

Aplicações Informáticas na Biomedicina

Urgências num Hospital

Trabalho Prático

João Pimentel (a80874)

Pedro Gonçalves (a82313)

Nuno Rei (a81918)

Fábio Senra (a82108)

Dezembro de 2019

Universidade do Minho
Mestrado Integrado em Engenharia Informática

Resumo

O presente trabalho visou o desenvolvimento de um *data warehouse*, no qual se aplicariam métodos de *Business Intelligence* para posteriormente analisar dados. Para tal recorreu-se a um esquema em estrela como base do *data warehouse* e utilizou-se um processo *ETL* para o povoamento do mesmo. Nesta vertente, estabeleceu-se uma comparação de um método "manual", isto é, um *script SQL*, com uma ferramenta *ETL*.

Numa outra fase do projeto, com base num determinado conjunto de dados hospitalares, foram definidos vários indicadores clínicos cujo objetivo consistia na otimização do funcionamento do Hospital em estudo.

Através do desenvolvimento deste trabalho foi possível visualizar o proveito da utilização destas ferramentas de *data warehousing* para o estudo analítico sobre um determinado negócio, de modo a permitir um funcionamento mais fluído.

Conteúdo

1	Introdução	1
2	Análise e Especificação	2
2.1	Descrição do Projeto	2
2.2	Especificação de Requisitos	2
3	Concepção da Resolução	3
3.1	Extração Registos do Ficheiro CSV	3
3.2	Esquema em Estrela	3
3.3	Povoamento das Tabelas de Dimensão	4
3.3.1	Com um <i>Script SQL</i>	4
3.3.2	Com <i>Jobs</i> no <i>Talend</i>	4
3.4	Povoamento Tabela de Factos	5
3.4.1	Com um <i>Script SQL</i>	5
3.4.2	Com <i>Jobs</i> no <i>Talend</i>	5
3.5	Comparação dos Métodos	6
3.6	Indicadores Clínicos	6
3.6.1	Relatórios	7
3.6.1.1	Causas Mais Frequentes	7
3.6.1.2	Urgências Consoante o Género dos Pacientes	7
3.6.1.3	Urgências Tendo em Conta a Data de Nascimento dos Pacientes	8
3.6.1.4	Urgências por Hora do Dia	8
3.6.1.5	Urgências por Dia da Semana	9
3.6.1.6	Número de Urgências Consoante Horas de Atendimento	10
3.6.1.7	Tempo de Internamento dos Pacientes Tendo em Conta a Sua Idade	10
3.6.1.8	Tempo de Internamento dos Pacientes Tendo em Conta a Sua Idade, Consoante o Género	11
3.6.1.9	Tempo de Internamento dos Pacientes Tendo em Conta a sua Idade, Consoante a Causa	12
3.6.1.10	Tempo de Internamento dos Pacientes Tendo em Conta a sua Idade, Consoante a Causa e Género	12
3.6.1.11	Fluxo de Número de Registos por Data de Nascimento, com Doença como Causa	13
3.6.1.12	Urgências por Especialidade	14
3.6.1.13	Urgências por Proveniência	14
3.6.1.14	Médias de Idades por Causa	15
3.6.1.15	Altas por Hora	15
3.6.1.16	Acidentes de Trabalho por Género	16
3.6.2	<i>Dashboard's</i>	16
3.6.2.1	<i>Dashboard</i> Temporal	16
4	Aplicação Informática para Monitorização das Urgências	18
4.1	Página Inicial	18
4.2	Número de Urgências nos Últimos Dias	19
4.3	Causas Mais Comuns	19
4.4	Especialidades com Mais Urgências	20
4.5	Número de Urgências por Horas	20
5	Conclusões e Trabalho Futuro	21

Lista de Tabelas

1	Tabela 1 - Alguns Registos do <i>Dataset</i> .	2
---	--	---

Lista de Figuras

1	Tabela <i>urg_inform_geral</i>	3
2	Esquema em Estrela	3
3	<i>Job Talend</i> para Povoamento de Tabela de <i>Causa</i>	4
4	<i>Job Talend</i> para Povoamento de Tabela de <i>Data_Hora</i>	5
5	<i>Query</i> para Povoamento de Tabela de Factos	5
6	<i>Job Talend</i> para Povoamento de Tabela de Factos	6
7	Causas Mais Frequentes	7
8	Urgências Consoante o Género dos Pacientes	7
9	Urgências Tendo em Conta a Data de Nascimento dos Pacientes	8
10	Urgências por hora do dia	9
11	Urgências por dia da semana	9
12	Número de urgências consoante o número de horas de atendimento.	10
13	Tempo de internamento dos pacientes tendo em conta a sua idade.	11
14	Tempo de internamento dos pacientes tendo em conta a sua idade, consoante o género.	11
15	Tempo de internamento dos pacientes tendo em conta a sua idade, consoante a causa.	12
16	Tempo de internamento dos pacientes tendo em conta a sua idade, consoante a causa e género.	13
17	Fluxo de número de regtos por data de nascimento, com doença como causa.	14
18	Urgências por especialidade.	14
19	Urgências por proveniência.	15
20	Médias de idades por causa.	15
21	Altas por hora.	16
22	Acidentes de trabalho por género.	16
23	Análise temporal de diversos fatores.	17
24	Página inicial	18
25	Número de urgências nos últimos dias	19
26	Causas mais comuns	19
27	Especialidades com mais urgências	20
28	Número de urgências por horas	20

Lista de Extratos

1	Extrato 1 - Povoamento da Tabela de Dimensão <i>Causa</i>	4
2	Extrato 2 - Povoamento da Tabela de Dimensão <i>Data_Hora</i>	4
3	Extrato 3 - Povoamento da Tabela de Factos.	5

1 Introdução

Diariamente todos os hospitais têm de lidar com milhares de novos episódios de urgências, milhares de novos registo a serem inseridos na Base de Dados. Estes novos dados gerados devem ser tratados e analisados posteriormente para que se possa obter informação relevante para fins tal como, por exemplo, suporte a decisão. Perante as necessidades enunciadas, a implementação de um *DataWarehouse* torna-se uma solução bastante apelativa. Estes modelos estão desenhados para favorecer relatórios, análise de grandes volumes de dados e obter informações estratégicas para ajudar na tomada de decisões.

O trabalho prático proposto consistiu na criação de um esquema em formato estrela que permita albergar os dados de um *dataset* com registo de episódios de urgências de um hospital. Posteriormente, foi proposto que se realizasse uma análise dos dados de forma a obter indicadores e criar *dashboards* com informação relevante na área da biomedicina.

Para o processo da análise dos dados foi utilizado *Business Intelligence* que consiste em tecnologias as quais suportam uma grande quantidade de dados e ajudam a identificar e criar novas estratégias para o negócio.

2 Análise e Especificação

2.1 Descrição do Projeto

O *dataset* disponibilizado contém dados reais de uma lista de realização de urgências gerais num determinado hospital nacional. A informação neste contida é guardada numa base de dados *ORACLE* nos sistemas de informação hospitalares da instituição de saúde. Para o problema em questão foram extraídos 1000 registos de uma tabela da base de dados para um ficheiro no formato *CSV*. A informação contida neste ficheiro inclui o número de episódios da urgência (identificador único), a data e hora de admissão do paciente nas urgências, bem como a data e hora da sua alta. Além disso, cada registo possui a descrição da especialidade da alta, do local, da proveniência e da causa de entrada no hospital. Por fim, é representado o género do paciente, bem como a sua data de nascimento.

Note-se que, por razões de confidencialidade, informações que possam identificar os pacientes foram retiradas do *dataset*.

Sendo assim, tenha-se a Tabela 1, onde se podem ver alguns registos contidos no *dataset*, bem como os campos que os caracterizam.

Tabela 1 - Alguns Registos do *Dataset*.

URG_EPISODIO	15051576	15051577	15051578	15051579
DATAHORA_ADM	2017-01-09 13:24:06	2017-01-09 13:25:44	2017-01-09 13:33:30	2017-01-09 13:35:43
DATAHORA_ALTA	2017-01-09 14:24:29	2017-01-09 15:46:43	2017-01-09 17:13:27	2017-01-09 14:39:43
ALTA_DES_ESPECIALIDADE	CLINICA GERAL	CLINICA GERAL	CLINICA GERAL	CLINICA GERAL
DES_LOCAL	URGENCIA GERAL	URGENCIA GERAL	URGENCIA GERAL	URGENCIA GERAL
DES_PROVENIENCIA	EXTERIOR	EXTERIOR	EXTERIOR	EXTERIOR
SEXO	F	F	F	F
DTA_NASCIMENTO	1992-12-26	1987-04-02	1935-05-26	1958-11-08
DES_CAUSA	DOENCA	DOENCA	DOENCA	DOENCA

2.2 Especificação de Requisitos

Deste modo, tendo como base o *dataset* apresentado, pretende-se realizar a execução das tarefas seguintes:

1. Criar um *schema* para o problema, fazendo a importação dos dados no *dataset* para uma tabela.
2. Modelar um modelo dimensional no formato de esquema em estrela (*EER Diagram*) – uma tabela de factos ligada às tabelas de dimensão definidas.
3. Gerar o modelo físico e povoar o *data warehouse* a partir da tabela gerada com o *dataset*:
 - Com um *script SQL*.
 - Com *jobs* no *Talend*.
4. Comparar os processos de povoamento utilizados.
5. Definir e criar indicadores clínicos e de desempenho com o *Microsoft Power BI*, justificando e descrevendo a relevância e a utilidade de cada um dos indicadores.
6. Sugerir e descrever as diversas interfaces e funcionalidades de uma aplicação informática que poderia incluir *dashboards* com os indicadores clínicos e de desempenho definidos, bem como as tecnologias necessárias à sua implementação.

3 Concepção da Resolução

Neste Capítulo será abordado todo o processo de resolução de cada uma das tarefas enunciadas previamente em 2.2.

3.1 Extração Registos do Ficheiro CSV

De modo a extrair os registos contidos no *dataset*, povoando uma tabela com os mesmos, foi utilizada a funcionalidade *Table Data Import Wizard* do MySQL. Assim, o resultado foi uma tabela (*urg_inform_geral*), como se pode ver na Figura 1.

URG_EPISODIO	DATAHORA_ADM	DATAHORA_ALTA	ALTA_DES_ESPECIALIDADE	DES_LOCAL	DES_PROVENIENCIA	SEXO	DTA_NASCIMENTO	DES_CAUSA
15051576	2017-01-09 13:24:06	2017-01-09 14:24:29	CLINICA GERAL	URGENCIA GERAL	EXTERIOR	F	1992-12-26	DOENCA
15051577	2017-01-09 13:25:44	2017-01-09 15:46:43	CLINICA GERAL	URGENCIA GERAL	EXTERIOR	F	1987-04-02	DOENCA
15051578	2017-01-09 13:33:30	2017-01-09 17:13:27	CLINICA GERAL	URGENCIA GERAL	EXTERIOR	F	1935-05-26	DOENCA
15051579	2017-01-09 13:35:43	2017-01-09 14:39:43	CLINICA GERAL	URGENCIA GERAL	EXTERIOR	F	1958-11-08	DOENCA
15051580	2017-01-09 13:50:26	2017-01-09 14:50:18	CLINICA GERAL	URGENCIA GERAL	EXTERIOR	M	1956-09-08	DOENCA
15051581	2017-01-09 13:56:01	2017-01-09 15:14:11	CLINICA GERAL	URGENCIA GERAL	EXTERIOR	F	1989-03-05	DOENCA
15051582	2017-01-09 14:00:49	2017-01-09 15:57:32	CLINICA GERAL	URGENCIA GERAL	EXTERIOR	M	1931-07-28	DOENCA
15051583	2017-01-09 14:03:52	2017-01-09 15:00:02	CLINICA GERAL	URGENCIA GERAL	EXTERIOR	F	1979-01-07	DOENCA

Figura 1 - Tabela *urg_inform_geral*.

3.2 Esquema em Estrela

A principal característica de um esquema em estrela é a presença de dados altamente redundantes, permitindo um melhor desempenho do *Data Warehouse*.

