



# **Business Intelligence em Dados da Indústria Alimentar**

**JEAN CARLOS FERNANDES RAMALHO VIDAL SABENÇA**

Outubro de 2022

# **Business Intelligence em Dados da Indústria Alimentar**

**Jean Carlos Sabença**

**Dissertação para obtenção do Grau de Mestre em  
Engenharia Informática, Área de Especialização em  
Sistemas de Informação e Conhecimento**

**Orientador: Paulo Oliveira**

**Júri:**

Presidente:

Vogais:

Porto, outubro 2022



# Dedicatória

Dedico esta dissertação à minha filha Maria, ao meu filho Henrique, à minha esposa Elisabete e aos meus Pais.



# Resumo

Os dados que serviram como caso de estudo para a concretização desta dissertação foram fornecidos por uma empresa da área alimentar, mais concretamente no setor da indústria de panificação e pastelaria. A mesma armazena os dados da sua operação num ERP, incluindo os dados das vendas.

A informação resultante das vendas torna-se relevante para a melhoria do desempenho da área comercial, assim como para outras áreas. A análise dessa informação serve de apoio ao processo de tomada de decisões operacionais e estratégicas.

Foram elaborados uma série de indicadores acerca dos comerciais, produtos e clientes, que possibilitam retirar conclusões como a variação da quantidade de uma categoria de produtos, que produtos são adquiridos pelos clientes, qual o volume de vendas de um comercial ou de uma determinada equipa de comerciais, que tipos de produtos são mais vendidos, entre outras.

Foi desenvolvida uma solução centralizada de Business Intelligence com dados obtidos de um cubo de vendas, com uma boa performance e dados confiáveis. A solução está preparada de forma a receber futuras atualizações e melhorias.

Foram desenvolvidos e disponibilizados alguns *dashboards* que permitem aos utilizadores realizarem análises acerca das categorias de produtos, clientes, tipos de cliente, produtos por cliente, localizações, comerciais e equipas.

Para finalizar, para avaliar o sistema desenvolvido, foram realizados teste unitários e elaborado um inquérito de satisfação aos utilizadores. Com base nos resultados obtidos, a solução implementada cumpre os objetivos inicialmente definidos.

**Palavras-chave:** ERP, vendas, indicadores, comerciais, produtos, clientes, Business Intelligence, cubo, dashboards.



# Abstract

The data that served as a case study for the realization of this dissertation were provided by a food company, more specifically in the bakery and pastry industry. It stores your operation data in an ERP, including sales data.

The resulting sales information becomes relevant for the improvement of the performance of the commercial area, as well as for other areas. The analysis of this information supports the process of making operational and strategic decisions.

A series of indicators were elaborated about the salespersons, products, and customers, which make it possible to draw conclusions such as the variation of the quantity of a category of products, which products are purchased by customers, what the sales volume of a salesperson or a particular team of salespersons, which types of products are sold, among others.

A centralized Business Intelligence solution was developed with data obtained from a sales cube, with good performance and reliable data. The solution is prepared to receive future updates and improvements.

Some dashboards have been developed and made available to allow users to perform analyses about product categories, customers, customer types, products per customer, locations, salespersons, and teams.

Finally, to evaluate the developed system, a unit test was performed, and a user satisfaction survey was conducted. Based on the results obtained, the implemented solution meets the objectives initially defined.

**Keywords:** ERP, sales, indicators, salespersons, products, customers, Business Intelligence, cube, dashboards.



# Agradecimentos

A concretização desta dissertação só foi possível devido ao apoio e incentivo recebido por parte de algumas pessoas e organizações.

Quero agradecer em primeiro lugar à minha esposa Elisabete pelo apoio, incentivo e compreensão durante este percurso. Tenho de agradecer imensamente à minha filha Maria e ao meu filho Henrique, que irá nascer no próximo mês de dezembro, por toda a motivação. Sem o apoio e motivação da minha família tudo isto não teria sido possível.

Agradeço também ao meu orientador, o professor Paulo Oliveira por me acompanhar nesta jornada. Obrigado pela motivação, disponibilidade, sugestões e ensinamentos.

Também quero agradecer à Prodite Zeelandia por ter possibilitado e criado condições para a realização desta dissertação. Agradeço aos meus colegas o apoio e cooperação.

Por fim, quero deixar um agradecimento ao ISEP, principalmente ao DEI pelo contributo dado ao longo do meu percurso académico.



*“Sem ambição, não se começa nada. Sem trabalho, não se termina nada. O prémio não será enviado para si! Tem de vencer.”*

**Ralph Waldo Emerson**



# Índice

<b>1. Introdução</b> .....	<b>1</b>
1.1 Contexto .....	1
1.2 Problema .....	2
1.3 Motivação .....	2
1.4 Objetivos .....	2
1.5 Abordagem .....	3
1.6 Estrutura do documento .....	3
<b>2. Estado da Arte</b> .....	<b>5</b>
2.1 Business Intelligence .....	5
2.2 Componentes de um Sistema de Business Intelligence .....	6
2.2.1 Processo de ETL .....	7
2.2.2 Staging Area .....	7
2.2.3 Armazém de Dados .....	8
2.2.4 Data Mart .....	10
2.2.5 Modelação Dimensional .....	10
2.2.6 Factos .....	11
2.2.7 Dimensões .....	11
2.2.8 Esquemas de modelos de dados .....	12
2.2.9 Modelo Multidimensional .....	14
2.3 Ferramentas de Business Intelligence .....	16
2.3.1 Ferramentas de ETL .....	17
2.3.2 Ferramentas de análise de dados .....	23
<b>3. Análise de Valor</b> .....	<b>33</b>
3.1 Processo de Inovação .....	33
3.1.1 Identificação da oportunidade .....	35
3.1.2 Análise da oportunidade .....	35
3.1.3 Geração de ideias .....	36
3.1.4 Seleção da Ideia .....	37
3.1.5 Definição do conceito .....	37
3.2 Análise de valor .....	38
3.2.1 Valor, valor para o cliente e valor percebido .....	38
3.3 Proposta de Valor .....	39
3.4 Modelo de negócio Canvas .....	39
3.5 Cadeia de valor de Porter .....	40
3.6 Método de Análise Hierárquica para Comparação de Tecnologias .....	42
<b>4. Análise e Design</b> .....	<b>49</b>

4.1	Domínio do problema.....	49
4.2	Identificação dos atores.....	50
4.3	Engenharia de requisitos.....	50
4.3.1	Requisitos não funcionais .....	50
4.3.2	Requisitos funcionais .....	51
4.4	Opções de arquitetura.....	53
4.5	Arquitetura escolhida .....	54
4.5.1	Fontes de dados .....	55
4.5.2	Logical Data Map .....	56
4.5.3	Staging Area .....	56
4.5.4	Modelação do Data Warehouse.....	58
4.5.5	Dimensões.....	59
4.5.6	Tabela de Factos .....	64
<b>5.</b>	<b>Implementação.....</b>	<b>67</b>
5.1	Tecnologias Utilizadas.....	67
5.2	Processo de ETL .....	67
5.2.1	Extração de dados .....	67
5.2.2	Transformação e Carregamento de dados.....	69
5.2.3	Cubo de vendas.....	73
5.3	Análise de dados .....	74
<b>6.</b>	<b>Apresentação e Análise de Resultados .....</b>	<b>77</b>
6.1	Dashboards.....	77
6.1.1	Análise de vendas por categoria de produto .....	77
6.1.2	Análise de vendas por cliente .....	78
6.1.3	Análise de vendas por tipo cliente.....	79
6.1.4	Análise de referências por cliente.....	80
6.1.5	Análise de vendas por localização .....	81
6.1.6	Análise de vendas por distribuidor.....	82
6.1.7	Análise de vendedor .....	83
6.1.8	Análise por equipa .....	85
6.2	Análises Excel.....	86
<b>7.</b>	<b>Avaliação de Resultados.....</b>	<b>87</b>
7.1	Grandezas em Avaliação .....	87
7.2	Metodologias de Avaliação.....	87
7.3	Resultados obtido dos testes unitários.....	89
7.3.1	Resultados dos testes de conformidade.....	89
7.3.2	Resultados dos testes do tempo de execução da atualização de dados.....	90
7.4	Resultados obtido do inquérito aos utilizadores.....	91
<b>8.</b>	<b>Conclusão .....</b>	<b>93</b>
8.1	Objetivos Atingidos .....	93

8.2	Dificuldades Sentidas .....	94
8.3	Trabalho Futuro .....	94



# Lista de Figuras

Figura 1: Processo de <i>Business Intelligence</i> (Larion, 2021).....	6
Figura 2: Abordagem de Inmon (Moscoso-Zea, Paredes-Gualtor, and Luján-Mora, 2018).....	9
Figura 3: Abordagem de Kimball (Moscoso-Zea, Paredes-Gualtor, and Luján-Mora, 2018) .....	10
Figura 4: Modelo em Estrela (1&1 IONOS, 2021) .....	13
Figura 5: Modelo em Floco de Neve (1&1 IONOS, 2021).....	14
Figura 6: Modelo Multidimensional (1&1 IONOS, 2021) .....	15
Figura 7: Quadrante Mágico relativo a ferramentas de integração de dados (Gartner, 2021). 16	
Figura 8: Informatica PowerCenter Designer (Informatica, 2022).....	19
Figura 9: IBM DataStage (IBM, 2022).....	20
Figura 10: Microsoft SSIS (Microsoft, 2022a).....	21
Figura 11: Hitachi Vantara PDI (Hitachi Vantara, 2022).....	22
Figura 12: Quadrante Mágico relativo a ferramentas de análise de dados (Gartner, 2022).....	24
Figura 13: Principais elementos do Microsoft Power BI (Microsoft, 2022b) .....	25
Figura 14: Tableau (Tableau, 2022).....	26
Figura 15: Qlik (QlikTech International 2022) .....	27
Figura 16: MicroStrategy Workstation (MicroStrategy, 2022) .....	28
Figura 17: Oracle Business Intelligence (Oracle, 2022).....	29
Figura 18: Interface SAS (SAS Institute Inc., 2022).....	30
Figura 19: Interface Pentaho (Hitachi Vantara, 2022) .....	31
Figura 20: Modelo de New Concept Development (Belliveau, P. and Griffin, A. 2002) .....	34
Figura 21: Modelo de negócio Canvas .....	40
Figura 22: Cadeia de valor de Porter (Nagy, J. <i>et al.</i> 2018).....	41
Figura 23: Estrutura hierárquica AHP.....	43
Figura 24: Diagrama de casos de uso.....	51
Figura 25: Diagrama de sequência visualizar dashboard.....	52
Figura 26: Opção de arquitetura 1.....	53
Figura 27: Opção de arquitetura 2 ROLAP .....	54
Figura 28: Opção de arquitetura 3 MOLAP .....	54
Figura 29: Parte do modelo relacional da base de dados (DynamicsDocs, 2022) .....	55
Figura 30: Modelo da <i>Staging Area</i> .....	57
Figura 31: Dimensão Data.....	60
Figura 32: Dimensão Envio.....	61
Figura 33: Dimensão Expedidor .....	61
Figura 34: Dimensão Fornecedor.....	62
Figura 35: Dimensão Localização .....	62
Figura 36: Dimensão Vendedor.....	63
Figura 37: Tabela de Factos Vendas.....	64
Figura 38: Data Mart Vendas .....	65
Figura 39: Etapa de extração na arquitetura da solução .....	68
Figura 40: Vistas criadas na base de dados operacional.....	68

Figura 41: Carregamento das tabelas da <i>Staging Area</i> .....	69
Figura 42: Etapa de transformação e carregamento na arquitetura da solução .....	70
Figura 43: Carregamento da tabela DimProduto .....	71
Figura 44: Carregamento da tabela FactVendas .....	72
Figura 45: Cubo de Vendas Construído .....	73
Figura 46: Hierarquia de produto .....	74
Figura 47: Hierarquia de data .....	74
Figura 48: Etapa de análise de dados na opção de arquitetura implementada .....	74
Figura 49: Análise de vendas por categoria de produto .....	77
Figura 50: <i>Drill down</i> quantidade base por categoria de produto .....	78
Figura 51: Análise de vendas por cliente .....	79
Figura 52: Análise de vendas por tipo de cliente .....	80
Figura 53: Análise de referências por cliente .....	81
Figura 54: Análise de vendas por localização .....	82
Figura 55: Análise de vendas por distribuidor .....	83
Figura 56: Análise de vendedor .....	84
Figura 57: <i>Drill down</i> análise vendedor .....	84
Figura 58: Análise por equipa .....	85
Figura 59: Exemplo de análise realizada em Excel .....	86
Figura 60: Inquérito enviado aos utilizadores .....	89
Figura 61: Comparação valores ERP e valores obtidos do cubo de vendas .....	90



# Lista de Tabelas

Tabela 1: Comparação de ferramentas ETL .....	22
Tabela 2: Comparativo entre ferramentas de análise de dados .....	32
Tabela 3: Escala fundamental de Saaty (Saaty, 1987).....	44
Tabela 4: Matriz de prioridades entre critérios .....	44
Tabela 5: Matriz de critérios normalizada, com prioridade relativa.....	44
Tabela 6: Valores de IR para matrizes quadradas de ordem n .....	45
Tabela 7: Matriz de comparação do critério Performance .....	46
Tabela 8: Matriz de comparação do critério Integração.....	46
Tabela 9: Matriz de comparação do critério Usabilidade .....	46
Tabela 10: Matriz final do AHP para escolha de ferramenta de BI .....	47
Tabela 11: Dimensão Cliente.....	59
Tabela 12: Dimensão Produto.....	63
Tabela 13: Objetivos e filtros aplicados a cada vista.....	68
Tabela 14: Escala de Likert .....	88
Tabela 15: Tempos de execução do processo de integração e atualização dos dados .....	90
Tabela 16: Frequência de respostas dos utilizadores .....	91



# Acrónimos

## Lista de Acrónimos

<b>AHP</b>	<i>Analytic Hierarchy Process.</i>
<b>BI</b>	<i>Business Intelligence.</i>
<b>ERP</b>	<i>Enterprise Resource Planning.</i>
<b>DW</b>	<i>Data Warehouse.</i>
<b>ERP</b>	<i>Enterprise Resource Planning.</i>
<b>ETL</b>	<i>Extract, Transform, Load.</i>
<b>OLAP</b>	<i>Online Analytical.</i>
<b>MOLAP</b>	<i>Multidimensional Online Analytical Processing.</i>
<b>ROLAP</b>	<i>Relational Online Analytical Processing.</i>
<b>SCD</b>	<i>Slowly Changing Dimension.</i>





# 1. Introdução

Neste capítulo é feito o enquadramento do trabalho, nomeadamente a descrição do tema de estudo realizado no âmbito da unidade curricular Tese de Mestrado em Engenharia Informática, lecionada no Instituto Superior de Engenharia do Porto. Será apresentado o problema que deu origem a esta dissertação, quais os objetivos a atingir, que abordagem será utilizada para resolver o problema e, por fim, é descrita a estrutura do documento.

## 1.1 Contexto

Atualmente as organizações lidam cada vez mais com um grande volume de dados, surgindo uma necessidade cada vez maior de análise desses mesmos dados por parte de gestores e dos mais diversos colaboradores. Uma boa gestão e análise dos dados é elementar para suportar a tomada de decisão, para isso a informação disponibilizada ao utilizador deve estar estruturada e resumida.

Com o aumento da competitividade no mercado, torna-se fundamental disponibilizar informação fidedigna de uma forma rápida para permitir às organizações melhorarem o seu desempenho e assim alcançarem uma vantagem competitiva face à concorrência. Se uma organização conseguir transformar a informação que detém em conhecimento pode possibilitar uma melhoria importante do seu negócio.

Na área alimentar, tal como em muitas outras, torna-se importante perceber se os objetivos e metas a atingir definidos pela organização estão a ser cumpridos. O recurso a *dashboards* é uma das formas de visualizar o seu estado e desempenho. O desempenho do departamento comercial de uma organização pode influenciar positivamente ou negativamente o sucesso da mesma.

Neste contexto, as ferramentas de Business Intelligence assumem um papel de relevo no processo de recolha, preparação e representação de dados, ajudando desta forma na tomada de decisão e gestão estratégica das organizações.

Este trabalho desenvolve-se na organização Prodite Zeelandia, a qual pretende uma ferramenta de BI, que permita obter análises a partir das vendas da equipa comercial. Essas análises devem ser obtidas de uma forma célere e com dados confiáveis.

## 1.2 Problema

Cada vez mais as organizações lidam com grandes volumes de dados, surgindo uma maior necessidade de análise desses mesmos dados por parte de gestores e dos mais diversos colaboradores. A capacidade de gestão e análise de dados analíticos de uma organização são pontos fundamentais para o desenvolvimento do negócio, uma vez que a tomada de decisão é suportada pela informação adquirida a partir dos mesmos.

Como ponto de partida para esta dissertação tem-se o problema com que uma organização da área alimentar se depara, que é a necessidade de ter um sistema de BI que permita analisar dados das vendas da equipa comercial, que neste momento é feita com recurso a ficheiros Excel, com ligações à base de dados transacional ou através de mapas tabulares desenvolvidos no *Enterprise Resource Planning* (ERP).

Os dois métodos trazem bastantes inconvenientes. Os ficheiros Excel com ligações à base de dados transacional tornam o processo de obtenção dos dados bastante lento, obrigam o utilizador a conhecer quais as tabelas da base de dados que contêm a informação relevante e não permitem guardar histórico de alterações. O desenvolvimento dos mapas no ERP torna o processo dispendioso e limitativo quanto à forma de apresentação das análises. Com os processos atuais torna-se difícil garantir a confiabilidade dos dados.

## 1.3 Motivação

A possibilidade de encontrar uma solução que permita resolver as dificuldades atuais no processo de análise de dados, adquirindo assim experiência na área de BI.

O desafio de poder desenvolver uma solução que permita resolver os problemas anteriormente descritos na secção 1.2, assim como a possibilidade de resolver um problema real aplicando os conhecimentos adquiridos ao longo do percurso académico. A oportunidade de poder desenvolver um projeto na área de *Business Intelligence* (BI), pela qual o autor deste trabalho demonstra grande interesse e na qual nunca teve a oportunidade de desenvolver um projeto, sendo um desafio pessoal, motivador e fundamental para a escolha deste projeto.

## 1.4 Objetivos

Considerando o problema descrito anteriormente, as necessidades identificadas e as motivações acima referidas os objetivos são:

- Desenvolvimento de um armazém de dados com os dados das vendas da organização, que crie um histórico baseado nas transações registadas no ERP da organização, de modo a possibilitar análises de dados mais detalhadas.

- Elaborar um processo de ETL – Extração, Transformação e Carregamento de dados.
- Construção de um cubo de dados a partir do armazém de dados criado com os dados da organização.
- Criação de *dashboards* simples e intuitivos que facilitem a visualização e assimilação da informação.
- Desenvolver uma solução que permita realizar análises mais expeditas, detalhadas e confiáveis.

## 1.5 Abordagem

Para resolver o problema exposto e atingir os objetivos do presente trabalho, vai ser utilizada uma abordagem composta pelos passos seguintes:

- Efetuar um estudo sobre o conceito de *Business Intelligence*, os componentes que o compõem e a sua arquitetura.
- Realizar um levantamento do estado da arte das ferramentas de ETL, ferramentas de BI e analisar as suas funcionalidades.
- Desenhar e implementar um armazém de dados e respetivo cubo analítico.
- Desenhar e implementar uma solução, com recurso a uma ferramenta de BI, que responda às necessidades de análise dados da organização.
- Avaliar o desenvolvimento de todo o processo.

## 1.6 Estrutura do documento

O presente documento encontra-se dividido em oito capítulos.

É iniciado com um capítulo de Introdução, onde é contextualizado e descrito o problema que deu origem a esta dissertação.

O segundo capítulo refere-se ao Estado da Arte, onde serão apresentados conceitos relativos à área de estudo em que este projeto se enquadra, diversas ferramentas de processamento, análise e representação dos dados, que soluções existem no mercado e indicadores de performance utilizados na área comercial.

Segue-se o terceiro capítulo com a Análise de Valor, onde é detalhado que valor é criado com a realização deste projeto e o modelo de negócio em que se baseia.

No quarto capítulo encontra-se a Análise e Design, onde é realizado o levantamento de requisitos (funcionais e não funcionais), identificação dos atores, apresentação de opções de arquitetura e respetivos componentes.

O quinto capítulo contém a descrição da implementação do sistema, desde o processo de ETL, até à criação do cubo de dados.

No sexto capítulo são descritos os *dashboards* que foram desenvolvidos, assim como analisados os dados apresentados.

No sétimo capítulo descreve-se a avaliação da solução, através de teste unitários e inquéritos de satisfação.

O oitavo capítulo contém as conclusões obtidas nesta dissertação, incluindo os objetivos atingidos, dificuldades sentidas e trabalho futuro que poderá ser desenvolvido.

## 2. Estado da Arte

Neste capítulo será apresentado o estudo que foi realizado de forma a suportar o trabalho realizado. São abordados vários temas na área em que este trabalho se insere, nomeadamente conceitos relacionados com a área de *Business Intelligence* (BI), processo de ETL, arquiteturas, tecnologias, ferramentas e comparação entre softwares. Também é feita a comparação na qual se baseia a seleção da ferramenta para a construção do armazém de dados e respetiva ferramenta para análise e apresentação dos dados. As escolhas tiveram em consideração vários critérios, tais como a facilidade de utilização, integração, performance, variedade de fontes de dados, suporte por parte da comunidade que as utiliza e documentação existente. Foi também tido em conta o valor de licenciamento/aquisição de cada ferramenta assim como a regularidade com que recebem atualizações (melhoramentos, correções).

### 2.1 Business Intelligence

Existem inúmeros conceitos e definições relativas ao *Business Intelligence* (BI), mas de uma forma simplista, o BI pode ser visto como o processo que fornece informações confiáveis no momento certo às pessoas certas de forma a tomarem a decisão mais rápida e acertada. De acordo com a definição de Howson, para gerir o negócio, melhorar o seu desempenho, descobrir oportunidades e trabalhar de forma eficiente, o BI permite que pessoas de todos os níveis de uma organização acedam, interajam e analisem dados (C. Howson 2008).

O BI consiste num processo de recolha, organização, partilha, monitorização e análise de grandes quantidades de informação. Os processos organizacionais geram uma quantidade significativa de dados. Os sistemas de BI tiram partido desses dados, transformando-os em informação útil, de forma a suportar a tomada de decisão. Promovem o aumento do nível de conhecimento da organização, com base em análises de dados, permitindo assim a melhoria da estratégia organizacional.

Os sistemas de BI marcam presença em vários setores de negócio, por exemplo, o retalho, o setor financeiro, telecomunicações, setor dos transportes, educação, saúde, comércio eletrónico e seguros. A partir desta interpretação de tendências e através da análise de dados ou da utilização de técnicas de previsão (e.x., *data mining*) os sistemas de BI conseguem desenvolver estratégias e oportunidades de negócios, tendo em vista (a curto prazo) a vantagem competitiva no mercado e (a longo prazo) a estabilidade. Assim sendo, através do uso do BI, as organizações garantem a qualidade e promovem a criação e disseminação de conhecimento (CCG, 2018).

O BI é utilizado para entender os recursos disponíveis da empresa, tendências de mercado, ações dos concorrentes e implicações destas ações (Negash, 2004).

Pode-se afirmar que os sistemas de *Business Intelligence* são considerados sistemas de apoio à decisão (DSS) orientados por dados, que utilizam tecnologia para poder analisá-los. É a partir dos resultados obtidos que se consegue obter informações vantajosas para o negócio e apoiar na tomada de decisão (Teixeira, Filho and Costa, 2018). O elevado volume de dados sofre um processo de transformação, através de ferramentas ETL (Extract, Transform, Load) para que mais tarde esses mesmos dados possam ser representados de uma forma visualmente apelativa, evidenciando os pontos essenciais do negócio, nomeadamente tendências e problemas que eram até então desconhecidos. Isto torna possível às organizações medirem as suas ações perante os resultados e desenvolver estratégias competitivas essenciais para se manterem no mercado.

## 2.2 Componentes de um Sistema de Business Intelligence

Segundo Sherman (Sherman, 2015), os três principais blocos que constituem um sistema de *Data Warehousing/BI* são: Integração dos dados, Armazenamento de dados e Apresentação e análise da informação. Na Figura 1 é possível ver como estes se relacionam.

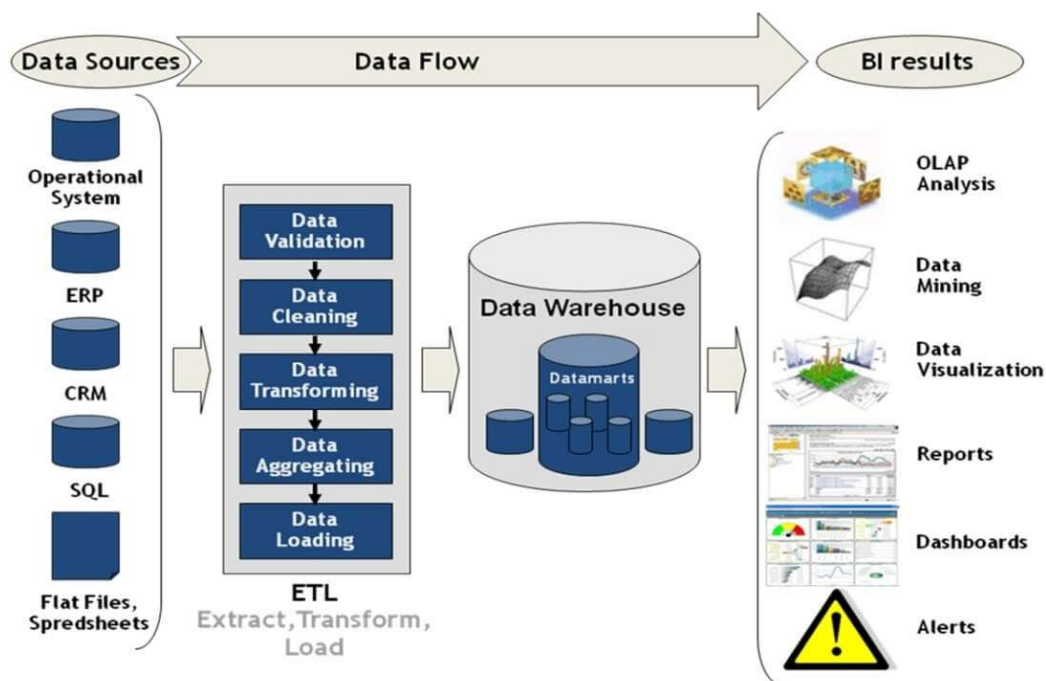


Figura 1: Processo de *Business Intelligence* (Larion, 2021)

O bloco de Integração dos dados (ETL na Figura 1) é responsável pelas etapas:

- Extração de dados de diversas fontes de dados.
- Combinação e tratamento de dados.
- Armazenamento de dados no Armazém de Dados (AD).

No AD o processo de armazenamento é separado das transações de negócio sendo otimizado o acesso e análise de dados. Finalizado o processo de BI são realizadas representação dos dados que o negócio utiliza para adquirir conhecimento e tomar as melhores opções para a empresa (Sherman, 2015).

### 2.2.1 Processo de ETL

ETL é um tipo de integração de dados que se refere às três etapas (extração, transformação, carregamento) usadas para misturar dados de origens diversas. É usado para construir um armazém de dados. Durante o processo, os dados são retirados (extraídos) de um sistema de origem, convertidos (transformados) num formato no qual podem ser analisados e armazenados (carregados) num armazém de dados ou outro sistema (SAS Institute Inc., 2022).

As três fases que constituem o ETL são:

- **Extract** (extração) - Os dados podem ser extraídos de diversas fontes. Os dados extraídos devem ser consolidados e filtrados para facilitar a realização de operações. Os dados são limpos de forma a estarem livres de qualquer informação desnecessária ou imprevisível para que a eficiência dos resultados obtidos seja a melhor.
- **Transform** (transformação) - Os dados extraídos e limpos precisam de ser transformados num formato normal para poderem cumprir os requisitos do sistema de destino. A informação transformada é recolhida num repositório.
- **Load** (carregamento) - Carrega os dados num sistema de destino, pode ser num *Data Warehouse* ou num *Data Mart*, sendo estes construídos com o objetivo de reduzir o tempo de resposta e prolongar a escalabilidade (Srivastava, G. et al. 2022).

### 2.2.2 Staging Area

A *Staging Area* é uma área de armazenamento intermédio fazendo parte do processo de ETL. Auxilia na transição dos dados das diversas origens, não comprometendo a integridade dos mesmos, para o destino final no *Data Warehouse*. Na *Staging Area* os dados são limpos e sofrem uma série de alterações, com o intuito de eliminar informação desnecessária e uniformizar os dados obtidos a partir das diversas fontes. Os principais motivos para o uso de uma *Staging Area* são aumentar a eficiência dos processos de ETL, garantir a integridade e ajudar a manter a qualidade dos dados. As principais funcionalidades da *Staging Area* são (Qora, 2022):

- **Consolidação:** Uma das principais funções desempenhadas por uma *Staging Area* é a consolidação de dados de diferentes origens. Funciona como o repositório agregador

no qual os dados das várias fontes podem ser colocados temporariamente para posterior processamento.

- **Alinhamento:** O alinhamento de dados inclui a normalização dos dados e a validação de relações entre os mesmos.
- **Deteção de alterações:** Esta funcionalidade é particularmente útil quando os sistemas de origem não suportam formas confiáveis de deteção de alterações, como registo de data e hora ou registo de alterações.
- **Limpeza:** A limpeza de dados inclui identificação e remoção (ou atualização) de dados inválidos dos sistemas de origem.
- **Arquivo de dados e resolução de problemas:** O arquivo de dados pode ser realizado ou suportado por uma *Staging Area*. Esta pode ser usada para manter um histórico dos dados durante o processo de carregamento, ou pode ser usada para enviar os dados para uma outra estrutura de arquivo. Além disso, os dados podem ser mantidos por longos períodos de forma a ajudar na resolução de problemas que possam ocorrer no processo de ETL.

### 2.2.3 Armazém de Dados

*Data Warehousing* é o processo de armazenamento e preparação de dados. Um Armazém de Dados (*Data Warehouse*) é uma base de dados em que os dados são apenas de leitura, reunidos e tornados consistentes. Uma base de dados que guarda históricos de dados e é orientada para o assunto (por exemplo vendas) (Sherman, 2015).

Devido à não volatilidade dos dados e à complexidade das consultas *ad-hoc*, um armazém de dados é fortemente influenciado pela modelação dos dados, ao contrário das bases de dados tradicionais onde o foco reside na definição de um modelo de transações eficiente. As técnicas tradicionais de modelação de dados são inadequadas para um armazém de dados pelas seguintes razões:

- A estrutura da base de dados resultante é muito complexa para ser compreendida pelo utilizador final. No melhor das hipóteses, o modelo gerado é composto por uma centena de tabelas interligadas por um conjunto complexo de relações. Até mesmo as consultas mais simples requerem o estabelecimento de junções entre múltiplas tabelas, que se torna num mecanismo propenso a erros e inapropriado ao acesso eficiente ao grande volume de dados armazenado no armazém de dados;
- Os modelos tradicionais enfatizam a normalização como forma de evitar dados redundantes e consequentes anomalias de atualização. Num armazém de dados os dados não são atualizados e a sua redundância demonstra ser um artifício para maximizar a eficiência das consultas ao repositório (Kimball & Ross, 2013).

Existem duas formas distintas para o design de um armazém de dados. Uma das formas é a abordagem defendida por Bill Inmon (Moscoso-Zea, Paredes-Gualtor, and Luján-Mora, 2018) e a outra a abordagem defendida por Ralph Kimball (Kimball & Ross, 2013). Em ambas abordagens é defendido que o armazém de dados é o repositório central. A forma como os dados são modelados, carregados e armazenados é que difere (Moscoso-Zea, Paredes-Gualtor, and Luján-Mora, 2018).

### 2.2.3.1 Abordagem de Inmon

Inmon propõe uma abordagem de cima para baixo, que fornece uma visão dimensional muito consistente dos dados em *Data Marts* (consultar secção 2.2.4), pois todos os *Data Marts* são carregados do armazém de dados central. A abordagem de cima para baixo é flexível para dar suporte á gestão de mudanças, devido a visão da organização como um todo. Na Figura 2 encontra-se uma representação da abordagem de Inmon.

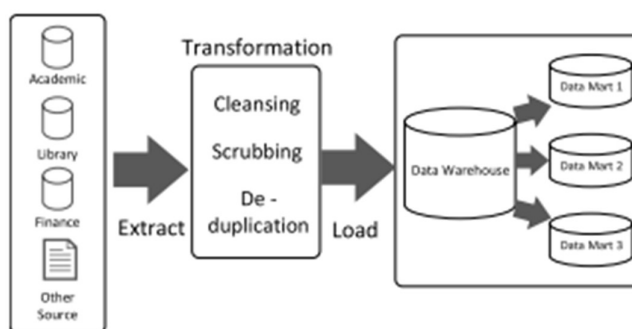


Figura 2: Abordagem de Inmon (Moscoso-Zea, Paredes-Gualtor, and Luján-Mora, 2018)

### 2.2.3.2 Abordagem de Kimball

Por outro lado, Kimball defende uma abordagem de baixo para cima pois considera que é necessário ter em conta os processos de negócio mais importantes da organização. Por essa mesma razão advoga que se deve primeiro criar os *Data Marts*, que detêm os dados mais importantes sobre cada área do negócio e a combinação desses *Data Marts* constitui o *Data Warehouse*. A Figura 3 ilustra a abordagem de Kimball.

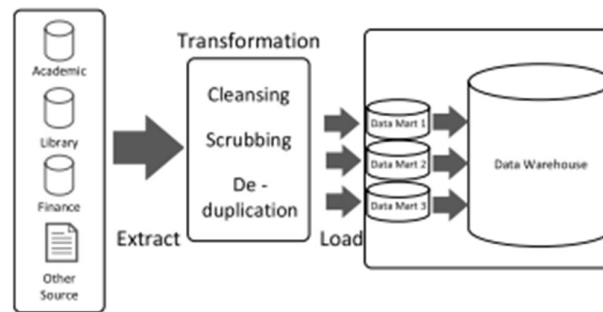


Figura 3: Abordagem de Kimball (Moscoso-Zea, Paredes-Gualtor, and Luján-Mora, 2018)

As duas abordagens são válidas e podem ser utilizadas no design de um *Data Warehouse* a ser implementado numa organização, tendo em consideração os objetivos delineados pela mesma. A abordagem de Inmon pode ser aplicada em organizações que dispõem de bastante tempo para a implementação do sistema de BI e economicamente estáveis, enquanto a abordagem de Kimball pode ser aplicada em organizações que pretendem demorar menos tempo na implementação e cujo foco é a otimização (Moscoso-Zea, Paredes-Gualtor, and Luján-Mora, 2018).

#### 2.2.4 Data Mart

Um *Data Mart* é um subconjunto de dados de um *Data Warehouse*. Um *Data Mart* fornece um acesso mais fácil aos dados mais relevantes para uma determinada equipa ou a um departamento em específico dentro de uma organização.

Os *Data Marts* obtêm os dados desnormalizados e indexados a partir do DW possibilitando desta forma múltiplas consultas em simultâneo e simplificar o processo de pesquisa, tornando-o mais rápido.

Estes repositórios podem ser produzidos a partir de um *Data Warehouse* (abordagem de Inmon) ou podem ser contruídos através de bases de dados operacionais internos ou externos (abordagem de Kimball) (Srivastava, G. et al. 2022).

#### 2.2.5 Modelação Dimensional

A modelação dimensional é amplamente aceite como a técnica preferida para apresentação de dados analíticos porque atende simultaneamente a dois requisitos:

- Disponibiliza dados que são perceptíveis aos utilizadores corporativos;
- Fornece consultas rápidas e de elevado desempenho.

Modelação dimensional é uma técnica usada há vários anos para tornar as bases de dados mais simples. Embora os modelos dimensionais sejam instanciados em sistemas de gestão de base

de dados relacionais, não se encontram na terceira forma normal, em que não pode haver redundância. Um modelo dimensional apresenta redundância, caso contrário seria lento e apresentaria demasiadas tabelas (Kimball & Ross, 2013). Um modelo dimensional é constituído por uma tabela de factos e por tabelas de dimensão.

### 2.2.6 Factos

Segundo Sherman (Sherman, 2015), um fato é uma medida de uma atividade, pode ser um evento ou uma transação do negócio, também é referido como medida de desempenho organizacional. São exemplos de factos as vendas de uma organização e o stock de produtos. Tipicamente representam dados numéricos, com contagens ou valores.

As tabelas de factos são normalizadas e têm pouca redundância, visto que a quantidade de dados apresentados neste tipo de tabelas tem tendência a aumentar. Nas tabelas de factos é onde se encontra maior parte dos dados do modelo dimensional (Sherman, 2015).

