

**UNIVERSIDADE FEDERAL DE SANTA CATARINA  
DEPARTAMENTO DE INFORMÁTICA E ESTATÍSTICA**

Giancarlo Souza de Freitas, Ronan Romeu Knob, Sabrina Schütz de  
Oliveira, Valdir Luiz Hofer Arnhold

**PROPOSTA PARA DESENVOLVIMENTO DE UM DATA  
WAREHOUSE**

Florianópolis

2017



Giancarlo Souza de Freitas, Ronan Romeu Knob, Sabrina Schütz de  
Oliveira, Valdir Luiz Hofer Arnhold

## **PROPOSTA PARA DESENVOLVIMENTO DE UM DATA WAREHOUSE**

Trabalho submetido ao Curso de Sistemas  
de Informação para a obtenção da apro-  
vação parcial na disciplina de Data Wa-  
rehouse.

Orientador: Prof. Jose Leomar Todesco,  
Dr.

Florianópolis

2017



## **LISTA DE FIGURAS**

Figura 1	Dimensões. ....	16
Figura 2	Fato - Versão 1.....	16
Figura 3	Fato - Versão 2.....	17
Figura 4	Resposta da pergunta 1. ....	20
Figura 5	Resposta da pergunta 2. ....	21



## LISTA DE TABELAS

Tabela 1	<i>Payback Period Analysis</i> .....	10
Tabela 2	Riscos. ....	11
Tabela 3	Recursos Humanos.....	13
Tabela 4	Estimativa de tempo.....	14
Tabela 5	Indicadores.....	19
Tabela 6	Variáveis.....	19





## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>9</b>
1.1	CONTEXTUALIZAÇÃO .....	9
1.2	PERGUNTAS ESTRATÉGICAS .....	9
1.3	ESCOPO .....	9
1.4	JUSTIFICATIVA .....	10
1.5	FATORES CRÍTICOS DE SUCESSO .....	11
1.6	EXCLUSÕES DE ESCOPO .....	11
1.7	RISCOS .....	11
<b>2</b>	<b>PLANEJAMENTO .....</b>	<b>13</b>
2.1	DEFINIÇÃO DA EQUIPE .....	13
2.2	CRONOGRAMA .....	13
2.3	CUSTOS .....	14
<b>3</b>	<b>MODELAGEM DIMENSIONAL .....</b>	<b>15</b>
3.1	PROCESSO DO NEGÓCIO .....	15
3.2	GRANULARIDADE .....	15
3.3	DIMENSÕES .....	15
3.4	FATOS .....	16
<b>4</b>	<b>INDICADORES .....</b>	<b>19</b>
4.1	EXEMPLOS DE RELATÓRIOS .....	20
<b>4.1.1</b>	<b>Pergunta 1 .....</b>	<b>20</b>
<b>4.1.2</b>	<b>Pergunta 2 .....</b>	<b>21</b>
<b>5</b>	<b>CONCLUSÃO .....</b>	<b>23</b>



# 1 INTRODUÇÃO

## 1.1 CONTEXTUALIZAÇÃO

A locadora Ronbuster, fundada em 1972, possui 4 filiais na grande Florianópolis e, preocupada com o crescimento dos meios digitais de transmissão de filmes, decidiu implantar uma data mart para analisar o histórico de vendas, buscando novas estratégias para continuar crescendo no mercado.

O público da locadora é composto em sua maioria, por pessoas acima de 40 anos. Seu acervo é composto, em grande parte, por DVD's e Blu-Rays. Também possui em acervo VHS, e outras mídias, mas em pequena quantidade.

As categorias de preço variam em função da data de chegada da mídia na locadora. Ao chegar na locadora, uma mídia é classificada como lançamento, e o empréstimo deve ser retornado em até 24 horas, por um preço de 8 reais. Depois de 60 dias do lançamento, a categoria da mídia muda para “vermelha”, e sua devolução pode ser feita em até 72 horas após o empréstimo, pelo valor de 6 reais. Passados 120 dias na categoria vermelha, a mídia passa para a categoria “verde” e o seu empréstimo pode ser devolvido em até 96 horas, com valor de 4 reais.

