



ESCOLA NAVAL



talant de bi-faire

Jéssica Sofia Teles Machado

Business Intelligence da Atividade Operacional da Marinha Portuguesa

**Processamento e Análise de Mensagens de Texto Formatado e
Mensagens Estruturadas**

Dissertação para obtenção do grau de Mestre em Ciências Militares Navais,
na especialidade de Marinha

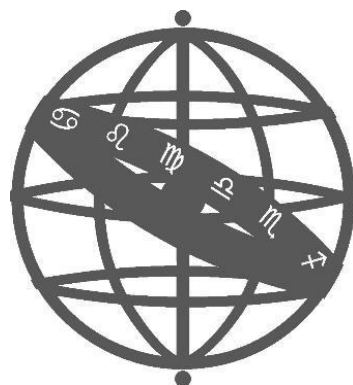


Alfeite
2018



ESCOLA NAVAL

talant de bi-faire



Jéssica Sofia Teles Machado

*Business Intelligence da Atividade Operacional da Marinha Portuguesa
Processamento e Análise de Mensagens de Texto Formatado e Mensagens
Estruturadas*

**Dissertação para obtenção do grau de Mestre em Ciências Militares Navais, na
especialidade de Marinha**

Orientação de: Professor Anacleto Cortez e Correia

Co-orientação de: CTEN EN-AEL Andrade Gonçalves

O Aluno Mestrando

O Orientador

Jéssica Machado

ASPOF Teles Machado

Anacleto Cortez e Correia

Professor Cortez e Correia

**Alfeite
2018**

Epígrafe

“A morte é algo inevitável. Quando um Homem fez o que considera seu dever para com seu povo e seu país, pode descansar em paz.”

Nelson Mandela

Dedicatória

Ao meu porto de abrigo e em quem deposito o meu maior orgulho, que são os meus pais e a minha irmã, que me orientaram a seguir o caminho genuíno graças a todo o apoio, presença, acompanhamento, força e amor incondicional.

Agradecimentos

Uma das fases mais importantes da vida de um futuro oficial de marinha é representada pelo término da dissertação de mestrado. Como tal, o auxílio e apoio de várias pessoas foi crucial. Sendo assim, expresso, de seguida, a minha gratidão a todos aqueles que, de alguma forma, contribuíram para a elaboração da presente dissertação de mestrado.

Primeiramente, gostaria de agradecer às duas pessoas que foram sem dúvida os meus alicerces para a realização deste trabalho, ao meu orientador e ao meu coorientador, Senhor Professor Anacleto Cortez e Correia e Capitão-tenente Rui Manuel Andrade Gonçalves, por todo o apoio, paciência, conselhos, correções, disponibilidade e incentivo.

Os meus sinceros agradecimentos à Capitão-tenente Colaço Ferreira e ao Capitão-tenente Tusto Cachucho por todo o tempo que dispenderam comigo para me transmitirem os seus conhecimentos, por todo o apoio, disponibilidade, paciência e amabilidade.

Aos oficiais da Direção de Análise e Gestão da Informação, quero agradecer pela forma atenciosa e pelo incentivo demonstrado durante a realização da presente dissertação de mestrado.

Agradecer ao Primeiro-tenente Cardoso Godinho por todo o auxílio prestado, essencialmente pela introdução aos conceitos, à doutrina e à disponibilidade demonstrada, no âmbito da fiscalização da pesca.

Ao Capitão-tenente Neves Simões e à Capitão-tenente Pereira Lopes quero agradecer o apoio, a disponibilidade, as sugestões e a partilha de conhecimento que me proporcionaram bases para a análise de comunicados operacionais.

Agradecer ao Capitão-de-fragata Nuno Miguel Castro Guimarães Palmeiro Ribeiro, aos restantes oficiais e à guarnição do NRP D. Carlos I, pela preocupação, encorajamento, partilha de conhecimentos e do zelo pelo bem-estar a bordo, numa das fases mais importantes.

Aos meus pais, irmã, avô, tio e tia, por toda a força, apoio, compreensão, e amor prestado que me torna uma pessoa ainda mais completa.

Ao Rui, um especial agradecimento, porque a gratidão é a memória do coração.

Ao Curso “CMG Henrique Quirino da Fonseca”, pela camaradagem, amizade e companheirismo ao longo destes anos.

Resumo

O volume de dados que as unidades da Marinha Portuguesa, no âmbito da atividade operacional, trocam entre si é amplo e variado. O modo como cada unidade trata os dados é dissemelhante, o que dificulta o processo de análise e de tomada de decisão de forma unificada. A difusão dos dados ocorre através de Mensagens de Texto Formatado e Mensagens Estruturadas, sendo que provoca dificuldades no tratamento de dados por serem mensagens com formatos distintos. Perante isto, procedeu-se a uma análise dos dados pertencentes aos dois tipos de mensagem referidos acima, onde se obtiveram conclusões para uma estrutura de mensagem que abrange os dados mais importantes e que possibilita um tratamento automático e uniforme entre as unidades.

Esses referidos dados que circulam entre as unidades, necessitam de um tratamento que se baseia na deteção e correção de incongruências e, desta forma, origina informação. No entanto, verifica-se que nem toda a informação tem um carácter relevante, isto é, que influencie no processo de tomada de decisão, o que constitui uma necessidade de gestão de informação relevante.

Atualmente, estamos perante uma confluência de novas tecnologias, sendo que muitas delas são fundamentadas em sistemas de informação e possuem ferramentas que conformam um auxílio significativo quer no tratamento de dados, quer na visualização de informação e conseqüentemente, um impacto na tomada de decisão.

Deste modo, foi identificada a necessidade de conceber, desenhar e implementar um sistema que permita a recolha, tratamento, integração, consolidação, e análise dos dados relevantes, contidos nas mensagens, para que estes possam contribuir para a mais eficaz e eficiente tomada de decisão.

Com o intuito de promover a visualização da informação relevante, sumarizada e de interesse, a presente dissertação tem como finalidade a conceção de uma solução que contribua para que as unidades da Marinha obtenham uma visão unificada, que compreenda apenas a informação relevante.

Para esse efeito, foi edificado um sistema *Online Analytical Processing*, suportado por *Data Warehouse*, *Data Mart* e por componentes de uma solução de *Business Intelligence*, que tem como prova de conceito a fiscalização de pesca e como aplicação assistir às necessidades de informação e tomada de decisão de utilizadores a diferentes níveis da organização.

Palavras-chave: *Business Intelligence*, Marinha Portuguesa, Mensagens de Texto Formatado, Mensagens Estruturadas, Tomada de Decisão, Gestão e Visualização de Informação Relevante.

Abstract

The volume of data that the units of the Portuguese Navy, under the operational activity, exchange with each other is wide and varied. The way each unit handles the data is dissimilar, what hinders the process of analysis and decision-making of unified form. The dissemination of data occurs through Message Text Formats and Structured Messages and causes difficulties in the processing of data by being messages with different formats. Given this, an analysis of the data pertaining to the two types of message referred to above, where obtained conclusion for a message structure that covers the most important data and enabling automatic and uniform treatment between the units.

These data circulating between units, require a treatment which is based on detection and correction of inconsistencies and in this way, information originates. However, it turns out that not all the information has a relevant character, this is that influence in the decision-making process, which is a need for relevant information management.

Currently, we are facing a confluence of new technologies, and that many of them are based on information systems and have tools that make up a significant aid both in data processing, both on information visualization and therefore an impact on decision-making.

Thus, it was identified the need to conceive, design and implement a system for the collection, processing, integration, consolidation, and analysis of the relevant data contained in the messages, so that they can contribute to the more effective and efficient decision-making.

In order to promoting the visualization of relevant information, summarized and of interest, the present dissertation has as purpose the design of a solution that contribute so that the units of the Navy get a unified view, which includes only relevant information.

For this effect, was built a system Online Analytical Processing supported by Data Warehouse, Data Mart and by components of a solution of Business Intelligence, that has a proof concept the fisheries inspection and as an application watch the information needs and user decision making at different levels of the organization.

Keywords: Business Intelligence, Portuguese Navy, Message Text Formats, Structured Messages, Decision-making, Management and Visualization of Relevant Information.

Índice

Epígrafe	v
Dedicatória	vii
Agradecimentos	ix
Resumo	xi
Abstract	xiii
Índice	xv
Índice de figuras	xvii
Índice de tabelas e gráficos	xix
Lista de abreviaturas, siglas e acrónimos	xxi
Introdução	1
Enquadramento	1
Definição do Problema	2
Objetivo e Questões de Investigação	3
Metodologias de Investigação	3
Estrutura	5
1. Capítulo 1 – Revisão da Literatura	9
1.1. Estado da Arte	9
1.1.1. Introdução ao <i>Business Intelligence</i>	9
1.1.2. <i>Data Warehouse</i>	10
1.1.3. Modelação Dimensional	16
1.1.4. Visualização	20
1.1.5. Tendências e perspetivas futuras do <i>Business Intelligence</i>	22
1.2. Levantamento de Ferramentas	23
2. Capítulo 2 – Caracterização do Domínio do Problema	27
2.1. Medidas de Interesse	27
2.2. Fiscalização da Pesca	27
2.2.1. Organismos com competência legal	28
2.2.2. Áreas de Atuação	29
2.2.3. Recursos Haliêuticos	32
2.2.4. Controlo e Fiscalização das Atividades de Pesca	34
2.2.5. Ferramentas de apoio no âmbito da fiscalização da pesca	35

3. Capítulo 3 – Recolha e Tratamento de Dados	41
3.1. Conceito e objetivo das MTF e das Mensagens Estruturadas	41
3.2. Identificação e Tratamento das Fontes de Dados	42
3.3. Recolha de Dados.....	45
3.4. Análise e tratamento dos dados recolhidos	48
3.5. Base de Dados	51
4. Capítulo 4 – Conceção da Solução	59
4.1. Modelação Dimensional.....	59
4.2. Processo Extração, Transformação e Carregamento	60
4.3. Cubo Multidimensional.....	64
4.4. Visualização da Informação	66
5. Capítulo 5 – Análise de Resultados	73
5.1. Avaliação do Artefacto.....	73
5.2. Análise dos Resultados Obtidos	74
5.2.1. Características da Amostra	74
5.2.2. Análise detalhada dos resultados obtidos	76
5.2.3. Análise da eficácia do sistema edificado.....	83
Conclusões e Recomendações	87
Conclusões	87
Recomendações e Trabalhos Futuros.....	88
Bibliografia	89
Glossário	101
Apêndices	103
Apêndice A – Conceitos fundamentais sobre <i>Business Intelligence</i>	105
Apêndice B – Identificação dos dados contidos nas mensagens, relativos à fiscalização da pesca.....	111
Apêndice C – Identificação das mensagens e dos dados referentes às medidas de interesse	113
Apêndice D – Tratamento de incongruências detetadas nos dados recolhidos	117
Apêndice E – Classificação do tipo de dados.....	119
Apêndice F – Modelação dimensional: esquema em estrela.....	121
Apêndice G – Modelação dimensional aplicada nas restantes medidas de interesse	123
Apêndice H – Processo ETL referente ao fato	125
Apêndice I – <i>Dashboard</i> e relatórios complementares referentes à fiscalização da pesca ...	129
Apêndice J – Inquérito para avaliação do artefacto produzido	131