Os dados são modelados em tabelas de dimensão, as quais possuem características de eventos, ligadas a uma tabela de factos. Esta última armazena os factos ocorridos, bem como as chaves para os eventos correspondentes, presentes nas várias tabelas de dimensão.

Como se pode ver na Figura 2, a tabela de factos é a tabela *factos_urgencia*, sendo que todas as restantes representam tabelas de dimensão.

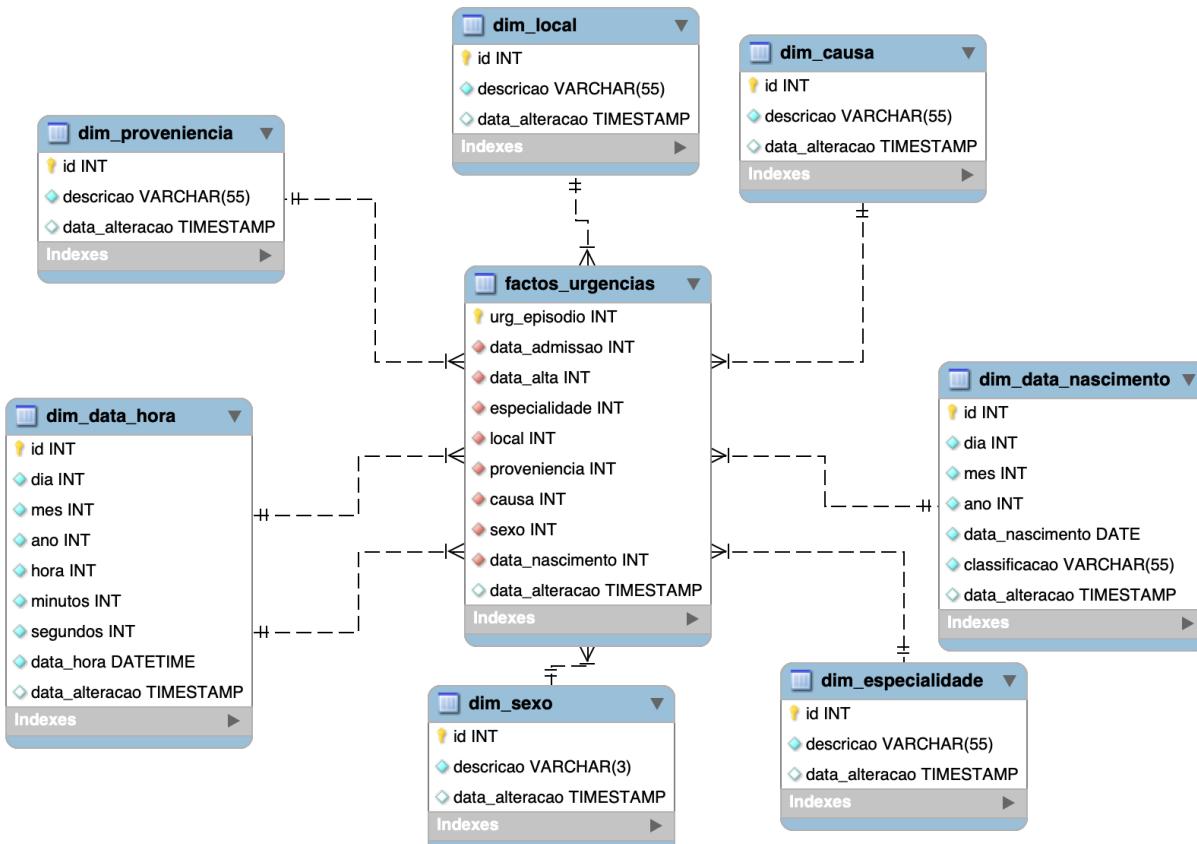


Figura 2 - Esquema em Estrela.

3.3 Povoamento das Tabelas de Dimensão

3.3.1 Com um *Script SQL*

O povoamento das tabelas de dimensão segue uma lógica simples – encontrar todos os valores distintos de um campo útil para o *Data Warehouse* e fazer a sua inserção na respetiva tabela. No caso do Extrato 1, pode-se ver que, para a tabela referente à *causa*, apenas vai ser necessário inserir o valor da *descrição*. Tal deve-se ao facto de o identificador ser calculado pelo sistema (auto-incremento) e a data de alteração ser gerada, também, pelo sistema, associando o valor exato do momento da inserção. Note-se que os valores selecionados para o povoamento provêm da tabela referente ao *dataset*.

```
-- Povoamento tabela Causa
```

```
insert into dim_causa (descricao)
    select distinct DES_CAUSA from urg_inform_geral;
```

Extrato 1 - Povoamento da Tabela de Dimensão *Causa*.

Tenha-se, agora, o caso do povoamento de uma tabela referente a valores de data e hora. Como se pode ver pelo Extrato 2, é necessária a divisão desta data em vários campos, desde anos até segundos, bem como a adição do seu valor original, em formato *DATETIME*. Os restantes campos, ou seja, o identificador e a data de alteração, seguem o método mencionado anteriormente.

```
-- Povoamento tabela Data_Hora
```

```
-- Datas de Admissão
```

```
insert into dim_data_hora (ano, mes, dia, hora, minutos, segundos, data_hora)
    select distinct SUBSTRING(DATAHORA_ADM,1,4), SUBSTRING(DATAHORA_ADM,6,2),
        SUBSTRING(DATAHORA_ADM,9,2), SUBSTRING(DATAHORA_ADM,12,2), SUBSTRING(
        DATAHORA_ADM,15,2), SUBSTRING(DATAHORA_ADM,18,2), STR_TO_DATE(DATAHORA_ADM, '%Y
        -%m-%d %H:%i:%s') from urg_inform_geral;
```

Extrato 2 - Povoamento da Tabela de Dimensão *Data_Hora*.

3.3.2 Com *Jobs* no *Talend*

Povoar cada uma das tabelas de dimensão, através de *jobs* no *Talend* é, também, um processo simples. Sendo que apenas se torna mais complexo no caso das datas que envolvem horas, devido à existência de duas colunas em *urg_inform_geral* que povoarão uma mesma tabela de dimensão. Assim, observe-se a Figura 3, onde se efetua o povoamento da tabela *causa*. Através da conexão à Base de Dados (*DW_URG_Talend*) é selecionado um *input*, neste caso a tabela povoada a partir do *dataset*, na qual é efetuado um mapeamento. Este mapeamento seleciona apenas a coluna *causa* de cada registo. Em seguida é colocado um filtro para ficar com os valores distintos. Por fim, o segundo mapeamento utiliza cada descrição de causa distinta, associa-lhe um identificador e uma data de alteração, tendo em conta as regras do *schema*, guardando todos os registos num determinado *output*.

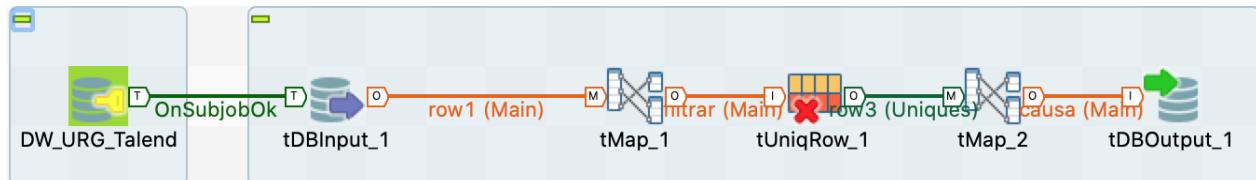


Figura 3 - *Job Talend* para Povoamento de Tabela de *Causa*.

Já no caso da tabela *data_hora*, primeiramente, é necessário efetuar uma união entre as colunas referentes às datas de admissão e de alta. Para isso, são necessários dois *inputs*, aos quais é realizada uma filtragem onde um dos *inputs* apenas fica com a data de admissão e o outro com a data de alta. Tendo ambas as datas, com igual nome (obtidas a partir do mapeamento), é possível efetuar uma união. Note-se que esta operação funciona como um *union all* em termos de *SQL*, sendo necessário retirar valores repetidos. Obtidos os valores distintos, são guardados todos os campos necessários

a partir de cada data, desde o seu ano, até ao segundo. Por fim, os dados são guardados no *Data Warehouse*, como se pode ver na Figura 4.

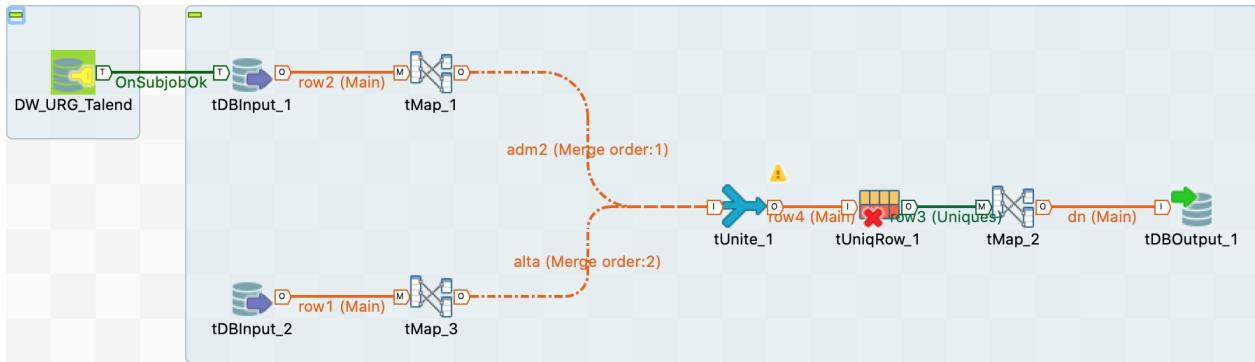


Figura 4 - *Job Talend* para Povoamento de Tabela de *Data_Hora*.

3.4 Povoamento Tabela de Factos

3.4.1 Com um *Script SQL*

O povoamento da tabela de factos é efetuado após todas as tabelas de dimensão estarem completamente preenchidas, já que esta vai utilizar os valores contidos nas últimas.

Assim, a ideia passa por efetuar um conjunto de uniões que permitam relacionar a cada linha da tabela do *dataset*, uma entrada na tabela de factos, associando esta às chaves que representam os valores das tabelas de dimensão.

Note-se que, como o campo *URG_EPISODIO* funciona como um identificador único, pode ser utilizado como identificador de cada entrada na tabela. Além disso, é de destacar que são feitas duas ligações à tabela de dimensão *dim_data_hora*, sendo estas referentes às datas de admissão e alta. É, ainda, merecedor de destaque o facto de não existirem valores nulos no *dataset*, fazendo com que não seja necessária a definição de um valor especial para estes casos.

```
-- Povoamento tabela de Factos
insert into factos_urgencias (urg_episodio , causa , local , proveniencia , especialidade
, sexo , data_nascimento , data_admissao , data_alta)
select L.URG_EPISODIO, C.id , Loc.id , P.id , E.id , S.id , DN.id , DH1.id , DH2.id
from urg_inform_geral as L
left join dim_causa as C on L.DES_CAUSA = C.descricao
left join dim_local as Loc on L.DES_LOCAL = Loc.descricao
left join dim_proveniencia as P on L.DES_PROVENIENCIA = P.descricao
left join dim_especialidade as E on L.ALTA DES_ESPECIALIDADE = E.descricao
left join dim_sexo as S on L.SEXO = S.descricao
left join dim_data_nascimento as DN on STR_TO_DATE(L.DTA_NASCIMENTO, '%Y-%m-%d') = DN.data_nascimento
left join dim_data_hora as DH1 on STR_TO_DATE(L.DATAHORA_ADM, '%Y-%m-%d %H:%i:%s') = DH1.data_hora
left join dim_data_hora as DH2 on STR_TO_DATE(L.DATAHORA_ALTA, '%Y-%m-%d %H:%i:%s') = DH2.data_hora;
```

Extrato 3 - Povoamento da Tabela de Factos.

