



RELATÓRIO DE PROJETO

SISTEMA DE APOIO À DECISÃO PARA O DEPARTAMENTO DE GESTÃO DA FARMACIA DE UM HOSPITAL

Gestão de Produtos Farmacêuticos para Regulação de disponibilidade e Otimização de Custos

Discente: Maria Raquel Quintão, nº2023134

Docentes: Prof. Patrícia Narciso

Prof. Pedro Nascimento

Prof. Marta Sousa

Relatório de Projeto

Índice

Apresentação de Projeto	2
Enquadramento.....	2
Requisitos e Análises pretendidas.....	3
Métricas	5
Dimensões	6
Matriz de requisitos.....	8
Modelo dimensional.....	8
Desenvolvimento - Extract, Transform and Load.....	10
Staging Area.....	10
Data Warehouse.....	18
Considerações finais – desenvolvimento ETL e Data Warehouse.....	28
Visualização de dados – ferramenta FrontEnd em Power BI.....	29
Importação de dados e Desenvolvimento do Modelo.....	29
Desenvolvimento Dashboard.....	31
Conclusões da análise do Dashboard.....	39
Considerações finais – Visualização de Dados.....	40

Relatório de Projeto

Projeto - Gestão de Produtos Farmacêuticos para Regulação de disponibilidade e Otimização de Custos

Enquadramento do processo

O sistema de gestão de stock de farmácia e realização de encomendas num hospital é fundamental para garantir que os medicamentos e produtos farmacêuticos necessários estejam disponíveis de forma adequada, ao mesmo tempo em que os custos sejam controlados.

Os últimos anos refletiram necessidades e gastos particulares excessivos, devido à COVID-19, e pretende-se rever as necessidades de produtos farmacêuticos e otimizar os custos associados com base no último ano (2022).

Através deste ano, pretende-se reajustar a disponibilização e os custos associados a esta “nova realidade”.

Descrição do processo

No seguimento de uma reunião interna chegou-se à conclusão que seria necessário otimizar e regular o sistema de stocks e prever custos de acordo com as dispensas realizadas nesse ano.

Todos os serviços têm um stock próprio sempre disponível, sendo necessária a reposição de stock diária garantindo um tratamento ininterrupto e adequado às necessidades previstas.

Nesse sentido, pretende-se também possibilitar uma análise anual que permita ajustar as necessidades dos produtos disponíveis nos armários de stock próprio, ajustados aos produtos mais dispensados por cada serviço.

Cada serviço tem a sua especialidade, e capacidade, havendo necessidades diferentes baseadas na utilização diária, com custos diferentes.

O departamento de gestão da farmácia hospitalar tem a responsabilidade de verificar, interpretar e regulamentar as necessidades mais adequadas e eficientes do hospital com ênfase nos custos para o hospital.

Relatório de Projeto

É da responsabilidade interna do hospital conseguir ter disponíveis os recursos para que a dispensa ao utente seja possível. As entidades responsáveis são quem comparticipa e devolve ao hospital os custos efetivos gastos por cada dispensa, dependendo do utente (por ex: SNS, ADSE, Independentes).

Requisitos e Análises pretendidas

Os requisitos vão estar associados às quantidades dispensadas e aos custos associados.

Pretende-se:

Analisar gastos do ano anterior para planear stocks para os próximos anos e custos associados.

- Quantidades gastas de cada produto farmacêutico e o seu custo.
- Custo total no ano anterior.

Perceber que produtos farmacêuticos cada serviço deve ter de stock próprio.

- Quais os produtos mais dispensados de cada serviço.

Entender os gastos em produtos farmacêuticos de cada serviço, por utente.

- Calcular o gasto médio de cada serviço, por utente.

Entender quais as entidades que mais comparticipam os produtos farmacêuticos.

- Calcular o gasto total por cada entidade.

Relatório de Projeto

Fonte de dados:

A partir de um ficheiro Excel disponibilizado pelo departamento de gestão de farmácia, apresentam-se 70978 registos com dispensas de produtos farmacêuticos.

A fonte de dados apresenta 17 colunas, respetivamente:

- A: Doente (Identificador do utente)
- B: Nome (completo do utente - dados anonimizados para o presente trabalho)
- C: Data (da dispensa do produto farmacêutico)
- D: N° Doc (n° requisição médica)
- E: Medicamento (identificação do produto farmacêutico)
- F: Descrição (nome do produto farmacêutico e apresentação, podendo este ser desde medicamentos, pensos, Fluidoterapia, Alimentação Parentérica, Alimentação entérica, etc)
- G: Serviço (identificador do serviço)
- H: Descrição.1 (nome do serviço)
- I: Tipo Episódio
- J: Episódio (identificador do episódio)
- K: Descrição.2 (percurso utente)
- L: Lote (do produto farmacêutico)
- M: Qtd. Cons. (quantidade consumida)
- N: Unidade (Tipo de apresentação)
- O: Preço Unit. (Preço unitário)
- P: Valor (valor da dispensa)
- Q: Responsável (Entidade Responsável)

Relatório de Projeto

Métricas – necessárias aos requisitos pretendidos:

- Quantidade dispensada
- Custo unitário
- Custo total por dispensa de medicamento = quantidade dispensada x custo

Categorização

Ao classificar as métricas como aditivas, não aditivas e semi-aditivas, está-se a categorizar como essas métricas podem ser agregadas ao longo das dimensões no modelo dimensional.

QUANTIDADE DISPENSADA:

Classificação: Aditiva.

Justificação: A quantidade dispensada é uma métrica aditiva, pois faz sentido somar as quantidades dispensadas em diferentes dimensões (por exemplo, somar a quantidade dispensada por produto, por serviço, por data, etc.).

CUSTO UNITÁRIO:

Classificação: Semi-aditiva.

Justificação: A soma direta dos custos unitários não apresenta significado até porque os custos unitários podem variar bastante. Contudo, a média do custo unitário ao longo do tempo pode ser considerada para representar uma tendência, podendo ser usada em análises em períodos específicos, o que torna a métrica semi-aditiva.

CUSTO TOTAL POR DISPENSA DE MEDICAMENTO:

Classificação: Aditiva.

Justificação: O custo total por dispensa de medicamento é uma métrica aditiva, pois é o resultado da multiplicação da quantidade dispensada pelo custo unitário. A soma dos custos totais pode ser feita ao longo das dimensões.

Relatório de Projeto

Definições:

Aditiva: Pode ser agregada ou somada ao longo de todas as dimensões. A soma da métrica para uma dimensão específica é igual à soma das métricas individuais dentro dessa dimensão.

Não Aditiva: Não pode ser agregada ou somada ao longo de todas as dimensões. A métrica mantém sua individualidade em cada dimensão, e a soma em níveis mais altos de agregação não tem significado.

Semi-Aditiva: Algumas dimensões podem permitir agregação, enquanto outras não. Em outras palavras, a métrica é aditiva em algumas dimensões e não aditiva em outras. Podendo ser aditiva em alguns contextos, mas não em todos.

Estas classificações ajudam na construção de consultas analíticas e na compreensão de como as métricas podem ser combinadas em diferentes níveis de granularidade no sistema de apoio à decisão.

Dimensões – necessárias à análise e classificação:

A escolha do tipo de Slowly Changing Dimension dependerá da natureza dos dados e dos requisitos específicos do projeto. Foi necessário considerar fatores como a frequência de mudanças, a necessidade de rastrear históricos e o impacto no desempenho das consultas ao projetar o modelo dimensional.

1. PRODUTO FARMACÊUTICO:

Atributos: Chave do produto, Nome, Tipo de unidade.

Classificação: Tipo 2 (SCD2) - Mudanças históricas importantes são preservadas. Novos registos são criados para cada alteração no produto.

2. UTENTE:

Atributos: Chave do Utente, Nome.

Classificação: Se informações pessoais como nome mudassem ao longo do tempo, então Tipo 2 (SCD2) poderia ser a classificação mais apropriada. No entanto o contexto atual não beneficia de tal histórico e por isso será classificada com Tipo 1 (SCD1).

Relatório de Projeto

3. SERVIÇO:

Atributos: Chave do Serviço, Nome do Serviço.

Classificação: Pode ser Tipo 1 (SCD1) se os atributos não mudarem ou Tipo 2 (SCD2) se houver mudanças históricas importantes.

No contexto atual irei classificar como Tipo 2 uma vez que, se um serviço mudar de nome deverá estar relacionado com uma reestruturação e como tal deverá ser analisado como um novo serviço.

4. ENTIDADE RESPONSÁVEL:

Atributos: Chave da Entidade Responsável, Nome da Entidade Responsável

Classificação: A Entidade Responsável poderá dividir-se em vários ramos ou protocolos ao longo do tempo. Nos dados atuais existem evidências desses acontecimentos relativamente a diferentes protocolos do SNS, nesse sentido classifico como Tipo 2 (SCD2). Assim será fácil distinguir entre os vários protocolos ao longo do tempo.

5. DATA:

Atributos: Chave da Data, Dia, Mês, Ano.

Classificação: A data classifico como Tipo 1 (SCD1), pois a data em si não muda.

Definições:

Tipo 1 (SCD1): Atualiza o atributo existente sem manter o histórico. A mudança é simplesmente refletida no registo atual.

Tipo 2 (SCD2): Mantém o histórico de mudanças, criando novos registos para refletir as diferentes versões da dimensão ao longo do tempo.

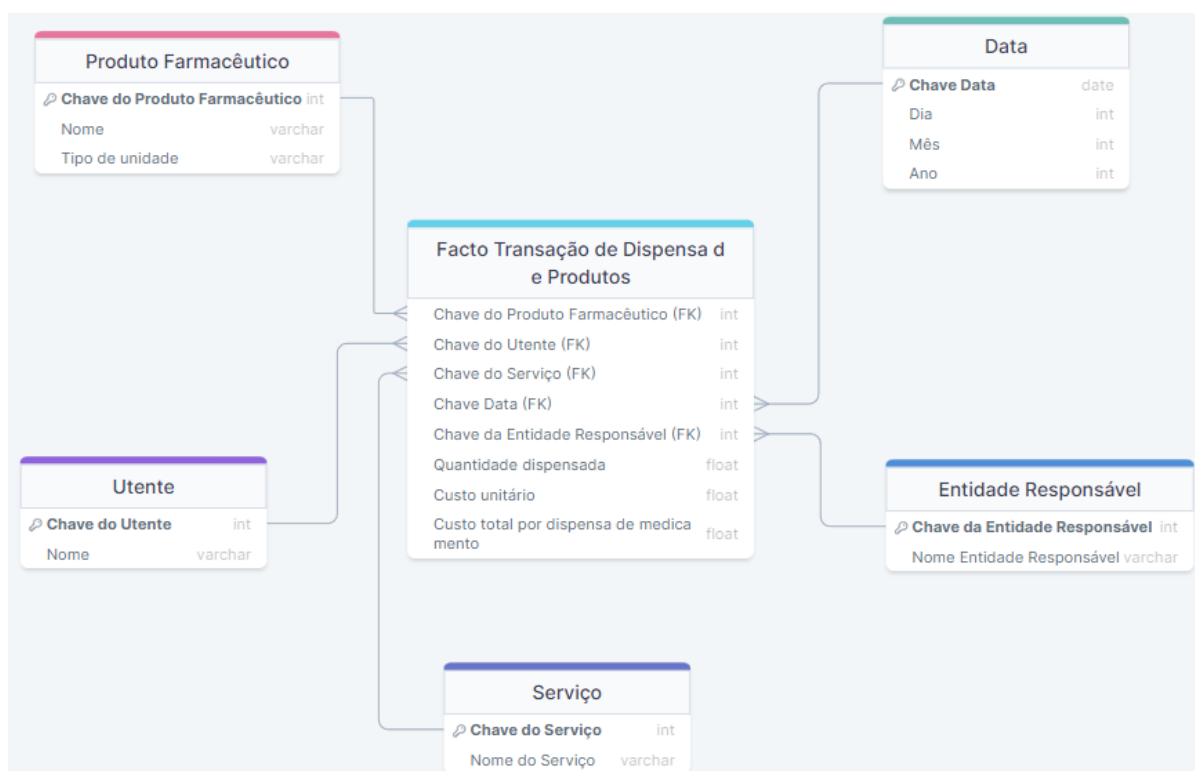
Tipo 3 (SCD3): Mantém uma tabela extra para armazenar versões anteriores da dimensão, permitindo que os usuários consultem dados históricos sem comprometer o desempenho das consultas.