### 2.2.7 Dimensões

Uma dimensão tem o objetivo de descrever. É uma entidade que estabelece contexto quando relacionada com factos do negócio. São exemplos de dimensões a data e os produtos de uma fatura. Permite que não haja redundância nas tabelas de factos, tendo em conta que armazena os atributos que estão relacionados com os factos. Esta relação entre tabelas é realizada através de uma chave, sendo que a tabela de dimensão apresenta a chave primária e a tabela de factos a chave estrangeira. Tendo em conta a perspetiva de negócio, uma tabela de dimensões permite filtrar atributos para analisar dados segundo medidas de desempenho (Sherman, 2015).

#### 2.2.7.1 Slowly Changing Dimensions

Uma dimensão pode estar sujeita a sofrer alterações, quer sejam correções a dados introduzidos erradamente, ou alguma mudança que ocorra no sistema operacional. Existe então o conceito de *slowly changing dimension* (SCD), que define uma estratégia para lidar com essas alterações.

Existem quatro tipos de SCD:

- **SCD tipo 1:** atualiza o valor, não guardando histórico;
- **SCD tipo 2:** Insere um novo registo e “desativa” o antigo. É a técnica predominante;
- **SCD tipo 3:** Adiciona uma coluna para o atributo em questão;
- **SCD tipo híbrido:** Insere um novo registo e adiciona uma coluna para o atributo em questão;

O tipo 1 não permite guardar histórico, devendo ser utilizado apenas para fazer correções aos dados existentes. O tipo 2 e 3 permitem guardar histórico, sendo que o tipo 3 permite fazer consultas como se a alteração não tivesse existido enquanto o tipo 2 permite apenas consultas num intervalo de tempo, existindo apenas um "elemento atualizado". O tipo híbrido é uma espécie de combinação entre o tipo 2 e 3, permitindo consultas apenas considerando os estados iniciais dos dados e análises em intervalos de tempo.

## **2.2.8 Esquemas de modelos de dados**

São identificados três tipos de esquemas de modelos de dados (Brito, 2015), são estes:

- Modelo em estrela (*star schema*);
- Modelo em floco de neve (*snowflake schema*);
- Modelo multidimensional (*multidimensional schema*).

### **2.2.8.1 Modelo em Estrela**

O modelo em estrela é o mais utilizado. Neste tipo de modelo a tabela de factos é associada às múltiplas dimensões. Cada tabela desnormalizada do modelo de dados representa uma dimensão em que um atributo está associado a uma tabela normalizada. Um esquema em estrela é intuitivo para o utilizador e apresenta um desempenho de consulta mais rápido (Brito, 2015). Na Figura 4 tem-se a representação do modelo em estrela.

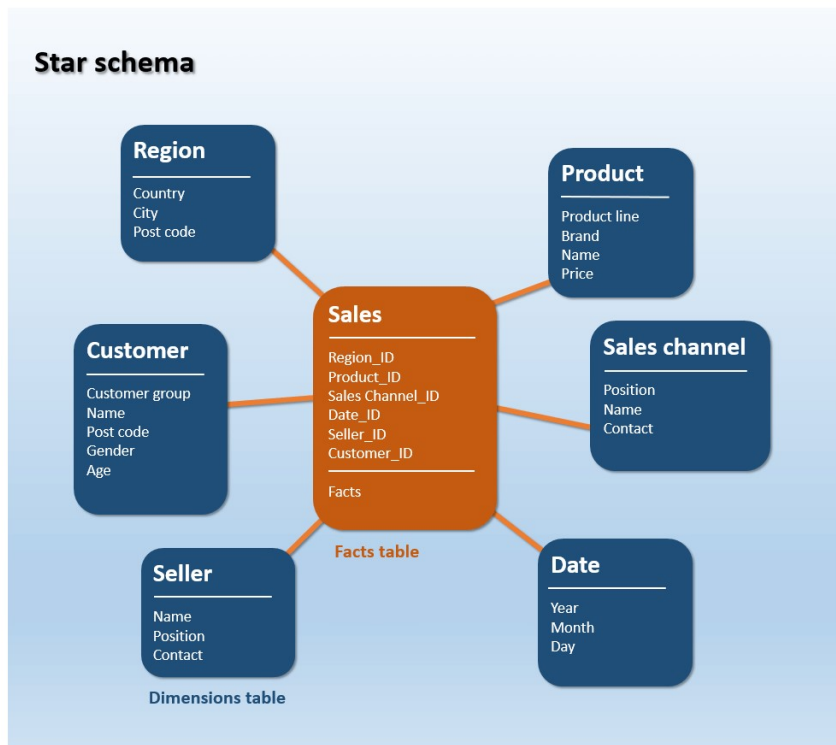


Figura 4: Modelo em Estrela (1&1 IONOS, 2021)

### 2.2.8.2 Modelo em Floco de Neve

No modelo em floco de neve, ou *snowflake*, a tabela de factos está igualmente associada a dimensões, só que neste caso pelo menos uma das dimensões encontra-se na forma normalizada e apresenta uma hierarquia dimensional. Ou seja, uma das dimensões encontra-se ligada a outra dimensão, criando assim uma hierarquia, como se pode verificar na Figura 5. Este tipo de abordagem é utilizado quando uma dimensão apresenta vários atributos.

Apesar deste modelo ser mais fácil de ser projetado e desenvolvido, é mais complicado para realizar análises o que normalmente leva à procura de outro tipo de solução (Javatpoint, 2021).

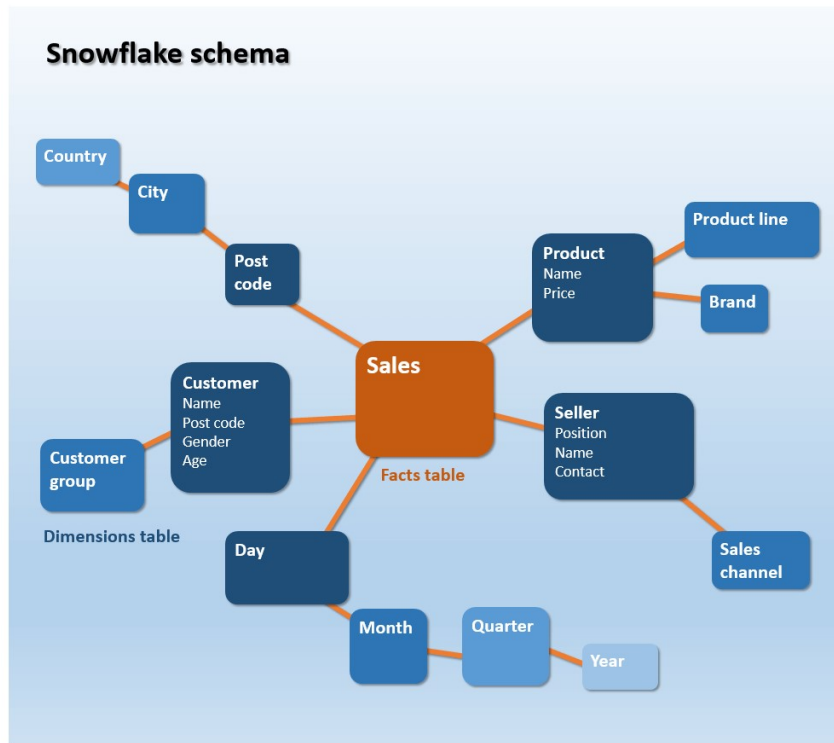


Figura 5: Modelo em Floco de Neve (1&1 IONOS, 2021)

### 2.2.9 Modelo Multidimensional

O modelo multidimensional consiste numa estrutura que combina diferentes dimensões com várias medidas. Este tipo de modelo é representado numa base de dados multidimensional designada por cubo de processamento analítico on-line (OLAP – Online Analytical Processing). Este tipo de base de dados agrega dados com diferentes níveis de hierarquias (Javatpoint, 2021). A Figura 6 ilustra o modelo multidimensional.

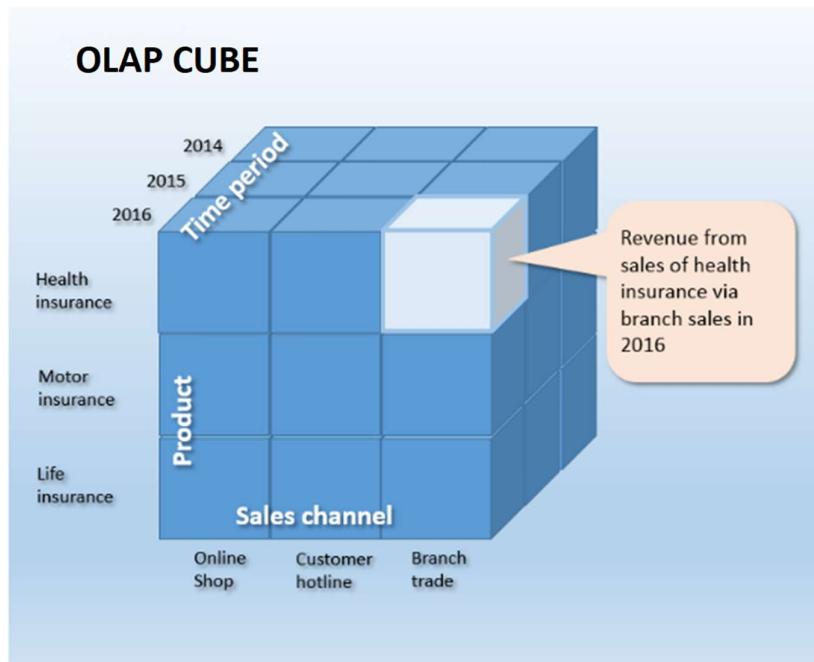


Figura 6: Modelo Multidimensional (1&1 IONOS, 2021)

### 2.2.9.1 Cubo MOLAP

Um cubo MOLAP (*Multidimensional Online Analytical Processing*) é uma estrutura de dados que elimina as limitações das bases de dados relacionais, fornecendo uma análise rápida de dados. Os cubos MOLAP podem apresentar e somar grandes quantidades de dados, enquanto fornecem aos utilizadores acesso pesquisável a qualquer ponto dos dados para que os dados possam ser agregados, segmentados e repartidos conforme necessário para lidar com uma grande variedade de perguntas por parte do utilizador (Microsoft, 2021).

### 2.2.9.2 Cubo ROLAP

O ROLAP é um modelo Relational Online Analytical Processing, no qual os dados são armazenados como na base de dados relacional, ou seja, linhas e colunas no armazém de dados. Num cubo ROLAP, os dados são apresentados ao utilizador na forma multidimensional. Para apresentar os dados, de forma multidimensional, é criada uma camada semântica de metadados que faz o mapeamento da dimensão para as tabelas relacionais. Os metadados também suportam a agregação dos dados. Para consultas complexas o cubo ROLAP obtém dados do armazém de dados principal e cria dinamicamente uma visualização multidimensional para o utilizador (Velho, S. V. M. 2011).

O cubo ROLAP armazena os dados em bases de dados relacionais, enquanto o MOLAP armazena em bases de dados multidimensionais, sendo esta a principal diferença entre os dois.

## 2.3 Ferramentas de Business Intelligence

Independentemente da dimensão de uma organização, a utilização de ferramentas de BI permite criar uma visão do negócio mais completa através da recolha, análise e conversão de dados.

Com o recurso a este tipo de ferramentas as organizações são capazes de definir e medir os seus Indicadores Chave de Performance (KPI) mais relevantes e dessa forma delimitar estratégias para alcançar o sucesso.

A escolha do software é um aspeto fundamental, isto porque é a partir do mesmo que as organizações adquirem conhecimento sobre o seu crescimento, observam as tendências e comportamentos dos seus clientes. De certa forma, este tipo de ferramenta serve de apoio à tomada de decisão.

As necessidades das organizações são diferentes, assim como as características de cada software de *Business Intelligence*, posto isto é indispensável perceber as características de cada software, estabelecendo uma comparação das vantagens e desvantagens e aquilo que elas podem trazer à organização.

A comparação das ferramentas pode ser feita mediante a definição de vários critérios. Para esta dissertação foi tido em conta o “Quadrante Mágico” 2021 da Gartner (Gartner, 2021), representado na Figura 7, para a definição dos critérios a comparar.



Figura 7: Quadrante Mágico relativo a ferramentas de integração de dados (Gartner, 2021)

A história da Gartner teve início em 1979, criada por Gideon Gartner (1935-2020), um empreendedor e filantropo, conhecido por ser um patriarca da indústria de TI. O objetivo da Gartner sempre foi trazer o conhecimento da tecnologia disruptiva para as empresas que querem abraçar a inovação, e que nela observam, a oportunidade de um progresso sustentável. Ocupa assim uma posição de destaque na realização de estudos avançados de mercado, que apresentam mundialmente, as tendências tecnológicas para o futuro próximo (Navita, 2021).

O “Quadrante Mágico” da Gartner é constituído pelos seguintes pontos:

1. **Líderes:** As principais companhias do mercado, são aquelas que já atuam com o que há de mais inovador e ditam as tendências;
2. **Desafiadores:** São aquelas que ainda não são consideradas líderes, entretanto, adotam o que é novo e possuem facilidade em lidar com a inovação;
3. **Visionários:** Atuam diretamente com pesquisa e desenvolvimento, mas ainda não possuem grandes mudanças nas suas infraestruturas;
4. **Concorrentes de nicho:** Empresas que fazem o básico do que se dispõem a fazer, sem procurarem ir além, com tecnologias disruptivas. Têm o foco num determinado segmento de mercado.

### 2.3.1 Ferramentas de ETL

A Gartner define a integração de dados como a disciplina que compreende os padrões de arquitetura, ferramentas e metodologias que permitem às organizações acederem, harmonizar, transformar, processar e mover dados abrangendo vários pontos de origem/destino e em qualquer infraestrutura. Posto isto, a Gartner define alguns critérios a considerar no momento de escolher uma ferramenta de integração de dados (Gartner 2021). Seguem alguns exemplos:

- **Engenharia de dados:** Construção, gestão e operacionalização de pipelines de dados em apoio a diversos pedidos de análise e ciência de dados (por exemplo, *data warehouse* lógico, *analytics/business intelligence* e *machine learning*, entre outros casos de uso analítico) seguindo padrões de design, ferramentas e metodologias definidas.
- **Suporte para entrega otimizada de análises:** Capacidade de fornecer acesso a várias fontes heterogêneas de e, em seguida, extrair dados dessas fontes para um repositório de destino otimizado para fornecer dados integrados para análise.
- **Migração para a nuvem (cloud):** Possuir ferramentas de integração de dados para apoiar a migração e modernização de cargas de trabalho de dados e análises para infraestrutura na nuvem.
- **Integração de dados operacionais:** Suporte a integração de dados operacionais/transacionais, como gestão de dados e aquisição e partilha de dados interempresas (incluindo a capacidade de criar *data hubs* para integração quando necessário).

- **Partilha de dados interempresa:** As organizações são cada vez mais obrigadas a partilhar dados com parceiros comerciais externos (clientes, fornecedores, parceiros de negócios e outros).
- **Consistência de dados entre aplicações operacionais:** As ferramentas de integração de dados devem ser capazes de garantir a consistência ao nível da base de dados entre aplicações, tanto internamente quanto interempresa, e de forma bidirecional ou unidirecional.

Utilizando o Quadrante Mágico como referência, foram selecionadas três ferramentas de ETL que figuram no quadrante “Líderes”, nomeadamente o Informatica PowerCenter, o IBM DataStage e o Microsoft SQL Integration Services. A referência ao Pentaho Data Integration deve-se ao fato de ser uma ferramenta *open source*. Nas subsecções seguintes são descritas as ferramentas de ETL selecionadas.

#### **2.3.1.1 Informatica PowerCenter**

O Informatica PowerCenter (Informatica, 2022) é na sua essência um software cuja função é a extração de dados de diversos sistemas, limpeza de dados e transformação conforme as regras de negócio, para posteriormente serem carregados num *Data Mart* ou num *Data Warehouse*.

Uma das características do PowerCenter é a sua portabilidade, não existindo a necessidade de uma base de dados específica para que a ferramenta funcione pois é capaz de aceder a praticamente qualquer tipo de arquitetura, inclusive a mainframes.

A estrutura do PowerCenter suporta processamento de dados em massa assim como em tempo real. A plataforma da ferramenta permite o acesso em tempo real à base de dados transacional para que os dados sejam atualizados com a velocidade e necessidade exigida pelo negócio (Bi4all, 2021). Na Figura 8 é possível visualizar a interface do PowerCenter.

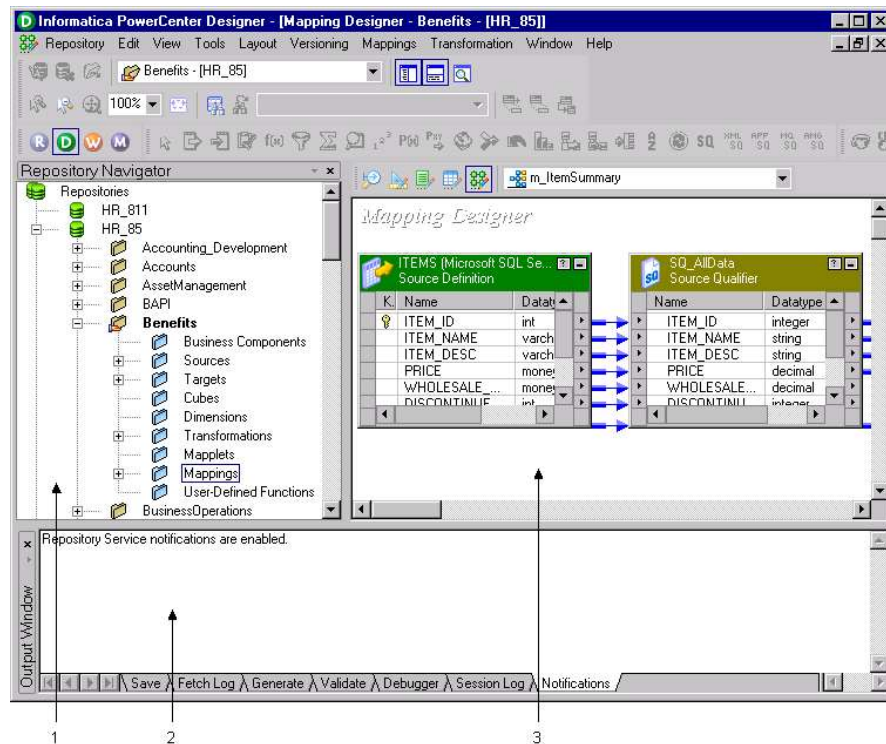


Figura 8: Informatica PowerCenter Designer (Informatica, 2022)

### 2.3.1.2 IBM DataStage

O IBM DataStage é uma ferramenta que integra dados de múltiplos sistemas com recurso a uma estrutura paralela de alto desempenho que suporta a gestão de metadados e a conectividade corporativa. Com uma plataforma escalável que fornece uma integração mais flexível de todos os tipos de dados, incluindo *Big Data* (baseado em Hadoop) ou Stream (baseado em fluxo), em plataformas distribuídas e mainframe (Cetax, 2022).

O Data Stage fornece os seguintes recursos e benefícios:

- **Plataforma ETL escalável:** suporta a extração, integração e transformação de grandes volumes de dados, desde as estruturas de dados mais simples às mais complexas.
- **Suporte a Big Data<sup>1</sup> e Hadoop<sup>2</sup>:** permite acesso direto a *Big Data* num sistema de ficheiros distribuído e permite de uma forma mais eficiente acesso a novas origens de dados fornecendo suporte JSON<sup>3</sup> e JDBC<sup>4</sup>.

<sup>1</sup> Big Data são dados com maior variedade que chegam em volumes crescentes e com velocidade cada vez maior.

<sup>2</sup> Apache Hadoop é uma coleção de utilitários de software de código aberto.

<sup>3</sup> JSON (JavaScript Object Notation) é um formato de troca de dados.

<sup>4</sup> Java Database Connectivity (JDBC) é uma interface de programação (API) para a linguagem de programação Java.

- **Integração de dados em tempo real:** assim como conectividade entre origens de dados e aplicações.
- **Gestão de regras de carga de trabalho:** ajuda otimizar a utilização de hardware e priorizar tarefas essenciais.
- **Facilidade de utilização:** ajuda a melhorar a velocidade, flexibilidade e efetividade na construção, implementação, atualização e gestão da infraestrutura de integração de dados.

Segue-se na Figura 9 a interface do DataStage.

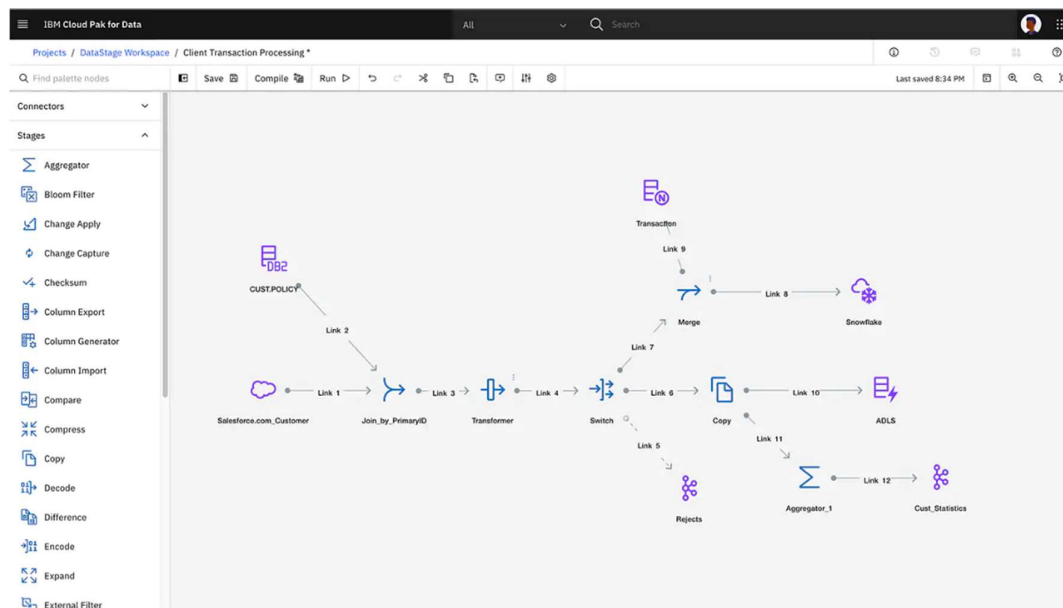


Figura 9: IBM DataStage (IBM, 2022)

### 2.3.1.3 Microsoft SQL Server Integration Services (SSIS)

O Microsoft SQL Server Integration Services é uma plataforma para a construção de soluções de integração de dados e transformações de dados a nível empresarial. Tem a capacidade de resolver problemas empresariais complexos recorrendo à cópia de dados, carregando os mesmos para armazéns de dados, permitindo também a limpeza e mineração de dados.

As capacidades mais relevantes do SSIS são a sua interface gráfica de fácil utilização, visível na Figura 10, a rapidez de implementação e a facilidade de integração com outras ferramentas da Microsoft, como por exemplo o SQL Server e o serviço Azure na *cloud* (Microsoft, 2022).

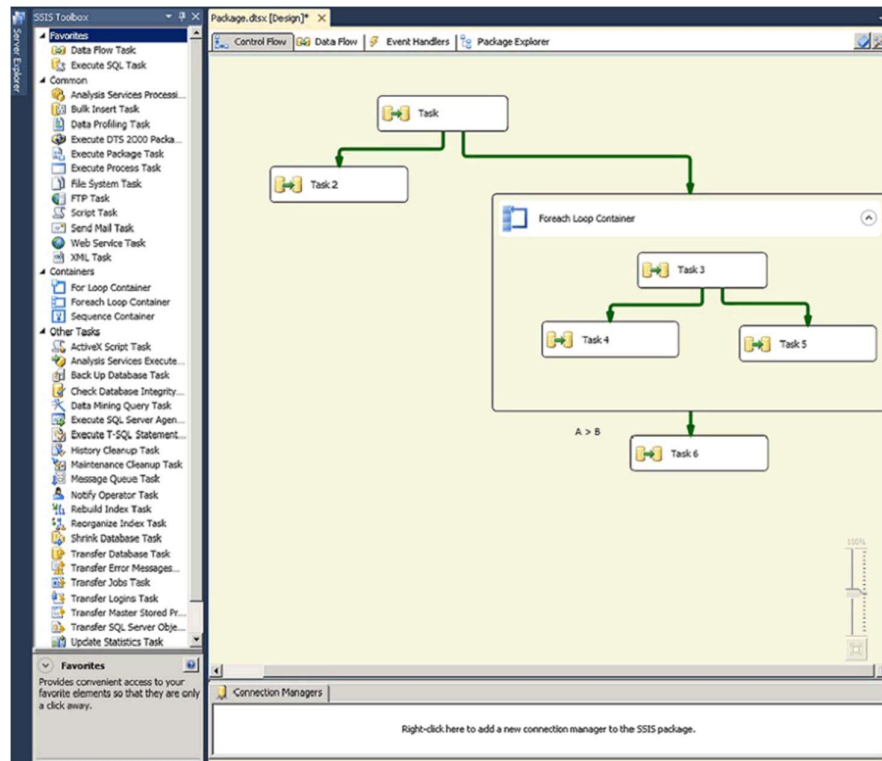


Figura 10: Microsoft SSIS (Microsoft, 2022a)

#### 2.3.1.4 Pentaho Data Integration

Também conhecido como Kettle (Hitachi Vantara, 2022) e uma das mais conhecidas ferramentas ETL *open source*, o PDI (Pentaho Data Intagracion) é um motor de integração de dados, escrito em Java e com um interface gráfico chamado Spoon, visível na Figura 11. O interface gráfico confere uma das grandes vantagens duma ferramenta como o PDI, programação visual. O PDI permite ao utilizador fazer transformações aos dados assim como criar jobs de integração (Hitachi Vantara, 2022).

As características principais do PDI são:

- Possibilidade de coordenar e combinar transformações de dados, incluindo a geração de notificações e alertas;
- Ferramentas de *debug* para testes em tarefas do pipeline;
- Apresentação de vistas para a modelação e visualização de dados em tempo de processamento dos dados;

Na Figura 11 tem-se a interface do PDI.

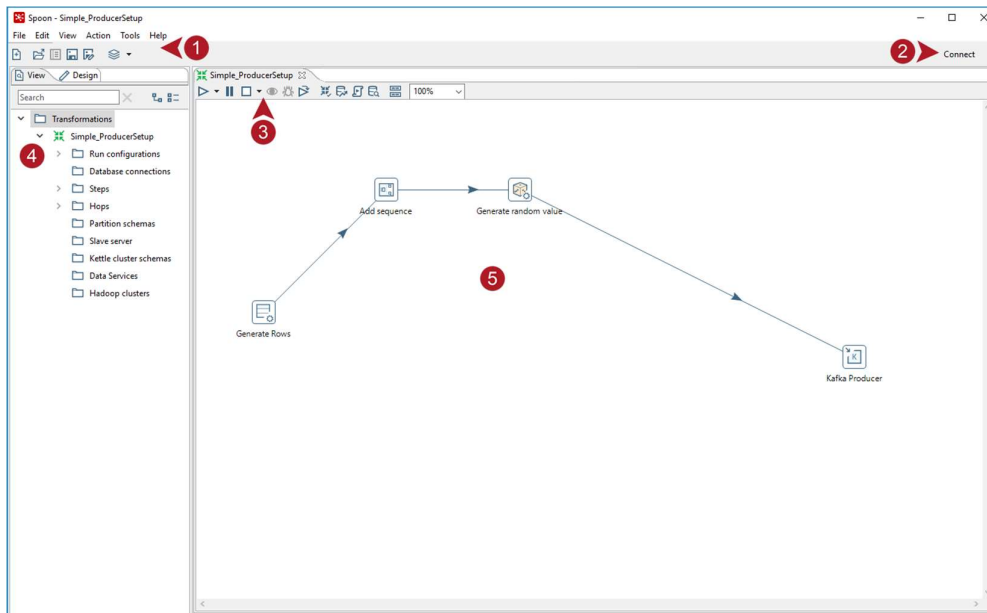


Figura 11: Hitachi Vantara PDI (Hitachi Vantara, 2022)

### 2.3.1.5 Análise Comparativa

Na tabela 1 é feita uma análise comparativa entre as ferramentas ETL anteriormente referidas.

Tabela 1: Comparação de ferramentas ETL

	Informatica PowerCenter	IBM DataStage	Microsoft SSIS	Pentaho PDI
Engenharia de dados	✓	✓	✓	✓
Suporte a várias fontes	✓	✓	✓	✓
Migração para a cloud	✓	✓	✓	
Integração de dados	✓	✓	✓	✓
Partilha de dados	✓	✓	✓	
Consistência de dados	✓	✓	✓	✓
Open source				✓

Na tabela anterior pode-se verificar que todas as ferramentas selecionadas, à exceção do Pentaho PDI, possuem as mesmas características quando aplicados os critérios propostos pela Gartner. Também é possível verificar que o Pentaho PDI é a única ferramenta *open source*, mas também a única que não cumpre todos os restantes critérios. Desta forma, qualquer uma das outras ferramentas de ETL pode ser eleita para o desenvolvimento desta dissertação.

Mesmo tendo em conta a análise anterior, considera-se que a escolha da ferramenta de ETL irá depender da escolha da ferramenta de análise de dados e dos valores de licenciamento de cada solução. Posto isto, só após a análise das ferramentas de análise é possível ter uma opinião mais clara e consciente acerca das ferramentas a utilizar na elaboração desta dissertação.

### 2.3.2 Ferramentas de análise de dados

As plataformas de análise de dados já não são diferenciadas pelas suas capacidades de visualização de dados, que agora são comuns. Todas as plataformas de análise podem construir *dashboards* de indicadores de desempenho interativos (KPI) utilizando gráficos comuns (barra/coluna, linha/área, dispersão, tarte e mapas geográficos) e baseando-se numa vasta gama de fontes de dados. A diferenciação mudou para a forma como as plataformas suportam a análise aumentada. A análise aumentada utiliza *machine learning* (ML) e a inteligência artificial (IA) na preparação de dados, geração de insights<sup>5</sup> e explicação de *insight* para ajudar os empresários e analistas de dados a explorar e analisar os dados de forma mais eficaz do que podiam manualmente. Em vez de ser um recurso discreto, a análise aumentada é agora ligada através de plataformas à medida que o ML é aplicado no *workflow* de dados (Gartner, 2022).

De acordo com a Gartner existem 12 critérios para avaliar ferramentas de análise de dados:

- **Segurança:** Capacidades que permitem a segurança da plataforma, administração de utilizadores, auditoria de acesso e autenticação.
- **Capacidade de gestão:** recursos que monitorizam a utilização da plataforma de BI e gerem como as informações são partilhadas (e por quem).
- **Análise na nuvem:** a capacidade de dar suporte à criação, implementação e gestão de análises na nuvem, com base em dados armazenados na nuvem e no local.
- **Ligação a fontes de dados:** recursos que permitem que os utilizadores se liguem, consultem e explorem os dados, otimizando o desempenho.
- **Preparação de dados:** Suporte para arrastar e soltar, combinação orientada para o utilizador de dados e criação de modelos analíticos (como medidas, conjuntos, grupos e hierarquias definidos pelo utilizador).
- **Catálogo:** A capacidade de gerar e organizar automaticamente um catálogo pesquisável de conteúdo analítico.
- **Insights automatizados:** um atributo central da análise aumentada é a aplicação de técnicas de ML de forma a gerar automaticamente conhecimento para os utilizadores

---

<sup>5</sup> Insight - compreensão, percepção ou revelação repentina. Capacidade de introspeção ou de autoconhecimento

finais (por exemplo, identificando os atributos mais importantes num conjunto de dados).

- **Visualização de dados:** Suporte para *dashboards* altamente interativos e exploração de dados através da manipulação de imagens e gráficos.
- **Storytelling dos dados:** a capacidade de combinar visualização interativa de dados com técnicas narrativas para empacotar e fornecer conteúdo analítico de forma atraente e de fácil compreensão para suportar a tomada decisão.
- **Consulta de linguagem natural (NLQ):** permite que os utilizadores façam perguntas e consultem dados e conteúdo analítico utilizando termos inseridos numa caixa de pesquisa ou falados.
- **Geração de linguagem natural (NLG):** A criação automática de descrições linguisticamente ricas de respostas, dados e conteúdo analítico. Dentro do contexto de análise, à medida que o utilizador interage com os dados, a narrativa muda dinamicamente para explicar as principais descobertas ou o significado de gráficos ou painéis.
- **Relatórios:** A capacidade de criar e distribuir (ou “explodir”) relatórios de várias páginas com um layout adequado para os utilizadores de forma programada.

Na Figura 12 apresenta-se o “Quadrante Mágico” referente às ferramentas de análise de dados.



Figura 12: Quadrante Mágico relativo a ferramentas de análise de dados (Gartner, 2022)

Utilizando o Quadrante Mágico como referência, foram selecionadas sete ferramentas de BI, três do quadrante “Líderes”, uma do quadrante “Desafiadores”, duas que figuram no quadrante “Visionários” e uma que não se encontra representada, mas por ser *open source* foi selecionada. Seguem-se nas subsecções seguintes as ferramentas que foram tidas em conta para comparação.

### 2.3.2.1 Microsoft Power BI

O Microsoft Power BI é um conjunto de serviços de software que transforma dados não relacionados em dados coesos e interativos. Capaz de extrair dados de variadas fontes de dados, o Power BI, pode ser utilizado como uma ferramenta pessoal de *report* e visualização de dados ou como mecanismo de análise e decisão a nível corporativo.

É um software de utilização intuitiva, capaz de responder a vários problemas de análise de dados, como modelação extensiva e análises em tempo real (Microsoft, 2022).

Na Figura 13, estão representados os 3 componentes do Power BI: *desktop*, *service* e *mobile*. Estes elementos são utilizados para criar, partilhar e visualizar relatórios de negócio.

O Microsoft Power BI figura no quadrante “Líderes”.

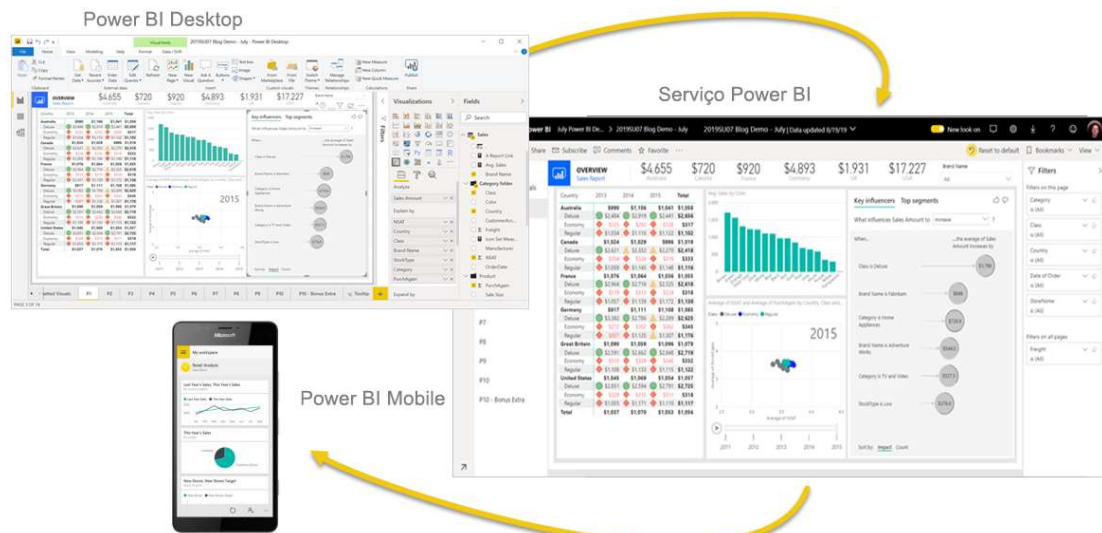


Figura 13: Principais elementos do Microsoft Power BI (Microsoft, 2022b)

### 2.3.2.2 Tableau

O Tableau é uma plataforma self-service de *Business Intelligence* que utiliza tecnologia inovadora que permite que qualquer pessoa consiga simplesmente arrastando e soltando

dados com o intuito de analisá-los. Dessa forma, rapidamente os dados são acedidos e visualizados, podendo ser criadas folhas de cálculo interativas através de alguns cliques. Oferece uma experiência de utilização simples, baseada em exploração visual, não recorrendo a assistentes e *scripts* que iriam dificultar o uso da plataforma. Sendo assim, qualquer utilizador que esteja ambientado com o Microsoft Excel, por exemplo, consegue utilizar o Tableau (Tableau, 2022). Na Figura 14 tem-se a interface do Tableau.

O Tableau figura no quadrante “Líderes”.