Índice de figuras

Figura 1- Missões e tarefas da Marinha.....	2
Figura 2 - Etapas da metodologia de investigação	4
Figura 3 - Estrutura da dissertação	6
Figura 4 – Processo de Extração-Transformação-Carregamento	12
Figura 5 – Arquiteturas de <i>data warehouse</i>	15
Figura 6 – Alternativas de implementação do <i>data warehouse</i>	15
Figura 7 - Esquemas da modelação multidimensional	19
Figura 8 - Algumas práticas a ter em conta na construção de um <i>dashboard</i>	21
Figura 9 - Arquitetura comum de um sistema BI	21
Figura 10 - Quadrante mágico de Gartner (comparação entre 2016 e 2017)	23
Figura 11 - Comparação de ferramentas <i>open source</i>	24
Figura 12 - Objetivos das diferentes fases do processo de fiscalização da pesca	28
Figura 13 - Extensão das zonas marítimas sob soberania ou jurisdição nacional	30
Figura 14 - Espaços marítimos sob soberania, jurisdição e responsabilidade nacional .	30
Figura 15 - Áreas de pesca (Divisão FAO)	31
Figura 16 - Exemplo de interpretação dos códigos ao respetivo significado	47
Figura 17 – Primeiro passo para evitar confusão na interpretação de dados.....	52
Figura 18 - Atualização da tabela principal, nomeadamente o campo referente à unidade naval	52
Figura 19 – Inserção de valores na tabela auxiliar que diz respeito à unidade naval	53
Figura 20 - <i>Query</i> que estabelece a relação entre o tipo de embarcação e o subtipo de embarcação	55
Figura 21 - Esquema em estrela, com as respetivas relações às dimensões	60
Figura 22 - Ferramentas utilizadas para o carregamento dos dados nas respetivas dimensões	62
Figura 23 - Tratamento da incoerência referente à diferenciação entre unidades navais e capitánias	62
Figura 24 - <i>Control flow</i> referente à tabela do fato	63
Figura 25 - Comando SQL para eliminar os dados existentes na tabela do fato.....	63
Figura 26 – Hierarquia respeitante à dimensão do tempo	65
Figura 27 - Informação organizada no cubo multidimensional, através do fato e respetivas dimensões	66
Figura 28 - Estrutura do <i>dashboard</i> , com a respetiva informação relevante distribuída	67
Figura 29 - Estrutura do dashboard com a informação selecionada para os <i>datasets</i>	68
Figura 30 - Informação selecionada para os <i>datasets</i> referente aos relatórios complementares ao dashboard.....	69
Figura 31 - Parâmetros estabelecidos para o <i>dashboard</i> e para os relatórios complementares	69
Figura 32 - Parâmetros estabelecidos para o relatório das artes de pesca	69
Figura 33 - Conjunto de gráficos complementares à ligação relativa à atividade da pesca (Perguntas 2 a 7, Secção B)	76
Figura 34 - Resultados da análise do gráfico referente aos tipos de embarcação	77

Figura 35 - Verificação das respostas no que diz respeito às nacionalidades inspeccionadas.....	77
Figura 36 - Análise de respostas obtidas relativamente às embarcações portuguesas vistoriadas.....	78
Figura 37 - Respostas obtidas referentes à solução edificada (Perguntas 1 e 2, Secção D)	81
Figura 38 - Análise de questões relativas à organização do <i>dashboard</i> e respetiva interface gráfica (Perguntas 3 e 4, Secção D).....	81
Figura 39 - Resultados obtidos no que diz respeito à envolvente do sistema (Perguntas 5,6,7 e 8, Secção D).....	82
Figura 40 - Resultados obtidos acerca da aplicação de soluções automatizadas (Perguntas 9, 10 e 11, Secção D).....	82
Figura 41 – Pirâmide do conhecimento	105
Figura 42 - Dados de interesse referentes ao consumo de combustível, a nível nacional	113
Figura 43 - Dados de interesse respeitantes às limitações operacionais.....	114
Figura 44 - Dados de interesse relativos às operações de busca e salvamento	115
Figura 45 - Dados de interesse respeitantes à navegação	116
Figura 46 - Tratamento de dados relativos a incongruências entre chave de ligação e respetiva expressão	117
Figura 47 - Tratamento dos dados referente à presença de dados decimais.....	118
Figura 48 - Esquema em estrela, pertencente ao fato da fiscalização da pesca.....	121
Figura 49 - Modelação dimensional referente às limitações operacionais	123
Figura 50 - Modelação dimensional aplicada nas operações de busca e salvamento... ..	123
Figura 51 - Modelação Dimensional relativa à navegação.....	124
Figura 52 - <i>Query</i> realizada para extrair os dados da fonte de dados para o fato.....	125
Figura 53 - Processo com as respetivas ferramentas utilizadas para o carregamento de dados na tabela do fato	126
Figura 54 - <i>Query</i> realizada para transformar valores nulos em zero, no que diz respeito às medidas do fato	127
Figura 55 - <i>Dashboard</i> referente à fiscalização da pesca	129
Figura 56 - Relatórios referente a fiscalizações realizadas por capitánias e por unidades navais	130
Figura 57 - Relatórios com informação complementar do subtipo de embarcação e das artes de pesca	130
Figura 58 - Relatório referente ao resultado das vistorias	130
Figura 59 - Relatório com informação complementar do tipo e quantidade de pescado	130

Índice de tabelas e gráficos

Tabela 1 - Comparação entre sistemas OLTP e sistemas OLAP	14
Tabela 2 - Quantitativos de infrações averiguadas pela Marinha Portuguesa, nos períodos entre 2015 e janeiro de 2018.....	35
Tabela 3 - Comunicados que contêm pelo menos um campo referente à fiscalização ..	43
Tabela 4 - Comunicados relativos à fiscalização e a respetiva finalidade.....	43
Tabela 5 - Formato do comunicado FISCREP	44
Tabela 6 - Organização dos dados por dimensão e respetivos atributos, referente ao fato	45
Tabela 7 - Dados recolhidos com a respetiva descrição	46
Tabela 8 - Tabelas auxiliares que dispõem de dados complementares	47
Tabela 9 - Dados recolhidos relativos ao tempo	48
Tabela 10 - Dados recolhidos com a verificação da <i>primary key</i> nas tabelas auxiliares	49
Tabela 11 - Ligação entre questões derivadas e questões do questionário, com o respetivo grau	84
Tabela 12 - Análise das respostas dos inquiridos em função das questões derivadas....	85
Tabela 13 - Definições de <i>Business Intelligence</i>	106
Tabela 14 - Identificação de dados e respetivos campos e mensagens, relativos à fiscalização	111
Tabela 15 - Compilação dos dados de interesse relativos ao fato	112
Tabela 16 - Classificação do tipo de dados relativos à tabela principal e às tabelas auxiliares.....	119
Gráfico 1 - Capturas nominais de pescado fresco ou refrigerado (ton), por arte de pesca nos períodos de 2015 e 2016	34
Gráfico 2 - Categoria dos inquiridos (Pergunta 3, Secção A)	75
Gráfico 3 - Ligação à atividade da fiscalização da pesca (Pergunta 1, Secção B)	75
Gráfico 4 - Verificação dos meses com maior número de fiscalizações	76
Gráfico 5 - Respostas obtidas relativamente à eficácia da fiscalização	78
Gráfico 6 - Análise de respostas obtidas acerca das áreas que compreendem presumíveis infratores.....	79
Gráfico 7 - Respostas obtidas referentes aos quantitativos dos presumíveis infratores .	79
Gráfico 8 - Resultados obtidos no que concerne aos resultados das vistorias.....	80
Gráfico 9 - Resultados obtidos numa tomada de decisão	80

Lista de abreviaturas, siglas e acrónimos

3NF	<i>Third Normal Form</i>
AIS	<i>Automatic Identification System</i>
AMN	Autoridade Marítima Nacional
APPS	Acordos de Parceria no domínio da Pesca Sustentável
BI	<i>Business Intelligence</i>
BPM	<i>Business Performance Management</i>
C2	Comando e Controlo
CAPSI	Conferências da Associação Portuguesa de Sistemas de Informação
CDC	<i>Changed Data Capture</i>
CECAF	<i>Central East Comission Atlantic Fishing</i>
CEMA	Chefe do Estado-Maior da Armada
CGPM	Comando-Geral da Polícia Marítima
CN	Comando Naval
CNUDM	Convenção das Nações Unidas sobre o Direito do Mar
CPS	<i>Cyber-Physical Systems</i>
CPU	<i>Central processing Unit</i>
CRM	<i>Customer Relationship Management</i>
CRP	Constituição da República Portuguesa
CTOI	Comissão dos Atuns do Oceano Índico
DAGI	Direção de Análise e Gestão da Informação
DDS	<i>Dimensional Data Store</i>
DGAM	Direção-Geral da Autoridade Marítima
DGRM	Direção-Geral de Recursos Naturais, Segurança e Serviços Marítimos
DM	<i>Data Mart</i>
DOLAP	<i>Desktop Online Analytical Processing</i>
DSR	<i>Design Science Research</i>

DW	<i>Data Warehouse</i>
E/R	<i>Entity Relationship</i>
EDW	<i>Enterprise Data Warehouse</i>
EIS	<i>Executive Information Systems</i>
ERP	<i>Enterprise Resource Planning</i>
ETL	<i>Extract, Transform and Load</i>
FA	Força Aérea
FAO	<i>Food and Agriculture Organization of the United Nations</i>
FEAMP	Fundo Europeu dos Assuntos Marítimos e das Pescas
GNR	Guarda Nacional Republicana
HOLAP	<i>Hybrid Online Analytical Processing</i>
ICCAT	Comissão Internacional para a Conservação do Atum Atlântico
ICCC	<i>International Conference on Intelligent Computing, Communication & Convergence</i>
INE	Instituto Nacional de Estatística
IoT	<i>Internet of Things</i>
KPI	<i>Key Performance Indicator</i>
LOMAR	Lei Orgânica da Marinha
MDB	<i>Multidimensional Databases</i>
MOLAP	<i>Multidimensional Online Analytical Processing</i>
MONICAP	Sistema de Monitorização Contínua da Atividade de Pesca
MP	Marinha Portuguesa
MTF	Mensagens de Texto Formatado
NAFO	Organização da Pesca do Atlântico Noroeste
N/A	Não Aplicável
N/D	Não Definido
NEAFC	Comissão da Pesca do Atlântico Nordeste
ODS	<i>Operational Data Stores</i>

OLAP	<i>Online Analytical Processing</i>
OLTP	<i>Online Transaction Processing</i>
OSI	<i>Open Source Initiative</i>
ORGP	Organizações Regionais de Gestão das Pescas
PCP	Política Comum das Pescas
QD	Questão Derivada
RAA	Região Autónoma dos Açores
RAM	Região Autónoma da Madeira
RCAAP	Repositórios Científicos de Acesso Aberto de Portugal
RCD	Rede de Comunicação de Dados
ROLAP	<i>Relational Online Analytical Processing</i>
RUN	Repositório da Universidade Nova de Lisboa
SA	<i>Staging Area</i>
SAD	Sistemas de Apoio à Decisão
SADAP	Sistema de Apoio à Decisão para a Atividade de Patrulha
SAR	<i>Search and Rescue</i>
SCD	<i>Slowly Changing Dimensions</i>
SciELO	<i>Scientific Electronic Library Online</i>
SDW	<i>Spatial Data Warehouse</i>
SI	Sistemas de Informação
SGBD	Sistema de Gestão de Bases de Dados
SIFICAP	Sistema de Fiscalização e Controlo das Atividades da Pesca
SME	<i>Small and Medium Enterprises</i>
SQL	<i>Structured Query Language</i>
TI	Tecnologias de Informação
UE	União Europeia
UN	Unidades Navais
WOLAP	<i>Web Online Analytical Processing</i>

ZEE	Zona Económica Exclusiva
ZMA	Zona Marítima dos Açores
ZMC	Zona Marítima do Centro
ZMM	Zona Marítima da Madeira
ZMN	Zona Marítima do Norte
ZMS	Zona Marítima do Sul

Introdução

O mar distingue Portugal, de forma política, económica, cultural e securitária, das outras nações (Lopes, 2011). A Marinha Portuguesa (MP) tem como missão principal, contribuir para que Portugal use o mar, na medida dos seus interesses (Marinha, 2018).

