



UNIVERSIDADE CATÓLICA DO SALVADOR – UCSAL
MBA EM TECNOLOGIAS E APLICAÇÕES DE BUSINESS INTELLIGENCE

TRABALHO FINAL – CONSTRUÇÃO DATA WAREHOUSE

FOLHA DE PAGAMENTO

Disciplina: Modelagem e Construção Multidimensional em
Data Warehouse.

Prof.: Grimaldo Lopes

Aluno: Camila da Silva Oliveira

Salvador, Março de 2019

Sumário

TRABALHO – CONSTRUÇÃO DATA WAREHOUSE FOLHA DE PAGAMENTO 3

- 1. INTRODUÇÃO 3
- 2. OBJETIVO 4
- 3. ETAPAS DA CONSTRUÇÃO DO DATA WAREHOUSE 5
 - 3.1. *Matriz de Necessidades do Modelo* 5
 - 3.2. *Fonte de Dados do Modelo* 6
 - 3.3. *Modelo Dimensional* 7
 - 3.4. *Construção das Cargas: Staging, Dimensão e Fato* 8
- 4. CONCLUSÃO 12
- 5. BIBLIOGRAFIA 13

Trabalho – Construção Data Warehouse Folha de Pagamento

1. Introdução

Com a evolução da tecnologia e o crescimento da competitividade do mercado, é demandado das empresas a necessidade de serem mais ágeis e eficientes, precisando constantemente se adaptar as mudanças e desenvolver novas metodologias para apresentar os resultados necessários para sobreviver. Uma das soluções amplamente utilizada é o business intelligence (BI). Uma série de processos que tem como objetivo reunir e analisar dados, transformando-os em informações e conhecimento, possibilitando as tomadas de decisão com maior assertividade para obtenção de um melhor retorno e lucro.

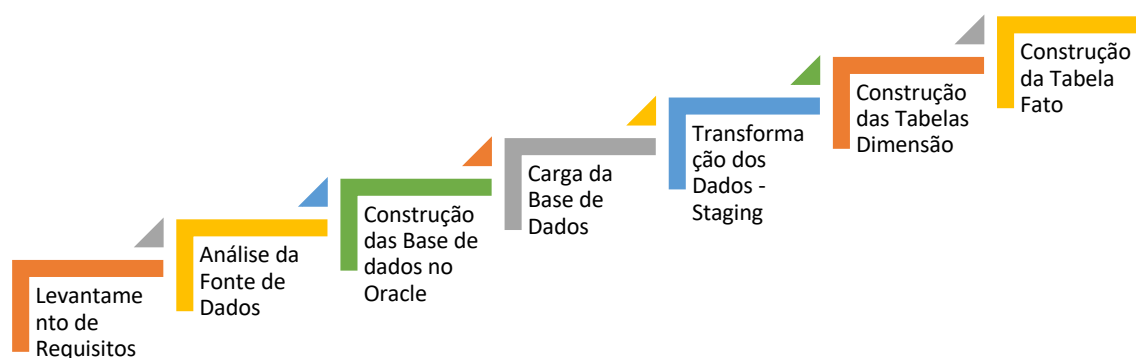
Para a implantação do BI pode-se utilizar diversas técnicas, tendo como objetivo principal a construção do Data Warehouse – depósito de dados. Coleção de dados orientados pelos principais assuntos da empresa, integrados, variante ao tempo podendo ser apresentados em tempos determinados e os dados existentes não devem ser modificados. Cada assunto de uma DW é chamado de Data Mart.

O trabalho tem como objetivo demonstrar a construção de uma solução de business intelligence, passando pela maioria das etapas de desenvolvimento. Foram fornecidos arquivos contendo informação referente a folha de pagamento de uma empresa fictícia.

A primeira etapa constitui de reuniões para entendimento do negócio e levantamento dos requisitos com os stakeholders – pessoas responsáveis e envolvida no negócio. Esta etapa foi substituída pela análise dos arquivos disponibilizados e desenvolvimento de perguntas que geralmente são feitas pelos stakeholders com expectativa de serem respondidas pelo DW.