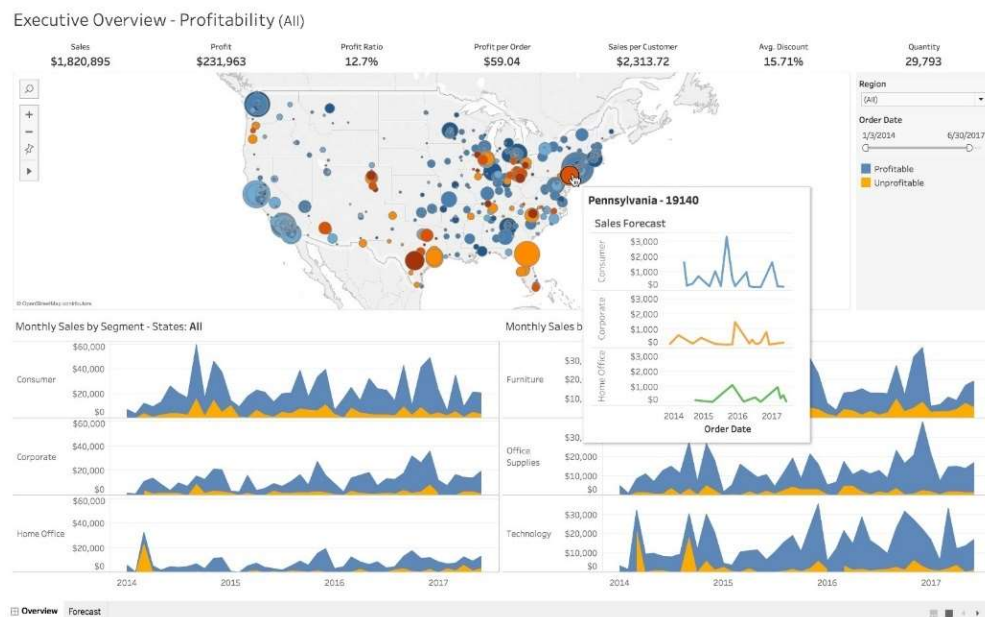


Figura 14: Tableau (Tableau, 2022)

### 2.3.2.3 Qlik

O Qlik é uma plataforma de análise proprietária de alto desempenho que representa como os dados devem ficar relacionados. Disponibiliza visualização self-service, análise orientada e incorporada. Usa indexação na memória para manter o relacionamento de dados e serve vários produtos como o Qlik Sense.

O Qlik Sense é uma plataforma proprietária gratuita usada para visualização interativa e decisões úteis. A plataforma disponibiliza três versões principais, como o Qlik Sense Desktop que é uma aplicação gratuita para criar visualizações e relatórios personalizados, o Qlik Sense Enterprise para fins empresariais e o Qlik Sense Cloud é uma aplicação baseado na cloud (QlikTech International AB, 2022). Na Figura 15 tem-se a interface do Qlik.

O Qlik figura no quadrante “Líderes”.

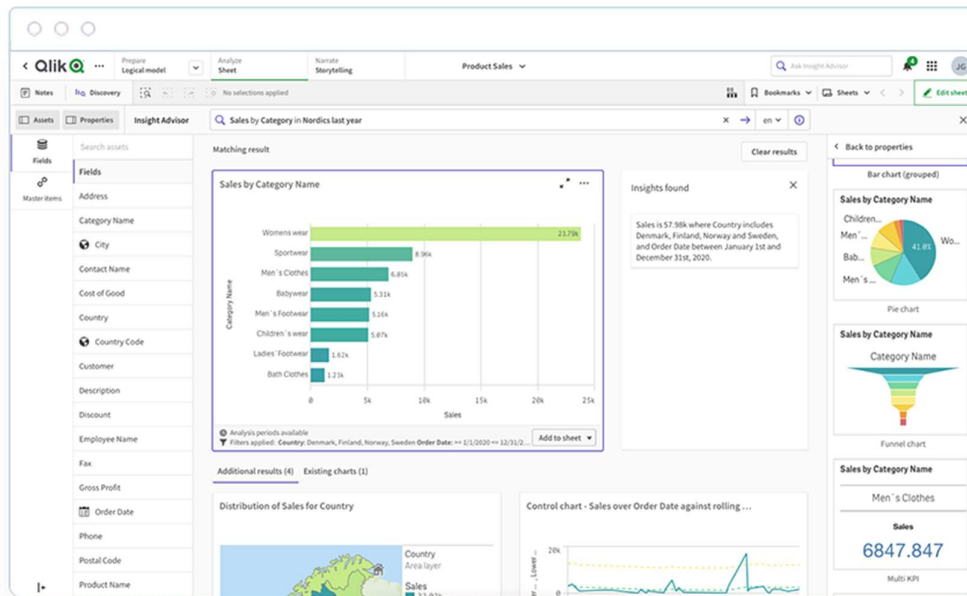


Figura 15: Qlik (QlikTech International 2022)

### 2.3.2.4 MicroStrategy

A plataforma MicroStrategy, contém um conjunto de produtos analíticos para interligação, visualização, relatórios e análises de dados avançadas, é complementada por produtos móveis, na nuvem, incorporados e de análise de identidade. Um dos seus pontos fortes é o HyperIntelligence, uma aplicação intuitiva, que utiliza um gráfico semântico para sobrepor e identificar dinamicamente *insights* predefinidos em aplicações existentes. É funcionalmente capaz numa ampla variedade de casos de uso e os seus recursos de consulta direta são adequados para utilização em *data warehouses* na nuvem (Gartner, 2022). Na Figura 16 tem-se a interface do MicroStrategy.

O MicroStrategy figura no quadrante “Desafiadores”.

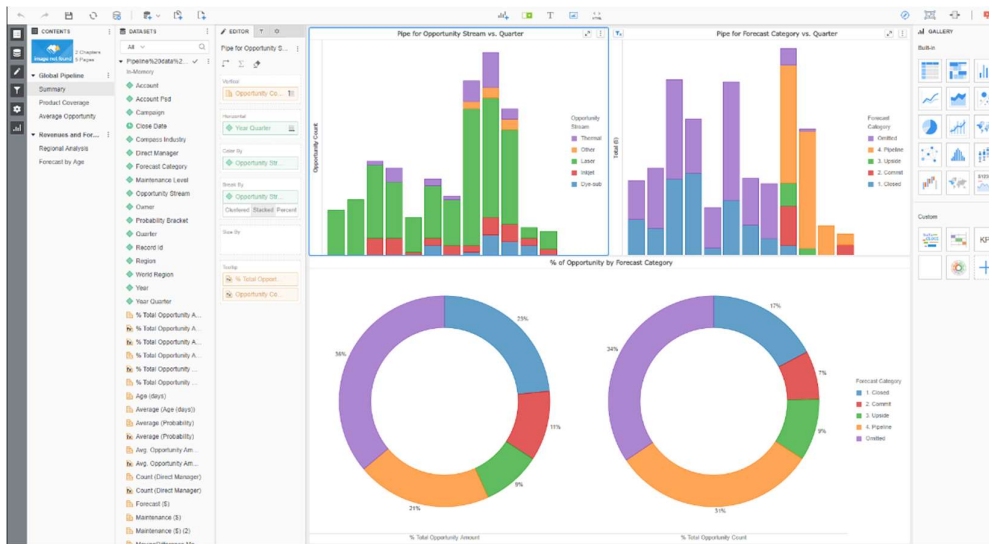


Figura 16: MicroStrategy Workstation (MicroStrategy, 2022)

### 2.3.2.5 Oracle (OBIEE + Endeca)

OBIEE é uma ferramenta *open-source* de *Business Intelligence* proprietária que fornece relatórios, análise de consulta ad hoc, processamento analítico online entre outras funcionalidades. Todas as análises, definições e cálculos de negócio são enquadrados num Modelo de Informações Corporativas Comum. O utilizador pode aceder a informação de várias formas, juntamente com o local de trabalho colaborativo e por meio de aplicações. Conhecido por fornecer aos utilizadores conhecimento detalhado e preciso. O Oracle Endeca Information Discovery é apresentado como uma solução completa para a descoberta ágil de dados. Ajuda a gerir o equilíbrio do negócio com a colaboração de TI. Fornece acesso rápido à análise de dados tradicionais e a dados não tradicionais. Mantém os investimentos empresariais atuais, reduz o consumo de tempo e ajuda a tomar decisões de negócio precisas (Oracle, 2022). Na Figura 17 tem-se a interface do Oracle Business Intelligence.

A Oracle figura no quadrante “Visionários”.

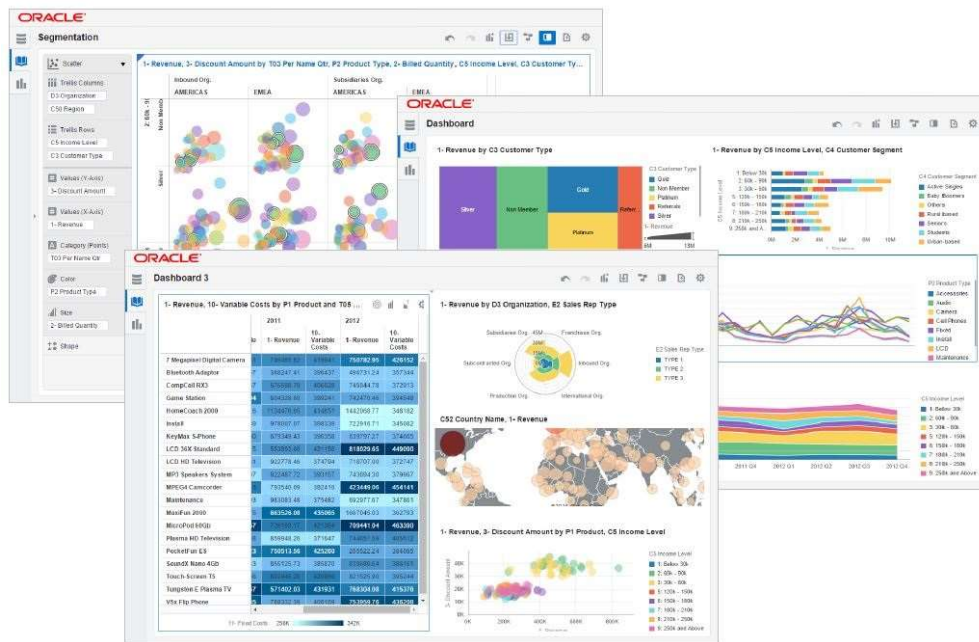


Figura 17: Oracle Business Intelligence (Oracle, 2022)

### 2.3.2.6 SAS

O SAS (SAS Institute Inc., 2022) é uma ferramenta proprietária conhecida por disponibilizar ao utilizador informação assertiva no momento certo. Fornece uma ligação de alta qualidade entre várias fontes de dados e o utilizador final. Ajuda a proteger dados comerciais com identificação e resolução de problemas de segurança. Suporta meta dados centralizados, governabilidade, escalabilidade e visualização de dados. Fornece análise em tempo real e permite aos utilizadores criarem os seus próprios relatórios (SAS Institute Inc., 2022). A Figura 18 ilustra a interface do SAS.

O SAS figura no quadrante “Visionários”.

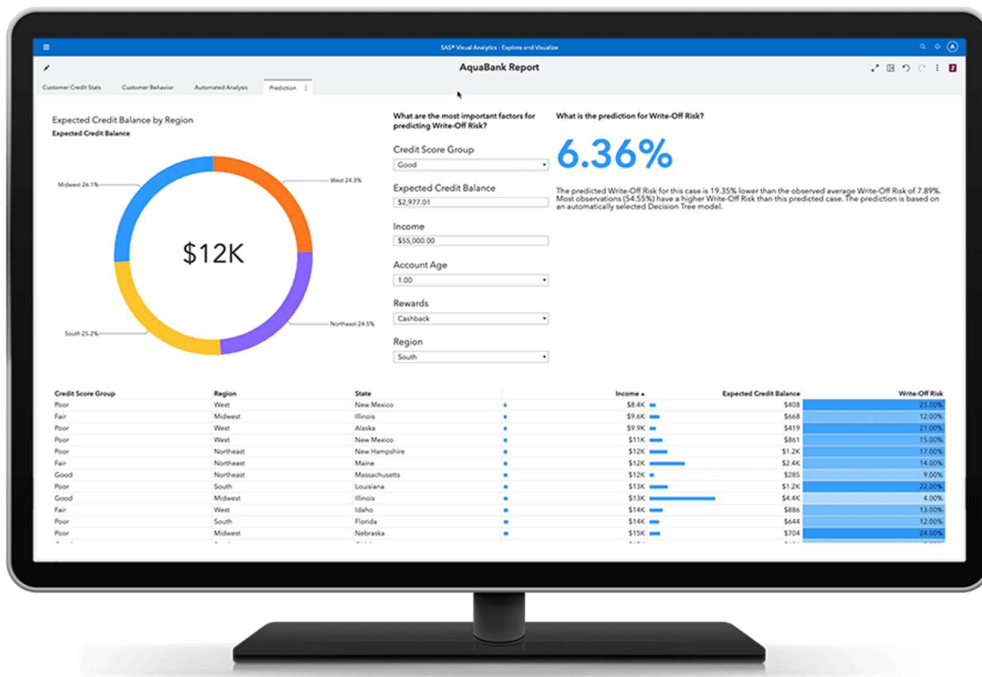


Figura 18: Interface SAS (SAS Institute Inc., 2022)

### 2.3.2.7 Pentaho

O Pentaho é um software para criar soluções de BI, com várias características interessantes. É baseado em tecnologia *open-source*, têm uma edição Community que é gratuita e uma edição Enterprise que tem custos. Suporta dados estruturados e não estruturados de diversas origens. É uma plataforma de BI completa constituída por vários componentes, que podem ser usados de forma independente ou integrada, conferindo-lhe uma grande flexibilidade. Pode-se dizer que é uma plataforma *end-to-end*, pois inclui componentes para todas as fases de um projeto típico de BI. Possui uma vasta comunidade e bastante ativa, muitos dos componentes atualmente existentes na plataforma nasceram de iniciativas da comunidade (Hitachi Vantara, 2022). Na Figura 19 tem-se a interface do Pentaho.

O Pentaho não se encontra representado no “Quadrante Mágico” da Gartner.

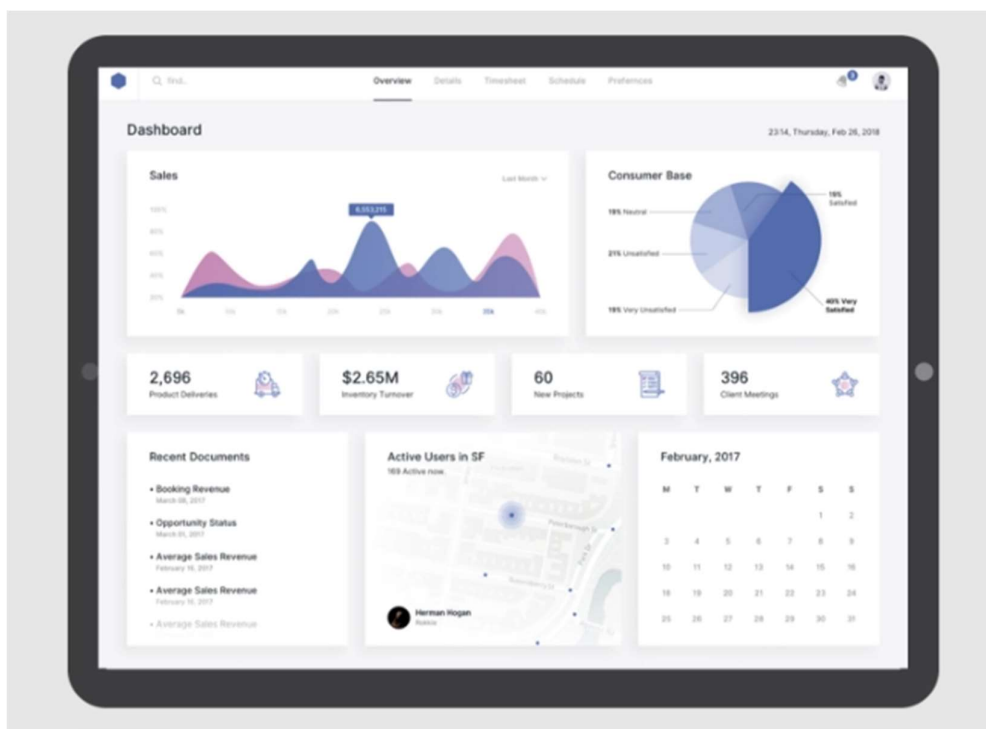


Figura 19: Interface Pentaho (Hitachi Vantara, 2022)

### 2.3.2.8 Análise Comparativa

Tendo em conta os critérios da Gartner, para esta dissertação serão utilizados os seguintes 9 critérios:

- Administração e Segurança
- *Cloud BI*
- Conectividade a fontes de dados
- ETL e armazenamento de dados
- *Reports e Dashboards*
- Descoberta de conhecimento nos dados
- Incorporação de Conteúdo
- Partilha de Conteúdo
- Facilidade de utilização

A Tabela 2 apresenta a comparação entre as ferramentas de análise de dados, utilizando os critérios acima apresentados.

Tabela 2: Comparativo entre ferramentas de análise de dados

	Power BI	Tableau	Qlik	MicroStrategy	Oracle	SAS	Pentaho
Administração e Segurança	✓	✓	✓	✓	✓	✓	✓
Cloud BI	✓		✓	✓	✓	✓	
Conetividade a fontes de dados	✓	✓	✓	✓	✓	✓	✓
ETL e armazenamento de dados	✓	✓	✓	✓	✓	✓	✓
Reports e Dashboards	✓	✓	✓	✓	✓	✓	✓
Descoberta de conhecimentos nos dados	✓		✓		✓	✓	
Incorporação de conteúdo	✓	✓	✓	✓	✓	✓	✓
Partilha de conteúdo	✓	✓	✓	✓	✓	✓	
Facilidade de utilização	✓	✓	✓	✓		✓	

Como é possível verificar existem várias semelhanças nas características das ferramentas estudadas. De salientar que embora o Microsoft Power BI tenha capacidade de ETL, a mesma pode ser complementada com a utilização do Microsoft SQL Integration Services (SSIS). Analisando a tabela verifica-se que apenas duas ferramentas cumprem os 9 critérios, O Power BI e o SAS.

## 3. Análise de Valor

Neste capítulo o objetivo passa por analisar e aplicar técnicas para se perceber se a ideia a desenvolver tem potencial para acrescentar valor ao negócio onde se insere. A partir da análise de valor é possível verificar se um caso de estudo pode ter ou não ter sucesso, justificando em vários casos a existência de uma determinada organização.

As organizações devem a sua existência à capacidade de criar valor. Nos casos em que não exista valor acrescentado, então não existe qualquer motivo para a existência da mesma.

### 3.1 Processo de Inovação

De uma maneira geral todas as organizações devem constantemente procurar inovar, ou criar produtos de forma que estes tragam à sociedade valor acrescentado. A inovação é uma construção complexa e que se sobrepõe a outros conceitos predominantes, como tecnologia, criatividade e mudança, sendo transversal a diversas áreas, incluindo negócios, economia, engenharia e administração pública. As condições externas e internas induzem a inovação e esta permite alterar a conduta organizacional e os resultados da organização (Lam, A. 2009).

O modelo de conceito de desenvolvimento apresentado por Peter Koen surge a partir do processo de inovação, essencial na inovação e criação de novos produtos e o mesmo encontra-se dividido em três fases:

- Fuzzy Front-End (FFE)
- Processo de desenvolvimento de novos Produtos (NPD)
- Comercialização

A FFE é a componente principal do processo de inovação e esta permite melhorar o mesmo, “aumentando valor, quantidade e probabilidade de sucesso no desenvolvimento e na comercialização de produtos” (Belliveau, P. and Griffin, A. 2002).

Porém, devido à inexistência de uma linguagem ou definição comum para os elementos-chave do FFE, surge o modelo New Concept Development (NCD), desenvolvido para fornecer uma linguagem comum.

Segundo Koen (Belliveau, P. and Griffin, A. 2002), o modelo NCD consiste em três componentes, tal como se verifica na Figura 20:

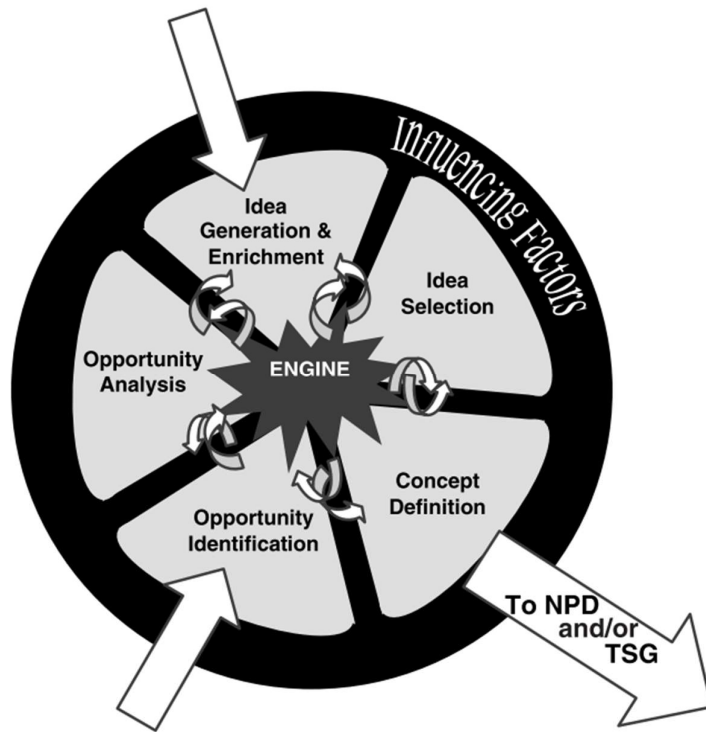


Figura 20: Modelo de New Concept Development (Belliveau, P. and Griffin, A. 2002)

- A parte do motor ou alvo é constituído pela liderança, a cultura e a estratégia de negócios que impulsiona os cinco elementos-chave que são controláveis pela organização.
- A área no raio interno define os cinco elementos de atividade controláveis (identificação de oportunidade, análise de oportunidade, geração e enriquecimento de ideias, seleção de ideias e definição de conceito) do FFE.
- Os fatores influenciadores consistem em capacidades organizacionais, no mundo exterior (canais de distribuição, lei, política governamental, clientes, concorrentes e clima político e económico), e nas ciências habilitantes (internas e externas) que podem estar envolvidas.

Estes fatores afetam todo o processo de inovação até a comercialização, sendo relativamente incontroláveis pela organização.

O modelo NCD, nomeadamente os cinco elementos de atividade controláveis são aplicados ao contexto do problema em questão, sendo que os resultados obtidos são apresentados nas secções seguintes.

### 3.1.1 Identificação da oportunidade

A fase de identificação de oportunidades é um dos pontos de entrada do modelo NCD e tem como objetivo identificar oportunidades exploráveis, sejam elas comerciais, no sentido de alocar corretamente os recursos em novos segmentos de mercado, otimização de processos ou aumento de eficiência dos métodos existentes.

Esta atividade é normalmente despoletada por objetivos comerciais, por exemplo, para dar resposta a uma nova concorrência, a possibilidade de inovar e obter vantagem competitiva, melhorar as operações e processos internos aumentando a eficiência ou simplesmente redução de custos.

Para executar esta atividade existem diferentes métodos, ferramentas e técnicas que possibilitam a execução, nomeadamente (Koen, P. A. et al. 2002):

- **Roadmapping:** captura as forças motrizes do negócio de forma gráfica, a fim de melhorar a comunicação e a perceção.
- **Análise das tendências tecnológicas:** dedica-se à pesquisa e análise de tendências tecnológicas que possam ser relevantes e tornarem-se em oportunidades competitivas para o projeto.
- **Análise de tendências do cliente:** prende-se com a pesquisa e análise das tendências numa vertente mais direcionada para os clientes.
- **Análise de inteligência competitiva:** elemento essencial na atividade de identificação de oportunidades, uma vez que permite obter, analisar e comunicar informação de tendências competitivas que ocorrem fora da organização, possibilitando a definição de uma estratégia de negócio.
- **Pesquisa de mercado:** pesquisa do mercado em que o projeto se insere.
- **Planeamento de cenário:** proporciona uma abordagem disciplinada para imaginar e preparar o futuro, estimulando as decisões que de outra forma seriam ignoradas. O objetivo será criar diferentes perspetivas do futuro e assim obter conhecimento daquele que poderá ser o ambiente futuro que envolve o projeto.

No contexto do problema a oportunidade surge da necessidade de melhorar o processo de criação de análises sobre os dados do negócio, nomeadamente obtenção de indicadores relativamente ao desempenho da equipa comercial.

### 3.1.2 Análise da oportunidade

Nesta fase, é onde uma organização identifica as oportunidades que pretende realizar, orientadas aos objetivos do negócio.

Os métodos para análise de oportunidades segundo (Koen, P. A. et al. 2002) são:

- **Enquadramento estratégico:** determina como uma oportunidade se enquadra no mercado, forças tecnológicas, lacunas e ameaças da empresa.
- **Avaliação do segmento de mercado:** descrição detalhada do segmento de mercado, mostrando por que razão representa uma grande oportunidade. A análise da dimensão do mercado, as taxas de crescimento e a quota de mercado dos concorrentes são determinadas. São também avaliados fatores económicos, culturais, demográficos, tecnológicos e regulamentares que impactem o segmento de mercado.
- **Análise da concorrência:** determina quem são os principais concorrentes no segmento de mercado identificado. Determina o tipo de novos produtos necessários para obter vantagens competitivas. Avalia as estratégias e capacidades dos concorrentes e o estatuto das patentes recentes nesta área.
- **Avaliação do cliente:** determina quais as principais necessidades do cliente que não estão a ser satisfeitas pelos produtos atuais.

Com a implementação deste projeto, a ferramenta proposta será capaz de providenciar aos colaboradores da organização uma análise clara, coerente e confiável dos indicadores de performance associados ao negócio, oferecendo-lhe as formas mais adequadas para a análise dos dados e conseqüentemente das suas oportunidades de negócio.

### 3.1.3 Geração de ideias

A geração de ideias e o seu enriquecimento prende-se com o nascimento, desenvolvimento e maturação de uma ideia em particular, sendo um dos pontos de entrada do modelo NCD, à semelhança da identificação de oportunidades (Koen, P. A. et al. 2002). As ideias podem passar por várias iterações e sofrer múltiplas alterações à medida que são analisadas, estudadas e discutidas com os restantes elementos do modelo NCD (Koen, P. A. et al. 2002). É comum, nesta fase, haver contato direto com os clientes e utilizadores, e para a realizar podem ser utilizados processos formais (sessões de brainstorming) ou informais (fornecedor a oferecer um novo material ou cliente que faz um pedido incomum), no sentido de gerar ideias novas ou refinadas para uma dada oportunidade (Koen, P. A. et al. 2002). Normalmente esta fase passa pela execução de técnicas de brainstorming para gerar ideias, ou a aplicação de métodos de enriquecimento que utilizam a Theory of Inventive Problem Solving (TRIZ). A TRIZ é uma metodologia que melhora a criatividade levando os indivíduos a pensar além da sua própria experiência e a resolver os problemas utilizando soluções de outras áreas de ciência (Koen, P. A. et al. 2002).

Como forma de solucionar o problema encontrado, surge a ideia de implementar um novo sistema que utilize ferramentas de ETL e de BI para conseguir apresentar aos colaboradores da organização informação sobre o negócio de uma forma intuitiva.

### **3.1.4 Seleção da Ideia**

A seleção de ideias tem como objetivo selecionar as ideias que podem trazer maior valor a um determinado negócio. Ou seja, o problema, normalmente, não é ter novas ideias, é saber quais poderão ser mais relevantes para o projeto (Koen, P. A. et al. 2002). Esta seleção pode ser determinante para o progresso do projeto e para o seu sucesso, no entanto, não há forma de garantir que a ideia selecionada foi a melhor. A maioria das fases de seleção de ideias é constituída por diferentes iterações de atividades que poderão incluir outras passagens pela identificação de oportunidades, análise de oportunidades, geração de ideias e enriquecimento, normalmente com novos conhecimentos dos fatores influenciadores (Koen, P. A. et al. 2002). A fase de seleção de ideias é caracterizada pela utilização de métodos como o portfólio de metodologias baseado em múltiplos fatores, processo de seleção de ideias formal com feedback para aqueles que submeteram a ideia, e a utilização da teoria das opções, embora nenhuma delas possa ser utilizada para todos os tipos de seleção de ideias. Assim, as organizações devem determinar qual o melhor método a adotar considerando o tipo de seleção que pretendem utilizar (Koen, P. A. et al. 2002).

A ideia apresentada no ponto anterior constitui uma solução de Business Intelligence. Através de tecnologias eficazes é construído um sistema que permite obter informação confiável, detalhada, de forma resumida e de fácil interpretação, uma vez que utiliza funcionalidades disponíveis por ferramentas, tais como tabelas, gráficos e relatórios.

### **3.1.5 Definição do conceito**

Na fase final do modelo NCD surge a definição de conceito. É nesta fase que o inovador realiza um argumento convincente para o investimento na oportunidade escolhida, conhecida por "win statement" ou "gate document". Esta fase tem por objetivo apresentar um caso de investimento que possui informação qualitativa e quantitativa utilizada para convencer os clientes (Koen, P. A. et al. 2002). Os critérios e requisitos variam, normalmente, dependendo da natureza do tipo de conceito assim como da atitude do decisor face ao risco. Se o caso apresentado não for convincente o suficiente, o "conceito" retorna para o modelo NCD em vez de prosseguir para o processo NPD ou TSG, obrigando a uma nova iteração do processo (Koen, P. A. et al. 2002). O processo de definição de conceito pode ser realizado, através das seguintes metodologias (Koen, P. A. et al. 2002):

- Abordagens de entrega de objetivos, que corresponde ao tempo despendido em definir cuidadosamente os objetivos do projeto e os resultados esperados.
- Definição de critérios para a organização descrevendo o que será um projeto atrativo.
- (em termos financeiros, crescimento de mercado, dimensão de mercado, etc.).
- Avaliação rápida de inovações com elevado potencial.
- Utilização rigorosa do processo TSG para projetos de alto risco.

- Perceber e determinar o limite da capacidade de performance da tecnologia.
- Participação precoce do cliente em testes reais dos produtos, ou seja, envolvimento do cliente mesmo antes do produto estar completo.
- Estabelecer parcerias fora das áreas centrais de competência.
- Perseguir abordagens científicas alternativas.

Com a utilização do sistema explicado anteriormente, a informação surge de forma resumida e de fácil interpretação porque através das funcionalidades disponibilizadas pelas ferramentas é possível obter indicadores sob a forma de *dashboards* e relatórios.

## **3.2 Análise de valor**

É através da análise de valor que se consegue entender se o tema selecionado é vantajoso e traz benefícios não só para a organização, mas também para os clientes que pretendem ter acesso ao produto ou serviço.

### **3.2.1 Valor, valor para o cliente e valor percebido**

As organizações geralmente apresentam um conjunto de valores os quais são defendidos para moldar a cultura da mesma. Apesar disto, o conceito é bastante subjetivo, porque os valores defendidos pela organização podem diferir dos valores que os clientes pretendem. De certa forma, traduz-se não só em valores monetários, como também em benefícios que os produtos e serviços comercializados oferecem e depende das necessidades que cada um tem. Pode-se definir o conceito de valor para o cliente como “a avaliação feita pelo cliente da relação entre benefícios e custos percebidos” (Toledo, 2016). De certa forma, o cliente paga para receber em troca outros benefícios, tais como tempo, esforço, energia, qualidade. Apesar disto, o valor percebido varia conforme o cliente. Cada cliente avalia o custo-benefício dos produtos ou serviços e entende se o que é cobrado em troca dos benefícios é justo.

A diferenciação pressupõe que uma organização ofereça, um produto/serviço que seja considerado único pelos clientes, ou seja, cujas características o distingam daqueles oferecidos pela concorrência (Dezordi, 2021).

No sistema desenvolvido, os benefícios superam os custos de implementação, na medida em que os colaboradores da organização possam utilizar a mesma para obterem análises de uma forma mais expedita, fiável e sendo visualmente mais apelativas.

### 3.3 Proposta de Valor

Segundo Alexander Osterwalder, pode-se definir o conceito de proposta de valor como “uma visão geral de um pacote de produtos e serviços da organização que, em conjunto, representam valor para um segmento específico de clientes” (Osterwalder, 2004).

Pode-se então alegar que, a proposta de valor descreve os benefícios que são entregues aos clientes ou até mesmo aos colaboradores e acionistas. Assim sendo, as organizações apresentam os motivos que as diferenciam da concorrência e a razão pela qual os clientes devem preferir os seus produtos.

Para se criar valor é necessário inovar e recriar, pois as necessidades dos clientes são cada vez exigentes e mudam constantemente (Osterwalder, 2004).

No atual contexto, pretende-se desenvolver um sistema baseado em Business Intelligence que permita obter indicadores acerca do desempenho da equipa comercial de uma organização. Utilizando esta solução, os colaboradores são capazes de obter a informação do negócio num único local, o que possibilita uma melhoria no processo de tomada de decisão.

As tecnologias utilizadas para implementar a solução permitem que o processo garanta a qualidade da informação e a transmita ao utilizador final com resultados confiáveis.

### 3.4 Modelo de negócio Canvas

O modelo CANVAS permite de uma forma fácil estruturar o negócio e visualizar o mesmo numa única página. É a partir deste que se adquire a visão geral sobre o negócio e se compreende como é criado valor para o cliente (Muhtaroglu, 2013). O modelo CANVAS assenta em nove blocos que devem ser preenchidos da seguinte forma:

- **Segmento de clientes:** que clientes estão dispostos a pagar pelo produto/serviço que se disponibiliza em troca do valor que é fornecido.
- **Proposta de valor:** “valor” /” benefício” que o produto ou serviço entrega ao cliente.
- **Canais:** formas de comunicação e distribuição pelas quais a organização entrega valor ao cliente.
- **Relacionamento com os clientes:** estratégias adotadas pela organização para conseguir reter clientes.
- **Receitas:** de que forma o cliente pagará pelos benefícios recebidos. (venda de produtos, assinatura, licença, leilões).
- **Recursos-chave:** ativos fundamentais para que o negócio funcione, uma vez que permitem chegar até aos clientes (máquinas e instalações, patentes, recursos humanos, programadores).
- **Atividades-chave:** atividades que a empresa deve fazer numa base regular para conseguir entregar valor ao cliente (resolução de problemas, gestão de aplicações, desenvolvimento de produtos).

- **Parceiros-chave:** organizações/parcerias que contribuem para a realização do negócio.
- **Estrutura de custos:** custos decorrentes da operacionalização do desenvolvimento da solução. A ideia de negócio que se pretende desenvolver foi apresentada num modelo de negócio CANVAS, que pode ser visualizado na Figura 21.



Figura 21: Modelo de negócio Canvas

### 3.5 Cadeia de valor de Porter

A cadeia de valor pode ser definida como um conjunto de atividades que uma organização realiza para gerar valor para os clientes.

Introduzido por Porter (Rejeb and Keogh, 2021) , a estrutura da cadeia de valor é uma ferramenta bem conhecida que ajuda a entender as origens da vantagem competitiva. É um método que distingue as atividades de negócios, subdividindo-as em partes mais pequenas e mais fáceis de gerir e, em seguida, analisando o impacto de cada atividade no custo e valor para otimizar a cadeia de valor (Rejeb and Keogh, 2021). Agrega uma empresa em atividades estrategicamente relevantes e interligadas.

Uma simples visualização da estrutura mostra que as atividades de uma empresa são divididas em dois grupos principais (conforme apresentado na Figura 22).

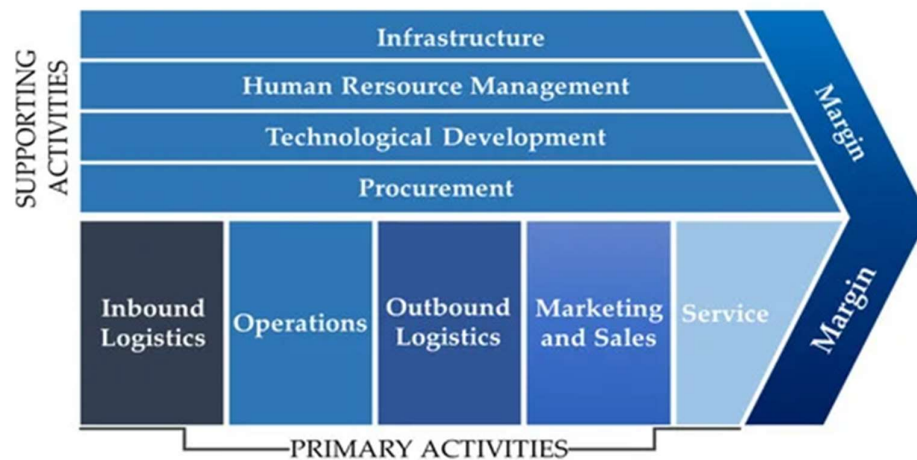


Figura 22: Cadeia de valor de Porter (Nagy, J. *et al.* 2018)

As atividades de suporte e primárias consistem nas seguintes (Rejeb and Keogh, 2021):

#### Atividades de suporte:

- **Infraestrutura:** sistemas que dão suporte à empresa e permitem manter as operações diárias. Para tal a organização deve definir uma estrutura organizacional composta por diferentes departamentos (financeiro, recursos humanos, administração, desenvolvimento).
- **Gestão de recursos humanos:** forma como uma organização recruta e mantém os seus funcionários. Existe um conjunto de ações que incentivam os recursos a permanecerem fiéis à organização nomeadamente: promoções, bom ambiente de trabalho e estágios.
- **Desenvolvimento tecnológico:** atividades relacionadas com a gestão e processamento de informações. Isto é possível através de utilização de tecnologias e ferramentas recentes no mercado e que contribuem para inovar os produtos.
- **Compras:** recursos que uma organização necessita para operar. Neste caso, é necessário adquirir servidores para alojar bases de dados, bem como software para análise dos dados.

