

DEPARTAMENTO DE INFORMÁTICA



**ESCOLA
SUPERIOR
DE TECNOLOGIA
E GESTÃO**

Tópicos avançados de armazenamento de dados

Medical Center Management System

Trabalho realizador por:

João Silva (8220024),

Luís Silva (8220025)

Estudantes do mestrado em Engenharia Informática

Supervisores/Orientadores

Responsável pelo projeto: Bruno Oliveira

ESCOLA SUPERIOR DE TECNOLOGIA E GESTÃO

Resumo

De forma a se colocar em prática todos os conhecimento teóricos adquiridos na Unidade Curricular de Tópicos Avançados de Armazenamento de Dados foi-nos proposto o seguinte trabalho. Este trabalho consistiu na realização de um caso de estudo de um domínio de uma empresa, suportado por um conjunto de requisitos de análise para a implementação de um sistema de *Data Warehousing*.

Tendo em conta esta proposta, o nosso grupo procedeu à escolha do tema, sendo este baseado na elaboração de um sistema de Gestão de um Posto de Centro médico. Esta escolha teve por base o interesse do grupo pelo tema, bem como a fácil distinção de processos existentes, isto é, ao nível de Processo de Marcação de Consultas, bem como o processo de Execução das Consultas.

A nível de software, procedeu-se à utilização do *Microsoft Sql Server*, onde se conseguiram implementar todos os processos de *Data Definition Language* (DDL), assim com os processos de *Data Manipulation Language* (DML), bem como a criação de todos os *procedures* necessários para a Base de Dados Operacional. Utilizou-se também o *SQL Server Integration Services* de forma a elaborar e a proceder a todos os processos de criação das Tabelas de Facto e dimensões do *Data Warehouse*, bem como os processos de *Extract-Transform-Load* (ETL).

Posteriormente, de forma a estudar todos os processos apresentados, bem como a tentar perceber todos os *benchmarks* que podem afetar a evolução e o entendimento do problema, procedeu-se à elaboração de *Dashboards*, com os dados provenientes do *Data Warehouse*, com o intuito de se responder às diversas questões de negócio. Esta implementação das *dashboards*, bem como a sua visualização, foi através da utilização do *PowerBi*.

Índice

| | | |
|----------|---|-----------|
| 1 | Introdução | 1 |
| 1.1 | Contextualização | 1 |
| 1.2 | Motivação | 1 |
| 1.3 | Processos | 2 |
| 1.3.1 | Processos principais | 2 |
| 1.4 | Objetivos | 4 |
| 1.5 | Entidade Relacionamento | 6 |
| 1.6 | Povoamento das Tabelas | 7 |
| 2 | Planeamento do DW | 7 |
| 2.1 | Modelação Dimensional | 7 |
| 2.1.1 | Métricas e Requisitos | 8 |
| 2.1.2 | Questões | 8 |
| 2.1.3 | Processos de Negócio | 8 |
| 2.2 | Idealização do ETL | 13 |
| 2.2.1 | <i>Source-to-target map</i> | 13 |
| 2.2.2 | BPMN referente ao ETL | 13 |
| 3 | Criação do <i>Data Warehouse</i> | 20 |
| 4 | <i>Dashboards de Business Intelligence</i> | 22 |
| 5 | Conclusão | 28 |
| | Referências | 29 |

Lista de Figuras

| | | |
|----|---|----|
| 1 | Processo de atendimento dos pacientes. | 2 |
| 2 | Processo de receção do paciente. | 3 |
| 3 | Processo da consulta do paciente. | 3 |
| 4 | Processo de pós consulta do paciente. | 4 |
| 5 | Diagrama Entidade Relacionamento do Domínio. | 7 |
| 6 | Caracterização de Dimensões do Processo de Marcação de consulta. . . . | 11 |
| 7 | Caracterização de Dimensões do Processo de consulta. | 11 |
| 8 | Esquema em estrela do Processo de Marcação de consulta e consulta. . . | 12 |
| 9 | <i>Source-to-target map book appointment.</i> | 13 |
| 10 | <i>Source-to-target map appointment.</i> | 13 |
| 11 | BPMN referente ao ETL de nível abstrato mais alto. | 14 |
| 12 | BPMN processo inicial. | 15 |
| 13 | BPMN para carregar a Dimensão Counter. | 15 |
| 14 | BPMN para carregar a Dimensão Patient. | 16 |
| 15 | BPMN para carregar a Dimensão Doctor. | 16 |
| 16 | BPMN para carregar a Dimensão Medical Center. | 17 |
| 17 | BPMN para carregar a Dimensão Payment. | 17 |
| 18 | BPMN para carregar a Dimensão Prescription. | 18 |
| 19 | BPMN para carregar a tabela de factos bookappointment. | 18 |
| 20 | BPMN para carregar a tabela de factos appointment. | 19 |
| 21 | <i>Package</i> para a criação do <i>Data Warehouse</i> e dimensões de dia e Hora. . . | 20 |
| 22 | <i>Package</i> para o povoamento do <i>Data Warehouse</i> | 21 |
| 23 | <i>PowerBI Dashboard</i> de relação paciente e doutor. | 22 |
| 24 | <i>PowerBI Dashboard</i> para análise paciente. | 23 |
| 25 | <i>PowerBI Dashboard</i> para análise de <i>Ticket</i> | 24 |
| 26 | <i>PowerBI Dashboard</i> para análise de consultas realizadas. | 25 |
| 27 | <i>PowerBI Dashboard</i> para análise de consultas realizadas. | 26 |
| 28 | <i>PowerBI Dashboard</i> para análise de prescrição de medicamentos. | 27 |

1 Introdução

Este capítulo tem como objetivo apresentar o enquadramento e a motivação do tema do trabalho, os objetivos e resultados esperados no seu desenvolvimento e a abordagem de investigação selecionada.

1.1 Contextualização

No âmbito da Unidade Curricular de Técnicas Avançadas de Análise de Dados (TAAD), do mestrado em Engenharia Informática, foi-nos proposta o planeamento e implementação de um Sistema de Bases de Dados Multidimensionais, ou seja, a implementação de um *Data Warehouse* (DW) e que, partindo deste, fosse possível a análise visual do mesmo, através de *Dashboards* recorrendo ao *PowerBi*.

Seguindo os conteúdos lecionados nas aulas teóricas, este projeto foi dividido em várias fases:

- Criação da Base de dados.
- Povoamento da Base de dados.
- Análise dos Dados.
- Planeamento do *Data Warehouse*.
- ETL e Implementação do *Data Warehouse*.
- Criação das *Dashboards*.
- Análise dos Resultados.

1.2 Motivação

De forma bastante resumida, um sistema de Gestão de um Centro de saúde é um sistema computacional organizado, que tem por base automatizar as operações e processos que se realizam diariamente num centro hospitalar.

O Centro de Saúde é a unidade básica do Serviço Nacional de Saúde (SNS) que realiza atendimento e prestação de cuidados de saúde à população. O Centro de Saúde é a primeira porta a que deve bater caso precise de cuidados médicos.

Estas operações e processos podem ser de vários níveis e tipos, tais como a interação com o paciente à sua entrada e posterior orientação, a verificação se este paciente tem ou não consulta efetuada, sendo que em caso negativo poderá fazer marcação. Também permitirá o processo de consulta, isto é, a interação do paciente com o doutor, bem como o diagnóstico e prescrição de medicamentos, caso necessário. Ainda no fluxo de consulta e interação do paciente com o doutor, o sistema deverá permitir a realização de pagamento das consultas, caso necessário, bem como a atualização do valor da consulta, caso o paciente apresente seguro de saúde.

Também deverá incluir o departamento de Recursos Humanos, que permitirá a realização de todos os pagamentos mensais aos funcionários do Centro Hospitalar.

1.3 Processos

Dentro do domínio de Gestão de um centro de saúde podem-se considerar diversos processos principais. Estes processos têm como objetivo automatizar e dinamizar o bom e regular funcionamento da organização.

1.3.1 Processos principais

Todos os processos descritos abaixo, foi utilizado uma notação da metodologia de gestão de processos de negócio, o Business Process Model and Notation (BPMN). O processo de atendimento dos pacientes resume o fluxo geral do negócio e está definido no diagrama da figura 1.

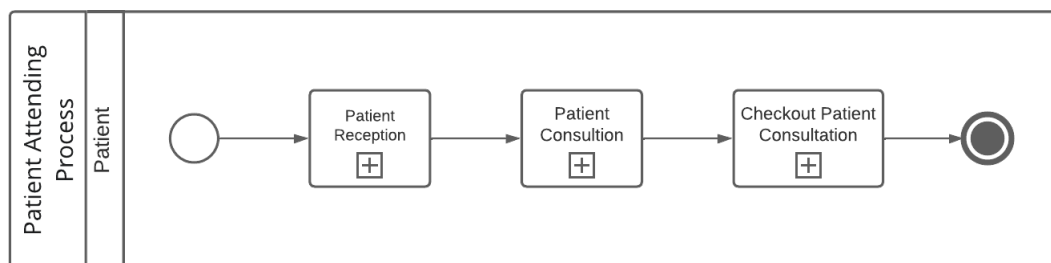


Figura 1: Processo de atendimento dos pacientes.

Este processo geral é composto por 3 grandes processos:

- Processo de receção do paciente.

- Processo da consulta do paciente.
- Processo de pós consulta do paciente.

Em relação ao processo de receção do paciente, este faz parte do negócio onde o paciente entra na receção do centro de saúde e está definido no diagrama da figura 2. De forma a simplificar a leitura foi fragmentado em 3 sub-processos: paciente espera pela chamada da sua senha, a receção chama senhas, e por fim, a receção confirma a entrada do paciente.

