

Painel de Acompanhamento da Folha de pagamento do Bolsa Família

Projeto final do curso:

**BI MASTER - BUSINESS INTELLIGENCE MASTER - SISTEMAS INTELIGENTES  
DE APOIO À DECISÃO EM NEGÓCIOS**

Orientador: Prof. Anderson Nascimento

prof.anderson@ica.ele.puc-rio.br

**Componentes do Projeto:**

Marilivia Melo de Sousa Mat. 221100845 – mariliviamelo@gmail.com

## Histórico de Versões

<b>Data</b>	<b>Versão</b>	<b>Descrição</b>	<b>Autor</b>
30/06/2024	1.0	Planejamento do trabalho	Marilivia Melo de Sousa
06/07/2024	1.1	Cargas da base	Marilivia Melo de Sousa
12/07/2024	1.2	Tratamento da base SQL	Marilivia Melo de Sousa
26/07/2024	1.3	Montagem banco transacional	Marilivia Melo de Sousa
01/08/2024	1.4	Montagem Data Warehouse	Marilivia Melo de Sousa
10/08/2024	1.5	ETL(PDI)	Marilivia Melo de Sousa
17/08/2024	1.6	Elaboração do Painel (Power BI)	Marilivia Melo de Sousa
18/08/2024	1.7	Revisão geral	Marilivia Melo de Sousa

## SUMÁRIO1

<b>1. INTRODUÇÃO.....</b>	<b>4</b>
<b>2. ESTUDO DE CASO .....</b>	<b>5</b>
<b>2.1 Descrição do Estudo de Caso.....</b>	<b>5</b>
<b>2.2 Descrição do Modelo Transacional.....</b>	<b>6</b>
<b>2.3 Fonte 1 – Fonte dados iniciais.....</b>	<b>6</b>
<b>2.4 Proposta de Processo de BI.....</b>	<b>7</b>
<b>3. MODELO MULTIDIMENSIONAL.....</b>	<b>8</b>
<b>4. ELABORAÇÃO DO DATA WAREHOUSE .....</b>	<b>9</b>
<b>4.1 Definição do DW .....</b>	<b>9</b>
4.1.1 ARQUITETURA.....	9
4.1.2 ABORDAGEM DE CONSTRUÇÃO .....	9
4.1.3 ARQUITETURA FÍSICA .....	9
<b>5. PROJETO DE ETL .....</b>	<b>10</b>
<b>5.1 Descrição do Projeto de ETL.....</b>	<b>10</b>
<b>6. DASHBOARD .....</b>	<b>12</b>
<b>7. CONCLUSÃO .....</b>	<b>14</b>
<b>8. ANEXOS.....</b>	<b>15</b>
<b>9. ARQUIVOS.....</b>	<b>16</b>

## **1. INTRODUÇÃO**

Este documento tem por finalidade coletar, analisar e definir as principais necessidades do projeto do estudo de caso para o desenvolvimento do Painel de Acompanhamento da Folha de pagamento do Bolsa Família.

## **2. ESTUDO DE CASO**

Nesta seção vamos apresentar uma visão resumida da necessidade e objetivo do projeto BI.

### **2.1 Descrição do Estudo de Caso**

O presente estudo de caso tem como objetivo o desenvolvimento de um painel gerencial para o acompanhamento da folha de pagamento de um programa social – Bolsa Família. A CAIXA por meio da lei n 14.601, de 19 de junho de 2023, foi designada agente operador do programa, ou seja, é a responsável pelo pagamento do benefício as famílias participantes do programa.

A gestão do processo de geração da folha de pagamento, está a cargo da Gerência Nacional GEFOL.

A equipe da GEFOL, é responsável por acompanhar todo processo de geração da folha de pagamento, além de fornecer informações para a diretoria da empresa sobre a folha tais como quantidade de famílias na folha de pagamento por referência, montante a ser pago, situação das famílias na geração da folha e canais de pagamento.

Atualmente, essas informações são disponibilizadas na forma de um relatório que precisa ser atualizado manualmente, o que demanda bastante tempo e atenção dos membros da equipe que tem que parar alguma de suas atividades para atualizar o reporte.

A fim de otimizar o tempo da equipe e mitigar possíveis falhas na elaboração do relatório a colaboradora Maria sugeriu a criação de um painel de acompanhamento da Folha de pagamento do Bolsa Família. Maria se dispôs a criar o dashboard pois, assim ela colocaria em pratica o conhecimento adquirido em uma das disciplinas da sua pós-graduação em Análise de Dados. O painel conterà todas as informações que constam hoje, no relatório de folha além disso, será construído rotina de atualização das informações excluindo a necessidade de atualizações manuais.