A Marinha é constituída por unidades em terra e no mar, que apoiam o cumprimento do conjunto de missões atribuídas à organização, e para o efeito produzem e trocam grandes quantidades de dados, na sua atividade operacional e no processo de tomada de decisão. Esses dados estão disponíveis através do conjunto amplo de fontes. No entanto, para que a organização deles beneficie, é necessário verificar a sua qualidade e tratamento, por forma a que a informação obtida seja fidedigna. Daí que, entre as linhas de ação da Diretiva de Planeamento da Marinha, tenha sido estabelecido “assegurar a evolução e/ou o desenvolvimento de sistemas de informação críticos para a atividade da Marinha, contribuindo para a operacionalização e automatização de processos, facilitando o acesso e a partilha de informação nas diferentes áreas funcionais, fatores essenciais para uma Marinha de futuro, que se quer cada vez mais eficiente na gestão de recursos, mais ágil no processo de decisão e mais eficaz na sua ação” (Marinha, 2017, p. 21).

A informação a recolher e a tratar deverá ser relevante para o processo de tomada de decisão. Para conduzir a recolha de dados, estabeleceram-se um conjunto de medidas de interesse, tendo em conta as necessidades de informação do Comando Naval (CN) e de outros organismos da Marinha.

Na presente dissertação, procede-se ao desenvolvimento das várias etapas que visam contribuir para a referida recolha e tratamento de informação para o processo de decisão.

A presente introdução é constituída por cinco secções. A primeira secção, visa o enquadramento do leitor na realidade da MP. Na segunda secção, é definido o problema em estudo. De seguida, enunciam-se os objetivos e as questões de investigação levantadas. A quarta secção justifica a metodologia de investigação escolhida e as atividades a desenvolver para a concretização dos objetivos estabelecidos. Na quinta secção, é apresentada a estrutura da dissertação, designadamente os vários capítulos em que está dividida e o respetivo conteúdo.

Enquadramento

A MP tem, de entre outros, como responsável pelo cumprimento das missões que lhe são atribuídas – o Comando Naval (Marinha, 2018).

O reconhecimento do papel da Marinha em Portugal e no Mundo, é obtido, essencialmente, pelas missões que realiza (vide Figura 1), nomeadamente: operações militares (combate à pirataria, por exemplo); fiscalização de pesca; ações de busca e salvamento, num espaço marítimo amplo; apoio à Autoridade Marítima Nacional (AMN);

e cooperação com a Direção-Geral de Recursos Naturais, Segurança e Serviços Marítimos (DGRM). A Marinha presta também apoio científico, em diversos projetos de Investigação, Desenvolvimento e Inovação através de parcerias com diversas entidades.



Figura 1- Missões e tarefas da Marinha¹

A missão da Marinha é pois, em síntese, “contribuir para que Portugal use o mar, na medida dos seus interesses”, fazendo “duplo uso”, de forma equilibrada e otimizada, dos seus meios (Lopes, 2011). O cumprimento da missão tem como requisito “efetivas capacidades de Comando e Controlo” (C2), sustentadas por Sistemas de Informação (SI)” (Marinha, 2005).

Definição do Problema

Laudon & Laudon (2011) definem SI como um “conjunto de elementos ou componentes inter-relacionados que reúnem, armazenam, processam e distribuem dados e informação”, com as seguintes finalidades: “apoiar as funções e a tomada de decisão, comunicação, coordenação, controlo, análise e visualização da organização”.

A decisão eficiente tem de ser suportada pelo SI da organização. Uma das principais funções do SI é disponibilizar informação correta e acessível *quando, onde, na qualidade e quantidade* necessárias à pessoa certa (Agarwal & Dhar, 2014; Bernus 1998, como referido em Vasconcelos, Caetano, Sinogas, Mendes, & Tribolet, 2002).

A tomada de decisão é um processo complexo que requer o panorama situacional mais clarificado possível. Uma forma de clarificar esse panorama é ter informação

¹ Fonte: Marinha, s.d.a, p.7.

objetiva e sumariada. Para o efeito, torna-se necessário a existência de ferramentas que facilitem o processo de tomada de decisão.

As unidades da Marinha recebem um grande volume de dados, que devem ser transformados em informação relevante para a tomada de decisão. A presente dissertação pretende, pois, endereçar o problema da produção e visualização de informação para a tomada de decisão nas unidades de Marinha.

Objetivo e Questões de Investigação

O objetivo principal da presente dissertação é a conceção de um sistema de informação de gestão (*Business Intelligence* – BI), solução que contribua para que, os diferentes níveis de gestão das unidades da Marinha, obtenham uma visão unificada da informação relevante para a tomada de decisão.

A questão de investigação que irá nortear o desenvolvimento da presente dissertação será a seguinte:

Como efetuar a criação de um sistema de informação de gestão, que permita a tomada de decisão, nas várias áreas de atuação da Marinha?

Para obtenção da resposta à questão proposta, foram formuladas diferentes questões derivadas (QD), designadamente:

- O sistema de informação de gestão concebido e implementado, fornece informação, relativamente à dimensão selecionada, necessária à tomada de decisão?
- A organização, visualização e interpretação da informação contida no *dashboard* é a adequada durante o processo de tomada de decisão?
- O sistema de informação de gestão possui limitações que afetem o potencial do sistema no apoio à tomada de decisão das atividades de fiscalização da pesca?

As respostas a estas questões derivadas contribuirão para verificar se foi obtida resposta à questão de investigação principal.

Metodologias de Investigação

O desenvolvimento da presente dissertação tem por base dois referenciais metodológicos: um científico e outro de desenvolvimento de um artefacto de *software*. É importante mencionar que um artefacto é a “organização dos componentes do ambiente interno para atingir objetivos num determinado ambiente externo” (Zaidan, Bax, & Parreiras, 2016, p. 3763).

O referencial científico justifica-se porque se está na presença de uma dissertação, e por isso tem de ter por base uma metodologia de investigação científica. O desenvolvimento do artefacto de *software*, será obtido através da utilização de uma metodologia de engenharia de software, em que as fases que a compõem iterativas, de forma a validar passo a passo a conceção, implementação e validação do artefacto.

Posto isto, considerando as necessidades acima referidas, foram escolhidas as metodologias: *Design Science Research* (DSR) e SCRUM.

O DSR é o método de investigação que serve de suporte ao desenvolvimento de artefactos, que integra soluções destinadas à resolução de um problema específico. O foco do DSR não é a ação em si, mas sim o conhecimento que se obtém com o desenvolvimento do projeto (Lacerda, Dresch, Proença, & Júnior, 2013; Zaidan *et al.*, 2016).

Desta forma, o que surge como resultado da aplicação desta metodologia são artefactos (Haj-bolouri, 2015). Esses podem ser classificados em: conceitos (Lacerda *et al.*, 2013) e/ou constructos, modelos, métodos e instanciações (como referido em Bax, 2013; Hevner, March, Park, & Ram, 2004⁹²). Na presente dissertação, o tipo de artefacto que se pretende obter é um sistema de informação de gestão (BI), pelo que se caracteriza como uma instanciação. As instanciações mostram que constructos, modelos e métodos podem ser implementados num sistema de trabalho, permitindo uma avaliação do impacto do artefacto no seu ambiente, ou seja, se é adequado à finalidade pré-definida e como os utilizadores interagem com o novo sistema (Hevner *et al.*, 2004, p. 74).

Apesar da diversidade das abordagens de autores² que formalizaram o método DSR, a estrutura a seguir nesta dissertação foi baseada em Roel Wieringa. Wieringa orienta a sua proposta fornecendo diretrizes, esclarecimentos e soluções para problemas (Zaidan *et al.*, 2016).

A resolução de problemas, surge numa estrutura lógica - o ciclo regulador. O ciclo regulador padronizado por Wieringa (2009), pode ser aplicado em quaisquer áreas (Zaidan *et al.*, 2016, p. 3761), e é composto por 5 etapas distintas. As etapas são aplicadas à presente dissertação. No entanto, por se estar na presença do desenvolvimento de um produto de *software*, as etapas que envolvem o desenvolvimento do artefacto serão suportadas pela metodologia SCRUM (Sverrisdottir, Ingason, & Jonasson, 2014, p. 264), conforme se pode observar na Figura 2.

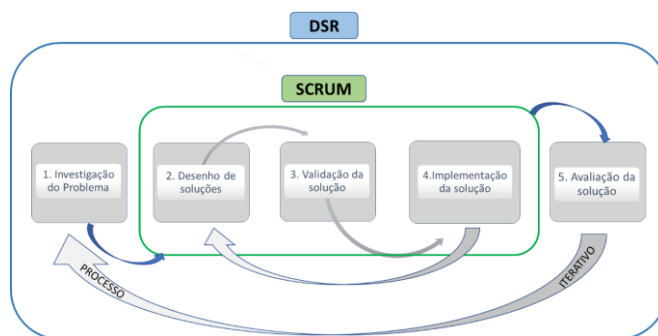


Figura 2 - Etapas da metodologia de investigação³

² Bunge (1980), Eekels & Roozenburg (1991), Manson (2006), Gregor & Jones (2007), Vaishnavi & Kuechler (2009), Aken (2011), entre outros.

³ Fonte: Baseado em Wieringa (2009; 2014).

As etapas da metodologia de investigação seguidas durante a elaboração da dissertação foram as seguintes:

1. Investigação do problema: formulação da questão de investigação e descrição do problema.

O resultado desta primeira etapa consta da *Introdução* e dos capítulos de *Revisão de Literatura* e *Caracterização do Domínio do Problema*.

2. Desenho da solução, Validação da Solução e Implementação da Solução

O resultado destas etapas consta dos capítulos de *Recolha e Tratamento de Dados* e da *Conceção da Solução*.

A realização destas etapas, seguiram os preceitos da SCRUM segundo, Schwaber & Sutherland (2013) designadamente:

- Os esboços da solução foram planeados e realizados em ciclos, designados de *sprints*⁴;
- Na *sprint*, foi realizado o trabalho estabelecido, designadamente: levantamento de requisitos; investigação de ferramentas; recolha de dados em folhas de cálculo, tratamento de dados; definição, validação e implementação (meta da *sprint*) do artefacto; entre outros;
- Para a *sprint* realizaram-se: (1) reuniões de planeamento, onde se efetuou uma análise de progressos e de obstáculos detetados; (2) reunião de revisão, *Sprint Review Meeting*, onde se observou o resultado atingido relativamente à implementação do artefacto.

3. Avaliação da implementação

Esta última etapa está documentada nos capítulos da *Análise de Resultados e Conclusão*;

A fase da avaliação consistiu essencialmente em verificar se os objetivos e os requisitos do artefacto correspondem às necessidades. Uma das formas de avaliar a funcionalidade do artefacto é verificar se atende às necessidades dos “utilizadores-alvo”.