## 1.2 PERGUNTAS ESTRATÉGICAS

As perguntas estratégicas levantadas para o desenvolvimento do Data Mart são as seguintes:

1. Em quanto tempo uma mídia física se paga?
2. Qual é o fluxo de empréstimos e devoluções das mídias por dia da semana?
3. Qual o número ideal de cópias de cada título por filial?

## 1.3 ESCOPO

O escopo do projeto consiste em planejar e desenvolver um Data Mart para a locadora Ronbuster e suas filiais. De forma específica, para o desenvolvimento do mercado e a geração de receita. O número máximo de usuários

suportados pelo sistema, para realizar a análise dos resultados gerados pelo *Data Mart* é de 10 pessoas.

#### 1.4 JUSTIFICATIVA

A gestão da informação, quando feita de forma correta, pode trazer inúmeros benefícios para as organizações. Com a implantação do Data Mart na locadora Ronbuster, se busca aumentar o lucro sobre as locações de mídias do acervo, gerando um maior fluxo de caixa para empresa. Além do mais, numa sociedade que cresce em competitividade a cada dia, um data mart pode ser um diferencial no mercado para responder rapidamente às demandas e tendências do mercado.

Para estimar o retorno do projeto, em um período anual, foi utilizado a técnica simples do *Payback Period Analysis*, onde foi feita uma estimativa do retorno estimado ao ano de impacto nas vendas, e diluído o valor do projeto em parcelas deste retorno anual, conforme é mostrado na Tabela 1.

Tabela 1 – *Payback Period Analysis*.

<b>Custo estimado do projeto do Ronbuster <i>Data Mart</i></b>	R\$ 52.000,00
<b>Custo do retorno anual</b>	R\$ 37.000,00
<b><i>Payback Period Analysis</i></b>	1,4 anos

Fonte: Os autores (2017).

Como visto na tabela acima, com a implantação do data mart na rede de locadoras Ronbuster foi calculado uma receita a mais de R\$13.500,00 para a empresa a cada ano. Isso é, após 2,6 anos, que é o período de Payback, a empresa lucrará 13,45% a mais, anualmente, pelo uso da ferramenta, considerando o levantamento de R\$ 275.000,00 aproximadamente de receita das filiais da locadora juntas, anualmente.

Além do próprio projeto se pagar num tempo aceitável, o valor das informações sobre o negócio, mesmo que não mensurados e usados no cálculo, devem ser levados em conta. Sendo assim, o projeto de instalação do data mart se torna viável e traz um bom retorno em médio prazo para a empresa.

## 1.5 FATORES CRÍTICOS DE SUCESSO

Os fatores críticos de sucesso, inicialmente levantados, são os seguintes:

- Prover uma fonte única de informações sobre os processos de negócio da vídeo locadora.
- Aumentar a eficácia das locações, em 25% , nos períodos de promoções.

## 1.6 EXCLUSÕES DE ESCOPO

Nesta etapa de desenvolvimento não será feita uma integração de dados com sistemas de terceiros. Além disso, na primeira versão do projeto não vai ser disponibilizado ferramentas para descoberta de dados utilizando data mining.

## 1.7 RISCOS

Os possíveis riscos levantados e seu plano de resposta e mitigação são apresentados na Tabela 2.

Tabela 2 – Riscos.

<b>Risco</b>	<b>Probabilidade</b>	<b>Impacto</b>	<b>Estratégia de resposta</b>	<b>Estratégia de mitigação</b>
Alteração do escopo	Alta	Alto	Redefinição do tempo	Validação dos requisitos iniciais
Dificuldade técnica	Média	Alto	Contratação de especialista	Capacitação técnica para a equipe

Fonte: Os autores (2017).