Relatório de Projeto

Matriz de requisitos:

Métricas/ Dimensão	Data	Produto Farmacêutico	Utente	Serviço	Entidade Responsável	Observações
TF Dispensa de Produto	X	X	X	X	X	Facto Transação
Quantidade dispensada	X	X	X	X	X	Métrica Aditiva (calculada em unidades)
Custo unitário	X	X	X	X	X	Métrica semi-aditiva (calculada em euros)
Custo total por dispensa de medicamento	X	X	X	X	X	Métrica Aditiva (calculada em euros)

Modelo Dimensional – Star Schema



Tipo de Tabela de Facto:

Uma tabela de factos, é uma parte central de um modelo dimensional em um *data warehouse*. Neste modelo, apresenta-se como uma tabela de facto de transação, pois armazena informações detalhadas sobre o evento específico – Dispensa de produto farmacêutico num determinado período de tempo (2022).

Relatório de Projeto

Uma tabela de fato de transação, num modelo dimensional, é uma estrutura de banco de dados que regista detalhes granulares sobre eventos ou transações num dado período de tempo, incluindo métricas numéricas, informações temporais e chaves estrangeiras que se relacionam com as tabelas de dimensão para uma análise detalhada e eficiente dos dados.

Através de um Facto de transação consegue-se a aplicação para uma visualização mais detalhada de produtos farmacêuticos a cada dispensa.

Granularidade:

- Uma dispensa de produto farmacêutico.

Menor granularidade vai dar mais flexibilidade para responder aos requisitos solicitados.

Desta forma, permite ao cliente a capacidade de fazer *drill down* em caso de deteção de anomalia de stock ou de custos.

Relatório de Projeto

Desenvolvimento - Extract, Transform and Load:

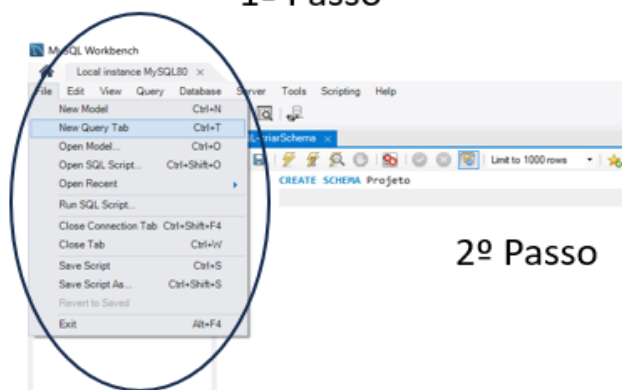
Staging Area

O objetivo de uma ferramenta de ETL é ajudar a trabalhar os dados desde o formato original até ao formato que pretendemos colocar no Data Mart (destino).

A origem para o processo de extração, neste projeto, foi a partir de 1 único ficheiro/fonte de dados Excel. A partir de vários processos de ETL com a mesma fonte, foram carregadas tabelas destino diferentes. Ou seja, foram realizados diferentes jobs com a mesma fonte para diferentes tabelas destino.

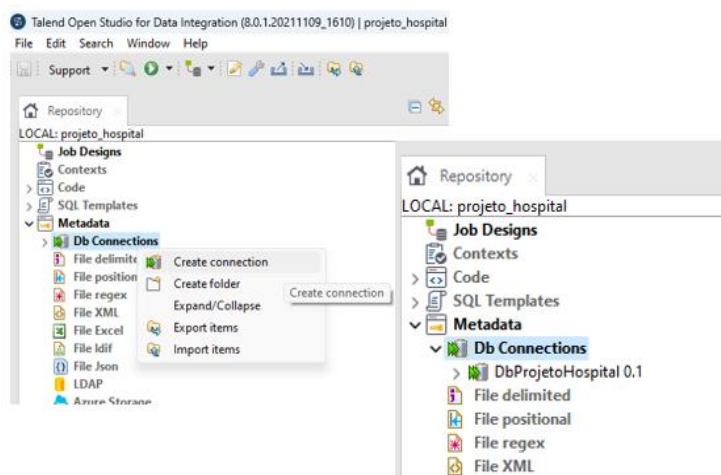
O processo ETL teve início com a criação de um novo Schema na plataforma MySQL, com o nome “Projeto”.

1º Passo



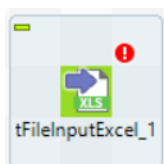
2º Passo

A partir da plataforma Talend foi criado o “projeto_hospital” e foi realizada a conexão à base de dados MySQL.

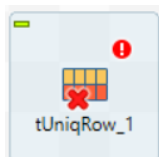


Relatório de Projeto

Neste sentido, elaborei no total 5 jobs de *staging area*, 4 para a criação das dimensões e 1 para a criação da Tabela de Facto. Em cada Job extraio a informação necessária e aplico métodos de transformação para o carregamento dos dados posteriormente (tabelas destino).



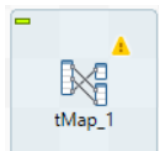
O componente tFileInputExcel, comum a todos os Jobs, permite a leitura de dados a partir de arquivos Excel durante o processo de integração de dados. As definições do mesmo são carregadas através de item File Excel na secção Metadata. Para cada Job criei um item File Excel de forma a selecionar as colunas necessárias para cada job.



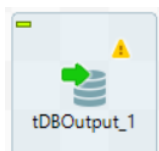
O componente tUniqRow, é comum em todos os jobs para criação das tabelas. Este permite a eliminação de duplicados resultando em registos únicos, tal como uma função distinct. Para tal acontecer é necessário selecionar a coluna ou colunas que pretendemos que não se repitam.



O componente tMySQLInput permite a leitura de dados a partir de uma base dados MySQL (representado como tDBInput). Nesta fase de *staging area* será apenas usado no job para a construção da tabela que servirá de core para a tabela de fato, e com o propósito de consultar a chave correspondente à Entidade Responsável.



O componente tMap, comum a todos os jobs, permite joins, filtragem de colunas e/ou registos, transformações e múltiplos outputs. Cada job terá a sua configuração de acordo com as necessidades.

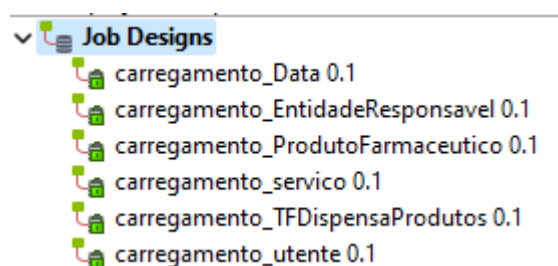


O componente tMySQLOutput, utilizado em todos os jobs permite o carregamento de dados processados pelo Talend numa base de dados MySQL (representado como tDBOutput). Neste projeto será usado sempre como item Db Connections

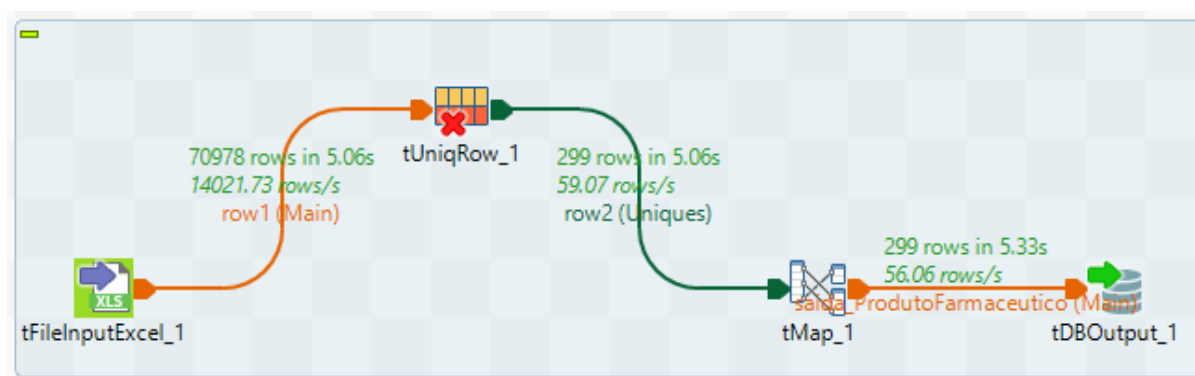
Relatório de Projeto

DbProjectoHospital mencionado anteriormente. Em todos os jobs será configurado com a “Action on table” a “Drop table if exists and create” para que sempre que se execute o job crie uma tabela nova e limpa de iterações anteriores, no entanto todos os jobs terão tabelas de output diferentes.

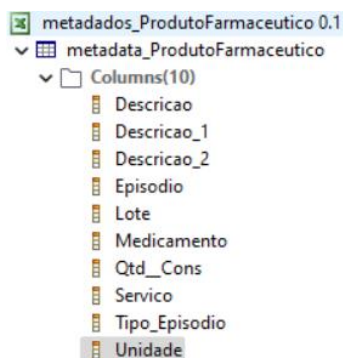
Jobs desenvolvidos:



Job carregamento_ProdutoFarmaceutico



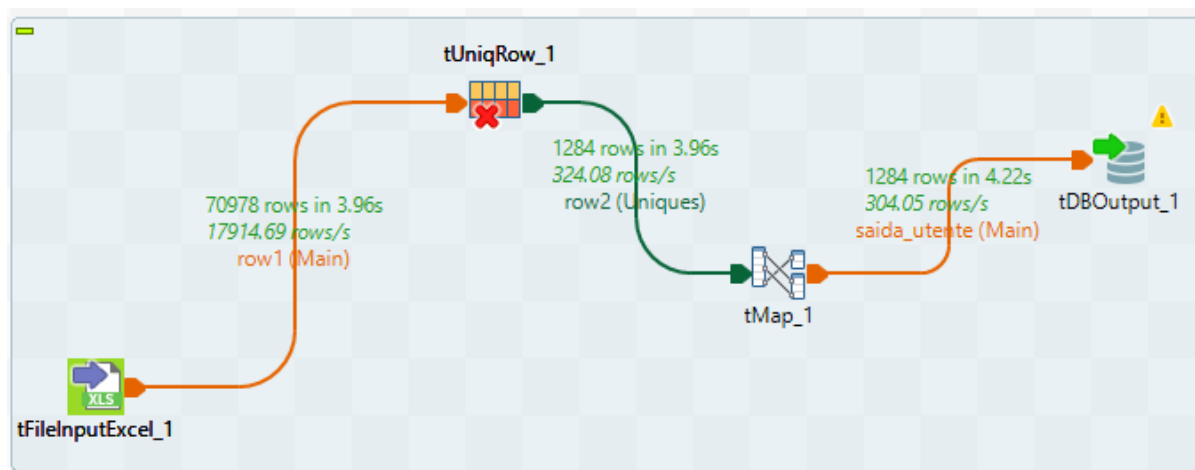
- Neste job usei no componente tFileInputExcel o seguinte item File Excel:



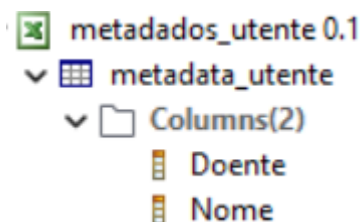
Relatório de Projeto

- Na componente tUniqRow seleccionei como key a coluna Medicamento, correspondente à coluna do identificador único do produto farmacêutico ligando ao tMap com a opção Uniques, para ter a lista de produtos farmacêuticos sem repetições.
- No tMap seleccionou-se unicamente as colunas necessárias à dimensão Medicamento, Descricao e Unidade e renomeando para ProdutoFarmaceutico_id, Nome_Produto e Tipo_Unidade respectivamente. Devido a erros de processamento na execução do job foi necessário adaptar os comprimentos das colunas de entrada e de saída.
- No componente tMySQLOutput, deu-se o nome à tabela de destino "STG_PRODUTOFARMACEUTICO".

Job carregamento_utente



- Neste job usei no componente tFileInputExcel o seguinte item File Excel:



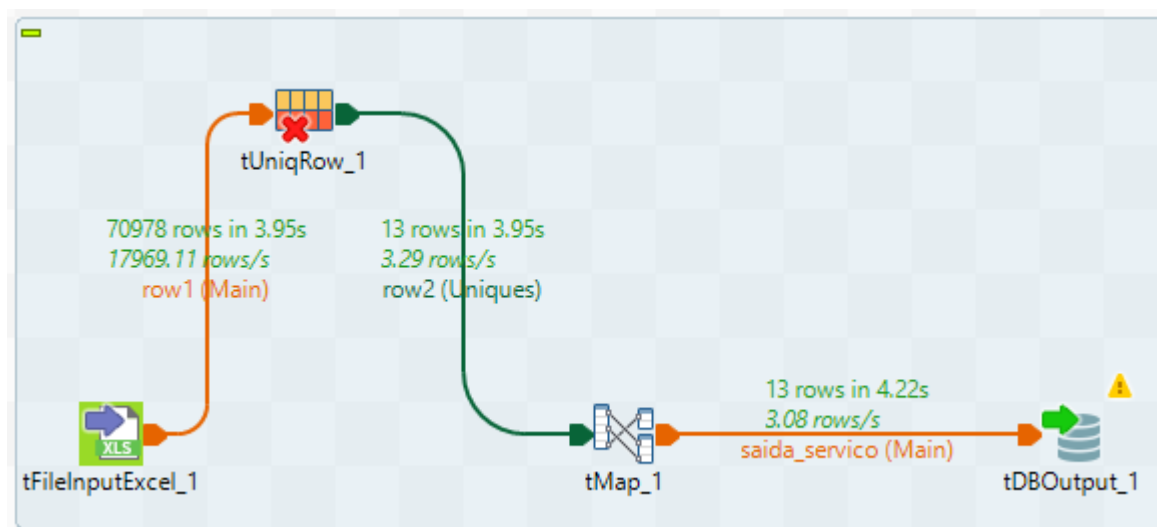
- Na componente tUniqRow seleccionei com key a coluna Doente, correspondente a coluna do identificador único do Utente ligando ao tMap com a opção Uniques, para ter a lista de todos os Utentes sem repetições.
- No tMap seleccionou-se unicamente as colunas necessárias para a posterior criação da dimensão, Doente e Nome e renomeando para Utente_id e Nome respetivamente.