Com base nas perguntas é elaborada a matriz de necessidade, com o mapeamento de todos os indicadores e métricas que irão existir na solução. Em seguida foi realizado o mapeamento da fonte, onde os arquivos são reavaliados e mapeados, com estes dados foi elaborado o modelo multidimensional – definição do modelo adequado para as consultas bases que serão feitas.

A próxima etapa foi a construção do DW constituída das etapas de transferência dos dados do operacional (staging area ou área auxiliar), transformação e limpeza de dados (ETL) para construção das tabelas dimensão e tabela fato.



Etapas de Desenvolvimento Solução do BI

2. Objetivo

A proposta principal do trabalho é efetuar a construção de uma Data Warehouse da Folha de Pagamento, através da análise e tratamento dos arquivos com os dados relacionados ao assunto e mais informação externa para enriquecimento das análises.

As fontes das informações disponibilizadas são os arquivos CSV (texto separados com vírgulas), descritos abaixo:

- Cargo.csv – Descrição dos cargos ocupados pelos funcionários da empresa;
- Departamento – Setor da empresa onde os colaboradores estão alocados;
- Divisão – Área de atuação de cada funcionário;
- Funcionários – Relação dos colaboradores, contendo o nome e sobrenome;
- Média Salário – Base de dados complementar com a informação dos salários praticados no mercado por cargo, incluindo a quantidade de horas da jornada de trabalho, piso salarial, média salarial, o maior salário pago e valor por hora.
- Cargo_Vinculo – Complementar a informação foi elaborado um arquivo com o de-para do descritivos dos cargos;
- Folha de Pagamento – Valor pago por colaborador em um determinado período de tempo.

Os dados fornecidos foram avaliados e elaborados questionamentos sobre a folha de pagamento. Que definiram as informações necessárias para a tomada de decisões mais efetivas, servindo de base para as demais etapas para a construção do Data Warehouse.

Questionamentos realizados:

N	PERGUNTAS
1	Qual o gasto de pessoal por departamento e divisão?
2	Qual o gasto mensal com funcionários?
3	Qual o gasto mensal por divisão, departamento e cargo?
4	Quais cargos estão acima da média salarial de mercado?
5	Quais cargos estão abaixo da média salarial de mercado?
6	Há salário pago acima do maior do mercado? Quantos funcionários?
7	Qual funcionário com maior e menor salário?
8	Quais os funcionários estão abaixo do piso?
9	Quantas pessoas por departamento, divisão e cargo?
10	Qual a média salarial por departamento, divisão e cargo?

3. Etapas da Construção do Data Warehouse

3.1. Matriz de Necessidades do Modelo

Na matriz de necessidade são relacionados os indicadores e métricas pontuadas nos questionamentos.

DADOS LEVANTADOS	INDICADORES	Folha de Pagamento
------------------	-------------	--------------------

	Salário - Valor
Cargo	
Descrição do Cargo	✓
Piso_Mercado	✓
Media_Mercado	✓
Maior_Mercado	✓
Departamento	
Descrição do Departamento	✓
Funcionário	
Nome	✓
Sobrenome	✓
Divisão	
Nome da Divisão	✓
TEMPO	
Descrição ano mês	✓
Descrição bimestre	✓
Descrição data dia	✓
Descrição do dia	✓
Descrição do mês	✓
Descrição mês ano	✓
Descrição mês ano numérico	✓
Descrição mês ano completo	✓
Descrição quadrimestre	✓
Descrição quinzena	✓
Descrição semestre	✓
Descrição trimestre	✓
Número ano	✓
Número bimestre	✓
Número dia	✓
Número mês	✓
Número nível	✓
Número quadrimestre	✓
Número quinzena	✓
Número semestre	✓
Número trimestre	✓
Status fim de semana	✓

3.2. Fonte de Dados do Modelo

Diagrama utilizado para mapear os campos levantados na Matriz de Necessidade que constam nos arquivos utilizados.