Estrutura

A presente dissertação de mestrado é constituída por sete partes (conforme Figura 3):

⁴ *Sprints* são iterações, ou seja, um ciclo repetitivo de trabalhos que produz um incremento de produto.

- **Introdução** – Breve enquadramento, identificação do problema em estudo, objetivos e questões de investigação, e metodologia de investigação.
- **Capítulo 1 – Revisão da Literatura.** Levantamento e análise do estado da arte sobre o tema, referência aos conceitos e ao processo de conceção da solução, e por fim, a escolha da ferramenta a ser usada na implementação da solução.
- **Capítulo 2 – Caracterização do Domínio do Problema.** Identificação das medidas de interesse escolhidas. Caracterização da medida de interesse selecionada: *Fiscalização da Pesca*.
- **Capítulo 3 – Recolha e Tratamento de Dados.** Levantamento dos dados através da definição do conteúdo das Mensagens de Texto Formatado (MTFs) e das Mensagens Estruturadas. Explicação dos diversos passos efetuados para a extração de informação relevante das fontes de dados.
- **Capítulo 4 – Conceção da Solução.** Instanciação do repositório de dados e da aplicação de visualização dos dados.
- **Capítulo 5 – Análise de Resultados.** Validação do artefacto, análise dos resultados obtidos, e resposta às questões de investigação.
- **Conclusões** – Elaboração das conclusões com base nos resultados obtidos e recomendações para trabalhos futuros.

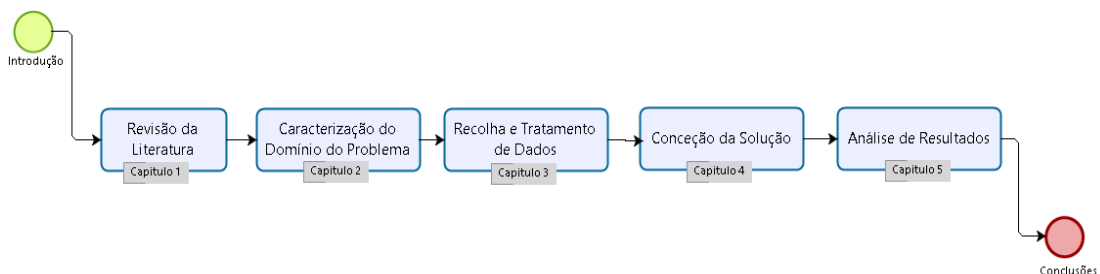


Figura 3 - Estrutura da dissertação

CAPÍTULO 1

REVISÃO DA LITERATURA

1.1. Estado da Arte

1.2. Levantamento de Ferramentas

1. Capítulo 1 – Revisão da Literatura

1.1. Estado da Arte

1.1.1. Introdução ao *Business Intelligence*

A tomada de decisão requer a análise se os objetivos definidos para a organização estão a ser alcançados e da eficiência da estratégia definida para os atingir. A forma de atingir os objetivos pode ter que ser alterada face a alterações na envolvente da organização. A decisão sobre que alterações efetuar enfrenta a dificuldade de obtenção de informação útil a partir de “uma montanha de dados” (Dalfovo & Tamborlim, 2017, p. 28). O *Business Intelligence* pretende ultrapassar a dificuldade da necessidade de informação relevante para a tomada de decisão, tanto nas grandes (Ivan, 2014, p. 176), pequenas e médias empresas (*Small and Medium Enterprises – SME*). (Ali, Khan, & Miah, 2017, p. 163; Guarda, Santos, Pinto, Augusto, & Silva, 2013, p. 187), como nas instituições públicas (Vo *et al.*, 2017, p. 1).

Para se entender o BI, é necessário compreender os seus conceitos base (vide para o efeito o Apêndice A).

Nelson (2007), considera que o termo BI está relacionado com “conhecimento prévio” (*foreknowledge*). Esse termo foi referido por Sun Tzu (autor da obra “A Arte da Guerra”, no ano 513 a.C.), que defendeu que, o que os governantes e generais sábios conseguiram conquistas através do “conhecimento prévio” (Giles, 2013, Capítulo XIII). Também os autores clássicos Nicolau Maquiavel (autor da obra “O Príncipe”, no ano 1513) e Carl von Clausewitz (autor da obra “Da Guerra”, no ano 1832), defendem que “conhecimento é vantagem”. Estes clássicos por via de uma análise estratégica, realçam a importância de conhecer as forças e fraquezas da organização e dos concorrentes, no meio em que se atua (Sezões, Oliveira, & Baptista, 2006, p. 8).

Devens utilizou pela primeira vez o termo BI ao descrever o caso do banqueiro, Sir Henry Furnese, que ao receber informação antecipada conseguiu obter vantagem sobre os seus concorrentes (Devens, 1865, p. 210). No entanto, a primeira utilização do termo BI, foi efetuada por Luhn (vide Resenha Histórica).

Para além da primeira definição, ao longo dos anos surgiram várias definições de BI, usadas nos mais diversos contextos. Apesar da diversidade de definições, realçam-se as seguintes:

- “conceito que engloba um vasto conjunto de aplicações de apoio à tomada de decisão que possibilitam um acesso rápido, partilhado e interativo da informação, bem como a sua análise e manipulação; através destas ferramentas, os utilizadores podem descobrir relações e tendências e transformar grandes quantidades de informação em conhecimento útil” (Sezões *et al.*, 2006, p. 10);

- “os processos, as tecnologias e as ferramentas necessárias para transformar dados em informação, informação em conhecimento, e conhecimento em planos de ação. BI engloba *data warehousing*, ferramentas analíticas e gestão de conteúdo/conhecimento” (Eckerson, 2002, p. 6);
- Todesco & Carretero & Duran declaram BI como: “termo contemporâneo usado para se referir a um conjunto de tecnologias de informação (plataformas, aplicações e processos), que visa facilitar a tomada de decisões em todos os níveis” (como referido em Lucas, Café, & Viera, 2016, p. 172).

O dia-a-dia numa organização é caracterizado por inúmeras ocorrências⁵, sendo cada uma delas representa uma transação. O armazenamento, controlo e processamento das transações, são realizados por Sistemas de Processamento Transaccional (*Online Transaction Processing – OLTP*) (Bog, 2014; Larson, 2009), também designados de Bases de Dados Operacionais (*Operational Data Stores – ODS*) (Cody, Kreulen, Krishna, & Spangler, 2002, p. 698; Santos & Ramos, 2009, p. 79). O OLTP é assim uma das fontes de dados operacionais de uma organização. Estes sistemas possibilitam a inserção e consulta de dados, suportando o rápido crescimento do volume de dados (Alvaro, 2014; Scheps, 2008 ; Cunha, 2008).

Os dados inseridos num OLTP são denominados de “dados primitivos” (Inmon, 2002, p. 15) e/ou “dados operacionais” (Santos & Ramos, 2009, p. 79). Esse sistema não são os mais adequados para os requisitos do BI. Esses requisitos requerem a “agregação das transações individuais” (Larson, 2009, p. 27). Por isso, surgiram outro tipo de sistemas - os Sistemas de Informação Executiva (*Executive Information Systems – EIS*).

Segundo Power (2006), os EIS evoluíram a partir dos Sistemas de Apoio à Decisão (SAD). Os EIS possuem alguns meios para tratar informação interna e externa relevante, em forma de relatórios dinâmicos multidimensionais, previsões, análise de tendências, entre outros (Dalfovo & Tamborlim, 2017, pp. 27 e 32).

1.1.2. Data Warehouse

Inmon, em 1996 (Santos & Ramos, 2009, p.80), definiu um outro sistema designado de Armazém de Dados (*Data Warehouse – DW*) como sendo “um conjunto de dados baseado em assuntos, integrado, não volátil e variável em relação ao tempo, que apoia na gestão de decisões” (como referido em Inmon & Krishnan, 2011, p. 42). Um sistema de DW tem como foco a incorporação de informação, vinda de várias fontes, para posteriormente a “tratar, formatar e consolidar numa única estrutura de dados” (Ferreira, Miranda, Abelha, & Machado, 2010, p. 757).

O DW é considerado a “peça chave do *Business Intelligence*” (Parra, Galeano, & B., 2010, p. 36).

⁵ Controlo da produção, prestação de serviços, pagamentos ou recebimentos.

Os dados que se pretendem colocar num DW estão presentes nos referidos sistemas OLTP. Como já foi referido, são sistemas fontes de dados que funcionam como bases de dados operacionais das organizações (Ballard, Farrell, Gupta, Mazuela, & Vohnik, 2006, p. 62; Vo *et al.*, 2017, p. 1). As fontes de dados (Moody & Kortink, 2000, p. 5 -1) também podem ser de Sistemas de Informação Empresariais (*Enterprise Resource Planning* – ERP) (dados internos), Sistema de Gestão da Relação com Clientes (*Customer Relationship Management* – CRM) (dados externos), e sistemas de gestão de documentos, processos de negócios, folhas de cálculos, entre outros (Cody *et al.*, 2002; Negash, 2004).

Para que os dados possam ser transferidos dos sistemas em que têm origem para um DW, é necessário recorrer ao processo de “Extração-Transformação-Carregamento” (*Extract-Transform-Load* – ETL) (Rainardi, 2008). Esse processo consome cerca de 70% do processo de implementação de um DW (Kimball & Caserta, 2004, p. 23) e permitem a deteção e a remoção de erros dos dados, a transformação desses dados em informação, assegurando a qualidade desses dados (Xavier & Moreira, 2013, p. 382). Os dados são assim, “a base do processo” (Denilson da Silva, Lemos, Paulino, & Kannenberg, 2016, p. 75) de edificação de um DW.

O processo de ETL consiste no seguinte conjunto de atividades:

- Ligação aos sistemas fontes de dados e **extração dos dados**. Conforme descrito por Kimball, Reeves, Ross, & Thornthwaite (em 1998), o processo da extração compreende duas etapas (como referido em El-Sappagh, Hendawi, & El Bastawissy, 2011, p. 93):
 - Extração Inicial (*Initial extraction*) – Os dados são inseridos no DW pela primeira vez. Esta fase apenas ocorre uma vez.
 - Alteração dos dados extraídos (*Changed data extraction/changed data capture* – CDC) – Os dados modificados e adicionados aos sistemas de origem desde a última extração são inseridos no DW. Esta fase ocorre consoante a periodicidade das atualizações e necessidades da organização.
- **Transformação dos dados** – os dados são disseminados para uma área temporária (*Staging Area* – SA), onde ocorre a transformação dos dados, que consiste na limpeza dos dados (*data cleansing*), isto é, a deteção, correção e consolidação dos dados englobando também a eliminação de dados indefinidos. Os dados ao sofrerem alterações, apenas ficam alterados no sistema de destino e não no sistema de origem (Ferreira *et al.*, 2010, p. 760). Esta fase é fundamental e uma das mais críticas pois envolve o processamento de grandes volumes de dados, estes que, muitas das vezes, são extraídos de duas ou mais fontes, com “codificações, estruturas e armazenamentos diferentes”. Nesta fase, os dados devem tornar-se precisos, corretos, consistentes e sem ambiguidade.

- **Carregamento dos dados** – os dados extraídos e transformados são introduzidos num sistema de destino – uma estrutura dimensional⁶. O sistema de destino é designado de Base de Dados Dimensional (*Dimensional Data Store – DDS*) (Rainardi, 2008, p. 2).

A Figura 4 representa o processo ETL, com base em: Denilson da Silva *et al.* (2016); El-Sappagh *et al.* (2011); Ferreira *et al.* (2010); Inmon (2002); Kimball & Caserta (2004); Kimball & Ross (2013); Rainardi (2008); Valacich & Schneider (2017).