Relatório de Projeto

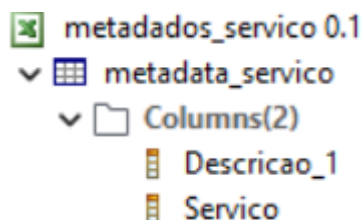
Devido a erros de processamento na execução do job foi necessário adaptar os comprimentos das colunas de entrada e de saída.

- No componente tMySQLOutput, deu-se o nome à tabela de destino "STG_UTENTE".

Job carregamento_servico



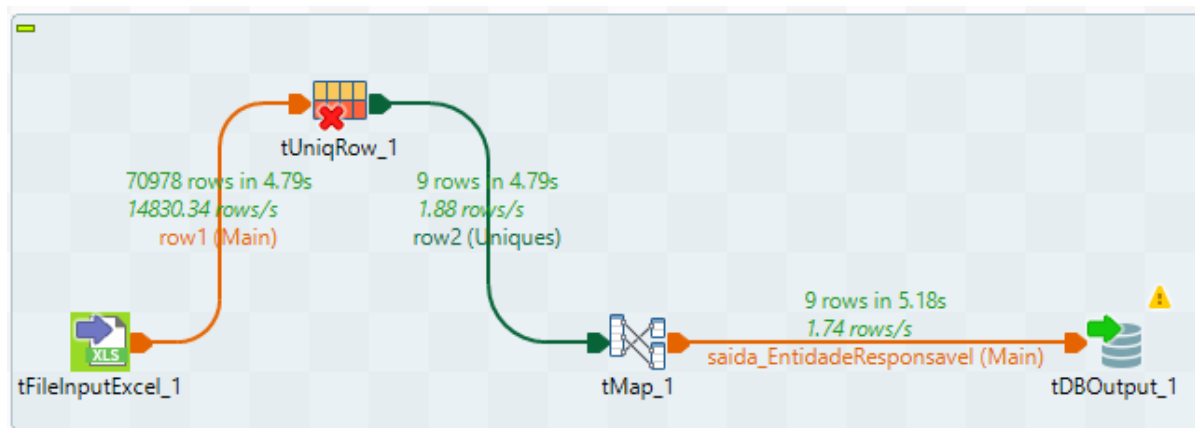
- Neste job usei no componente tFileInputExcel o seguinte item File Excel:



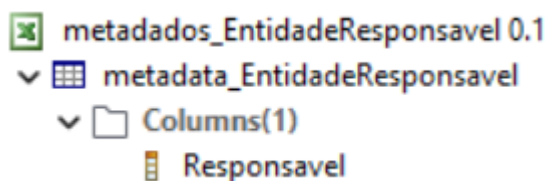
- Na componente tUniqRow seleccionei com key a coluna Servico, correspondente a coluna do identificador único do Serviço ligando ao tMap com a opção Uniques, para ter a lista de todos os Serviços sem repetições.
- No tMap seleccionou-se unicamente as colunas necessárias à dimensão, Servico e Descricao_1 e renomeando para Sevcico_id e Nome_servico respetivamente. Devido a erros de processamento na execução do job foi necessário adaptar os comprimentos das colunas de entrada e de saída.
- No componente tMySQLOutput, deu-se o nome à tabela de destino "STG_SERVICO".

Relatório de Projeto

Job carregamento_EntidadeResponsavel



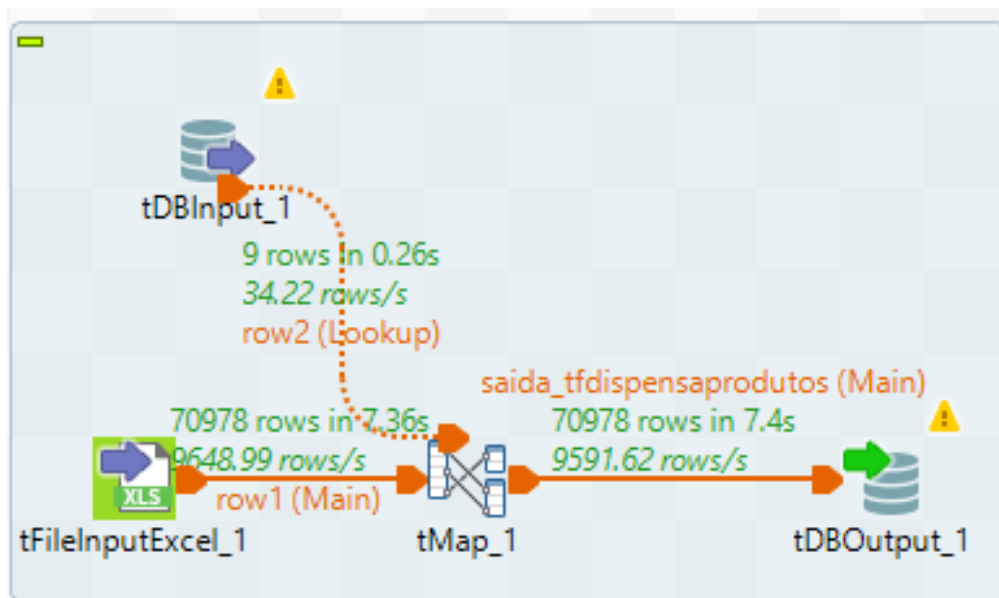
- Neste job usei no componente tFileInputExcel o seguinte item File Excel:



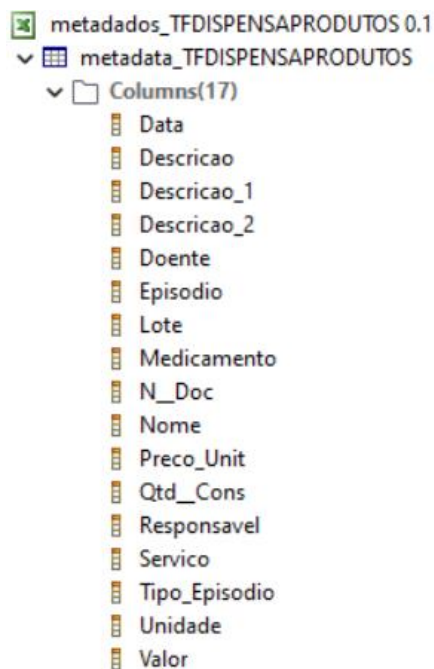
- Na componente tUniqRow seleccionei como key a coluna Responsavel, correspondente à coluna do nome da Entidade Responsável ligando ao tMap com a opção Uniques, para ter a lista de todas as Entidades sem repetições.
- No tMap seleccionou-se a coluna necessária para a posterior criação da dimensão, Responsavel e renomeando para Entidade_Responsavel. No sentido de criar um identificador único para cada Entidade, criou-se uma coluna Entidade_Responsavel_id à qual associou-se a função *Numeric.sequence("s1",1,1)* para gerar os identificadores numéricos inteiros, de modo sequencial para cada registo. Devido a erros de processamento na execução do job foi necessário adaptar os comprimentos das colunas de entrada e de saída.
- No componente tMysqlOutput, deu-se o nome à tabela de destino "STG_ENTIDADERESPONSABEL".

Relatório de Projeto

Job carregamento_TFDispensaProdutos



- Neste job usei no componente tFileInputExcel o seguinte item File Excel, ligando como Main:

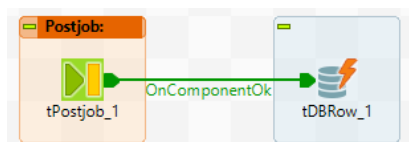


- Usei componente tMySQLInput para extrair da base de dados MySQL o identificador único gerado no job carregamento_EntidadeResponsavel. Ligando ao tMap como lookup. Usei a Query *"select Entidade_Responsavel_id, Entidade_Responsavel from stg_entidaderesponsavel"* para obter a tabela da Entidade Responsável.

Relatório de Projeto

- No tMap foi necessário criar manualmente as colunas na row2 (correspondente à tabela extraída da base dados MySQL), conforme o especificado na query Entidade_Responsavel_id e Entidade_Responsavel identificando a coluna Entidade_Responsavel_id como primary key. Para fazer a correspondência entre o id e o nome da Entidade foi necessário ligar as colunas row1.Responsavel e row2.Entidade_Responsavel. Para tabela de saída foram seleccionadas todas as colunas necessárias e criada a coluna Custo_total cujo valor corresponde à multiplicação das colunas row1.Preco_Unit e row1.Qtd_Cons, conforme demonstrado na imagem abaixo. No sentido de criar um identificador único para cada Dispensa de Produto, criou-se uma coluna DispensaProduto_id à qual associou-se a função *Numeric.sequence("s2",1,1)* para gerar os identificadores numéricos inteiros, de modo sequencial para cada registo, dado que neste contexto de modelo poderão existir combinações iguais nas FK. Devido a erros de processamento na execução do job foi necessário adaptar os comprimentos das colunas de entrada e de saída.

saida_tfdispensaproductos	
Expression	Column
Numeric.sequence("s2",1,1)	DispensaProduto_id
row1.Medicamento	ProdutoFarmaceutico_id
row1.Doente	Utente_id
row1.Servico	Servico_id
row2.Entidade_Responsavel_id	Entidade_Responsavel_id
row1.Data	Data
row1.Qtd_Cons	Quantidade_dispensada
row1.Preco_Unit	Custo_Unitario
row1.Preco_Unit*row1.Qtd_Cons	Custo_total



De modo a alterar o tipo de Data do ficheiro fonte - *varchar* para tipo *Date*, criou-se o Post-job com o componente tDBRow com a query:

Query

```
"UPDATE projeto.stg_tfdispensaproductos
SET Data=STR_TO_DATE(Data,\"%d/%m/%Y\")"
```

- No componente tMysqlOutput, deu-se o nome à tabela de destino "STG_TFDISPENSAPRODUTOS"

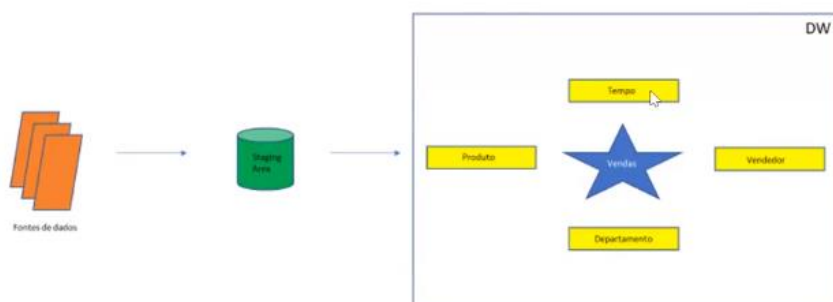
A partir deste processo está criada uma área intermédia para representar o que posteriormente será o Data Warehouse do modelo dimensional.

Relatório de Projeto

Data Warehouse:

Após o desenvolvimento dos jobs da Staging Area, no trimestre anterior, considera-se que se tem condições para seguir para a implementação do Data Warehouse.

A partir deste momento decorre a implementação na ferramenta ETL, carregando a TF e as dimensões parametrizadas à sua tipologia de Slowly Chaging Dimension.

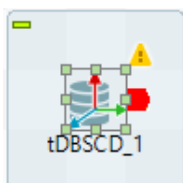


Neste sentido foram criados 7 novos jobs no Talend, 4 para as dimensões, 1 para a tabela de facto, 1 para fazer a limpeza da staging area e 1 para a automatização da execução dos jobs, assim como a criação manual em SQL de 6 tabelas na base de dados MySQL, para a dimensão Data e restantes dimensões.

Todas as tabelas criadas (à exceção da tabela da dimensão data) existe uma coluna SK (chave artificial), usada como *primary key* para facilitar a gestão das dimensões e assegurar um comportamento constante. A chave natural foi colocada como índice da tabela de forma a acelerar a ligação e junção das tabelas ao invés da SK porque não tem significado real.

