

UNIVERSIDADE DO MINHO

DEPARTAMENTO DE INFORMÁTICA

APLICAÇÕES INFORMÁTICAS NA BIOMEDICINA

---

## Trabalho Prático

Bernardo Silva (a77230)  
César Ribeiro (a79014)  
Mariana Fernandes (a81728)  
Paulo Jorge Mendes(a78203)  
Raul Vilas Boas (a79617)

---

19 de Dezembro de 2019

# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>2</b>
<b>2</b>	<b><i>Data Warehouse</i></b>	<b>2</b>
2.1	Povoamento com SQL . . . . .	3
2.2	Povoamento no Talend . . . . .	3
2.3	Vantagens e Desvantagens . . . . .	4
<b>3</b>	<b><i>Indicadores Clínicos</i></b>	<b>5</b>
3.1	Indicador Clínico 1 . . . . .	5
3.2	Indicador Clínico 2 . . . . .	6
3.3	Indicador Clínico 3 . . . . .	7
<b>4</b>	<b><i>Interface com Utilizador</i></b>	<b>8</b>
<b>5</b>	<b><i>Conclusão</i></b>	<b>9</b>

# 1 Introdução

Este relatório tem como finalidade expor a proposta de resolução do trabalho prático. O objectivo deste trabalho é desenvolver um *data warehouse* com dados provenientes de um serviço de urgência, para posteriormente extrair conhecimento fazendo uso de ferramentas de *Business Intelligence*.

Para tal, foi-nos fornecido um CSV com informação variada sobre idas ao serviço de urgência em questão, no entanto, no mesmo não estão presentes dados confidenciais ou identificadores dos pacientes.

## 2 Data Warehouse

A primeira tarefa a realizar para criar o *data warehouse* é a importação dos dados fornecidos em CSV. Para tal, foi criada uma tabela *urg\_inform\_geral*, na base de dados *bd\_urg*, onde cada registo dessa tabela é uma linha do CSV.

De seguida, construiu-se um modelo dimensional em estrela, constituído pelas tabelas de dimensão *dim\_sexo*, *dim\_des\_causa*, *dim\_alta\_des\_especialidade*, *dim\_des\_proveniencia*, *dim\_des\_local* e *dim\_data* e a tabela de factos *fact\_urgencia*. Através da ferramenta de

