



Novas Tecnologias em Banco de Dados - Projeto Prático

André Luiz Bonfatti, 408182

Caroline Pessoa Caetano, 408247

Hitalo Siqueira 408476

1. Escopo do projeto

Nosso projeto será desenvolvido baseado no feedback de consumidores mais especificamente nas reclamações realizadas pelos mesmos de diversos tipos de serviço e produto. Afim de analisarmos desde o tipo de problema que levou o cliente a realizar uma reclamação formal, até como o cliente foi atendido, como por exemplo qual foi o tempo de resposta da empresa fornecedora do bem de consumo.

Este trabalho será baseado em dados governamentais abertos emitidos e publicados oficialmente pelo Governo Federal.

As dimensões do projeto são :

- Faixa Etária (Idade máxima, Idade mínima)
- Tempo (views para dois roles: Data Abertura e Data Fechamento) - Role Playing
- Problema (Categoria, Origem do Contrato que gerou o problema, Área e Assunto do Problema)
- Empresa (Nome Fantasia) - Dimensão Degenerada
- Status (Status da Reclamação) - Dimensão Degenerada

2. Requisitos de negócio

O objetivo deste trabalho é explorar os dados coletados e analisá-los utilizando heurísticas e conceitos de *Data Warehousing* e *Online Analytical Processing*, afim de prover conclusões baseadas em fatos para serem utilizadas como base de tomadas de decisões por órgãos superiores. Tais decisões representam a relação consumidor-empresa e visam melhorar tal relação.

2.1. Consultas e Análise

Para este trabalho, propomos as seguintes análises:

- a. Notas de reclamações separadas e rankeadas por assunto do problema.
- b. Quantidade de reclamações agrupadas por mês e ano que foi aberta.
- c. Notas das reclamações e quantidade separadas por área do problema.
- d. Quantidade de reclamações realizadas no período de feriados importantes.

3. Fonte de dados

Para a criação de um *Data Warehouse* baseado em um banco de dados mais detalhados os dados serão coletados de mais de uma fonte. As duas principais serão os sites www.consumidor.gov.br e <http://dados.gov.br/>.

4. Modelo Banco de Dados Relacional

À partir da análise da fonte de dados foi gerado o modelo relacional de banco de dados, apresentado na **Figura 1**, que será utilizado para gerar nosso *Data Warehouse*.