Nas SCD do tipo II era necessário ter a capacidade de validar o registo consoante o dia da dispensa, por isso nessas dimensões as tabelas incluem as colunas DTC_INICIO e DTC_FIM, que indicam o período de validade do registo. Optou-se por seguir o exemplo da aula e começar todos os registos com DTC_INICIO em 1900-01-01 por ser uma data destaque para ser facilmente diferenciada das posteriores alterações guardadas em histórico. Neste sentido criou-se um Postjob em todos os jobs para as dimensões desta tipologia.

Para facilitar e garantir uma correta implementação de todos os campos nas tabelas usadas provenientes do MySQL foi realizada de forma prévia à colocação de componentes a funcionalidade *Retrieve Schema*.



Nos jobs criados, para popular as tabelas das dimensões, recorreu-se a um componente tMysqlSCD. Este componente diferencia-se do tMysqlOutput na medida em que, no tMysqlSCD parametriza-se a tipologia de SCD, Surrogate keys, Source keys assim como as colunas DTC_INICIO e DTC_FIM.

Relatório de Projeto

Tabelas de dimensão:

Dimensão Data

Na dimensão data houve a necessidade de criar uma nova tabela DIM_DATA, com a seguinte *query*:

```
1 • CREATE TABLE DIM_DATA
2   (SK_DATA INTEGER NOT NULL,
3    DES_DATA_DIA VARCHAR(25),
4    DTC_DATA DATE,
5    NUM_ANO INTEGER,
6    NUM_MES INTEGER,
7    NUM_DIA INTEGER,
8    NUM_QUADRIMESTRE INTEGER,
9    NUM_TRIMESTRE INTEGER,
10   DES_QUINZENA VARCHAR(12),
11   DES_QUADRIMESTRE VARCHAR(12),
12   NUM_BIMESTRE INTEGER,
13   DES_BIMESTRE VARCHAR(12),
14   DES_ANO_MES VARCHAR(8),
15   DES_DIA VARCHAR(7),
16   NUM_SEMESTRE INTEGER,
17   DES_MES_ANO_NUMERICO VARCHAR(7),
18   DES_TRIMESTRE VARCHAR(12),
19   IND_FINAL_SEMANA CHAR(1),
20   DES_MES VARCHAR(15),
21   NUM_QUINZENA INTEGER,
22   DES_MES_ANO_COMPLETO VARCHAR(30),
23   DES_SEMESTRE VARCHAR(12),
24   NUM_NIVEL INTEGER,
25   DES_MES_ANO VARCHAR(8),
26   CONSTRAINT DIM_DATA_DIM_KEY_PKT PRIMARY KEY (SK_DATA));
27 • CREATE INDEX TEMPO_IDX ON DIM_DATA (DES_DATA_DIA);
```

Posteriormente para popular os dados da tabela fez-se a criação de um *procedure* semelhante ao criado em aula, mas com os parâmetros adequados para este projeto.

```
43   SET ano_inicio = 2020;
44   SET ANO_FINAL = 2024;
```

Conforme as boas práticas, fez-se *truncate* à tabela antes de correr o *stored procedure*, garantindo que esta estava vazia antes da execução.

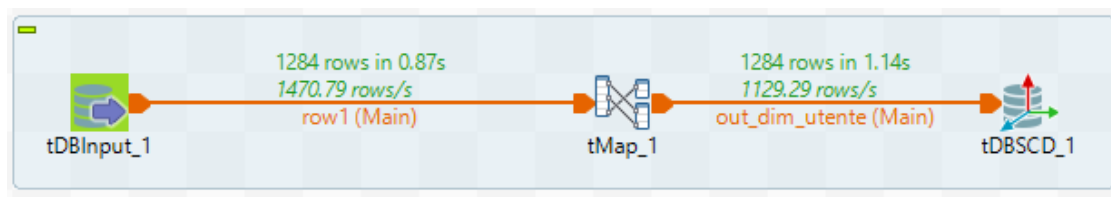
Dimensão Utente

Na dimensão utente houve a necessidade de criar uma nova tabela DIM_UTENTE, com a seguinte *query*:

```
3 • CREATE TABLE DIM_UTENTE
4   (SK_UTENTE INTEGER NOT NULL,
5    UTENTE_ID VARCHAR(20),
6    NOME VARCHAR(200),
7    PRIMARY KEY(SK_UTENTE)
8   );
9 • CREATE INDEX NUTEN_IDX ON DIM_UTENTE (UTENTE_ID);
```

Esta tabela será populada pelo job carregamento_dim_utente:

Relatório de Projeto



- No tDBInput são recolhidos da base de dados MySQL os registos que estão na staging area desta dimensão stg utente.

Database	MySQL	Apply
Property Type	Repository	DB (MYSQL):Db_hospital
DB Version	Mysql 5	
<input type="checkbox"/> Use an existing connection		
Host	"localhost"	
Username	"root"	
Schema	Built-In	Edit schema ...
Table Name	"stg_utente"	
Query Type	Built-In	Guess Query Guess schema
Query	"select * from stg_utente"	

- No tMap são realizadas as correspondências entre as colunas da tDBInput (staging area) com a tDBSCD (dimensão).

row1	Var	out_dim_utente
Column		Expression
Utente_id		SK_UTENTE
Nome		row1.Utente_id
		row1.Nome
		UTENTE_ID
		NOME

- No tDBSCD configurou-se a tabela de output DIM_UTENTE e parametrizou-se da seguinte forma:

[illegible]

Surrogate keys	
name	SK_UTENTE
creation	Table max + 1
complement	

Type 1 fields	
NOME	

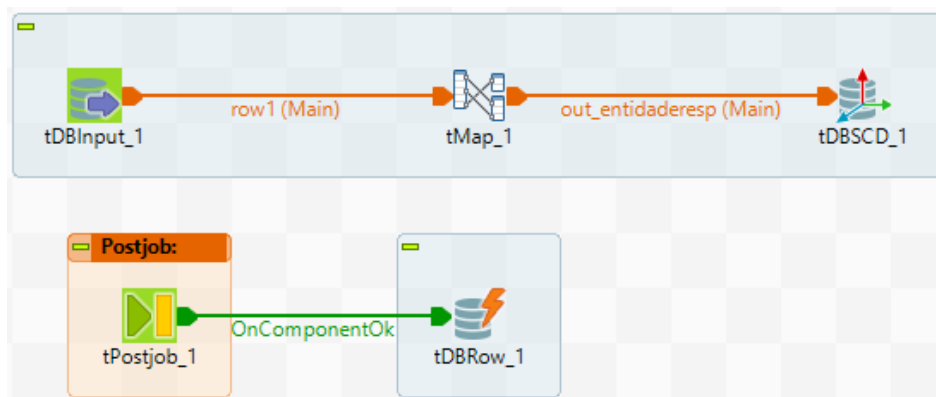
Relatório de Projeto

Dimensão Entidade Responsável

Na dimensão entidade responsável houve a necessidade de criar uma nova tabela DIM_ENTIDADERESPONSAVEL, com a seguinte *query*:

```
44 • CREATE TABLE DIM_ENTIDADERESPONSAVEL
45 • (SK_ENTIDADERESPONSAVEL INTEGER NOT NULL,
46 •   ENTIDADE_RESPONSAVEL_ID INTEGER,
47 •   ENTIDADE_RESPONSAVEL VARCHAR(200),
48 •   DTC_INICIO DATETIME,
49 •   DTC_FIM DATETIME,
50 •   PRIMARY KEY(SK_ENTIDADERESPONSAVEL)
51 • );
52 • CREATE INDEX NENT_IDX ON DIM_ENTIDADERESPONSAVEL (ENTIDADE_RESPONSAVEL_ID)
```

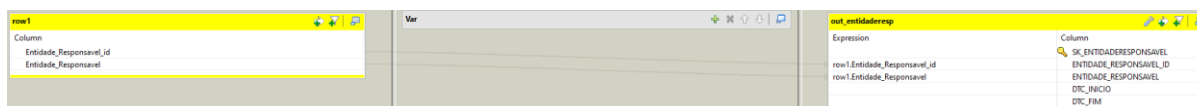
Esta tabela será populada pelo job carregamento_dim_entidaderesponsavel:



- No tDBInput são recolhidos da base de dados MySQL os registos que estão na *staging area* desta dimensão stg_entidaderesponsavel.

Database	MySQL	Apply
Property Type	Repository	DB (MYSQL):Db_hospital
DB Version	Mysql 5	
<input type="checkbox"/> Use an existing connection		
Host	"localhost"	
Username	"root"	
Schema	Built-In	Edit schema ...
Table Name	"stg_entidaderesponsavel"	
Query Type	Built-In	Guess Query Guess schema
Query	"select * from stg_entidaderesponsavel"	

- No tMap são realizadas as correspondências entre as colunas da tDBinput (staging area) com a tDBSCD (dimensão).



Relatório de Projeto

- No tDBSCD configurou-se a tabela de output DIM_ENTIDADERESPONSAVEL e parametrizou-se da seguinte forma:

Unused	
DTC_FIM	
DTC_INICIO	
SK_ENTIDADERESPONSAVEL	

Type 2 fields	
ENTIDADE_RESPONSAVEL	

Source keys	
ENTIDADE_RESPONSAVEL_ID	

Surrogate keys	
name	SK_ENTIDADERESPONSAVEL
creation	Table max + 1
complement	

Versioning			
type	name	creation	complement
start	DTC_INICIO	Job start time	
end	DTC_FIM	NULL	
<input type="checkbox"/> version	scd_version		
<input type="checkbox"/> active	scd_active		

Dimensão Produto Farmacêutico

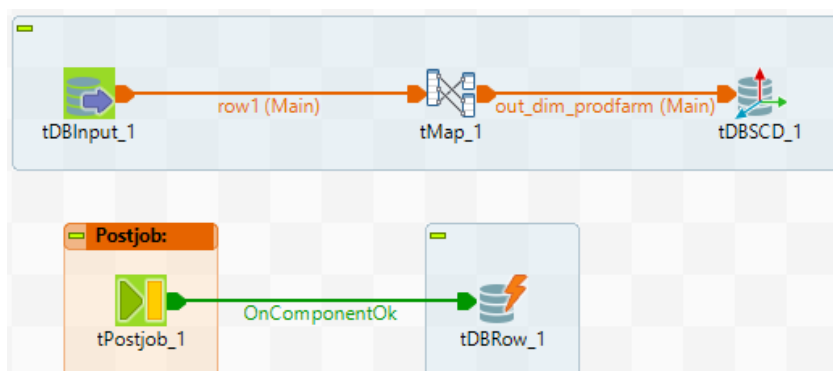
Na dimensão produto farmacêutico houve a necessidade de criar uma nova tabela DIM_PRODUTOFARMACEUTICO, com a seguinte *query*:

```

29 • CREATE TABLE DIM_PRODUTOFARMACEUTICO
30 • (SK_PRODUTOFARMACEUTICO INTEGER NOT NULL,
31 • PRODUTOFARMACEUTICO_ID INTEGER,
32 • NOME_PRODUTO VARCHAR(200),
33 • TIPO_UNIDADE VARCHAR(100),
34 • DTC_INICIO DATETIME,
35 • DTC_FIM DATETIME,
36 • PRIMARY KEY(SK_PRODUTOFARMACEUTICO)
37 • );
38 • CREATE INDEX NPROD_IDX ON DIM_PRODUTOFARMACEUTICO (PRODUTOFARMACEUTICO_ID);

```

Esta tabela será populada pelo job carregamento_dim_produtofarmaceutico:



- No tDBInput são recolhidos da base de dados MySQL os registos que estão na *staging area* desta dimensão stg_produtofarmaceutico.