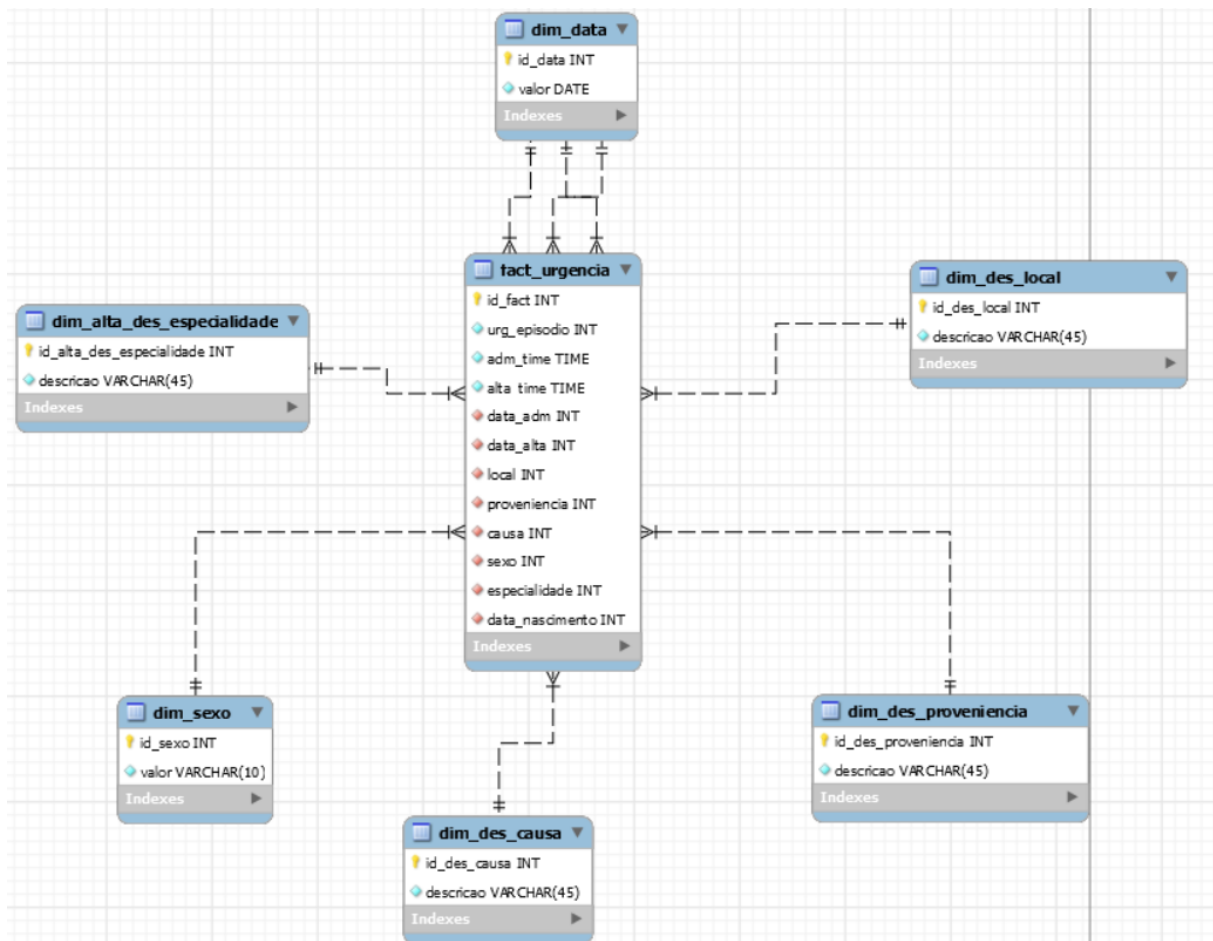


Fig. 1: Modelo em estrela

*forward engineer* do MySQL Workbench, foi gerado o modelo físico do *data warehouse* *dw\_urg*.

## 2.1 Povoamento com SQL

Em primeiro lugar, foi feito o povoamento das tabelas de dimensão, e só depois foi feito o da tabela de factos, para que nesta possam ser usadas as chaves estrangeiras que são chaves primárias nas tabelas de dimensão.

Este procedimento foi feito através de um *script* em SQL que lê da tabela *bd\_urg.urg\_inform\_geral* e insere nas respectivas tabelas de dimensão. De seguida o *script* povoa a tabela de factos agregando os dados da tabela original e das tabelas de dimensão, inserindo apenas as FK nos devidos casos.

## 2.2 Povoamento no Talend

Para fazer o povoamento da *data warehouse* com o Talend, foi necessário criar jobs específicos para tal, como podemos ver nas figuras seguintes. Para preencher a tabela de factos, em vez de uma inserção directa na tabela *fact\_urgencia*, tivemos de efectuar o equivalente ao inner join no mysql, utilizando para isso o tMap, ferramenta do Talend.