Dimensões	Base Original			
	TABELA	Atributo	Relac.	Regra/Domínio
Cargo				
	Folha_Pagamento	cod_cargo	CHAVE1	
	Cargo	cod_cargo	CHAVE1	
Descrição do Cargo	Cargo	cargo		
	Cargo	cargo	CHAVE2	
	Media_Salarial	cargo	CHAVE2	Necessário validar os cargos para realizar o cruzamento
Piso_Mercado	Media_Salarial	Piso		
Media_Mercado	Media_Salarial	Media		
Maior_Mercado	Media_Salarial	Maior		
Departamento				
	Folha_Pagamento	cod_departamento	CHAVE1	
	Departamento	cod_departamento	CHAVE1	
Descrição do Departamento	Departamento	departamento		
Funcionario				
	Folha_Pagamento	cod_funcionario	CHAVE1	
	Funcionário	cod_funcionario	CHAVE1	
Sobrenome	Funcionário	sobrenome		
Nome	Funcionário	nome		
Divisao				
	Folha_Pagamento	cod_divisao	CHAVE1	
	Divisao	cod_divisao	CHAVE1	
Nome da Divisão	Divisao	divisao		
Tempo				
Data Pagamento	Folha_Pagamento	Data		
Ano de Referência				
Semestre				
Quadrimestre				
Trimestre				
Bimestre				
Mês de Referência				
Quinzena de Ref.				
Dia				
Dia da Semana				

Com base nos dados fornecidas, foram definidas as métricas:

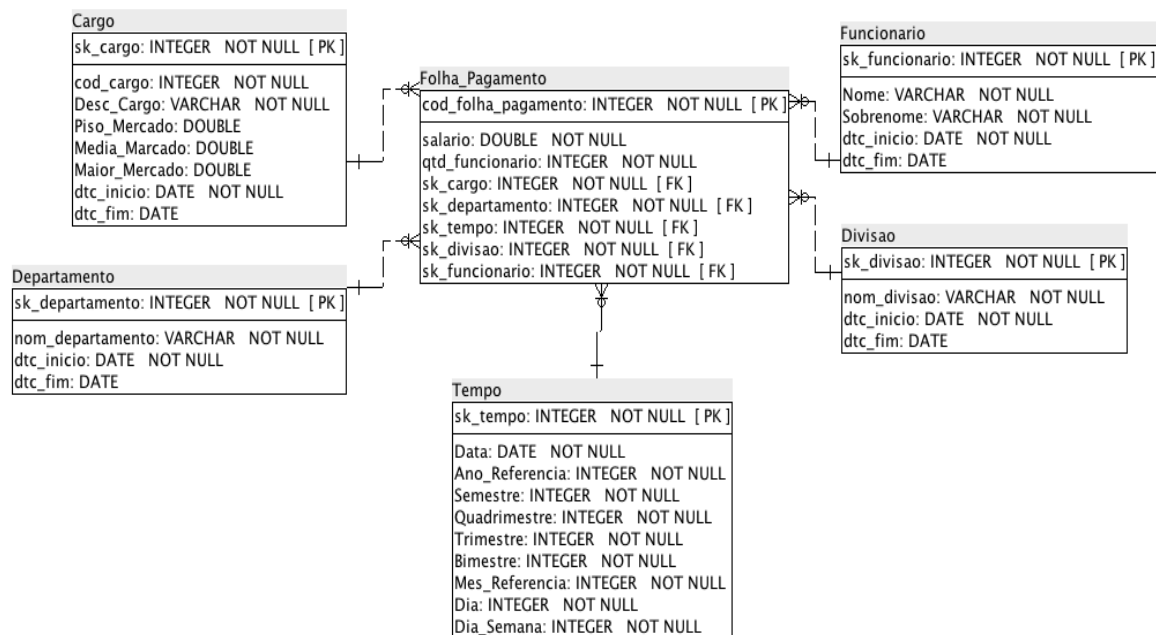
Métricas	FATO	Base Original		
		TABELA	Atributo	Regra
Salário	Folha de Pagamento	Folha	salario_valor	
Quantidade_Funcionario	Folha de Pagamento	Folha		Se cod_funcionario não é nulo considera 1, se não considera 0.

3.3. Modelo Dimensional

Com a definição da fonte de dados, seguiu-se com a definição do modelo dimensional. Foram definidas cinco dimensões e a tabela fato, organizados no modelo estrela – composição de uma tabela fato central, contendo as chaves artificiais e métricas, vinculadas as cinco dimensões, conforme diagrama. As dimensões elaboradas foram do tipo Slow Changing Dimension do tipo de controle de versões, que possibilita a manutenção do histórico.