Relatório de Projeto

Database: MySQL Apply

Property Type: Repository DB (MYSQL):Db_hospital

DB Version: Mysql 5

☐ Use an existing connection

Host: "localhost"

Username: "root"

Schema: Repository DB (MYSQL):Db_hospital - stg_

Table Name: "stg_produtofarmaceutico"

Query Type: Built-In Guess Query Guess schema

Query: "select * from stg_produtofarmaceutico"

- No tMap são realizadas as correspondências entre as colunas da tDBinput (staging area) com a tDBSCD (dimensão).

row1	Var	out_dim_produtofarm
Column		Column
Produtofarmacutico_id		SK_PRODUTOFARMACEUTICO
Nome_Produto		NOME_PRODUTO
Tipo_Unidade		TIPO_UNIDADE
		DTC_INICIO
		DTC_FIM

- No tDBSCD configurou-se a tabela de output DIM_PRODUTOFARMACEUTICO e parametrizou-se da seguinte forma:

Unused	
DTC_FIM	
DTC_INICIO	
SK_PRODUTOFARMACEUTICO	

Type 2 fields	
NOME_PRODUTO	
TIPO_UNIDADE	

Source keys	
PRODUTOFARMACEUTICO_ID	

Surrogate keys	
name	SK_PRODUTOFARMACEUTIC
creation	Table max + 1
complement	

Versioning				
type	name	creation	complement	
start	DTC_INICIO	Job start time		
end	DTC_FIM	NULL		
<input type="checkbox"/> version	scd_version			
<input type="checkbox"/> active	scd_active			

Dimensão Serviço

Na dimensão serviço houve a necessidade de criar uma nova tabela DIM_SERVICO, com a seguinte *query*:

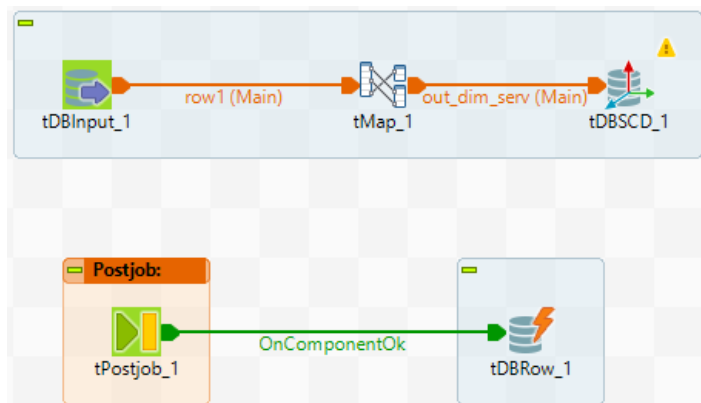
```

15 CREATE TABLE DIM_SERVICO
16 (SK_SERVICO INTEGER NOT NULL,
17  SERVICO_ID VARCHAR(4),
18  NOME_SERVICO VARCHAR(200),
19  DTC_INICIO DATETIME,
20  DTC_FIM DATETIME,
21  PRIMARY KEY(SK_SERVICO)
22 );
23 CREATE INDEX NSER_IDX ON DIM_SERVICO (SERVICO_ID);

```


Relatório de Projeto

Esta tabela será populada pelo job carregamento_dim_servico:



- No tDBInput são recolhidos da base de dados MySQL os registos que estão na *staging area* desta dimensão stg_produtofarmaceutico.

Database: MySQL Apply

Property Type: Repository DB (MYSQL):Db_hospital

DB Version: Mysql 5

☐ Use an existing connection

Host: "localhost" *

Username: "root"

Schema: Repository DB (MYSQL):Db_hospital - stg_servico

Table Name: "stg_servico"

Query Type: Built-In Guess Query Guess schema

Query: "select* from stg_servico"

- No tMap são realizadas as correspondências entre as colunas da tDBinput (staging area) com a tDBSCD (dimensão).

row1	Var	out_dim_serv
Column		Column
Service_id		SK_SERVICO
Nome_servico		SERVICO_ID
		NOME_SERVICO
		DTC_INICIO
		DTC_FIM

- No tDBSCD configurou-se a tabela de output DIM_SERVICO e parametrizou-se da seguinte forma:

Unused	Type 2 fields
DTC_FIM	NOME_SERVICO
DTC_INICIO	
SK_SERVICO	

Source keys	Versioning
SERVICO_ID	type
	start
	end
	version
	active

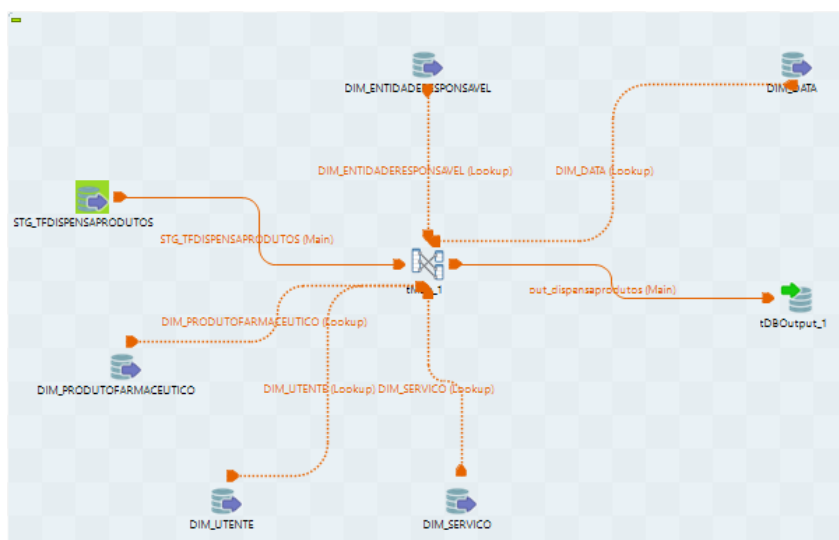
Surrogate keys	creation	complement
name	SK_SERVICO	
creation	Table max + 1	
complement		

name	creation	complement
DTC_INICIO	Job start time	
DTC_FIM	NULL	
scd_version		
scd_active		

Relatório de Projeto

Tabela de Facto:

Para a criação da tabela de facto foi necessário começar por importar a tabela core da informação da tabela de facto (tabela STG_TFDispensaProdutos), ligando-a ao Tmap com a opção Main. As restantes tabelas de dimensão ligam ao Tmap com a opção Lookup.



No tMap cria-se as ligações entre os id's da stg_TFDispensaProdutos às chaves naturais de cada dimensão respetivamente, para a obtenção da correcta SK.

Nas dimensões do tipo II, foi necessário acrescentar uma validação consoante a data da dispensa para fazer a recolha correta do registo válido à data. Conforme o exemplo para a dimensão serviço:

Expression

```
(STG_TFDispensaProdutos.Data.after(DIM_SERVICO.DTC_INICIO) ||
(STG_TFDispensaProdutos.Data.equals(DIM_SERVICO.DTC_INICIO))
&&
(STG_TFDispensaProdutos.Data.before((DIM_SERVICO.DTC_FIM==null)
?TalendDate.parseDate("dd-MM-yyyy", "31-12-2050"):DIM_SERVICO.DTC_FIM))|
```

Esta filter expression foi adaptada para as dimensões Entidade Responsável, Serviço e Produto Farmacêutico, conforme pode ser validado no Talend.

Posteriormente colocou-se na saída do tMap a DispensaProduto_id, as métricas assim como as SKs de cada dimensão. Neste caso foi necessário identificar a coluna DispensaProduto_id como primary key uma vez que no contexto deste projeto a combinação de FKs pode não ser única, esta coluna foi criada em staging area o que permite esta definição.

Relatório de Projeto

STG_TFDDISPENSAPRODUTOS	Var	out_dispensaproductos
Column		Expression
DispensaProduto_id		STG_TFDDISPENSAPRODUTOS.DispensaProduto_id
ProdutoFarmaceutico_id		(DIM_PRODUTOFARMACEUTICO.SK_PRODUTOFARMACEUTICO...
Utente_id		(DIM_UTENTE.SK_UTENTE == 0)?-1: DIM_UTENTE.SK_UTENTE
Servico_id		(DIM_SERVICO.SK_SERVICO == 0)?-1: DIM_SERVICO.SK_SERVICO
Entidade_Responsavel_id		(DIM_DATA.SK_DATA == 0)?-1: DIM_DATA.SK_DATA
Quantidade_dispensada		(DIM_ENTIDADE_RESPONSAVEL.SK_ENTIDADE_RESPONSAVEL == ...
Custo_Unitario		STG_TFDDISPENSAPRODUTOS.Quantidade_dispensada
Custo_Total		STG_TFDDISPENSAPRODUTOS.Custo_Unitario
		STG_TFDDISPENSAPRODUTOS.Custo_Total
		Column
		DispensaProduto_id
		SK_PRODUTOFARMACEUTICO
		SK_UTENTE
		SK_SERVICO
		SK_DATA
		SK_ENTIDADE_RESPONSAVEL
		Quantidade_dispensada
		Custo_Unitario
		Custo_Total

Para uma melhor visualização de erros de ligação entre a stg_TFDDISPENSAPRODUTOS e as Dimensões optei por usar uma expressão para o caso da SK ser igual a 0 e colocá-la a -1. Conforme o exemplo para a dimensão serviço:

`(DIM_SERVICO.SK_SERVICO == 0)?-1: DIM_SERVICO.SK_SERVICO` `SK_SERVICO`

O tDBOutput vai exportar o resultado do tMap para a tabela de facto na base dados MySQL, com as seguintes configurações:

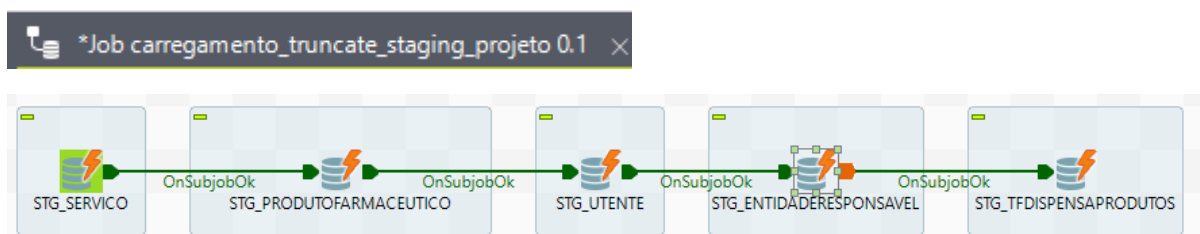
Database	MySQL	Apply
Property Type	Repository	DB (MYSQL):Db_hospital
DB Version	Mysql 5	
<input type="checkbox"/> Use an existing connection		
Host	"localhost"	
Database	"Projeto"	
Username	"root"	
Table	"FACTO_DISPENSAPRODUTOS"	
Action on table	Drop table if exists and create	Action on data Insert
Schema	Built-In	Edit schema ... Sync columns

Por fim, confirma-se no tDBOutput se está parametrizado o Commit a 10000, uma vez que se garante que os dados sejam carregados de forma mais rápida. Este valor dá-se por não ser um número demasiado grande, que possa criar bloqueios, e vai acelerar os registos carregados no MySQL de cada vez.

Jobs de Automatização:

Job carregamento_truncate_staging_projeto

Conforme as boas praticas, deve-se limpar a staging area antes da sua utilização. Nesse sentido foi criado este job para que todas as tabelas sejam eliminadas, ficando esta área livre de registos, o que permitirá a posterior atualização da staging area com um job posterior.



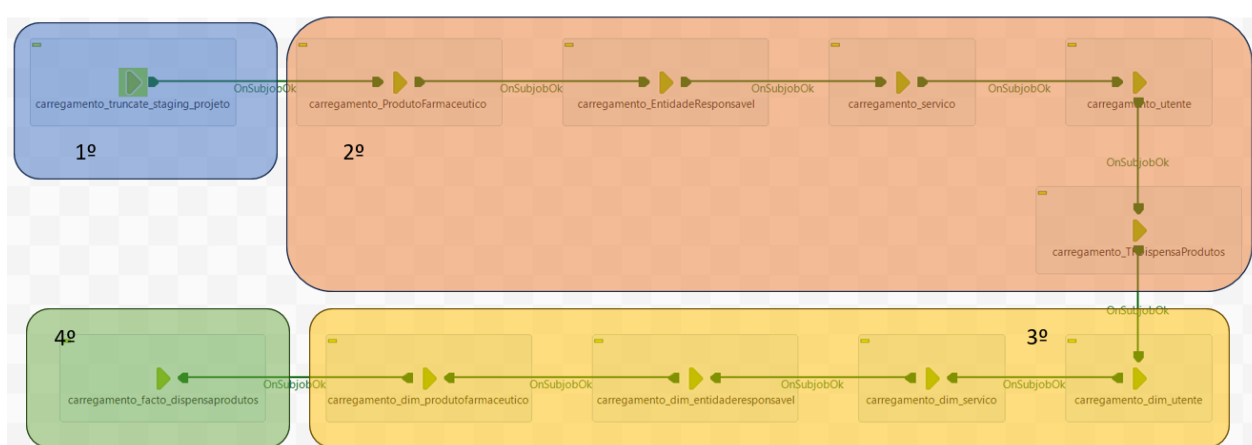
Foram utilizados componentes do tipo tMySQLRow para cada uma das tabelas da staging area com o comando SQL "TRUNCATE TABLE". Cada um destes componentes foram ligados de forma sequencial caso o componente anterior termine corretamente (Trigger>OnSubjobOk). Neste caso a ordem das execuções não é relevante.

Relatório de Projeto

Job carregamento_geral_projeto

Com o objectivo de sistematizar e automatizar o processo de ETL foi criado este job. Este job executa todos os jobs criados anteriormente para este projeto.

1. Começa por limpar a staging area usando o job carregamento_truncate_staging_projeto.
2. Executar todos os jobs que fazem leem a fonte de dados e populam a staging area.
3. Executar todos jobs que carregam as dimensões.
4. Executar o carregamento_facto_dispensaproductos para a criação/actualização da tabela de facto.



Considerações finais – desenvolvimento ETL e Data Warehouse:

Foram necessárias algumas adaptações, criando Post-job's assim como adicionando uma tabela para melhor adequar o processo na *Staging Area*.