#### Atividades primárias:

- **Logística de entrada:** atividades relacionadas com a receção, armazenamento e distribuição das entradas no produto. Relativamente à dissertação as entradas são os dados obtidos do ERP da organização.
- **Operações:** atividades que transformam entradas em saídas. No contexto atual, os dados iniciais são transformados em informações úteis para o utilizador final.
- **Logística de saída:** atividades que permitem a entrega do produto ou serviço ao cliente. No caso em questão, não existe entrega de um produto físico, sendo que o que se pode considerar é apenas a manutenção do sistema.

- **Marketing e vendas:** processos que diferenciam os produtos da concorrência persuadindo os clientes a comprar os mesmos. A criação de tutoriais que mostram o modo de utilização do sistema, assim como a facilidade de utilização podem ser aplicados neste contexto.
- **Serviço:** atividades relacionadas com a manutenção do valor do produto ou serviço após a compra. Neste ponto, a existência de um serviço de apoio e de gestão da infraestrutura associada ao produto são essenciais.

Através da análise acima referida, é evidenciado o potencial que o tema proposto tem na criação de valor e a possível adaptabilidade a uma utilização num contexto real de uma organização.

### 3.6 Método de Análise Hierárquica para Comparação de Tecnologias

Thomas Saaty, em 1980, introduziu um método multicritério designado de Analytic Hierarchy Process (AHP) (Saaty, 1987) para ajudar a resolver problemas complexos e apoiar na tomada de decisão. Este método permite estabelecer prioridades e dividir o problema complexo em menores de forma a conseguir obter todos aspetos, sejam eles objetivos ou subjetivos.

O problema complexo é decomposto em fatores, que por sua vez podem ainda ser decompostos até ao nível mais baixo, de forma a estabelecer relações e resumir a informação, comparando aspetos quantitativos e qualitativos de cada alternativa (Marins, Souza and Barros, 2009).

Este método foi aplicado nesta dissertação conforme as várias fases definidas pelo mesmo, para decidir qual a ferramenta de BI a utilizar, de modo que os dados que constituem o armazém de dados possam ser carregados para gerar análises visualmente mais apelativas. As quatro ferramentas a considerar serão descritas posteriormente no estado da arte.

Na primeira fase, o método AHP estrutura o problema em níveis hierárquicos, de forma a seguir compreender e avaliar o mesmo. O primeiro nível da hierarquia corresponde ao problema, o segundo aos critérios e o terceiro às alternativas. Aplicando o método AHP ao contexto atual, pode-se definir a estrutura presente na Figura 23.

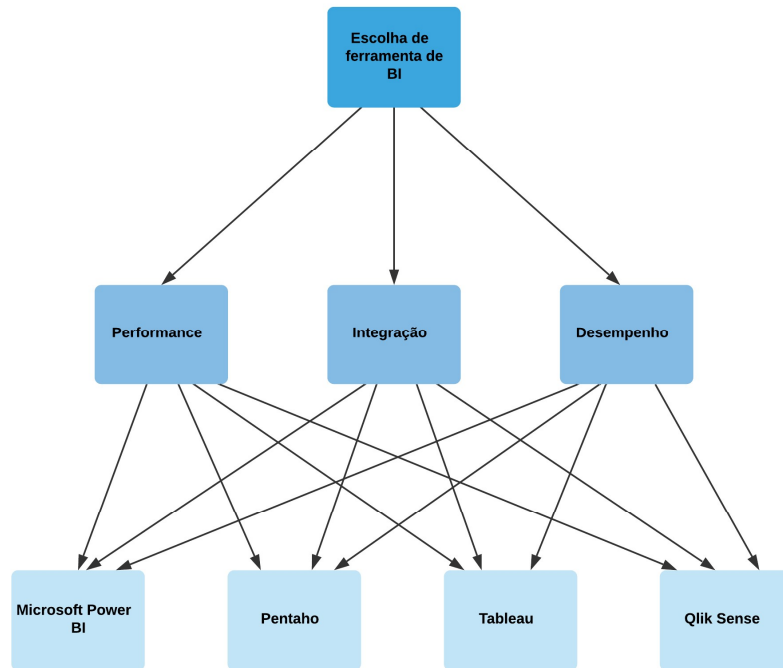


Figura 23: Estrutura hierárquica AHP

De forma a alcançar o objetivo definido, foram estabelecidos os seguintes critérios:

- Performance - Desempenho na execução de tarefas.
- Integração - Integração com várias fontes de dados.
- Usabilidade - Facilidade de utilização da ferramenta.

Na fase seguinte do método AHP são atribuídas prioridades aos critérios e alternativas apresentados anteriormente (divisão hierárquica). Para isso são definidas prioridades para os critérios e de seguida são definidas prioridades para cada alternativa em relação ao critério em causa. Desta maneira é possível perceber qual é a melhor solução para determinado critério. Antes de serem atribuídas as prioridades é preciso estabelecer uma escala de valores que, não devem ultrapassar um total de nove fatores, para ser possível comparar. Saaty sugere a utilização de uma escala padrão com o objetivo de avaliar de 1 a 9 os critérios e as alternativas no processo de decisão. Os valores numéricos dessa escala possuem uma definição qualitativa, tal como se pode visualizar na tabela 3.

Tabela 3: Escala fundamental de Saaty (Saaty, 1987)

Nível de importância	Definição	Explicação
1	Igual importância	As duas atividades contribuem igualmente para o objetivo.
3	Fraca importância	A experiência e o julgamento favorecem levemente uma atividade em relação à outra.
5	Forte importância	A experiência e o julgamento favorecem fortemente uma atividade em relação à outra.
7	Muito forte importância	Uma atividade é muito fortemente favorecida em relação à outra.
9	Importância absoluta	A evidência favorece uma atividade em relação à outra com o mais alto grau de certeza
2,4,6,8	Valores intermédios	Quando se procura uma condição de compromisso entre duas definições

Assim sendo, a tabela 4 representa as prioridades relativamente a cada critério.

Tabela 4: Matriz de prioridades entre critérios

	Performance	Integração	Usabilidade
Performance	1	3	5
Integração	1/3	1	4
Usabilidade	1/5	1/4	1

Como é possível verificar na tabela anterior o critério performance tem uma forte importância comparativamente com o critério usabilidade, sendo levemente favorecido relativamente ao critério integração. No que diz respeito à comparação entre os critérios integração e usabilidade, pode-se verificar que a integração tem maior importância comparativamente à usabilidade.

Prosseguindo com a aplicação do método AHP, na terceira fase obtém-se a matriz normalizada dos critérios de segundo nível (o objetivo é igualar todos os critérios à mesma unidade, onde cada valor da tabela anterior é dividido pelo total da coluna respetiva), com a prioridade relativa de cada critério, como podemos verificar na tabela 5.

Tabela 5: Matriz de critérios normalizada, com prioridade relativa

	Performance	Integração	Usabilidade	Prioridade relativa
Performance	0,6522	0,7059	0,5000	0,6194
Integração	0,2174	0,2353	0,4000	0,2842
Usabilidade	0,1304	0,0588	0,1000	0,0964

Analisando a tabela anterior, o cálculo da prioridade relativa é obtido através da média dos valores de cada linha da matriz normalizada, indicando a ordem de importância de cada critério. Pode-se verificar que a performance é o critério mais importante, seguindo-se a integração e por fim a usabilidade.

Na fase seguinte pretende-se avaliar a consistência das prioridades relativas. Para isso é necessário calcular a Razão de Consistência (RC). Se a RC for inferior a 0,1, pode-se concluir que os valores das prioridades relativas são consistentes.

Sendo assim, considera-se que:

$$[Ax = \lambda_{max} x]$$

onde  $x$  é o vetor próprio, obtido através das prioridades relativas da tabela anterior.

Dessa forma, obtém-se a seguinte matriz:

$$\lambda_{max} \cong \begin{bmatrix} 1,9541 \\ 0,8764 \\ 0,2913 \end{bmatrix}$$

Calculando a média dos valores obtidos pela matriz anterior e dividindo pelos valores da matriz de critérios normalizada com o valor da prioridade relativa de cada critério (tabela 3), obtemos o  $\lambda_{max}$ .

$$\lambda_{max} = \frac{1,9541}{0,6194} + \frac{0,8764}{0,2842} + \frac{0,2913}{0,0964} = 3,0867$$

Para calcular o Índice de Consistência (IC), sendo  $n$  igual ao número de critérios, aplica-se a seguinte fórmula:

$$IC = \frac{(\lambda_{max} - n)}{(n - 1)} = \frac{(3,0867 - 3)}{(3 - 1)} = 0,0433$$

Depois de calculado o IC, para se calcular a RC, é necessário saber qual é o valor do Índice Aleatório (IR). O valor de IR obtém-se a partir de uma tabela própria, tendo em conta o valor de  $n$ .

Tabela 6: Valores de IR para matrizes quadradas de ordem  $n$

n	1	2	3	4	5	6	7	8	9	10
IR	0,00	0,00	0,58	0,90	1,12	1,24	1,32	1,41	1,45	1,49

Como o número de critérios é 3, consultando a tabela anterior podemos verificar que o valor de IR é 0,58. Posto isto a RC é calculada da seguinte forma:

$$RC = \frac{IC}{IR} = \frac{0,0433}{0,58} = 0,0747$$

Como 0,0747 é menor que 0,1, os valores das prioridades relativas são considerados como estando consistentes.

Na fase cinco, constrói-se a matriz de comparação par a par relativamente a cada critério, levando em conta cada uma das alternativas selecionadas. Assim como anteriormente realizado na fase três, obtém-se a prioridade relativa de cada um dos critérios para as alternativas escolhidas, como se pode verificar nas Tabelas 7,8 e 9. O preenchimento das referidas tabelas foi realizado baseado em alguns artigos, tais como por exemplo “Power BI vs Tableau vs Qlik Sense: Which BI Tool is the Winner?” (SelectHub 2022) e “Power BI Vs Tableau: 10 Key Differences” (IntelliPaat 2022), assim como alguns testes realizados pelo autor deste trabalho.

Tabela 7: Matriz de comparação do critério Performance

	Power BI	Pentaho	Tableau	Qlik Sense	Prioridade relativa
Power BI	0,4615	0,3333	0,5217	0,4000	0,4292
Pentaho	0,1538	0,1111	0,0870	0,0667	0,1046
Tableau	0,2308	0,3333	0,2609	0,4000	0,3062
Qlik Sense	0,1538	0,2222	0,1304	0,1333	0,1600

Tabela 8: Matriz de comparação do critério Integração

	Power BI	Pentaho	Tableau	Qlik Sense	Prioridade relativa
Power BI	0,5357	0,4800	0,6316	0,3571	0,5011
Pentaho	0,1786	0,1600	0,1053	0,2857	0,1824
Tableau	0,1786	0,3200	0,2105	0,2857	0,2487
Qlik Sense	0,1071	0,0400	0,0526	0,0714	0,0678

Tabela 9: Matriz de comparação do critério Usabilidade

	Power BI	Pentaho	Tableau	Qlik Sense	Prioridade relativa
Power BI	0,4918	0,3846	0,5581	0,4091	0,4609
Pentaho	0,0984	0,0769	0,0698	0,0455	0,0726
Tableau	0,2459	0,3077	0,2791	0,4091	0,3104
Qlik Sense	0,1639	0,2308	0,0930	0,1364	0,1560

Como é possível verificar nas tabelas acima, tendo em conta os três critérios definidos a ferramenta *Microsoft Power BI* surge em primeiro lugar, seguindo-se as ferramentas *Tableau*, *Qlik Sense* e *Pentaho*.

Na fase 6 são obtidas as prioridades compostas das alternativas, recorrendo à multiplicação da matriz prioridade, sendo esta construída através dos valores das tabelas anteriores (Tabela 7, 8

e 9) com os valores das prioridades relativas de cada critério (matriz com o peso de cada critério – tabela 4).

$$\begin{bmatrix} 0,4292 & 0,5011 & 0,4609 \\ 0,1046 & 0,1824 & 0,0726 \\ 0,3062 & 0,2487 & 0,3104 \\ 0,1600 & 0,0678 & 0,1560 \end{bmatrix} * \begin{bmatrix} 0,6194 \\ 0,2842 \\ 0,0964 \end{bmatrix} = \begin{bmatrix} 0,4527 \\ 0,1237 \\ 0,2903 \\ 0,1334 \end{bmatrix}$$

Por fim, na última fase, escolhemos a alternativa com a prioridade relativa mais elevada, em função dos critérios definidos e das respectivas importâncias. Na Tabela 10 são apresentadas as prioridades relativas de cada critério e a prioridade composta de cada alternativa.

Tabela 10: Matriz final do AHP para escolha de ferramenta de BI

	Performance	Integração	Usabilidade	Prioridade composta
Power BI	0,4292	0,5011	0,4609	0,4527
Pentaho	0,1046	0,1824	0,0726	0,1237
Tableau	0,3062	0,2487	0,3104	0,2903
Qlik Sense	0,1600	0,0678	0,1560	0,1334

Analisando a Tabela 10, pode-se verificar que a alternativa com a prioridade composta mais elevada é a ferramenta *Microsoft Power BI*, o que mostra que é a melhor alternativa para o problema, visto ser a que apresenta valores mais elevados nos três critérios definidos, sendo esta a ferramenta selecionada para ser utilizada na solução.



## 4. Análise e Design

Neste capítulo é feita uma descrição sucinta das regras de negócio da área comercial de uma indústria alimentar assim como a descrição técnica que permite atingir os objetivos propostos nesta dissertação. De uma maneira simplificada, é apresentada uma visão do domínio da área comercial na área alimentar, apresentados os atores que vão usufruir do sistema, enumerados os requisitos funcionais e não funcionais assim como apresentadas propostas de arquitetura da solução.

Também neste capítulo e de forma a selecionar a melhor arquitetura para o sistema é feita uma comparação entre as arquiteturas possíveis, são apresentadas as fontes de dados utilizadas e quais as estruturas necessárias para conceber a arquitetura.

### 4.1 Domínio do problema

De forma a ser possível propor uma arquitetura para o sistema, é importante entender como funciona a área comercial do negócio. Os clientes encontram-se segmentados por tipo e por área geográfica. A cada cliente é atribuído um comercial, sendo essa atribuição feita conforme a área geográfica e o tipo de cliente. Os clientes são visitados pelo comercial mediante rotas de visita semanal, quinzenal ou mensal. Estas visitas podem servir para recolha de encomendas, apresentação de produtos, receitas e campanhas, recolha de cobranças, pedidos de devolução e pedidos de apoio técnico. Estas interações comerciais também podem ser feitas através de contato telefónico ou email.

A área comercial é constituída por várias equipas, sendo cada equipa constituída por um conjunto de comercias e liderada por um supervisor. O supervisor coordena e orienta o trabalho da sua equipa, fazendo também a ponte com os vários departamentos e reportando á direção comercial.

A organização encontra-se dividida em quatro localizações, com a seguinte disposição:

- Vila Nova de Gaia – sede
- Lisboa - delegação
- Madeira - delegação
- Algarve – delegação

As entregas das encomendas aos clientes regra geral são feitas mediante a proximidade a cada uma destas localizações, podendo ser realizadas pela distribuição interna ou com recurso a empresas de distribuição. Nada impede que um cliente possa ser fornecido por qualquer uma destas localizações.

## 4.2 Identificação dos atores

Utilizadores do sistema – são considerados utilizadores do sistema os colaboradores que integram o departamento comercial, departamento de marketing e a direção. Cada um dos intervenientes irá fazer uma utilização semelhante, variando somente o nível de detalhe a que cada um tem acesso. A consulta dos dados do sistema irá auxiliar a tomada de decisão dos vários utilizadores.

## 4.3 Engenharia de requisitos

Na engenharia de requisitos é efetuada a análise dos requisitos funcionais e não funcionais. Os requisitos são classificados segundo as categorias do modelo FURPS+ (COEPD, 2022).

A adoção deste modelo advém do seu reconhecimento histórico e da sua aprovação conseguida pelo sucesso de bastantes implementações.

### 4.3.1 Requisitos não funcionais

#### 4.3.1.1 Suportabilidade

Neste projeto foram identificados os seguintes requisitos de funcionalidade:

- Os *dashboards* e *reports* devem estar disponíveis em múltiplas plataformas.

#### 4.3.1.2 Usabilidade

Neste projeto foram identificados os seguintes requisitos de usabilidade:

- Os *dashboards* e *reports* devem apresentar um design apelativo e de leitura intuitiva.

#### 4.3.1.3 Desempenho

Neste projeto foram identificados os seguintes requisitos de desempenho:

- A Integração de dados e a atualização do cubo deve ser otimizada em termos de performance. Idealmente toda a operação não deve ultrapassar os 15 minutos.

#### 4.3.1.4 Suportabilidade

Neste projeto foram identificados os seguintes requisitos de suportabilidade:

- A interface dos *dashboards* e *reports* deve adaptar-se às características da plataforma em que estão a ser visualizados.

#### 4.3.1.5 Restrições adicionais (+)

##### 4.3.1.5.1 Restrições de implementação

Neste projeto foram identificados os seguintes requisitos:

- A solução deve ser desenvolvida com recurso a ferramentas da Microsoft.

##### 4.3.1.5.2 Restrições de segurança

Neste projeto foram identificados os seguintes requisitos:

- A solução deve ser capaz de implementar diferentes níveis de autoridade.

#### 4.3.2 Requisitos funcionais

Os requisitos funcionais determinam as principais características do sistema. Estes requisitos assumem somente os requisitos associados à implementação das características do sistema.

A sua representação é realizada na forma de casos de uso:

1. Visualizar *dashboard*.

##### 4.3.2.1 Casos de uso

O projeto possui apenas um caso de uso, conforme a Figura 24.



Figura 24: Diagrama de casos de uso

###### 4.3.2.1.1 Visualizar *dashboard*

O utilizador seleciona a lista de *dashboards*. O sistema mostra a lista de *dashboards* disponíveis. O utilizador seleciona o *dashboard*. O sistema mostra filtro de dados disponíveis caso existam. O utilizador seleciona filtros de dados. O sistema mostra o *dashboard* com os filtros aplicados.

Para representação do fluxo básico do caso de uso foi realizado o diagrama de sequência de sistema, possível de verificar na Figura 25. Neste diagrama estão representados os passos necessários para consultar um *dashboard*.



Figura 25: Diagrama de sequência visualizar dashboard

#### 4.3.2.2 Análises

Algumas das análises que foram solicitadas:

- Como utilizador eu pretendo obter a análise de vendas com a quantidade por categoria de produtos.
- Como utilizador eu pretendo obter a análise de vendas por clientes/produtos.
- Como utilizador eu pretendo obter a análise de vendas de quantidade vendida por tipo de cliente.
- Como utilizador eu pretendo obter a análise do nº de referências adquirida por cada cliente.
- Como utilizador eu pretendo obter a análise da quantidade vendida por cada comercial.
- Como utilizador eu pretendo obter a análise das vendas por cada equipa de comerciais.
- Como utilizador eu pretendo obter a análise da quantidade vendida por cada uma das localizações da empresa.
- Como utilizador eu pretendo obter a análise da quantidade entregue por cada um dos distribuidores logísticos.

Para além das mencionadas anteriormente, futuramente poderão ser solicitadas outros tipos de análises.

## 4.4 Opções de arquitetura

Na fase de design da solução, anterior à de implementação, foi imprescindível idealizar possíveis arquiteturas para o sistema. Dependendo dos dados de origem, o tratamento que é necessário aplicar sobre os mesmos ou mesmo até o instante em que se pretende visualizar os resultados, tudo isso irá influir na implementação da arquitetura.

Na conceção de uma arquitetura de BI, tendo em conta a metodologia que se pretende aplicar, poderão ser utilizadas um conjunto de tecnologias e componentes tais como ETL, *Data Mining*, *Data Warehouse*, Data Mart, entre outros (Elias 2018).

Se não for necessário realizar grandes limpezas e transformações aos dados de origem, os mesmos podem ser acedidos diretamente com recurso a ferramentas de BI que analisam e transmitem os dados ao utilizador final. Se pelo contrário, os dados de origem necessitarem de consideráveis operações de transformação e limpeza, as mesmas podem ser realizadas e posteriormente os dados apresentados ao utilizador. Assim, depois de os dados sofrerem um processo de limpeza e transformação são carregados num armazém de dados, podendo já nesta fase serem utilizadas ferramentas de BI para estabelecerem ligação a este repositório e apresentarem os dados ao utilizador. Outra alternativa é utilizar este Data Mart com os dados das vendas para criar um cubo OLAP a partir do mesmo. As ferramentas de BI estabelecem ligação ao cubo OLAP e é desta forma que os dados são transmitidos ao utilizador. Estas duas últimas opções trazem a vantagem de ser possível armazenar um histórico de alterações, o mesmo não é possível na primeira opção.

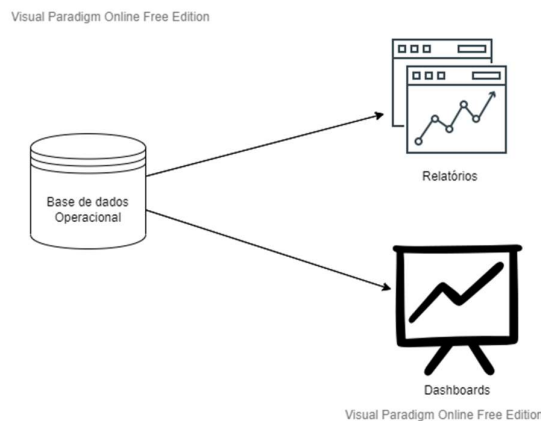


Figura 26: Opção de arquitetura 1

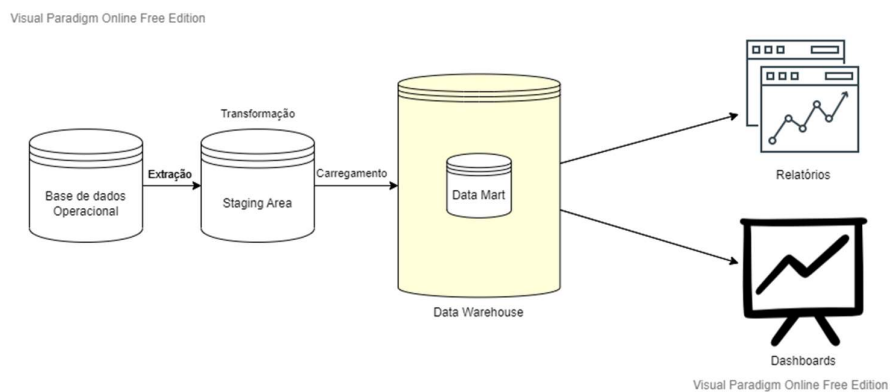


Figura 27: Opção de arquitetura 2 ROLAP

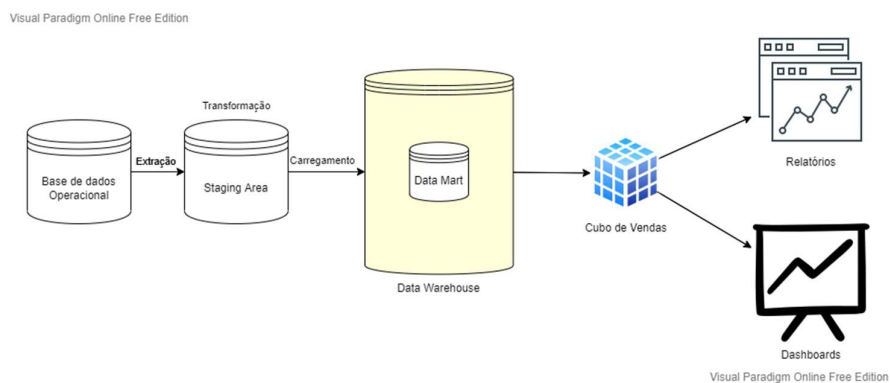


Figura 28: Opção de arquitetura 3 MOLAP

Estabelecendo uma comparação entre as três arquiteturas, a diferença encontra-se na forma como as ferramentas de BI vão estabelecer ligação ao repositório dos dados a ser analisados. Como se pode verificar na Figura 26 as ferramentas de BI efetuam *queries* diretamente á base de dados operacional. Na Figura 27, como se pode verificar os dados são obtidos pelas ferramentas de BI diretamente do Data Mart que contém a informação do negócio que se pretende analisar, neste caso as vendas da área comercial. Por seu turno na Figura 28 os dados são obtidos a partir do cubo MOLAP. Esta última opção revela-se mais vantajosa porque o cubo permite criar medidas adicionais às existentes no Data Mart, indicadores de performance (KPI) e hierarquias nas dimensões, permitindo desta forma ao utilizador obter os dados utilizando por exemplo o Excel para realizar análises produzidas por si, de uma forma rápida e já com todas as medidas, KPI e hierarquias. É necessário ter em conta as características de cada arquitetura de forma a escolher as tecnologias a aplicar sobre o sistema.

## 4.5 Arquitetura escolhida

Nesta solução optou-se por seguir abordagem preconizada por Ralph Kimball, em que os dados das vendas da área comercial são armazenados num Data Mart.

Uma das principais vantagens que existe ao seguir a abordagem de Ralhp Kimbal é a de que se futuramente for necessário acrescentar novas áreas à solução atual, ou seja a criação de novos *Data Marts*, isso não irá comprometer o Data Mart existente. Tem também a vantagem de em caso de modificação da estrutura do sistema, seja por acréscimo/remoção de funcionalidades, seja por motivos de manutenção, é possível fazê-lo de uma forma célere e sem dificuldades de maior. O sistema a implementar deve ter a capacidade de guardar o histórico dos dados, permitindo desta forma o auxílio na tomada de decisão baseada em factos adquiridos anteriormente. Por esse motivo e por se adequar às características pretendidas foi escolhida a arquitetura 3, ilustrada na figura 28.

#### 4.5.1 Fontes de dados

O ERP Microsoft Business Central, que armazena as transações numa base de dados SQL, serviu como fonte de dados das vendas da área comercial. Esta base de dados é composta por milhares de tabelas e centenas de vistas. Este tipo de aplicações (ERP) permitem uma gestão centralizada das várias áreas da organização, tais como faturação, contabilidade, produção, logística, recursos humanos, imobilizado, entre outras. Na Figura 29 temos a representação de uma ínfima parte do modelo relacional da base de dados, centrado na tabela *Item Category*, para se perceber as características da mesma.

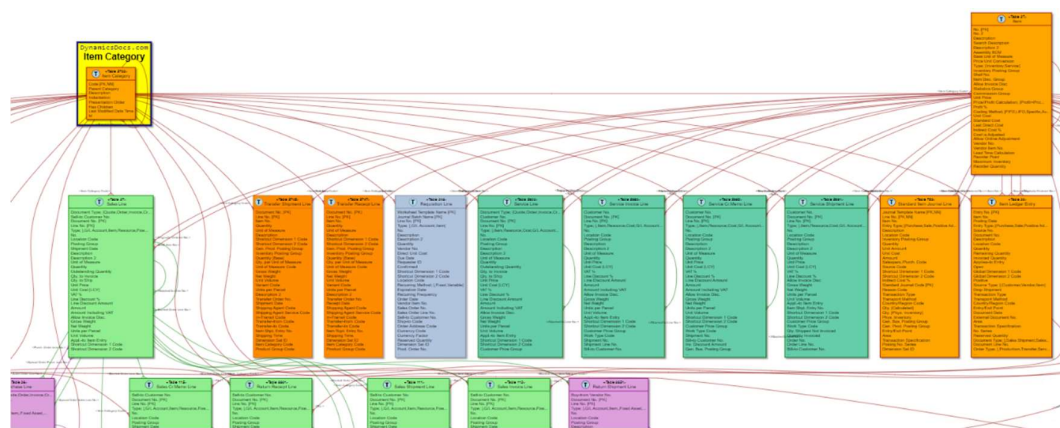


Figura 29: Parte do modelo relacional da base de dados (DynamicsDocs, 2022)

Os dados das vendas foram assim obtidos a partir dessa base de dados operacional SQL, o que obrigou a alguma análise e pesquisa de forma a perceber em que tabelas era armazenada a informação necessária e relevante para a construção do Data Mart e posterior construção do cubo de vendas. Após a análise foram identificadas as tabelas que continham a informação necessária assim como a necessidade de criação de algumas vistas de forma a facilitar a obtenção dos dados.

#### 4.5.2 Logical Data Map

O *Logical Data Map* é um artefacto no qual é realizado um mapeamento de cada atributo da base de dados operacional para a *Staging Area* e da mesma para o Data Mart. O *Logical Data Map* descreve a relação entre cada ponto de partida (fonte dos dados) e o seu extremo oposto (destino dos dados) de um processo de ETL (Power BI Expert, 2022).

Com recurso ao Excel, foi criado o *Logical Data Map* com os seguintes componentes:

- **Fonte:** identificação da base de dados origem, o nome da tabela, nome da coluna e o tipo de dados.
- **Destino:** identificação do nome da tabela de destino, o nome da coluna, o tipo de tabela (Dimensão ou Facto), tipo de dados e tipo de SCD.
- **Transformação:** conversão de dados necessária para o formato esperado. A mesma pode ser expressa em SQL ou pseudocódigo. Quando o campo não é preenchido, significa que os dados são carregados diretamente para o destino, não precisando de nenhuma transformação.

No Anexo 1 e no Anexo 2 encontra-se uma representação do *Logical Data Map* que foi criado como instrumento de suporte na criação do Data Mart.

#### 4.5.3 Staging Area

A *Staging Area* funciona como um repositório intermédio componente do processo de ETL. Para facilitar o carregamento de dados para este repositório não existe qualquer tipo de relação entre as tabelas. Na Figura 30 é possível visualizar a *Staging Area* considerada para o sistema desenvolvido.

Figura 30: Modelo da *Staging Area*

A estrutura da *Staging Area* foi idealizada de forma a armazenar a informação considerada necessária para posterior carregamento no Data Mart. Foram criadas as seguintes tabelas na *Staging Area*:

- SalesInvoiceHeader – tabela que contém os cabeçalhos das faturas.
- SalesInvoiceLine – tabela que contém as linhas das faturas.
- SalesCrMemoHeader – tabela que contém os cabeçalhos das notas de crédito.
- SalesCrMemoLine – tabela que contém as linhas das notas de crédito.
- Cust\_LedgerEntry – tabela que contém os movimentos dos clientes.
- ItemLedgerEntry – tabela que contém os movimentos dos produtos.
- Item – tabela que contém os dados dos produtos.
- Salesperson\_Purchaser – tabela que contém os dados dos comerciais.
- Vendor – tabela que contém os dados dos fornecedores.
- Location – tabela que contém os dados das localizações.
- ItemCategory – tabela que contém as categorias de produtos.
- Customer – tabela que contém os dados dos clientes.
- GL\_Account – tabela que contém dos dados das contas da contabilidade.
- ShippingAgent – tabela que contém os dados dos distribuidores.

#### 4.5.4 Modelação do Data Warehouse

Nesta dissertação é seguida a metodologia de Kimball. Um *Data Warehouse* pode ser composto por um único Data Mart ou por um conjunto de *Data Marts*, mas neste contexto, existe apenas um Data Mart referente às vendas da área comercial.

Na metodologia de Kimball, quando se concebe a modelação de um Data Mart é indispensável definir um conjunto de etapas que são fundamentais na sua arquitetura, e que de seguida são referidas:

- Identificação da área de negócio.
- Identificação do nível de granularidade/detalhe.
- Identificação das dimensões.
- Identificação dos factos.

A **área de negócio** é o que possibilita identificar o conteúdo das informações que constituem um determinado sistema numa organização. No atual contexto, cada linha da tabela de Factos no Data Mart representa uma venda da equipa comercial. Como tal, a área de negócio é vendas.

O **nível de Granularidade** define qual o nível mais alto que se pode obter as métricas na tabela de Factos, possibilitando desta forma ao utilizador uma maior ou menor capacidade analítica, isto porque a informação é apresentada com mais ou menos detalhe conforme o nível de granularidade. Na solução idealizada cada uma das linhas da tabela de Factos reflete a venda de um produto a um determinado cliente, num determinado dia, no âmbito de uma venda. Ou seja, o nível mais alto de granularidade da solução é ao dia.

Para dar resposta às necessidades dos utilizadores, foi realizada uma análise de forma a identificar as dimensões necessárias. Foram consideradas necessárias para o desenvolvimento da solução as seguintes dimensões:

- DimCliente
- DimData
- DimEnvio
- DimExpedidor
- DimFornecedor
- DimLocalizacao
- DimProduto
- DimVendedor

Foram considerados os seguintes atributos da tabela de Factos como sendo os mais relevantes atributos:

- Quantidade
- Quantidade Base
- Quantidade Vendida
- Quantidade Bónus
- Preço Unitário
- Valor Líquido
- % Desconto
- Valor Desconto
- Custo
- Data Pagamento
- Dias Pagamento
- Mês Pagamento
- Ano Pagamento
- Quantidade Campanha

#### 4.5.5 Dimensões

##### DimCliente

Toda a informação relativa ao Cliente encontra-se nesta dimensão, que se encontra representada na Tabela 11. Habitualmente nesta tabela só existe um registo por cliente, podendo existir vários registos para o mesmo cliente se existir alguma atualização num dos atributos que salvaguarda histórico e contém diversas informações como número, nome, endereço, contatos, código de vendedor, tipo de cliente, código de rota, território e localização geográfica.

Tabela 11: Dimensão Cliente

Atributo	Estratégia SCD
ClienteKey	-
No_ (Id do sistema)	-
Nome	1
Nome2	1
Alias	1
Endereco	1
Endereco2	1
Cidade	1
Pais	1
Distrito	1
CodigoPostal	1

Telefone	1
Email	1
CodigoVendedor	1
TipoCliente	1
CodigoRota	1
CodigoTerritorio	1
Geographic Location	1
CodigoVendedorHist	2
TipoClienteHist	2
CodigoRotaHist	2

Relativamente aos atributos “CodigoVendedorHist”, “TipoClienteHist” e “CodigoRotaHist”, foi aplicado SCD do tipo 2, visto que foi considerado relevante manter o histórico destes. Se um destes 3 atributos for atualizado é importante saber a data de alteração de forma a perceber que implicações traz para o cliente ou para o comercial. Se por exemplo um cliente mudar de comercial é relevante saber quando é que essa alteração aconteceu.

### DimData

A dimensão Data encontra-se representada na Figura 31. É extremamente importante num armazém de dados ser possível rastrear os dados armazenados ao longo do tempo, quer seja para pesquisa de dados, quer seja para verificação de histórico. O valor mais baixo a considerar nesta dimensão é o dia, mas é possível obter a informação a diversos níveis.

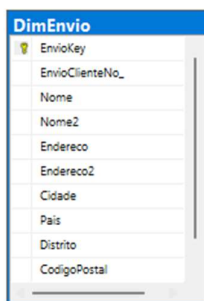
DimData	
DataKey	
DataCompleta	
Ano	
Semestre	
Quadrimestre	
Trimestre	
Mes	
MesNome	
Semana	
SemanaMes	
DiaNumeroAno	
DiaNumeroMes	
DiaNumeroSemana	
DiaSemana	
FimSemana	
Estacao	
UltimoDiaMes	

Figura 31: Dimensão Data

Visto que nesta dimensão os atributos são referentes a datas e não existe nenhum interesse em guardar histórico, nesta dimensão não foi aplicada nenhum tipo de SCD.

### **DimEnvio**

A dimensão Envio, representada na Figura 32, alberga toda a informação relativa aos dados de envio de um documento de faturação. É de salientar que os dados de faturação de um documento podem ser diferentes dos dados de envio, por esse motivo a existência desta dimensão torna-se relevante.



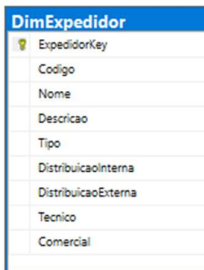
DimEnvio	
EnvioKey	
EnvioClienteNo_	
Nome	
Nome2	
Endereco	
Endereco2	
Cidade	
Pais	
Distrito	
CodigoPostal	

Figura 32: Dimensão Envio

Também na dimensão Envio não foi aplicado qualquer tipo de SCD.

### **DimExpedidor**

Na tabela da dimensão Expedidor representada na Figura 33, encontram-se os atributos que caracterizam quem é responsável por fazer a entrega, permitindo desta forma compreender melhor quais os níveis de distribuição interna e externa.



DimExpedidor	
ExpedidorKey	
Codigo	
Nome	
Descricao	
Tipo	
DistribicaoInterna	
DistribicaoExterna	
Tecnico	
Comercial	

Figura 33: Dimensão Expedidor

Na dimensão Expedidor também não foi aplicado qualquer tipo de SCD.

### **DimFornecedor**

Na tabela da Dimensão Fornecedor encontra-se toda a informação relativa ao fornecedor de um determinado produto. Permite identificar o fornecedor e qual a sua área geográfica. Na Figura 34 podemos ver a representação da dimensão Fornecedor.