Sendo assim, como conclusão deste trabalho a equipe vai obter um Data Warehouse, que será a Base para as informações mostradas em dashboard de Power BI, também desenvolvidos neste projeto.



## 2.4 Proposta de Processo de BI

O processo proposto no projeto consiste na criação de um repositório de dados (Data Warehouse) que será alimentado por meio de um processo de ETL, usando como ferramenta o Pentaho Data Integration, e visualização de dados através do Microsoft Power BI. A periodicidade da carga dos dados será semanal, às 00:00h de sábado. Os dados serão referentes as informações da posição de folha mais recente gerada.

**Figura 3.** representação processo de montagem BI

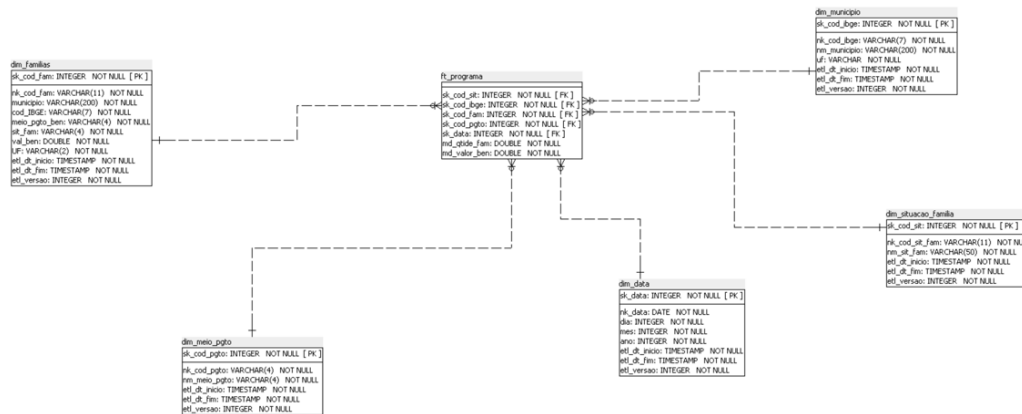


Fonte: Elaborado pelo autor (2024)

### 3. MODELO MULTIDIMENSIONAL

Para este estudo de caso foi elaborado um modelo estrela conforme ilustrado na figura abaixo.

**Figura 4.** Modelo multidimensional Data Warehouse



Fonte: Elaborado pelo autor (2024)

O Modelo será composto por 5 dimensões e a tabela fato:

- Dim\_familia: (sk\_cod\_familiar, nk\_cod\_familiar, municípios, cod\_ibge, uf, meio\_pgto\_ben, sit\_fam, val\_ben, etl\_dt\_inicio, etl\_dt\_fim, etl\_versao);
- Dim\_municipio (sk\_cod\_ibge, nk\_cod\_ibge, nm\_municipio, uf, etl\_dt\_inicio, etl\_dt\_fim, etl\_versao)
- Dim\_meio\_pgto: (sk\_cod\_pgto, nk\_cod\_pgto, etl\_dt\_inicio, etl\_dt\_fim, etl\_versao);
- Dim\_sit\_fam (sk\_cod\_sit, nk\_cod\_sit\_fam, etl\_dt\_inicio, etl\_dt\_fim, etl\_versao);
- Dim\_data (sk\_data, nk\_data, dia, mês, ano, etl\_dt\_inicio, etl\_dt\_fim, etl\_versao);
- Ft\_programa (sk\_cod\_ibge, sk\_cod\_pgto, sk\_cod\_sit, sk\_data, med\_qtde\_fam, med\_valor\_ben).



## **4. ELABORAÇÃO DO DATA WAREHOUSE**

O Data Warehouse será a fonte integradora de informações da empresa, a tecnologia será utilizada com o intuito de servir de base para a camada de aplicação que será responsável por fornecer dados para a tomada de decisão na organização.

### **4.1 Definição do DW**

#### **4.1.1 ARQUITETURA**

A arquitetura do Data Warehouse será INDEPENDENTE, ou seja, vai atender somente as necessidades específicas do departamento não havendo conectividade com outros departamentos.

#### **4.1.2 ABORDAGEM DE CONSTRUÇÃO**

Não será necessário trabalhar com o conceito de Data Marts, pois não vamos subdividir os dados em áreas, e o Data Warehouse será INDEPENDENTE. A granularidade de dados será de baixo nível.