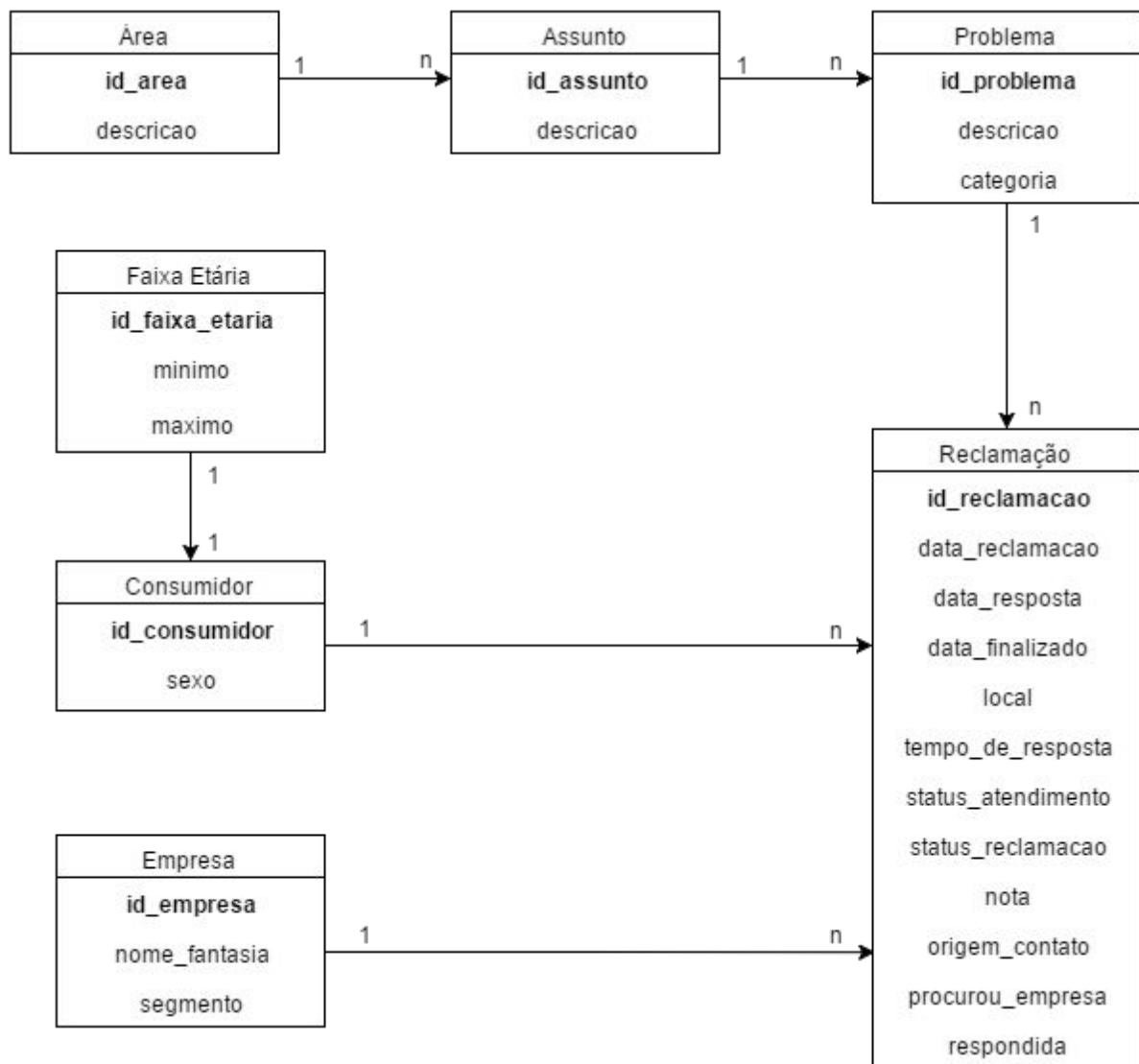


Figura 1. Modelo Entidade Relacionamento Simplificado do Projeto

5. Regras de Transformação

As seguintes regras de transformação serão necessárias para o projeto:

- Remover tuplas que não avaliaram o atendimento (onde o campo 'nota' é nulo).
- Criar desagregações para as tabelas de **Área** e **Assunto**: ambas serão representadas, no *Data Warehouse*, por campos dentro da tabela Problema.
- Criar desagregação para a tabela **Empresa** (nome fantasia) que será transformada em uma dimensão degenerada presente na tabela Fato no DW.

6. Modelo Data Warehouse

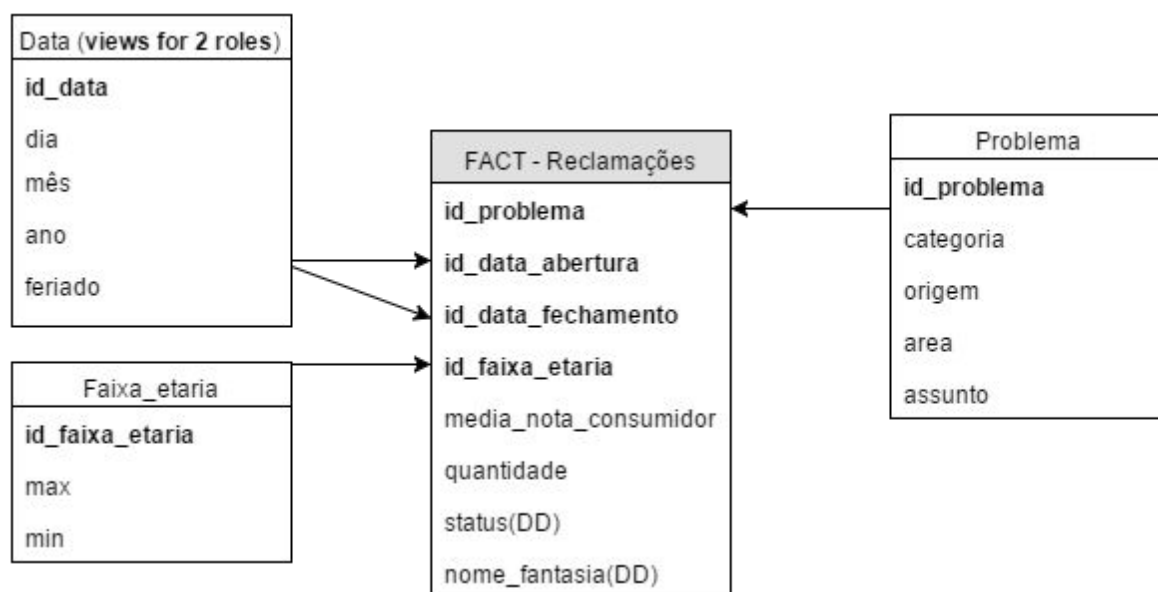


Figura 2. Modelagem 'Star' do Data Warehouse do Projeto

6.1. DW e Referências Gerais com o Modelo Relacional

A modelagem do *Data Warehouse* se assimila com o modelo relacional, centralizando as informações na tabela Fato, **Reclamações**, para que seja possível realizarmos as consultas de negócio definidas no escopo do projeto.