DimFornecedor	
FornecedorKey	
No_	
Nome	
Nome2	
Alias	
Endereco	
Endereco2	
Cidade	
CodigoPais	
Distrito	
CodigoPostal	
Telefone	

Figura 34: Dimensão Fornecedor

Também na dimensão Fornecedor não existe a necessidade de guardar qualquer tipo de histórico e por esse motivo não foi aplicado nenhum tipo de SCD.

### DimLocalizacao

A dimensão Localizacao ilustrada na Figura 35, agrega os atributos que caracterizam a localização responsável por fazer a entrega de uma determinada encomenda. Torna-se relevante para ajudar a perceber qual o nível de entrega de cada uma das localizações.

DimLocalizacao	
LocalizacaoKey	
Codigo	
Nome	
Nome2	
Cidade	
CodigoPostal	
Distrito	
AreaGeografica	

Figura 35: Dimensão Localização

Tal como referido anteriormente noutras dimensões, também na dimensão Localização não foi aplicado qualquer tipo de SCD.

### DimProduto

A tabela da dimensão Produto engloba todas as informações sobre os produtos, representada na Tabela 12. Esta tabela contém informações como o código, descrição, categoria, grupo, segmento, código de fornecedor e tipo. Sobre os atributos “CodCategoria” e “GrupoContabilistico” foi aplicado SCD do tipo 2 pela necessidade salvaguarda de histórico. Se um destes 2 atributos for atualizado é importante saber a data de alteração de forma a ajudar a perceber possíveis implicações.

Tabela 12: Dimensão Produto

Atributo	Estratégia SCD
ProdutoKey	-
No_ (Id do sistema)	-
Descricao	1
Descricao2	1
Alias	1
Bloqueado	1
CodigoCategoria	1
CategoriaDescricao	1
GrupoContabilistico	1
Cod_GrupoProduto	1
No_ProdutoZeelandia	1
Nº_PLM	1
Segmento	1
Cod_Fornecedor	1
Tipo	1

### DimVendedor

Na dimensão Vendedor, visível na Figura 36, encontra-se todos os atributos que caracterizam um comercial. Esta dimensão permite identificar quem foi o comercial responsável pela venda assim como o seu supervisor. Foi aplicado SCD do tipo 1 nos atributos “Codigo”, “Nome”, “CodigoSupervisor” e “AreaGeografica” por não existir necessidade de guardar histórico. Já sobre o atributo “CodigoSupervisorHist” foi aplicado SCD do tipo 2 precisamente para permitir ter um histórico.

DimVendedor	
VendedorKey	
Codigo	
Nome	
CodigoSupervisor	
CodigoSupervisorHist	
AreaGeografica	
EffectiveDate	
ExpireDate	
IsCurrent	

Figura 36: Dimensão Vendedor

#### 4.5.6 Tabela de Factos

Depois de identificadas todas as dimensões, o passo seguinte é identificar os factos, segundo a arquitetura preconizada por Kimball. Os mesmos estão disponíveis na tabela de Factos, representados por variáveis numéricas e referentes aos dados de uma determinada venda. Para a construção do Data Mart foi utilizado o modelo em estrela, no qual todas as dimensões estabelecem uma relação direta com a tabela de Factos. Para além de ser considerado o mais apropriado ao utilizador, este tipo de modelo permite também um melhor desempenho na consulta dos dados, tornando-se por isso mais vantajoso.

No presente contexto, a tabela de Factos designa-se **FactVendas** e na mesma encontram-se todas as chaves das dimensões, assim como valores numéricos relativos a quantidades, preço, valores da venda, valores de desconto, data de pagamento de um documento, quantidades de campanhas promocionais, entre outros. A tabela de factos pode ser visualizada na Figura 37.

Column Name	Condensed Type	Nullable
DataKey	int	No
ClienteKey	int	No
ProdutoKey	int	No
LocalizacaoKey	int	No
ExpedidorKey	int	No
EnvioKey	int	No
FornecedorKey	int	No
VendedorKey	int	No
DocumentoNo	bigint	No
NoLinha	int	No
TipoDocume...	nvarchar(15)	No
TipoLinha	nvarchar(5)	No
Tipo	nvarchar(7)	No
Quantidade	float	No
QuantidadeB...	float	No
QtdVendida	float	No
QtdBonus	float	No
PrecoUnitario	float	No
ValorLiquido	float	No
[Desconto%]	float	No
ValorDesconto	float	No
Custo	float	No
DataPagame...	date	Yes
DiasPagame...	int	Yes
AnoPagame...	int	Yes
MesPagame...	int	Yes
CodigoCamp...	nvarchar(20)	No
QtdCampanha	int	No

Figura 37: Tabela de Factos Vendas

Por último, é visível na Figura 38 o modelo do Data Mart que foi implementado e a partir do qual foi desenvolvido um Cubo de Vendas.

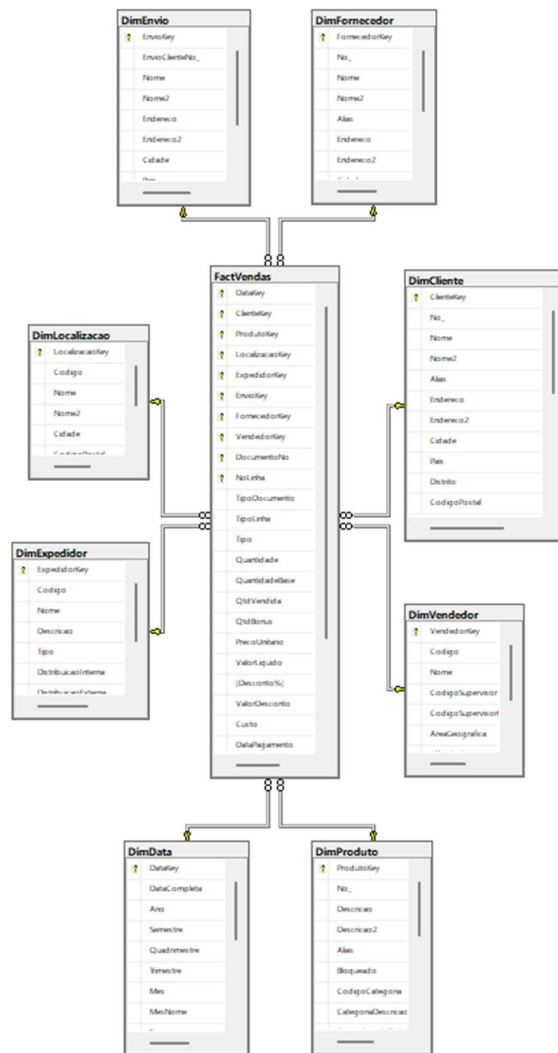


Figura 38: Data Mart Vendas



## 5. Implementação

A longo deste capítulo é realizada uma descrição do trabalho desenvolvido nos vários componentes da arquitetura. Na primeira secção é feita uma apresentação das tecnologias utilizadas durante a implementação da solução, sendo que nas secções seguintes são descritas as várias fases do processo de ETL da solução. As várias fases mostram as operações de extração, transformação/limpeza e carregamento dos dados. Por fim, é descrito como os utilizadores podem ter acesso às análises de dados.

### 5.1 Tecnologias Utilizadas

Durante o desenvolvimento e implementação da solução foram utilizadas algumas ferramentas do portefólio da Microsoft, particularmente o SQL Server Management Studio, o SQL Server Data Tools e o Analysis Services. O primeiro para a exploração de dados, criação de vistas, tabelas e do modelo de armazenamento de dados e o segundo no processo de integração de dados e na criação do cubo de dados. O SSIS tem a capacidade e as características para realizar todos os processos de ETL (Extração, Transformação e Carregamento) e facilitar a integração com o SQL Server e o Office. Como ferramenta de visualização e análise de dados utiliza-se o Power BI ou o Excel, que são outras ferramentas disponibilizadas pela Microsoft. A seleção das ferramentas é realizada baseada nas características apresentadas no capítulo dois e pelo fato de a organização já possuir licenciamento das mesmas.

### 5.2 Processo de ETL

O processo de ETL já foi abordado e descrito em capítulos anteriores. Resumindo, trata-se de um processo que compreende a extração de dados das diversas fontes, transformação dos mesmos considerando as regras de negócio, sendo concluído com o carregamento dos dados no repositório de destino.

#### 5.2.1 Extração de dados

A extração de dados é a primeira etapa no processo de ETL. É efetuada a ligação à base de dados SQL do ERP Microsoft Business Central, onde se encontra armazenada a informação, sendo extraída a informação para a *Staging Area*, onde a mesma permanecerá temporariamente. Na Figura 39 pode-se visualizar em destaque o processo de extração na opção de arquitetura escolhida para a solução.

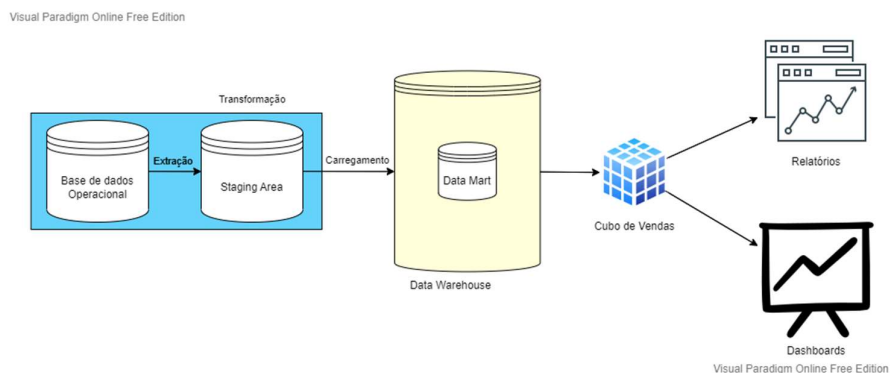


Figura 39: Etapa de extração na arquitetura da solução

Para simplificar o processo de carregamento de dados nas tabelas da *Staging Area*, foram criadas algumas vistas na base de dados operacional, visíveis na Figura 40. A criação dessas vistas tem como objetivo filtrar os dados da base de dados operacional antes dos mesmos serem carregados para a *Staging Area*.

- ⊕ [ícone] dbo.SSIS\_Copy\_CrMemoHeader\_ALL
- ⊕ [ícone] dbo.SSIS\_Copy\_CrMemoLine\_ALL
- ⊕ [ícone] dbo.SSIS\_Copy\_CustLedgerEntry\_ALL
- ⊕ [ícone] dbo.SSIS\_Copy\_ItemLedgerEntry\_ALL
- ⊕ [ícone] dbo.SSIS\_Copy\_SalesInvoiceHeader\_ALL
- ⊕ [ícone] dbo.SSIS\_Copy\_SalesInvoiceLine\_ALL
- ⊕ [ícone] dbo.SSIS\_Copy\_ValueEntry\_ALL

Figura 40: Vistas criadas na base de dados operacional

Na Tabela 13 pode-se verificar qual o objetivo pretendido com a criação de cada uma das vistas.

Tabela 13: Objetivos e filtros aplicados a cada vista

Vista	Objetivo	Filtros aplicados
SSIS_Copy_CrMemoHeader_All	Filtrar os dados da tabela Cabeçalhos de Notas de Crédito.	Data de registo superior a 31-12-2018.
SSIS_Copy_CrMemoLine_All	Filtrar os dados da tabela Linhas de Notas de Crédito.	Data de registo superior a 31-12-2018 e Tipo de Linha igual a Produto ou Conta.
SSIS_Copy_CustLedgerEntry_All	Filtrar os dados da tabela Movimentos Cliente.	Data de registo superior a 31-12-2018 e Tipo de Documento igual a Fatura ou Nota de Crédito.
SSIS_Copy_ItemLedgerEntry_All	Filtrar os dados da tabela Movimentos Produto.	Data de registo superior a 31-12-2018 e Tipo de Movimento igual a Venda.
SSIS_Copy_SalesInvoiceHeader_All	Filtrar os dados da tabela Cabeçalhos de Faturas.	Data de registo superior a 31-12-2018
SSIS_Copy_SalesInvoiceLine_All	Filtrar os dados da tabela Linhas de Faturas.	Data de registo superior a 31-12-2018 e Tipo de Linha igual a Produto ou Conta.
SSIS_Copy_ValueEntry_All	Filtrar os dados da tabela Movimentos Valor.	Data de registo superior a 31-12-2018

As tabelas sobre as quais foram criadas as vistas são tabelas com um tamanho considerável pois armazenam vários anos de histórico. Com as vistas que foram criadas reduziu-se o número de registos a serem obtidos da base de dados operacional para depois serem carregados para o repositório temporário, diminuindo desta forma o tempo desta operação. Após a criação das vistas na base de dados operacional o carregamento de todas as tabelas da *Staging Area* foi realizado de forma direta, ou seja, diretamente do repositório de origem para o repositório de destino. Na Figura 41 podemos visualizar o carregamento das tabelas da *Staging Area*.

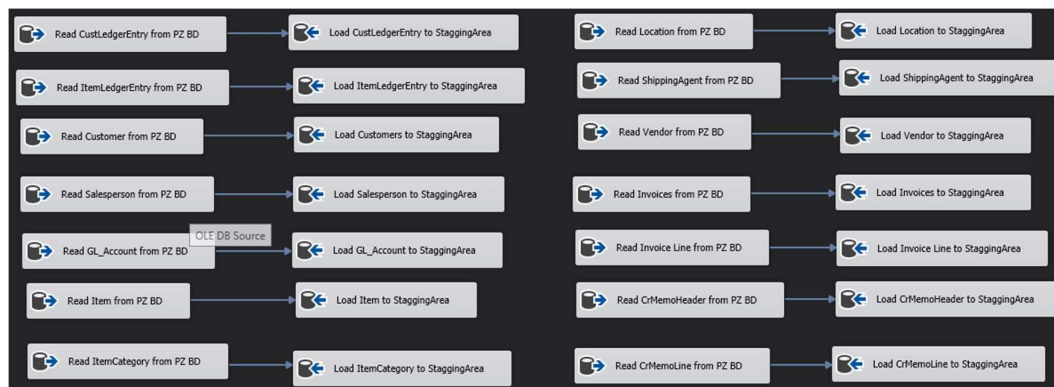


Figura 41: Carregamento das tabelas da *Staging Area*

## 5.2.2 Transformação e Carregamento de dados

Na etapa de transformação de dados é indispensável verificar se os dados armazenados na *Staging Area* estão em conformidade com as regras de negócio das Vendas e se não for o caso realizar as operações de transformação necessárias.

Na solução implementada não existiu a necessidade de realizar operações de transformação relevantes sobre os dados armazenados na *Staging Area*, isto deve-se ao facto de os dados serem obtidos da base de dados do ERP que já faz uma validação bastante completa dos mesmos. Essa validação por parte do ERP é ainda mais rigorosa na área da faturação e contabilidade devido a cumprimento de requisitos legais. Por exemplo todos os dados presentes num documento de faturação têm de ser válidos, desde as informações do cliente (nomes, moradas, NIF, etc) aos dados do produto (códigos produto, descrições, valores, etc). Não existe a possibilidade de campos obrigatórios não preenchidos ou um formato inválido.

Por opção, as ligeiras transformações que foram necessárias, foram realizadas na etapa de carregamento dos dados para as dimensões e para a tabela de factos, tal como podemos visualizar na Figura 42 o enquadramento das etapas de transformação e carregamento na opção de arquitetura escolhida para o sistema. A maioria das transformações consiste em atualização de valores nulos como por exemplo se o código de vendedor na linha for nulo é atualizado com o valor que está no cabeçalho da fatura. Ou na conversão de tipos de dados, como é o caso do nº de documento que é convertido de varchar para bigint.

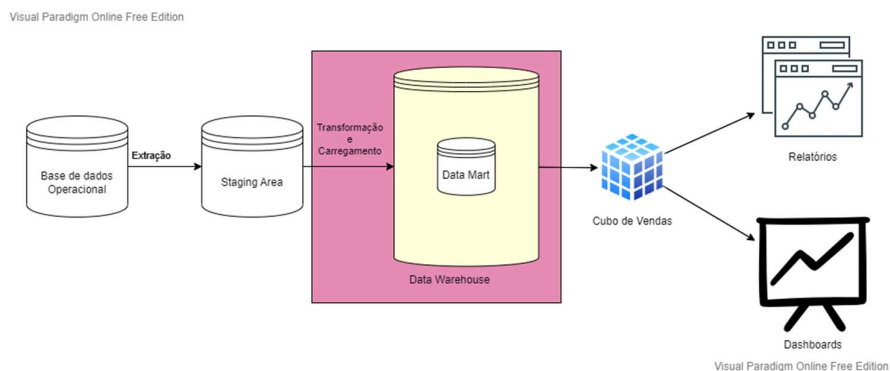


Figura 42: Etapa de transformação e carregamento na arquitetura da solução

Previamente à implementação do Data Mart, foi criado um *Logical Data Map*, tal como foi referido na secção *Logical Data Map*.

As principais transformações realizadas foram...

O carregamento dos dados da *Staging Area* para o Data Mart é dividido em duas etapas. Na primeira etapa são carregados os dados das dimensões e numa segunda etapa são carregados os dados da tabela de Factos, isto porque a tabela de Factos detém a chave estrangeira de cada uma das dimensões, logo tem uma dependência das mesmas.

Quando se efetua o carregamento dos dados em cada uma das dimensões é indispensável verificar se os atributos das mesmas devem ser atualizados com a informação mais atual, ou se pelo contrário devem manter o histórico. Na implementação do Data Mart da solução foram aplicados SCD do Tipo 1 e do Tipo 2.

Nas dimensões em que foi necessário aplicar SCD do Tipo 2 foram adicionados três atributos: "EffectiveDate", "ExpireDate" e "IsCurrent". Estes atributos tem o objetivo de guardar duas datas e um "marcador" a indicar se o registo de um cliente, produto ou vendedor é o mais atual.

Na solução implementada o carregamento das dimensões Expedidor, Localização, Fornecedor, Envio e Data foram diretas, ou seja, à exceção da dimensão Data todas as referidas anteriormente foram carregadas diretamente da *Staging Area* para o Data Mart.

No caso das dimensões Produto, Cliente e Vendedor, embora também tenham sido carregadas diretamente da *Staging Area*, como foi aplicado SCD do Tipo 2, as mesmas tiveram um carregamento diferente. Na Figura 43 pode-se visualizar o processo de carregamento da dimensão Produto. Durante esse processo são obtidos os registos das tabelas Item e GL\_Account da *Staging Area*, de seguida adiciona-se o tipo (produto ou conta) antes de se juntar o fluxo de dados com um "Union All". Depois é feito um "Lookup" a uma tabela da base de dados do ERP que contém atributos dos produtos de forma a obter-se o segmento do produto. De seguida junta-se novamente o fluxo de dados com um "Merge". Neste ponto já temos toda a informação que necessitamos para inserir ou atualizar a tabela DimProduto. Nos passos seguintes é feita a verificação dos atributos que guardam histórico. Se algum dos atributos que

guardam histórico for modificado é marcado o registo antigo com a data de expiração igual à data atual e é inserido na tabela um novo registo com a data efetiva igual à data atual.

O carregamento das tabelas da dimensão Cliente e da dimensão Vendedor foi efetuado de forma idêntica.



Figura 43: Carregamento da tabela DimProduto

Depois de terminado o carregamento das dimensões, inicia-se o processo de carregamento da tabela de Factos. Anteriormente ao início do processo de carregamento de dados da tabela de factos são eliminadas as restrições chave primária - chave estrangeira entre as dimensões e a tabela de factos de forma a facilitar o carregamento. As chaves foram criadas para permitir uma relação com as tabelas das dimensões. Durante este processo de carregamento de dados toda a informação é verificada, para que cumpra as regras do negócio, sendo carregada a informação numérica referente às vendas que seja considerada fundamental para obtenção de conclusões. Após o término deste processo são adicionadas novamente as restrições chave primária - chave estrangeira entre as dimensões e a tabela de factos.

A Figura 44 ilustra o processo de carregamento da tabela de Factos que foi implementado no Microsoft SSIS.



### 5.2.3 Cubo de vendas

Concluído a criação do Data Mart, já existem dados que podem ser utilizados por ferramentas de BI para a realização de análises, mas nesta solução optou-se pela criação de um Cubo utilizando os dados do Data Mart de Vendas.

A criação do Cubo permite por exemplo a adição novas medidas, KPI e hierarquias de uma forma rápida. O utilizador pode recorrer ao Excel ou ao Power BI para obter dados do Cubo, que os mesmos já estão agregados, validados e prontos a usar. Não é necessário o utilizador se preocupar com a criação de medidas ou de relações entre os dados para construir as suas análises.

Para a construção do Cubo foi utilizado o Microsoft SQL Server Data Tools e criado um projeto do tipo Analysis Services no qual são importadas as tabelas do Data Mart, como fonte de dados. Após a importação dos dados foram adicionadas ao cubo as dimensões e a tabela de factos. Foram também adicionadas hierarquias nas dimensões Data, Produto e Cliente. Na Figura 45 pode-se verificar o cubo construído.

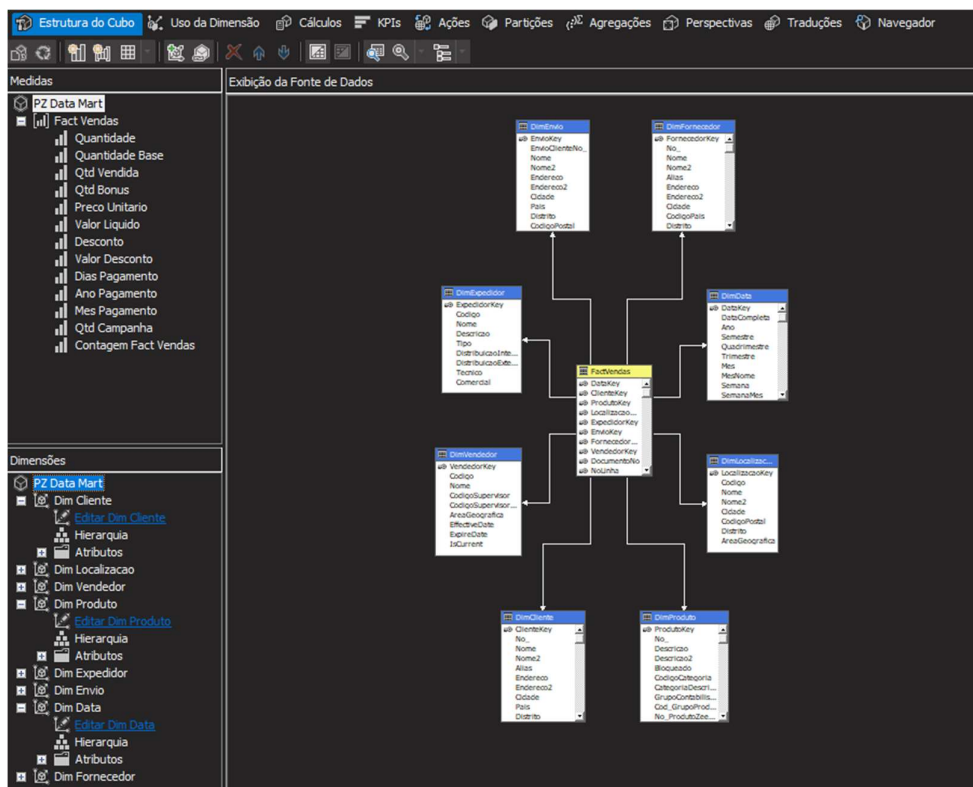


Figura 45: Cubo de Vendas Construído



Figura 46: Hierarquia de produto

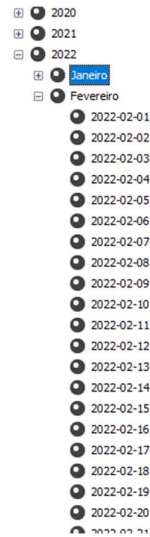


Figura 47: Hierarquia de data

Na Figura 46 pode-se visualizar a hierarquia criada para a DimProduto com dois níveis, código de categoria e nº de produto. Na Figura 47 pode-se visualizar a hierarquia criada para a DimData com três níveis, ano, mês e data. As hierarquias estabelecem uma relação do tipo pai-filho e permitem navegar entre níveis, ou seja, fazer o *drill down* e o *roll up*.

Para finalizar e de forma a permitir perceber como todos os processos referidos neste capítulo se interligam, é apresentado no Anexo 3 uma vista geral de todo o Fluxo de Controlo desenvolvido ao longo desta dissertação. Todos os processos apresentados anteriormente são apenas alguns exemplos do trabalho realizado ao nível do Fluxo de Controlo do Microsoft SQL Server Integration Services.

### 5.3 Análise de dados

A última etapa da arquitetura da solução desenvolvida é a de Análise de dados, a qual se encontra representada na Figura 48.

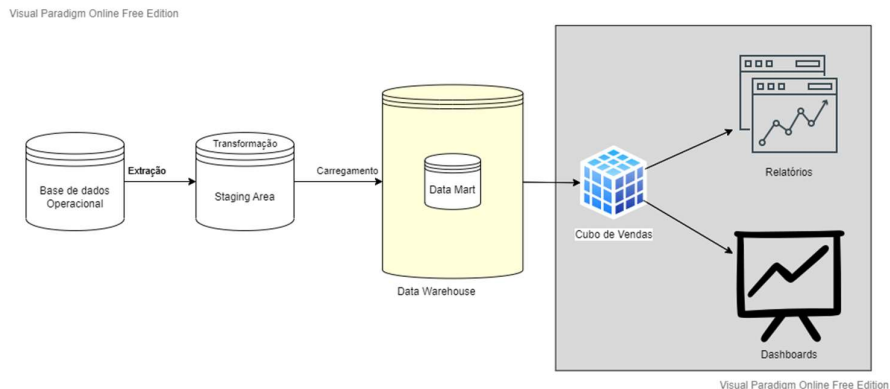


Figura 48: Etapa de análise de dados na opção de arquitetura implementada

Para ser possível visualizar os dados do Cubo desenvolvido podem ser usadas duas ferramentas, o Excel ou o Power BI da Microsoft. A primeira permite ao utilizador estabelecer ligação com o Cubo de Vendas, de uma forma simples e obter os seus dados para realizar as suas análises. O Power BI permite obter os dados do Cubo de Vendas e criar *reports* e *dashboards* para posteriormente serem apresentados ao utilizador final, possibilitando a este retirar conclusões e ajudar na tomada de decisão.

No capítulo seguinte são ilustradas algumas das análises que foram criadas a partir do Cubo de dados.



## 6. Apresentação e Análise de Resultados

No presente capítulo são apresentadas algumas análises que foram elaboradas com recurso à solução desenvolvida.

As análises implementadas têm como objetivo apoiar o processo de tomada de decisão do utilizador, sendo expectável que as mesmas sejam de fácil compreensão e permitam a assimilação da informação.

### 6.1 Dashboards

Um *dashboard* é essencialmente a interface gráfica do sistema que permite aos utilizadores terem acesso de uma forma visual e centralizada a indicadores de desempenho e métricas. Os *dashboards* foram desenvolvidos de forma a responder às necessidades dos utilizadores da organização, necessidades essas abordadas ao longo da presente dissertação. As subsecções seguintes ilustram os diversos *dashboards* que foram criados com os dados obtidos do cubo que contém a informação acerca das vendas. É de salientar que a informação apresentada nas diversas análises está filtrada de modo ser possível visualizar uma amostra dos dados, evitando assim a divulgação de informação considerada sensível por parte da organização.

#### 6.1.1 Análise de vendas por categoria de produto

Na Figura 49 encontra-se a visualização do *dashboard* que contém informação relativa a categoria do produto.

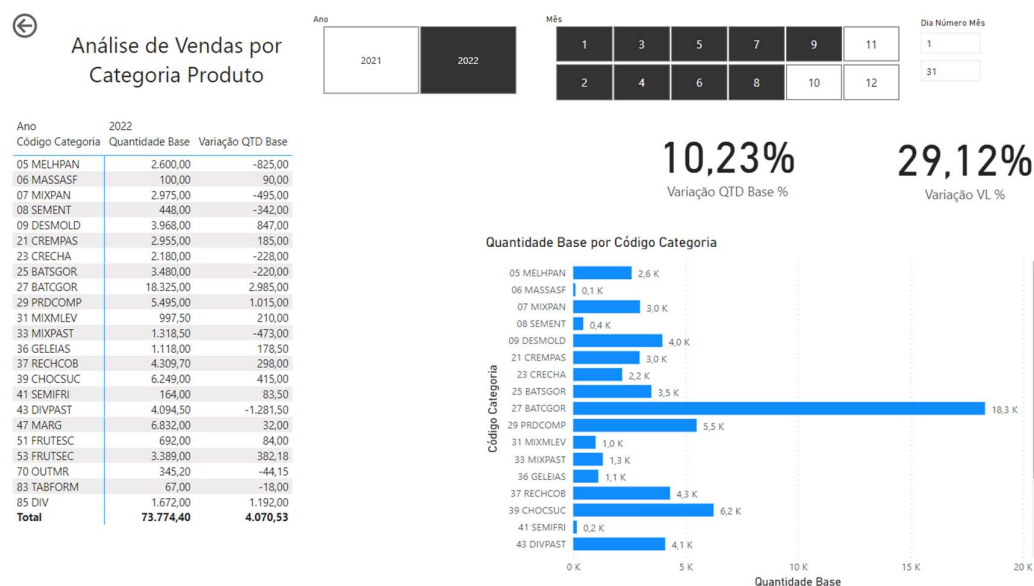


Figura 49: Análise de vendas por categoria de produto

Neste *dashboard* os dados estão filtrados ao nível do ano e tipo de cliente sendo possível visualizar a quantidade base por categoria de produto dos últimos dois anos de um determinado tipo de clientes. No topo do *dashboard* é possível selecionar qual o período que se pretende analisar no que diz respeito ao ano, mês e dia. Como se pode observar é apresentada a quantidade base e a variação face ao período homólogo do ano anterior por categoria de produto em formato tabular. Também pode-se ver a variação percentual da quantidade base e do valor líquido. Pode-se verificar que a categoria de produto com maior quantidade é a categoria 27 BATCGOR com 18.325 kg e uma variação de 2.985 kg. A variação percentual da quantidade é de 10,23% e a variação percentual do valor líquido é de 29,12%. Isto para o ano de 2022 de janeiro a setembro. No *dashboard* também se tem presente um gráfico de barras com as respetivas quantidades por categoria de produto, permitindo fazer o *drill down* para consultar quais as quantidades dos produtos de cada categoria. Na Figura 50 pode-se visualizar o *drill down* da categoria de produto 27 BATCGOR.

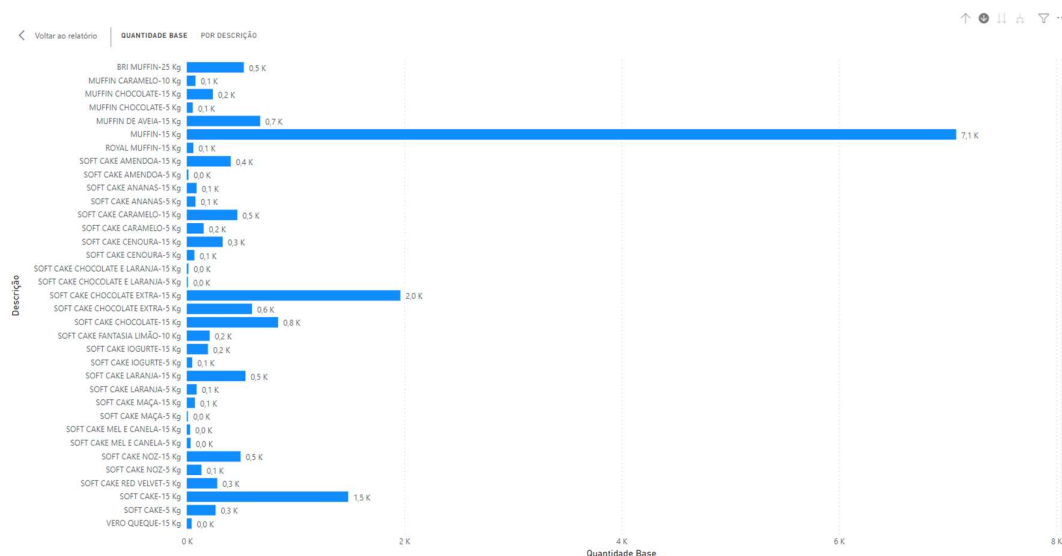


Figura 50: *Drill down* quantidade base por categoria de produto

Pode-se visualizar que o produto Muffin-15 Kg é o produto com maior quantidade vendida no período de janeiro a setembro de 2022. Os botões no canto superior do gráfico permitem navegar entre os diversos níveis de detalhe. Quando se passa com o rato por cima de cada uma das barras é mostrado o respetivo detalhe para ser possível saber qual a categoria ou o produto a que diz respeito.

### 6.1.2 Análise de vendas por cliente

O *dashboard* de análise de vendas por cliente permite perceber tal como o próprio nome indica as vendas de um determinado cliente num determinado período. Tal como no *dashboard* anterior a informação também se encontra filtrada ao nível do ano e tipo de cliente. A Figura 51 representa a análise de vendas por cliente.



Figura 51: Análise de vendas por cliente

Este é um *dashboard* que inclui dois elementos visuais em formato tabular que contém alguma informação que permite identificar e caracterizar melhor o cliente. Na tabela da esquerda da Figura 50 encontramos a identificação do cliente, o respetivo comercial, o tipo de cliente, a rota a quantidade base, a variação da quantidade base e a variação percentual da quantidade base. Enquanto na tabela da direita é apresentada a informação relativa aos produtos e respetiva quantidade. Estes dois visuais tabulares permitem saber quais os produtos que um determinado cliente compra assim como quais os clientes que compram um determinado produto. Isto porque as duas tabelas são reativas entre si, quando selecionado um cliente é mostrado na tabela da direita quais os produtos que ele compra, assim como quando selecionado um produto é visualizado na tabela da esquerda os clientes que o compram.

Esta análise permite perceber por exemplo que o cliente Padaria e Pastelaria Sebolido, Lda do comercial Mário Miranda adquiriu 4.183 kg, o que representa um aumento de quantidade de 1.640 kg e uma variação percentual de quantidade de 64,46% no período entre janeiro e setembro de 2022. Permite identificar também os produtos que este cliente adquiriu, neste caso foram 10 produtos, sendo o produto P.C. Tipo 2 Brioche-25 kg o de maior quantidade com 3.500 kg.

### 6.1.3 Análise de vendas por tipo cliente

Esta é uma análise que permite perceber qual a distribuição das vendas por tipo de cliente. De referir também que este *dashboard* também se encontra filtrado ao nível do ano e tipo de cliente. Na Figura 52 temos uma representação da análise de vendas por tipo de cliente.

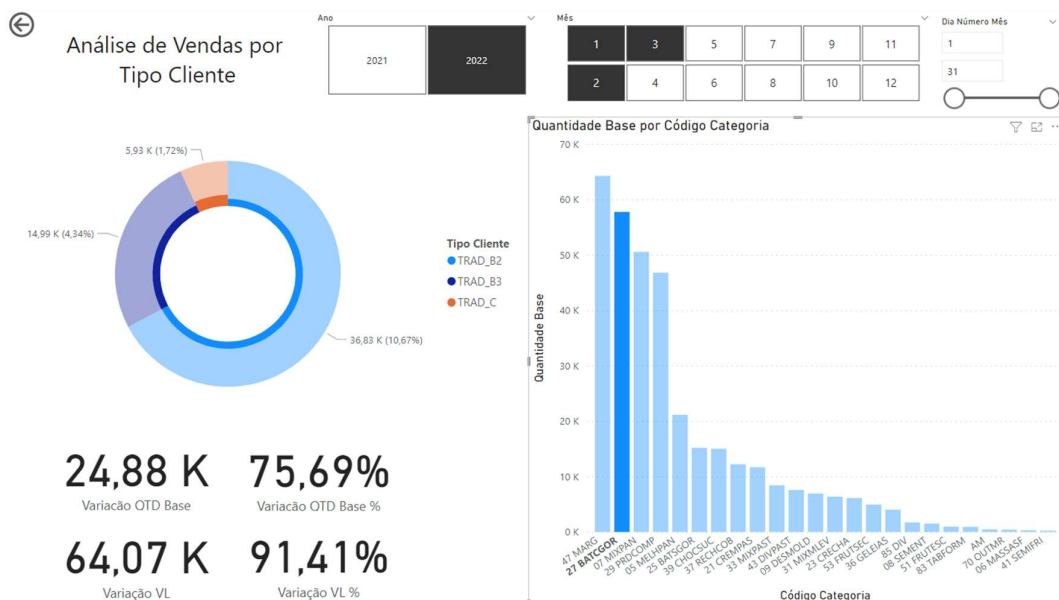


Figura 52: Análise de vendas por tipo de cliente

No topo da página do *dashboard* pode-se filtrar a informação que se pretende visualizar por ano, mês e dia. Contém um gráfico em forma de anel com a distribuição da quantidade vendida por tipo de cliente, um gráfico de barras com a quantidade por categoria de produto, um elemento com a variação da quantidade base, um elemento com a variação percentual da quantidade base, um elemento com a variação do valor líquido e um elemento com a variação percentual do valor líquido.