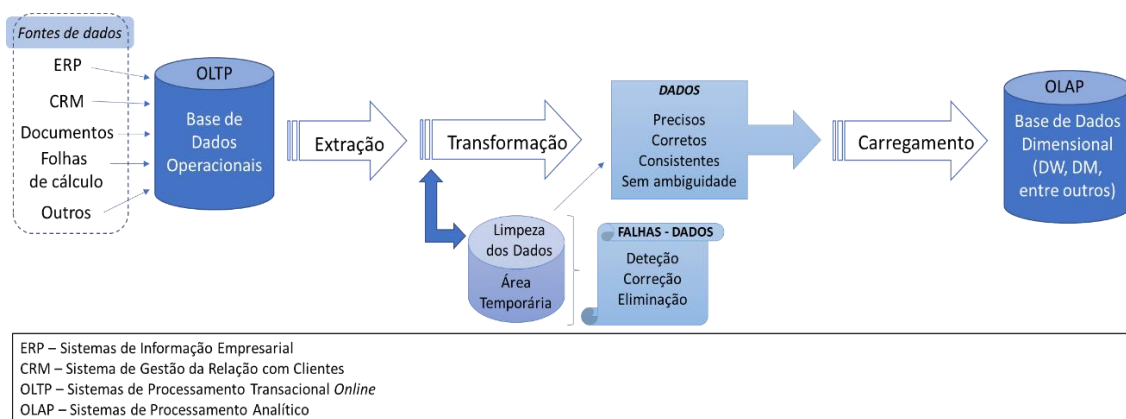


Figura 4 – Processo de Extração-Transformação-Carregamento

A razão pela qual se colocam os dados dos sistemas de origem na DDS e a consulta dos dados se realiza também nessa base de dados, é porque estando os dados integrados e organizados num formato dimensional, permitem uma análise mais adequada na tomada de decisão (Rainardi, 2008, p. 2).

O processo ETL estabelece a relação entre os Sistemas OLTP e os Sistemas de Processamento Analítico (*Online Analytical Processing – OLAP*) (Vo *et al.*, 2017, p. 3). Os dados do OLTP são processados e enviados para o sistema OLAP para possibilitar a análise da informação (Akita, 2006, p. 103; Braghittoni, 2017).

Para se criar essa transferência de dados, são aplicadas técnicas de modelação. Kimball defende que as técnicas tradicionais de modelação de sistemas transaccionais não devem ser utilizadas em ambientes de *data warehousing*, nomeadamente o Modelo Entidade-Relação (*Entity Relationship Model – E/R*) (Moody & Kortink, 2000, p. 5 -2). Com esta técnica, geralmente normaliza-se até à *Terceira Forma Normal (Third Normal Form – 3NF)*, sendo cada coluna funcionalmente dependente apenas da chave primária (Ballard *et al.*, 2006, pp. 49–51).

No entanto, Moody & Kortink (2000) defendem o contrário, reconhecendo a complexidade que existe em combinar essa técnica nos ambientes, transaccional e de *data warehousing*. O E/R permite obter estruturas normalizadas dos dados, típica dos sistemas

⁶ Não é necessariamente um DW, pode ser um *Data Mart* ou uma outra base de dados, por exemplo.

transacionais. Num ambiente de *data warehousing*, é benéfico efetuar a desnormalização dos dados. No entanto, é necessário o controlo adequado dos dados não normalizados.

Aplicada a técnica de modelação, os dados são enviados para o sistema OLAP, também intitulado de Sistema de Gestão de Bases de Dados Multidimensionais (*Multidimensional Database Management System – MDBMS*) (Inmon, 2002, p. 185). Trata-se de uma tecnologia destinada ao processamento e análise, de forma interativa, dos dados armazenados no DW (Rainardi, 2008, p. 14; Shen, Liu, Chen, & Wang, 2012, p. 398). O OLAP contém “dados analíticos” (Caldas, Scandelari, & Kovaleski, 2006, p. 2) e/ou “dados derivados”, que são os dados transacionais transformados para análise e suporte à decisão da organização (Caldas *et al.*, 2006, p. 2; Inmon, 2002, p. 15).

Os servidores de base de dados que suportam este tipo de processamento dos dados, designam-se por (Denilson da Silva *et al.*, 2016, p. 80; Santos & Ramos, 2009, p. 95):

- *Relational OLAP (ROLAP)* – usam como suporte uma base de dados relacional. Este tipo de base de dados permite flexibilidade no acesso e alteração dos dados;
- *Multidimensional OLAP (MOLAP)* – o processamento dos dados é feito através de uma base de dados multidimensional, representada por dimensões e fatos. Este tipo de base de dados possibilita o cruzamento de informação;
- *Hybrid OLAP (HOLAP)* – combinação das tecnologias ROLAP e MOLAP, usufruindo do melhor de cada uma delas. Por exemplo, um servidor HOLAP pode armazenar os dados numa base de dados relacional, e, por seu turno, as operações entre os dados podem ser sustentadas pelo servidor MOLAP.

Segundo Inmon, a tipologia da implementação das ferramentas OLAP pode ser ainda ser classificada em (como referido em Araújo, Batista, & Moreira De Magalhães, 2007):

- *Desktop OLAP (DOLAP)* – implementação que possibilita aos utilizadores o acesso à base de dados multidimensional ou a parte dela, uma vez que sendo criada no servidor é transferida para o *desktop* (Cavalcanti, Oliveira & Monteiro, em 1998 como referido em Araújo *et al.*, 2007);
- *Web OLAP (WOLAP)* – servidor com recurso às ferramentas OLAP, através da tecnologia de *internet* (Andreato, em 1999 como referido em Araújo *et al.*, 2007).

As principais diferenças entre os dois sistemas enunciados acima, OLTP e OLAP, estão resumidas na Tabela 1:

Tabela 1 - Comparação entre sistemas OLTP e sistemas OLAP⁷

OLTP	OLAP
Aplicações do negócio	Aplicações sobre o negócio
Dados possuem alto nível de detalhe	Dados sumarizados
Dados atuais e suscetíveis a modificação e remoção	Dados históricos que não sofrem alteração (exceto por motivo de erro ou falta de consistência da informação)
Atualização dos dados é feita no momento da transação – frequência muito alta	Atualização da informação é feita no processo de carregamento dos dados – frequência baixa
Objetivos operacionais	Registo histórico
Milhares de utilizadores	Poucos utilizadores (gestores e analistas com responsabilidade na tomada de decisão)
Acesso a poucos registos de cada vez	Acesso a muitos registos de cada vez
Permite leitura, inserção, modificação e eliminação dos dados	Permite a inserção e leitura dos dados
Consultas simples a complexas	Geralmente, consultas muito complexas
Análise da execução operacional	Análise da tomada de decisão

Ao repositório de dados, menor do que um DW, designa-se de *Data Mart*. Os DMs compreendem apenas a informação de uma determinada área da organização (Coelho *et al.*, 2016, p. 765), o que permite uma maior rapidez na respetiva consulta e análise (Vo *et al.*, 2017, p. 1). Um DW pode compreender centenas de *terabytes* de dados enquanto que, um DM cifra-se nos 10 a 100 *gigabytes* (Cody *et al.*, 2002, p. 698).

Os DMs podem ser independentes ou dependentes de um DW, consoante a arquitetura do DW (Ballard *et al.*, 2006, p. 40).

Segundo Ballard *et al.* (2006) a arquitetura DW-DMs (Figura 5) tem as seguintes configurações possíveis:

- *Enterprise Data Warehouse (EDW)* – repositório de dados, onde os dados estão logicamente centralizados e, por isso, pode também ser chamada de *Hub and Spoke*. Neste tipo de arquitetura, os DMs não estão inseridos num DW (Kimball & Caserta, 2004, p. 27);
- *Independent Data Marts* – os dados analíticos são organizados para dar resposta a necessidades específicas relativas a um determinado departamento. Neste tipo de

⁷ Fonte: Baseado em Caldas *et al.*, 2006; Denilson da Silva *et al.*, 2016; Teorey, Lightstone, Nadeu & Jagadish, 2013; Santos & Ramos, 2009.

arquitetura, não existe relação entre DMs, o que pode levar a resultados inconsistentes (Ballard *et al.*, 2006, p. 60; Kimball & Ross, 2013, p. 27);

- *Dependent Data Marts* – os DMs são implementados em vários departamentos, mas nesta arquitetura existe integração, o que possibilita uma visão unificada dos dados (Ballard *et al.*, 2006, p. 61).

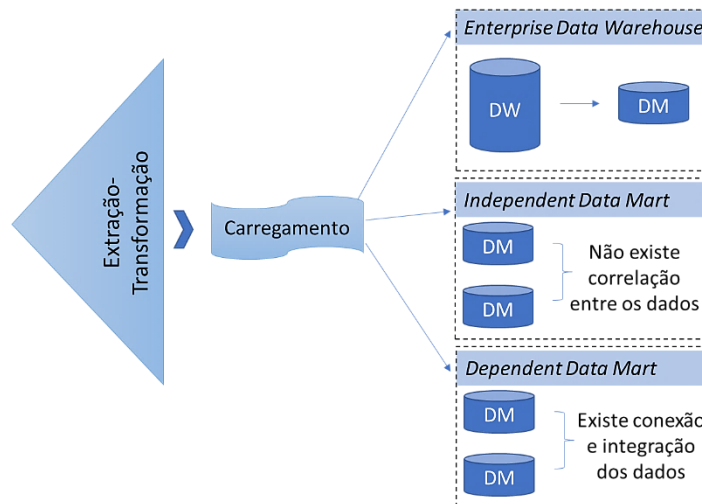


Figura 5 – Arquiteturas de *data warehouse*

A arquitetura de DW pode ser implementada de três formas distintas (ver Figura 6) (Vercellis, 2009, p. 52):

- *Top-down* (defende Inmon) – os dados dos DMs são extraídos de um DW integrado, ou seja, constrói-se primeiramente o DW. Apesar de ser o padrão geral, implica um desenvolvimento demorado (Denilson da Silva *et al.*, 2016, p. 79; Vercellis, 2009, p. 52);
- *Botton-up* (defende Kimball) – os dados inseridos nos vários DMs são enviados para o DW, com a finalidade de se construir um DW a partir de diversos DMs. Este tipo de implementação é mais rápido que o anterior, mas não proporciona uma visão global da arquitetura, e por os DMs serem independentes pode dificultar a integração (Denilson da Silva *et al.*, 2016; Vercellis, 2009);
- Combinação das duas anteriores (*Mixed*)– baseada no modelo usual de DW, com a diferença de que a implementação ocorre de forma sequencial consoante a necessidade das várias áreas de interesse (Vercellis, 2009, p. 53).

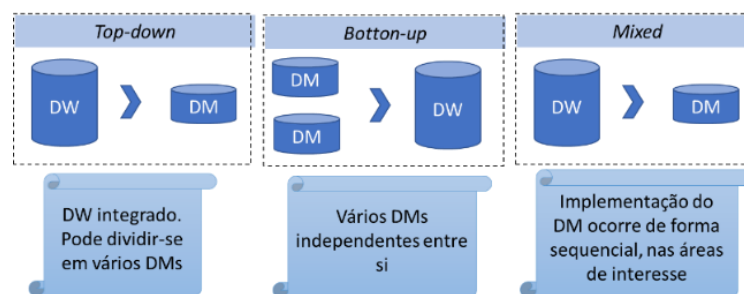


Figura 6 – Alternativas de implementação do *data warehouse*

1.1.3. Modelação Dimensional

O sistema de *Data Warehouse* assenta na modelação dimensional, (segundo Kimball em 1996, referido em Moody & Kortink, 2003, p. 8), também designada por modelação multidimensional (Santos & Ramos, 2009, p. 82). Nesta modelação, os dados estão organizados de modo a facilitar o processamento e interpretação de consultas em diferentes perspetivas. Essa modelação pode ser alcançada através da implementação de: cubos multidimensionais, esquemas em estrela, esquemas em floco de neve ou esquemas em constelação.

