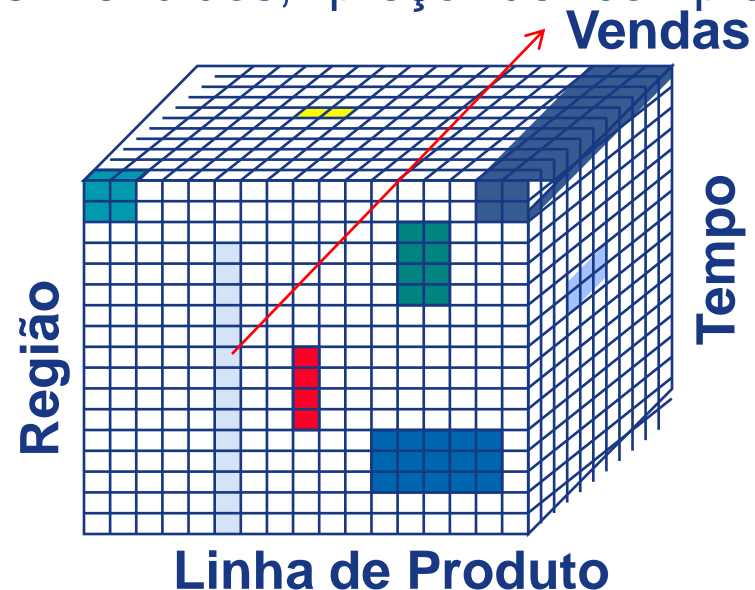




Sobre

- DIMENSÃO
 - Contém os descritores textuais do negócio.
 - Exemplo : Tempo , Cliente, Produto, Tipo de Embalagem, Situação, etc.
- FATO
 - Termo utilizado para a medição do negócio.
 - Exemplo: quantidade de produtos vendidos, preço de compra, preço de venda, lucro, etc.

- Dimensões
 - Região
 - Linhas de Produto
 - Tempo (obrigatória)
- Fato
 - Vendas