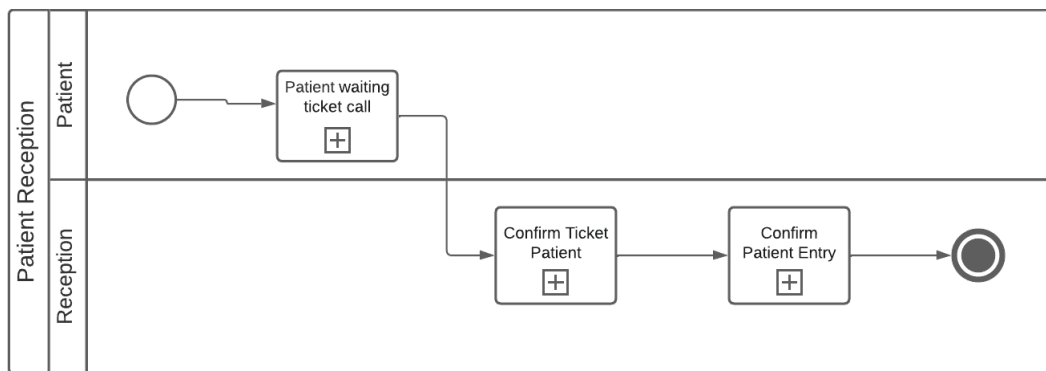


Figura 2: Processo de receção do paciente.

Relativamente ao processo da consulta do paciente é alusivo à parte do negócio onde o paciente entra na consulta médica e está definido no diagrama da figura 3. Seguindo a lógica do anterior este processo foi descrito com 2 sub-processos: o médico confirma se o paciente aparece na consulta, e, a consulta.

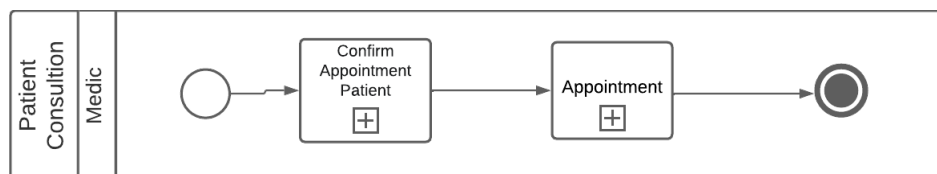


Figura 3: Processo da consulta do paciente.

Por fim o processo de pós consulta do paciente faz referência à parte do negócio onde o paciente dirige-se novamente à receção para marcar nova consulta, ou não 4.

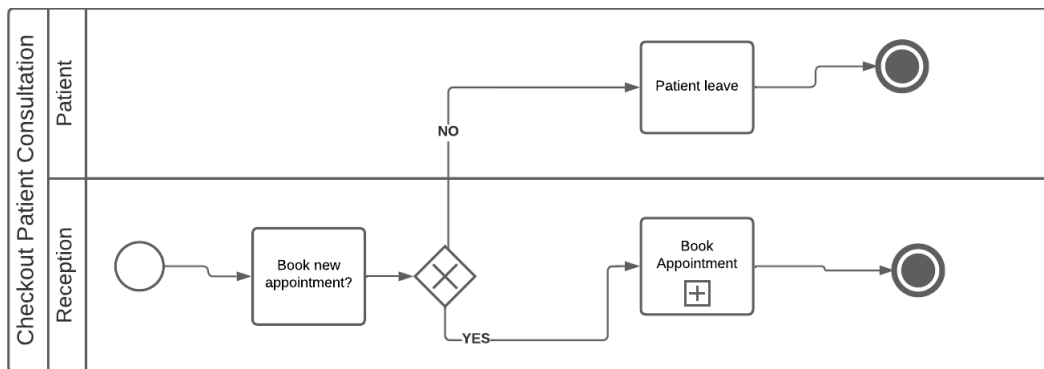


Figura 4: Processo de pós consulta do paciente.

1.4 Objetivos

Tendo em conta o contexto do domínio utilizado, todos os departamento representados no domínio e a diversidade de pacientes a frequentar o domínio é possível apresentar diversos objetivos. Estes objetivos podem estar ligados ao entendimento do grupo mais recorrente a nível de pacientes, seja este por faixa etária, por sexo ou profissão. Ainda neste âmbito, tentar aferir se existe uma doença que faça com que os pacientes recorram mais frequentemente ao Centro Hospitalar, ou mesmo, a quantidade de pacientes que não vai a consultas frequentemente.

No que concerne à lotação do Hospital, estudar o efeito temporal na mesma, isto é, se se verifica que em meses mais frios e chuvosos a frequência aumenta. Verificar também se existem horas ou dias da semana com maior efetividade de pacientes. Ainda neste Âmbito, tentar verificar se existirá uma influência significativa de acordo com os doutores em trabalho laboral nos dias em questão.

- Estações do ano com maior afluências de consultas.
- Dia e hora mais recorrente de consulta.
- Doutor mais solicitado em consultas

De acordo com o diagrama de Processos relativo ao processo de marcação de consultas, definiu-se como objetivo a análise e posterior intervenção nos seguintes pontos.

- Redução do tempo de espera na fila.
- Redução da quantidade de pessoas que vai embora.

- Cálculo dos dias em que as consultas são mais afluentes.

No que concerne ao Diagrama de Processos relativo ao processo da consulta, definiu-se como objetivo a análise e posterior intervenção nos seguintes pontos.

- Otimização do Tempo médio de consulta.
- Redução de tempos de atraso nas consultas.
- Faixa etária mais recorrente.
- Sexualidade mais recorrente.
- Profissão mais recorrente.
- Doença mais recorrente.
- Diagnóstico e sintomas mais recorrentes.
- Gestão de Stock de fármacos.
- Otimização do número de médicos a trabalhar por dia.

No processo de sistema ligado ao Pagamento das consultas deverá ser possível fazer um estudo à quantidade de pessoas que apresenta isenção de pagamento, se existe uma boa percentagem populacional com seguro de saúde, e qual o mais recorrente. Poderá ainda verificar-se qual o método de pagamento mais utilizado.

- Quantidade de pessoas com isenção de pagamento.
- Percentagem populacional com seguro de saúde.
- Seguro de saúde mais recorrente.
- Método de Pagamento mais utilizado.

Por último, uma análise ao staff do Hospital poderá ser realizada, aferindo sobre o contentamento ou descontentamento dos mesmos, tanto a nível profissional como de faixa etária, as razões para o descontentamento, bem como se a sexualidade tem influência.

- Contentamento ou Descontentamento dos funcionários.
- Influência da profissão no mesmo.
- Razões de descontentamento mais recorrente.
- Influência da sexualidade no bem-estar profissional.

1.5 Entidade Relacionamento

Os centros de saúde têm vários departamentos e cada um deles tende a manter seu próprio sistema individual para manutenção de registros. De forma a aumentar a coerência entre os departamentos propomos uma estrutura de dados centralizada. O Diagrama de Entidade Relacionamento do domínio proposto está presente na Figura 5.

Através da Análise do Diagrama, pode-se retirar algumas informações que permitem entender o modelo:

- A Base de Dados possuiu 14 Tabelas, que permitem resumir a Existência de Pacientes, Rececionistas, Médicos, Tickets na Recepção, Pagamentos mensais ao Médicos e Rececionista, Medicamentos, Consultas, Prescrições, Pagamentos de consultas, Diversos Sintomas e Seguros de Saúde.
- A Base de dados foca-se principalmente em Dois processos distintos: Marcação de Consultas e Execução de Consultas.
- Nestes Processos temos 3 Principais intervenientes: Os Pacientes, os Rececionistas e os Médicos.

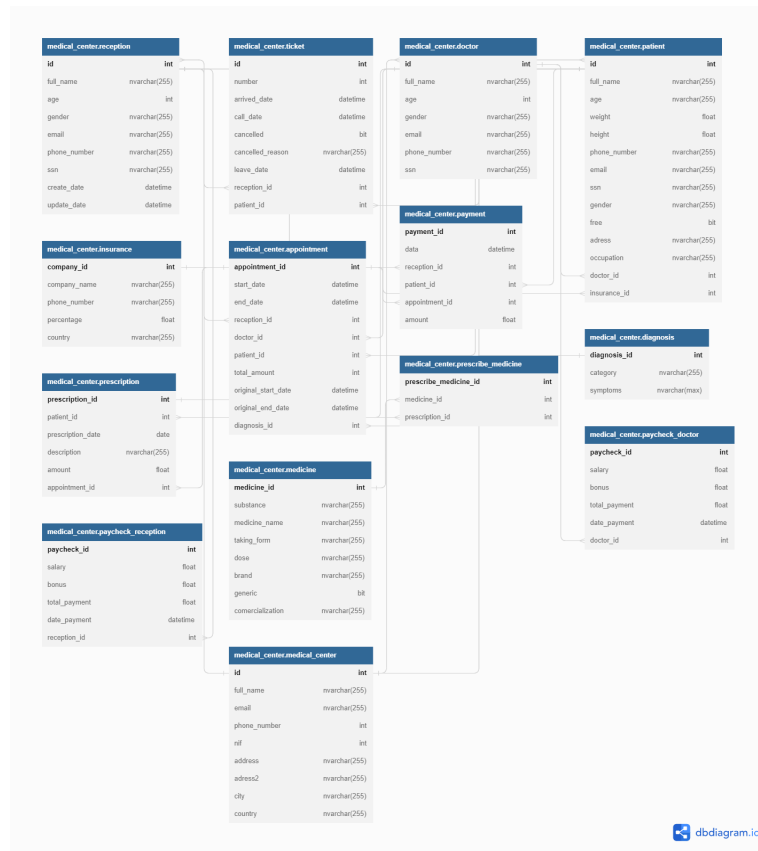


Figura 5: Diagrama Entidade Relacionamento do Domínio.

1.6 Povoamento das Tabelas

De forma a povoar as tabelas da Base de Dados operacional utilizou-se a plataforma *Mockaroo*. Esta plataforma permite codificar os dados que se pretende gerar e posteriormente realizar a sua exportação em formato *CSV* ou *SQL*.

Para as tabelas com informação de medicamentos e diagnósticos, encontrou-se documentos em formato *CSV* online. Posteriormente, foi feita a transformação do *CSV* para formato *SQL*, e realizada a sua importação.

2 Planeamento do DW

2.1 Modelação Dimensional

De forma a se proceder ao Processo de Modelação Dimensional, primeiramente estabelecem-se as Métricas e Requisitos que se pretende obter.

