



Universidade Federal de Santa Catarina
Centro Tecnológico - CTC
Departamento de Informática e de Estatística - INE

Data Warehouse

Modelagem Dimensional de uma Vídeo locadora

Graduandos:

Fernanda Narloch Rizzo Hahn (18102712)

Henrique Hideki Saito (12203704)

Letícia Amarante dos Santos (20150623)

Matheus Henrique Schaly (18200436)

Nicolas Antero Nunes (17203409)

Professor: José Leomar Todesco, Dr.

Florianópolis
2020

Conteúdo

1	Introdução	1
2	Contexto	2
3	Escopo do projeto	2
3.1	Planejamento do projeto	2
3.2	Requisitos de negócio	2
3.3	Exclusões do escopo do projeto	3
3.4	Fatores Críticos de Sucesso	3
3.5	Riscos	4
3.6	Justificativa	4
4	Identificação do projeto: Análise de Locações	5
5	Equipe	5
6	Modelagem Dimensional	5
6.1	Resposta aos Requisitos de Negócio	6
6.1.1	Tempo médio em que uma mídia se paga	6
6.1.2	Lucratividade das alocações:	7
6.1.3	Momento certo para realizar o descarte de uma mídia .	7
6.1.4	Melhor distribuição das mídias entre as locadoras . . .	8
7	Referências	9

1 Introdução

A empresa de vídeo locadora deseja aprimorar o conhecimento sobre seus próprios dados a fim de aumentar seu lucro. Atualmente há pouca informação sobre os custos envolvidos em cada mídia, sabe-se pouco sobre o tempo médio em que a mídia se paga, sua lucratividade e qual o melhor momento para realizar o descarte da mídia. Além disso, constantemente certas mídias permanecem não alocados em algumas filiais da locadora, enquanto nas demais filiais há demanda excessiva das mesmas mídias.

O projeto de Data Mart irá criar um ambiente de suporte a decisão contendo informações sobre a receita das mídias e seus índices de alocações. As equipes de vendas, marketing e gerenciamento terão acesso às informações para permitir melhor conhecimento acerca das mídias e melhor remanejamento das mídias entre as locadoras, conseqüentemente aumentando o retorno sobre o capital investido na aquisição de cada mídia.

2 Contexto

A video locadora conta atualmente com uma matriz e 5 filiais. São mais de 8 mil clientes e um acervo de mais de 3 mil títulos.

Os tomadores de decisão têm dificuldades de obter informações sobre o seu negócio. Para tal, através de entrevistas realizadas com gerentes e funcionários conseguiu-se ter um melhor entendimento do negócio e definidos requisitos a serem seguidos. Necessita-se saber o tempo médio em que uma mídia adquirida se paga, a lucratividade que se obtêm através das alocações, o momento certo para realizar o descarte de uma mídia e a melhor distribuição das mídias entre as 6 locadoras.

É proposto como solução para o problema apresentado, uma modelagem dimensional para suportar um Data Mart que atenda as demandas de informações e indicadores para apoio a tomada de decisão dessa organização, conforme os requisitos apresentados acima.

3 Escopo do projeto

Atrvés de coletas feitas previamente na empresa por meio de entrevistas, conseguimos reunir informações suficientes para que o escopo do projeto seja definido.

3.1 Planejamento do projeto

- Recursos: 5 indivíduos envolvidos no projeto;
- Estimativa de esforço: 40 horas semanais x 5 membros x 22 semanas = 4400 horas
- Estimativa de data de início: 19 de outubro de 2020;
- Estimativa de data de término: 19 de abril de 2020;
- Situação atual: 5% do trabalho concluído (definição dos requisitos de negócio e modelagem dimensional), esforço necessário para a conclusão: 4000 horas.

3.2 Requisitos de negócio

O projeto do Data Mart deve possibilitar a realização das seguintes análises:

- Tempo médio em que uma mídia se paga;

- A lucratividade das alocações;
- O momento certo para realizar o descarte de uma mídia;
- A melhor distribuição das mídias entre as locadoras;

E também deve permitir as seguintes funcionalidades:

- Filtro por título, data de lançamento, diretor, categoria e censura dos filmes;
- Filtro por locadora;

3.3 Exclusões do escopo do projeto

Não estão inclusos no escopo desse projeto as seguintes informações, pelo fato de tais informações não serem importante :

- Dados referentes a outras empresas de locação de mídias;
- Informações pessoais dos clientes, como nome e CPF;
- Informações referentes aos funcionários;
- Dados específicos sobre pagamentos (número do banco, agência, número do cartão de crédito, etc);

3.4 Fatores Críticos de Sucesso

Para que alcancemos o sucesso do projeto, foram designados os seguintes critérios a serem alcançados:

- Promover fonte única de informações para análise de informações sobre o negócio;
- Aumentar lucratividade em 10%;
- Redução do tempo de geração de relatórios;
- Acelerar o retorno de investimento em 8%;

3.5 Riscos

Falta de entendimento dos requisitos.