Uma das relações da modelagem do *DW*, porém, não existente inicialmente no modelo relacional é a relação **Tempo**. A mesma é uma dimensão com vários papéis, assim foram criadas duas views uma para Data de Abertura e outra para Data de Fechamento. Estas fontes de dados serão alimentadas na fase de *ETL*, pois são dados pertinentes a um dos requisitos de negócios do projeto.

6.2. Mapeamento de Entidades do DW para o Modelo Relacional

Esta sessão discorre sobre as entidades que são relacionadas entre os modelos *Relacional* e *Data Warehouse*. Todas as entidades, salvo o **Tempo**, são relacionáveis.

Empresa:

Para esta entidade, não seria necessário nenhum mapeamento no *DW*, pois não estamos considerando dados da **Empresa** relevantes para os requisitos

definidos no escopo. Entretanto, a **Empresa** compõe parte extremamente relevante no processo de reclamações, uma vez que é a mesma a prestadora do serviço falho. Por estas razões, manteremos a entidade no *DW*. Mas como o único interesse é o nome fantasia da Empresa essa foi mapeada como uma dimensão degenerada na tabela Fato.

Status:

O status da reclamação faz parte da tabela Reclamação do nosso modelo relacional. No *DW*, esta será uma dimensão degenerada, essencial para a formação de Data Cubes que cumpram os requisitos propostos.

Problema:

Esta entidade é essencial para os requisitos de negócio e é onde está centralizada a maior parte das consultas-chave do projeto.

No modelo relacional, o **Problema** é composto de um **Assunto** e de uma **Área**. No *DW*, estas três entidades são mapeadas para somente uma entidade, a Problema (operação de desagregação realizada durante a fase de ETL).

Consumidor:

Sobre o consumidor a única informação relevante é a Faixa Etária contendo a Idade máxima e mínima, sendo assim a chave da mesma agora também compõe a tabela Fato.

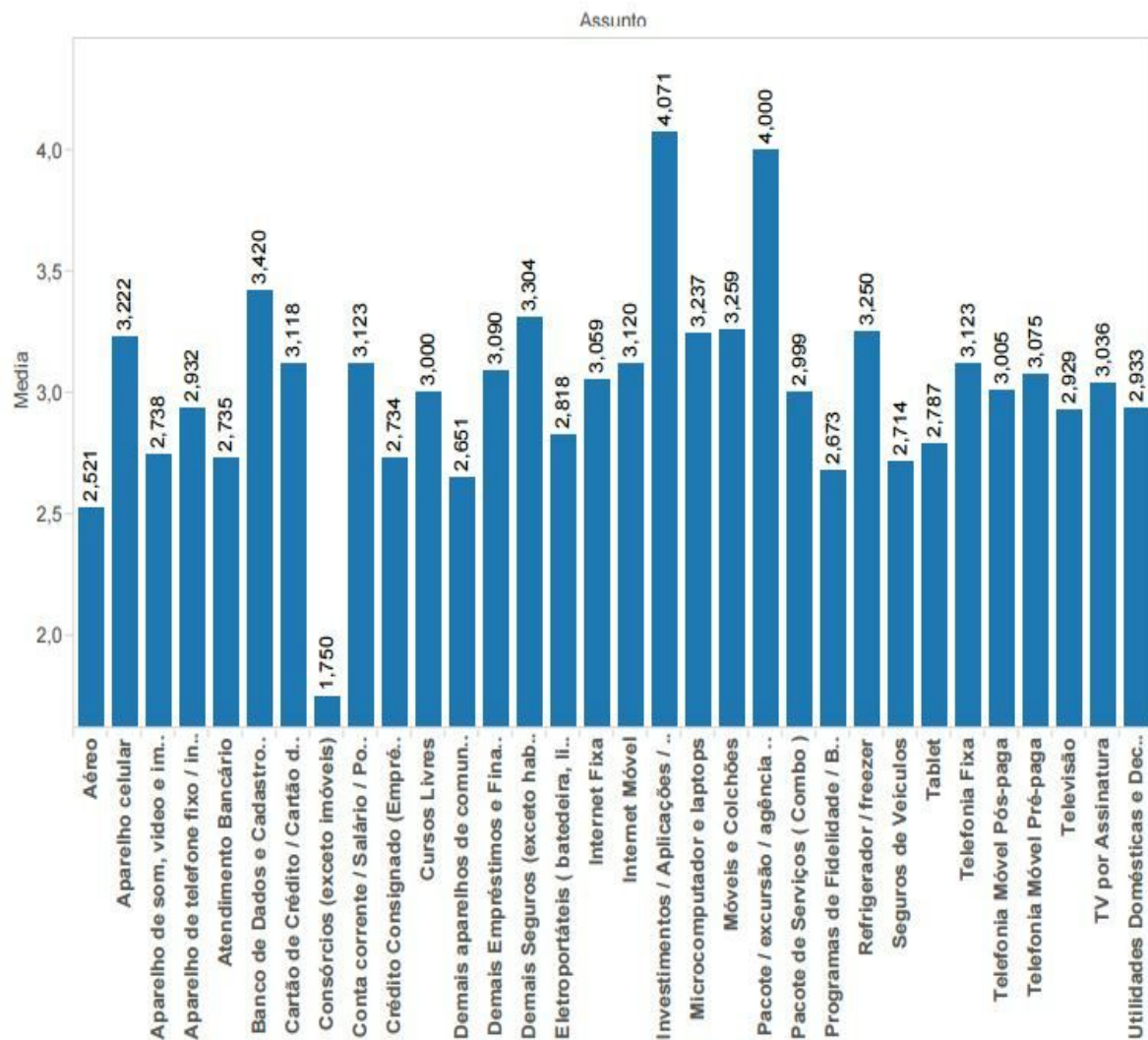
Reclamação - Fato:

Esta é a tabela Fato do projeto e sua transformação para o modelo de *DW* contém algumas características. Primeiramente, serão removidos os atributos desnecessários para os requisitos.

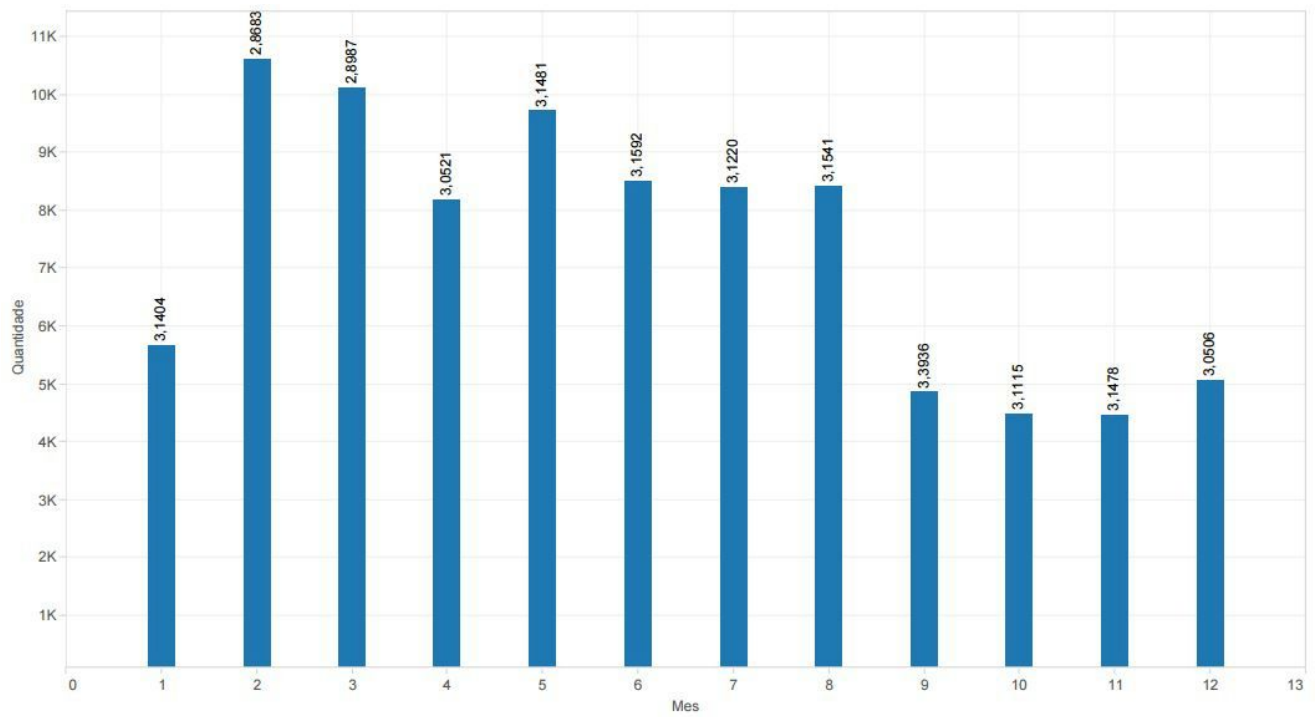
Na tabela Fato do *DW* irão existir dois atributos: (1) a média das notas dadas pelo consumidor sobre o atendimento e (2) quantidade de dados presentes na amostragem.