2.1.1 Métricas e Requisitos

- Melhorar a qualidade do serviço de Marcação de Consulta e do processo de Consulta;
- Identificar sazonalidades/tendências/periodicidade de maior afluência;
- Identificar grupo de pessoas que marcam mais consultas;
- Qual o dia da semana em que se registam mais consultas?
- Percentagem de consultas com pessoas isentas por dia?
- Atualização de Stock com base nos fármacos mais prescritos;
- Quantificação de Staff diário, com base em indicadores de Afluência;

2.1.2 Questões

- Número de consultas efetuadas por médico e por mês;
- Média de tentativas de marcação de consultas falhadas por mês;
- Média de consultas por mês;
- Percentagem de consultas por faixa etária por mês;
- Média de cada tipo de fármacos prescritos por mês;
- Qual o dia da semana em que se registam mais consultas?
- Percentagem de consultas com pessoas isentas por dia?
- Qual a hora em que se registam mais consultas por semana e mês?

2.1.3 Processos de Negócio

A diferenciação dos diferentes Processos de Negócio é relevante no seguinte projeto, dado que facilitam a elaboração dos diferentes *Data Marts*, que se tratam de pequenos *data warehouse* que contêm informação relativa a pelo menos um processo de negócio da empresa.

A nível de processos de Negócio, escolheram-se dois dos principais processos de negócio para análise:

- Processo de Marcação de Consulta
- Processo de Consulta

Estes processos não ocorrem simultaneamente, pelo que serão necessários dois Esquemas em Estrela para os descrever. De forma a se chegar ao Esquema em Estrela final, a técnica que se vai utilizar será o Método dos 4 passos, sendo que esta técnica será utilizada individualmente para cada um dos processos[1].

O Método dos passos segue-se da Seguinte forma para o processo de Marcação de Consulta.

1. Escolha do processo: Marcação de Consulta.
2. Identificação do Grão
 - (a) Cada linha de marcação corresponde a uma tentativa de marcação de consulta por cliente e Rececionista, num determinado centro de saúde, num determinado dia e hora
3. Identificação das Dimensões
 - (a) Quem está envolvido?
 - i. Cliente e Rececionista.
 - (b) O que foi realizado?
 - i. Marcação de Consulta.
 - (c) Quando aconteceu?
 - i. Determinado dia e hora.
 - (d) Onde foi realizado?
 - i. Determinado Centro de saúde.
4. Identificação dos Factos/medidas
 - (a) Tempo de espera (Aditivo)

O Método dos passos segue-se da Seguinte forma para o processo de Consulta.

1. Escolha do processo: Consulta.
2. Identificação do Grão
 - (a) Cada linha de consulta corresponde a uma consulta por cliente e doutor, num determinado centro de saúde, num determinado dia e hora
3. Identificação das Dimensões
 - (a) Quem está envolvido?
 - i. Cliente e Doutor.
 - (b) O que foi realizado?
 - i. Consulta.
 - (c) Quando aconteceu?
 - i. Determinado dia e hora.
 - (d) Onde foi realizado?
 - i. Determinado Centro de saúde.
4. Identificação dos Factos/medidas
 - (a) Tempo de atraso em minutos(Aditivo)
 - (b) Valor total pago(Aditivo)
 - (c) Valor total (Aditivo)
 - (d) Valor do Seguro participado(Aditivo)
 - (e) Número de fármacos prescritos (Aditiva)

Uma vez idealizados os esquemas, procedeu-se à caracterização das dimensões, tal como se pode verificar nas seguintes Figuras.

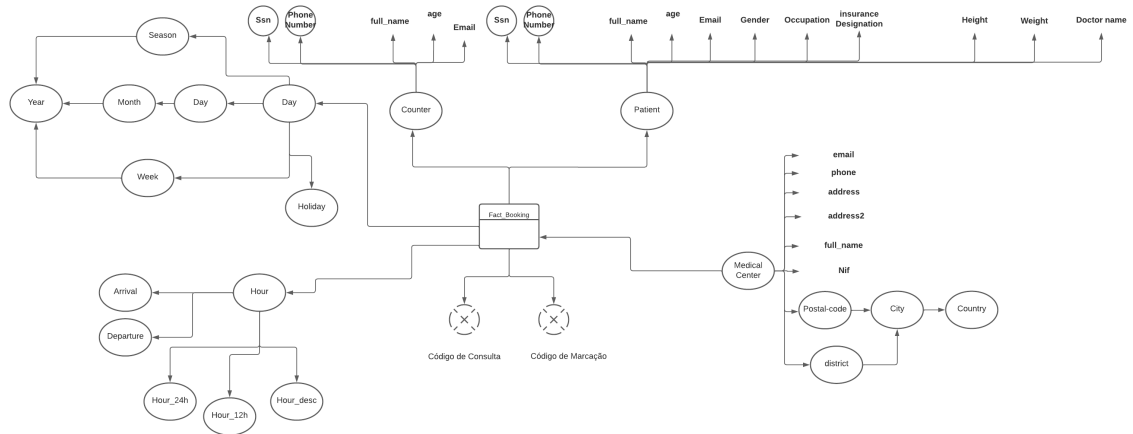


Figura 6: Caracterização de Dimensões do Processo de Marcação de consulta.

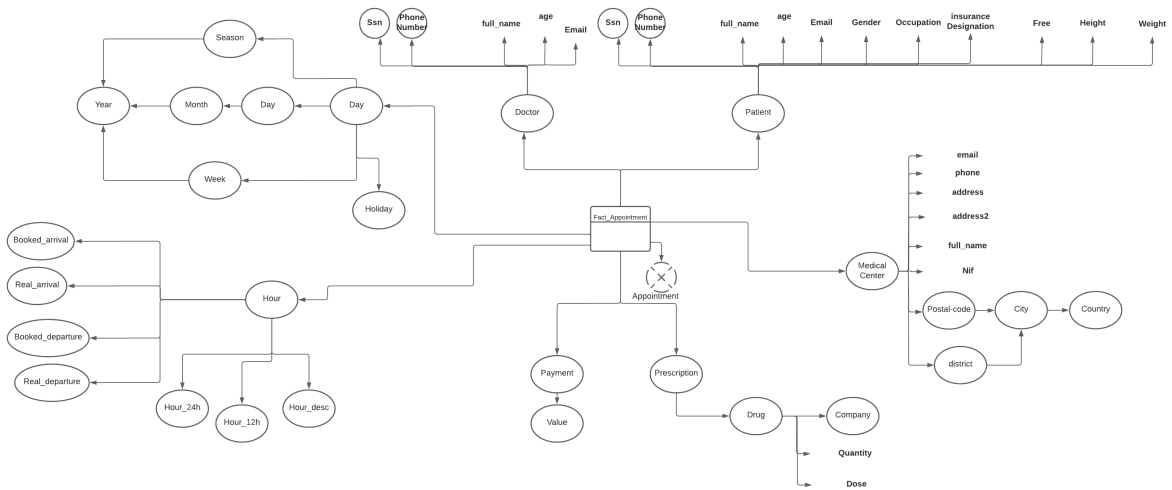


Figura 7: Caracterização de Dimensões do Processo de consulta.

Procedeu-se então à elaboração dos Esquemas em Estrela para os dois processos previamente descritos. Nesta fase, escolheu-se a Modelação dos Esquemas em Estrela em detrimento dos Esquemas em *Snow-Flake*, uma vez que estes últimos tornam as consultas na Base de dados mais complexas, podendo assim causa uma diminuição da rapidez de consultas, ou seja, uma diminuição do desempenho.

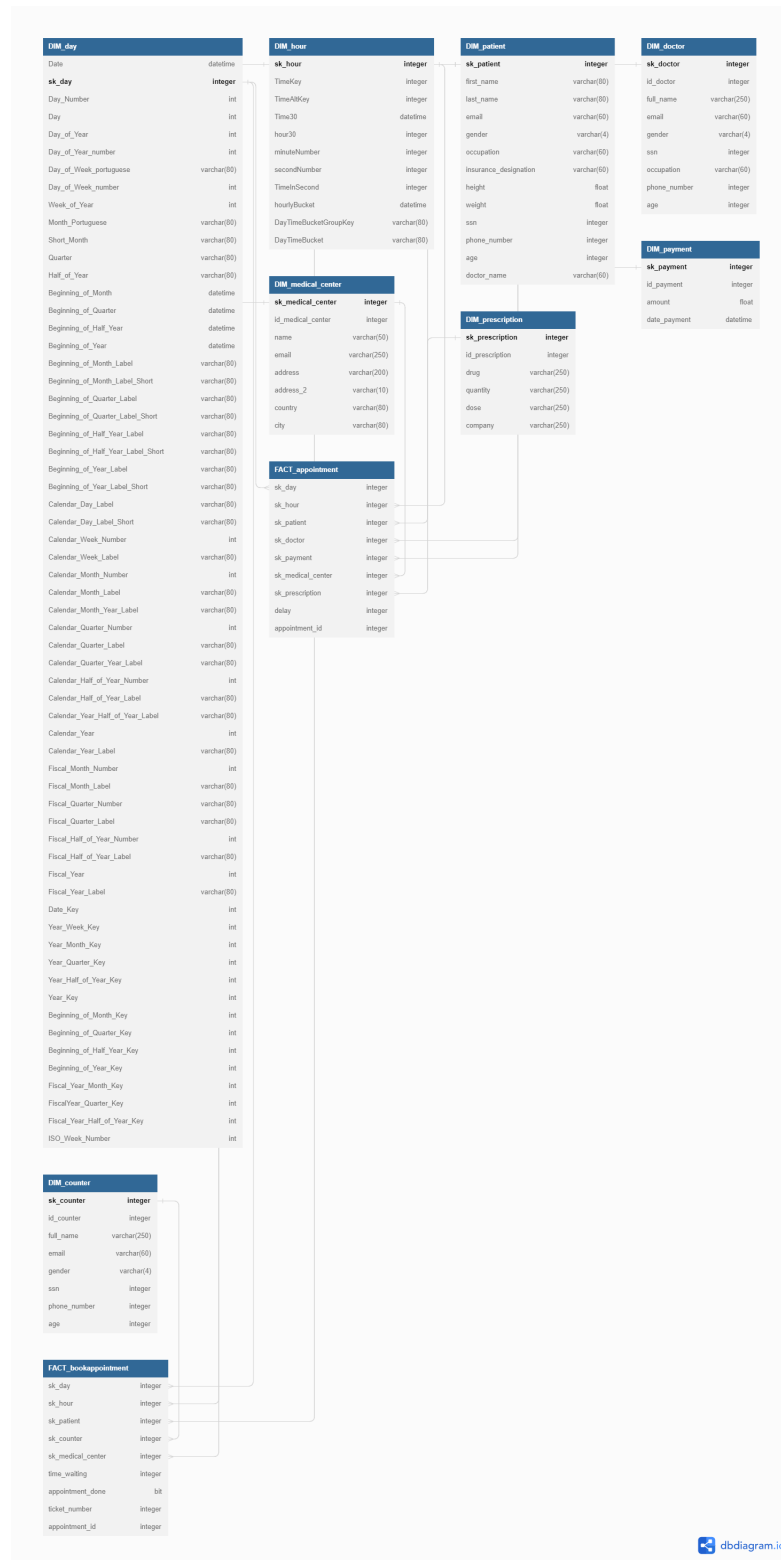


Figura 8: Esquema em estrela do Processo de Marcação de consulta e consulta.