## 2 PLANEJAMENTO

### 2.1 DEFINIÇÃO DA EQUIPE

Os recursos humanos e seus respectivos papéis envolvidos no projeto são definidos na Tabela 3 .

Tabela 3 – Recursos Humanos.

<b>Papel</b>	<b>Recurso Humano</b>
Diretor de DW	Ronan Romeu Knob
Gerente de projetos de DW	Sabrina Schütz de Oliveira
Analista de negócios	Giancarlo Souza de Freitas
Arquiteto de DW	Giancarlo Souza de Freitas
Equipe técnica	Valdir Luiz Hofer Arnhold

Fonte: Os autores (2017).

### 2.2 CRONOGRAMA

O projeto de *Data Mart* Ronbuster foi pensado para ter início no mês de março. Sua duração estimada foi estipulada em 6 meses, contemplando as fases apresentadas na tabela x. A escolha deste período de início foi devido a que é o período de término das férias escolares, o que deve baixar o movimento das filiais, estando pronto também antes do final do ano, novo grande ciclo de férias.

Sendo assim, as datas iniciais de início e término do projeto são:

- Início do projeto - 01/03/2017;
- Término do projeto - 31/08/2017;
- Dias úteis neste período: 129; e
- Horas do projeto: 1032h

Abaixo temos uma estimativa das fases do projeto, e seu tempo esperado. Os valores apresentados não consideram feriados e sábados e domingos. Cálculo da hora é feito por dias úteis \* 8h.

Tabela 4 – Estimativa de tempo.

<b>Fase</b>	<b>Data de início</b>	<b>Data de término</b>	<b>Duração</b>
Planejamento do projeto	01/03/2017	24/03/2017	144 h
Entrevistas	27/03/2017	07/04/2017	80 h
Definição do esquema estrela e plano de ação	10/04/2017	20/04/2017	64 h
Definição da equipe do DW	24/04/2017	26/04/2017	24 h
Implementação do DW	27/04/2017	09/06/2017	248 h
Criação de portal de dashboards e relatórios	12/06/2017	04/08/2017	320 h
Treinamento	07/08/2017	11/08/2017	40 h
Homologação e ajustes pontuais	14/08/2017	31/08/2017	112 h
<b>Total</b>			<b>1032 h</b>

Fonte: Os autores (2017).

### 2.3 CUSTOS

Com base no Cronograma mostrado e nas horas estimadas o custo projeto é de R\$ 52.000,00.



### 3 MODELAGEM DIMENSIONAL

Para a construção do modelo dimensional, 4 passos são necessários, sendo eles:

1. Decisão de qual processo de negócio deve-se modelar, com base na combinação do conhecimento do negócio com o conhecimento dos dados disponíveis.
2. Definição do grão do processo do negócio, o qual é o nível fundamental atômico de dados que representará o processo na tabela de fatos.
3. Escolha das dimensões que serão aplicadas a cada registro da tabela de fatos.
4. Escolha dos fatos mensuráveis que irão popular cada registro da tabela de fatos.

A seguir, temos a utilização desses passos para a criação da modelagem dimensional em questão, para auxiliar a resolução das perguntas estratégicas levantadas.

#### 3.1 PROCESSO DO NEGÓCIO

O processo do negócio que será modelado é o movimento diário de item, nos permitindo acompanhar quais filmes estão sendo emprestados, em que lojas, a que preço e em que dias.