Como mitigar:

- Interagir constantemente com o consumidor.
- Priorizar tarefas.

Estimativa imprecisa para entrega do produto.

Como mitigar:

- Considerar que o desenvolvedor pode estar atuando em mais de um projeto.
- Considerar a incerteza da estimativa.

Falta de conhecimento do usuário.

Como mitigar:

- Interagir constantemente com o usuário;
- Realizar testes e pesquisas com o usuário;
- Disponibilizar versão Beta para alguns usuários.

Baixa qualidade dos dados.

Como mitigar:

- Limpar os dados na fonte;
- Limpar os dados durante o processo de ETL.

3.6 Justificativa

Considerando as seguintes 3 técnicas, o projeto é viável e justifica os investimentos a serem feitos com base nos benefícios que trará ao negócio.

- Período de retorno: 1 ano;
- Taxa de retorno interno: 10%;
- Valor Presente Líquido: 170.

Considerando as 3 técnicas o projeto é viável e traz benefícios para o cliente.

4 Identificação do projeto: Análise de Locações

O projeto Análise de Locações (AL) tem como objetivo auxiliar a video-locadora com a implementação de um Data Mart através de uma oferta de uma visão mais analítica do negócio. A seguir estão descritas informações sobre a equipe participante do projeto e um desenvolvimento mais detalhado do plano.

5 Equipe

A equipe é composta por 5 membros, alguns assumindo múltiplos papéis dentro do projeto. Os papéis assumidos por cada membro estão listados abaixo:

- Diretor do projeto: Henrique Hideki Saito
- Gerente do projeto: Fernanda Narloch
- Analista de negócios: Nicolas Antero Nunes
- Arquiteto do DW: Letícia Amarante dos Santos
- Equipe técnica:
 - Desenvolvedores das aplicações: Fernanda Narloch
 - Desenvolvedores de extração, limpeza, transformação e carga dos dados: Nicolas Antero Nunes
 - Educadores: Matheus Henrique Schaly
 - Suporte técnico: Matheus Henrique Schaly

6 Modelagem Dimensional

Apresentação e explicação da modelagem dimensional desenvolvida de forma a responder as necessidades do cliente.