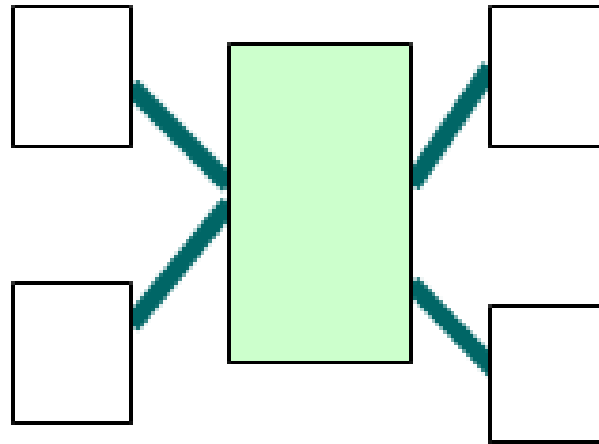




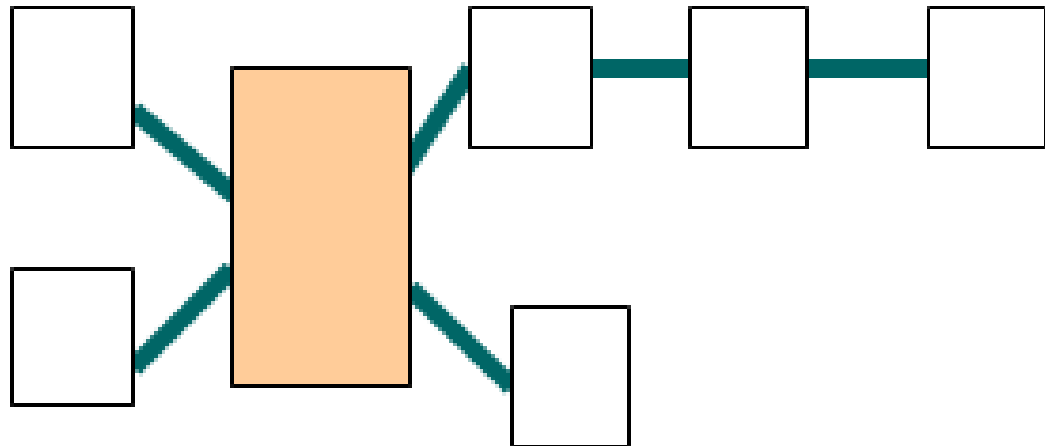
Sobre

- MODELAGEM MULTIDIMENSIONAL

STAR SCHEMA



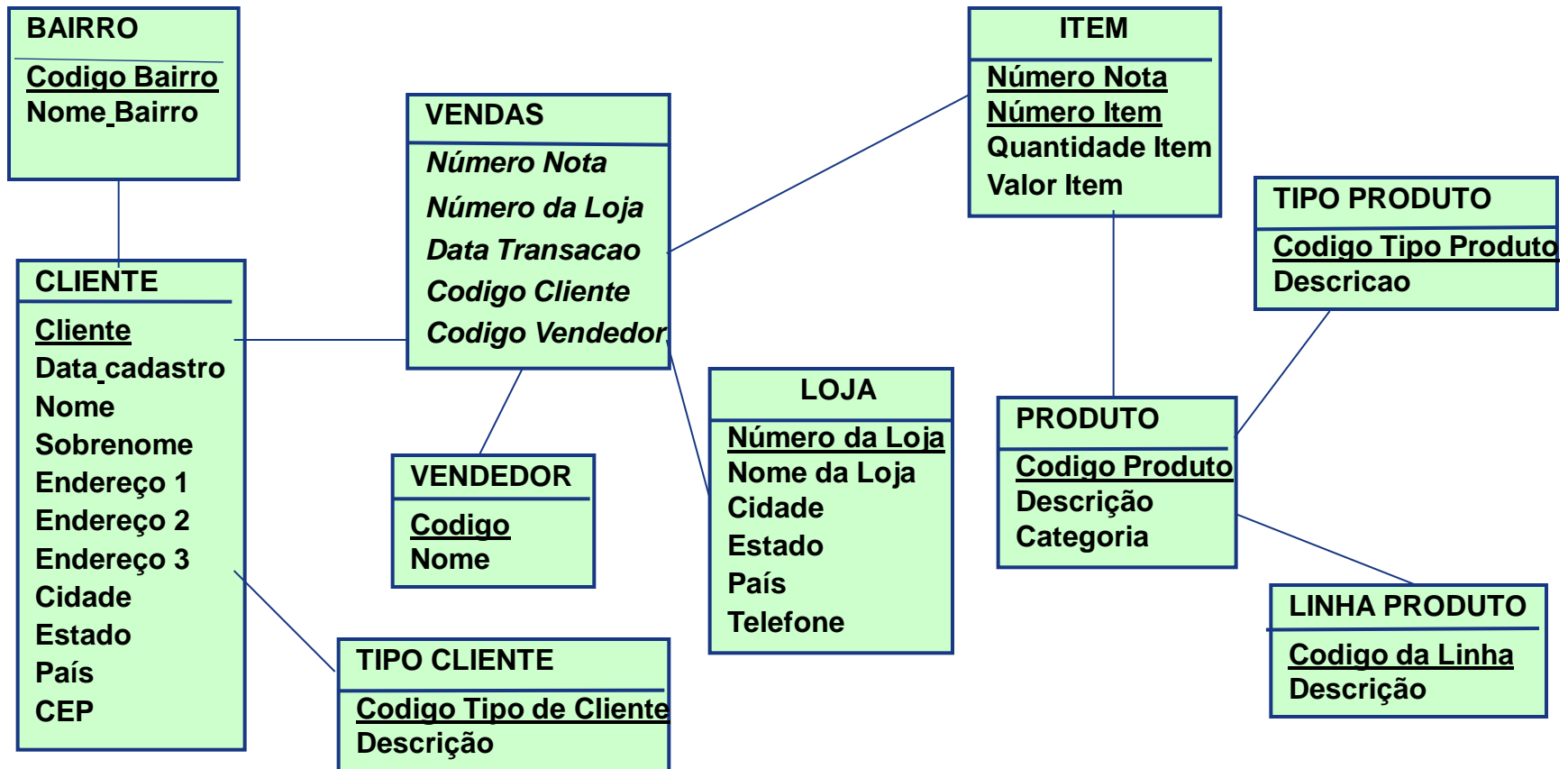
SNOW FLAKE





Sobre

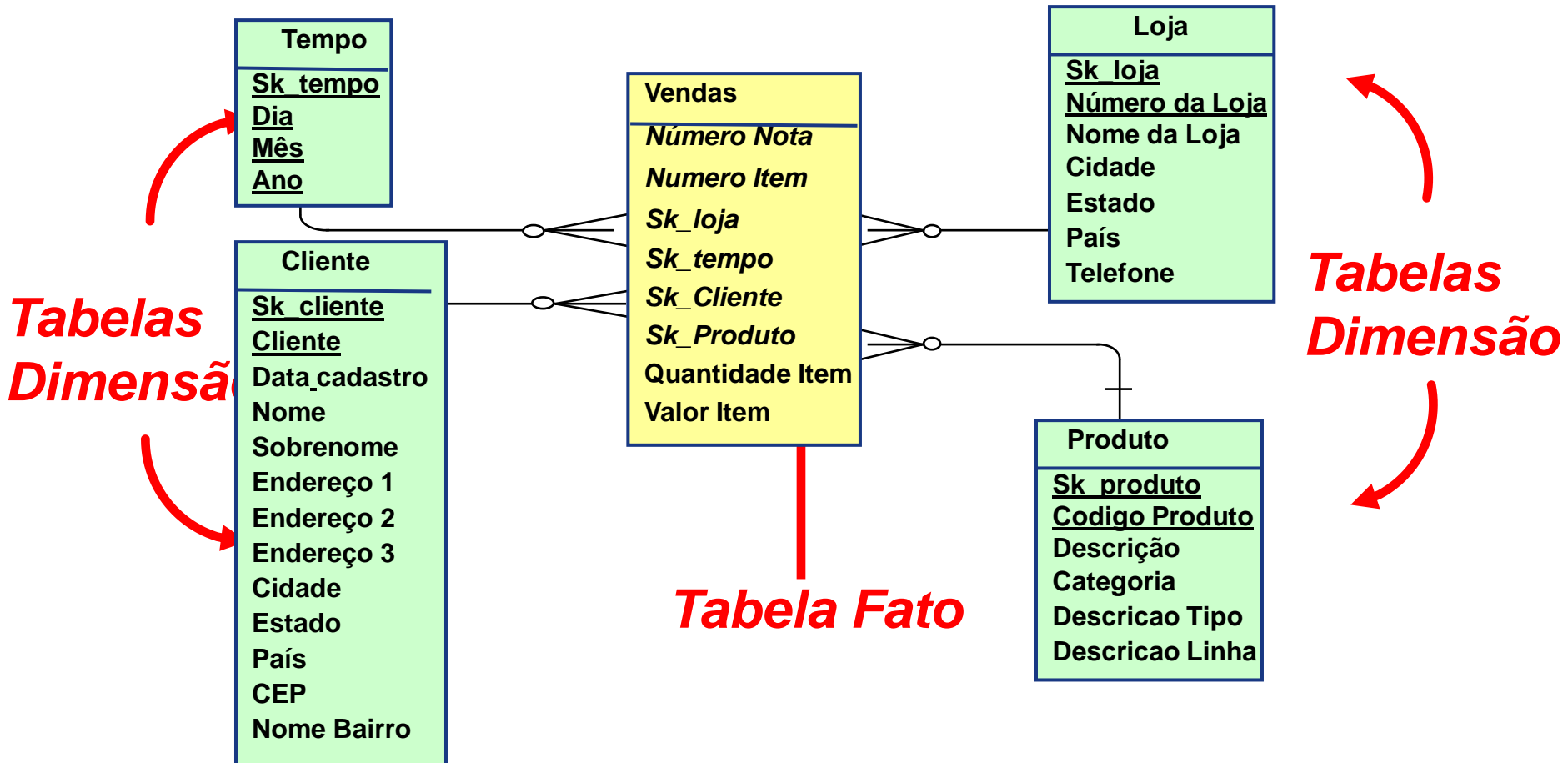
- TABELAS OPERACIONAIS





Sobre

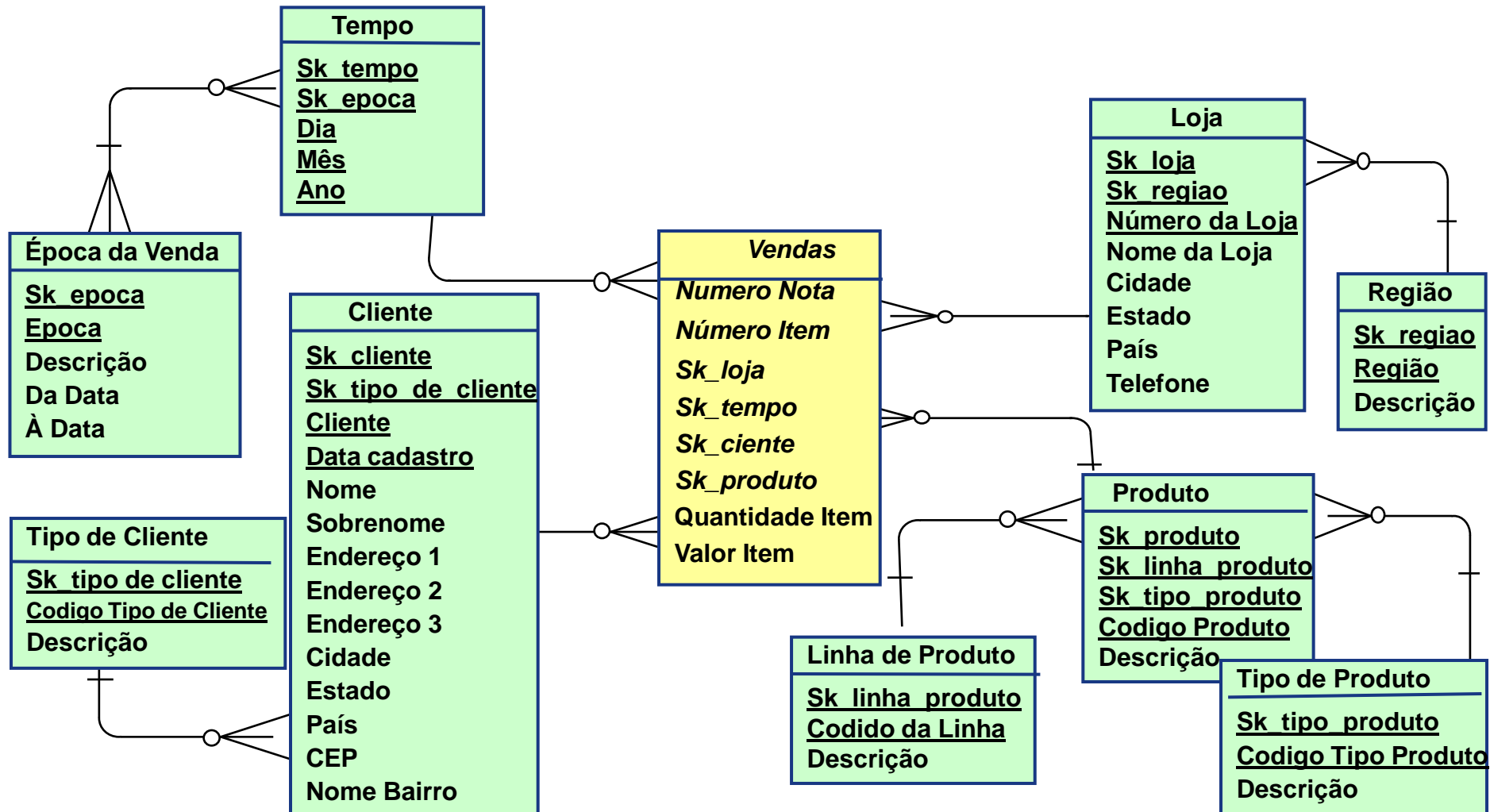
- MODELO MULTIDIMENSIONAL – STAR SCHEMA





Sobre

- MODELO MULTIDIMENSIONAL – SNOW FLAKE





Construção do Data Warehouse

- Importando Arquivos



- Criando Staging



- Criando Dimensões





Construção do Data Warehouse

- Carga Staging



- Carga Dimensões



- Criação da Fato



- Carga da Fato





COMPONENTES DO TALEND DATA INTEGRATION

tFileArchive: Compacta Arquivos

tFileUnarchive: Descompacta Arquivos

tFileCompare: Compara dois arquivos