Dimensões – Tabelas cargo, departamento, funcionário, divisão e tempo.

Fato – Tabela folha de pagamento.



Modelo Dimensional

3.4. Construção das Cargas: Staging, Dimensão e Fato

A construção do Data Mart Folha de Pagamento foi realizada no Talend e os dados armazenados no Banco de Dados Oracle. A primeira etapa da construção foi a elaboração das tabelas onde os dados do Staging serão armazenados.

Na base do staging foi adicionada a fonte de dados de média salarial. Para possibilitar o cruzamento das informações da base com a média salarial foi criada a tabela de vínculo relacionado o cargo.

Tabelas criadas para a fonte do Staging

REP_OWB.TRAB_STG_CARGO	
P *	CARGOID NUMBER (*,0)
*	DESC_CARGO VARCHAR2 (50 BYTE)
PK_CARGO_CARGOID (CARGOID)	
PK_CARGO_CARGOID (CARGOID)	

REP_OWB.TRAB_STG_CARGO_VINCULO	
P *	CARGOID NUMBER (*,0)
*	CBO NUMBER (*,0)
PK_CARGO_VINC (CARGOID)	
PK_CARGO_VINC (CARGOID)	

REP_OWB.TRAB_STG_DIVISAO	
P *	DIVISAOID NUMBER (*,0)
	DESC_DIVISAO VARCHAR2 (50 BYTE)
PK_DIVISAO_DIVISAOID (DIVISAOID)	
PK_DIVISAO_DIVISAOID (DIVISAOID)	

REP_OWB.TRAB_STG_FOLHA_PAGAMENTO	
P *	FUNCIONARIOID NUMBER (*,0)
P *	CARGOID NUMBER (*,0)
P *	DEPARTAMENTOID NUMBER (*,0)
P *	DIVISAOID NUMBER (*,0)
*	DATA_PG DATE
	SALARIO FLOAT (126)
PK_FOLHAPAG_ID (FUNCIONARIOID, CARGOID, DEPARTAMENTOID, DIVISAOID, DATA_PG)	
PK_FOLHAPAG_ID (FUNCIONARIOID, CARGOID, DEPARTAMENTOID, DIVISAOID, DATA_PG)	

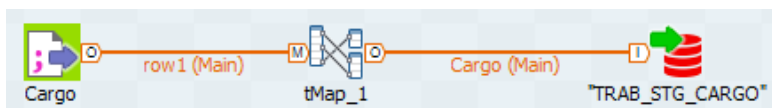
REP_OWB.TRAB_STG_FUNCIONARIO	
P *	FUNCIONARIOID NUMBER (*,0)
	NOME_FUNCIONARIO VARCHAR2 (50 BYTE)
	SOBRENOME_FUNCIONARIO VARCHAR2 (50 BYTE)
PK_FUNCIONARIO_FUNCIONARIOID (FUNCIONARIOID)	
PK_FUNCIONARIO_FUNCIONARIOID (FUNCIONARIOID)	

REP_OWB.TRAB_STG_MEDIA	
P *	CBO NUMBER (*,0)
P *	CARGO VARCHAR2 (200 BYTE)
	JORNADA NUMBER (*,0)
	PISO FLOAT (126)
	MEDIA FLOAT (126)
	MAIOR FLOAT (126)
	SALARIO_HORA FLOAT (126)
PK_MEDIA_CBO (CBO, CARGO)	
PK_MEDIA_CBO (CBO, CARGO)	

REP_OWB.TRAB_STG_DEPARTAMENTO	
P *	DEPARTAMENTOID NUMBER (*,0)
	DESC_DEPARTAMENTO VARCHAR2 (50 BYTE)
PK_DEPARTAMENTO_DEPARTAMENTOID (DEPARTAMENTOID)	
PK_DEPARTAMENTO_DEPARTAMENTOID (DEPARTAMENTOID)	

Após a construção das tabelas no Oracle, realizada o desenvolvimento das regras para carga das tabelas do staging com base nos arquivos fontes no Talend. No desenvolvimento foram criados 7 Jobs de Upload do Staging:

- Carga Staging Cargo



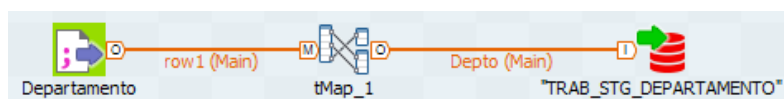
- Carga Staging Cargo Vinculo



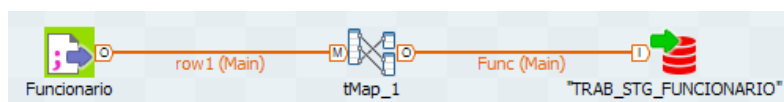
- Carga Staging Divisão



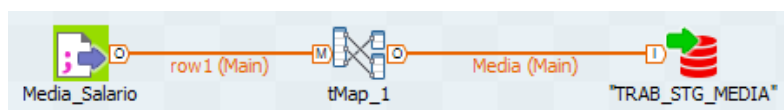
- Carga Staging Departamento



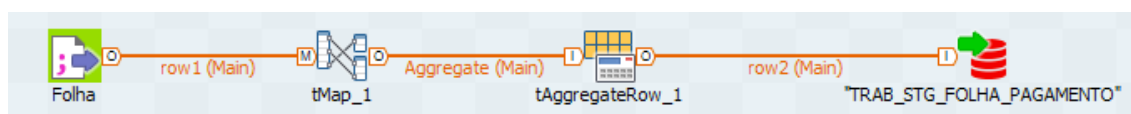
- Carga Staging Funcionário



- Carga Staging Media



- Carga Staging Folha Pagamento



Uma vez concluída a carga do Staging é feita a construção da base física das tabelas dimensões e fato, para posterior carga das tabelas.

Tabelas criadas para a fonte da Dimensão e Fato

REP_OWB.TRAB_DIM_CARGO	
* SK_CARGO	NUMBER
P * CARGOID	NUMBER
* DESC_CARGO	VARCHAR2 (50 BYTE)
PISO	FLOAT (126)
MEDIA	FLOAT (126)
MAIOR	FLOAT (126)
DTC_INICIO	DATE
DTC_FIM	DATE
PK_DIM_SK_CARGO (CARGOID)	
PK_DIM_SK_CARGO (CARGOID)	

REP_OWB.TRAB_DIM_DEPARTAMENTO	
* SK_DEPARTAMENTO	NUMBER
P * DEPARTAMENTOID	NUMBER
* DESC_DEPARTAMENTO	VARCHAR2 (50 BYTE)
DTC_INICIO	DATE
DTC_FIM	DATE
PK_DIM_SK_DEPTO (DEPARTAMENTOID)	
PK_DIM_SK_DEPTO (DEPARTAMENTOID)	

REP_OWB.TRAB_DIM_DIVISAO	
* SK_DIVISAO	NUMBER
P * DIVISAOID	NUMBER
* DESC_DIVISAO	VARCHAR2 (50 BYTE)
DTC_INICIO	DATE
DTC_FIM	DATE
PK_DIM_SK_DIVISAO (DIVISAOID)	
PK_DIM_SK_DIVISAO (DIVISAOID)	

REP_OWB.TRAB_DIM_TEMPO	
P * SK_TEMPO	NUMBER
DES_DATA_DIA	VARCHAR2 (25 BYTE)
DTC_DATA	DATE
NUM_ANO	NUMBER (4)
NUM_MES	NUMBER (3)
NUM_DIA	NUMBER (3)
NUM_QUADRIMESTRE	NUMBER (3)
NUM_TRIMESTRE	NUMBER (2)
DES_QUINZENA	VARCHAR2 (12 BYTE)
DES_QUADRIMESTRE	VARCHAR2 (12 BYTE)
NUM_BIMESTRE	NUMBER (3)
DES_BIMESTRE	VARCHAR2 (12 BYTE)
DES_ANO_MES	VARCHAR2 (8 BYTE)
DES_DIA	VARCHAR2 (7 BYTE)
NUM_SEMESTRE	NUMBER (3)
DES_MES_ANO_NUMERICO	VARCHAR2 (7 BYTE)
DES_TRIMESTRE	VARCHAR2 (12 BYTE)
IND_FINAL_SEMANA	CHAR (1 BYTE)
DES_MES	VARCHAR2 (15 BYTE)
NUM_QUINZENA	NUMBER (3)
DES_MES_ANO_COMPLETO	VARCHAR2 (30 BYTE)
DES_SEMESTRE	VARCHAR2 (12 BYTE)
NUM_NIVEL	NUMBER (1)
DES_MES_ANO	VARCHAR2 (8 BYTE)
TRAB_DIM_TEMPO_KEY_PK (SK_TEMPO)	
TRAB_DIM_TEMPO_KEY_PK (SK_TEMPO)	