2.2 Idealização do ETL

A implementação do ETL pode ser uma tarefa complicada, se não for idealizado e arquitetado previamente. Com isto em mente, utilizando os requisitos identificados anteriormente nos *Data Marts*, foi elaborada a arquitetura para a implementação do ETL para integrar o *data warehousing* no Processo de Negócio.

2.2.1 Source-to-target map

Procedeu-se à elaboração da documentação do *source-to-target map* para os dois Processos de Negócio descritos, como se pode confirmar nas seguintes figuras.

| | | | | | | | | | | | | | | | | | |
|--------------|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|
| Table Name | tf_bookappointment | | | | | | | | | | | | | | | | |
| Table Type | Fact | | | | | | | | | | | | | | | | |
| View Name | book appointment | | | | | | | | | | | | | | | | |
| Display Name | book appointment | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| Description | Cada linha de marcação corresponde a uma tentativa de marcação de consulta por cliente e Rececionista, num determinado centro de saúde, num determinado dia e hora | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | |
| | </ | | | | | | | | | | | | | | | | |

Figura 9: *Source-to-target map book appointment.*

[illegible]

Figura 10: *Source-to-target map appointment.*

2.2.2 BPMN referente ao ETL

Para a representação do ETL, foi realizado um BPMN para a modelação conceptual dos principais processos de povoamento. Para isto, foi inicialmente desenhado um BPMN de nível abstrato mais alto, *Process level*, como se pode confirmar na figura 11.

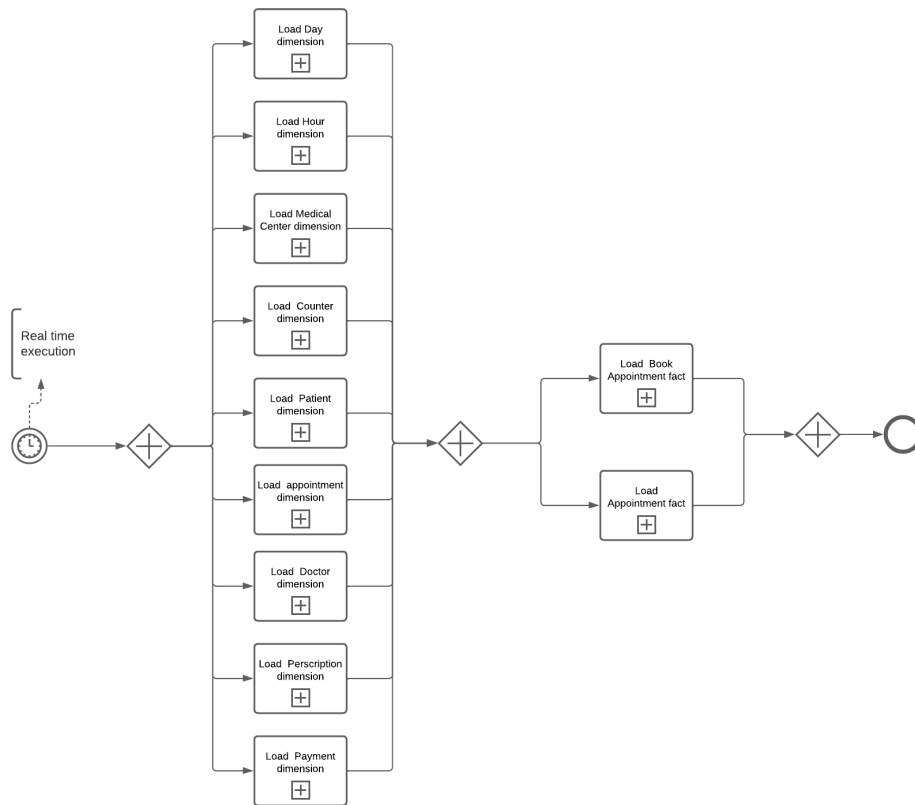


Figura 11: BPMN referente ao ETL de nível abstrato mais alto.

À medida que o desenvolvimento deste avançava, o BPMN foi melhorado e para isto usamos o nível seguinte para detalhar o processo, *Pattern level*.

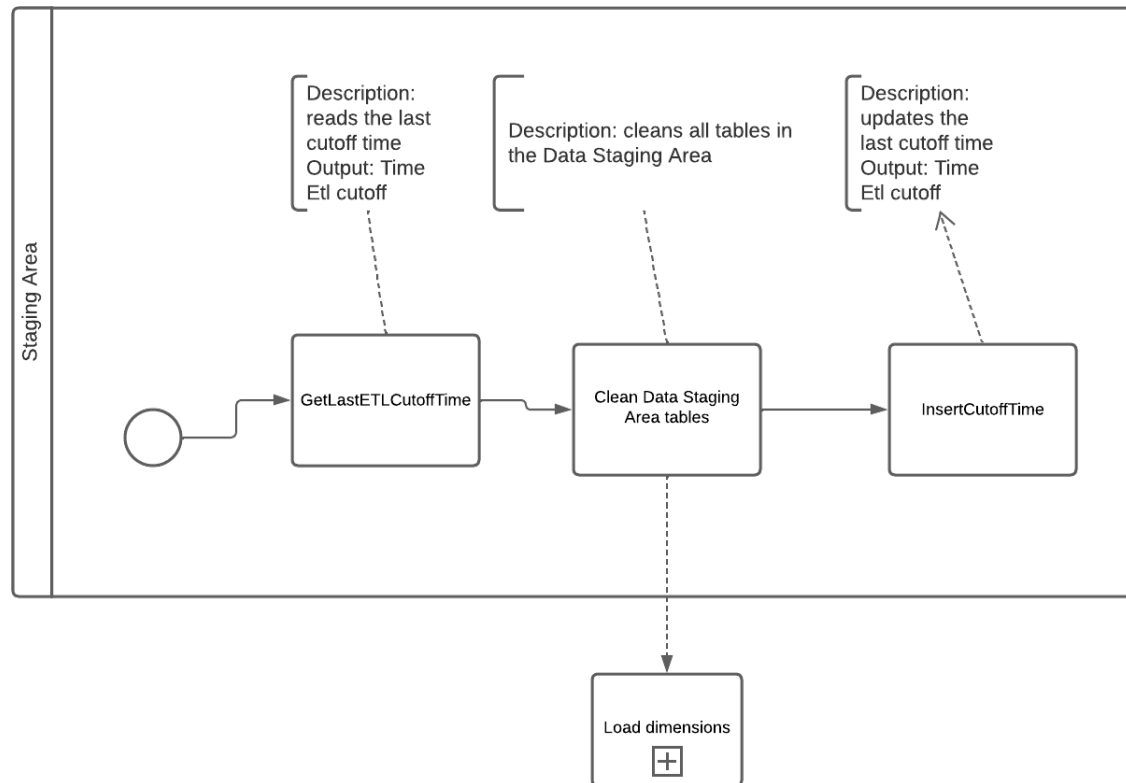


Figura 12: BPMN processo inicial.

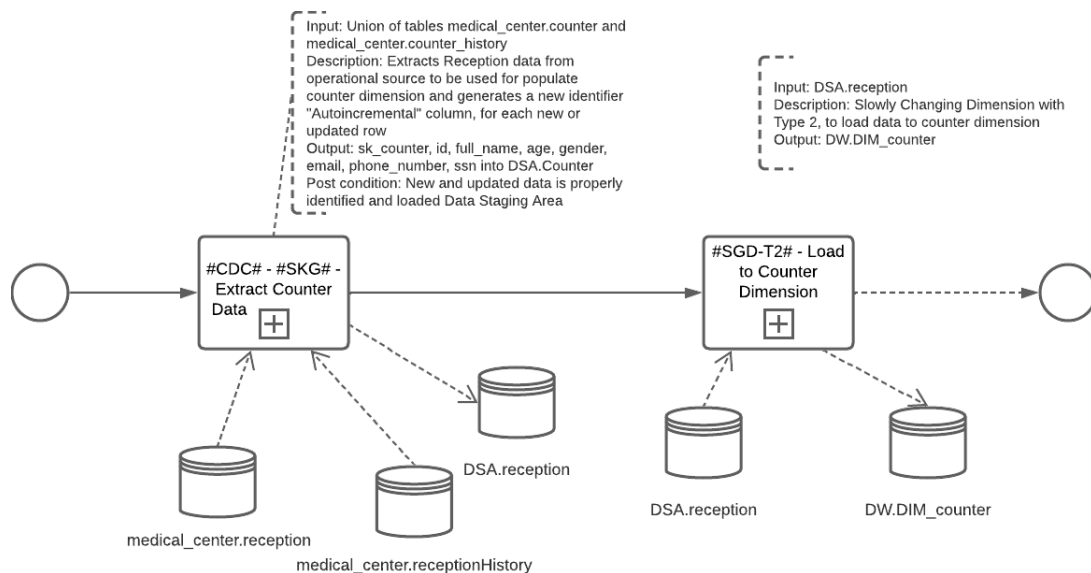


Figura 13: BPMN para carregar a Dimensão Counter.