Este *dashboard* mostra a informação relativa ao período que decorre entre janeiro e março de 2022. Verifica-se que nesse período a distribuição de quantidade por tipo de cliente para a categoria de produto 27 BATCGOR foi de 5.930 kg em clientes do tipo TRAD\_C, 14.990 Kg em clientes do tipo TRAD\_B3 e de 36.830 kg em clientes do tipo TRAD\_B2. A quantidade vendida nestes três tipos de cliente da categoria 27 BATCGOR foi de cerca de 57.800 kg o que representa um aumento de quantidade de 24.880 kg (75,69%) e um aumento de valor líquido de 64.070€ (91,41%) face ao período homologo do ano anterior.

#### 6.1.4 Análise de referências por cliente

Na Figura 53 pode-se visualizar a análise de referências por cliente. O *dashboard* está também filtrado ao nível do tipo de cliente e ano.

No topo da página é possível filtrar a informação a visualizar por ano, mês e dia. Este *dashboard* é constituído principalmente por dois visuais tabulares. Um deles com a identificação do cliente, o nº de referências distintas, a quantidade base e a variação percentual da quantidade base. O outro visual tabular contém a identificação do comercial, a quantidade base, a variação percentual da quantidade base e o nº de referências distintas. Possui ainda um indicador do nº de clientes e um indicador com o nº de documentos.

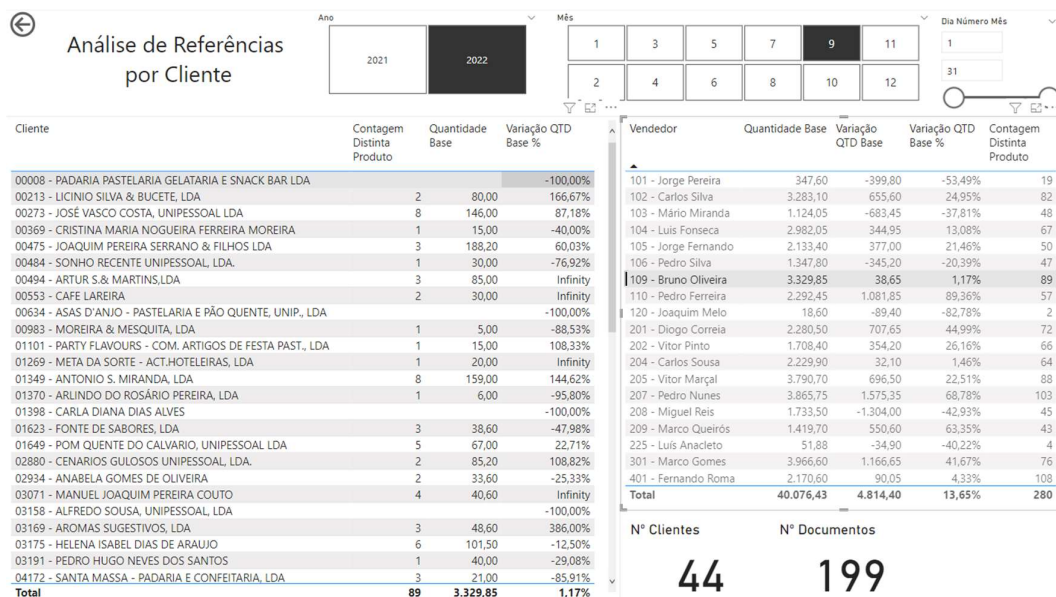


Figura 53: Análise de referências por cliente

Com este *dashboard* pode-se apurar que o comercial Bruno Oliveira no mês de setembro de 2022 vendeu uma quantidade de 3.329,85 kg, com um aumento de 38,65 kg, o que representa uma variação de 1,17% face ao período homólogo do ano anterior e 89 referências distintas. Pode-se ainda verificar que nesse período teve 44 clientes a comprar e emitidos 199 documentos de faturação. Para o mesmo comercial, apurar quais foram os clientes ativos, o nº de referências que cada um adquiriu, a quantidade base e a variação percentual. Verifica-se por exemplo o cliente José Vasco Costa Unipessoal Lda adquiriu 8 referências, o que representou 146 kg e uma variação de 87,18% face ao período homólogo.

### 6.1.5 Análise de vendas por localização

Esta análise permite perceber qual a distribuição de quantidade vendida por localização. Tal como os *dashboards* anteriores, este *dashboard* está filtrado ao nível do ano e tipo de cliente. Na Figura 54 pode-se visualizar a análise de vendas por localização.

No topo da página deste *dashboard* é possível selecionar qual o período a analisar, quer seja ao nível do ano, mês ou dia. Neste *dashboard* existem alguns elementos visuais como um visual tabular com a identificação de cada localização, a quantidade base e a variação. Têm também um gráfico anelar com a distribuição da quantidade base por localização e um gráfico de barras horizontais com a quantidade por categoria de produto e o produto, permitindo assim fazer o *drill down* pela categoria até ao produto.

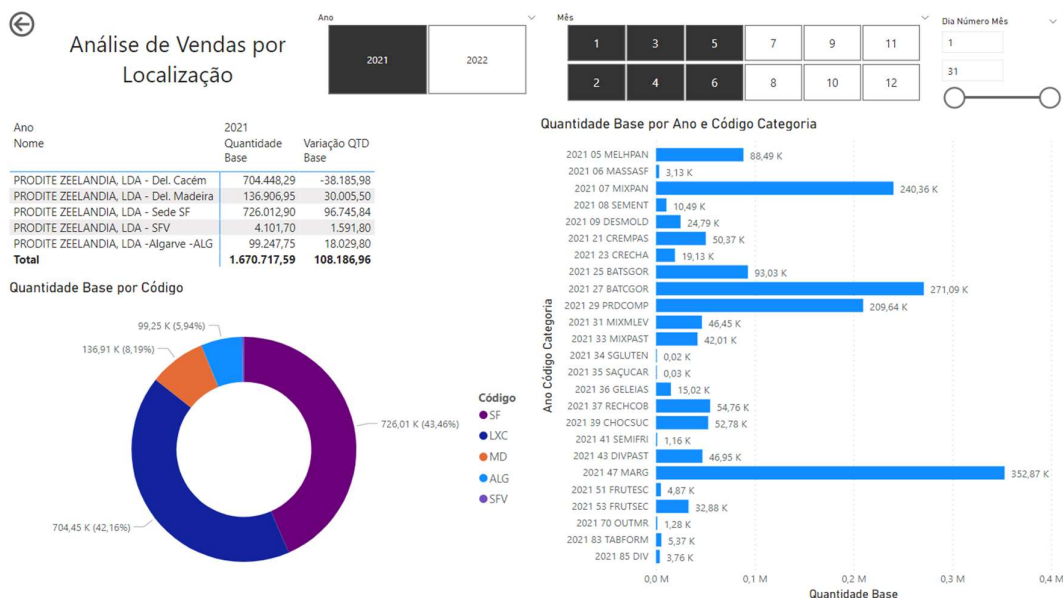


Figura 54: Análise de vendas por localização

Com esta análise pode-se estabelecer uma comparação da quantidade vendida entre as diversas localizações num determinado período. Neste *dashboard* podemos concluir que a localização do Algarve no período de janeiro a junho de 2021 teve um volume de vendas de 99.247,75 kg o que representa um aumento de 18.029,80 face ao período homólogo do ano anterior. Verifica-se também que no mesmo período a quantidade vendida pela sede e pela delegação de Lisboa (Cacém) foi muito semelhante, com 43,46% e 42,16% respetivamente. Por fim de acordo com o gráfico de barras, verifica-se que a quantidade vendida da categoria 47 MARG foi de 352.870 kg. O *dashboard* permite selecionar uma das localizações para filtrar o gráfico de barras com a quantidade base por categoria de produto e efetuar o *drill down* até aos produtos.

### 6.1.6 Análise de vendas por distribuidor

Com esta análise é possível aferir qual a quantidade entregue por cada distribuidor. Na Figura 55 encontra-se ilustrada a análise de vendas por distribuidor. Este *dashboard* tal como os anteriores está filtrado ao nível do tipo de cliente e ano.

O *dashboard* permite filtrar o período em análise recorrendo aos controlos existentes no topo da página. É constituído por um gráfico em forma de anel e por uma lista que permite filtrar o tipo de distribuidor, ou seja, se é distribuição interna, externa, comercial, técnico ou outro.

De acordo com o gráfico, verifica-se que no período compreendido entre janeiro e junho de 2022 o distribuidor Transportes Francisco Oliveira foi responsável por cerca de 74,4% das entregas (33% + 21,14% + 15,73% + 4,49%), assim como o distribuidor Expresserpa foi responsável por 17,68% das entregas.



Figura 55: Análise de vendas por distribuidor

### 6.1.7 Análise de vendedor

Esta análise pretende auxiliar a medição do desempenho de um comercial. A análise de vendedor encontra-se representada na Figura 56. Tal como nos *dashboards* anteriores este *dashboard* encontra-se filtrado ao nível do ano e do tipo de cliente.

No topo do *dashboard* encontram-se elementos que permitem filtrar qual o período que é pretendido analisar (ano, mês e dia) e assim como filtrar o supervisor. O *dashboard* é composto por vários elementos, tais como uma tabela com a identificação do vendedor, identificação dos clientes, a quantidade base, a variação da quantidade base e a variação percentual da quantidade base. Contém um gráfico de barras com a quantidade base e a quantidade base do ano anterior por categoria de produto e produto. Contém também um elemento visual com o nº de clientes, outro com o nº de produto e outro com o nº de documentos.

Neste *dashboard* ao selecionar um vendedor da tabela consegue-se perceber qual o volume de vendas para o período selecionado e para o período homólogo. Ao analisar-se o *dashboard* verifica-se por exemplo que o comercial Jorge Fernando para o período compreendido entre janeiro e março de 2021 vendeu 4.696,60 kg, o que representa uma diminuição de -350,65 kg e uma variação negativa de -6,95%. Verifica-se também que no período em questão vendeu em 35 clientes distintos, 72 produtos diferentes em 221 documentos de faturação. Também se pode concluir que existiu uma diminuição acentuada na categoria 27 BATCGOR e um aumento significativo na categoria 47 MARG.

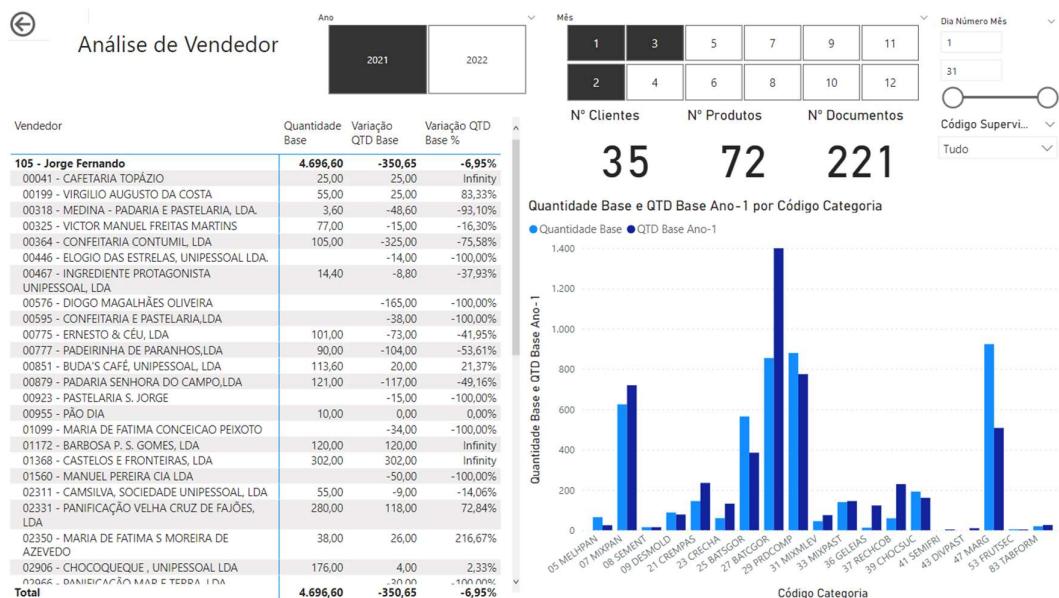


Figura 56: Análise de vendedor

Este *dashboard* permite ainda realizar o *drill down* desde o vendedor até ao produto, para que dessa forma seja possível identificar os clientes de um comercial e quais os produtos que esses clientes compram. A Figura 57 representa o resultado de um *drill down* no *dashboard*.

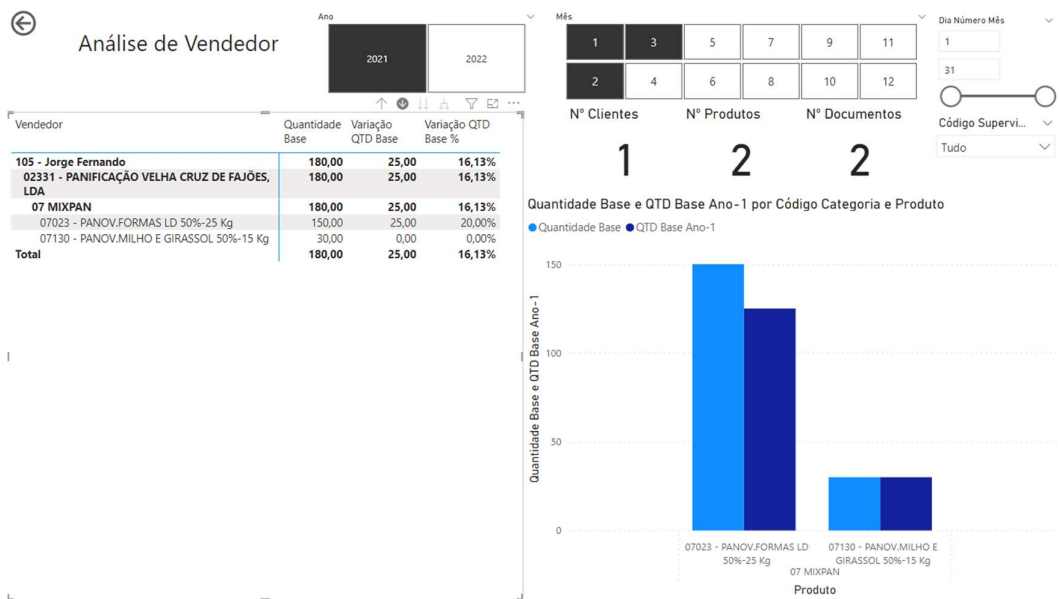


Figura 57: Drill down análise vendedor

Tal como se pode visualizar na imagem é possível verificar que o cliente Panificação Velha Cruz de Fajões Lda do comercial Jorge Fernando, no período de janeiro a março de 2021 adquiriu 150 kg do produto PANOV. FORMAS LD 50%-25 kg e 30 kg do produto PANOV. MILHO E GIRASSOL 50%-15 Kg.

Se for necessário voltar aos níveis acima basta utilizar os botões em forma de seta que se encontram no topo direito do elemento visual.

### 6.1.8 Análise por equipa

A análise por equipa permite auxiliar na medição de desempenho de um conjunto de comerciais assim como o respetivo supervisor. Tal como os anteriores *dashboards*, este está filtrado ao nível do ano e do tipo de cliente. Na Figura 58 pode-se visualizar a representação da análise por equipa.

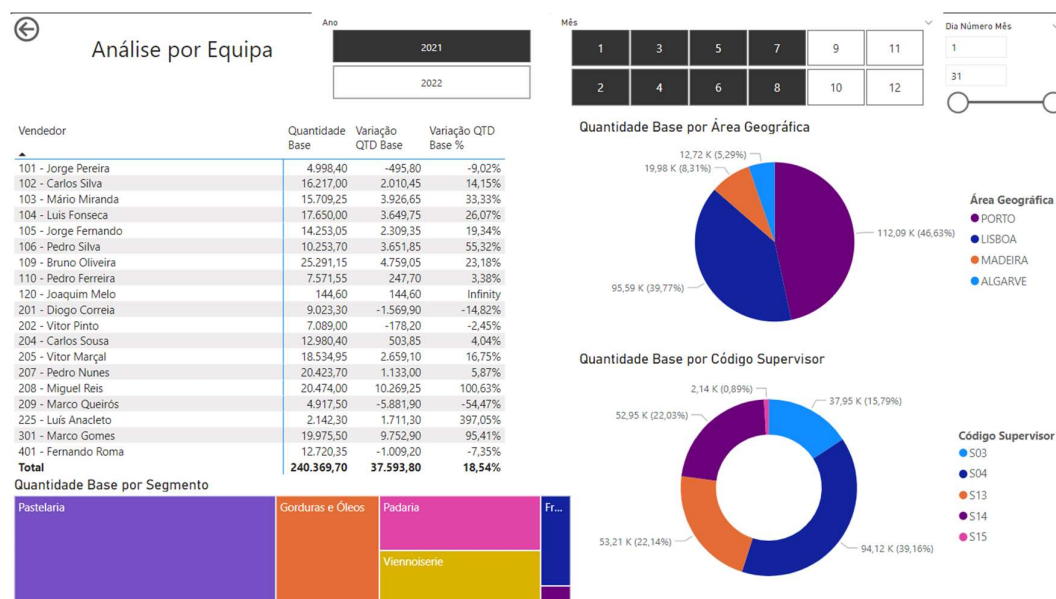


Figura 58: Análise por equipa

A análise por equipa é composta por vários elementos. No topo da página estão disponíveis os elementos que permitem filtrar a informação por período (ano, mês e dia). A *dashboard* é composta por elemento tabular com a informação dos vendedores, quantidade, variação e percentual de variação. Contém também um gráfico circular com a quantidade por área geográfica do comercial, um gráfico em anel com a quantidade por código de supervisor e ainda um *treemap* com a quantidade por segmento de produto.

Ao analisar o *dashboard* verifica-se para o período compreendido entre janeiro e agosto de 2021 as equipas do Porto foram responsáveis por 46,63% das vendas e que a equipa do supervisor S04 teve a percentagem mais elevada das vendas com 39,16%. Pode-se concluir também que os produtos de pastelaria representaram cerca de metade da quantidade vendida. Outra conclusão que se pode verificar é que o comercial Bruno Oliveira foi o que teve maior volume de vendas com 25.291,15 kg.

## 6.2 Análises Excel

Para além dos *dashboards* desenvolvidos no Power BI, a solução implementada permite que se utilize o Excel para realizar análises.

Um utilizador com permissões de acesso ao Analysis Services pode obter facilmente os dados do cubo e construir análises no Excel.

Na figura 59 é possível visualizar um exemplo de uma dessas análises construída em Excel.

Cód Fornecedor	(Itens múltiplos)						
Quantidade Base	Rótulos de Coluna						Total Geral
Rótulos de Linha	1	2	3	4	5	6	
MR							
BE	102	84	4386	1506	168	150	6396
CZ	14816	14328	14834	15580	14818	14804	89180
ES	66831	87552	133068	87533	100155	95071	570210
IE	11484	7572	12672	7380	18660	26328	84096
IT	1	5	4	1	6	497	514
NL	26210,3	25520,4	33548,1	26684,5	31762,7	31520,8	175246,8
PT	11335,5	9197	12092,5	10320,5	12257	14461,5	69664
PA							
NL	1099	96	115	1104,5	711,5	1852,5	4978,5
PT	128	25	103	153	153	231	793
<b>Total Geral</b>	<b>132006,8</b>	<b>144379,4</b>	<b>210822,6</b>	<b>150262,5</b>	<b>178691,2</b>	<b>184915,8</b>	<b>1001078,3</b>

Figura 59: Exemplo de análise realizada em Excel

Na imagem pode-se visualizar as vendas por código de país de origem do fornecedor, no período de janeiro a junho de 2022. Nesta análise a informação revelada é diminuta por se tratar de informação sensível, mas a mesma serve para demonstrar outra alternativa para a construção de análises.

## 7. Avaliação de Resultados

Este capítulo é dedicado à apresentação da forma como será realizada a avaliação do atual projeto nas diversas grandezas identificadas. Assim, o capítulo inicia com a identificação das grandezas que devem ser avaliadas e o que é pretendido avaliar em cada uma. Segue-se a apresentação da metodologia de avaliação para cada uma das grandezas, finalizando com a apresentação das experiências a realizar e quais os resultados esperados.

### 7.1 Grandezas em Avaliação

Tendo em conta o presente projeto e o conjunto de funcionalidades que a solução propõe, as grandezas a considerar dizem respeito ao tempo necessário para integração dos dados e atualização do cubo, à exatidão e confiabilidade desses mesmos dados e à avaliação dos utilizadores quanto à satisfação com a solução apresentada.

O processo de ETL e atualização do cubo de dados é um processo que deverá ocorrer diariamente, preferencialmente após o fim do horário laboral. Nesse período são executadas uma série de rotinas na base de dados operacional, de maneira que todo o processo de atualização dos dados não deverá ter uma duração que possa perturbar esses mesmos processos.

A autenticidade dos valores envolve a necessidade de apresentar valores confiáveis aos utilizadores para dessa forma possibilitar uma correta análise dos indicadores de desempenho da organização e a sua evolução ao longo do tempo.

Por outro lado, terá de se ter em conta a satisfação do utilizador uma vez que a solução propõe agilizar o processo de criação de análises de dados, assim como a sua posterior visualização, tornando a simplicidade e intuição fatores fundamentais para minimizar o esforço necessário à implementação da metodologia e da utilização da própria solução no dia-a-dia da organização.

### 7.2 Metodologias de Avaliação

No geral, o sistema foi testado com vários testes unitários e de integração para verificar se cada recurso está implementado corretamente.

De forma a avaliar o tempo despendido durante todo o processo de integração e atualização, foram realizadas várias execuções do mesmo de forma a obter uma média do tempo de execução. O objetivo é permitir compreender qual a duração esperada e qual o melhor horário de forma a não interferir com nenhum dos outros processos que ocorrem no mesmo período.

Os dados do sistema desenvolvido foram validados com recurso aos mapas já existentes no ERP. O objetivo é perceber se os dados do sistema desenvolvido estão conforme e são confiáveis para uso na criação das diversas análises.

Para avaliar a satisfação dos utilizadores com a utilização global do sistema, foi enviado um inquérito de satisfação a um grupo de utilizadores para obter uma avaliação das várias funcionalidades, aplicando um sistema de classificação a cada questão do inquérito. Na secção seguinte são descritos os cálculos necessários para verificar a classificação da solução.

O grupo de utilizadores que responderam ao inquérito fazem parte do departamento comercial da organização. A estes utilizadores foram apresentadas algumas das análises desenvolvidas com a informação das vendas. A opinião destes utilizadores é importante pois permite perceber se o sistema desenvolvido corresponde as suas necessidades.

De forma a permitir uma fácil e rápida disseminação dos inquéritos, o mesmo foi realizado e disponibilizado online. No inquérito de satisfação foi usado um grupo de afirmações que foram classificadas com recurso à escala de Likert, para se perceber qual o nível de concordância do utilizador com as mesmas.

A escala de Likert (SurveyMonkey, setembro 2022) permite medir o grau de satisfação do utilizador com recurso a níveis de classificação. Para medir o grau de satisfação dos utilizadores com o sistema implementado foram aplicados cinco níveis de satisfação, de forma a ser possível medir a concordância dos mesmos com as afirmações. No inquérito de satisfação foi utilizada a escala representada na Tabela 14.

Tabela 14: Escala de Likert

Escala	Descrição
1	Discordo totalmente
2	Discordo
3	Não concordo nem discordo
4	Concordo
5	Concordo totalmente

Na Figura 60 pode-se ver a representação do inquérito com 6 questões e as respetivas opções de resposta que foi enviado aos utilizadores.

### Inquérito de satisfação com novo sistema de BI

1. A aplicação é de fácil utilização?
- Discordo totalmente  Concordo
- Discordo  Concordo totalmente
- Não concordo nem discordo
2. O layout das análises apresentadas é mais apelativo do que o atual sistema?
- Discordo totalmente  Concordo
- Discordo  Concordo totalmente
- Não concordo nem discordo
3. As análises apresentadas são relevantes?
- Discordo totalmente  Concordo
- Discordo  Concordo totalmente
- Não concordo nem discordo
4. A informação apresentada é de fácil leitura e compreensão?
- Discordo totalmente  Concordo
- Discordo  Concordo totalmente
- Não concordo nem discordo
5. Considera o novo sistema uma ferramenta útil de apoio à decisão?
- Discordo totalmente  Concordo
- Discordo  Concordo totalmente
- Não concordo nem discordo
6. Considera que o sistema desenvolvido cumpre os objetivos propostos?
- Discordo totalmente  Concordo
- Discordo  Concordo totalmente
- Não concordo nem discordo

Concluído

Figura 60: Inquérito enviado aos utilizadores

## 7.3 Resultados obtido dos testes unitários

Foram realizados dois tipos de testes unitários, testes de conformidade dos dados e testes de tempo de execução da integração e atualização dos dados.

### 7.3.1 Resultados dos testes de conformidade

De forma a verificar a conformidade dos dados foram realizadas diversas comparações entre os dados obtidos do ERP e os dados obtidos da solução implementada.

Na Figura 61 é possível visualizar uma demonstração de uma comparação entre os valores apresentados no ERP e os valores obtidos do sistema implementado.

Análise Vendas - Família

Filtros 01-01-2020 31-01-2020

Nome	Grupo	Ano	Varição	Var %	Ano-1
05 Melhorantes p/ Panificação		25 470	-140	-0,55	25 610
06 Massas Fermentadas		14 040	1 428	11,32	12 613
07 Mixes p/ Panificação		261 775	-12 740	-4,64	274 515
08 Sementes e Decorações		13 434	-767	-5,40	14 201
09 Desmoldantes e Lubrif.		13 620	-41	-0,30	13 662
21 Cremes Pasteleiros e Simil.		16 160	-1 810	-10,07	17 970
23 Cremes tipo Chantilly		12 286	-10 518	-46,12	22 804
25 Batidos sem Gordura		50 913	8 238	19,30	42 675
27 Batidos com Gordura		120 265	12 800	11,91	107 465
29 Produtos Compostos		123 655	15 535	14,37	108 120
31 Mixes para Massas Lêvedas		11 160	-2 385	-17,61	13 545
33 Mixes para Pastelaria		31 170	10 102	47,95	21 069
34 Produtos sem Glúten		0	0	0,00	0
35 Produtos sem adição Açúcar		1 710	1 125	192,31	585
36 Brilhos e Geleias		8 756	-163	-1,83	8 919
37 Recheios e Coberturas		14 923	-3 499	-18,99	18 422
39 Chocolates, Sucedenãos e Cre...		10 916	-2 102	-16,14	13 018
41 Semi-Frios e Pastas		325	42	14,64	284
43 Diversos para Pastelaria		8 775	2 026	30,01	6 750
47 Margarinas		91 644	12 444	15,71	79 200
51 Frutas Escorridas		2 600	-860	-24,86	3 460
53 Frutas Secas		19 036	-967	-4,83	20 003
70 Outras Mercadorias		408	-44	-9,82	452
83 Tabuleiros e Formas		1 909	-25	-1,29	1 934
84 Artigos para Decoração		1	-8	-88,89	9
85 Diversos		961	-167	-14,80	1 128
Centro Tecnológico		0	0	0,00	0
Descontos Diversos		0	0	0,00	0
TOTAL		855 911	27 501	3,32	828 410

Rótulos de Linha	Quantidade Base	Varição QTD Base %	QTD Base Ano-1
05 MELHPAN	25470	-0,55%	25610
06 MASSASF	14040	11,32%	12612,5
07 MIXPAN	261775	-4,64%	274515
08 SEMENT	13433,5	-5,40%	14200,5
09 DESMOLD	13620,4	-0,30%	13661,6
21 CREMPAS	16160	-10,07%	17970
23 CRECHA	12286	-46,12%	22804
25 BATSGOR	50912,5	19,30%	42675
27 PRDTCOMP	120265	11,91%	107465
29 PRDTCOMP	123655	14,37%	108120
31 MIXMLEV	11160	-17,61%	13545
33 MIXPAST	31170	47,95%	21069,5
35 SAÇUCAR	1710	192,31%	585
36 GELEIAS	8756	-1,83%	8919
37 RECHCOB	14923,05	-18,99%	18422,2
39 CHOCSUC	10916	-16,14%	13017,5
41 SEMIFRI	325	14,64%	283,5
43 DIVPAST	8775	30,01%	6749,5
47 MARG	91644	15,71%	79200
51 FRUTESC	2600	-24,86%	3460
53 FRUTSEC	19036	-4,83%	20003,04
70 OUTMR	408	-9,82%	452,45
83 TABFORM	1909	-1,29%	1934
84 ARTDEC	1	-88,89%	9
85 DIV	961	-14,80%	1128
Total Geral	855911,45	3,32%	828410,29

Figura 61: Comparação valores ERP e valores obtidos do cubo de vendas

Analisando os dados apresentados pode-se concluir que os dados do cubo de vendas estão em conformidade com os dados apresentados pelo ERP. Isto é muito importante porque garante que o sistema implementado permite obter dados confiáveis e credíveis. O exemplo apresentado é representativo dos vários testes que foram realizados aos dados.

Conclui-se assim que objetivo de desenvolver um sistema que permitisse obter dados confiáveis foi atingido.

### 7.3.2 Resultados dos testes do tempo de execução da atualização de dados

Para ser possível aferir qual o tempo de execução compreendido entre o início do processo de integração e o término da atualização do cubo de dados, foram realizadas 6 repetições do processo e registados os tempos de início e fim de cada uma. Na Tabela 15 podemos ver um resumo das medições realizadas e respetivo tempo médio de execução.

Tabela 15: Tempos de execução do processo de integração e atualização dos dados

Nº Execução	Hora Inicial	Hora Final	Duração em minutos
1	14:02:52	14:07:24	04:33
2	14:09:42	14:14:03	04:23
3	14:15:18	14:20:07	04:50
4	14:23:34	14:28:27	04:54
5	14:29:58	14:34:33	04:35
6	14:38:43	14:42:58	04:15
		<b>Média</b>	<b>04:35</b>

Consultando a Tabela 15 verifica-se que a duração média de execução de todo o processo de integração e atualização de dados é de 4 minutos e 35 segundos. Podemos concluir que o objetivo definido para a duração máxima deste processo foi atingido, isto porque o tempo que foi definido como limite máximo foi de 15 minutos e a média das execuções ficou bem abaixo, situando-se nos 4 minutos e 35 segundos.

## 7.4 Resultados obtido do inquérito aos utilizadores

O inquérito de satisfação realizado aos utilizadores do sistema contou com a participação de 18 participantes, tendo este grupo demonstrado satisfação com o sistema desenvolvido.

O resultado obtido a cada uma das questões colocadas aos utilizadores está disponível no Anexo 4. Na Tabela 16 é possível consultar um resumo da frequência de resposta a cada uma das afirmações incluídas no inquérito de satisfação.

Tabela 16: Frequência de respostas dos utilizadores

Nível	1	2	3	4	5
Afirmação					
1	0% (0)	0% (0)	5,56% (1)	44,44% (8)	50% (9)
2	0% (0)	0% (0)	0% (0)	66,67% (12)	33,33% (6)
3	0% (0)	0% (0)	11,11% (2)	66,67% (12)	22,22% (4)
4	0% (0)	0% (0)	38,89% (7)	44,44% (8)	16,67% (3)
5	0% (0)	0% (0)	33,33% (6)	44,44% (8)	22,22% (4)
6	0% (0)	0% (0)	11,11% (2)	83,33% (15)	5,56% (1)

Ao realizar a análise dos resultados verifica-se que, na primeira afirmação 9 participantes classificaram a mesma com o nível 5 (Concordo totalmente), 6 classificaram com o nível 4 (Concordo) e 1 participante classificou com o nível 3 (Não concordo nem discordo). Por esse motivo é possível concluir que a aplicação é de fácil utilização.

Quanto à segunda questão, 6 participantes optaram pelo nível 5 (Concordo totalmente), 12 participantes optaram pelo nível 4 (Concordo). Pode-se concluir que para a maioria dos utilizadores as análises desenvolvidas são mais apelativas do que as existentes atualmente no ERP.

Relativamente à terceira questão, 4 participantes classificaram com o nível 5 (Concordo totalmente), 12 classificaram com o nível 4 (Concordo) e 2 classificaram com o nível 3 (Não concordo nem discordo). Desta forma é possível afirmar que grande parte dos inquiridos concordam que as análises apresentadas são relevantes.

No que diz respeito à quarta questão, 7 participantes escolheram o nível 3 (Não concordo nem discordo), 8 participantes escolheram o nível 4 (Concordo) e 3 participantes escolheram o nível

5 (Concordo totalmente). Por esse motivo é possível afirmar que mais de metade dos inquiridos concordam que a informação apresentada é de fácil leitura e compreensão.

Na questão 5, 8 participantes classificaram com o nível 4 (Concordo), 6 participantes classificaram com o nível 3 (Não concordo nem discordo) e 4 participantes classificaram com o nível 5 (Concordo totalmente). Também nesta questão é possível afirmar que mais de metade dos inquiridos considera o novo sistema uma ferramenta útil de apoio à decisão.

Por fim, na sexta e última questão, 15 participantes optaram pelo nível 4 (Concordo), 2 participantes optaram pelo nível 3 (Não concordo nem discordo) e um participante optou pelo nível 5 (Concordo totalmente). É assim possível concluir que os inquiridos consideram que o sistema desenvolvido cumpre os objetivos estabelecidos.

De uma forma geral e com base nas respostas dos utilizadores ao inquérito de satisfação, que avaliaram a solução desenvolvida, é possível afirmar que o grau de satisfação é elevado.

## 8. Conclusão

O presente capítulo inclui as conclusões que foram adquiridas ao longo desta dissertação. Nele são referidos os objetivos atingidos, as dificuldades sentidas durante a realização da mesma, finalizando com o trabalho futuro a ser desenvolvido sustentado no trabalho já desenvolvido no âmbito desta dissertação.

### 8.1 Objetivos Atingidos

O objetivo desta dissertação era o desenvolvimento de um sistema de apoio à decisão baseado em *Business Intelligence* que permitisse à organização gerir a informação que possui sobre as vendas e que possibilitasse o fornecimento de informação aos utilizadores, servindo assim de apoio no processo de tomada de decisão.

A presente dissertação foi realizada numa organização ligada à indústria alimentar, tendo os dados das vendas da mesma servido como base para o caso de estudo.

No decorrer desta dissertação, após a contextualização do problema, foram apresentados conceitos, componentes e arquiteturas de *Business Intelligence*. Foram também estudadas uma série de ferramentas de ETL e de BI com o intuito de se perceber quais seriam as mais apropriadas às necessidades apresentadas. O Microsoft SQL Server Integration Services foi a ferramenta escolhida para realizar o processo de ETL e o Microsoft Power BI foi escolhido como ferramenta de BI. A utilização conjunta destas duas ferramentas permitiu desenvolver uma solução com uma boa performance e visualmente apelativa para o utilizador.

A solução desenvolvida possui uma arquitetura que permite facilmente realizar modificações, quer seja a alteração de atributos, adição de novas medidas ou até a integração de diferentes fontes de dados e ainda a capacidade de guardar histórico.

Com o desenvolvimento de *dashboards* intuitivos foi possível ao utilizador analisar os dados com mais facilidade, reter a informação necessária mais rapidamente e retirar conclusões de uma forma imediata.

Por último e com o objetivo de validar o cumprimento dos requisitos funcionais enunciados nesta dissertação, foram realizados testes unitários para aferir a confiabilidade dos dados e o tempo de atualização dos dados do sistema. Também foi realizado um inquérito de satisfação aos utilizadores de forma a obter uma opinião sobre o sistema desenvolvido.

Os resultados do inquérito de satisfação revelam que os objetivos propostos nesta dissertação foram cumpridos.

## 8.2 Dificuldades Sentidas

Uma das dificuldades sentidas ao longo de toda a dissertação foi o de conseguir estabelecer tempo para a realização da mesma em horário laboral. Por vezes tornou-se difícil dedicar tempo para o desenvolvimento da solução. A gestão do tempo foi a principal dificuldade sentida.

O facto de não existir ninguém dentro da organização com conhecimentos de BI. Seria sempre mais enriquecedor ter alguém que partilhasse algum do seu conhecimento.

## 8.3 Trabalho Futuro

Após o término desta dissertação existe uma série de trabalhos que podem vir a ser desenvolvidos, de forma a terminar e melhorar o trabalho que foi desenvolvido.

Ao longo desta dissertação todo o trabalho desenvolvido foi realizado num ambiente de testes, com acesso aos dados da base de dados operacional. Um dos primeiros trabalhos a realizar será o de transferir a solução desenvolvida do ambiente de testes para o ambiente de produção.

Os *dashboards* desenvolvidos podem e devem ser melhorados, assim como deve ser avaliada a possibilidade de criação de novos *dashboards*, de forma a permitir novas e melhores análises.

Seria interessante alargar o âmbito do sistema de BI a outras áreas da organização. Por exemplo a criação de um Data Mart com dados da produção ou da logística e criar um cubo de dados sobre os mesmos. Iria permitir a construção de *dashboards* também para essas áreas.