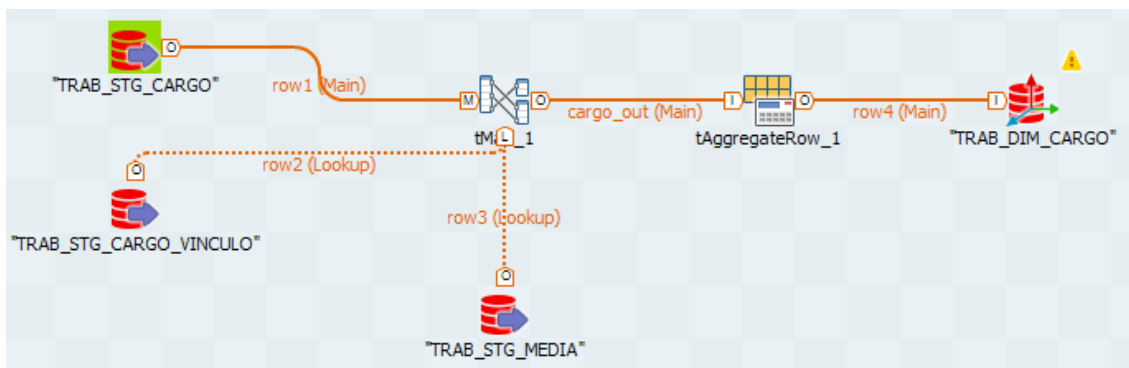
REP_OWB.TRAB_FATO_FOLHA	
P * SK_CARGO	NUMBER
P * SK_FUNCIONARIO	NUMBER
P * SK_DIVISAO	NUMBER
P * SK_DEPARTAMENTO	NUMBER
P * SK_TEMPO	NUMBER
SALARIO_VALOR	FLOAT (126)
* QTD_FUNCIONARIOS	NUMBER
PK_FATO_FOLHAID (SK_CARGO, SK_FUNCIONARIO, SK_DIVISAO, SK_DEPARTAMENTO, SK_TEMPO)	
PK_FATO_FOLHAID (SK_CARGO, SK_FUNCIONARIO, SK_DIVISAO, SK_DEPARTAMENTO, SK_TEMPO)	

REP_OWB.TRAB_DIM_FUNCIONARIO	
* SK_FUNCIONARIO	NUMBER
P * FUNCIONARIOID	NUMBER
NOME_FUNCIONARIO	VARCHAR2 (50 BYTE)
SOBRENOME_FUNCIONARIO	VARCHAR2 (50 BYTE)
DTC_INICIO	DATE
DTC_FIM	DATE
PK_DIM_SK_FUNC (FUNCIONARIOID)	
PK_DIM_SK_FUNC (FUNCIONARIOID)	

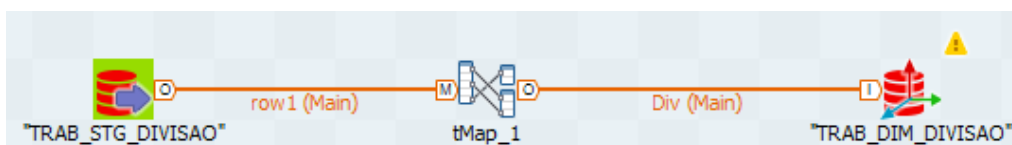
A tabela de dimensão foi preenchida através de uma procedure específica que preenche com base em um período específico.

O processo de carga da tabela física da dimensão foi realizado através dos Jobs no Talend, nesta parte foi realizado o tratamento, limpeza e transformação dos dados. Posteriormente foi realizado a carga da tabela fato.