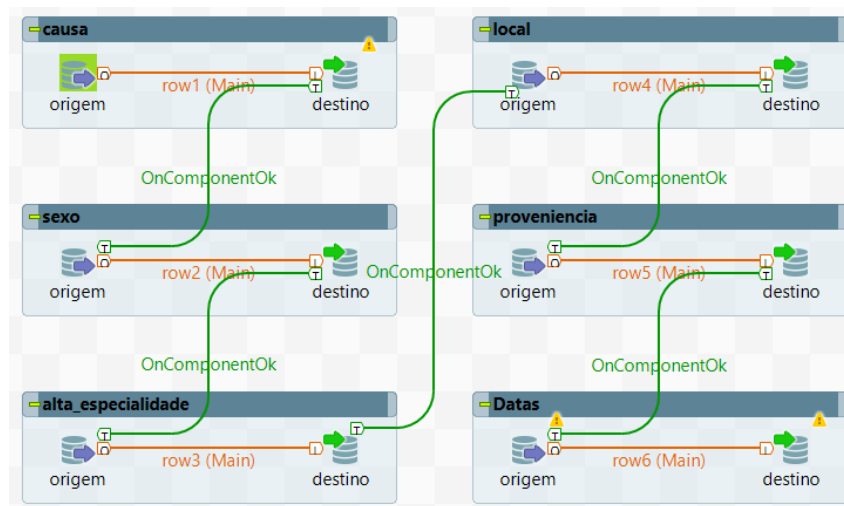


Fig. 2: Povoamento das tabelas de dimensão

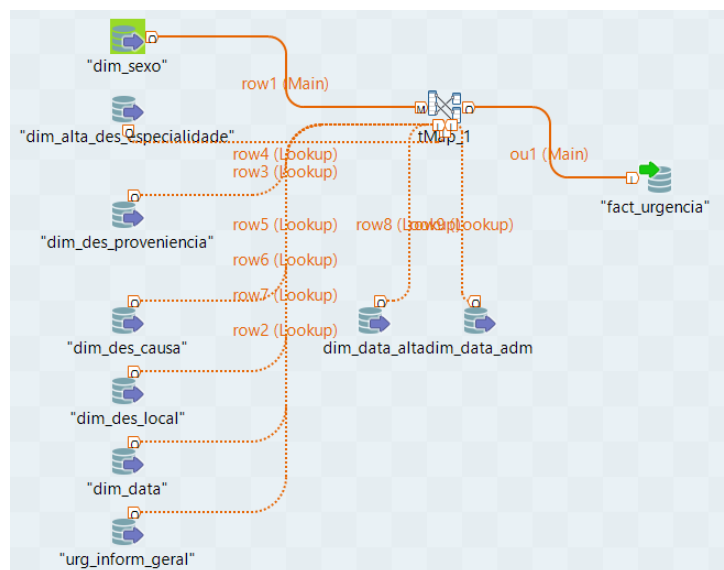


Fig. 3: Povoamento da tabela de factos

## 2.3 Vantagens e Desvantagens

No fundo, o processo desenrola-se de forma similar em ambas as tecnologias, no entanto, achamos muito mais fácil fazê-lo em SQL, pois estamos habituados a escrever comandos nessa linguagem, pelo que conseguimos fazê-lo com bastante destreza, ao passo que trabalhar com a interface gráfica do Talend se revelou mais trabalhoso pois não estávamos familiarizados com a mesma, e é muitas vezes necessário seguir uma série de passos para realizar uma tarefa que é igualmente alcançável com um simples comando SQL.

### 3 *Indicadores Clínicos*

#### 3.1 Indicador Clínico 1

Em primeiro lugar, escolhemos visualizar qual a contagem de urgências por sexo. Este pode ser um óptimo indicador que vai possibilitar à administração do hospital prever e organizar as necessidades logísticas, de modo a disponibilizar um maior numero de camas no internamento e enfermaria feminina, assim como gerir a atribuição de enfermeiros, médicos ou funcionários, de acordo com as necessidades de cada secção. A distribuição dos utentes por género pode também evidenciar factores ambientais ou genéticos que estão a afectar indivíduos de um dado sexo mais frequentemente do que outro.

Contagem de urgências por sexo

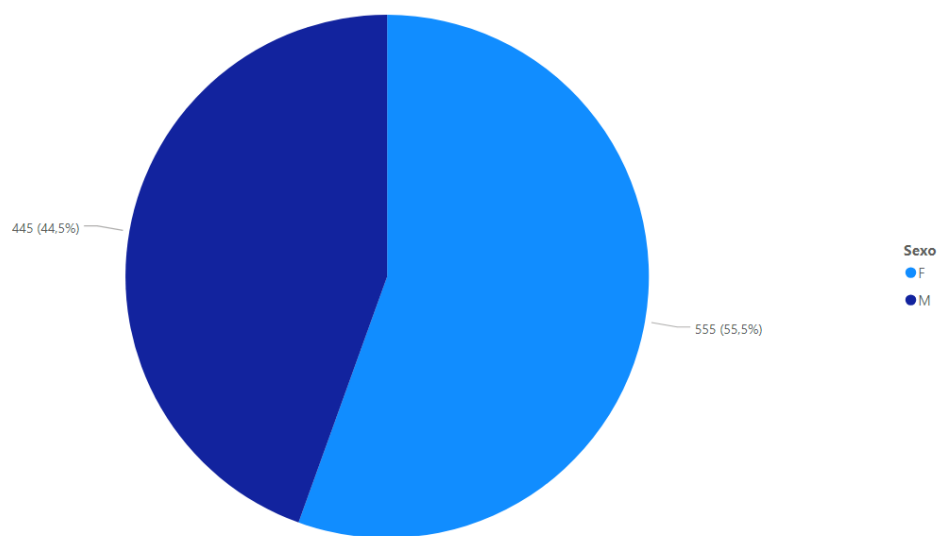


Fig. 4: Contagem de urgências por sexo

### 3.2 Indicador Clínico 2

Sabendo a contagem de urgências por sexo, decidimos agora perceber os motivos que levaram à entrada dos pacientes na urgência do hospital em questão. Se observarmos a informação contida na Fig. 5, podemos constatar que a esmagadora maioria dos casos de urgência, são causados por motivos de doença.