# Referências

- 1&1 IONOS (27 de dezembro de 2021). What is a data warehouse? Obtido de Web site Digital Guide: <https://www.ionos.com/digitalguide/online-marketing/web-analytics/dwh-what-is-a-data-warehouse/>.
- Belliveau, P. and Griffin, A. (2002) 'ToolBook for New Product Edited by'.
- Bi4all (27 de dezembro 2021). O Informatica PowerCenter como ferramenta de ETL. Obtido de Web Site de bi4all: <https://www.bi4all.pt/noticias/blog/powercenter-como-ferramenta-de-etl/>.
- Brito, E. (2015) Modelagem Dimensional, Consultor em TI.
- C. Howson, Successful Business Intelligence: Secrets to Making BI a Killer App, McGraw-Hill Osborne Media, 2008.
- CCG. (08 de dezembro de 2021). Business Intelligence: um bilhete de identidade das empresas?
- Cetax (02 de janeiro de 2022). Ferramenta ETL: Data Stage IBM. Obtido de Web Site de Cetax: <https://www.cetax.com.br/blog/datastage-ibm-ferramenta-de-etl/>.
- COEPD (30 de maio 2022). What is FURPS+? Obtido de Web Site de Center of Excellence for Professional Development: <https://businessanalysttraininghyderabad.wordpress.com/2014/08/05/what-is-furps/>.
- Dezordi, A. P. D. R. (2021). 'GESTÃO ESTRATÉGICA DE CUSTOS NO SEGMENTO DE MICROCERVEJARIAS ARTESANAIS: VARIÁVEIS DE PRECIFICAÇÃO DO PRODUTO dissertação Mestrado', *Unijuí – Universidade Regional Do Noroeste Do Estado Do Rio Grande Do Sul*, pp. 0–82. Elsevier Inc.
- DynamicsDocs (08 de outubro de 2022). Magic Quadrant for Analytics and Business Intelligence Platforms. Obtido de Web Site de DynamicsDocs: <https://dynamicsdocs.com/nav/2018/w1/table/item-category>.
- Elias, D. (2018) Arquitetura do Business Intelligence, BI na Prática.
- Gartner (28 de dezembro de 2021). Magic Quadrant for Data Integration Tools. Obtido de Web Site de Gartner: [https://www.gartner.com/doc/reprints?\\_\\_hstc=22819812.5c299c0a103ca66582272587c124308f.1644494217380.1644494217380.1644494217380.1&\\_\\_hssc=22819812.1.1644494217381&\\_\\_hsfp=3039884659&id=1-](https://www.gartner.com/doc/reprints?__hstc=22819812.5c299c0a103ca66582272587c124308f.1644494217380.1644494217380.1644494217380.1&__hssc=22819812.1.1644494217381&__hsfp=3039884659&id=1-)

27C89LW8&ct=210831&st=sb&utm\_source=web&utm\_medium=lp&utm\_campaign=Gartner%20MQ%202021&submissionGuid=cff26955-8965-4295-992d-85193b6cf2ca.

Gartner (02 de janeiro de 2022). Magic Quadrant for Analytics and Business Intelligence Platforms. Obtido de Web Site de Gartner: <https://www.gartner.com/doc/reprints?id=1-26QIXY7E&ct=210707&st=sb>.

Hitachi Vantara (04 de janeiro de 2022). Turn Data into Actionable Insights. Obtido de Web Site de Hitachi Vantara: <https://www.hitachivantara.com/en-us/products/data-management-analytics/pentaho/download-pentaho.html>.

Hitachi Vantara (02 de janeiro de 2022). Data Integration – Kettle. Obtido de Web Site de Hitachi Vantara: <https://community.hitachivantara.com/docs/DOC-1009855-data-integration-kettle>.

IBM (03 de janeiro de 2022). IBM DataStage. Obtido de Web Site de IBM: <https://www.ibm.com/br-pt/products/datastage>.

Informatica (05 de janeiro de 2022). PowerCenter Designer. Obtido de Web Site de Informatica: [https://docs.informatica.com/pt\\_pt/data-integration/powercenter/10-2/\\_guia-de-introducao\\_powercenter\\_10-2\\_ditamap/visao\\_geral\\_do\\_produto/cliente\\_do\\_powercenter/powercenter\\_designer.html](https://docs.informatica.com/pt_pt/data-integration/powercenter/10-2/_guia-de-introducao_powercenter_10-2_ditamap/visao_geral_do_produto/cliente_do_powercenter/powercenter_designer.html).

IntelliPaat (20 de fevereiro de 2022). Power BI Vs Tableau: 10 Key Differences. Obtido de Web Site de IntelliPaat: <https://intellipaat.com/blog/power-bi-vs-tableau-difference/>.

Javatpoint (15 de dezembro de 2021). What is Dimensional Modeling? Obtido de Web Site de Javatpoint: <https://www.javatpoint.com/data-warehouse-what-is-dimensional-modeling>.

Kimball, R., & Ross, M. (2013). *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*. Estados Unidos da América: John Wiley & Sons, Inc.

Koen, P. A. *et al.* (2002) 'Fuzzy Front End: and Techniques', *Industrial Research*, pp, pp. 5–35. Available at: [http://www.stevens.edu/cce/NEW/PDFs/FuzzyFrontEnd\\_Old.pdfNEW/PDFs/FuzzyFrontEnd\\_Old.pdf](http://www.stevens.edu/cce/NEW/PDFs/FuzzyFrontEnd_Old.pdfNEW/PDFs/FuzzyFrontEnd_Old.pdf).

Koen, P. et al. (2001) 'Providing clarity and a common language to the "fuzzy front end"', *Research Technology Management*, 44(2), pp. 46–55. doi: 10.1080/08956308.2001.11671418.

Lam, A. (2009) *Organizational Innovation, The Oxford Handbook of Innovation*. doi: 10.1093/oxfordhb/9780199286805.003.0005.

- LARION. (10 de dezembro de 2021). Components of Business Intelligence. Obtido de Web site da LARION: <https://larion.com/resource/whats-difference-business-intelligence-big-data/components-of-business-intelligence/>.
- Marins, C. S., Souza, D. de O. and Barros, M. da S. (2009) 'O Uso Do Método De Análise Hierárquica (AHP) Na Tomada De Decisões Gerenciais - Um Estudo', *Xli Sbp*, p. 11.
- Microsoft (05 de janeiro de 2022) SQL Server Integration Services. Obtido de Web site da Microsoft Docs: <https://docs.microsoft.com/en-us/sql/integration-services/sql-server-integration-services?view=sql-server-ver15> .
- Microsoft (04 de janeiro de 2022) O que é Power BI? Obtido de Web site do Power BI: <https://powerbi.microsoft.com/pt-pt/>.
- Microsoft (27 de dezembro de 2021). Overview of Service Manger OLAP cubes for advanced analytics. Obtido de Web site da Microsoft Docs: <https://docs.microsoft.com/pt-pt/system-center/scsm/olap-cubes-overview?view=sc-sm-2019>.
- MicroStrategy (05 de janeiro de 2022) Introdução ao MicroStrategy Workstation. Obtido de Web site da MicroStrategy: [https://www2.microstrategy.com/producthelp/Current/Workstation/WebHelp/Lang\\_1046/Content/About\\_MicroStrategy.htm](https://www2.microstrategy.com/producthelp/Current/Workstation/WebHelp/Lang_1046/Content/About_MicroStrategy.htm).
- Moscoso-Zea, O., Paredes-Gualtor, J. and Luján-Mora, S. (2018) 'A Holistic View of Data Warehousing in Education', *IEEE Access*, 6, pp. 64659–64673. doi: 10.1109/ACCESS.2018.2876753.
- Muhtaroglu, F. C. P. *et al.* (2013) 'Business model canvas perspective on big data applications', *Proceedings - 2013 IEEE International Conference on Big Data, Big Data 2013*, pp. 32–37. doi: 10.1109/BigData.2013.6691684.
- Nagy, J. *et al.* (2018) 'The role and impact of industry 4.0 and the internet of things on the business strategy of the value chain-the case of hungary', *Sustainability (Switzerland)*, 10(10). doi: 10.3390/su10103491.
- Navita (15 dezembro de 2021) O que é Gartner? Saiba sua importância no mundo corporativo! Obtido de Web site da Navita: <https://navita.com.br/blog/o-que-e-gartner-saiba-sua-importancia/>.
- Negash, S. (fevereiro de 2004). Business Intelligence. Communications of the Association for Obtido de Web site do CCG: <http://www.ccg.pt/business-intelligence-empresas/>
- Oracle (03 de janeiro de 2022). Business Analytics. Obtido de Web site da Oracle: <https://www.oracle.com/solutions/business-analytics/>

- Osterwalder, A. (2004) 'THE BUSINESS MODEL ONTOLOGY A PROPOSITION IN A DESIGN SCIENCE APPROACH', *Ecole des Hautes Etudes Commerciales de l'Université de Lausanne*, 2004, p. 352. Available at:  
<http://onlinelibrary.wiley.com/doi/10.1002/cbdv.200490137/abstract>.
- Power BI Expert (01 de outubro de 2022). What is staging in ETL Data Warehouse ETL Toolkit – Chapter 03 – Part 1 – Extracting. Obtido do Web site Qora Power BI Expert:  
<https://powerbixpert.com/2022/03/18/data-warehouse-etl-toolkit-chapter-03/>
- QlikTech International AB (03 de janeiro de 2022). The end-to-end data analytics platform without blind spots. Obtido de Web site da QlikTech International AB:  
<http://www.qlik.com/products/qlik-sense>
- Qora (02 de junho de 2022). What is staging in ETL? Obtido do Web site Qora:  
<https://www.quora.com/What-is-staging-in-ETL>
- Rejeb, A. and Keogh, J. G. (2021) '5G Networks in the Value Chain', *Wireless Personal Communications*. Springer US, 117(2), pp. 1577–1599. doi: 10.1007/s11277-020-07936-5.
- Saaty, R. W. (1987) 'The analytic hierarchy process-what it is and how it is used', *Mathematical Modelling*, 9(3–5), pp. 161–176. doi: 10.1016/0270-0255(87)90473-8.
- SAS Institute Inc. (03 de janeiro de 2022). SAS VISUAL ANALYTICS. Obtido de  
[https://www.sas.com/pt\\_pt/software/visual-analytics.html](https://www.sas.com/pt_pt/software/visual-analytics.html)
- SelectHub (20 de fevereiro de 2022). Power BI vs Tableau vs Qlik Sense: Which BI Tool is the Winner? Obtido de Web Site de SelectHub: <https://www.selecthub.com/business-intelligence/tableau-vs-qlikview-vs-microsoft-power-bi/>
- Sherman, R. (2015). *Business Intelligence Guidebook: From data integration to analytics*.
- Srivastava, G. et al. (2022) 'A review of the state of the art in business intelligence software', *Enterprise Information Systems*. Taylor & Francis, 16(1), pp. 1–28. doi: 10.1080/17517575.2021.1872107.
- SurveyMonkey (18 de setembro de 2022). SurveyMonkey. Obtido de  
<https://pt.surveymonkey.com/mp/likert-scale/>.
- Tableau (04 de janeiro de 2022) Why Tableau, Obtido de Web Site da Tableau:  
<https://www.tableau.com/pt-br/why-tableau>.
- Teixeira, J. B., Santos Filho, M. M. dos and Costa, C. A. D. (2018) 'Business Intelligence para uma análise da qualidade da entrega dos objetos postais', *Revista de Engenharia e Pesquisa Aplicada*, 3(3). doi: 10.25286/rep.v3i3.944.

Toledo, G. L. and Moretti, S. L. do A. (2016) 'Valor para o Cliente e Valor do Cliente: Conceitos e Implicações para o Processo de Marketing', *Desenvolvimento em Questão*, 14(35), p. 400. doi: 10.21527/2237-6453.2016.35.400-419.

Velho, S. V. M. (2011) 'ROLAP e MOLAP', *ROLAP e MOLAP: Comparação e Avaliação Prática*.



## A1 Anexo 1 – Logical Data Map (Prodite Zeelandia – Staging Area)

No Anexo 1 pode-se ver o Logical Data Map da base de dados operacional para a Staging Area, que foi desenvolvido durante a fase de design. Na tabela abaixo identifica-se a origem e o destino de todos os atributos assim como o tipo de cada um deles e as transformações aplicadas.

Destino					Origem				
Nome Tabela	Nome Coluna	Tipo Dados	Tipo Tabela	Tipo SCD	Nome Base de Dados	Nome Tabela	Nome Coluna	Tipo Dados	Transformação
Customer	No_	nvarchar(20)	Stagging		Prodite Zeelandia	Customer	No_	Texto Breve	
Customer	Name	nvarchar(50)	Stagging		Prodite Zeelandia	Customer	Name	Texto Breve	
Customer	Name 2	nvarchar(50)	Stagging		Prodite Zeelandia	Customer	Name 2	Texto Breve	
Customer	Search Name	nvarchar(50)	Stagging		Prodite Zeelandia	Customer	Search Name	Texto Breve	
Customer	Adress	nvarchar(50)	Stagging		Prodite Zeelandia	Customer	Adress	Texto Breve	
Customer	Adress 2	nvarchar(50)	Stagging		Prodite Zeelandia	Customer	Adress 2	Texto Breve	
Customer	City	nvarchar(30)	Stagging		Prodite Zeelandia	Customer	City	Texto Breve	
Customer	Country_Region Code	nvarchar(10)	Stagging		Prodite Zeelandia	Customer	Country_Region Code	Texto Breve	
Customer	County	nvarchar(30)	Stagging		Prodite Zeelandia	Customer	County	Texto Breve	
Customer	Post Code	nvarchar(20)	Stagging		Prodite Zeelandia	Customer	Post Code	Texto Breve	
Customer	Phone No_	nvarchar(30)	Stagging		Prodite Zeelandia	Customer	Phone No_	Texto Breve	
Customer	E-Mail	nvarchar(80)	Stagging		Prodite Zeelandia	Customer	E-Mail	Texto Breve	
Customer	Salesperson Code	nvarchar(10)	Stagging		Prodite Zeelandia	Customer	Salesperson Code	Texto Breve	
Customer	Customer Type	nvarchar(10)	Stagging		Prodite Zeelandia	Customer	Customer Type	Texto Breve	
Customer	Cod_Rota	nvarchar(10)	Stagging		Prodite Zeelandia	Customer	Cod_Rota	Texto Breve	
Customer	Territory Code	nvarchar(10)	Stagging		Prodite Zeelandia	Customer	Territory Code	Texto Breve	
Customer	Geographic Location	nvarchar(40)	Stagging		Prodite Zeelandia	Customer	Geographic Location	Texto Breve	

Cust_LedgerEntry	Entry No_	int	Stagging		Prodite Zeelandia	Cust_LedgerEntry	Entry No_	Número Inteiro	
Cust_LedgerEntry	Customer No_	nvarchar(20)	Stagging		Prodite Zeelandia	Cust_LedgerEntry	Customer No_	Texto Breve	
Cust_LedgerEntry	Posting Date	datetime	Stagging		Prodite Zeelandia	Cust_LedgerEntry	Posting Date	Data/Hora	
Cust_LedgerEntry	Document Type	int	Stagging		Prodite Zeelandia	Cust_LedgerEntry	Document Type	Número Inteiro	
Cust_LedgerEntry	Document No_	nvarchar(20)	Stagging		Prodite Zeelandia	Cust_LedgerEntry	Document No_	Texto Breve	
Cust_LedgerEntry	Salesperson Code	nvarchar(10)	Stagging		Prodite Zeelandia	Cust_LedgerEntry	Salesperson Code	Texto Breve	
Cust_LedgerEntry	Open	tinyint	Stagging		Prodite Zeelandia	Cust_LedgerEntry	Open	Número Inteiro	
Cust_LedgerEntry	Closed by Entry No_	int	Stagging		Prodite Zeelandia	Cust_LedgerEntry	Closed by Entry No_	Número Inteiro	
Cust_LedgerEntry	Closed at Date	datetime	Stagging		Prodite Zeelandia	Cust_LedgerEntry	Closed at Date	Data/Hora	
GL_Account	timestamp	timestamp	Stagging		Prodite Zeelandia	GL_Account	timestamp	Data/Hora	
GL_Account	No_	nvarchar(30)	Stagging		Prodite Zeelandia	GL_Account	No_	Texto Breve	
GL_Account	Name	nvarchar(50)	Stagging		Prodite Zeelandia	GL_Account	Name	Texto Breve	
GL_Account	Search Name	nvarchar(50)	Stagging		Prodite Zeelandia	GL_Account	Search Name	Texto Breve	
GL_Account	Blocked	tinyint	Stagging		Prodite Zeelandia	GL_Account	Blocked	Número Inteiro	
Item	No_	nvarchar(20)	Stagging		Prodite Zeelandia	Item	No_	Texto Breve	
Item	No_2	nvarchar(20)	Stagging		Prodite Zeelandia	Item	No_2	Texto Breve	
Item	Description	nvarchar(50)	Stagging		Prodite Zeelandia	Item	Description	Texto Breve	
Item	Search Description	nvarchar(50)	Stagging		Prodite Zeelandia	Item	Search Description	Texto Breve	
Item	Description 2	nvarchar(50)	Stagging		Prodite Zeelandia	Item	Description 2	Texto Breve	
Item	Inventory Posting Group	nvarchar(20)	Stagging		Prodite Zeelandia	Item	Inventory Posting Group	Texto Breve	
Item	Vendor No_	nvarchar(20)	Stagging		Prodite Zeelandia	Item	Vendor No_	Texto Breve	
Item	Blocked	tinyint	Stagging		Prodite Zeelandia	Item	Blocked	Número Inteiro	
Item	Item Category Code	nvarchar(20)	Stagging		Prodite Zeelandia	Item	Item Category Code	Texto Breve	
Item	Product Group Code	nvarchar(10)	Stagging		Prodite Zeelandia	Item	Product Group Code	Texto Breve	

Item	Zeelandia Product No_	nvarchar(20)	Stagging		Prodite Zeelandia	Item	Zeelandia Product No_	Texto Breve	
Item	PLM No_	nvarchar(20)	Stagging		Prodite Zeelandia	Item	PLM No_	Texto Breve	
ItemCategory	Code	nvarchar(20)	Stagging		Prodite Zeelandia	ItemCategory	Code	Texto Breve	
ItemCategory	Parent Category	nvarchar(20)	Stagging		Prodite Zeelandia	ItemCategory	Parent Category	Texto Breve	
ItemCategory	Description	nvarchar(50)	Stagging		Prodite Zeelandia	ItemCategory	Description	Texto Breve	
ItemLedgerEntry	Entry No_	int	Stagging		Prodite Zeelandia	ItemLedgerEntry	Entry No_	Número Inteiro	
ItemLedgerEntry	Item No_	nvarchar(20)	Stagging		Prodite Zeelandia	ItemLedgerEntry	Item No_	Texto Breve	
ItemLedgerEntry	Posting Date	datetime	Stagging		Prodite Zeelandia	ItemLedgerEntry	Posting Date	Data/Hora	
ItemLedgerEntry	Entry Type	int	Stagging		Prodite Zeelandia	ItemLedgerEntry	Entry Type	Número Inteiro	
ItemLedgerEntry	Source No_	nvarchar(20)	Stagging		Prodite Zeelandia	ItemLedgerEntry	Source No_	Texto Breve	
ItemLedgerEntry	Document No_	nvarchar(20)	Stagging		Prodite Zeelandia	ItemLedgerEntry	Document No_	Texto Breve	
ItemLedgerEntry	Location Code	nvarchar(10)	Stagging		Prodite Zeelandia	ItemLedgerEntry	Location Code	Texto Breve	
ItemLedgerEntry	Quantity	decimal(38,20)	Stagging		Prodite Zeelandia	ItemLedgerEntry	Quantity	Número Decimal	
ItemLedgerEntry	Aplies-to Entry	int	Stagging		Prodite Zeelandia	ItemLedgerEntry	Aplies-to Entry	Número Inteiro	
ItemLedgerEntry	Open	tinyint	Stagging		Prodite Zeelandia	ItemLedgerEntry	Open	Número Inteiro	
ItemLedgerEntry	Source Type	int	Stagging		Prodite Zeelandia	ItemLedgerEntry	Source Type	Número Inteiro	
ItemLedgerEntry	Document Date	datetime	Stagging		Prodite Zeelandia	ItemLedgerEntry	Document Date	Data/Hora	
ItemLedgerEntry	Document Type	int	Stagging		Prodite Zeelandia	ItemLedgerEntry	Document Type	Número Inteiro	
ItemLedgerEntry	Document Line No_	int	Stagging		Prodite Zeelandia	ItemLedgerEntry	Document Line No_	Número Inteiro	
ItemLedgerEntry	Completely Invoiced	tinyint	Stagging		Prodite Zeelandia	ItemLedgerEntry	Completely Invoiced	Número Inteiro	
Location	Code	nvarchar(10)	Stagging		Prodite Zeelandia	Location	Code	Texto Breve	
Location	Name	nvarchar(50)	Stagging		Prodite Zeelandia	Location	Name	Texto Breve	
Location	Name 2	nvarchar(50)	Stagging		Prodite Zeelandia	Location	Name 2	Texto Breve	

Location	Adress	nvarchar(50)	Stagging		Prodite Zeelandia	Location	Adress	Texto Breve	
Location	Adress 2	nvarchar(50)	Stagging		Prodite Zeelandia	Location	Adress 2	Texto Breve	
Location	City	nvarchar(30)	Stagging		Prodite Zeelandia	Location	City	Texto Breve	
Location	Post Code	nvarchar(20)	Stagging		Prodite Zeelandia	Location	Post Code	Texto Breve	
Location	County	nvarchar(30)	Stagging		Prodite Zeelandia	Location	County	Texto Breve	
Location	Country_Region Code	nvarchar(10)	Stagging		Prodite Zeelandia	Location	Country_Region Code	Texto Breve	
SalesCrMemoHeader	No_	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	No_	Texto Breve	
SalesCrMemoHeader	Sell-to Customer No_	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Sell-to Customer No_	Texto Breve	
SalesCrMemoHeader	Bill-to Customer No_	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Bill-to Customer No_	Texto Breve	
SalesCrMemoHeader	Bill-to Name	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Bill-to Name	Texto Breve	
SalesCrMemoHeader	Bill-to Name 2	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Bill-to Name 2	Texto Breve	
SalesCrMemoHeader	Bill-to Adress	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Bill-to Adress	Texto Breve	
SalesCrMemoHeader	Bill-to Adress 2	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Bill-to Adress 2	Texto Breve	
SalesCrMemoHeader	Bill-to City	nvarchar(30)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Bill-to City	Texto Breve	
SalesCrMemoHeader	Bill-to Contact	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Bill-to Contact	Texto Breve	
SalesCrMemoHeader	Ship-to Code	nvarchar(10)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Ship-to Code	Texto Breve	
SalesCrMemoHeader	Ship-to Name	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Ship-to Name	Texto Breve	
SalesCrMemoHeader	Ship-to Name 2	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Ship-to Name 2	Texto Breve	
SalesCrMemoHeader	Ship-to Adress	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Ship-to Adress	Texto Breve	
SalesCrMemoHeader	Ship-to Adress 2	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Ship-to Adress 2	Texto Breve	
SalesCrMemoHeader	Ship-to City	nvarchar(30)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Ship-to City	Texto Breve	
SalesCrMemoHeader	Ship-to Contact	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Ship-to Contact	Texto Breve	
SalesCrMemoHeader	Posting Date	datetime	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Posting Date	Data/Hora	
SalesCrMemoHeader	Shipment Date	datetime	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Shipment Date	Data/Hora	
SalesCrMemoHeader	Location Code	nvarchar(10)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Location Code	Texto Breve	
SalesCrMemoHeader	Salesperson Code	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Salesperson Code	Texto Breve	
SalesCrMemoHeader	Sell-to Customer Name	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Sell-to Customer Name	Texto Breve	

SalesCrMemoHeader	Sell-to Customer Name 2	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Sell-to Customer Name 2	Texto Breve	
SalesCrMemoHeader	Sell-to Adress	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Sell-to Adress	Texto Breve	
SalesCrMemoHeader	Sell-to Adress 2	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Sell-to Adress 2	Texto Breve	
SalesCrMemoHeader	Sell-to City	nvarchar(30)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Sell-to City	Texto Breve	
SalesCrMemoHeader	Sell-to Contact	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Sell-to Contact	Texto Breve	
SalesCrMemoHeader	Bill-to Post Code	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Bill-to Post Code	Texto Breve	
SalesCrMemoHeader	Bill-to County	nvarchar(30)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Bill-to County	Texto Breve	
SalesCrMemoHeader	Bill-to Country_Region Code	nvarchar(10)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Bill-to Country_Region Code	Texto Breve	
SalesCrMemoHeader	Sell-to Post Code	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Sell-to Post Code	Texto Breve	
SalesCrMemoHeader	Sell-to County	nvarchar(30)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Sell-to County	Texto Breve	
SalesCrMemoHeader	Sell-to Country_Region Code	nvarchar(10)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Sell-to Country_Region Code	Texto Breve	
SalesCrMemoHeader	Ship-to Post Code	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Ship-to Post Code	Texto Breve	
SalesCrMemoHeader	Ship-to County	nvarchar(30)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Ship-to County	Texto Breve	
SalesCrMemoHeader	Ship-to Country_Region Code	nvarchar(10)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Ship-to Country_Region Code	Texto Breve	
SalesCrMemoHeader	Document Date	datetime	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Document Date	Data/Hora	
SalesCrMemoLine	Document No_	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Document No_	Texto Breve	
SalesCrMemoLine	Line No_	int	Stagging		Prodite Zeelandia	SalesCrMemoLine	Line No_	Número Inteiro	
SalesCrMemoLine	Sell-to Customer No_	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Sell-to Customer No_	Texto Breve	
SalesCrMemoLine	Type	int	Stagging		Prodite Zeelandia	SalesCrMemoLine	Type	Número Inteiro	
SalesCrMemoLine	No_	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	No_	Texto Breve	
SalesCrMemoLine	Location Code	nvarchar(10)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Location Code	Texto Breve	
SalesCrMemoLine	Description	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Description	Texto Breve	
SalesCrMemoLine	Description 2	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Description 2	Texto Breve	
SalesCrMemoLine	Quantity	decimal(38,20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Quantity	Número Decimal	

SalesCrMemoLine	Unit Price	decimal(38,20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Unit Price	Número Decimal	
SalesCrMemoLine	Unit Cost	decimal(38,20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Unit Cost	Número Decimal	
SalesCrMemoLine	Line Discount_	decimal(38,20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Line Discount_	Número Decimal	
SalesCrMemoLine	Line Discount Amount	decimal(38,20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Line Discount Amount	Número Decimal	
SalesCrMemoLine	Amount	decimal(38,20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Amount	Número Decimal	
SalesCrMemoLine	Amount Including VAT	decimal(38,20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Amount Including VAT	Número Decimal	
SalesCrMemoLine	Bill-to Customer No_	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Bill-to Customer No_	Texto Breve	
SalesCrMemoLine	Line Amount	decimal(38,20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Line Amount	Número Decimal	
SalesCrMemoLine	Posting Date	datetime	Stagging		Prodite Zeelandia	SalesCrMemoLine	Posting Date	Data/Hora	
SalesCrMemoLine	Qty_per Unit of Measure	decimal(38,20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Qty_per Unit of Measure	Número Decimal	
SalesCrMemoLine	Unit of Measure Code	nvarchar(10)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Unit of Measure Code	Texto Breve	
SalesCrMemoLine	Quantity (Base)	decimal(38,20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Quantity (Base)	Número Decimal	
SalesCrMemoLine	Item Category Code	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Item Category Code	Texto Breve	
SalesCrMemoLine	Line Type	int	Stagging		Prodite Zeelandia	SalesCrMemoLine	Line Type	Número Inteiro	
SalesCrMemoLine	Campaign No_	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Campaign No_	Texto Breve	
SalesCrMemoLine	Campaign Qty_	int	Stagging		Prodite Zeelandia	SalesCrMemoLine	Campaign Qty_	Número Inteiro	
SalesInvoiceHeader	No_	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	No_	Texto Breve	
SalesInvoiceHeader	Sell-to Customer No_	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Sell-to Customer No_	Texto Breve	
SalesInvoiceHeader	Bill-to Customer No_	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Bill-to Customer No_	Texto Breve	
SalesInvoiceHeader	Bill-to Name	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Bill-to Name	Texto Breve	
SalesInvoiceHeader	Bill-to Name 2	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Bill-to Name 2	Texto Breve	
SalesInvoiceHeader	Bill-to Adress	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Bill-to Adress	Texto Breve	

SalesInvoiceHeader	Bill-to Adress 2	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Bill-to Adress 2	Texto Breve	
SalesInvoiceHeader	Bill-to City	nvarchar(30)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Bill-to City	Texto Breve	
SalesInvoiceHeader	Bill-to Contact	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Bill-to Contact	Texto Breve	
SalesInvoiceHeader	Ship-to Code	nvarchar(10)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Ship-to Code	Texto Breve	
SalesInvoiceHeader	Ship-to Name	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Ship-to Name	Texto Breve	
SalesInvoiceHeader	Ship-to Name 2	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Ship-to Name 2	Texto Breve	
SalesInvoiceHeader	Ship-to Adress	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Ship-to Adress	Texto Breve	
SalesInvoiceHeader	Ship-to Adress 2	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Ship-to Adress 2	Texto Breve	
SalesInvoiceHeader	Ship-to City	nvarchar(30)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Ship-to City	Texto Breve	
SalesInvoiceHeader	Ship-to Contact	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Ship-to Contact	Texto Breve	
SalesInvoiceHeader	Posting Date	datetime	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Posting Date	Data/Hora	
SalesInvoiceHeader	Shipment Date	datetime	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Shipment Date	Data/Hora	
SalesInvoiceHeader	Location Code	nvarchar(10)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Location Code	Texto Breve	
SalesInvoiceHeader	Salesperson Code	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Salesperson Code	Texto Breve	
SalesInvoiceHeader	Sell-to Customer Name	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Sell-to Customer Name	Texto Breve	
SalesInvoiceHeader	Sell-to Customer Name 2	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Sell-to Customer Name 2	Texto Breve	
SalesInvoiceHeader	Sell-to Adress	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Sell-to Adress	Texto Breve	
SalesInvoiceHeader	Sell-to Adress 2	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Sell-to Adress 2	Texto Breve	
SalesInvoiceHeader	Sell-to City	nvarchar(30)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Sell-to City	Texto Breve	
SalesInvoiceHeader	Sell-to Contact	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Sell-to Contact	Texto Breve	
SalesInvoiceHeader	Bill-to Post Code	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Bill-to Post Code	Texto Breve	
SalesInvoiceHeader	Bill-to County	nvarchar(30)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Bill-to County	Texto Breve	
SalesInvoiceHeader	Bill-to Country_Region Code	nvarchar(10)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Bill-to Country_Region Code	Texto Breve	
SalesInvoiceHeader	Sell-to Post Code	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Sell-to Post Code	Texto Breve	
SalesInvoiceHeader	Sell-to County	nvarchar(30)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Sell-to County	Texto Breve	
SalesInvoiceHeader	Sell-to Country_Region Code	nvarchar(10)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Sell-to Country_Region Code	Texto Breve	

SalesInvoiceHeader	Ship-to Post Code	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Ship-to Post Code	Texto Breve	
SalesInvoiceHeader	Ship-to County	nvarchar(30)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Ship-to County	Texto Breve	
SalesInvoiceHeader	Ship-to Country_Region Code	nvarchar(10)	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Ship-to Country_Region Code	Texto Breve	
SalesInvoiceHeader	Document Date	datetime	Stagging		Prodite Zeelandia	SalesCrMemoHeader	Document Date	Data/Hora	
SalesInvoiceLine	Document No_	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Document No_	Texto Breve	
SalesInvoiceLine	Line No_	int	Stagging		Prodite Zeelandia	SalesCrMemoLine	Line No_	Número Inteiro	
SalesInvoiceLine	Sell-to Customer No_	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Sell-to Customer No_	Texto Breve	
SalesInvoiceLine	Type	int	Stagging		Prodite Zeelandia	SalesCrMemoLine	Type	Número Inteiro	
SalesInvoiceLine	No_	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	No_	Texto Breve	
SalesInvoiceLine	Location Code	nvarchar(10)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Location Code	Texto Breve	
SalesInvoiceLine	Description	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Description	Texto Breve	
SalesInvoiceLine	Description 2	nvarchar(50)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Description 2	Texto Breve	
SalesInvoiceLine	Quantity	decimal(38,20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Quantity	Número Decimal	
SalesInvoiceLine	Unit Price	decimal(38,20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Unit Price	Número Decimal	
SalesInvoiceLine	Unit Cost	decimal(38,20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Unit Cost	Número Decimal	
SalesInvoiceLine	Line Discount_	decimal(38,20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Line Discount_	Número Decimal	
SalesInvoiceLine	Line Discount Amount	decimal(38,20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Line Discount Amount	Número Decimal	
SalesInvoiceLine	Amount	decimal(38,20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Amount	Número Decimal	
SalesInvoiceLine	Amount Including VAT	decimal(38,20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Amount Including VAT	Número Decimal	
SalesInvoiceLine	Bill-to Customer No_	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Bill-to Customer No_	Texto Breve	
SalesInvoiceLine	Line Amount	decimal(38,20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Line Amount	Número Decimal	
SalesInvoiceLine	Posting Date	datetime	Stagging		Prodite Zeelandia	SalesCrMemoLine	Posting Date	Data/Hora	

SalesInvoiceLine	Qty_per Unit of Measure	decimal(38,20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Qty_per Unit of Measure	Número Decimal	
SalesInvoiceLine	Unit of Measure Code	nvarchar(10)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Unit of Measure Code	Texto Breve	
SalesInvoiceLine	Quantity (Base)	decimal(38,20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Quantity (Base)	Número Decimal	
SalesInvoiceLine	Item Category Code	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Item Category Code	Texto Breve	
SalesInvoiceLine	Line Type	int	Stagging		Prodite Zeelandia	SalesCrMemoLine	Line Type	Número Inteiro	
SalesInvoiceLine	Campaign No_	nvarchar(20)	Stagging		Prodite Zeelandia	SalesCrMemoLine	Campaign No_	Texto Breve	
SalesInvoiceLine	Campaign Qty_	int	Stagging		Prodite Zeelandia	SalesCrMemoLine	Campaign Qty_	Número Inteiro	
Salesperson_Purchaser	Code	nvarchar(10)	Stagging		Prodite Zeelandia	Salesperson_Purchaser	Code	Texto Breve	
Salesperson_Purchaser	Name	nvarchar(50)	Stagging		Prodite Zeelandia	Salesperson_Purchaser	Name	Texto Breve	
Salesperson_Purchaser	Cod_Supervisor	nvarchar(10)	Stagging		Prodite Zeelandia	Salesperson_Purchaser	Cod_Supervisor	Texto Breve	
Salesperson_Purchaser	Cod_Equipa	nvarchar(10)	Stagging		Prodite Zeelandia	Salesperson_Purchaser	Cod_Equipa	Texto Breve	
Shipping Agent	Code	nvarchar(10)	Stagging		Prodite Zeelandia	Shipping Agent	Code	Texto Breve	
Shipping Agent	Name	nvarchar(50)	Stagging		Prodite Zeelandia	Shipping Agent	Name	Texto Breve	
Shipping Agent	Tech	tinyint	Stagging		Prodite Zeelandia	Shipping Agent	Tech	Número Inteiro	
Shipping Agent	Type	nvarchar(30)	Stagging		Prodite Zeelandia	Shipping Agent	Type	Texto Breve	
Shipping Agent	Internal Distribution	tinyint	Stagging		Prodite Zeelandia	Shipping Agent	Internal Distribution	Número Inteiro	
Shipping Agent	Sales	tinyint	Stagging		Prodite Zeelandia	Shipping Agent	Sales	Número Inteiro	
Shipping Agent	External Distribution	tinyint	Stagging		Prodite Zeelandia	Shipping Agent	External Distribution	Número Inteiro	
Vendor	No_	nvarchar(20)	Stagging		Prodite Zeelandia	Vendor	No_	Texto Breve	
Vendor	Name	nvarchar(50)	Stagging		Prodite Zeelandia	Vendor	Name	Texto Breve	
Vendor	Search Name	nvarchar(50)	Stagging		Prodite Zeelandia	Vendor	Search Name	Texto Breve	
Vendor	Name 2	nvarchar(50)	Stagging		Prodite Zeelandia	Vendor	Name 2	Texto Breve	
Vendor	Adress	nvarchar(50)	Stagging		Prodite Zeelandia	Vendor	Adress	Texto Breve	

Vendor	Adress 2	nvarchar(50)	Stagging		Prodite Zeelandia	Vendor	Adress 2	Texto Breve	
Vendor	City	nvarchar(30)	Stagging		Prodite Zeelandia	Vendor	City	Texto Breve	
Vendor	Phone No_	nvarchar(30)	Stagging		Prodite Zeelandia	Vendor	Phone No_	Texto Breve	
Vendor	Country_Region Code	nvarchar(10)	Stagging		Prodite Zeelandia	Vendor	Country_Region Code	Texto Breve	
Vendor	Post Code	nvarchar(20)	Stagging		Prodite Zeelandia	Vendor	Post Code	Texto Breve	
Vendor	County	nvarchar(30)	Stagging		Prodite Zeelandia	Vendor	County	Texto Breve	

## A2 Anexo 2 – Logical Data Map (Staging Area – Data Mart)

Tal como no anexo anterior o Anexo 2, também realizado durante a fase de design, contém a origem e o destino de cada atributo, mas neste caso a origem é a Staging Area e o destino é o Data Mart. Para além da origem, destino, tipo e transformações de cada atributo é identificada o tipo de SCD que será aplicado aos atributos das dimensões.