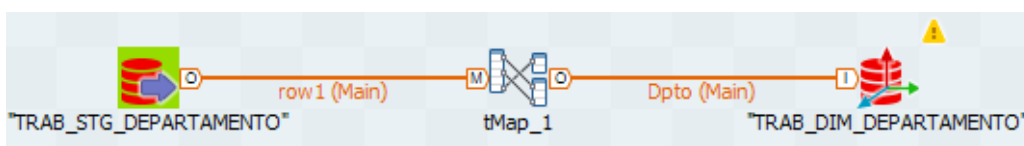
- Dimensão Cargo



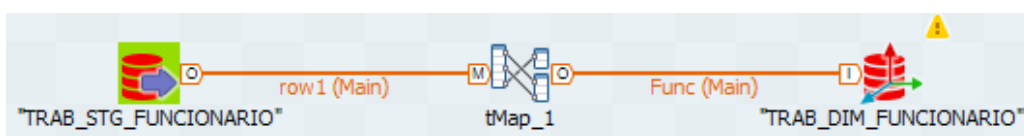
- Dimensão Divisão



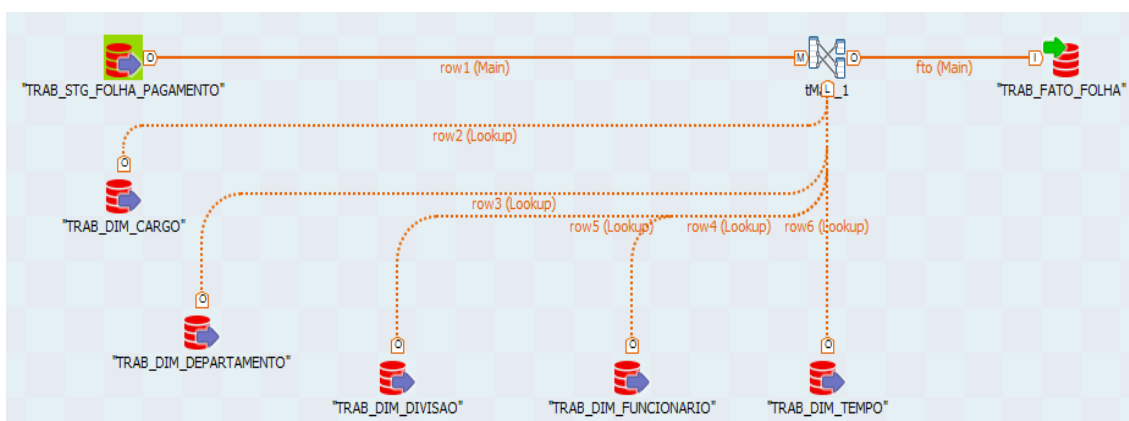
- Dimensão Departamento



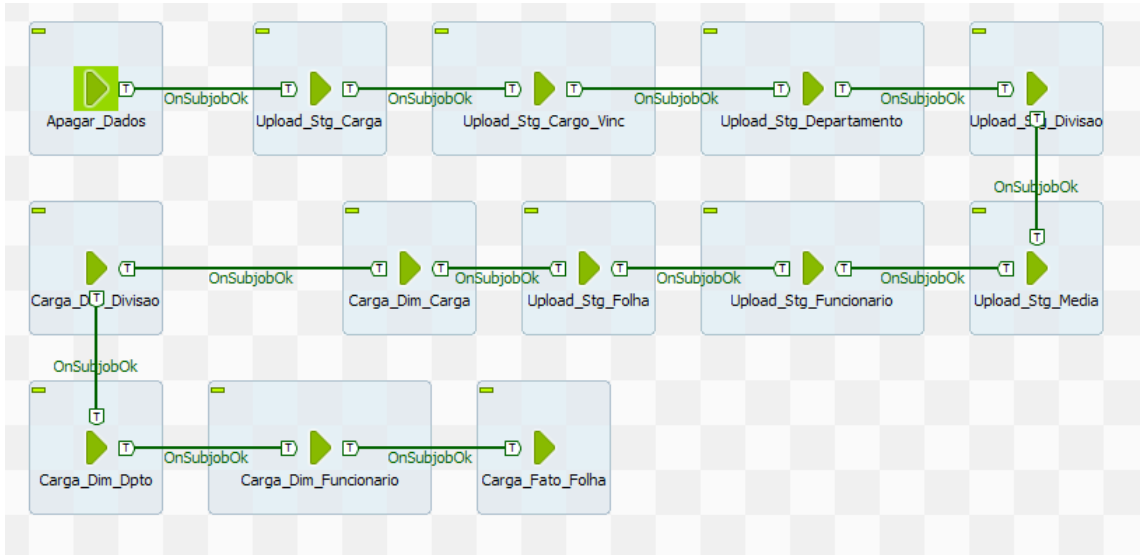
- Dimensão Funcionário



A última etapa foi a construção da tabela fato.