#### **4.1.3 ARQUITETURA FÍSICA**

Arquitetura física utilizada foi On-Premises visto que os servidores ficam na própria gerência e só pessoas lotadas na unidade e na Diretoria terão acesso o painel.

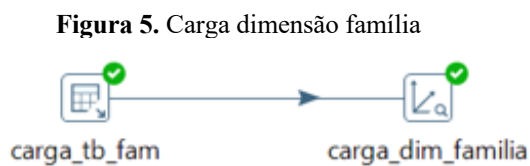
O Processo de atualização de dados semanal do Data Warehouse será executado, em servidor local por meio ferramenta de agendamento já utilizada na empresa este processo executará os processos de ETL e atualizará o Data Warehouse, deixando as informações até segunda feira atualizadas e à disposição da equipe da GEFOL e da Diretoria.

## 5. PROJETO DE ETL

### 5.1 Descrição do Projeto de ETL

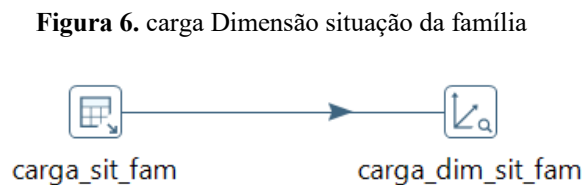
Foi realizada a extração dos dados transacionais para o modelo multidimensional utilizando o software Pentaho Data Integration (PDI).

- Para dimensão família foram realizadas a carga do código familiar sendo que este já havido sofrido tratamento de anonimização na carga do banco transacional, além dos campos ref\_temp, sit\_fam, val\_ben, IBGE, meio\_pgto, município, uf.



Fonte: Elaborado pelo autor (2024)

- Para dimensão situação da família foram carregados apenas dois campos cod\_sit\_fam e tipo\_sit (descrição da situação da família).



Fonte: Elaborado pelo autor (2024)

- Para dimensão meio de pagamento foram carregados os campos código de pagamento e meio de pagamento.

**Figura 7.** Carga da Dimensão meios de pagamento



Fonte: Elaborado pelo autor (2024)

- Para dimensão data o preenchimento da dimensão foi feito script específico em SQL e para fins de cronologia foi considerado o campo `ref_temp` com ano, mês e dia.
- Para dimensão município foram carregados os seguintes campos `códigos_ibge`, `município`, `uf` sendo considerado o IBGE do município onde a família beneficiária reside.

**Figura 8.** Carga dimensão Município



Fonte: Elaborado pelo autor (2024)

Para geração da tabela fato programa foi realizado cálculo para o preenchimento dos campos medidas quantidades de famílias e medidas dos valores pagos. Este cálculo foi realizado a partir da tabela família por meio da função `group by`.

**Figura 9.** Carga Tabela Fato



Fonte: Elaborado pelo autor (2024)