Após essa fase, na execução das Dimensões e Tabela de Facto, assim como todo o processo de automatização do processo de ETL obteve sucesso na execução de todos os Jobs, culminando na visualização dos dados na plataforma MySQL. Nesta fase fui testando algumas alterações, para verificar o correto processo na base de dados e verifiquei que corriam de forma adequada com o esperado, guardando histórico ou não de acordo com a tipologia da dimensão, por exemplo. Por fim, deixei conforme os registos originais para manter a base para a seguinte parte de visualização dos dados.

Relatório de Projeto

É de destacar a necessidade de execução dos Jobs por ordem adequada, sendo fundamental para o êxito do projeto. Cada etapa desempenha um papel vital, desde a extração e transformação até ao carregamento final, garantindo a integridade e consistência dos dados.

Esta abordagem permitiu a execução dos dados na tabela de facto, permitindo análises detalhadas e atuais, com possibilidade de atualizações e alterações realizadas ao longo do tempo. Desta forma foi possível de criar uma base de dados com uma adequada relação dinâmica que permitirá ter uma visualização apropriada em ferramentas de visualização.

Relatório de Projeto

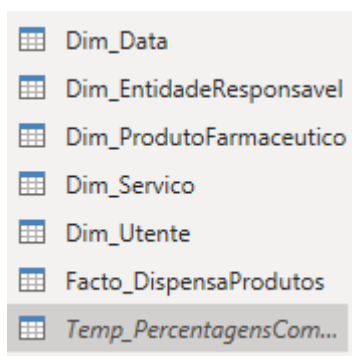
Visualização de dados – ferramenta FrontEnd em Power BI:

Nesta fase de visualização dos dados, carregados na ferramenta MySQL, pretende-se que seja possível de visualizar informação relevante que permitirá dar resposta aos requisitos propostos pelo cliente.

Para esta fase, e de forma a dar uma visão direcionada à realidade aplicada ao enquadramento de projeto prévio, foi desenvolvido este contexto a um hospital fictício – Hospital AlmaCare. Para este cliente pretenderá dar-se resposta aos requisitos propostos através desta ferramenta Power BI. O utilizador frequente desta ferramenta será o gestor do departamento de gestão de farmácia hospitalar.

Importação de dados e Desenvolvimento do Modelo

Para o desenvolvimento do dashboard em Power BI foi realizada a importação de dados em MySQL através de Power Query, selecionando as tabelas Dim_Data, Dim_EntidadeResponsavel, Dim_ProdutoFarmaceutico, Dim_Utente, e Facto_DispenzaProdutos.



Posteriormente foi necessário verificar o perfil, distribuição e a qualidade dos dados, de forma a perceber alguma anormalidade nos dados importados. Este processo foi realizado ativando as opções adequadas para esse processo na aba View do Power Query.

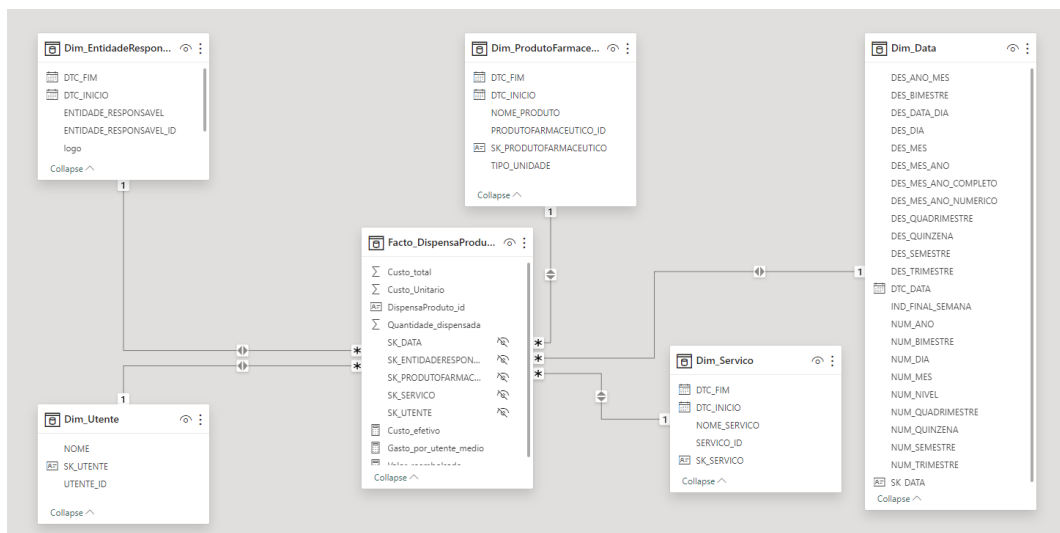
Adicionalmente realizei um “Enter Data” para introduzir uma nova tabela Temp_PercentagensComparticipação para criar dados fictícios relativos às participações de cada Entidade Responsável, dado que não foi possível ter acesso a esses dados. Realizei então um “Merge Queries” para adicionar a coluna Percentagem Participada à Dim_EntidadeResponsável. Optei por adicionar estes dados fictícios para ser possível adaptar esta visualização a hospitais privados em que as seguradoras participam parte dos tratamentos.

Relatório de Projeto

A ^B _C Entidade Responsavel		A ^B _C Percentagem Comparticipada	
Valid	100%	Valid	100%
Error	0%	Error	0%
Empty	0%	Empty	0%
9 distinct, 9 unique		7 distinct, 5 unique	
1	1	0.55	
2	2	0.85	
3	3	0.7	
4	4	0	
5	5	0	
6	6	0.75	
7	7	0.75	
8	8	0.9	
9	9	0.8	

Para evitar alterações ao modelo por parte desta tabela, foi desativada a opção “enable load”, escondendo-a do modelo dimensional (Model view).

De forma a facilitar e prevenir erros posteriores na análise dos dados e transformações realizadas foi primeiramente preparado o modelo dimensional, verificando as ligações geradas automaticamente pela ferramenta, ajustando os nomes das tabelas, as customizações dos dados agregáveis e a sua tipologia e por fim adequar as relações do modelo de forma a permitir e verificar as corretas filtragens, posteriormente, no momento da análise. De acordo com o modelo em imagem:



Relatório de Projeto

Desenvolvimento Dashboard

Recordando, o cliente apresentou os seguintes requisitos:

Pretende-se:

Analisar gastos do ano anterior para planear stocks para os próximos anos e custos associados.

- Quantidades gastas de cada produto farmacêutico e o seu custo.
- Custo total no ano anterior.

Perceber que produtos farmacêuticos cada serviço deve ter de stock próprio.

- Quais os produtos mais dispensados de cada serviço.

Entender os gastos em produtos farmacêuticos de cada serviço, por utente.

- Calcular o gasto médio de cada serviço, por utente.

Entender quais as entidades que mais comparticipam os produtos farmacêuticos.

- Calcular o gasto total por cada entidade.

Para conseguir responder aos requisitos estipulados foi necessário criar as seguintes métricas em Power BI através das seguintes fórmulas:

Custo efetivo:

`SUM(Facto_DispensaProdutos[Custo_total]) - [Valor_reembolsado]`

-Permitindo ao cliente ter a possibilidade de verificar os custos reais que acarreta do valor gasto total.

Valor reembolsado:

`SUMX(Facto_DispensaProdutos, Facto_DispensaProdutos[Custo_total]*RELATED(Dim_EntidadeResponsavel[Percentagem Comparticipada]))`

- Calcular o gasto total por cada entidade.

Gasto por utente médio:

`SUM(Facto_DispensaProdutos[Custo_total])/DISTINCTCOUNT(Dim_Utente[UTENTE_ID])`

- Calcular o gasto médio de cada serviço, por utente.

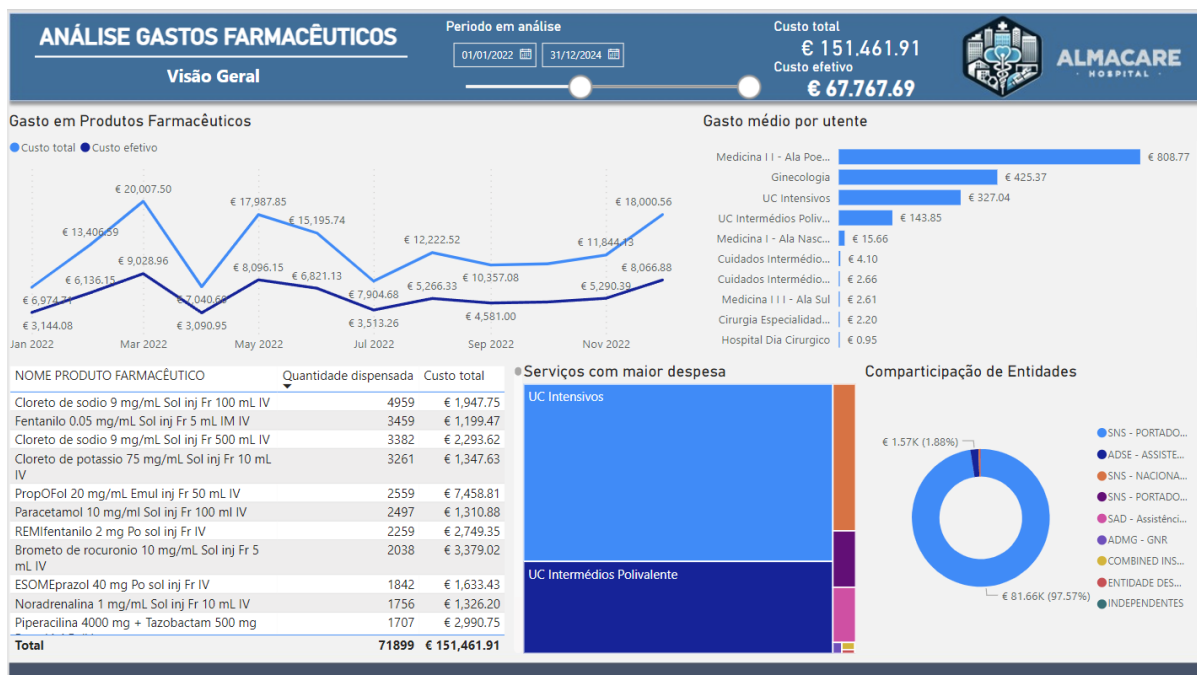
Relatório de Projeto

Para as métricas anteriormente geradas foram agregadas da seguinte forma:

Σ Custo_Unitario	Agregada por: Average (média)
Σ Custo_total	Agregada por: Sum (somatório)
Σ Quantidade_dispensada	Agregada por: Sum (somatório)

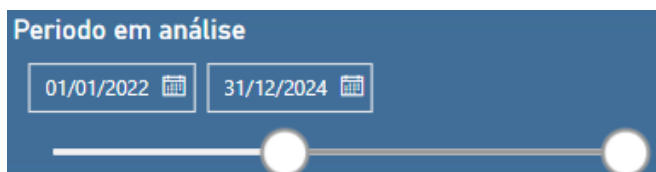
Para melhor responder aos requisitos do cliente foi criado um relatório de Power BI com 3 páginas:

- **Visão Geral** - permite ao cliente uma visualização dos principais gastos, serviços e entidades responsáveis com maior peso no orçamento.



Relatório de Projeto

- O Cliente pode efetuar um filtro por período temporal para melhor entender a variação de dispensas ao longo do ano.



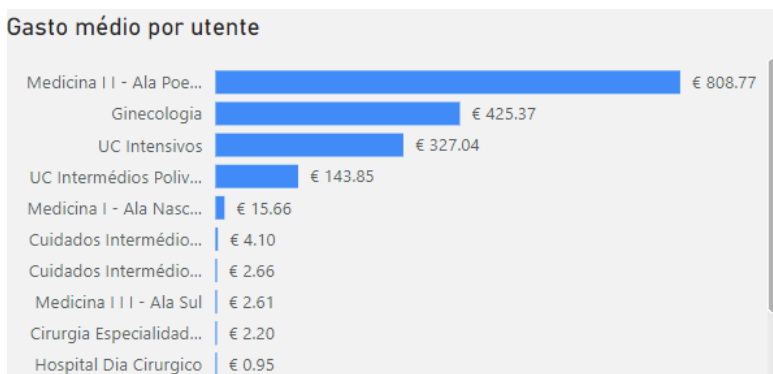
- Pode também rapidamente comparar referente ao período lecionado o custo total em produtos farmacêuticos e o seu custo efetivo.



- Assim como a sua variação ao longo do período selecionado no gráfico de linhas.