Para tornar o processo dinâmico e automático, foi construída o Job agrupando todos os processos anteriores.

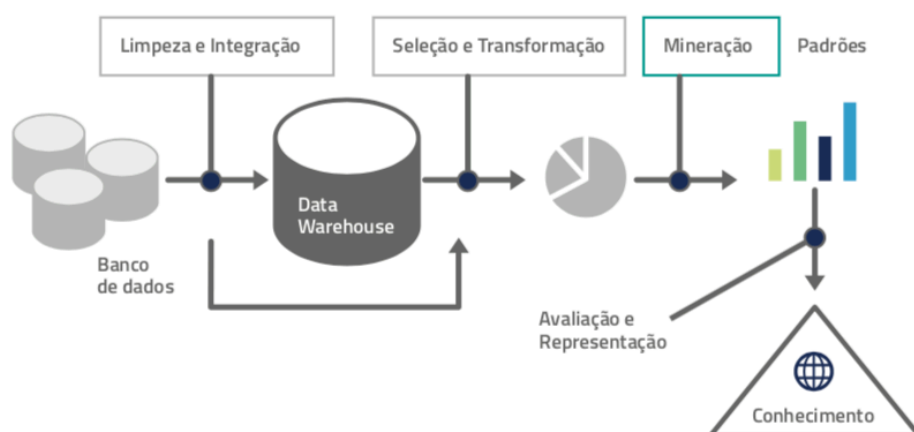


4. Conclusão

Finalizado desenvolvimento da Data Warehouse e implantando, a empresa terá acesso as informações estratégicas de forma estruturada, que poderá ser utilizada para acompanhamento dos gastos com pessoais sob várias vertentes. Alguns ganhos que poderão ser obtidos a partir da análise:

- Elaborar relatório de acompanhamento da folha de pagamento;
- Atuar nas áreas que estão com gastos acima ou abaixo da média de mercado para servir de subsidio para adoção de política de pessoal como plano de cargo e salários ou plano de carreiras;
- Verificar as disparidades de pagamento de salário para futuro ajuste e prevenção de futuras as judiciais;
- Analisar a quantidade de colaboradores por área e valor gasto com pessoal para tomada de decisão de futuros investimentos.

Outro ganho seria o cruzamento dos dados com informações de outras áreas da empresa, possibilitando incremento dos ganhos. Além das análises com base nas informações disponibilizadas, a empresa poderá investir e efetuar mineração de dados para identificar padrões nos dados utilizados e produzir novos conhecimentos.



Fluxo mineração de dados – Fonte: BI como deve ser.

O conhecimento adquirido através da implantação de uma solução de business intelligence possibilita a tomada de decisão mais assertiva e com velocidade, além de contribuir para aumentar ganhos e corrigir disparidades.

5. Bibliografia

Oliveira, Diego Elias; Oliveira, Grimaldo Lopes de. BI Como Deve Ser – O Guia Definitivo. 2ª ed. Salvador: 2016.

A importância do Data Warehouse em um projeto de BI < <http://igti.com.br/blog/a-importancia-do-data-warehouse-em-um-projeto-de-bi/>> Acessado em 04/03/19.

O que é business intelligence <<https://rafaelpiton.com.br/blog/o-que-e-bi-business-intelligence/>> Acessado em 04/03/19.

Conhecendo a arquitetura de Data Warehouse <<https://canaltech.com.br/business-intelligence/conhecendo-a-arquitetura-de-data-warehouse-19266/>> Acessado em 04/03/19.

Tabela Salarial 2019 < <https://www.salario.com.br/tabela-salarial/>> Acessado em 04/03/19.

As etapas do Business Intelligence < <https://www.binapratice.com.br/etapas-bi>> Acessado em 04/03/19.