Para compreender os vários aspetos referidos e entender como é organizada essa modelação, é necessário clarificar alguns conceitos, tais como:

- Tabela de Fatos – contempla atributos numéricos (fatos) e chaves estrangeiras (*foreign keys*) (Ballard *et al.*, 2006, p. 53), que expressam a ligação entre a tabela de fatos e as tabelas de dimensões que lhe estão associadas (Kimball & Ross, 2002, pp. 16–19; Santos & Ramos, 2009, p. 83);
 - Fatos – assunto que se pretende analisar. Associados a acontecimentos, isto é, valores numéricos que o utilizador pretende analisar relativo a determinadas medidas (Santos & Ramos, 2009, p. 87; Valacich & Schneider, 2017, p. 269);
- Medidas – correspondem às linhas que compõem a tabela de fatos. Devem ter o mesmo nível de granularidade da informação armazenada (Kimball & Ross, 2002, p. 17; Santos & Ramos, 2009, p. 87).
- Tabela de Dimensões – relacionadas com a tabela de fatos, da qual se pretende uma análise específica (quem, quando, onde, porquê, como). É constituída por atributos (Santos & Ramos, 2009, p. 83), devendo conter poucos códigos e privilegiar atributos descritivos (Kimball & Ross, 2002, p. 21).
 - Dimensões – cada dimensão é definida por uma chave primária relacionada com a tabela de fatos (Kimball & Ross, 2002, p. 19).
 - Atributos – correspondem às linhas da tabela de dimensões. Principal fonte de restrições quando se efetuam consultas ou relatórios. Devem estar organizados de forma hierárquica para possibilitar uma análise mais detalhada. “*O data warehouse é melhor quanto melhor os atributos de dimensão*” (Kimball & Ross, 2002, p. 20).

Para além disso, há a salientar algumas considerações relativamente aos conceitos explicitados, que são:

Fatos:

- A análise pode ser feita da mesma forma para todos os fatos, se estes possuírem as mesmas características, mas por vezes isso não acontece. Por isso, dividem-se da seguinte forma (Santos & Ramos, 2009, pp. 87–90):

- Fatos aditivos – podem ser agregados⁸ por todas as dimensões que estão conectadas às tabelas de fatos;
- Fatos semi-aditivos – podem ser agregados apenas por uma ou mais das dimensões existentes, mas não por todas;
- Fatos não-aditivos – não podem serem agregados por nenhuma das dimensões presentes na estrutura, pois não têm um sentido lógico para tais operações. Exemplos: preços, percentagens, valores unitários, entre outros.

Dimensões:

- Os atributos que compõem a tabela de dimensões são alvos de uma constante análise. Esses, por sua vez, podem alterar ao longo do tempo, apesar de não ser frequente e, por isso recorre-se a técnicas de modificação de dimensões (*Slowly Changing Dimensions* – SCD). Isto porque a partir do momento que os dados são inseridos no DW, a única operação possível sobre os dados é de leitura (exceto nas fases de carregamento e atualização do DW, já explicitada no processo ETL).
- Essas alterações devem ser pontuais. As técnicas para a atualização dos atributos, isto é, as SCD, são as seguintes (Kimball & Ross, 2002, pp. 95–102; Santos & Ramos, 2009, p. 91):
 - Tipo 1 – Substituição do Valor (*Overwrite the Value*) – substituição de um ou mais atributos, sendo sempre uma atualização. O processo ETL escolhe esta técnica para correção dos dados quando não há interesse em armazenar valores históricos (Kimball & Caserta, 2004, p. 185). Implementação simples, mas não guarda qualquer valor de atributo anterior (Kimball & Ross, 2002, p. 97);
 - Tipo 2 – Acréscimo de uma linha na Dimensão (*Add a Dimension Row e/ou Creating another dimension record*) – a alteração é feita através da inserção de uma nova linha, isto é, um novo atributo da dimensão, sendo assim possível verificar o valor anterior e a nova atualização. Para se distinguir ambos os atributos, aplicam-se duas chaves de substituição diferentes. Implementação complexa, mas guarda todo o histórico (Santos & Ramos, 2009, p. 92);
 - Tipo 3 – Acréscimo de uma coluna da Dimensão (*Add a Dimension Column e/ou Creating a new current value field*) – acrescenta um número definido de colunas que são utilizadas aquando da alteração, ou seja, existe a coluna com o valor inserido e outra coluna sobre o mesmo assunto, caso ocorra algum tipo de alteração. Técnica que possibilita observar o novo atributo e o anterior, mas não guarda atributos históricos caso ocorram mais alterações (Santos & Ramos, 2009, p. 92). Por isso, não permite fazer uma análise do impacto das alterações ocorridas, caso estas existam (Kimball & Ross, 2002, p. 101).

⁸ O termo “agregados” consiste em operações (somatório, contagem, entre outros), sendo que a sua aplicação depende do objetivo da análise.

- Outras técnicas, mais complexas, flexíveis e com um custo mais elevado que as técnicas anteriores, são as designadas de técnicas híbridas (*hybrid techniques*) (Kimball & Ross, 2002, p. 102):
 - *Predictable Changes with Multiples Version Overlays* – sucessão de alterações a realizar num determinado atributo. O tipo 2 seria o mais aconselhável, mas não é possível quando os requisitos das respetivas alterações são muito elaborados. Exige uma combinação entre os tipos 2 e 3 (Kimball & Ross, 2002, pp. 102–103);
 - *Unpredictable Changes with Single-Version Overlay* – Utilizada quando existe a necessidade de preservar dados históricos, conseguindo comparar e compreender a evolução através do acesso aos valores atuais. Para isso, recorre-se a duas colunas, a coluna com os valores atuais e, a outra coluna com os valores históricos. Assim, utilizam-se: o tipo 2 para criar a nova linha, que contém o valor anterior e o atual, empregando diferentes chaves; o tipo 3 que adiciona a coluna para inserir a nova atualização; e o tipo 1 que escreve por cima da anterior, visto que o valor histórico já pertence a outra coluna (Kimball & Ross, 2002, pp. 103–105).

Após a contextualização dos conceitos relativos à modelação multidimensional, analisam-se os diferentes tipos de análise de dados, que os vários esquemas (Figura 7) proporcionam:

- Esquema em estrela – esquema mais comum, principalmente quando se trata de consultas volumosas. Compreende uma única tabela de fatos (centro da estrela), e estabelece ligação com as várias tabelas de dimensões associadas, sendo que estas não são normalizadas (Ballard *et al.*, 2006, p. 52; Santos & Ramos, 2009, p. 82);
- Esquema em floco de neve – em termos do conteúdo de dados é equivalente ao esquema em estrela. No entanto, as dimensões são normalizadas, o que evita o armazenamento de informação redundante. Por ter uma estrutura mais complexa, isto é, especifica a estrutura de cada dimensão, pode dificultar a compreensão do utilizador e, por seu turno, é vantajosa porque explicita as dimensões de forma hierárquica (Moody & Kortink, 2003, p. 9; Santos & Ramos, 2009, p. 85);
- Esquema em constelação – Pode ser visto como um conjunto de esquemas em estrela, que possui múltiplas tabelas de fatos unidas através das dimensões em comum (Ballard *et al.*, 2006, p. 57; Santos & Ramos, 2009, p. 86).

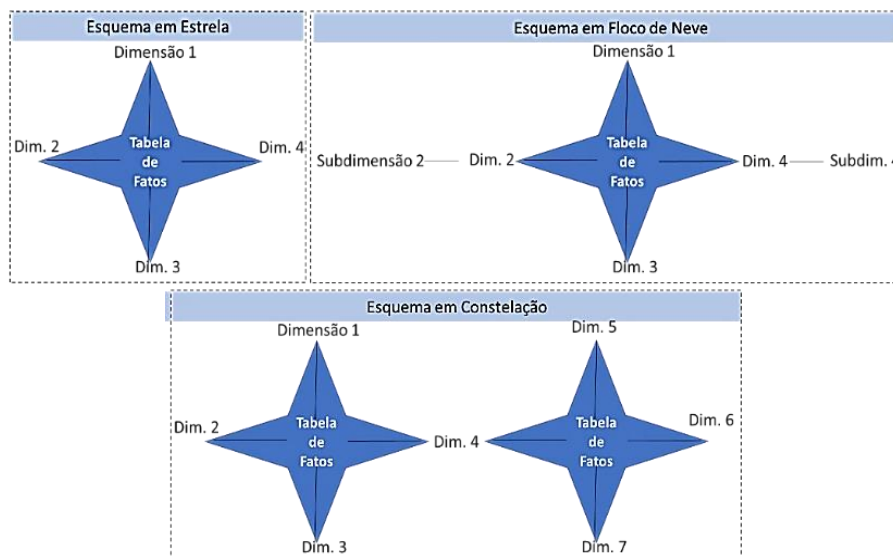


Figura 7 - Esquemas da modelação multidimensional⁹

Assim, podemos constatar que os três modelos derivam do mais usual, que é o que possui a estrutura mais fácil de entender – esquema em estrela. Este esquema pode ser visualizado como um “cubo”, em que cada uma das tabelas de dimensões do esquema em estrela corresponde a uma face do cubo (Santos & Ramos, 2009, p. 83).

Os dados da DDS estão em bases de dados multidimensionais (*Multidimensional Databases* – MDB), isto é, nos cubos multidimensionais (Korobko & Penkova, 2012; Rainardi, 2008). Esses cubos representam o DW e permitem um acesso rápido e dinâmico (Fernandes & Kantorski, 2001, p. 182).

O acesso dinâmico aos dados pode ser feito através de ferramentas OLAP (Santos & Ramos, 2009, p. 95), pelas seguintes operações:

- *Drill-up and drill-down* – analisar o assunto em diferentes níveis, de forma hierárquica (Ballard *et al.*, 2006, p. 90);
 - *Drill-down/Roll-down* – visualizar os dados de um modo geral para um modo mais detalhado (Santos & Ramos, 2009, p. 96);
 - *Drill-up/Roll-up* – o contrário de *drill-down* (Ballard *et al.*, 2006, p. 92), ou seja, a visualização dos dados começa pelo nível menos elevado para o mais elevado da agregação¹⁰ (Santos & Ramos, 2009, p. 96).
- *Drill-across* – quando se passa um dos atributos, dentro da mesma dimensão. Por exemplo, a dimensão “Tempo” é compreendida por ano, semestre, mês e dia. Esta operação consiste no utilizador passar de ano para mês (Aquad, s.d., p. 15);
- *Slice and Dice* - possibilita restringir a informação a visualizar, utilizando o corte e a redução de um conjunto de dados (Valacich & Schneider, 2017, p. 269);

⁹ Fonte: Adaptado de Santos & Ramos, 2009.

¹⁰ Por exemplo: dados > trimestre > ano.