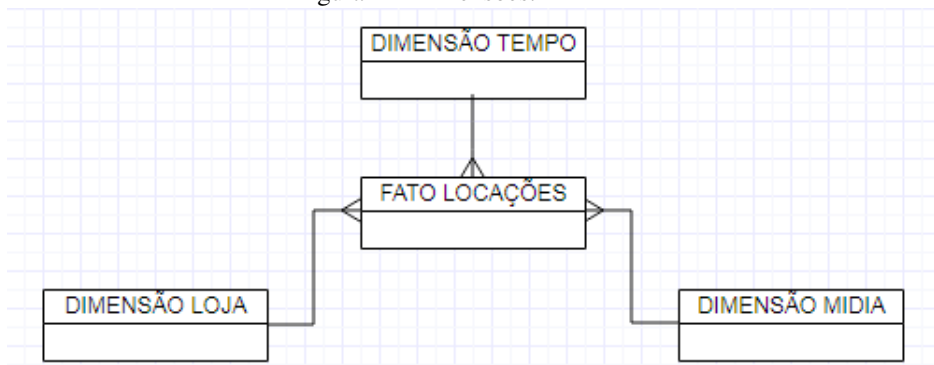
#### 3.2 GRANULARIDADE

A granularidade, representa o nível de detalhamento da modelagem. Para este projeto, se buscou alcançar o maior nível de granularidade possível. Por exemplo, para o mapeamento da dimensão do tempo é utilizado a hora, dessa forma se torna possível fazer consultas por dia ou semanas.

#### 3.3 DIMENSÕES

As dimensões escolhidas são Tempo, Midia e Loja , conforme mostra a Figura 1.

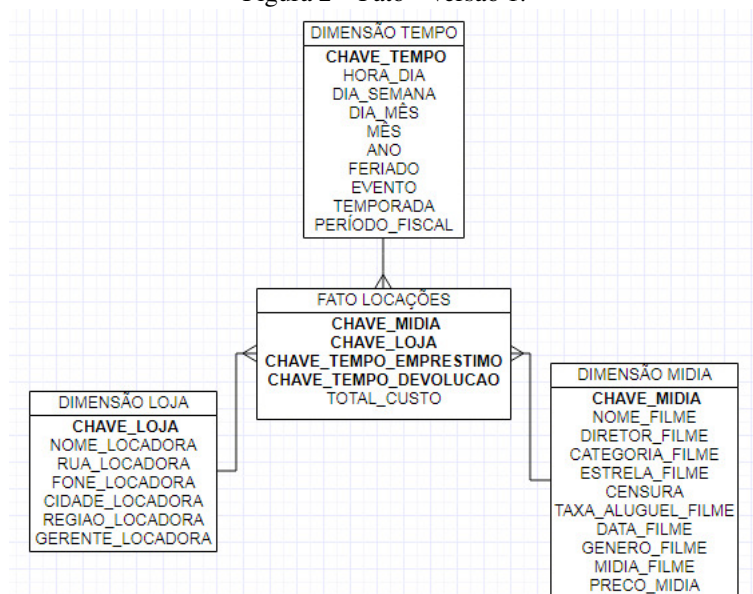
Figura 1 – Dimensões.



### 3.4 FATOS

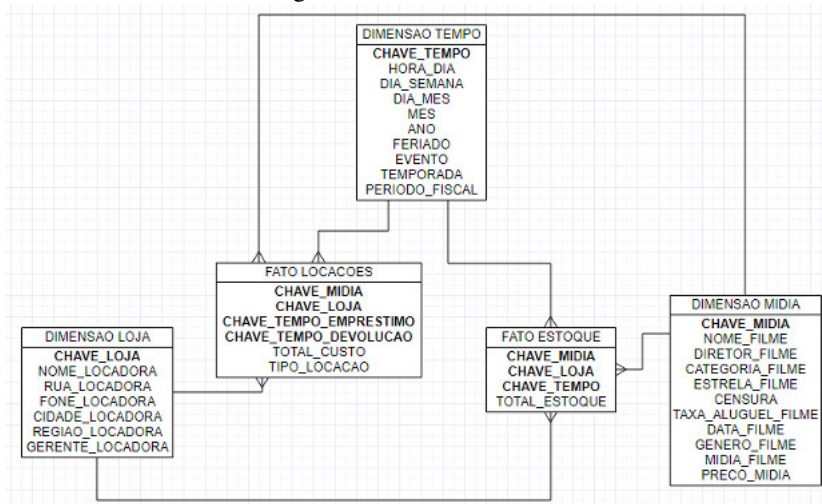
Inicialmente foi modelada a seguinte a tabela de fatos de Locações, conforme mostra a Figura 2.