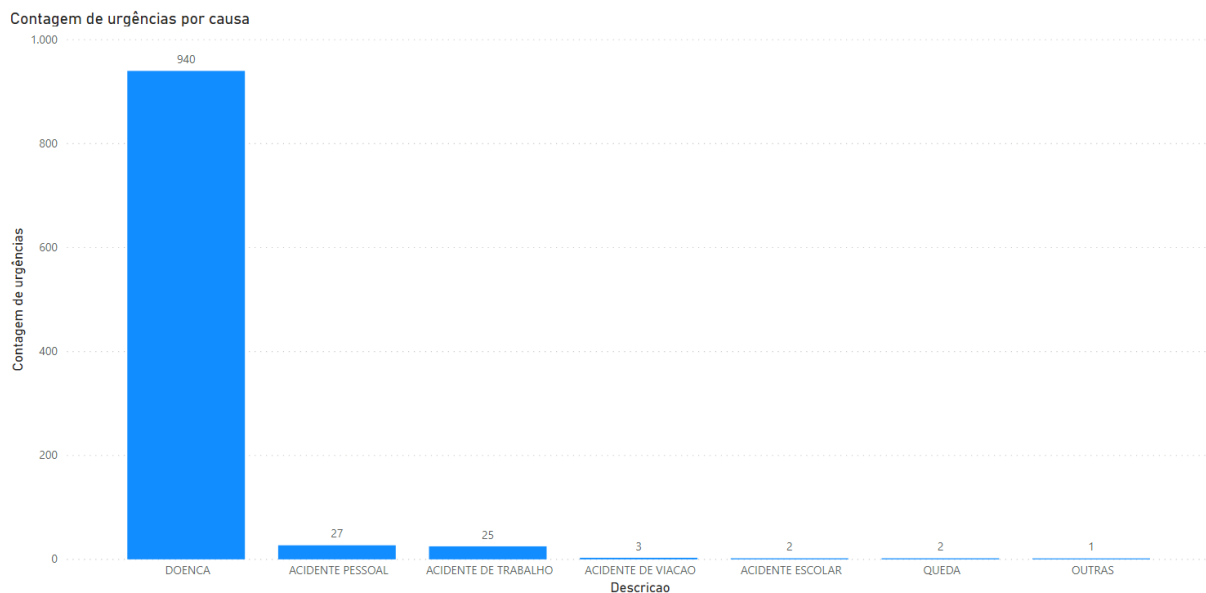


Fig. 5: Número de pacientes por causa

### 3.3 Indicador Clínico 3

Outro indicador importante é a distribuição dos utentes por idade. Este possibilita investigar sobre a possível existência de um ou mais factores que afetem com maior frequência, de forma sistemática, indivíduos de uma determinada faixa etária.