- *Slice* – seleção de um ou mais atributos de uma dimensão, em se consolida com as restantes tabelas de dimensões (Ballard *et al.*, 2006, p. 88);
- *Dice* – permite a inter-relação entre atributos de diferentes dimensões, através da colocação de um ou mais atributos nos eixos, obtendo-se um subcubo (Ballard *et al.*, 2006, p. 89; Santos & Ramos, 2009, p. 96).
- *Pivot (rotate)* – rotação dos eixos (Santos & Ramos, 2009, p. 97), que possibilita a permuta de linhas com colunas e vice-versa. Os dados são os mesmos, mas com uma visualização diferente (Ballard *et al.*, 2006, p. 90).

Estas funcionalidades são importantes para atender às necessidades de informação numa organização. Com a informação organizada num DW, e representada quer seja através de esquemas ou de cubos, é essencial recorrer a ferramentas que possibilitem a criação de relatórios para visualização da informação relevante. Alguns exemplos dessas ferramentas são: *folhas de cálculo; aplicações de relatórios; ferramentas de exploração de dados* (Ballard *et al.*, 2006, p. 94); entre outras.

1.1.4. Visualização

Os mecanismos mais comuns para visualizar a informação são os *dashboards*¹¹ e *scoreboards* (Ballard *et al.*, 2006, p. 98). Um *dashboard* é uma aplicação de BI que consolida, num único écran, a informação mais importante para atingir objetivos (segundo Few em 2003, referido em Oliveira, Cardoso, & Santana, 2015, p. 339). Esta aplicação apresenta *scoreboards*, através de gráficos (Ballard *et al.*, 2006, p. 98).

Os *dashboards* podem ser classificados como: operacionais (monitorização e acompanhamento dos vários processos do negócio); táticos (acompanhamento de processos departamentais); e/ou estratégicos (ou *scorecards*, que acompanha a execução dos objetivos estratégicos) (Eckerson, 2006, p. 7).

Para que o *dashboard* seja uma ferramenta proveitosa, existem alguns princípios que se devem ter em conta aquando da sua elaboração, tais como: necessidades do público-alvo, uso de detalhes ou de informação que não é essencial no painel, exibição numa tela inteira, destaque de informação importante, uso de paleta de cores que sejam de fácil interpretação, entre outras que estão ilustradas na Figura 8.

¹¹ *Dashboards* são conjuntos de relatórios, que podem ser, ou não, dinâmicos.

DASHBOARD: Painel que conta uma história e mantém o público-alvo na tela

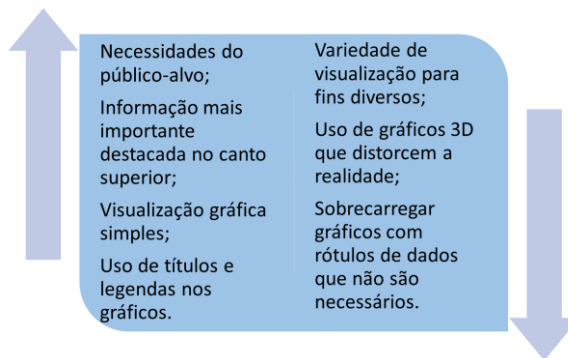


Figura 8 - Algumas práticas a ter em conta na construção de um *dashboard*¹²

Para além de *dashboards*, as aplicações de BI podem ser agrupadas em outras cinco categorias: *reporting applications*, *analytic applications*, *data mining applications*, *alerts* e *portal*. Pode-se também realizar consultas *ad-hoc*. Estas consultas baseiam-se num acesso com parâmetros, definidos consoante as necessidades da organização (Inmon, 2002).

Assim, é possível verificar a existência dos dois tipos de arquitetura do BI (vide Figura 9): *back-end*, caracterizada pelo agrupamento e organização de dados; e *front-end*, onde os dados são analisados e expostos para o utilizador (Vo *et al.*, 2017, p. 1).

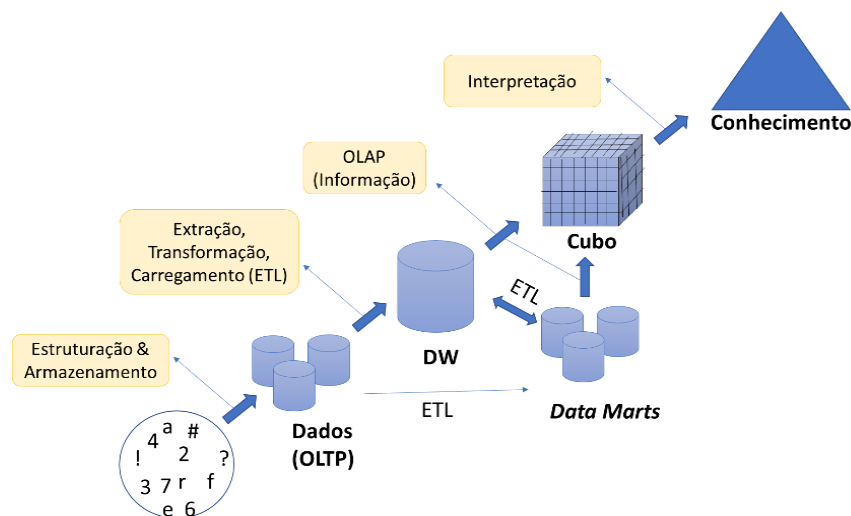


Figura 9 - Arquitetura comum de um sistema BI¹³

Apesar de se poder obter informação relevante numa visualização, pode existir necessidade de refinar. Os *key performance indicators* (KPI)¹⁴ permitem avaliar o estado

¹² Fonte: <https://docs.microsoft.com/pt-br/power-bi/service-dashboards-design-tips>, (obtido em junho de 2018).

¹³ Fonte: Adaptado de Pascoal *et al.*, 2009.

¹⁴ Exemplos de KPI: *lead conversion rate* (in sales); *inventory turnover* (in inventory management).

atual de concretização dos objetivos, mostrando o grau de alcance das metas estabelecidas (Ballard *et al.*, 2006, p. 31-35; Sandu, 2010, p. 33).

Segundo Vo *et al.* (2017), os sistemas tradicionais BI baseiam-se na arquitetura das três camadas (*three-tier*): camada de dados (*Database Layer*), camada da aplicação (*Application Layer*), e camada da apresentação (*Presentation Layer*). Esta arquitetura inclui os processos referidos anteriormente, desde a construção do DW (Sistemas OLTP>ETL>Integração dos dados), à tecnologia OLAP, que possibilita os diversos tipos de análise de dados (esquemas e cubo), e ao produto final que consiste na visualização dos dados tratados, ou seja, da informação que influenciará na tomada de decisão.

1.1.5. Tendências e perspectivas futuras do *Business Intelligence*

A tendência atual é evoluir tecnologicamente a arquitetura dos sistemas tradicionais. O volume dos dados é cada vez maior, influenciado pela integração crescente de *Internet of Things* (IoT) (Gomes & Quintela, 2016, p. 300) e do *Big Data*. De entre as tendências tecnológicas que, segundo Vo *et al.* (2017), podem provocar mudanças assinaláveis na arquitetura de BI, destaca-se: *Big Data*, *Cloud Computing*, *Machine Learning*, *Business Analytics* e *Data Mining*. E alguns conceitos também associados a BI: Inteligência Competitiva (*Competitive Intelligence*) e Gestão do Conhecimento (*Knowledge Management*). A avaliação do impacto das tecnologias no BI, está fora do âmbito da presente dissertação.

Um dos desafios atuais é incluir outras fontes de dados, tais como: mensagens enviadas através de *intranets* da empresa; perfis pessoais dos funcionários e clientes da Web; dispositivos móveis; e outros dados de sensor. Apesar do crescimento de dados ser proveitoso, o fato de algumas dessas fontes de dados não estarem estruturadas, dificulta o processo (Vo *et al.*, 2017, p. 1). Uma forma de armazenar, tratar, analisar e até processar cálculos relativos a grandes volumes de dados, de forma mais rápida, é através de *In-Memory Database* (IMDB) que visa um alto desempenho em BI (Ivan, 2014, p. 19).

Atualmente, os “motores” OLTP possuem “centenas de núcleos de processamento e *Terabytes* de memória”, implementados em servidores *multicore* e, no entanto, continuam a revelar limitações. Por exemplo, quando são definidas demasiadas restrições, as operações entram em conflito pois os motores OLTP não podem executar essas transações em paralelo nos diferentes núcleos (Appuswamy, Anadiotis, Porobic, Iman, & Ailamaki, 2017, p. 121).

O BI evidencia assim, a “simbiose entre gestão e tecnologia” (Sezões *et al.*, 2006, pt. Introdução), baseando-se num conjunto de ferramentas que fornecem uma visão global da organização (Denilson da Silva *et al.*, 2016, p. 75). As suas principais funções são *reunir, analisar, compreender, gerir, consultar, relatar* e *analisar* dados categorizados com *qualidade, confiança* e *importância* para auxiliar a **tomada de decisão** em níveis estratégico, tático e operacional (Rainardi, 2008, p. 12; Botelho & Filho, 2014, p.57).

1.2. Levantamento de Ferramentas

A realização desta secção teve como fontes de informação, essencialmente, sites das ferramentas, de empresas, entre outros.

Perante o processo de BI, surge a necessidade de avaliar as ferramentas disponíveis. Para definir as ferramentas de BI a serem analisadas, estabeleceram-se as seguintes funcionalidades que deverão dispor: modelação de dados para de Sistemas de Gestão de Base de Dados (SGBDs); processo ETL; construção do DW; criação de cubos OLAP e visualização de dados, em forma de relatórios e *dashboards*.

O Quadrante mágico da Gartner¹⁵ possibilita a visualização do posicionamento relativo dos diferentes atores que possuem ferramentas. É composto por dois eixos e quatro quadrantes. O eixo horizontal, que abrange a visão da empresa em relação à tecnologia; enquanto que o eixo vertical demonstra a capacidade de executar o que se estabelece. Os quatro quadrantes: *leaders* (visão atual, bom posicionamento para o futuro e têm uma boa execução); *visionaries* (tendências do mercado, e grande inovação, mas não conseguem concluir tudo o que foi determinado); *niche players* (concentração num segmento pequeno, e possuem falta de inovação); e *challengers* (empresas que podem dominar um grande segmento, mas não possibilitam conclusões na direção do mercado).

De acordo com o relatório anual de Gartner (analisado pela empresa inQuesti), em fevereiro de 2017, constatava-se uma movimentação no quadrante dos líderes. No entanto, os fornecedores do ano de 2016 mantêm-se: Qlik, Tableau e Microsoft (PowerBI), como se pode observar na Figura 10.



Figura 10 - Quadrante mágico de Gartner (comparação entre 2016 e 2017)

¹⁵ Fonte: https://www.gartner.com/technology/research/methodologies/research_mq.jsp, (obtido em janeiro de 2018).

As ferramentas BI, nomeadamente Oracle, SAP, SAS, IBM e as referidas anteriormente como líderes, apesar de fornecerem grandes potencialidades, têm custos elevados. Por isso, sucedem-se as ferramentas *open source*¹⁶, por vezes designadas de “*suites open source*”, quando integram um conjunto de funcionalidade BI.

As ferramentas *open source* podem não ser completamente gratuitas. Ou seja, existem ferramentas que possuem versões *Community*, e/ou *Enterprise* e/ou *Mobile*. A primeira versão contempla apenas uma parte das funcionalidades de uma determinada ferramenta, e é totalmente gratuita. Enquanto que a segunda versão, apesar de conter todas as funcionalidades, envolve custos de utilização. A versão *Mobile* permite a utilização do *software* em dispositivos móveis (Tereso & Bernardino, 2011).