3.4.2 Com *Jobs* no *Talend*

Tirando proveito das funcionalidades do *Talend*, é possível executar uma *query* sobre a base de dados, obtendo-se uma tabela. Assim, veja-se a Figura 5. A tabela obtida a partir da *query* é posteriormente utilizada como *input* do *job* em questão.

Query	<pre>"select L.URG_EPISODIO, C.id, Loc.id, P.id, E.id, S.id, DN.id, DH1.id, DH2.id from urg_inform_geral as L left join dim_causa as C on L.DES_CAUSA = C.descricao left join dim_local as Loc on L.DES_LOCAL = Loc.descricao"</pre>
-------	--

Figura 5 - *Query* para Povoamento de Tabela de Factos.

Assim, executando a seleção existente dentro do Extrato 3, é possível efetuar um redirecionamento direto do *input* para o *output*, como se vê na Figura 6. Deste modo, este *job* não passa de um mapeamento direto do *input* para a tabela de factos existente no *schema*.

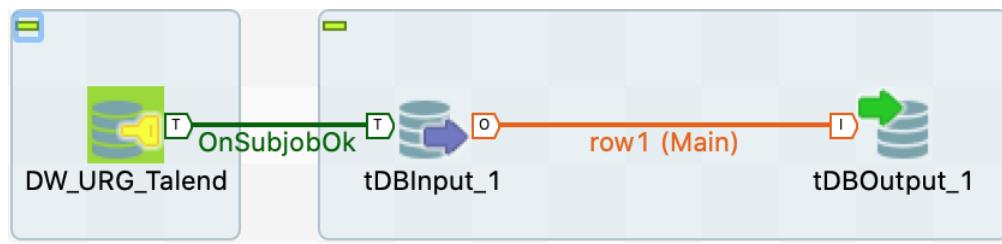


Figura 6 - *Job Talend* para Povoamento de Tabela de Factos.

Note-se que poderia ter sido utilizado um componente *tMap* para relacionar todas as tabelas, não sendo necessária a utilização da *query*. No entanto, esta solução não seria tão legível como a realizada.

3.5 Comparação dos Métodos

Após a implementação de um processo *ETL* com uma ferramenta apropriada e do mesmo processo tirando proveito de um *script SQL*, é possível concluir que as ferramentas apropriadas permitem a conclusão da tarefa, necessitando de pouco conhecimento sobre bases de dados. Quer isto dizer que permitem efetuar ações mais complexas com um número reduzido de parâmetros, visto que a ferramenta utilizada permite encadear vários componentes, apenas sendo necessário um conhecimento prévio das alterações necessárias. Além disso, devido à sua implementação, reduzem o tamanho das transações, permitindo um controlo de erros mais simples.

Apesar disto, um *script SQL* pode tirar partido de contagem de linhas, chaves estrangeiras, únicas, dimensões, entre outros componentes, de modo a otimizar as operações sobre os dados. Note-se, ainda, que, tendo um conhecimento elevado em *SQL*, torna-se relativamente simples a implementação destes processos, pelo que o uso de ferramentas *ETL* acaba por ser posto em segundo plano, apesar dos seus pontos positivos.

3.6 Indicadores Clínicos

Um indicador clínico ideal deve possuir as seguintes características:

- Ser baseado em definições pré-definidas, ser exaustivamente e exclusivamente descritivo;
- Detetar poucos falsos positivos e falsos negativos;
- Ser válido e de confiança;
- Discriminar bem os dados;
- Ser relevante para a prática clínica;
- Permitir comparações úteis;
- Ser baseado em evidências.

Além disso, cada indicador deve ser definido de forma detalhada, com especificações de dados, de modo a ser específico e sensível ao caso em estudo.

Dito isto, para a resolução do projeto proposto, o grupo optou por definir os indicadores clínicos apresentados de seguida. Note-se que a sua justificação e interpretação será efetuada no próximo capítulo.

Indicadores clínicos propostos:

- Causas mais frequentes;
- Género que mais urgências tem;

- Número de urgências, tendo em conta as datas de nascimento dos pacientes;
- Fluxo de urgências por hora;
- Número de urgências por dia da semana;
- Análise do tempo entre admissão e alta;
- Idades dos pacientes associadas ao seu tempo de internamento;
- Idades e géneros dos pacientes associados ao seu tempo de internamento;
- Idades dos pacientes associadas ao seu tempo de internamento e causa;
- Idades dos pacientes associadas ao seu tempo de internamento, tendo em conta a causa e o seu género;
- Urgência por doença e data de nascimento;
- Urgências por especialidade;
- Urgências por proveniência;
- Média de idades por causa;
- Altas por hora;
- Acidentes de trabalho por sexo.

3.6.1 Relatórios

3.6.1.1 Causas Mais Frequentes

De modo a conhecer as causas que mais levam os pacientes às urgências, uma análise sobre essas mesmas causas é crucial. Consoante os resultados, pode permitir ao Hospital informar os seus pacientes sobre medidas preventivas perante certas causas, reduzindo o número de urgências. Assim, como se pode observar pela Figura 7, doença é a causa que leva a grande maioria das pessoas às urgências. Note-se, ainda, que, das restantes causas, acidentes de trabalho e acidentes pessoais distanciam-se, também, das outras. Dito isto, uma maior informação sobre como tratar certas doenças de menor gravidade, pode permitir aos pacientes uma menor necessidade de se deslocarem às urgências do Hospital.

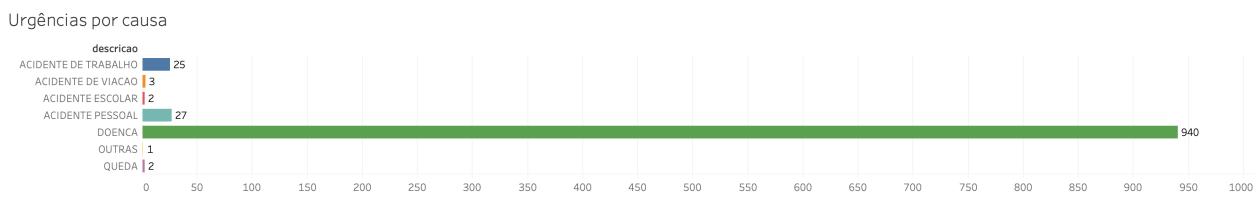


Figura 7 - Causas Mais Frequentes.

3.6.1.2 Urgências Consoante o Género dos Pacientes

Outro indicador clínico interessante passa por conhecer a relação entre o género dos pacientes e as suas idas à urgência. Sendo assim, como se pode ver pela Figura 8, existe um maior número de urgências associadas a pacientes do sexo feminino do que pacientes do sexo masculino.

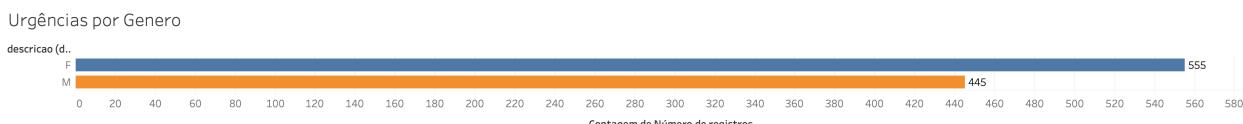


Figura 8 - Urgências Consoante o Género dos Pacientes.

3.6.1.3 Urgências Tendo em Conta a Data de Nascimento dos Pacientes

Para ter conhecimento da faixa etária dos pacientes que mais frequentam as urgências, este indicador torna-se bastante importante. Dito isto, pode permitir ao hospital alocar trabalhadores mais especializados nas urgências, consoante a idade do paciente, ou seja, se as urgências tiverem uma grande afluência de pacientes com menos de 5 anos de idade, o Hospital deve colocar especialistas nesta faixa etária nas urgências.

Ora, como se pode ver pela Figura 9, existe uma grande afluência de pacientes na casa dos 80 e na casa dos 10 anos de idade. Assim, devem ser colocados especialistas em pacientes séniores e em crianças, de modo a permitir um melhor tratamento nas urgências. Note-se que, apesar disto, existe, também, uma grande variedade nas idades dos pacientes.



Figura 9 - Urgências Tendo em Conta a Data de Nascimento dos Pacientes.

3.6.1.4 Urgências por Hora do Dia

De modo a ter sempre especialistas prontos a tratar de uma urgência, o Hospital tem que se prever para as horas em que há um maior congestionamento. Assim, o estudo do número de urgências por hora do dia permite destacar mais ou menos trabalhadores para o serviço de urgência, consoante o necessário.

Dito isto, como se pode ver pela Figura 10, existe um aumento no número de urgências desde madrugada até meio da manhã, sendo que este número diminui posteriormente até à hora de almoço, voltando a subir de forma exponencial logo de seguida. Note-se, ainda, que este valor volta a reduzir, gradualmente, até ao final do dia. Estas variações podem ser justificadas pela disponibilidade das pessoas em se deslocar até ao Hospital nas horas onde os picos são atingidos, ou seja, 10h e 14h. Consequentemente, devem ser destacados mais especialistas nestas horas, de modo a atender o maior número possível de pacientes, não pondo em risco a saúde dos mesmos.

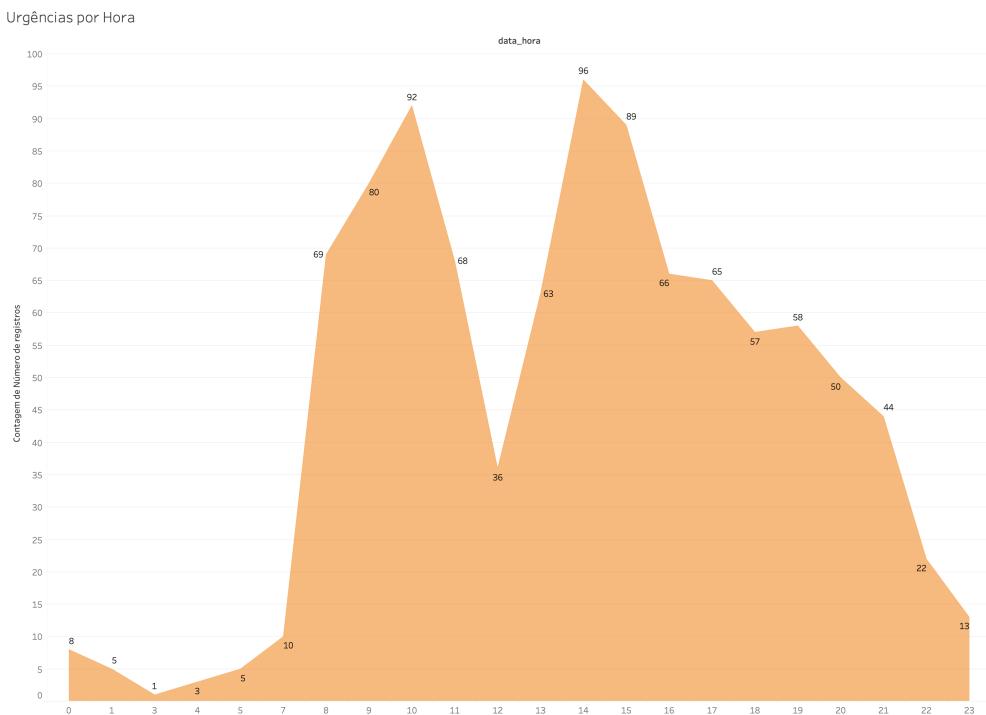


Figura 10 - Urgências por hora do dia.

3.6.1.5 Urgências por Dia da Semana

De forma a alocar o número de especialistas para cada dia da semana, de forma eficiente, é necessário perceber quais os dias onde existe mais movimento nas urgências. Assim sendo, um estudo do número de urgências nos diferentes dias da semana, permite destacar os funcionários necessários para garantir o normal funcionamento do Hospital a cada dia.

Posto isto, ao analisar a Figura 11, é possível perceber que o número de urgências aumenta desde segunda-feira atingindo o pico de afluência à quarta-feira. A partir desse dia, a afluência vai decrescendo nos restantes dias, atingindo o mínimo no sábado. Assim sendo, é necessário destacar mais especialistas para os dias com maior afluência, de modo a garantir que o Hospital continue a funcionar de forma expectável.