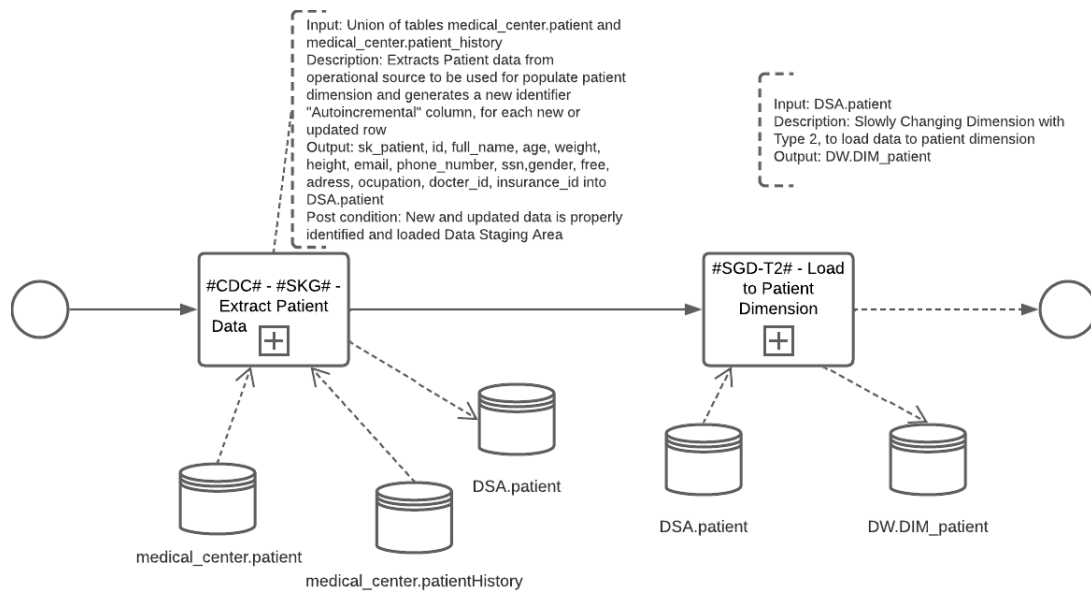


Figura 14: BPMN para carregar a Dimensão Patient.

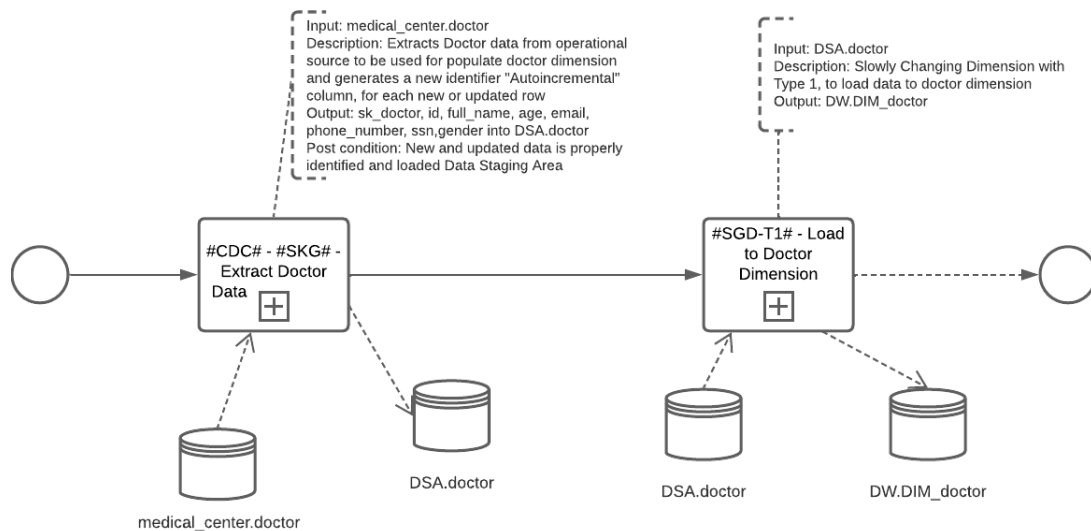


Figura 15: BPMN para carregar a Dimensão Doctor.

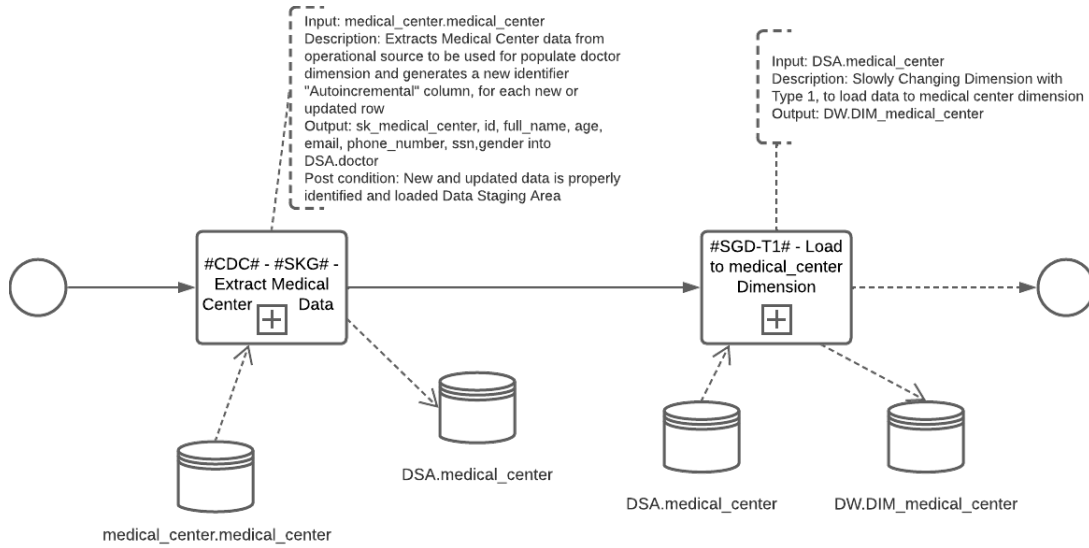


Figura 16: BPMN para carregar a Dimensão Medical Center.

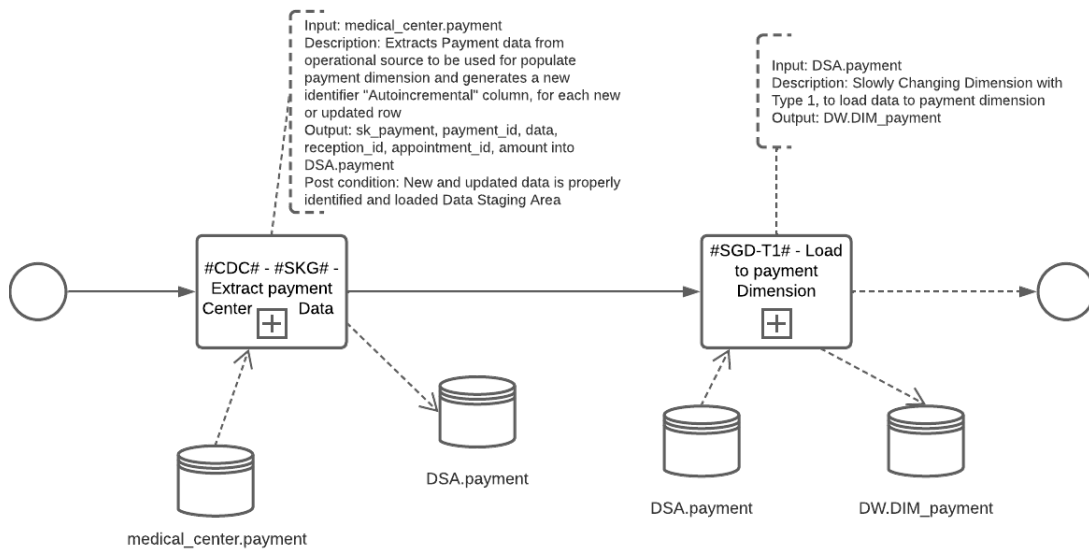


Figura 17: BPMN para carregar a Dimensão Payment.

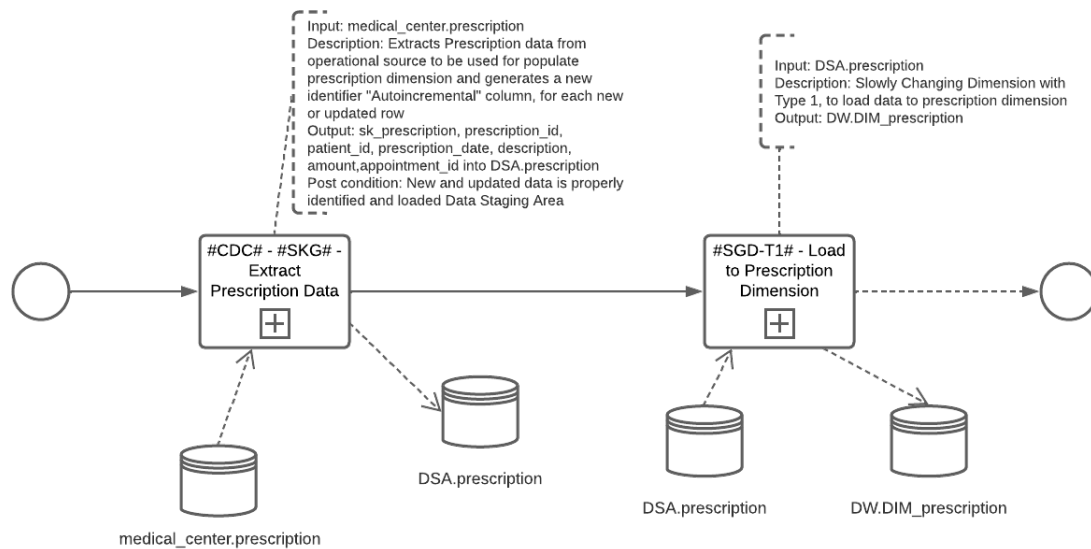


Figura 18: BPMN para carregar a Dimensão Prescription.

Depois de todas as dimensões estarem criadas e carregadas com dados, passamos para o processo final do ETL, que será carregar a tabela de factos com dados das dimensões.