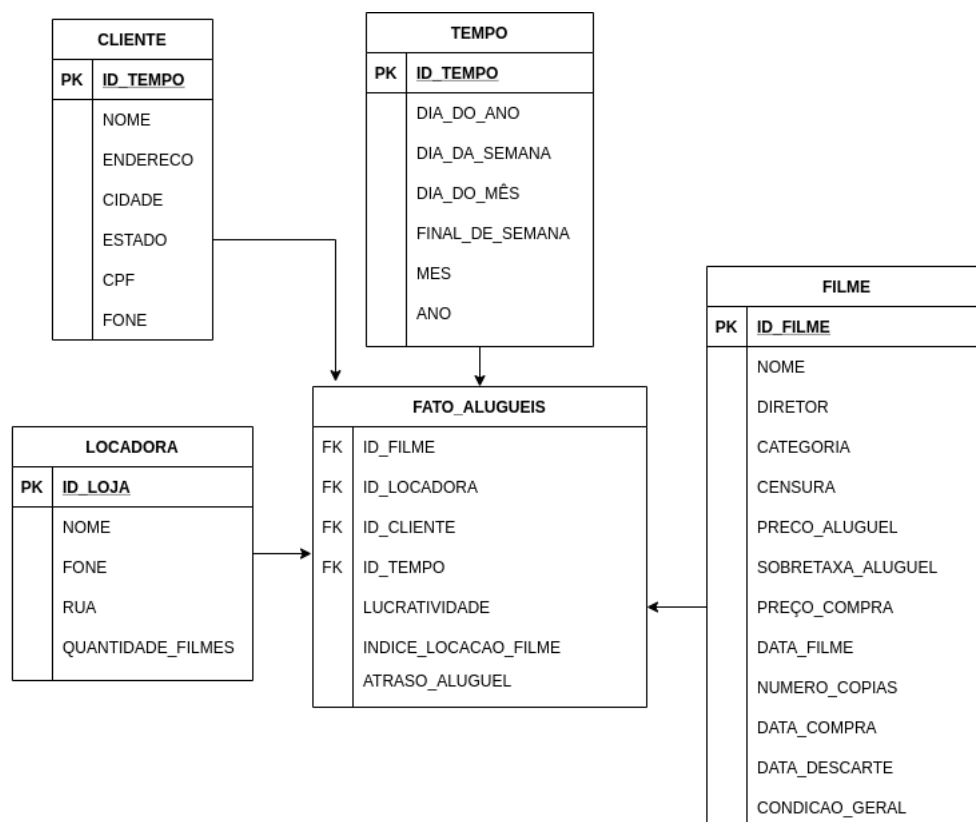


Figura 1: Modelagem Dimensional

6.1 Resposta aos Requisitos de Negócio

6.1.1 Tempo médio em que uma mídia se paga

Campo utilizado:

- FILME.ID_FILME
- FATO.LUCRATIVIDADE
- FILME.DATA_COMPRA

Como calcular: Quando a inserção de um aluguel resultar em uma Lucratividade 0 ou positiva, quer dizer que a mídia se pagou. E, com isso, sabemos que o tempo em que a mídia se pagou é igual a diferença da data da inserção com a data da compra da mídia. Já a média do tempo pode ser feita a partir de uma média simples de todas as mídias filtradas por locadora.

6.1.2 Lucratividade das alocações:

Campo utilizado:

- FILME_ID_FILME
- FILME_PRECO_ALUGUEL
- FILME_SOBRETAXA_ALUGUEL
- FILME_PRECO_COMPRA
- FATO_ATRASO_ALUGUEL

Medida criada:

- FATO_LUCRATIVIDADE

Como calcular:

- $$\begin{aligned} \text{FATO_LUCRATIVIDADE} = & \text{SUM}(\text{FILME_PRECO_ALUGUEL} \text{ filtrado pelo FILME_ID_FILME}) + \\ & \text{SUM}(\text{FILME_SOBRETAXA_ALUGUEL} \text{ filtrado pelo FILME_ID_FILME}) \times \\ & \text{FATO_ATRASSO_ALUGUEL} - \\ & \text{FILME_PRECO_COMPRA} \text{ filtrado pelo FILME_ID_FILME} \end{aligned}$$

Isto é, a soma dos preços pagos tanto por aluguel quanto por sobretaxa menos o custo na compra do filme.

6.1.3 Momento certo para realizar o descarte de uma mídia

Campo utilizado:

- FILME_ID_FILME
- FILME_NUMERO_COPIAS
- TEMPO_MES

Índice criado:

- FATO_INDICE_LOCACAO_FILME

Como calcular:

- $$\text{FATO_INDICE_LOCACAO_FILME} = \frac{\text{COUNT (filtrado pelo FILME_ID_FILME e pelo mês atual TEMPO_MES)}}{\text{FILME_NUMERO_COPIAS} \times 30 \text{ filtrado pelo FILME_ID_FILME}}$$

Quando o índice for menor que 0.2 podemos descartar o filme. Isto é, o filme está sendo alugado menos de 6 vezes ao mês (20% de 30). Temos a atualização desse índice ao final de cada mês

6.1.4 Melhor distribuição das mídias entre as locadoras

Campo utilizado:

- FILME_FILME_ID
- FILME_NUMERO_COPIAS
- TEMPO_MES

Índice já existente:

- FATO_INDICE_LOCACAO_FILME

Quanto maior o valor do índice, mais filmes estão sendo alugados e/ou menor a quantidade de cópias do filme. Podemos realizar o ranking desse índice e mover o filme da locadora com menor índice para o de maior índice. Só poderíamos mover o filme se a locadora de menor índice possui mais de uma cópia do filme. E podemos realizar tal movimentação apenas se a locadora de menor índice possuir o valor do índice menor que 0.5 e a locadora de maior índice possuir um valor de índice maior que 0.90.

7 Referências

- <https://docente.ifrn.edu.br/josecunha/disciplinas/adbd/pdfs/data-warehouse>
- The Data Warehouse Lifecycle Toolkit KIMBALL, R. et al. The Data Warehouse Lifecycle Toolkit.