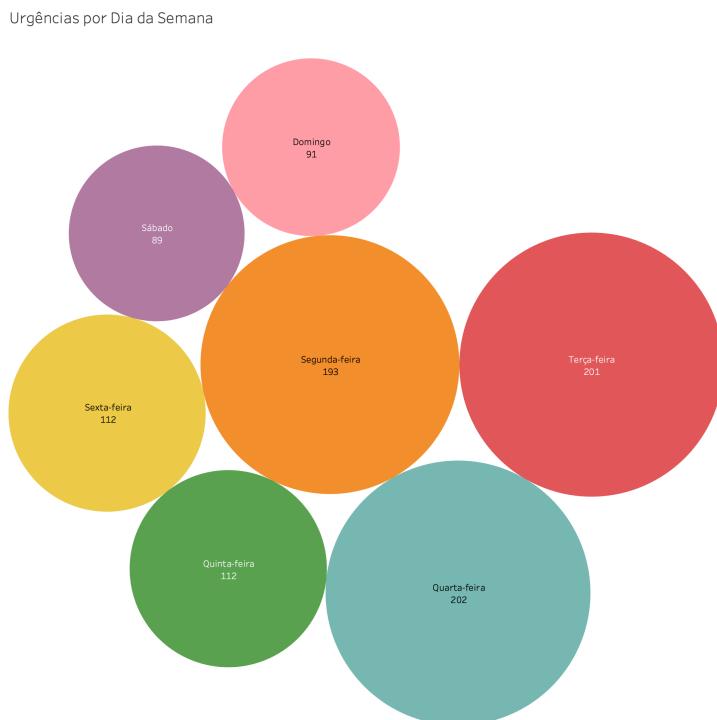


Figura 11 - Urgências por dia da semana.

3.6.1.6 Número de Urgências Consoante Horas de Atendimento

Com o intuito de perceber qual a eficiência do Hospital no atendimento aos pacientes, é útil estudar o número de horas de atendimento disponibilizado para cada indivíduo. Este estudo permite perceber se é necessário reforçar o número de funcionários nas urgências consoante os resultados obtidos, bem como permitir melhorar as condições dos pacientes durante o atendimento.

Analizando a Figura 12, é possível perceber que a grande maioria dos pacientes fica, no máximo, duas horas a ser atendidos. Além disso, é de notar que são poucos os casos em que um paciente se encontre mais do que 5 horas em atendimento. Assim, é necessário manter os locais onde os pacientes são atendidos, regularmente, devido ao reduzido tempo de atendimento, bem como ter profissionais prontos a ajudá-los, caso necessário.

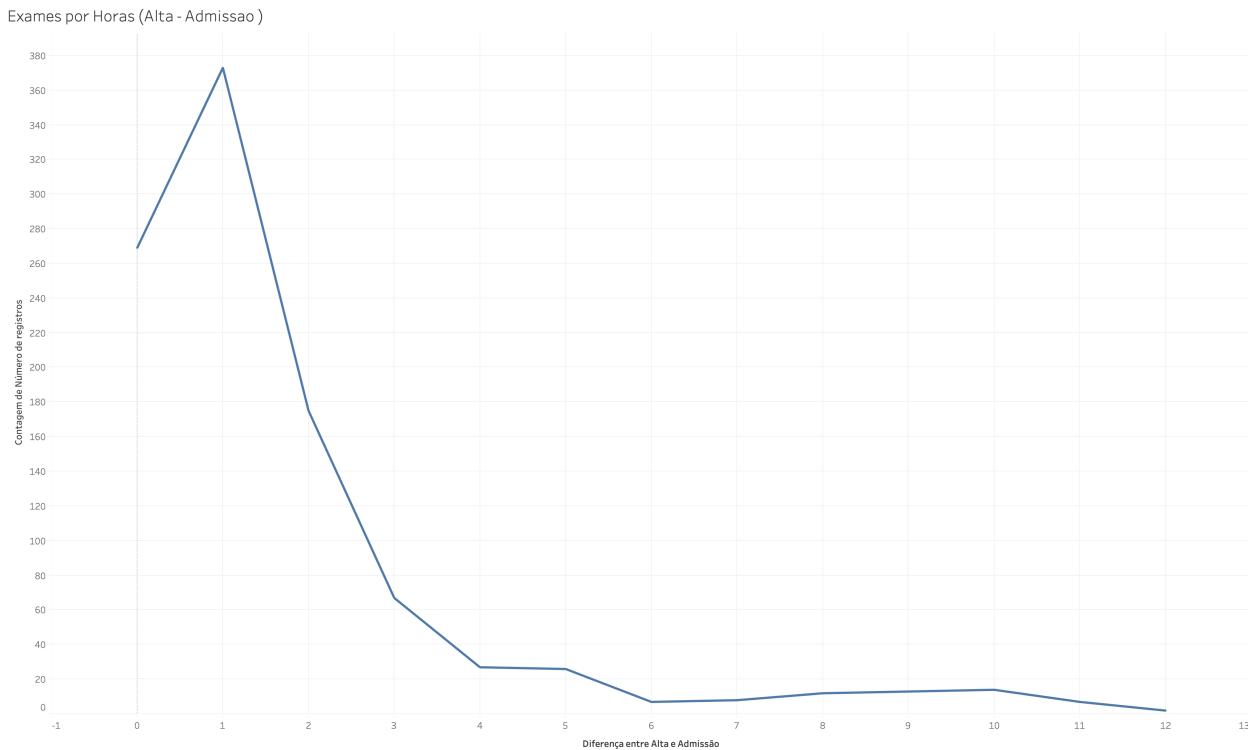


Figura 12 - Número de urgências consoante o número de horas de atendimento.

3.6.1.7 Tempo de Internamento dos Pacientes Tendo em Conta a Sua Idade

De modo a saber quais as idades dos pacientes que mais, ou menos, tempo estão no serviço de Urgência, em atendimento, este indicador torna-se crucial. Consoante os resultados, pode permitir que, com um conhecimento prévio da idade do paciente, os profissionais se preparem para tratar do paciente durante o determinado período de tempo, alocando os recursos necessários.

Dito isto, como se pode ver pela Figura 13, e como já confirmado anteriormente, a grande maioria dos pacientes não fica em análise por mais de 2 horas, sendo que até 3 horas ainda existe um número elevado de registo. Além disso, pode-se ver que não existe uma grande variação nas idades dos pacientes por cada hora. Quer isto dizer que, de momento, a idade dos pacientes não é um fator que influencie em grande medida o seu tempo de internamento.

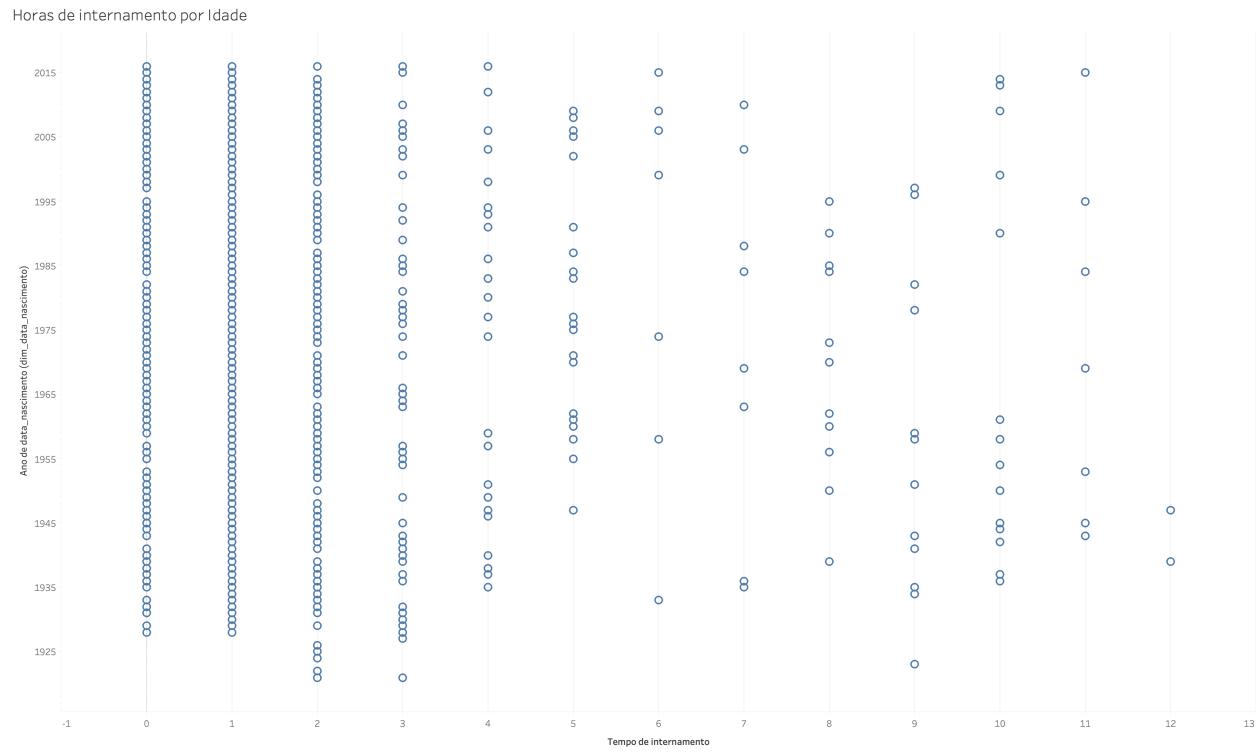


Figura 13 - Tempo de internamento dos pacientes tendo em conta a sua idade.

3.6.1.8 Tempo de Internamento dos Pacientes Tendo em Conta a Sua Idade, Consoante o Género

No seguimento do indicador anterior, a ideia deste é semelhante. No entanto, adicionou-se o género como fator de análise. Dito isto, pela visualização da Figura 14, pode-se ver que, no caso do sexo feminino, existe uma tendência em que com o aumento do tempo de internamento, também aumenta a idade. Note-se que isto não acontece no caso dos pacientes do sexo masculino, não sendo visível qualquer padrão, algo que poderia mudar com o aumento do número de registo

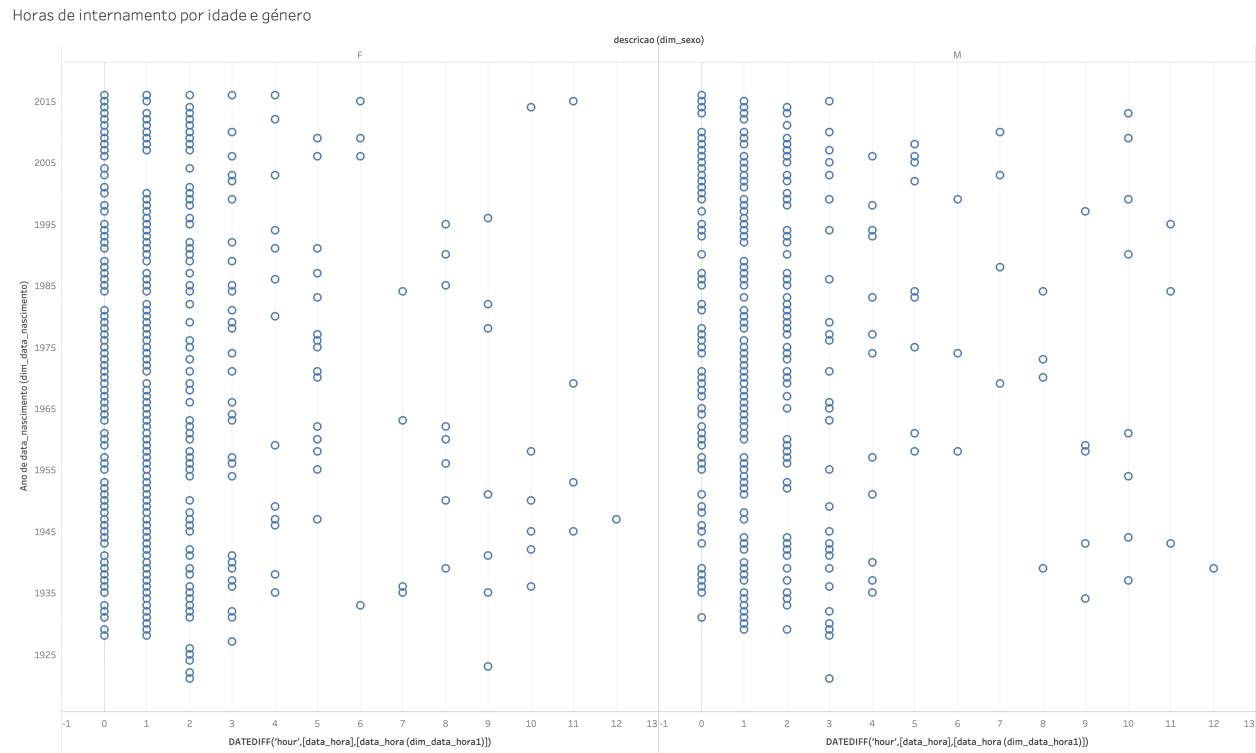


Figura 14 - Tempo de internamento dos pacientes tendo em conta a sua idade, consoante o género.