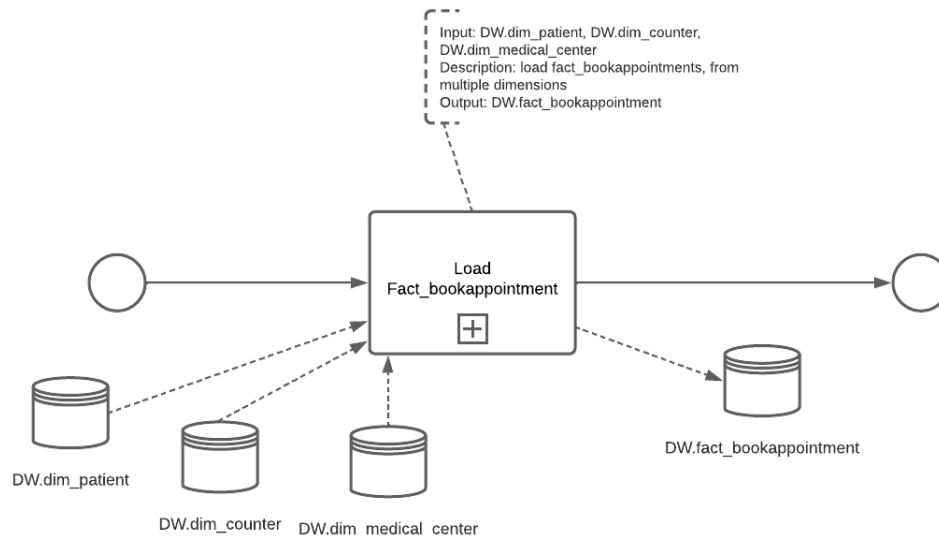


Figura 19: BPMN para carregar a tabela de factos bookappointment.

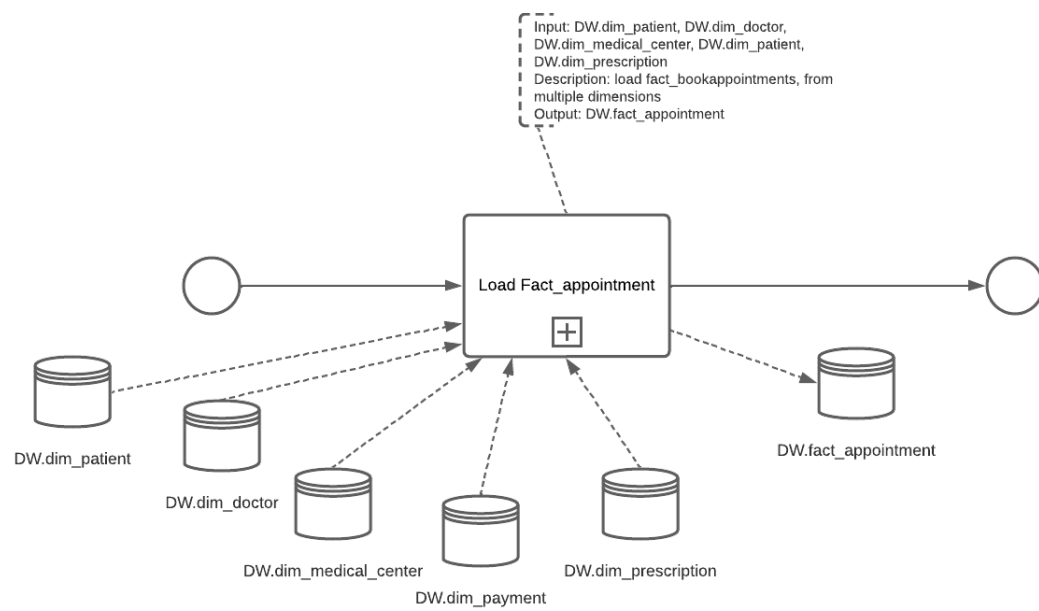


Figura 20: BPMN para carregar a tabela de factos appointment.

Ao finalizar a idealização do ETL, começamos a desenvolver o mesmo, com as ferramentas aprendidas na unidade curricular.

3 Criação do *Data Warehouse*

Para a criação do *Data Warehouse*, dividiu-se o problema em duas partes distintas, o que resultou na criação de dois *packages* no *SQL Server Integration Services* (SSIS).

O primeiro *package* foi criado com o intuito de se proceder à criação do *Data Warehouse* caso este ainda não exista. Neste sentido, começa com uma verificação prévia da existência ou não existência do mesmo. Caso seja comprovada a não existência, procede à criação das diversas dimensões, bem como das tabelas de facto respetivas.

Ainda neste sentido, faz também a população das dimensões de Dia e de Hora. Estas dimensões permitem a verificação de dias e horas do mês onde se verificam certos índices.

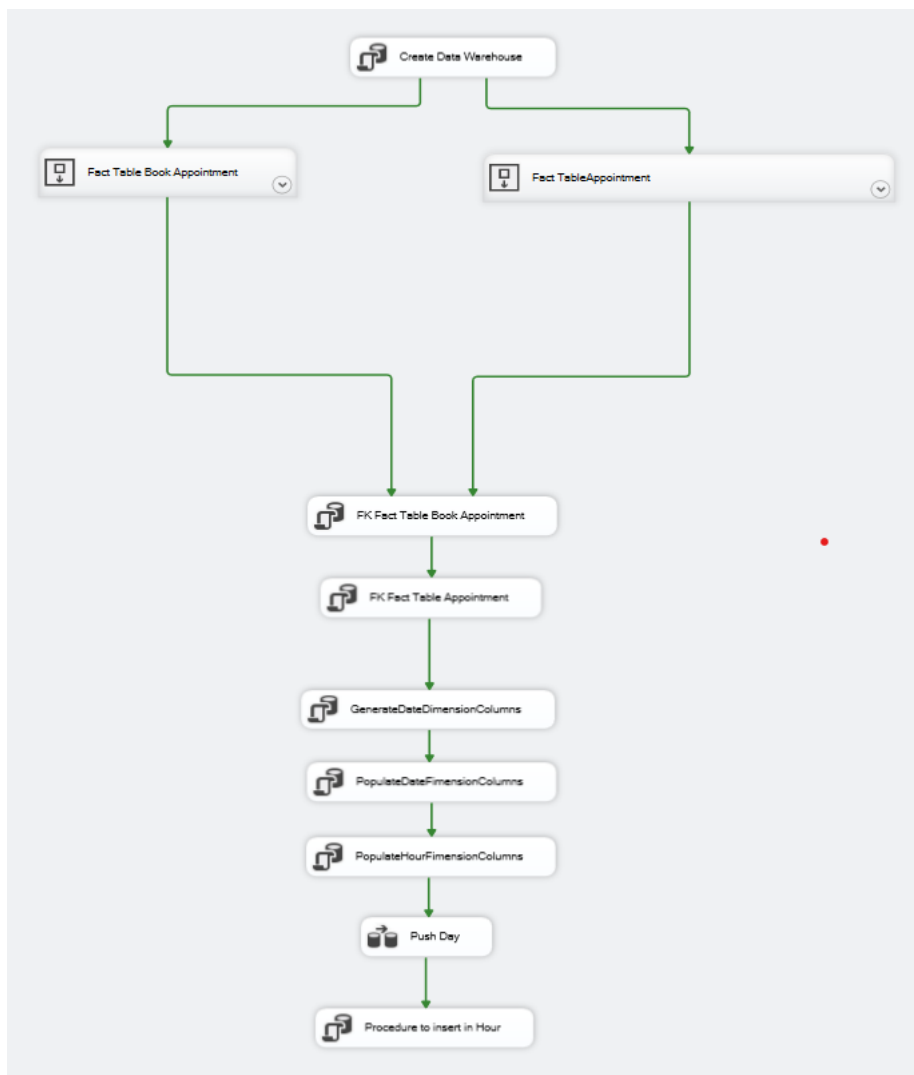


Figura 21: *Package* para a criação do *Data Warehouse* e dimensões de dia e Hora.

Uma vez idealizado e criado o *Data Warehouse*, tem de se criar mecanismo de po-

voamento do mesmo, sendo que para o povoamento do mesmo, utiliza-se uma camada intermédia (*Staging Area*).

Neste sentido, criou-se o segundo Package, que de forma resumida, permite a realização do Povoamento Inicial e do Povoamento incremental do *Data Warehouse*. A estrutura do *Package* verifica-se na Figura 22.

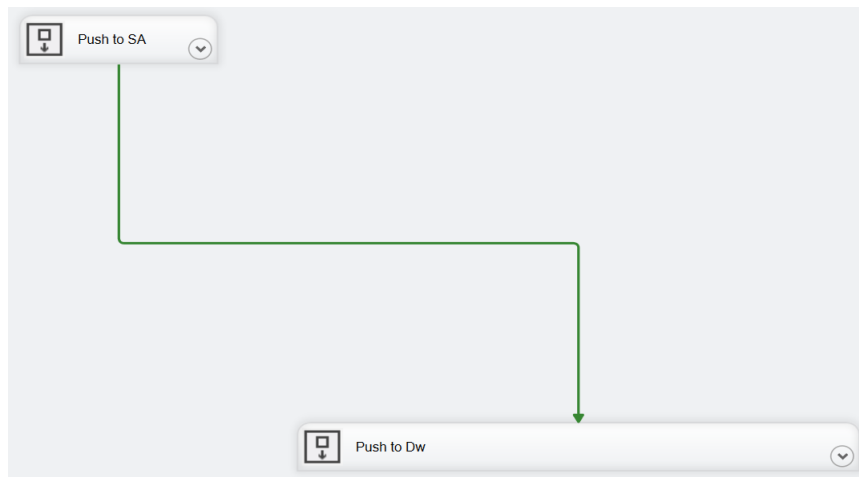


Figura 22: *Package* para o povoamento do *Data Warehouse*.

No primeiro *container* verifica-se a lógica da Implementação do Povoamento incremental, enquanto que o seguinte *container* implementar a lógica de povoamento do *Data Warehouse*, tendo por base a informação presente na *Staging Area*.

Para a implementação do processo de povoamento incremental é necessária a definição de *Change Data Capture*, neste caso através de *Temporal Tables* na Base de Dados Operacional, bem como mecanismos que permitem automatizar o processo, com base na *timestamp* da última Data de Povoamento.

Este mecanismo permite ter a *Staging Area* com dados que apenas foram atualizados ou inseridos após a última data de Povoamento.

4 *Dashboards de Business Intelligence*

O *PowerBi* é um *software* que permite a visualização de dados de forma interativa, desenvolvido pela *Microsoft*. Os dashboards são criados a partir de um conjunto de relatórios, que permitem a obtenção de *insights* e informações úteis a partir dos dados[2]. Posto isto, em seguida, seguem-se os *Dashboards* desenvolvidos no *PowerBi*, que respondem aos objetivos definidos anteriormente.