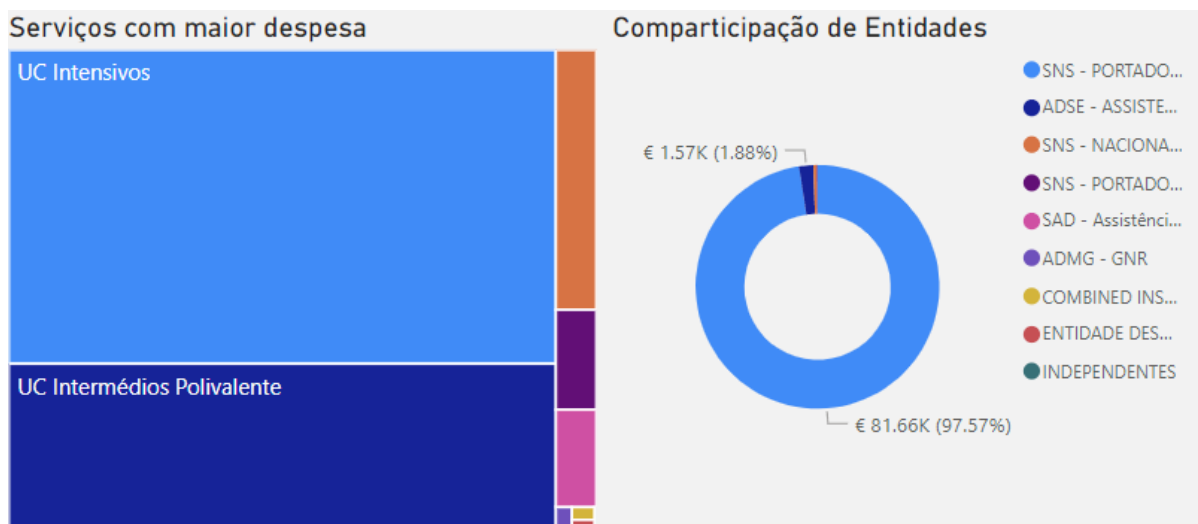


- Para permitir a comparação equitativa entre serviços e os seus gastos escolhi representar num gráfico de barras agrupadas com custo médio por utente.



Relatório de Projeto

- No sentido de facilitar uma rápida observação dos serviços que mais gastam e entidades que mais estão a compartilhar os produtos farmacêuticos, foram criados um mapa de árvore e um gráfico de donut. Esta escolha deve-se ao facto de o cliente querer entender quais os principais representantes de cada um deles, embora se tenha criado condições noutras páginas para uma análise mais detalhada.



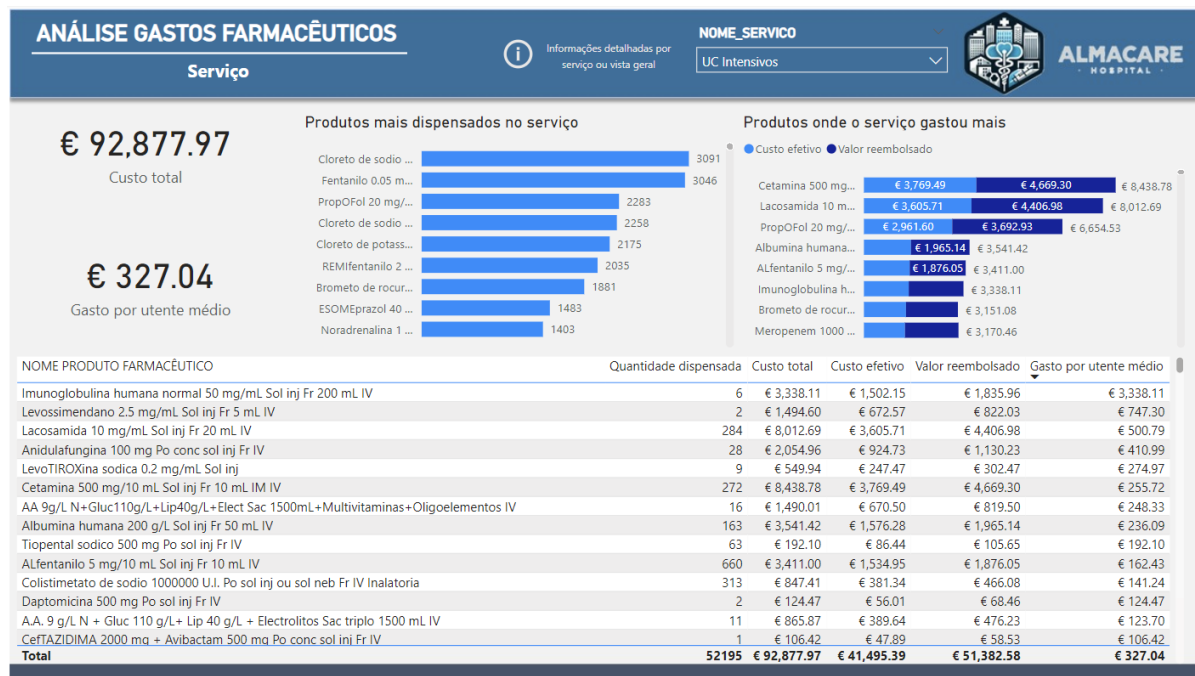
- Para responder ao pedido do cliente sobre identificar quais as quantidades gastas de cada um dos produtos farmacêuticos e o seu custo foi disponibilizada uma tabela.

NOME PRODUTO FARMACÊUTICO	Qtd. dispensada	Custo total
Cloreto de sodio 9 mg/mL Sol inj Fr 100 mL IV	4959	€ 1,947.75
Fentanilo 0.05 mg/mL Sol inj Fr 5 mL IM IV	3459	€ 1,199.47
Cloreto de sodio 9 mg/mL Sol inj Fr 500 mL IV	3382	€ 2,293.62
Cloreto de potassio 75 mg/mL Sol inj Fr 10 mL IV	3261	€ 1,347.63
PropOFol 20 mg/mL Emul inj Fr 50 mL IV	2559	€ 7,458.81
Paracetamol 10 mg/ml Sol inj Fr 100 ml IV	2497	€ 1,310.88
REMI fentanilo 2 mg Po sol inj Fr IV	2259	€ 2,749.35
Brometo de rocuronio 10 mg/mL Sol inj Fr 5 mL IV	2038	€ 3,379.02
ESOMEprazol 40 mg Po sol inj Fr IV	1842	€ 1,633.43
Noradrenalina 1 mg/mL Sol inj Fr 10 mL IV	1756	€ 1,326.20
Piperacilina 4000 mg + Tazobactam 500 mg Po sol inj Fr IV	1707	€ 2,990.75
LevETIRacetam 100 mg/mL Sol inj Fr 5 mL IV	1458	€ 2,085.24
Total	71899	€ 151,461.91

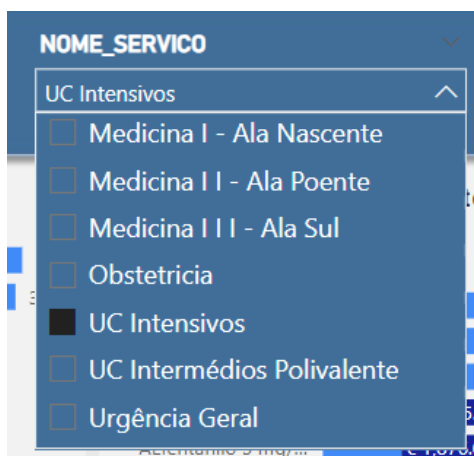
- Com este dashboard consegue-se dar resposta aos pontos:
 - Analisar gastos do ano anterior para planear stocks para os próximos anos e custos associados.
 - Entender os gastos em produtos farmacêuticos de cada serviço, por utente.

Relatório de Projeto

- **Serviço** - permite ao cliente uma visualização mais detalhada por serviço.



- O cliente pode filtrar os dados por serviço para permitir verificar as suas necessidades individuais.

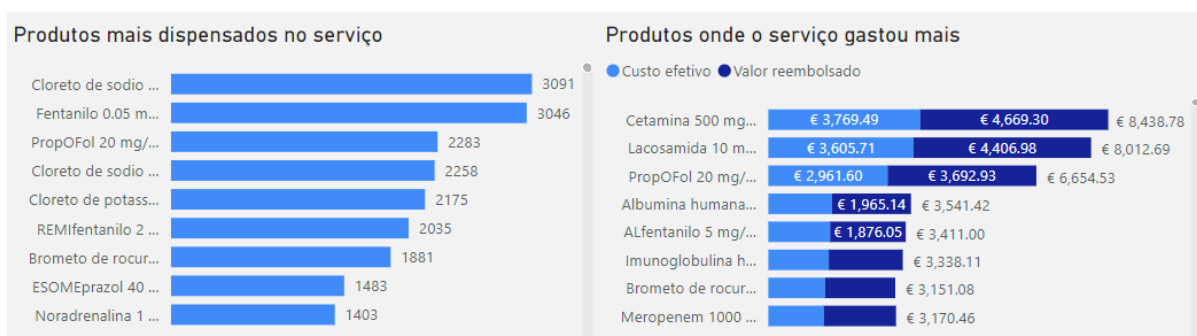


- Imediatamente consegue consultar o Custo total do serviço e correspondente gasto por utente.



Relatório de Projeto

- Para uma fácil visualização dos produtos que o serviço mais tem vindo a consumir, foi criado o gráfico de barras agrupadas. Assim como um gráfico de barras empilhadas para os produtos onde houve maior gastos. Estas duas opções dão ao cliente a possibilidade de consultar pelo que mais lhe é relevante, se o valor gasto ou se a quantidade dispensada. Desta forma optei por adicionar o valor reembolsado e o efetivo para uma visualização mais detalhada e interessante ao cliente.



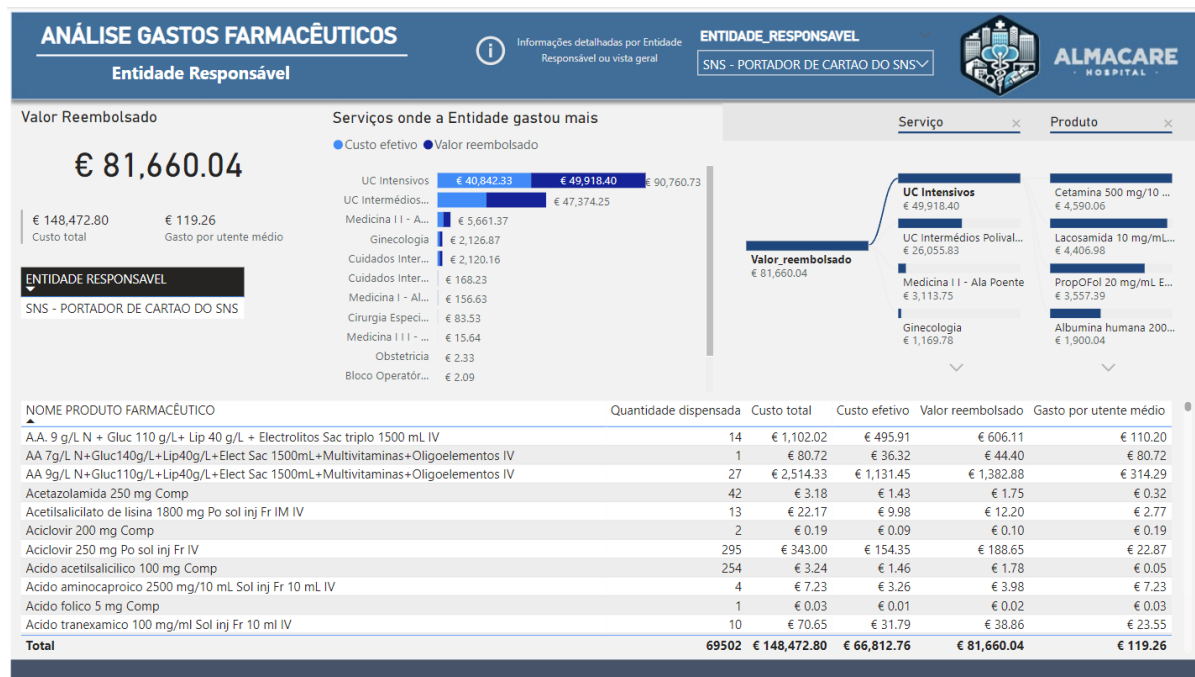
- A tabela disponível nesta página permite uma visão mais detalhada dos gastos e quantidades dispensadas, conforme pretendido pelo cliente.