3.6.1.9 Tempo de Internamento dos Pacientes Tendo em Conta a sua Idade, Consoante a Causa

Na tentativa de associar a cada causa uma estimativa de tempo de internamento tendo, ainda, em conta a idade dos pacientes, a ideia passou por juntar estes três parâmetros num indicador. Esta análise pode permitir, tal como os indicadores anteriores, alocar o número de especialistas necessários em cada causa por um determinado número de horas, mantendo o Hospital em funcionamento constante e sem atrasos.

Desta forma, como se pode ver pela Figura 15, os acidentes de trabalho e de viação ocorrem a pacientes numa determinada faixa de idades, excluindo crianças e idosos, como seria de esperar. Os acidentes pessoais ocorrem em qualquer faixa etária e tem uma duração de tratamento reduzida. Por fim, a maioria dos registos deve-se a doença, onde se pode ver que, independentemente da idade, o número de registos é elevado. Note-se que o tempo de atendimento varia bastante neste caso, sendo que a maior parte dos pacientes tem alta em menos de 5 horas.

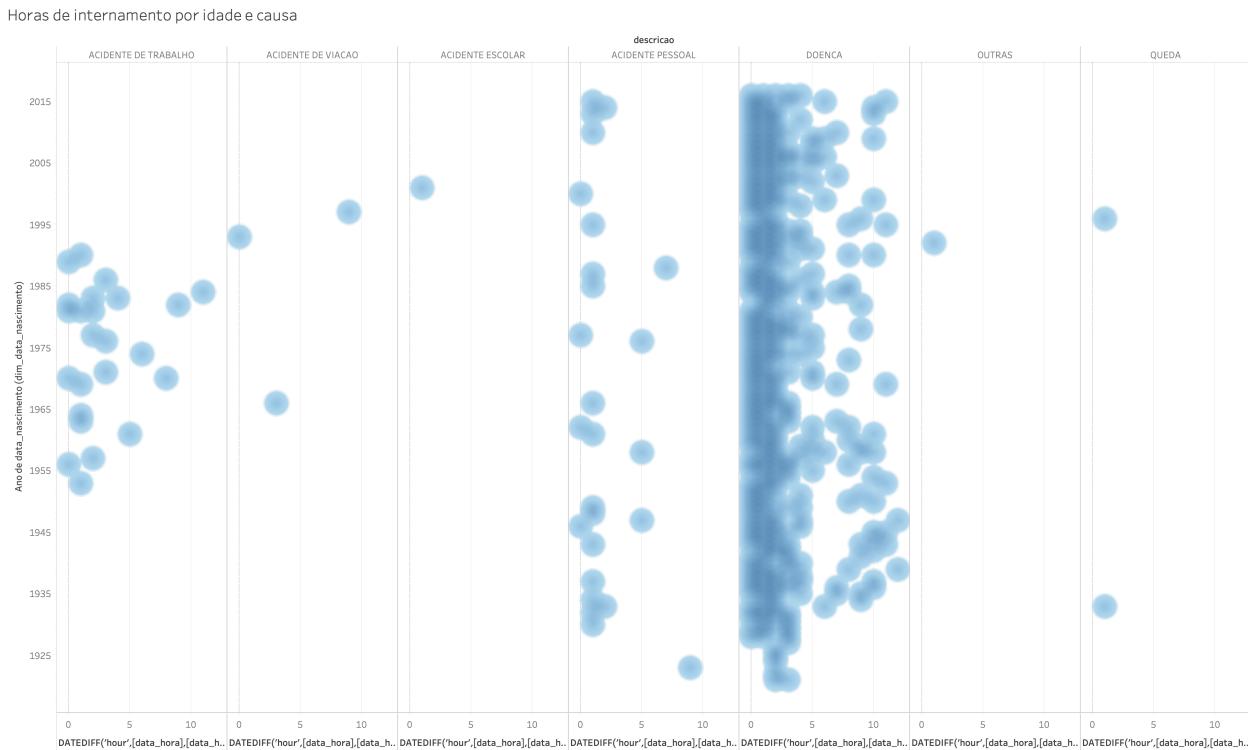


Figura 15 - Tempo de internamento dos pacientes tendo em conta a sua idade, consoante a causa.

3.6.1.10 Tempo de Internamento dos Pacientes Tendo em Conta a sua Idade, Consoante a Causa e Género

Mais uma vez, tendo em conta as ideias base dos indicadores anteriores, a ideia passou por ver de que forma o tempo de internamento variava consoante a causa e, agora, com o género dos pacientes.

Assim, pela análise da Figura 16, pode-se constatar que existem mais acidentes de trabalho e viação associados a pacientes do sexo masculino. No que toca às doenças, tanto pacientes do sexo masculino, como do sexo feminino possuem uma elevada variedade nas idades, sendo os tempos de tratamento relativamente baixos. Note-se, apenas, que o número de urgências associadas a pacientes do sexo masculino é mais baixo e, ainda, que, no que toca aos pacientes do sexo feminino, parece existir uma tendência em que o tempo de internamento aumenta com o aumento da idade.

Horas de Internamento por idade, género e causa

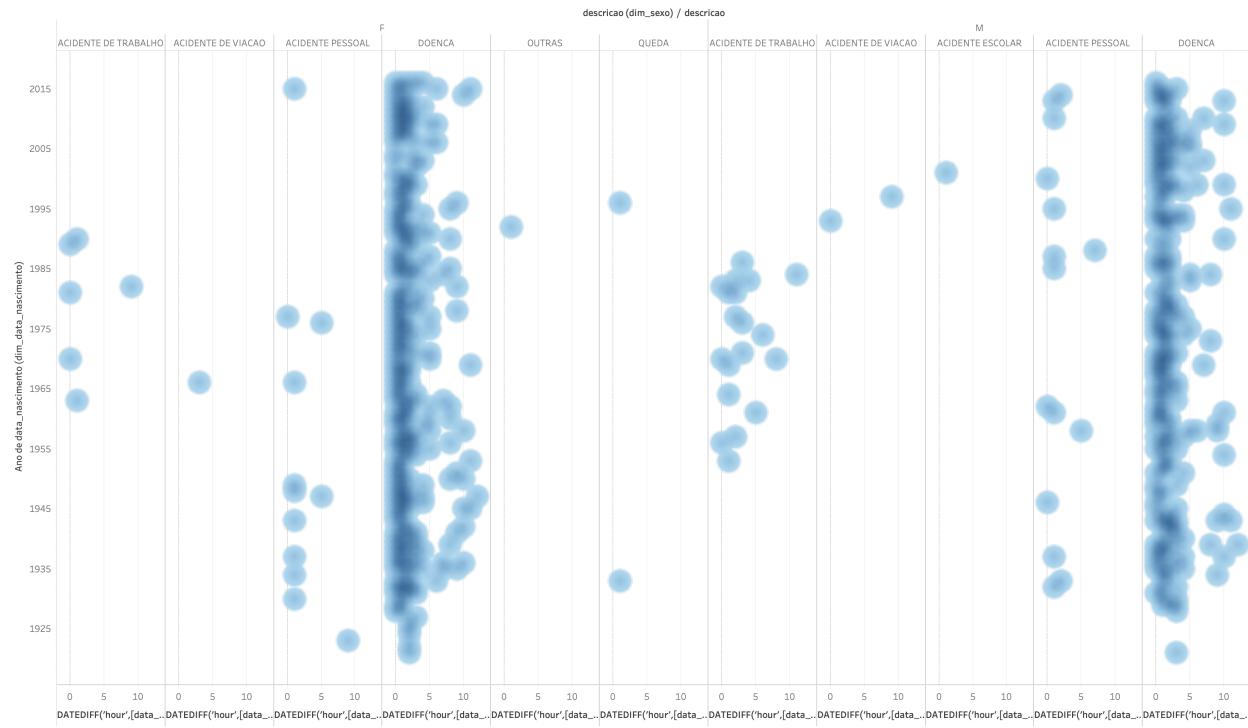


Figura 16 - Tempo de internamento dos pacientes tendo em conta a sua idade, consoante a causa e género.

3.6.1.11 Fluxo de Número de Registos por Data de Nascimento, com Doença como Causa

De modo a compreender o fluxo de registos associados à causa Doença, tendo em conta as idades dos pacientes, definiu-se este indicador. Este poderia permitir uma especialização dos profissionais de saúde em determinadas faixas etárias, melhorando a estadia dos pacientes nas urgências, deixando-os o mais à vontade possível.

Assim, pela visualização da Figura 17, é possível ver que existem dois grandes picos associados a pacientes com idade consideravelmente elevada e a pacientes de idade baixa, sendo que existem picos menores nas restantes idades. Note-se, também, que a partir de 1928, sensivelmente, o número de urgências é quase nulo, o que se pode dever à elevada idade dos pacientes, associada à esperança média de vida dos mesmos. Além disso, é de notar que os profissionais de saúde das urgências devem estar preparados para tratar um paciente de qualquer idade, no que toca à causa de doença, pelo que uma formação extra não se torna crucial.

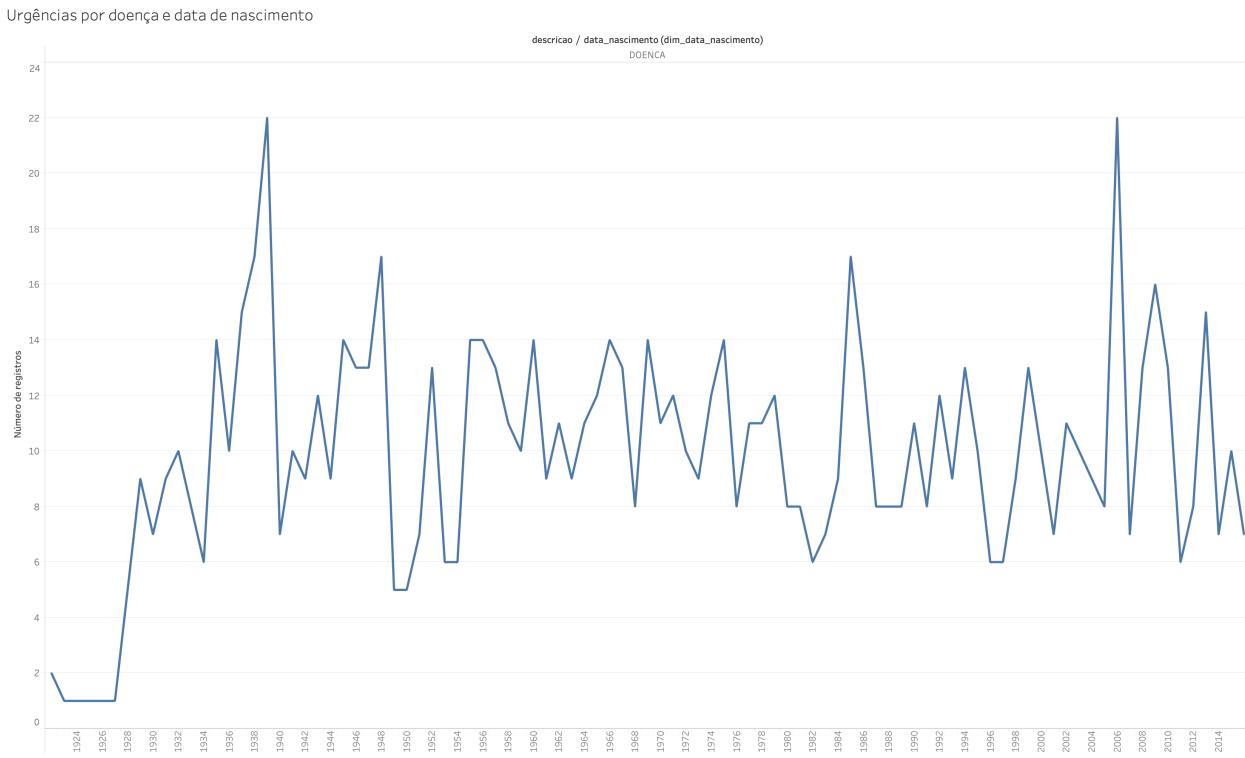


Figura 17 - Fluxo de número de registo por data de nascimento, com doença como causa.