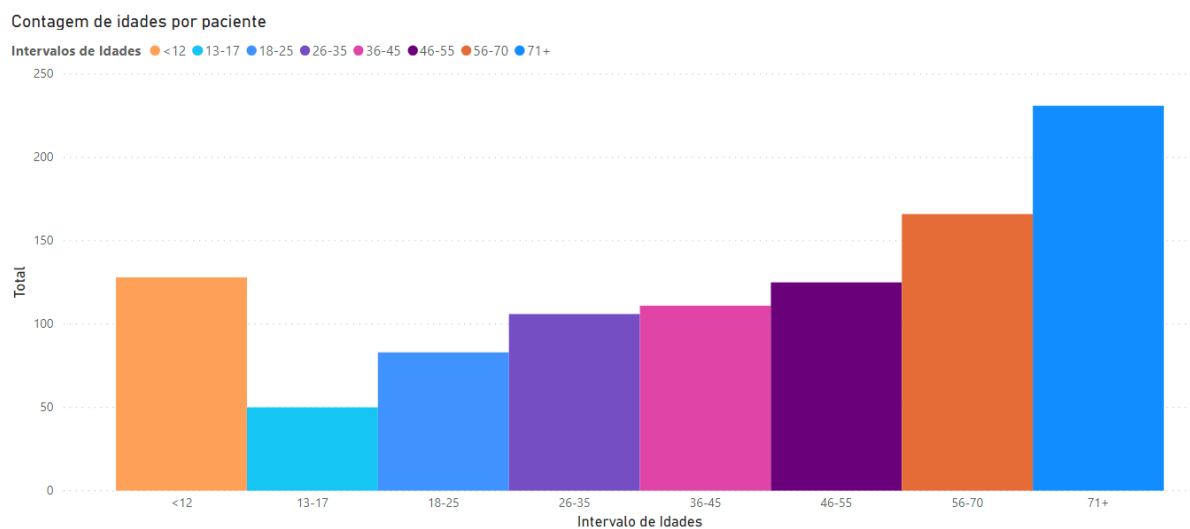


Fig. 6: Contagem de urgências por idade



## 4 Interface com Utilizador

Os indicadores clínicos que foram retirados servem principalmente como estimativas para estudar o estado da população e dessa maneira melhor apropriar os recursos do centro hospitalar. Dessa maneira propomos uma interface que permita aos profissionais aceder às estatísticas de forma simples e directa.

Esta funcionalidade estatística não teria de ser necessariamente a única utilidade da aplicação, pois isto a tornaria um pouco limitada, podendo estar embebida no sistema de gestão das urgências.

A aplicação iria utilizar HL7 como suporte de envio e recepção de mensagens entre dispositivos de forma rápida e segura.

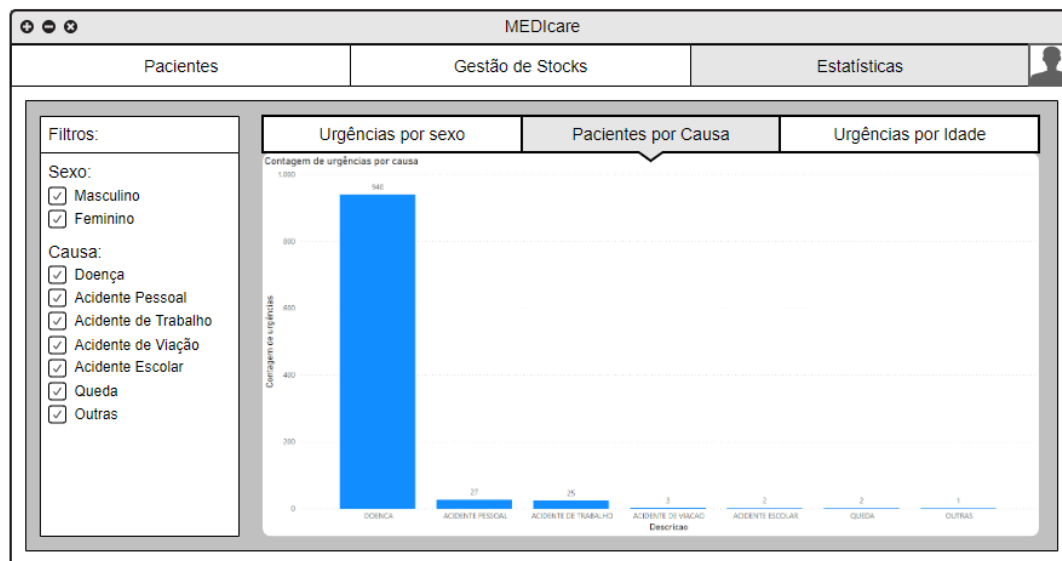


Fig. 7: Interface referente à observação de estatísticas do número de pacientes por causa da urgência

Esta mesma interface também estaria disponível na forma de uma aplicação para tablet e smartphone por exemplo, de forma a garantir o acesso portátil e fácil aos dados quando necessário.

## 5 *Conclusão*

Este projeto ajudou-nos a entender a importância que tem as ferramentas abordadas para a implementação de uma aplicação bem desenvolvida num data warehouse. Damos um especial destaque às ferramentas usadas ,Talend Open Studio e MySQL, em que é possível ver como os programas diferentes extraíram, transformaram e carregaram os dados para o utilizador poder analisar e desenvolver. Damos assim por concluída mais uma etapa da unidade curricular em que foram aplicados todos os conhecimentos adquiridos ao longo do semestre.