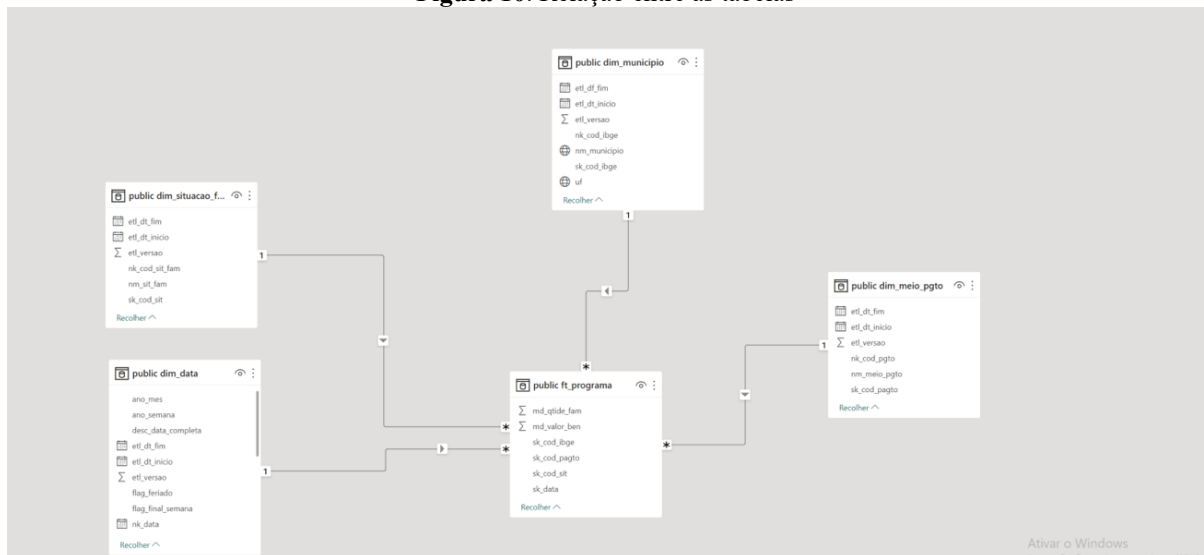
## 6. DASHBOARD

### 6.1 Descrição da Elaboração

O dashboard foi elaborado utilizando a ferramenta Power BI Desktop, o principal objetivo na construção foi é demonstrar de forma clara e simples as principais informações relacionadas a folha de pagamento que são quantitativo de famílias por referência de pagamento, montante a ser pago, situação das famílias e canais de pagamento foram incluídos filtros para que o usuário possa explorar os dados por ano, mês, UF e municípios. O processo de construção do dashboard foi feito sob a supervisão da equipe GEFOL e foi valida pelos gestores.

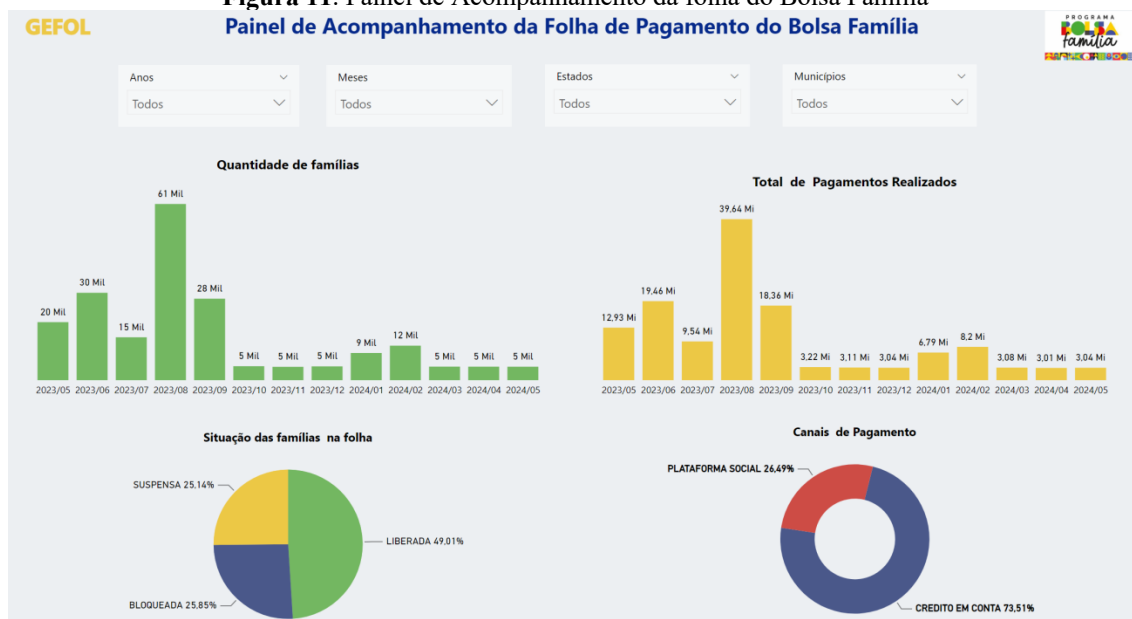
### 6.2 Telas do Dashboard

Figura 10. Relação entre as tabelas



Fonte: Elaborado pelo autor (2024)

**Figura 11. Painel de Acompanhamento da folha do Bolsa Família**  
**Painel de Acompanhamento da Folha de Pagamento do Bolsa Família**



Fonte: Elaborado pelo autor (2024)

## **7. CONCLUSÃO**

O estudo de caso foi realizado com base nas melhores práticas de BI do mercado. Nas tabelas transacionais foram realizados alguns tratamentos de anonimização de dados sensíveis. No Data Warehouse foi realizado a modelagem multidimensional tipo estrela (tabelas dimensões e fato) após a carga das dimensões e fato no Data Warehouse foi desenvolvido o painel.

Para este estudo de caso foram utilizadas as ferramentas Postgres para criação dos bancos de dados, Power Architect para a modelagem do banco transacional e Data Warehouse, Pentaho Data integration para realização dos ETLs e para o criação do Painel foi utilizado o Microsoft Power BI.

A partir da implantação do painel foi possível a exclusão do relatório de folha que era realizado manualmente conforme era esperado pela equipe da GEFOL.

## **8. ANEXOS**

## **9. ARQUIVOS**