Destino					Origem				
Nome Tabela	Nome Coluna	Tipo Dados	Tipo Tabela	Tipo SCD	Nome Base de Dados	Nome Tabela	Nome Coluna	Tipo Dados	Transformação
DimCliente	ClienteKey	int	Dimensão	-					Chave Substituição
DimCliente	No_	nvarchar(20)	Dimensão	1	PZ_StaggingArea	Customer	No_	nvarchar(20)	select [No_] from PZ_StaggingArea.dbo.Customer
DimCliente	Nome	nvarchar(50)	Dimensão	1	PZ_StaggingArea	Customer	Name	nvarchar(50)	select [Name] from PZ_StaggingArea.dbo.Customer
DimCliente	Nome2	nvarchar(50)	Dimensão	1	PZ_StaggingArea	Customer	Name 2	nvarchar(50)	select [Name 2] from PZ_StaggingArea.dbo.Customer
DimCliente	Alias	nvarchar(50)	Dimensão	1	PZ_StaggingArea	Customer	Search Name	nvarchar(50)	select [Search Name] from PZ_StaggingArea.dbo.Customer
DimCliente	Endereco	nvarchar(50)	Dimensão	1	PZ_StaggingArea	Customer	Adress	nvarchar(50)	select [Adress] from PZ_StaggingArea.dbo.Customer
DimCliente	Endereco2	nvarchar(50)	Dimensão	1	PZ_StaggingArea	Customer	Adress 2	nvarchar(50)	select [Adress 2] from PZ_StaggingArea.dbo.Customer
DimCliente	Cidade	nvarchar(30)	Dimensão	1	PZ_StaggingArea	Customer	City	nvarchar(30)	select [City] from PZ_StaggingArea.dbo.Customer
DimCliente	Pais	nvarchar(10)	Dimensão	1	PZ_StaggingArea	Customer	Country_Region Code	nvarchar(10)	select [Country_Region Code] from PZ_StaggingArea.dbo.Customer
DimCliente	Distrito	nvarchar(30)	Dimensão	1	PZ_StaggingArea	Customer	County	nvarchar(30)	select [County] from PZ_StaggingArea.dbo.Customer
DimCliente	CodigoPostal	nvarchar(20)	Dimensão	1	PZ_StaggingArea	Customer	Post Code	nvarchar(20)	select [Post Code] from PZ_StaggingArea.dbo.Customer

DimCliente	Telefone	nvarchar(30)	Dimensão	1	PZ_StaggingArea	Customer	Phone No_	nvarchar(30)	select [Phone No_] from PZ_StaggingArea.dbo.Customer
DimCliente	Email	nvarchar(80)	Dimensão	1	PZ_StaggingArea	Customer	E-Mail	nvarchar(80)	select [E-Mail] from PZ_StaggingArea.dbo.Customer
DimCliente	CodigoVendedor	nvarchar(10)	Dimensão	1	PZ_StaggingArea	Customer	Salesperson Code	nvarchar(10)	select [Salesperson Code] from PZ_StaggingArea.dbo.Customer
DimCliente	TipoCliente	nvarchar(10)	Dimensão	1	PZ_StaggingArea	Customer	Customer Type	nvarchar(10)	select [Customer Type] from PZ_StaggingArea.dbo.Customer
DimCliente	CodigoRota	nvarchar(10)	Dimensão	1	PZ_StaggingArea	Customer	Cod_Rota	nvarchar(10)	select [Cod_Rota] from PZ_StaggingArea.dbo.Customer
DimCliente	CodigoTerritorio	nvarchar(10)	Dimensão	1	PZ_StaggingArea	Customer	Territory Code	nvarchar(10)	select [Territory Code] from PZ_StaggingArea.dbo.Customer
DimCliente	Geographic Location	nvarchar(40)	Dimensão	2	PZ_StaggingArea	Customer	Geographic Location	nvarchar(40)	select [Geographic Location] from PZ_StaggingArea.dbo.Customer
DimCliente	CodigoVendedor Hist	nvarchar(10)	Dimensão	2	PZ_StaggingArea	Customer	Salesperson Code	nvarchar(10)	select [Salesperson Code] from PZ_StaggingArea.dbo.Customer
DimCliente	TipoClienteHist	nvarchar(10)	Dimensão	2	PZ_StaggingArea	Customer	Customer Type	nvarchar(10)	select [Customer Type] from PZ_StaggingArea.dbo.Customer
DimCliente	CodigoRotaHist	nvarchar(10)	Dimensão	2	PZ_StaggingArea	Customer	Cod_Rota	nvarchar(10)	select [Cod_Rota] from PZ_StaggingArea.dbo.Customer
DimProduto	ProdutoKey	int	Dimensão	-					Chave Substituição
DimProduto	No_	nvarchar(20)	Dimensão	1	PZ_StaggingArea	Item	No_	nvarchar(20)	select dbo.Item.[No_] from dbo.Item inner join dbo.ItemCategory on dbo.Item.[Item Category Code] = dbo.ItemCategory.Code
DimProduto	No_2	nvarchar(20)	Dimensão	1	PZ_StaggingArea	Item	No_2	nvarchar(20)	select dbo.Item.[No_2] from dbo.Item inner join dbo.ItemCategory on dbo.Item.[Item Category Code] = dbo.ItemCategory.Code

DimProduto	Descricao	nvarchar(50)	Dimensão	1	PZ_StaggingArea	Item	Description	nvarchar(50)	select dbo.Item.[Description] from dbo.Item inner join dbo.ItemCategory on dbo.Item.[Item Category Code] = dbo.ItemCategory.Code
DimProduto	Descricao2	nvarchar(50)	Dimensão	1	PZ_StaggingArea	Item	Description 2	nvarchar(50)	select dbo.Item.[Description 2] from dbo.Item inner join dbo.ItemCategory on dbo.Item.[Item Category Code] = dbo.ItemCategory.Code
DimProduto	Alias	nvarchar(50)	Dimensão	1	PZ_StaggingArea	Item	Search Description	nvarchar(50)	select dbo.Item.[Search Description] from dbo.Item inner join dbo.ItemCategory on dbo.Item.[Item Category Code] = dbo.ItemCategory.Code
DimProduto	Bloqueado	tinyint	Dimensão	1	PZ_StaggingArea	Item	Blocked	tinyint	select dbo.Item.[Blocked] from dbo.Item inner join dbo.ItemCategory on dbo.Item.[Item Category Code] = dbo.ItemCategory.Code
DimProduto	CodigoCategoria	nvarchar(20)	Dimensão	2	PZ_StaggingArea	Item	Item Category Code	nvarchar(20)	select dbo.Item.[Item Category Code] from dbo.Item inner join dbo.ItemCategory on dbo.Item.[Item Category Code] = dbo.ItemCategory.Code
DimProduto	CategoriaDescriç ao	nvarchar(50)	Dimensão	1	PZ_StaggingArea	Item	CategoryDescripti on	nvarchar(50)	select dbo.ItemCategory.[Description] as CategoryDescription from dbo.Item inner join dbo.ItemCategory on dbo.Item.[Item Category Code] = dbo.ItemCategory.Code

DimProduto	GrupoContabilistico	nvarchar(20)	Dimensão	2	PZ_StaggingArea	Item	Inventory Posting Group	nvarchar(20)	select dbo.Item.[Inventory Posting Group] from dbo.Item inner join dbo.ItemCategory on dbo.Item.[Item Category Code] = dbo.ItemCategory.Code
DimProduto	Cod_GrupoProduto	nvarchar(10)	Dimensão	1	PZ_StaggingArea	Item	Product Group Code	nvarchar(10)	select dbo.Item[Product Group Code] from dbo.Item inner join dbo.ItemCategory on dbo.Item.[Item Category Code] = dbo.ItemCategory.Code
DimProduto	No_ProdutoZeelandia	nvarchar(20)	Dimensão	1	PZ_StaggingArea	Item	Zeelandia Product No_	nvarchar(20)	select dbo.Item.[Zeelandia Product No_] from dbo.Item inner join dbo.ItemCategory on dbo.Item.[Item Category Code] = dbo.ItemCategory.Code
DimProduto	No_PLM	nvarchar(20)	Dimensão	1	PZ_StaggingArea	Item	PLM No_	nvarchar(20)	select dbo.Item.[PLM No_] from dbo.Item inner join dbo.ItemCategory on dbo.Item.[Item Category Code] = dbo.ItemCategory.Code

DimProduto	Segmento	nvarchar(40)	Dimensão	1	Prodite Zeelandia	Item, Item Attribute Value Mapping, Item Attribute Value	Value	nvarchar(20)	SELECT dbo.[Prodite\$Item Attribute Value Mapping].No_, dbo.[Prodite\$Item Attribute Value].Value FROM dbo.[Prodite\$Item Attribute Value Mapping] INNER JOIN dbo.[Prodite\$Item Attribute Value] ON dbo.[Prodite\$Item Attribute Value Mapping].[Item Attribute ID] = dbo.[Prodite\$Item Attribute Value].[Attribute ID] AND dbo.[Prodite\$Item Attribute Value Mapping].[Item Attribute Value ID] = dbo.[Prodite\$Item Attribute Value].ID WHERE (dbo.[Prodite\$Item Attribute Value Mapping].[Item Attribute ID] = 18)
DimProduto	Vendor No_	nvarchar(20)	Dimensão	1	PZ_StaggingArea	Item	Vendor No_	nvarchar(20)	select dbo.Item.[Vendor No_] from dbo.Item inner join dbo.ItemCategory on dbo.Item.[Item Category Code] = dbo.ItemCategory.Code
DimProduto	Tipo	nvarchar(7)	Dimensão	1	PZ_StaggingArea	Item			Se tabela==Item então Tipo="Produto" Senão Tipo="Conta"
DimEnvio	EnvioKey	int	Dimensão	-	PZ_StaggingArea				Chave Substituição
DimEnvio	EnvioClienteNo	nvarchar(20)	Dimensão	1	PZ_StaggingArea	SalesInvoiceHeader, SalesCrMemoHeader	Sell-to Customer No_	nvarchar(20)	select [Sell-to Customer No_] from PZ_StaggingArea.dbo.SalesInvoiceHeader
DimEnvio	Nome	nvarchar(50)	Dimensão	1	PZ_StaggingArea	SalesInvoiceHeader, SalesCrMemoHeader	Ship-to Name	nvarchar(50)	select [Ship-to Name] from PZ_StaggingArea.dbo.SalesInvoiceHeader

DimEnvio	Nome2	nvarchar(50)	Dimensão	1	PZ_StaggingArea	SalesInvoiceHeader, SalesCrMemoHeader	Ship-to Name 2	nvarchar(50)	select [Ship-to Name 2] from PZ_StaggingArea.dbo.SalesInvoiceHeader
DimEnvio	Endereco	nvarchar(50)	Dimensão	1	PZ_StaggingArea	SalesInvoiceHeader, SalesCrMemoHeader	Ship-to Adress	nvarchar(50)	select [Ship-to Adress] from PZ_StaggingArea.dbo.SalesInvoiceHeader
DimEnvio	Endereco2	nvarchar(50)	Dimensão	1	PZ_StaggingArea	SalesInvoiceHeader, SalesCrMemoHeader	Ship-to Adress 2	nvarchar(50)	select [Ship-to Adress 2] from PZ_StaggingArea.dbo.SalesInvoiceHeader
DimEnvio	Cidade	nvarchar(50)	Dimensão	1	PZ_StaggingArea	SalesInvoiceHeader, SalesCrMemoHeader	Ship-to City	nvarchar(50)	select [Ship-to City] from PZ_StaggingArea.dbo.SalesInvoiceHeader
DimEnvio	Pais	nvarchar(20)	Dimensão	1	PZ_StaggingArea	SalesInvoiceHeader, SalesCrMemoHeader	Ship-to Country_Region Code	nvarchar(20)	select [Ship-to Country_Region Code] from PZ_StaggingArea.dbo.SalesInvoiceHeader
DimEnvio	Distrito	nvarchar(30)	Dimensão	1	PZ_StaggingArea	SalesInvoiceHeader, SalesCrMemoHeader	Ship-to County	nvarchar(30)	select [Ship-to County] from PZ_StaggingArea.dbo.SalesInvoiceHeader
DimEnvio	CodigoPostal	nvarchar(20)	Dimensão	1	PZ_StaggingArea	SalesInvoiceHeader, SalesCrMemoHeader	Ship-to Post Code	nvarchar(20)	select [Ship-to Post Code] from PZ_StaggingArea.dbo.SalesInvoiceHeader
DimExpedidor	ExpedidorKey	int	Dimensão	1	PZ_StaggingArea				Chave Substituição
DimExpedidor	Codigo	nvarchar(10)	Dimensão	-	PZ_StaggingArea	ShippingAgent	Code	nvarchar(10)	select [Code] from PZ_StaggingArea.dbo.ShippingAgent
DimExpedidor	Nome	nvarchar(50)	Dimensão	1	PZ_StaggingArea	ShippingAgent	Name	nvarchar(50)	select [Name] from PZ_StaggingArea.dbo.ShippingAgent
DimExpedidor	Descricao	nvarchar(250)	Dimensão	1	PZ_StaggingArea	ShippingAgent	Internet Adress	nvarchar(250)	select [Ship-to Name 2] from PZ_StaggingArea.dbo.ShippingAgent

DimExpedidor	Tipo	nvarchar(30)	Dimensão	1	PZ_StaggingArea	ShippingAgent	Type	nvarchar(30)	select [Internet Adress] from PZ_StaggingArea.dbo.ShippingAgent
DimExpedidor	DistribuicaoInterna	tinyint	Dimensão	1	PZ_StaggingArea	ShippingAgent	Internal Distribution	tinyint	select [Internal Distribution] from PZ_StaggingArea.dbo.ShippingAgent
DimExpedidor	DistribuicaoExterna	tinyint	Dimensão	1	PZ_StaggingArea	ShippingAgent	External Distribution	tinyint	select [External Distribution] from PZ_StaggingArea.dbo.ShippingAgent
DimExpedidor	Tecnico	tinyint	Dimensão	1	PZ_StaggingArea	ShippingAgent	Tech	tinyint	select [Tech] from PZ_StaggingArea.dbo.ShippingAgent
DimExpedidor	Comercial	tinyint	Dimensão	1	PZ_StaggingArea	ShippingAgent	Sales	tinyint	select [Sales] from PZ_StaggingArea.dbo.ShippingAgent
DimFornecedor	FornecedorKey	int	Dimensão	-					Chave Substituição
DimFornecedor	No_	nvarchar(20)	Dimensão	1	PZ_StaggingArea	Vendor	No_	nvarchar(20)	select [No_] from PZ_StaggingArea.dbo.Vendor
DimFornecedor	Nome	nvarchar(50)	Dimensão	1	PZ_StaggingArea	Vendor	Name	nvarchar(50)	select [Name] from PZ_StaggingArea.dbo.Vendor
DimFornecedor	Nome2	nvarchar(50)	Dimensão	1	PZ_StaggingArea	Vendor	Name 2	nvarchar(50)	select [Name 2] from PZ_StaggingArea.dbo.Vendor
DimFornecedor	Alias	nvarchar(50)	Dimensão	1	PZ_StaggingArea	Vendor	Search Name	nvarchar(50)	select [Search Name] from PZ_StaggingArea.dbo.Vendor
DimFornecedor	Endereco	nvarchar(50)	Dimensão	1	PZ_StaggingArea	Vendor	Adress	nvarchar(50)	select [Adress] from PZ_StaggingArea.dbo.Vendor
DimFornecedor	Endereco2	nvarchar(50)	Dimensão	1	PZ_StaggingArea	Vendor	Adress 2	nvarchar(50)	select [Adress 2] from PZ_StaggingArea.dbo.Vendor
DimFornecedor	Cidade	nvarchar(30)	Dimensão	1	PZ_StaggingArea	Vendor	City	nvarchar(30)	select [City] from PZ_StaggingArea.dbo.Vendor
DimFornecedor	CodigoPais	nvarchar(10)	Dimensão	1	PZ_StaggingArea	Vendor	Country_Region Code	nvarchar(10)	select [Country_Region Code] from PZ_StaggingArea.dbo.Vendor

DimFornecedor	Distrito	nvarchar(30)	Dimensão	1	PZ_StaggingArea	Vendor	County	nvarchar(30)	select [County] from PZ_StaggingArea.dbo.Vendor
DimFornecedor	CodigoPostal	nvarchar(20)	Dimensão	1	PZ_StaggingArea	Vendor	Post Code	nvarchar(20)	select [Post Code] from PZ_StaggingArea.dbo.Vendor
DimFornecedor	Telefone	nvarchar(30)	Dimensão	1	PZ_StaggingArea	Vendor	Phone No_	nvarchar(30)	select [Phone No_] from PZ_StaggingArea.dbo.Vendor
DimLocalizacao	LocalizacaoKey	int	Dimensão	-					Chave Substituição
DimLocalizacao	Codigo	nvarchar(10)	Dimensão	1	PZ_StaggingArea	Location	Code	nvarchar(10)	select [Code] from PZ_StaggingArea.dbo.Location
DimLocalizacao	Nome	nvarchar(50)	Dimensão	1	PZ_StaggingArea	Location	Name	nvarchar(50)	select [Name] from PZ_StaggingArea.dbo.Location
DimLocalizacao	Nome2	nvarchar(50)	Dimensão	1	PZ_StaggingArea	Location	Name 2	nvarchar(50)	select [Name 2] from PZ_StaggingArea.dbo.Location
DimLocalizacao	Cidade	nvarchar(30)	Dimensão	1	PZ_StaggingArea	Location	City	nvarchar(30)	select [City] from PZ_StaggingArea.dbo.Location
DimLocalizacao	CodigoPais	nvarchar(10)	Dimensão	1	PZ_StaggingArea	Location	Country_Region Code	nvarchar(10)	select [Country_Region Code] from PZ_StaggingArea.dbo.Location
DimLocalizacao	Distrito	nvarchar(30)	Dimensão	1	PZ_StaggingArea	Location	County	nvarchar(30)	select [County] from PZ_StaggingArea.dbo.Location
DimLocalizacao	AreaGeografica	nvarchar(10)	Dimensão	1	PZ_StaggingArea	Location	Country_Region Code	nvarchar(10)	select [Country_Region Code] from PZ_StaggingArea.dbo.Location
DimVendedor	VendedorKey	int	Dimensão	1					Chave Substituição
DimVendedor	Codigo	nvarchar(10)	Dimensão	1	PZ_StaggingArea	Salesperson_Purchaser	Code	nvarchar(10)	select [Code] from PZ_StaggingArea.dbo.Salesperson_Purchaser
DimVendedor	Nome	nvarchar(50)	Dimensão	1	PZ_StaggingArea	Salesperson_Purchaser	Name	nvarchar(50)	select [Name] from PZ_StaggingArea.dbo.Salesperson_Purchaser

DimVendedor	CodigoSupervisor	nvarchar(10)	Dimensão	1	PZ_StaggingArea	Salesperson_Purchaser	Cod_Supervisor	nvarchar(10)	select [Cod_Supervisor] from PZ_StaggingArea.dbo.Salesperson_Purchaser
DimVendedor	CodigoSupervisorHist	nvarchar(10)	Dimensão	2	PZ_StaggingArea	Salesperson_Purchaser	Cod_Supervisor	nvarchar(10)	select [Cod_Supervisor] from PZ_StaggingArea.dbo.Salesperson_Purchaser
DimVendedor	AreaGeografica	nvarchar(20)	Dimensão	1	PZ_StaggingArea	Salesperson_Purchaser	Cod_Equipa	nvarchar(10)	[Cod_Equipa] == "PT" ? "PORTO" : [Cod_Equipa] == "LX" ? "LISBOA" : [Cod_Equipa] == "MD" ? "MADEIRA" : [Cod_Equipa] == "AL" ? "ALGARVE" : ""
DimData	DataKey	int	Dimensão	-					
DimData	DataCompleta	date	Dimensão	1					
DimData	Ano	int	Dimensão	1					
DimData	Semestre	tinyint	Dimensão	1					
DimData	Quadrimestre	tinyint	Dimensão	1					
DimData	Trimestre	tinyint	Dimensão	1					
DimData	Mes	tinyint	Dimensão	1					
DimData	MesNome	nvarchar(10)	Dimensão	1					
DimData	Semana	tinyint	Dimensão	1					
DimData	SemanaMes	tinyint	Dimensão	1					
DimData	DiaNumeroAno	int	Dimensão	1					
DimData	DiaNumeroMes	tinyint	Dimensão	1					
DimData	DiaNumeroSemana	tinyint	Dimensão	1					
DimData	DiaSemana	nvarchar(20)	Dimensão	1					
DimData	FimSemana	nvarchar(3)	Dimensão	1					
DimData	Estacao	nvarchar(10)	Dimensão	1					
DimData	UltimoDiaMes	nvarchar(3)	Dimensão	1					

FactVendas	DataKey	int	Factos	N.A.	PZ_StaggingArea, PZ_DataMart	DimData, SalesInvoiceLine, SalesCrMemoLine	DataKey	int	where PZ_StaggingArea.dbo.SalesInvoiceLine. [Posting Date] = PZ_DataMart.DimData.[DataCompleta]
FactVendas	ClienteKey	int	Factos	N.A.	PZ_StaggingArea, PZ_DataMart	DimCliente, SalesInvoiceLine, SalesCrMemoLine	ClienteKey	int	where PZ_StaggingArea.dbo.SalesInvoiceLine. [Bill-to Customer No_] = PZ_DataMart.DimCliente.[No_]
FactVendas	ProdutoKey	int	Factos	N.A.	PZ_StaggingArea, PZ_DataMart	DimProduto, SalesInvoiceLine, SalesCrMemoLine	ProdutoKey	int	where PZ_StaggingArea.dbo.SalesInvoiceLine. [No_] = PZ_DataMart.DimCliente.[No_]
FactVendas	LocalizacaoKey	int	Factos	N.A.	PZ_StaggingArea, PZ_DataMart	DimLocalizacao, SalesInvoiceLine, SalesCrMemoLine	LocalizacaoKey	int	where PZ_StaggingArea.dbo.SalesInvoiceLine. [Location Code] = PZ_DataMart.DimLocalizacao.[Codigo]
FactVendas	ExpedidorKey	int	Factos	N.A.	PZ_StaggingArea, PZ_DataMart	DimExpedidor, SalesInvoiceLine, SalesCrMemoLine	ExpedidorKey	int	where PZ_StaggingArea.dbo.SalesInvoiceLine. [Shipping Agent Code] = PZ_DataMart.DimExpedidor.[Codigo]
FactVendas	EnvioKey	int	Factos	N.A.	PZ_StaggingArea, PZ_DataMart	DimEnvio, SalesInvoiceLine, SalesCrMemoLine	EnvioKey	int	where PZ_StaggingArea.dbo.SalesInvoiceLine. [Seel-to Customer No_] = PZ_DataMart.DimEnvio.[EnvioClienteNo_]
FactVendas	FornecedorKey	int	Factos	N.A.	PZ_StaggingArea, PZ_DataMart	DimFornecedor, Item	FornecedorKey	int	where PZ_StaggingArea.dbo.Item.[Vendor No_] = PZ_DataMart.Fornecedor.[No_]
FactVendas	VendedorKey	int	Factos	N.A.	PZ_StaggingArea, PZ_DataMart	DimVendedor, SalesInvoiceLine, SalesCrMemoLine	VendedorKey	int	where PZ_StaggingArea.dbo.SalesInvoiceLine. [SalesPerson Code] = PZ_DataMart.DimVendedor.[Codigo]

FactVendas	DocumentoNo	bigint	Factos	N.A.	PZ_StaggingArea	SalesInvoiceLine, SalesCrMemoLine	Document No_	nvarchar(20)	select convert(bigint,PZ_StaggingArea.dbo.Sal esInvoiceLine.[Document No_]) from PZ_StaggingArea.dbo.SalesInvoiceLine
FactVendas	NoLinha	int	Factos	N.A.	PZ_StaggingArea	SalesInvoiceLine, SalesCrMemoLine	Line No_	int	select [Line No_] from PZ_StaggingArea.dbo.SalesInvoiceLine
FactVendas	TipoDocumento	nvarchar(15)	Factos	N.A.	PZ_StaggingArea	SalesInvoiceHeader, SalesCrMemoHeader			se tabela==SalesInvoiceHeader então TipoDocumento="Fatura" Senão TipoDocumento="Nota de Crédito"
FactVendas	TipoLinha	nvarchar(5)	Factos	N.A.	PZ_StaggingArea	SalesInvoiceLine, SalesCrMemoLine	Quantity (Base)	decimal(38,20)	[Quantity (Base)] != 0 && [Line Amount] == 0 ? "Bónus" : " "
FactVendas	Tipo	nvarchar(7)	Factos	N.A.	PZ_StaggingArea	SalesInvoiceLine, SalesCrMemoLine	Type	int	Type == "1" ? "Conta" : "Produto"
FactVendas	Quantidade	float	Factos	N.A.	PZ_StaggingArea	SalesInvoiceLine, SalesCrMemoLine	Quantity	decimal(38,20)	select Quantity from PZ_StaggingArea.dbo.SalesInvoiceLine
FactVendas	QuantidadeBase	float	Factos	N.A.	PZ_StaggingArea	SalesInvoiceLine, SalesCrMemoLine	Quantity (Base)	decimal(38,20)	select [Quantity (Base)] from PZ_StaggingArea.dbo.SalesInvoiceLine
FactVendas	QtdVendida	float	Factos	N.A.	PZ_StaggingArea	SalesInvoiceLine, SalesCrMemoLine	Quantity (Base)	decimal(38,20)	[Quantity (Base)] != 0 && [Line Amount] != 0 ? [Quantity (Base)] : 0
FactVendas	QtdBonus	float	Factos	N.A.	PZ_StaggingArea	SalesInvoiceLine, SalesCrMemoLine	Quantity (Base)	decimal(38,20)	[Quantity (Base)] != 0 && [Line Amount] == 0 ? [Quantity (Base)] : 0
FactVendas	PrecoUnitario	float	Factos	N.A.	PZ_StaggingArea	SalesInvoiceLine, SalesCrMemoLine	Unit Price	decimal(38,20)	select [Unit Price] from PZ_StaggingArea.dbo.SalesInvoiceLine
FactVendas	ValorLiquido	float	Factos	N.A.	PZ_StaggingArea	SalesInvoiceLine, SalesCrMemoLine	Line Amount	decimal(38,20)	select [Line Amount] from PZ_StaggingArea.dbo.SalesInvoiceLine
FactVendas	Desconto%	float	Factos	N.A.	PZ_StaggingArea	SalesInvoiceLine, SalesCrMemoLine	Line Discount_	decimal(38,20)	select [Line Discount_] from PZ_StaggingArea.dbo.SalesInvoiceLine
FactVendas	ValorDesconto	float	Factos	N.A.	PZ_StaggingArea	SalesInvoiceLine, SalesCrMemoLine	Line Discount Amount	decimal(38,20)	select [Line Discount Amount] from PZ_StaggingArea.dbo.SalesInvoiceLine
FactVendas	Custo	float	Factos	N.A.	PZ_StaggingArea	SalesInvoiceLine, SalesCrMemoLine		decimal(38,20)	[Unit Cost] * Quantity

FactVendas	DataPagamento	date	Factos	N.A.	PZ_StaggingArea	SalesInvoiceLine, SalesCrMemoLine, Cust_LedgerEntry	Closed at Date	date	select [Closed at Date] from PZ_StaggingArea.dbo.Cust_LedgerEntr y where PZ_StaggingArea.dbo.SalesInvoiceLine. [Document No_] = PZ_StaggingArea.dbo.Cust_LedgerEntr y.[Closed at Date]
FactVendas	DiasPagamento	int	Factos	N.A.	PZ_StaggingArea	SalesInvoiceLine, SalesCrMemoLine, Cust_LedgerEntry	Posting Date, Closed at Date	int	DATEDIFF("day", (DT_DATE)[Posting Date], (DT_DATE)[Closed at Date])
FactVendas	AnoPagamento	int	Factos	N.A.	PZ_StaggingArea	SalesInvoiceLine, SalesCrMemoLine, Cust_LedgerEntry	Posting Date, Closed at Date	int	YEAR([Closed at Date])
FactVendas	MesPagamento	int	Factos	N.A.	PZ_StaggingArea	SalesInvoiceLine, SalesCrMemoLine, Cust_LedgerEntry	Posting Date, Closed at Date	int	MONTH([Closed at Date])
FactVendas	CodigoCampanh a	nvarchar(20)	Factos	N.A.	PZ_StaggingArea	SalesInvoiceLine, SalesCrMemoLine	Campaign No_	nvarchar(20)	select [Campaign No_] from PZ_StaggingArea.dbo.SalesInvoiceLine
FactVendas	QtdCampanha	int	Factos	N.A.	PZ_StaggingArea	SalesInvoiceLine, SalesCrMemoLine	Campaign Qty_	int	select [Campaign Qty_] from PZ_StaggingArea.dbo.SalesInvoiceLine

## A3 Anexo 3 – Fluxo de Controlo da solução implementada

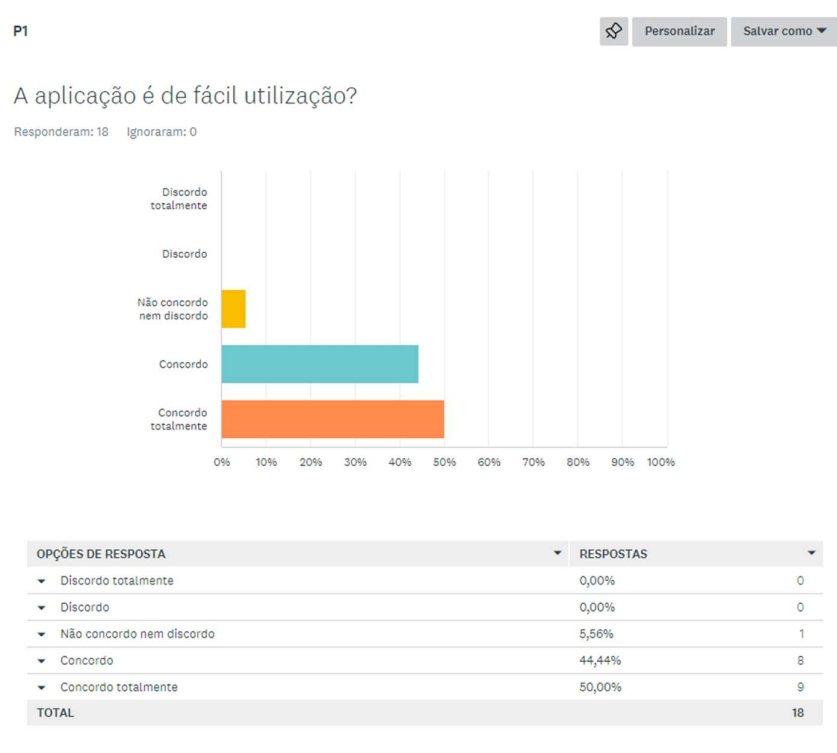
No anexo 3 podemos ver ilustrado o fluxo de controlo desenvolvido no Microsoft SSIS. Desta forma é possível visualizar como todos os componentes se interligam. Na imagem pode-se todos os passos desde a criação das bases de dados PZ\_StagingArea e PZ\_DataMart caso não existam, seguindo-se a criação das tabelas de ambas as bases de dados, o carregamento dos dados da Staging Area e da dimensão Data, seguindo-se o carregamento das restantes dimensões. Após o carregamento de todas as dimensões são eliminadas as restrições chave primária - chave estrangeira entre as dimensões e a tabela de factos, seguindo-se o carregamento da tabela de factos e são recriadas novamente as restrições chave primária - chave estrangeira. Por fim procede-se à atualização do cubo e das suas dimensões.





# A4 Anexo 4 – Resultados inquérito de satisfação aos utilizadores

Neste anexo podemos visualizar os resultados das respostas do inquérito realizado aos utilizadores do sistema, que contou com a participação de 18 utilizadores.

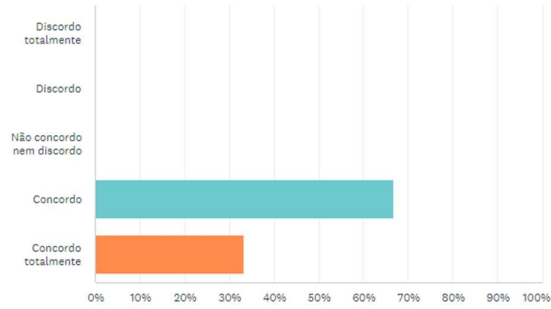


P2

Personalizar Salvar como

O layout das análises apresentadas é mais apelativo do que o atual sistema?

Responderam: 18 Ignoraram: 0



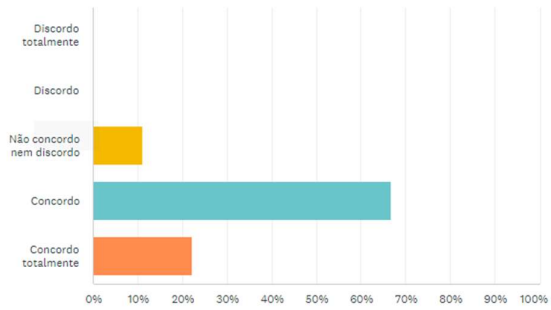
OPÇÕES DE RESPOSTA	RESPOSTAS
Discordo totalmente	0,00% 0
Discordo	0,00% 0
Não concordo nem discordo	0,00% 0
Concordo	66,67% 12
Concordo totalmente	33,33% 6
<b>TOTAL</b>	<b>18</b>

P3

Personalizar Salvar como

As análises apresentadas são relevantes?

Responderam: 18 Ignoraram: 0



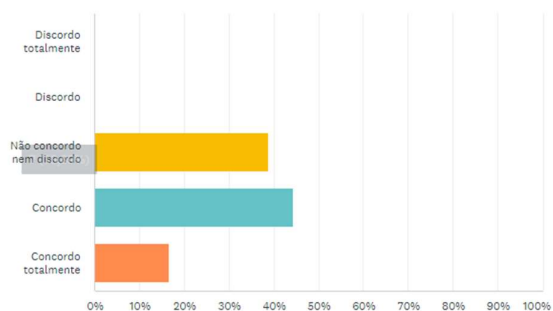
OPÇÕES DE RESPOSTA	RESPOSTAS
Discordo totalmente	0,00% 0
Discordo	0,00% 0
Não concordo nem discordo	11,11% 2
Concordo	66,67% 12
Concordo totalmente	22,22% 4
<b>TOTAL</b>	<b>18</b>

P4

Personalizar Salvar como

A informação apresentada é de fácil leitura e compreensão?

Responderam: 18 Ignoraram: 0



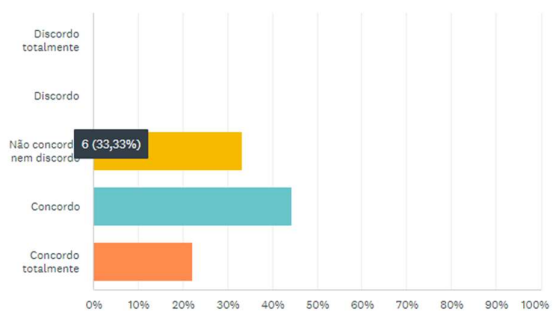
OPÇÕES DE RESPOSTA	RESPOSTAS
Discordo totalmente	0,00% 0
Discordo	0,00% 0
Não concordo nem discordo	38,89% 7
Concordo	44,44% 8
Concordo totalmente	16,67% 3
<b>TOTAL</b>	<b>18</b>

P5

Personalizar Salvar como

Considera o novo sistema uma ferramenta útil de apoio à decisão?

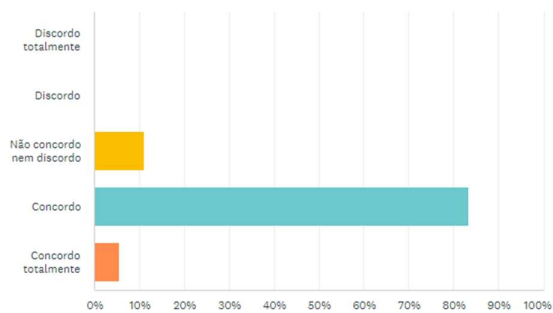
Responderam: 18 Ignoraram: 0



OPÇÕES DE RESPOSTA	RESPOSTAS
Discordo totalmente	0,00% 0
Discordo	0,00% 0
Não concordo nem discordo	33,33% 6
Concordo	44,44% 8
Concordo totalmente	22,22% 4
<b>TOTAL</b>	<b>18</b>

Considera que o sistema desenvolvido cumpre os objetivos propostos?

Responderam: 18 Ignoraram: 0



OPÇÕES DE RESPOSTA	RESPOSTAS	
▼ Discordo totalmente	0,00%	0
▼ Discordo	0,00%	0
▼ Não concordo nem discordo	11,11%	2
▼ Concordo	83,33%	15
▼ Concordo totalmente	5,56%	1
<b>TOTAL</b>		<b>18</b>