3.6.1.12 Urgências por Especialidade

Este indicador é muito importante para saber quais as especialidades com mais afluência no que toca a urgências. Através da análise da Figura 18 é possível concluir que todas as entradas nas urgências são da especialidade de clínica geral. Neste caso específico apenas existe esta especialidade. Porém, caso existissem vários tipos de entradas seria possível concluir quais as especialidades que teriam maior impacto nas urgências e, deste modo, seria mais fácil saber quais os médicos de cada especialidade que deveriam ser destacados para estarem no serviço de urgências.

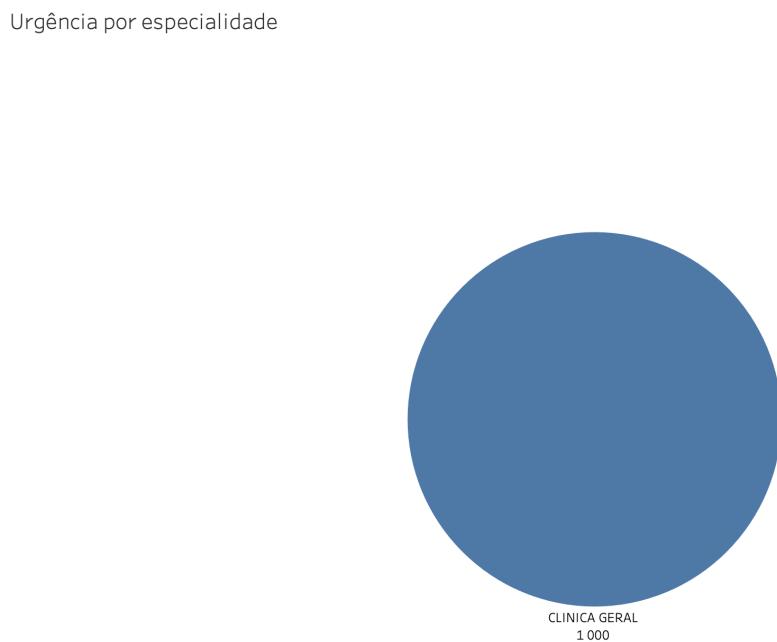


Figura 18 - Urgências por especialidade.

3.6.1.13 Urgências por Proveniência

De modo a saber quais as origens das urgências criou-se o indicador que mostra todas as proveniências das urgências que chegam ao hospital. Neste caso em concreto apenas temos uma proveniência Exterior (Figura 19), mas se houvesse outro tipo de origens seria possível concluir quais as proveniências que resultam em mais urgências.



Figura 19 - Urgências por proveniência.

3.6.1.14 Médias de Idades por Causa

O seguinte indicador é importante para saber quais as idades mais propensas a determinadas causas, ou seja, saber se alguém mais jovem ou mais velho é mais suscetível a um determinado tipo de doenças. Pela análise da Figura 20 é possível retirar que as quedas são a causa de urgência com média de idades superior (54,5 anos). Já a causa que regista menor média de idades são os acidentes escolares (18 anos), visto que são os mais jovens que frequentam os respetivos estabelecimentos.

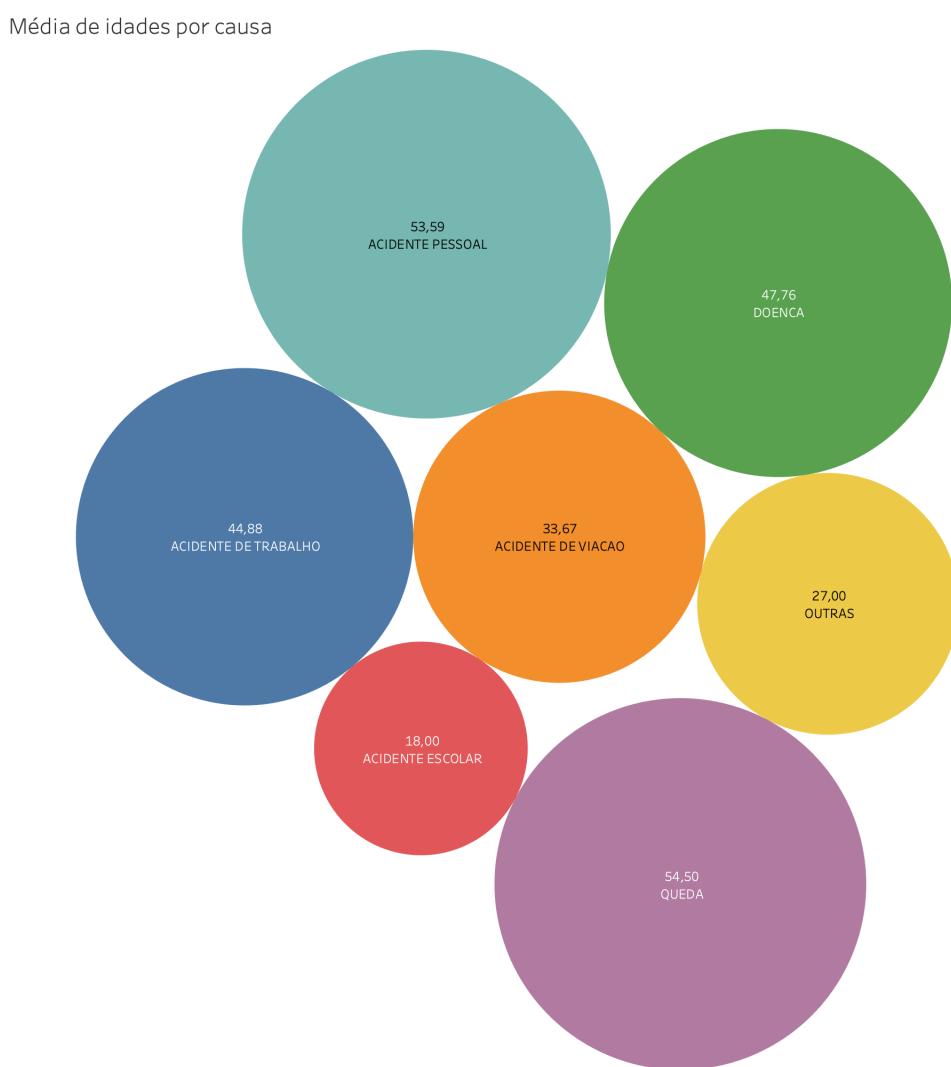


Figura 20 - Médias de idades por causa.

3.6.1.15 Altas por Hora

O indicador de altas por hora permite saber as horas às quais o maior número de pacientes tem alta e de que forma essas horas têm impacto na vida das famílias dos pacientes. Assim, através da Figura 21 verificam-se os seus picos durante a manhã, entre as 10h e as 12h, e durante a noite, volta das 23h. Posteriormente, volta a atingir um novo pico à 1h da manhã.

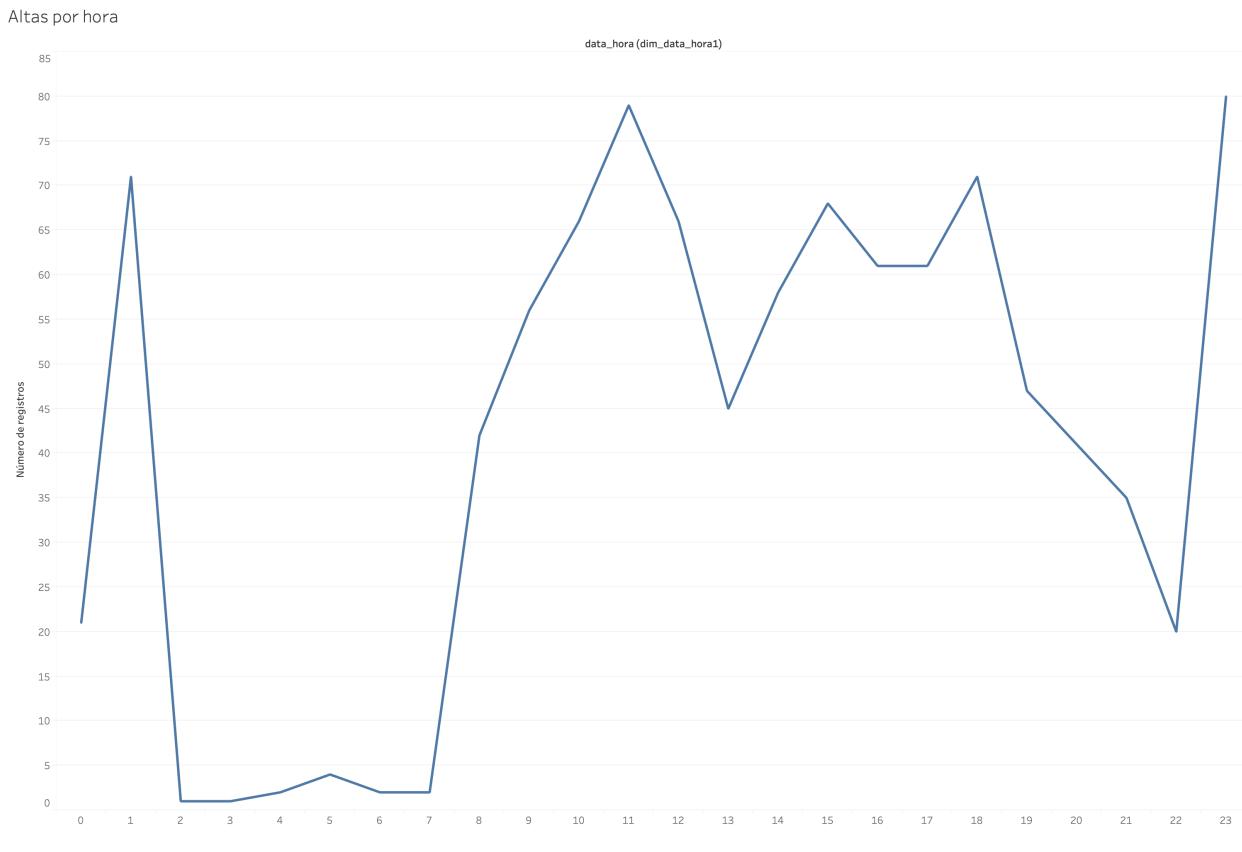


Figura 21 - Altas por hora.

3.6.1.16 Acidentes de Trabalho por Género

O presente indicador permite abordar de forma mais aprofundada a causa de urgência Acidentes de trabalho, possibilitando concluir qual dos géneros tem trabalhos de maior risco. No caso em estudo, como se pode ver na figura 22, o género com mais acidentes de trabalho é o masculino, pelo que se pode alegar que este possui trabalhos de maior risco.