NOME PRODUTO FARMACÊUTICO	Quantidade dispensada	Custo total	Custo efetivo	Valor reembolsado	Gasto por utente médio
Imunoglobulina humana normal 50 mg/mL Sol inj Fr 200 mL IV	6	€ 3,338.11	€ 1,502.15	€ 1,835.96	€ 3,338.11
Levossimendano 2.5 mg/mL Sol inj Fr 5 mL IV	2	€ 1,494.60	€ 672.57	€ 822.03	€ 747.30
Lacosamida 10 mg/mL Sol inj Fr 20 mL IV	284	€ 8,012.69	€ 3,605.71	€ 4,406.98	€ 500.79
Anidulafungina 100 mg Po conc sol inj Fr IV	28	€ 2,054.96	€ 924.73	€ 1,130.23	€ 410.99
LevoTIROXina sodica 0.2 mg/mL Sol inj	9	€ 549.94	€ 247.47	€ 302.47	€ 274.97
Cetamina 500 mg/10 mL Sol inj Fr 10 mL IM IV	272	€ 8,438.78	€ 3,769.49	€ 4,669.30	€ 255.72
AA 9g/L N+Gluc110g/L+Lip40g/L+Elect Sac 1500mL+Multivitaminas+Oligoelementos IV	16	€ 1,490.01	€ 670.50	€ 819.50	€ 248.33
Albumina humana 200 g/L Sol inj Fr 50 mL IV	163	€ 3,541.42	€ 1,576.28	€ 1,965.14	€ 236.09
Tiopental sodico 500 mg Po sol inj Fr IV	63	€ 192.10	€ 86.44	€ 105.65	€ 192.10
AL fentanilo 5 mg/10 mL Sol inj Fr 10 mL IV	660	€ 3,411.00	€ 1,534.95	€ 1,876.05	€ 162.43
Colistimetato de sodio 1000000 U.I. Po sol inj ou sol neb Fr IV Inalatoria	313	€ 847.41	€ 381.34	€ 466.08	€ 141.24
Daptomicina 500 mg Po sol inj Fr IV	2	€ 124.47	€ 56.01	€ 68.46	€ 124.47
A.A. 9 g/L N + Gluc 110 g/L + Lip 40 g/L + Electrolitos Sac tripro 1500 mL IV	11	€ 865.87	€ 389.64	€ 476.23	€ 123.70
CeTTAZIDIMA 2000.mg + Avibactam 500.mg Po conc sol inj Fr IV	1	€ 106.42	€ 47.89	€ 58.53	€ 106.42
Total	52195	€ 92,877.97	€ 41,495.39	€ 51,382.58	€ 327.04

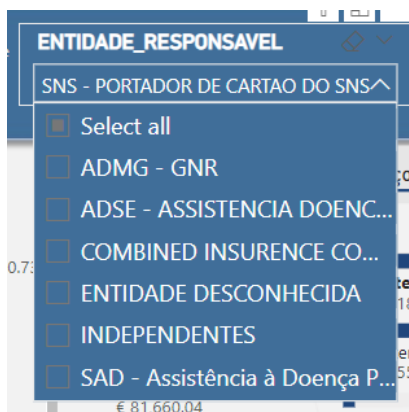
- Com este dashboard consegue-se dar resposta aos pontos:
 - Perceber que produtos farmacêuticos cada serviço deve ter de stock próprio.
 - Entender os gastos em produtos farmacêuticos de cada serviço, por utente.

Relatório de Projeto

- **Entidade Responsável** - permite ao cliente uma visualização mais detalhada por Entidade Responsável permitindo compreensão de valor reembolsado e custo efetivo por produtos e serviços.



- O cliente nesta página poderá filtrar por entidade responsável verificando dados mais detalhados para a mesma.

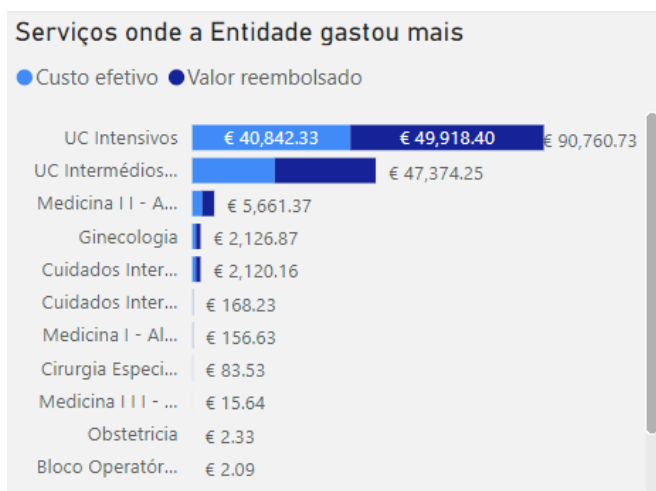


- Imediatamente consegue consultar o Valor Reembolsado assim como o Custo total dos utentes provenientes por esta entidade e o seu gasto médio por utente. A baixo dos mesmos está uma tabela que serve de placeholder para o logotipo da entidade ainda por implementar.

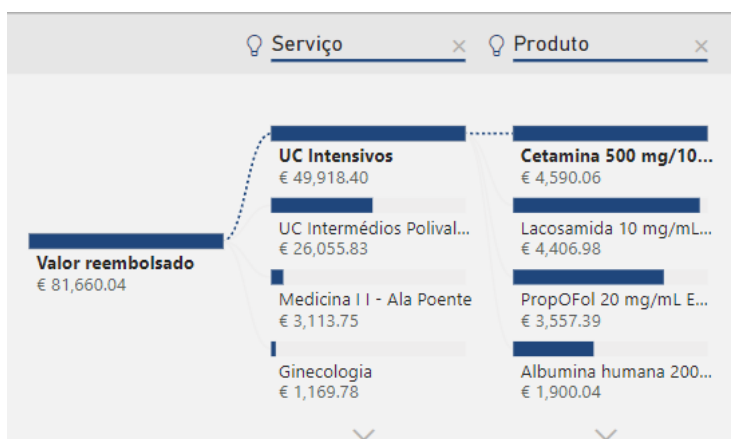
Relatório de Projeto



- Embora todos os requisitos relativos à entidade responsável estejam já respondidos com os dados anteriores, considerei oportuno adicionar mais visualizações para permitir uma análise mais detalhada e dinâmica.
- Nesse sentido criei um gráfico de barras empilhadas onde se terá a capacidade de entender quais os serviços onde a entidade está a ter uma ação maior.



- Assim como uma *decomposition tree* para uma análise mais minuciosa da área de interesse do analista permitindo melhor entender a distribuição do valor reembolsado. Isto significa que o cliente pode analisar em pormenor os produtos mais ou menos participados em cada serviço pela entidade.



Relatório de Projeto

- Por fim foi adicionada a tabela com todos os produtos e os seus dados relativos à entidade responsável, quantidade dispensada, custo total, custo efetivo, valor reembolsado e gasto por utente médio.

NOME PRODUTO FARMACÊUTICO	Quantidade dispensada	Custo total	Custo efetivo	Valor reembolsado	Gasto por utente médio
AA 9 g/L N + Gluc 110 g/L + Lip 40 g/L + Electrolitos Sac triplo 1500 mL IV	14	€ 1,102.02	€ 495.91	€ 606.11	€ 110.20
AA 7g/L N+Gluc140g/L+Lip40g/L+Elect Sac 1500mL+Multivitaminas+Oligoelementos IV	1	€ 80.72	€ 36.32	€ 44.40	€ 80.72
AA 9g/L N+Gluc110g/L+Lip40g/L+Elect Sac 1500mL+Multivitaminas+Oligoelementos IV	27	€ 2,514.33	€ 1,131.45	€ 1,382.88	€ 314.29
Acetazolamida 250 mg Comp	42	€ 3.18	€ 1.43	€ 1.75	€ 0.32
Acetilsalicilato de lisina 1800 mg Po sol inj Fr IM IV	13	€ 22.17	€ 9.98	€ 12.20	€ 2.77
Aciclovir 200 mg Comp	2	€ 0.19	€ 0.09	€ 0.10	€ 0.19
Aciclovir 250 mg Po sol inj Fr IV	295	€ 343.00	€ 154.35	€ 188.65	€ 22.87
Acido acetilsalicílico 100 mg Comp	254	€ 3.24	€ 1.46	€ 1.78	€ 0.05
Acido aminocaproico 2500 mg/10 mL Sol inj Fr 10 mL IV	4	€ 7.23	€ 3.26	€ 3.98	€ 7.23
Acido fólico 5 mg Comp	1	€ 0.03	€ 0.01	€ 0.02	€ 0.03
Acido tranexámico 100 mg/ml Sol inj Fr 10 mL IV	10	€ 70.65	€ 31.79	€ 38.86	€ 23.55
Total	69502	€ 148,472.80	€ 66,812.76	€ 81,660.04	€ 119.26

- Com este dashboard consegue-se dar resposta ao ponto:
 - Entender quais as entidades que mais participam os produtos farmacêuticos.

Conclusões da análise do Dashboard

Através da análise do dashboard consegue-se concluir que o cliente é capaz de:

- Analisar gastos do ano anterior para planear stocks para os próximos anos e custos associados.
 - No dashboard geral o cliente consegue analisar os gastos que fez durante o ano, em produtos farmacêuticos assim como verificar as quantidades dispensadas de cada produto.
 - Nesse sentido verifica-se que Cloreto de sódio 9 mg/mL Sol inj Fr 100 mL IV, Fentanilo 0.05 mg/mL Sol inj Fr 5 mL IM IV e Cloreto de sódio 9 mg/mL Sol inj Fr 500 mL IV são os produtos farmacêuticos mais dispensados pelo hospital.
 - Já a Imunoglobulina humana normal 50 mg/mL Sol inj Fr 200 mL IV, a Cetamina 500 mg/10 mL Sol inj Fr 10 mL IM IV e a Lacosamida 10 mg/mL Sol inj Fr 20 mL IV são os produtos farmacêuticos que representaram maior despesa.
- Perceber que produtos farmacêuticos cada serviço deve ter de stock próprio.
 - No dashboard Serviço o cliente pode verificar serviço a serviço quais os produtos mais utilizados e assim decidir que produtos deverá manter em stock próprio desse serviço
 - Usando como exemplo o serviço de UC Intensivos verificasse uma maior utilização de Cloreto de sódio 9 mg/mL Sol inj Fr 100 mL IV, Fentanilo 0.05 mg/mL Sol inj Fr 5 mL IM IV e PropOFol 20 mg/mL Emul inj Fr 50 mL IV como tal deveriam ter stock próprio.

Relatório de Projeto

- Entender os gastos em produtos farmacêuticos de cada serviço, por utente.
 - Quer no dashboard geral quer no dashboard serviço esta informação é apresentada. No dashboard geral o gráfico disponível permite uma mais fácil comparação entre serviços uma vez que estão todos os serviços representados com o seu respetivo gasto médio por utente.
 - Nesse sentido identificamos que os três serviços com maior gasto por utente são Medicina II – Ala Poente, Ginecologia e UC Intensivos.
- Entender quais as entidades que mais participam os produtos farmacêuticos.
 - Esta informação está disponível no dashboard geral onde é facilmente identificável através de um gráfico de donut as principais entidades, no entanto no dashboard Entidade Responsável é possível uma visualização mais pormenorizada dos valores reembolsados pelas entidades.
 - Através desta análise compreende-se que a Entidade Responsável com maior predomínio de participação ao hospital é o SNS – Portador cartão SNS, seguindo-se da ADSE, não tendo as restantes um valor expressivo.

Considerações finais – Visualização de Dados:

A partir dos dados obtidos, que representam apenas uma fração do total, foi possível avançar para uma base sólida de análises precisas, visualização adequada e dinâmica.

Ao concluir a análise do dashboard desenvolvido no Power BI para o hospital, considero que a ferramenta proporcionará ao cliente, gestor do departamento de gestão de farmácia hospitalar, uma visão detalhada e abrangente dos dados, permitindo-lhe tomar decisões informadas e estratégicas.

Através da visualização dos principais gastos, serviços e entidades responsáveis, o cliente pode facilmente identificar padrões, tendências e áreas de foco para otimização de recursos e melhorias no atendimento ao utente. O filtro por período temporal possibilita uma análise histórica e prospetiva, enquanto os gráficos e tabelas fornecem uma representação clara e concisa dos dados.

Além disso, as métricas calculadas, como custo efetivo, valor reembolsado e gasto por utente médio, oferecem insights valiosos sobre o desempenho financeiro e operacional do hospital, contribuindo para uma gestão mais eficiente.

Relatório de Projeto

Em suma, a utilização do Power BI como ferramenta de visualização de dados revelou-se fundamental para atender aos requisitos do cliente e proporcionar uma análise detalhada e orientada para a tomada de decisões no contexto específico da gestão de farmácia hospitalar.

É também importante destacar que esta UC de Análise e Visualização de Dados proporcionou as ferramentas necessárias para a realização deste projeto, bem como para futuras aplicações em diferentes contextos. As técnicas e metodologias exploradas não apenas atenderam aos requisitos específicos deste projeto, mas também criaram uma base sólida para a aplicação em outras áreas e situações.

É fundamental reconhecer que, apesar dos avanços alcançados neste projeto, há ainda um vasto espaço para expandir os conhecimentos e explorar novas práticas e tecnologias na área de análise e visualização de dados.

Portanto, este projeto não apenas demonstra a aplicação prática dos conceitos aprendidos na disciplina, mas também serve como um ponto de partida para uma jornada contínua de aprendizagem e aprimoramento na área de análise e visualização de dados, contribuindo para uma gestão mais eficiente e orientada por dados não apenas no contexto hospitalar, mas também em outros setores e cenários.