Figura 2 – Fato - Versão 1.



"Há somente uma garantia para qualquer projeto de DW e é a que ele será mudado." Depois de analisarmos novamente a modelagem, e com as perguntas estratégicas em mente, percebemos que uma nova tabela fato deveria ser adicionada para a resolução dessas perguntas, conforme é mostrado na Figura 5.

Figura 3 – Fato - Versão 2.





## 4 INDICADORES

Para verificar se as perguntas estratégicas levantadas no projeto da locadora Ronbuster estão sendo respondidas, é necessário criar indicadores. Os indicadores visam monitorar as variáveis necessárias para solucionar uma pergunta estratégica, e criando informação a partir delas, que responde aos questionamentos da gestão da empresa. Os indicadores são chamados indicadores estratégicos, que são os indicadores que podem dar ampla visão dos objetivos e metas mais gerais da empresa, bem como permitir a comparação de ações anteriores com resultados atuais.

Quanto a locadora Ronbuster, para responder as perguntas que criamos, foram levantados os indicadores mostrados na Tabela 5 e as variáveis mostradas na Tabela 6.

Tabela 5 – Indicadores.

Pergunta	Indicador
Em quanto tempo uma mídia física se paga?	$(V1 \cdot v4) / (V2 \cdot V3)$
Qual é o fluxo de empréstimos e devoluções das mídias por dia da semana?	$(V5 \cdot 24)$ e $(V5 \cdot 24)$
Qual o número ideal de cópias de cada título por filial?	V7

Fonte: Os autores (2017).

Tabela 6 – Variáveis.

V1 - Valor da locação de uma mídia
V2 - Custo da mídia (custo de compra de uma unidade)
V3 - Quantidade de mídias compradas
V4 - Número de locações de uma mídia
V5 - Número de locações por hora
V6 - Número de devoluções por hora
V7 - Número de títulos disponíveis por hora do título

Fonte: Os autores (2017).

## 4.1 EXEMPLOS DE RELATÓRIOS

### 4.1.1 Pergunta 1

A consulta para responder a pergunta 1, pode ser como acima. Nela, é apresentado o título da mídia, preço desta. Para resolver quanto tempo, foi calculado a média do custo da locação, considerando que pode haver diferença de custo por algum motivo, dividido pela média de dias que ela esteve locada. Considerando então uma locação de 24h, ou seja, 1 locação por dia, temos, dividindo o preço por esta média, o número de dias em que a mídia se pagará.