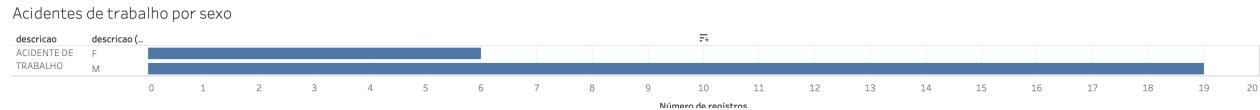


Figura 22 - Acidentes de trabalho por género.

3.6.2 Dashboard's

Nesta secção será apresentado um *dashboard* desenvolvido, onde são associados vários indicadores, de modo a permitir uma análise mais concisa dos dados.

Note-se que poderiam ser desenvolvidos mais exemplos de *dashboards*, no entanto, devido ao elevado número de indicadores, o grupo optou por apresentar apenas um *dashboard*, permitindo observar de que forma é possível agrupar indicadores.

3.6.2.1 Dashboard Temporal

Tendo em conta as análises efetuadas em cada indicador, esta junção permite visualizar os picos de trabalho nas urgências, quer em termos de admissões, quer em termos de altas. Assim, como se pode ver pela Figura 23, os máximos de admissões ocorrem próximos das 11 e das 14 horas, contrariamente às altas que variam um pouco mais ao longo do dia, tendo os seus picos às 1, 11 e 23 horas.

Além disso, é notório que a diferença entre a admissão e a alta tende a ser baixa, sendo que a maioria dos casos recebe alta ao fim de sensivelmente uma hora. Note-se, ainda, que ao longo da semana, segunda, terça e quarta são os dias com mais consultas a si associadas, sendo que estas reduzem consideravelmente nos restantes dias.

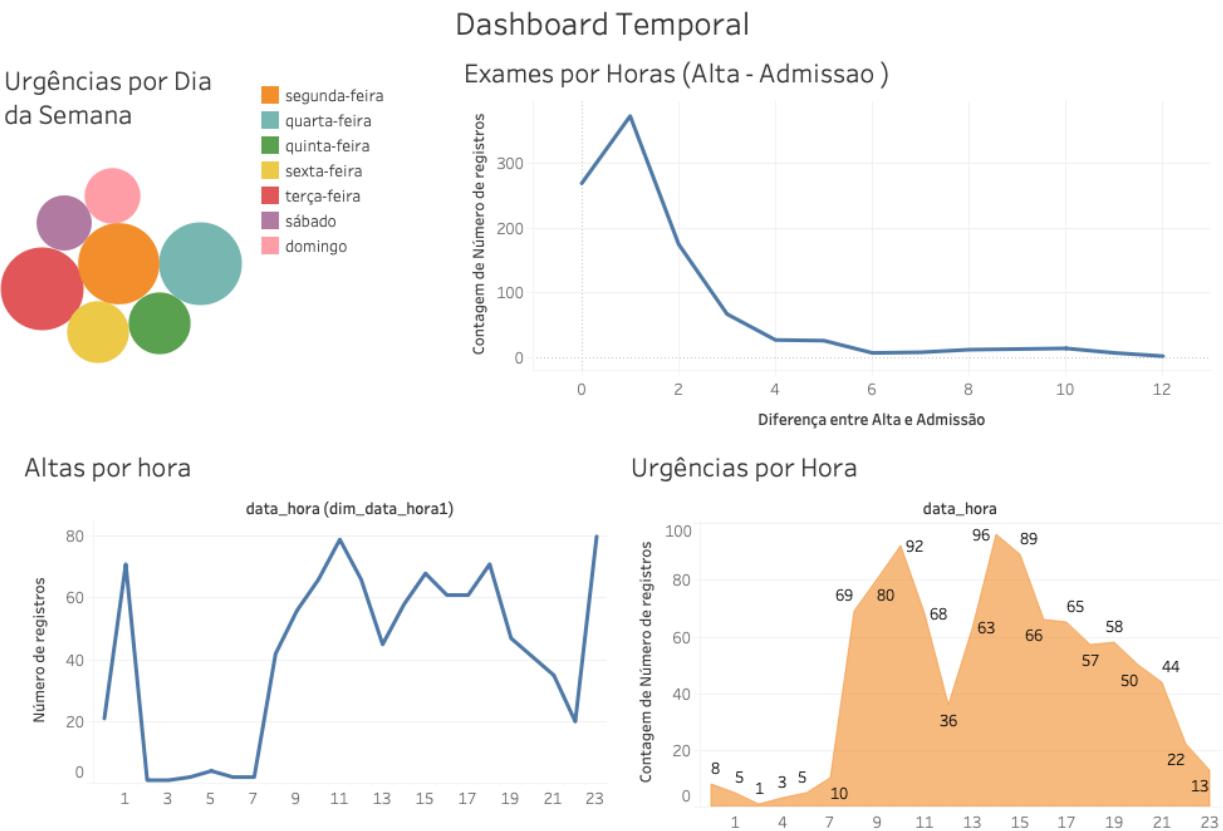


Figura 23 - Análise temporal de diversos fatores.

4 Aplicação Informática para Monitorização das Urgências

A aplicação informática a desenvolver para monitorizar urgências tinha como principal objetivo ser possível verificar o estado das urgências e a forma como estas entradas nas urgências se foram desenvolvendo ao longo do tempo. Esta permitirá facilitar o trabalho de quem gera o serviço de urgências tendo como funcionalidades:

- Estado atual das urgências com *dashboard*;
- Ver quais as urgências pela sua causa;
- Ver especialidades com mais urgências;
- Ver tempos de internamento resultantes das urgências;
- Saber quais os dias da semana em que se verificam maior número de urgências;
- Médias de idades em cada causa de urgência;
- Saber a que horas existem mais urgências.

Todas estas funcionalidades têm por base os indicadores e *dashboards* supracitados, pois através de gráficos será a forma mais intuitiva para os utilizadores fazerem a sua análise. As funcionalidades acima permitirão a quem as utiliza saber quais as áreas que deve reforçar para dar melhores respostas e garantir assim um melhor serviço para os seus utentes.

A aplicação seria desenvolvida com o intuito de ser utilizada a partir de um motor de busca, para que qualquer utilizador tivesse a possibilidade de aceder à aplicação em qualquer lugar, vendo o estado das urgências de forma mais confortável.

Esta aplicação teria uma base de dados *Oracle* que permitisse guardar todas as informações que se pretende mostrar aos utilizadores. Posteriormente, seria desenvolvida recorrendo a *NodeJS*, onde se faria a ligação à base de dados. A informação recolhida em JSON seria depois transformada para que a visualização se tornasse mais amigável para quem está a utilizar a aplicação, através da biblioteca *ChartJS*.

Para uma visão inicial de como seria a aplicação foram feitos os seguintes mockups.

4.1 Página Inicial

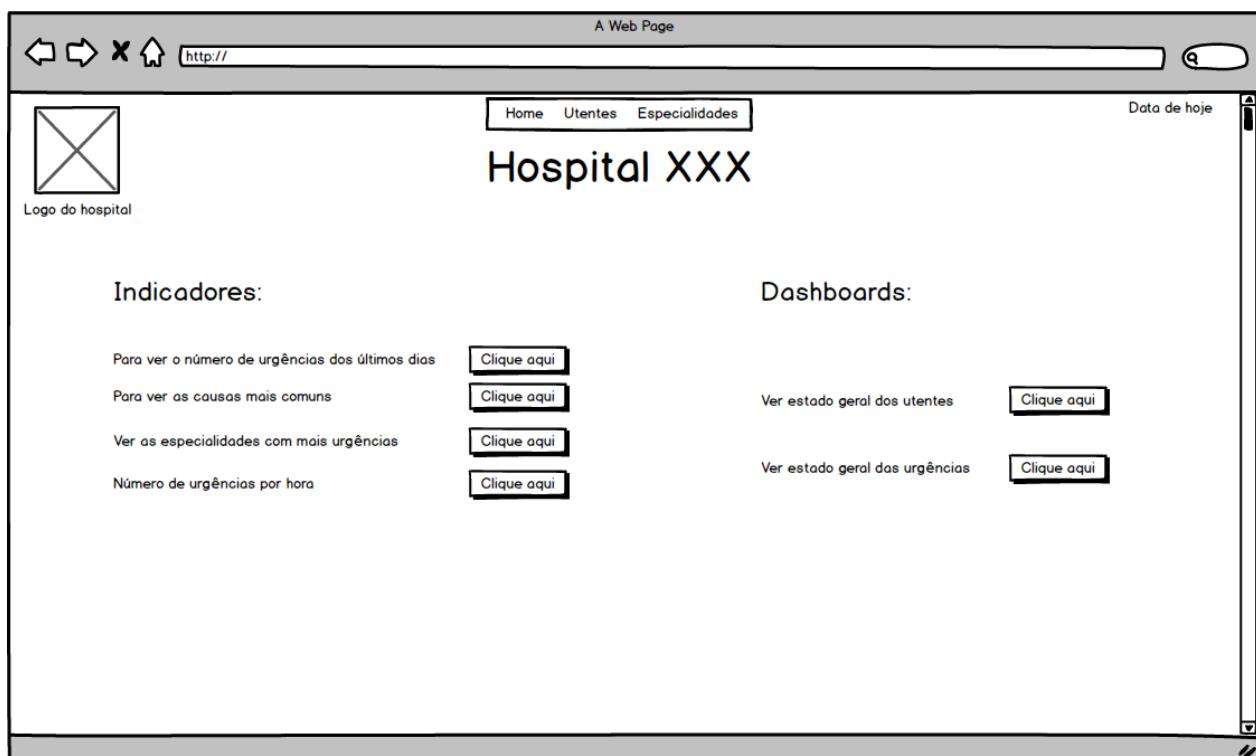


Figura 24 - Página Inicial.

4.2 Número de Urgências nos Últimos Dias

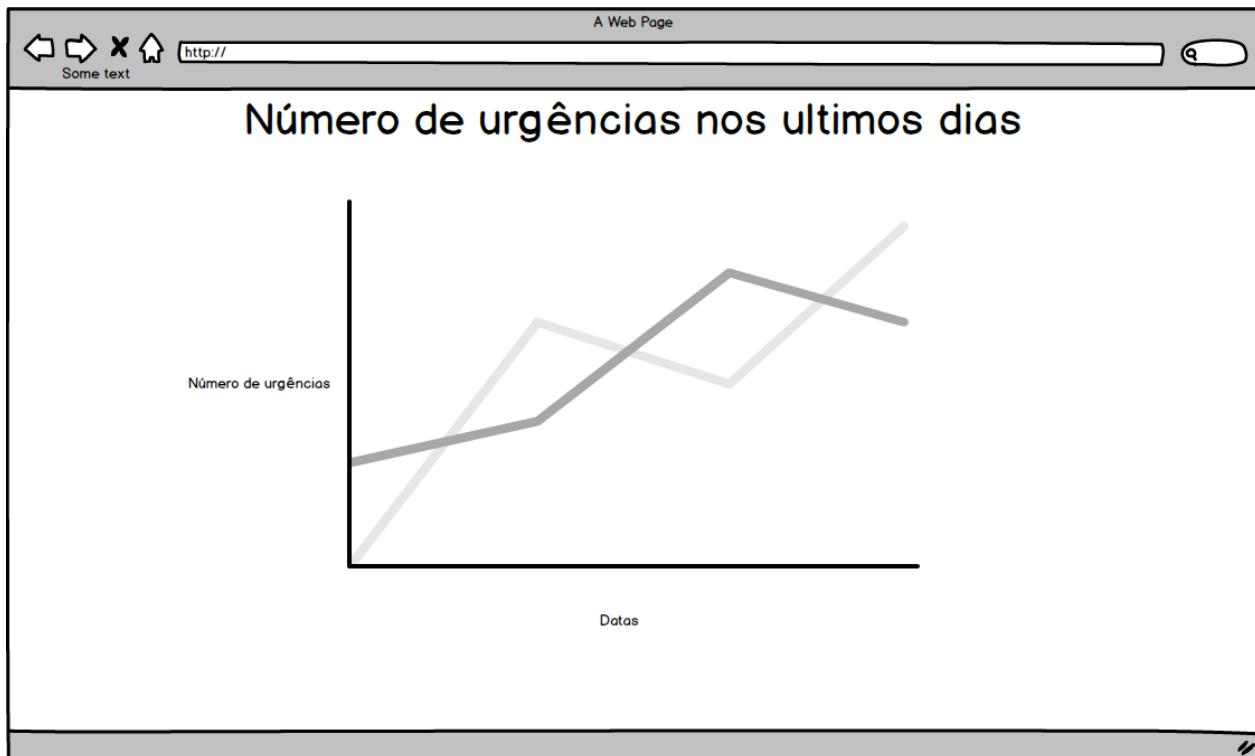


Figura 25 - Número de urgências nos últimos dias.