7. Visualização Agregados

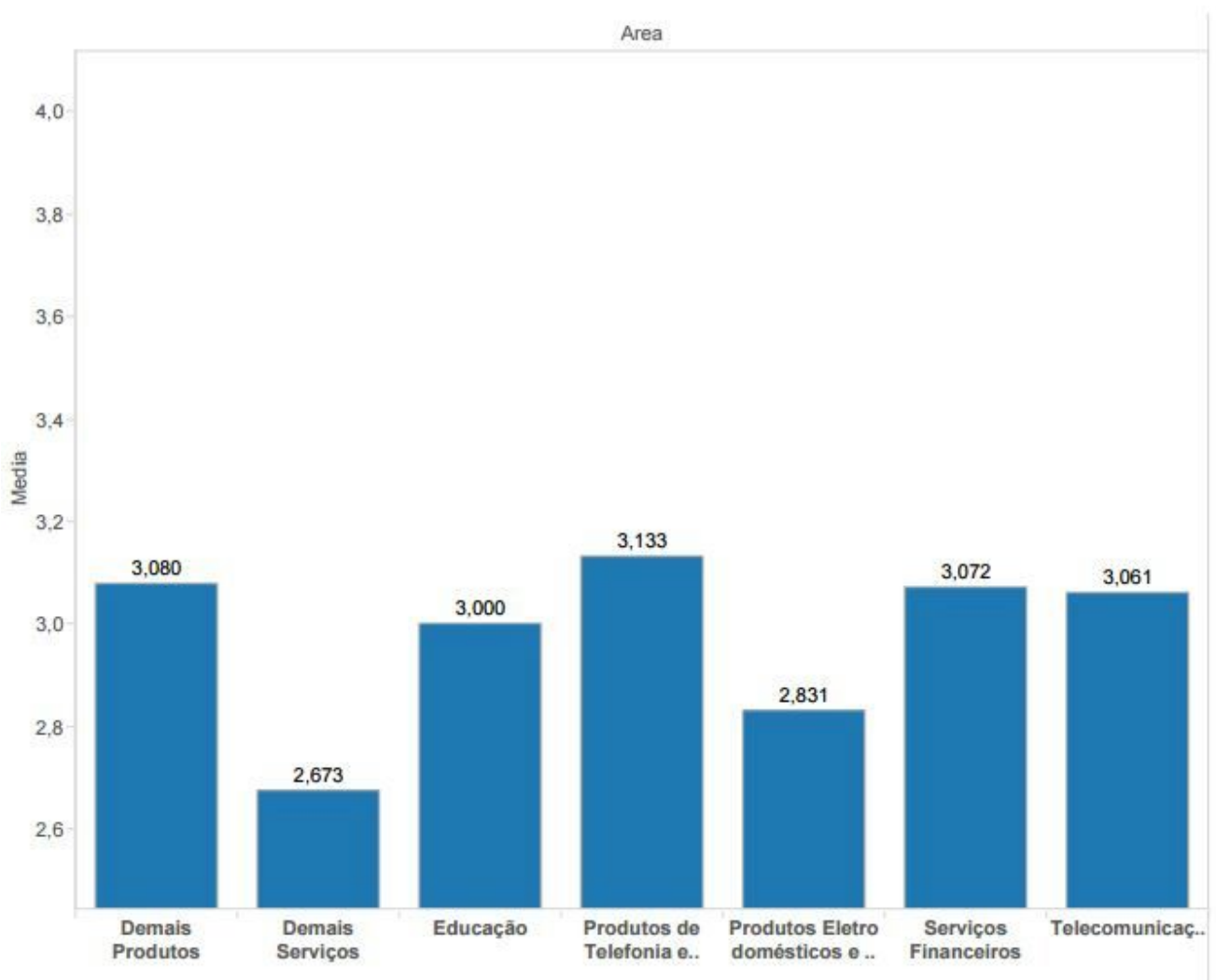
7.1. Notas de reclamações separadas e rankeadas por assunto do problema.



7.2.Quantidade de reclamações agrupadas por mês e ano que foi aberta



7.3. Notas das reclamações e quantidade separadas por área do problema.



7.4.Quantidade de reclamações realizadas no período de feriados importantes.