Neste primeiro *Dashboard*, confirma-se, o número de pacientes e número de doutores existentes no Hospital do caso de estudo, tal como a quantidade de pacientes por doutor e as idades referentes a cada doutor.

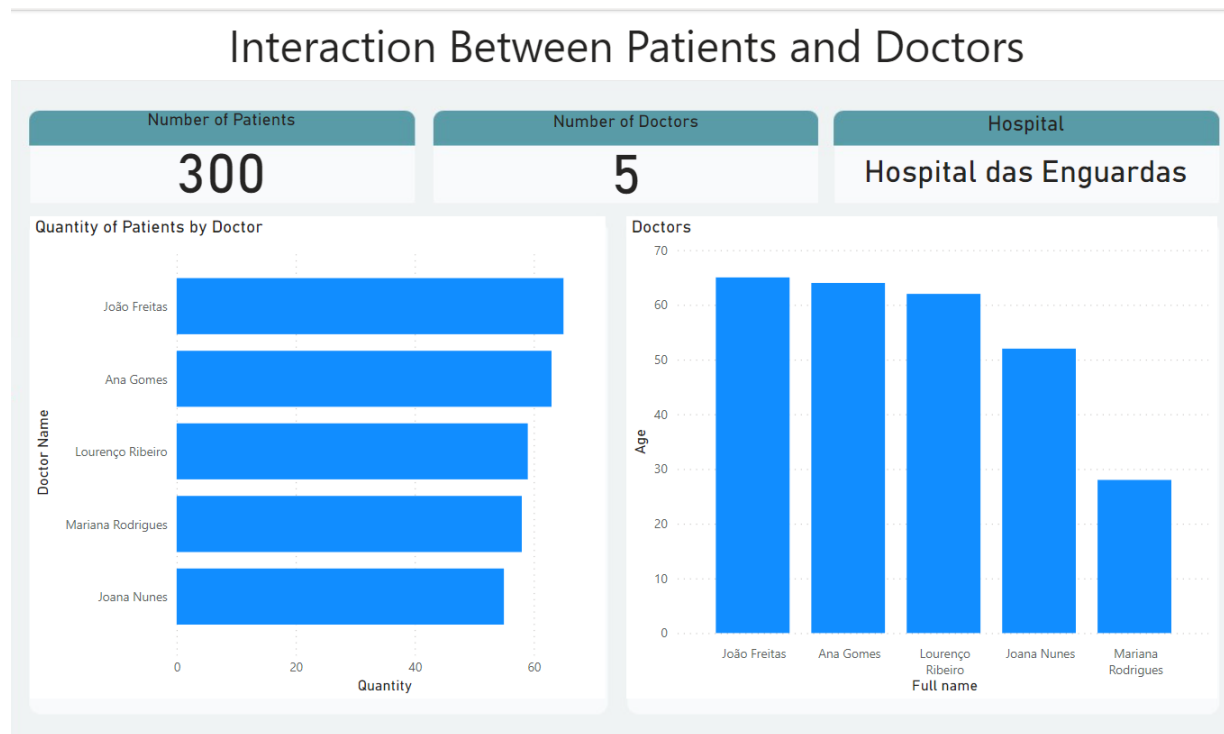


Figura 23: *PowerBI Dashboard* de relação paciente e doutor.

No *Dashboard* seguinte, analisa-se algumas informações sobre os pacientes, tal como, a quantidade de pacientes isentos de pagamento da consulta, a sexualidade mais recorrente e a faixa etária mais recorrente em consultas. Também se verifica alguns dados relativos unicamente ao paciente e informações sobre o ocupação profissional mais recorrente com consultas realizadas.

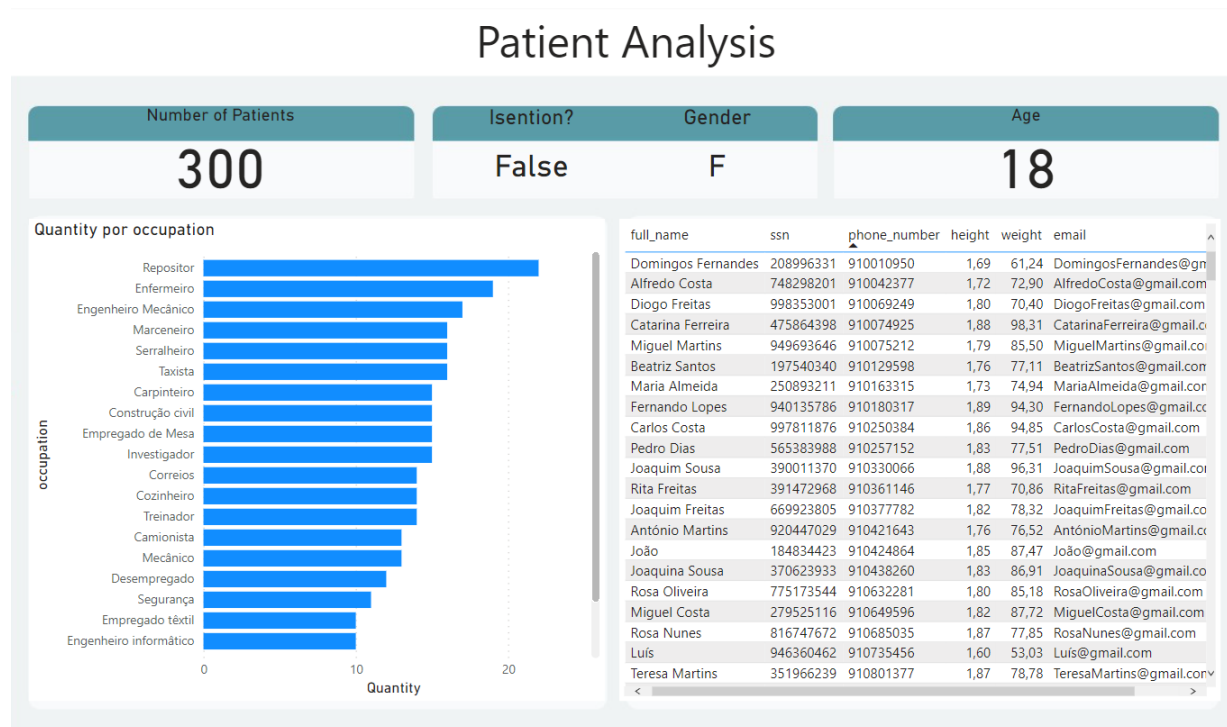


Figura 24: *PowerBI Dashboard* para análise paciente.

Em seguida, no *Dashboard* para análise de *Ticket*, verifica-se alguns dados, como o número total de tentativas de marcação de consulta, o número de tentativas de marcação de consulta realizadas com sucesso e também as sem sucesso. Neste *Dashboard*, também se analisa o número de *Tickets* retirados por dia da semana. No gráfico em anel, pode se verificar os dias de trabalho de cada rececionista, como também o número das tentativas de marcação de consulta com sucesso e sem sucesso, que cada um realizou.

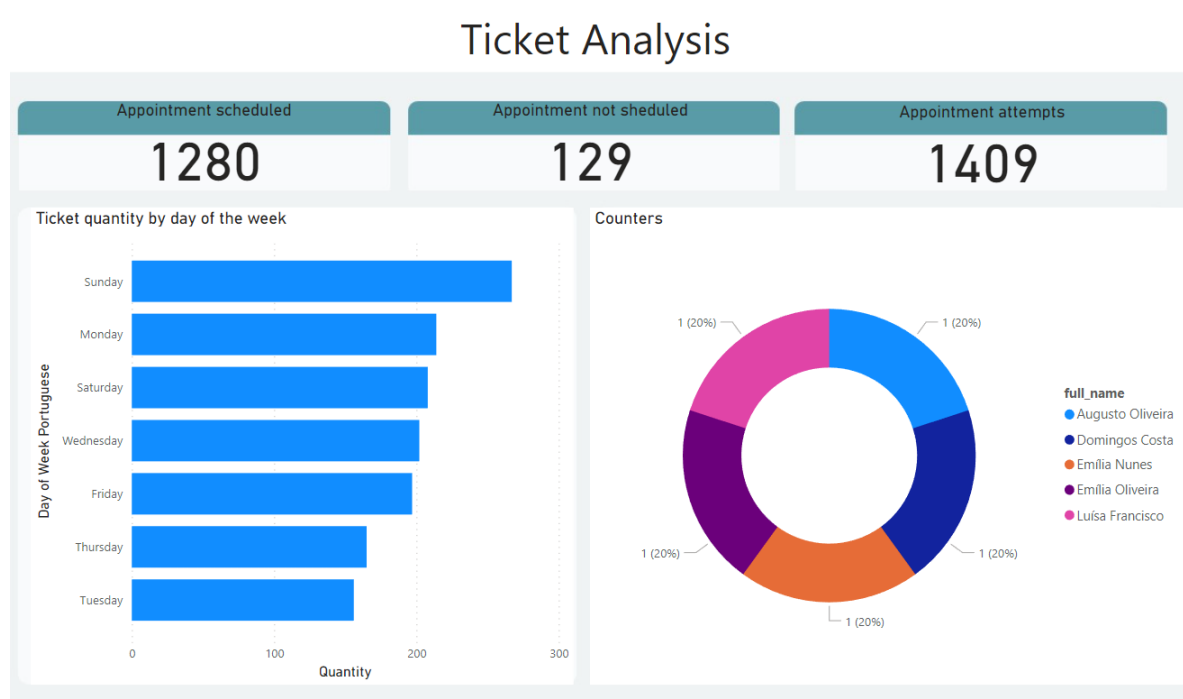


Figura 25: *PowerBI Dashboard* para análise de *Ticket*.

Nas seguintes figuras Figura 26 e Figura 27, verificam-se dados relativos às consultas realizadas no caso de estudo. Neste primeiro *Dashboards*, observam-se alguns dados como a quantidade de consultas realizadas pelo dia de semana, intervalo de horas mais recorrente para a realização de consulta e por fim, faixa etária com mais consultas realizadas.