4.3 Causas Mais Comuns

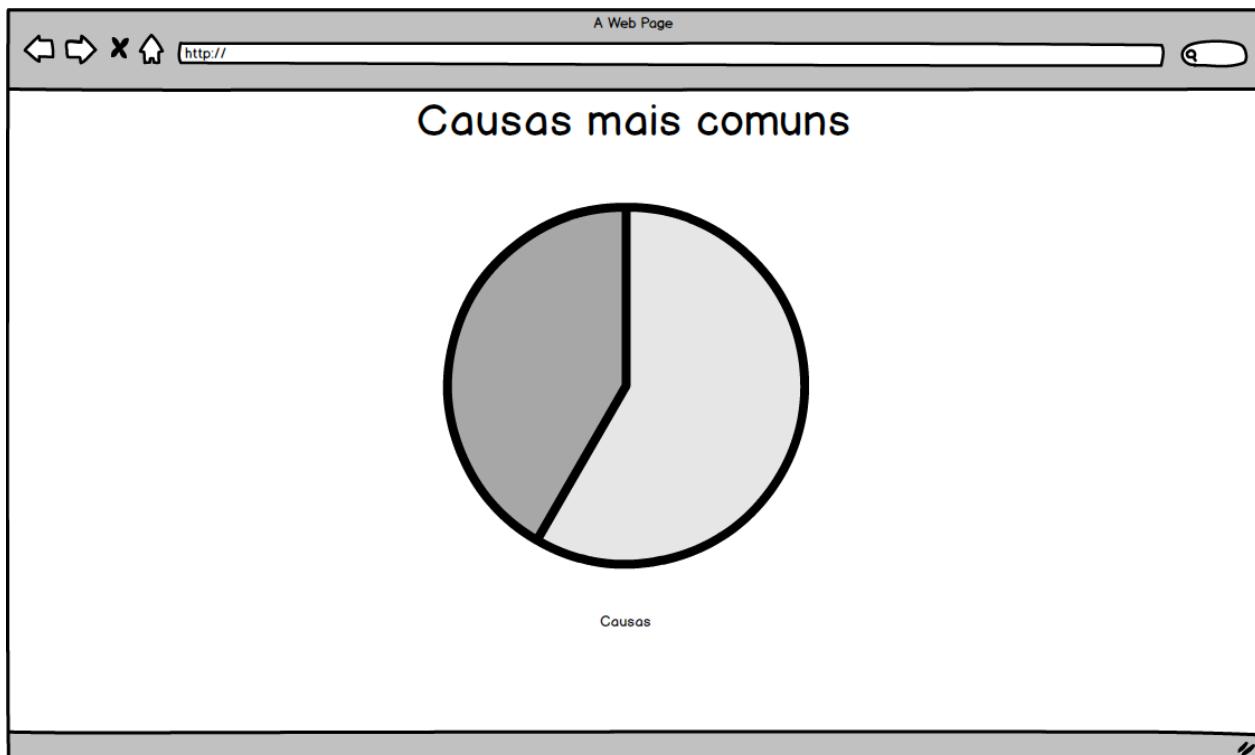


Figura 26 - Causas mais comuns.

4.4 Especialidades com Mais Urgências

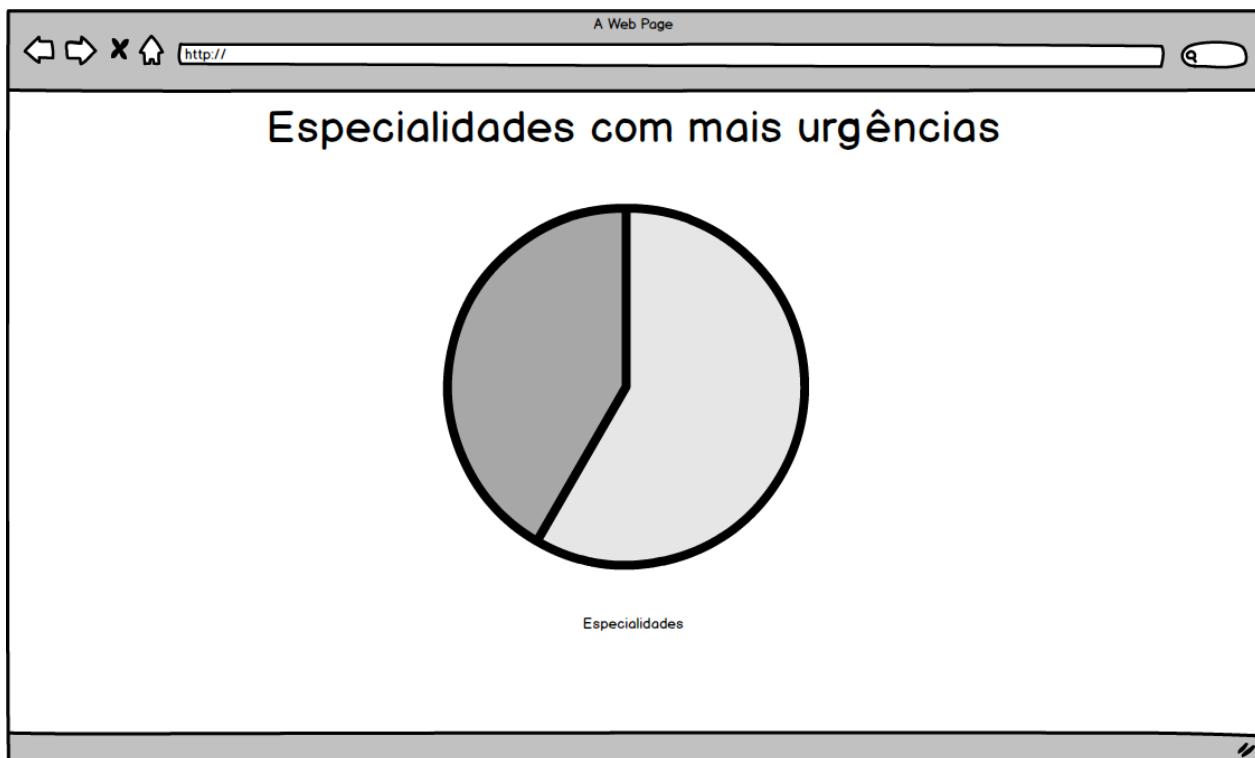


Figura 27 - Especialidades com mais urgências.

4.5 Número de Urgências por Horas

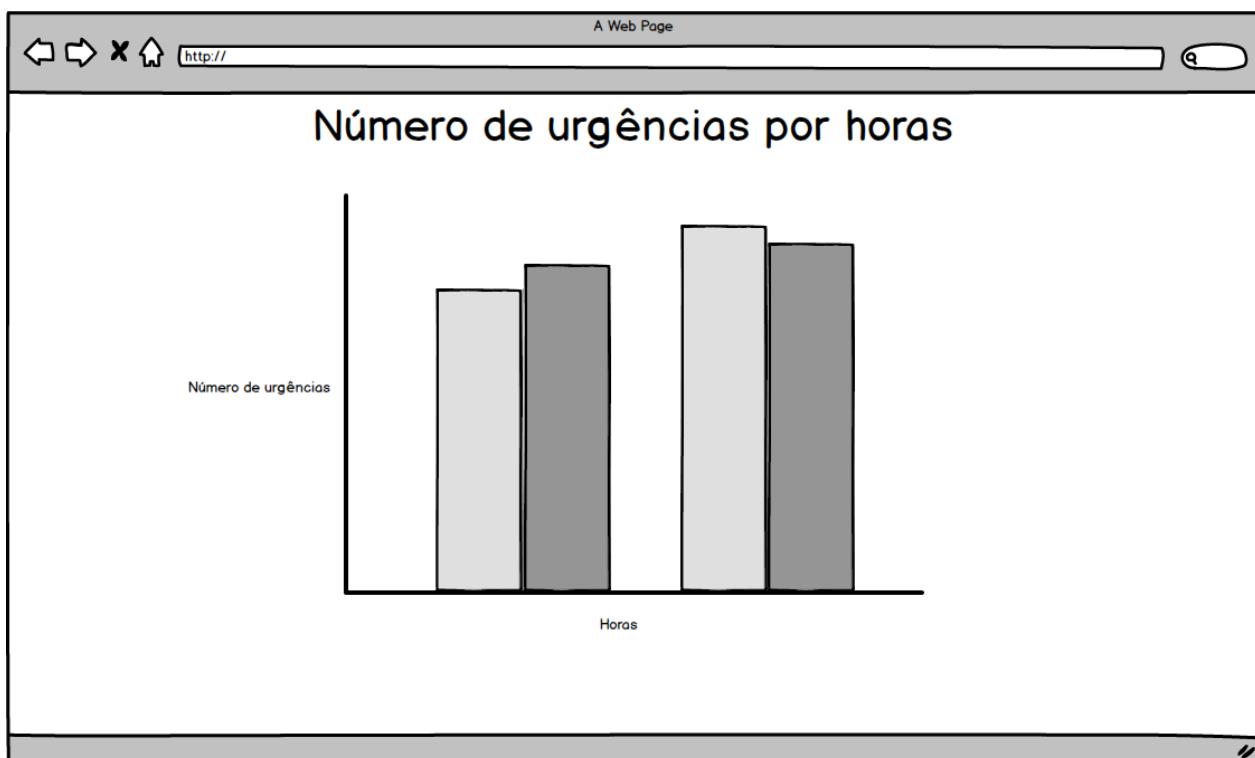


Figura 28 - Número de urgências por horas.

5 Conclusões e Trabalho Futuro

Este projeto permitiu ao grupo solidificar os seus conhecimentos sobre a importância de *Data Warehouses*, o processo de *ETL*, bem como *Business Intelligence*. Assim, através da implementação de todos estes métodos num caso de estudo referente a um ambiente clínico, foi possível extrapolar medidas que ajudariam a um melhor funcionamento do Hospital.

No que toca ao *Data Warehouse* em si, este foi desenhado com base num esquema em estrela, permitindo uma análise bastante rápida e eficiente dos dados. O povoamento do mesmo teve por base duas vias distintas: um *script SQL* ou uma ferramenta *ETL*. No primeiro caso, o grupo sentiu-se bastante confortável, visto que possuía conhecimentos avançados de como manipular os dados desta forma. No que toca ao segundo caso, a ferramenta permitiu obter representações gráficas do fluxo dos dados desde a sua obtenção, até ao seu carregamento. Além disso, permitiu explorar as potencialidades deste tipo de *software*, devido à sua simplicidade e facilidade de implementação.

Por fim, a definição de indicadores clínicos permitiu analisar diversas áreas do Hospital. Esta análise tornou-se bastante simples de implementar, graças à facilidade em representar graficamente dados, tirando proveito de ferramentas *OLAP*, juntamento com o *Data Warehouse* desenvolvido. Dito isto, foi possível observar quais os dias da semana mais movimentados, quais as horas com mais afluência nas urgências, de que forma a idade influência as idas ao Hospital, entre outros.

Em suma, a realização deste trabalho exigiu a aplicação de todos os conhecimentos adquiridos no decorrer das aulas, permitindo que o grupo cumprisse todos os objetivos propostos no enunciado, conciliando assim a teoria e a prática.