Figura 4 – Resposta da pergunta 1.

```

1  SELECT
2
3  BASE.TITULO AS 'TITULO MIDIA',
4  BASE.PRECO_MIDIA AS 'PREÇO DA CÓPIA',
5  BASE.MEDIA_CUSTO_LOCAÇÃO AS 'MÉDIA CUSTO LOCAÇÃO',
6  BASE.PRECO_MIDIA / BASE.DIAS_LOCAÇÃO_MIDIA AS 'QTD DE DIAS PARA SE PAGAR MIDIA'
7
8  FROM(
9
10     SELECT
11
12     DIMENSAO_MIDIA.TITULO AS TITULO,
13     DIMENSAO_MIDIA.PRECO_MIDIA AS PRECO_MIDIA,
14     AVG(FATO_LOCACOES.TOTAL_CUSTO) AS MEDIA_CUSTO_LOCAÇÃO
15     SUM(DIFERENÇA CHAVE_TEMPO_DEVOLUÇÃO(CONCATENA(DIA+MES+ANO+HORA)
16         - CHAVE_TEMPO_EMPRESTIMO.(CONCATENA(DIA+MES+ANO+HORA))))/24 AS DIAS_LOCAÇÃO_MIDIA
17
18     FROM FATO_LOCACOES
19     JOIN DIMENSAO_TEMPO EMPRESTIMOS ON FATO_LOCACOES.CHAVE_TEMPO_EMPRESTIMO = DIMENSAO_TEMPO.CHAVE_TEMPO
20     JOIN DIMENSAO_TEMPO DEVOLUCOES ON FATO_LOCACOES.CHAVE_TEMPO_DEVOLUCOES = DIMENSAO_TEMPO.CHAVE_TEMPO
21     JOIN DIMENSAO_MIDIA ON FATO_LOCACOES.CHAVE_MIDIA = DIMENSAO_MIDIA.CHAVE_MIDIA
22     JOIN DIMENSAO_LOJA ON FATO_LOCACOES.CHAVE_LOJA = DIMENSAO_LOJA.CHAVE_LOJA
23
24     GROUP BY
25     DIMENSAO_MIDIA.TITULO,
26     DIMENSAO_MIDIA.PRECO_MIDIA,
27     FATO_LOCACOES.TOTAL_CUSTO,
28     DIMENSAO_TEMPO.DIA_MES,
29     DIMENSAO_TEMPO.MES,
30     DIMENSAO_TEMPO.ANO,
31     DIMENSAO_TEMPO.HORA_DIA
32
33 ) AS BASE

```

As funções de diferença de tempo e concatenação de campos para formar as datas foram omitidas por apresentarem peculiaridades técnicas pertencentes a cada SGBD.

### 4.1.2 Pergunta 2

Para a pergunta estratégica 2: Qual é o fluxo de empréstimos e devoluções das mídias por dia da semana?

Figura 5 – Resposta da pergunta 2.

```

1  SELECT
2
3  DIMENSAO_MIDIA.NOME_FILME AS TITULO,
4  DIMENSAO_LOJA.NOME_LOCADORA AS FILIAL,
5  COUNT(*) AS NUMERO_EMPRESTIMO,
6  DIMENSAO_TEMPO.DIA_SEMANA AS DIA_SEMANA
7
8
9  FROM FATO_LOCACOES
10 JOIN DIMENSAO_TEMPO ON FATO_LOCACOES.CHAVE_TEMPO_EMPRESTIMO = DIMENSAO_TEMPO.CHAVE_TEMPO
11 JOIN DIMENSAO_MIDIA ON FATO_LOCACOES.CHAVE_MIDIA = DIMENSAO_MIDIA.CHAVE_MIDIA
12 JOIN DIMENSAO_LOJA ON FATO_LOCACOES.CHAVE_LOJA = DIMENSAO_LOJA.CHAVE_LOJA
13 WHERE FATO_LOCACOES.TIPO_LOCACAO = 'EMPRESTIMO'
14
15 GROUP BY
16 DIMENSAO_MIDIA.NOME_FILME,
17 DIMENSAO_LOJA.NOME_LOCADORA,
18 DIMENSAO_TEMPO.DIA_SEMANA

```

Essa pesquisa retorna os dias da semana e a quantidade de cópias emprestadas, separados por título e filial. A agregação poderia ser por qualquer unidade de tempo. Para saber a quantidade de devoluções, como a pergunta 2 também exige, precisaria apenas mudar, na linha 13, a palavra **'EMPRESTIMO'** para **'DEVOLUCAO'**, e a chave da tabela de **'FATO\_LOCACOES'** para **'CHAVE\_TEMPO\_DEVOLUCAO'**.





## **5 CONCLUSÃO**

De forma geral, conforme mostrado nas seções anteriores, com a modelagem dimensional, os indicadores criados e os modelos de consulta, é possível responder as perguntas estratégicas elaboradas. Com isso, se esperar aumentar a rentabilidade da organização através de uma fonte única de dados, organizados com qualidade.