Appointment Analysis

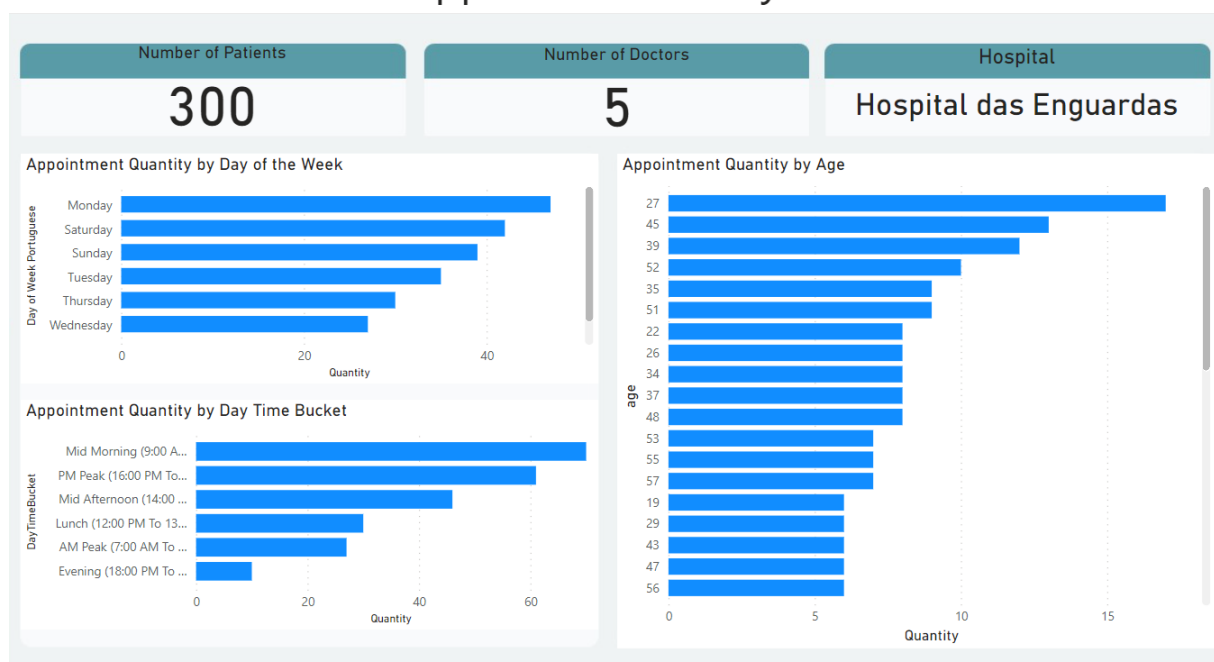


Figura 26: *PowerBI Dashboard* para análise de consultas realizadas.

No seguinte *Dashboard*, podemos verificar informações relativas á profissão mais recorrente na realização de consultas, como também a quantidade de pacientes isentos de pagamento das mesmas.

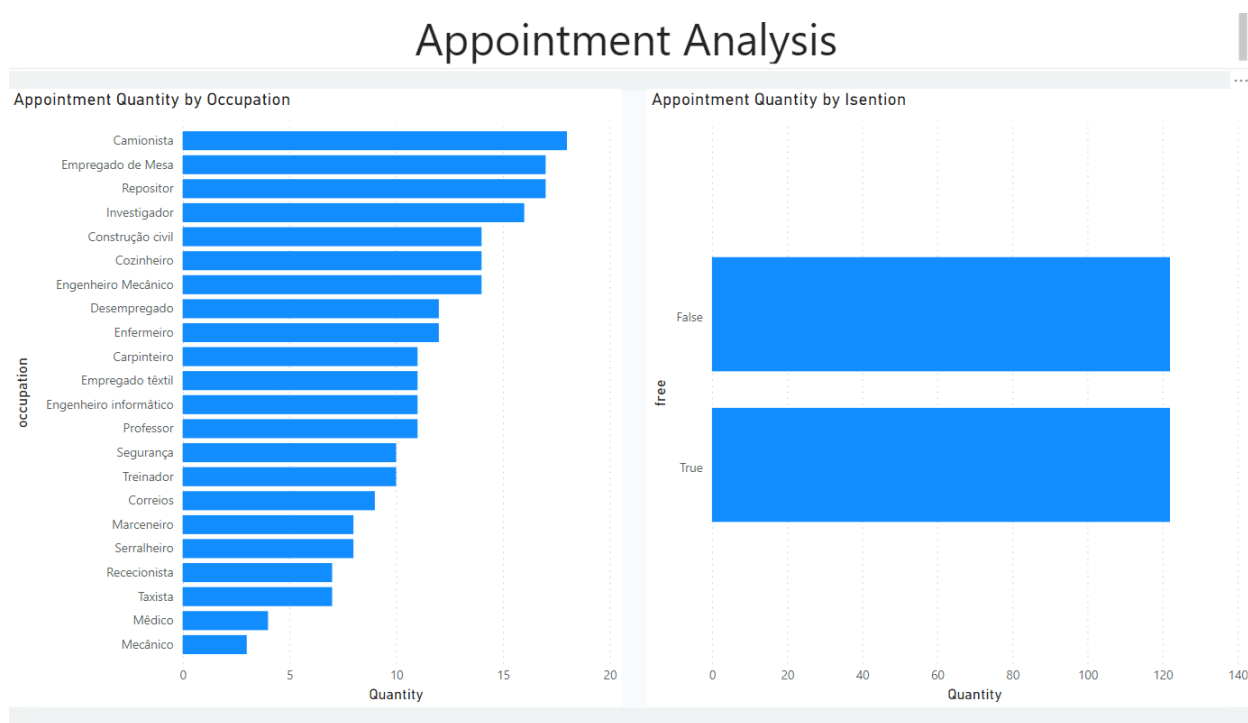


Figura 27: *PowerBI Dashboard* para análise de consultas realizadas.

Por fim, no último *Dashboard*, podemos verificar os dados relativos á analise de prescrição de medicamentos, estes dados englobam o número de prescrições realizadas, os medicamentos existentes e o seu fabricante, quantidade prescrita de medicamentos e a sua dosagem.

Prescription Analysis

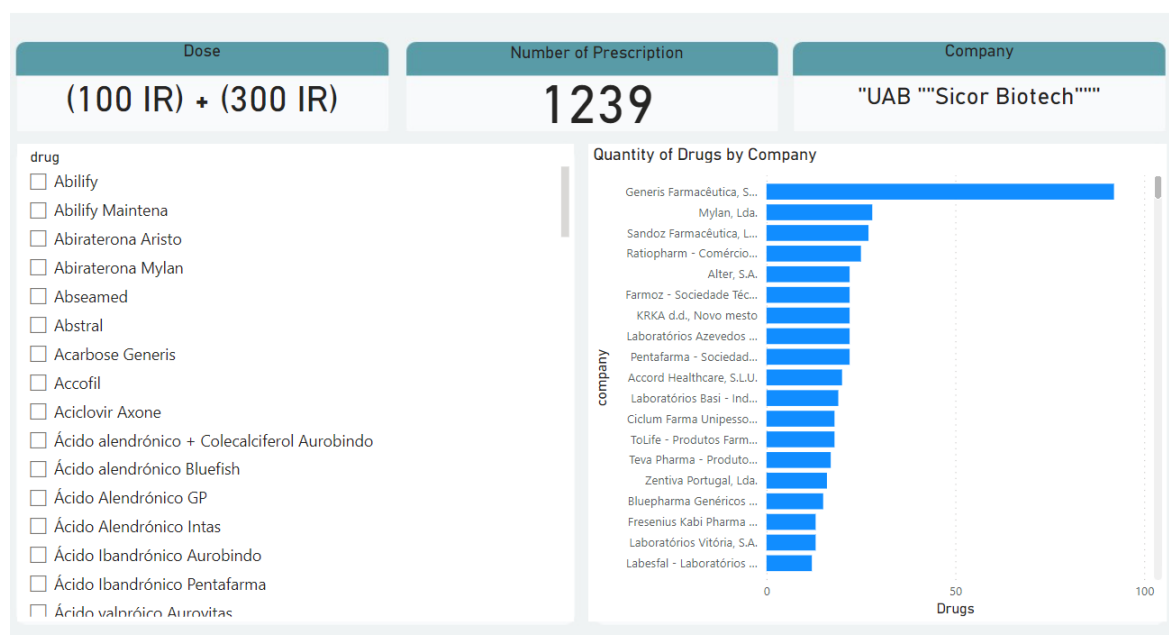


Figura 28: *PowerBI Dashboard* para análise de prescrição de medicamentos.

5 Conclusão

Em retrospectiva, consegue-se afirmar que o seguinte projeto permitiu a colocação em prática de todos os conceitos e conteúdos lecionados ao longo do semestre na Unidade Curricular de Técnicas Avançadas de Análise de Dados, salvo exceção para a utilização de Bases de dados não relacionais, que poderiam ter sido utilizadas na camada da *Staging Area*.

Este projeto foi bastante enriquecedor, no sentido de proporcionar diversas fases distintas e relacionadas, uma vez que nos permitiu a criação de uma Base de dados Operacional de um tema que nos interessava, bem como a modelação de um *Data Warehouse* que nos permitisse responder a diversas perguntas de negócio. Ainda neste sentido, o povoamento inicial e povoamento incremental do *Data Warehouse* recorrendo a uma *Staging Area*. Na perspetiva de análise de Dados de Negócio, isto é, a resposta a perguntas de negócio previamente escolhidas, bem como da análise dos dados de uma forma mais simples e facilitada, utilizou-se o *PowerBi*, o que foi bastante interessante, uma vez que permite a obtenção de *Dashboards* diretamente de dados obtidos do *Data Warehouse*, de forma bastante prática e eficaz.

Contudo, ainda existem diversas melhorias que poderiam ter sido implementadas. Primeiramente e possivelmente a mais fulcral, uma melhoria nos dados Operacionais, o que permite uma análise mais realista do problema em questão, bem como a obtenção de melhores métricas. Por outro lado, a melhoria dos Diagrama de processos relativos ao ETL, assim como a implementação do mesmo.

Em suma, o seguinte trabalho foi bastante enriquecedor, dado que nos permitiu perceber acerca das diversas etapas de um *Sistema de Business Intelligence*, das diversas arquiteturas possíveis para a obtenção do mesmo, bem como perceber as diversas questões e requisitos que se devem colocar do ponto de vista funcional e de negócio.

Referências

- [1] Bruno Oliveira. Modelação dimensional - dimensões e factos, 2022 – 2023.
- [2] Ana Henriques. Introdução ao *PowerBi*, 2023.