tFileOutput: Grava a saída de um arquivo

tFileCopy: Copia um ou mais arquivos entre diretórios

tFileList: Seleciona um ou mais arquivos

tFileDelete: Deleta um ou mais arquivos

tJava: Executa um script em java script

tFileExist: Verifica se um ou mais arquivos existem

tMsgBox: Exibe uma mensagem na tela



COMPONENTES DO TALEND DATA INTEGRATION

tFileRow: Verifica a quantidade de registros de um arquivo

tAggregateRow: Agrega dados em um arquivo ou tabela

tSortRow: Ordena dados em uma tabela

tLogRow: Exibe dados em uma saída de tela

tReplace: Efetua uma operação de pesquisa e substituição

tFilterColumn: Muda a ordem de apresentação dos resultados

tFilterRow: Filtra registros em uma tabela ou arquivo

tSplitRow: Transforma linhas em colunas em uma tabela ou arquivo

tFileInputXML: Ler um arquivo XML e transforma em um arquivo de dados



COMPONENTES DO TALEND DATA INTEGRATION

tRowGeneration: Gera registros fictícios

tDie: Informa uma mensagem de alerta ou erro

tLogCatcher: Recebe a mensagem do componente tDie



TALEND DATA INTEGRATION – Importante Saber

• Variável de Contexto

Create / Edit a context group

Step 2 of 2

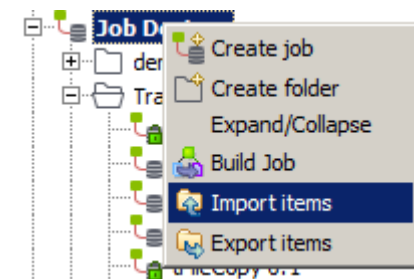
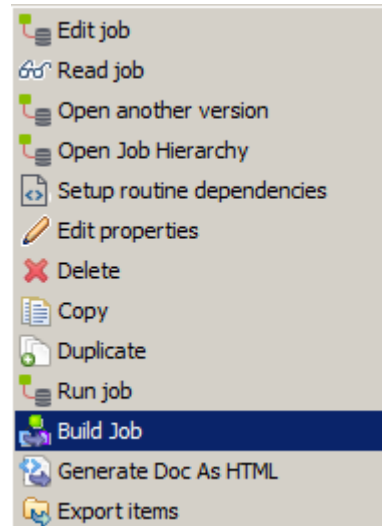
Define the contexts, variables and values

	Name	Type	Comment	Default Value	opcao2 Value
1	diretorio	String		"C:/TEMP"	"D:\Arquivo Treinamento Aprenda Virtual"
2	arquivo	String		"*.xlsx"	"*.csv"

Default context environment: Default

< Back Next > Finish Cancel

- Exportar JOB
- Exportar Projeto
- Importar Projeto



Certificado



Exercício Final

- O aluno deverá construir um projeto de DW do zero utilizando a ferramenta Talend Data Integration.
- O projeto deve contemplar: Tabelas Staging, tabelas Dimensão, tabela Fato, cargas Staging, cargas Dimensão e carga Fato.
- O aluno deve executar as cargas e popular as tabelas de Dimensão e Fato.
- O aluno deverá enviar o script de criação do banco de dados multidimensional do projeto ao professor e toda a estrutura de metadados do projeto.
- O aluno utilizará os dados de uma Folha de Pagamento para realizar o projeto, o arquivo encontrasse na área do aluno no site do curso.
- AO FINAL DEVEM SER CRIADOS:
- DIMENSÃO: Cargo, Departamento, Divisão, Funcionário
- FATO: Folha de Pagamento



OBRIGADO!!!

Dúvidas entre em contato.

**Avalie nosso curso e deixe uma
mensagem no fórum de notícias.**