A quantidade de ferramentas BI *open source* disponíveis é elevada e, por isso, as ferramentas que são especializadas em apenas uma ou duas das funcionalidades referidas acima, nomeadamente apenas em termos de visualização, não foram analisadas de forma aprofundada, tais como: Eclipse BIRT, ClicData, Helical insight, SQL Power Wabbit, SealReport, entre outras.

Procede-se à comparação de ferramentas, apenas com versão *Community*, que estão listadas na Figura 11.

		FERRAMENTAS								
		Jedox	JasperSoft BI Community	Open! Community	Pentaho BI Suite CE	SpagoBI	Vanilla	SQL Power Wabbit		
CARACTERÍSTICAS	Funcionalidades	ETL	✓	✓	NÃO	✓	✓	+/-	NÃO	
		OLAP	✓	✓	✓	✓	✓	✓	✓	
		Gráficos	✓	+/-	✓	✓	✓	✓	✓	
		Dashboards	✓	+/-	✓	✓	✓	✓	✓	
		KPI	✓	NÃO	NÃO	✓	✓	✓	NÃO	
		Consultas <i>ad-hoc</i>	✓	✓	NÃO	✓	✓	✓	✓	
	Sistema Operativo		Linux	✓	✓	✓	✓	✓	✓	NÃO
			Windows	✓	✓	✓	✓	✓	✓	✓
			Unix	NÃO	NÃO	NÃO	✓	✓	✓	✓

Figura 11 - Comparação de ferramentas *open source*

As *suite open source* que revelam mais potencialidades, face aos objetivos da presente dissertação, são: Jedox, Pentaho e SpagoBI. No entanto, é importante referir que o levantamento de ferramentas foi elaborado para uma noção atual das ferramentas no mercado, e não pelo seu emprego na presente dissertação.

¹⁶ Ferramentas sem fins lucrativos que seguem regras estabelecidas pela *Open Source Initiative* (OSI), cujo acesso ao software é gratuito.

CAPÍTULO 2

CARACTERIZAÇÃO DO DOMÍNIO DO PROBLEMA

2.1. Medidas de Interesse

2.2. Fiscalização da Pesca

2. Capítulo 2 – Caracterização do Domínio do Problema

2.1. Medidas de Interesse

Atendendo às necessidades de informação, e de se adquirir um panorama clarificado das diversas áreas que a Marinha tem influência, foram definidas medidas de interesse.

As medidas de interesse escolhidas para análise e tratamento de dados são:

- Fiscalização da Pesca;
- Consumo de Combustível;
- Consumo de Munições;
- Limitações Operacionais;
- Operações de Busca e Salvamento;
- Navegação.

A seleção de informação relevante é feita nas medidas de interesse escolhidas, mas apenas uma delas servirá como prova de conceito: a **Fiscalização da Pesca**. Perante isto, sucede-se um enquadramento dos tópicos circundantes da referida medida de interesse para, aquando do tratamento e análise dos dados, estes sejam passíveis de uma interpretação.

2.2. Fiscalização da Pesca

“Mar-Portugal é um desígnio nacional cujo potencial será concretizado pela valorização económica, social e ambiental do oceano e das zonas costeiras, para benefício de todos os Portugueses” (Governo de Portugal, s.d.).

Portugal, é um país de dimensão consideravelmente pequena, se for examinada apenas a dimensão terrestre. Por seu turno, quando se considera a dimensão marítima, Portugal adota a posição de ser “um dos grandes países marítimos do mundo”, e possui uma quantidade incalculável de recursos naturais (Governo de Portugal, s.d., p. 25).

O conceito de fiscalização baseia-se em verificar a forma como uma atividade é realizada, designadamente se está de acordo com a regulamentação prevista, em termos administrativos. Ao aplicarmos este conceito, em específico, às atividades de pesca obtém-se a fiscalização da pesca (Mourinha, 2012, p. 98).

A fiscalização da pesca consiste em averiguar se os pescadores cumprem a legislação estabelecida para este meio da pesca, que assenta em três vetores: Legislação, Fiscalização e Decisão Processual (Figura 12) (Mourinha, 2012, p. 99). Entende-se por Decisão Processual, todo o processo que após a fiscalização e a deteção de uma

presumível infração, conduz ao estabelecimento de medidas coercivas ou à ilibação dos presumíveis infratores (neste caso inocentes) (Mourinha, 2012, p. 118).

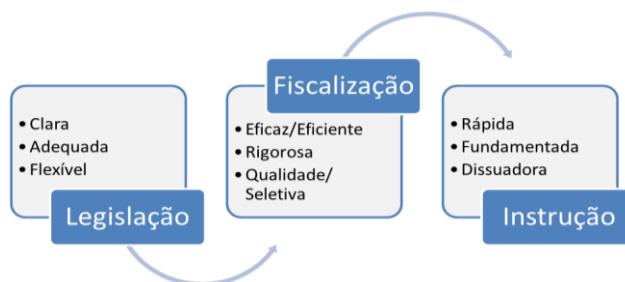


Figura 12 - Objetivos das diferentes fases do processo de fiscalização da pesca¹⁷

2.2.1. Organismos com competência legal

As espécies marinhas¹⁸ estão na base das tarefas de fiscalização, porque as atividades de captura, desembarque, cultura e comercialização dizem respeito a essas espécies.

Segundo o definido no n.º 1, do artigo 2.º, do DL n.º 49-A/2012, de 29 de fevereiro, na sua atual redação, a autoridade que tem como atribuição garantir a fiscalização, âmbito políticas de pesca, está conferida à DGRM.

Os organismos com competência legal, a nível nacional, para realizar tarefas de ações de vigilância, fiscalização e controlo das atividades da pesca, para uma melhor conservação e gestão dos recursos, estão enunciados no artigo 2.º do DL n.º 79/2001, de 5 de março, designadamente, Inspeção-Geral das Pescas, Marinha, Força Aérea (FA), Guarda Nacional Republicana (GNR), Região Autónoma dos Açores (RAA) e Região Autónoma da Madeira (RAM). É importante salientar que a Inspeção-Geral das Pescas é um órgão extinto desde 2002, passando a assegurar as suas responsabilidades a Direção-Geral de Pescas e Aquicultura (DGPA) (DL n.º 246/2002, de 8 de novembro), entretanto integrada na DGRM (DL n.º 49-A/2012, de 29 de fevereiro).

Para clarificar as competências dos respetivos órgãos, o Decreto Regulamentar n.º 86/2007, de 12 de dezembro, visa esclarecer a articulação entre autoridades de polícia e demais entidades competentes, de forma a garantir uma maior eficácia na atuação nos espaços marítimos sob soberania e jurisdição nacional. Assim, as competências da Marinha/AMN são “coordenar, no âmbito operacional, as ações de vigilância e fiscalização das atividades de pesca e culturas marinhas exercidas em espaços sob soberania e jurisdição nacional” (DR n.º 86/2007, de 12 de dezembro, art.º 3º).

¹⁷ Fonte: Adaptado de DGAM, 2018.

¹⁸ As espécies marinhas são todos os animais ou plantas que passam uma parte significativa do seu ciclo de vida na água salgada ou na água salobra (DL n.º 383/98, de 27 de novembro, art.º 2º).

A MP é um ramo das Forças Armadas, que obedece a todas as medidas¹⁹ declaradas no DL n.º 185/2014, de 29 de dezembro, de onde conta a aprovação da Lei Orgânica da Marinha (LOMAR). A este ramo competem diversas competências, que requerem uma atuação tanto a nível nacional como internacional, bem como a nível militar e não militar. Das quais, condutas não militares são, por exemplo: segurança marítima, exercício da autoridade pública no mar, investigação no âmbito das ciências do mar e preservação da cultura marítima (Monteiro & Mourinha, 2011).

Uma dessas competências é garantir o cumprimento das missões reguladas por legislação própria, nomeadamente, “exercer a autoridade do Estado nas zonas marítimas sob soberania ou jurisdição nacional e no alto mar” (DL n.º 185/2014, de 29 de dezembro, art.º 2º). Como extensão ao cumprimento das missões reguladas por legislação própria, compete ao CN apoiar o exercício do comando por parte do Chefe do Estado-Maior da Armada (CEMA), de forma a “garantir, no seu âmbito, a fiscalização nos espaços marítimos sob soberania ou jurisdição nacional, tendo em vista o exercício da autoridade do Estado através da adoção das medidas e ações necessárias” (DR n.º 10/2015, de 31 de julho, capítulo IV, art.º 61º, n.º 1, alínea a).

Assim, a Marinha realiza a sua competência, no âmbito da fiscalização da pesca, através das estruturas do CN, da Direção-Geral da Autoridade Marítima (DGAM) e do Comando-Geral da Polícia Marítima (CGPM), pelo que a fiscalização nos espaços oceânicos e costeiros é sobretudo efetuada pelas unidades navais do CN (Mourinha, 2012, p. 108).

2.2.2. Áreas de Atuação

Os espaços marítimos nacionais ocupam uma área ampla²⁰, na qual o Estado tem alguns deveres a cumprir, tais como: garantir a segurança, exercer a sua autoridade e assegurar uma presença equilibrada. A Marinha tem um papel notório neste processo, tendo em conta a disponibilidade de meios navais oceânicos, o saber, a experiência e a amplitude de responsabilidade inerente a este órgão. Por isso, para que o Estado consiga um empenhamento coerente de meios nos espaços marítimos sob soberania ou jurisdição nacional torna-se necessário definir as competências e os limites dos espaços marítimos (Monteiro & Mourinha, 2011).

Nos termos da Convenção das Nações Unidas sobre o Direito do Mar (CNUDM), assinada a 10 de dezembro de 1982, em *Montego Bay*, e ratificada por Portugal em 1997 (Resolução da Assembleia da República n.º 60-B/97, de 03 de abril), a intervenção do Estado português pode ser exercida através dos navios da Marinha²¹.

¹⁹ Missão, estrutura, organização, competências, entre outras.

²⁰ Superior a 1,7 milhões de km².

²¹ A legitimidade para atuar reside, em grande parte, no Direito Internacional que é consagrado através do art.º 8º da Constituição da República Portuguesa (CRP).

As zonas marítimas sob soberania ou jurisdição nacional, em conformidade com a CNUDM, são: Águas Interiores, Mar Territorial, Zona Contígua, Zona Económica Exclusiva (ZEE) (Lei n.º 34/2006, de 28 de julho, arts.º 2º e 11º) (Figura 13 e Figura 14).

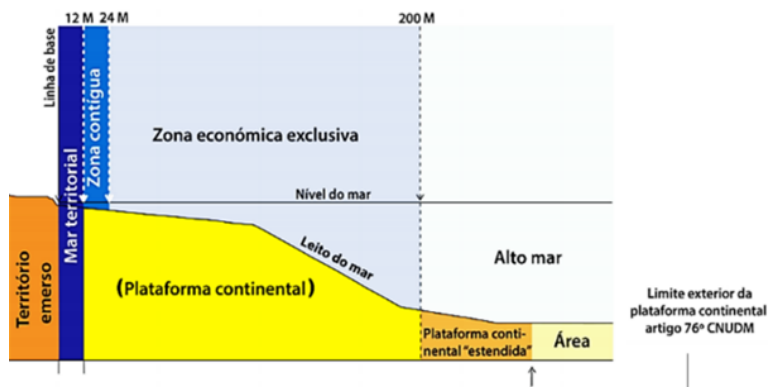


Figura 13 - Extensão das zonas marítimas sob soberania ou jurisdição nacional²²

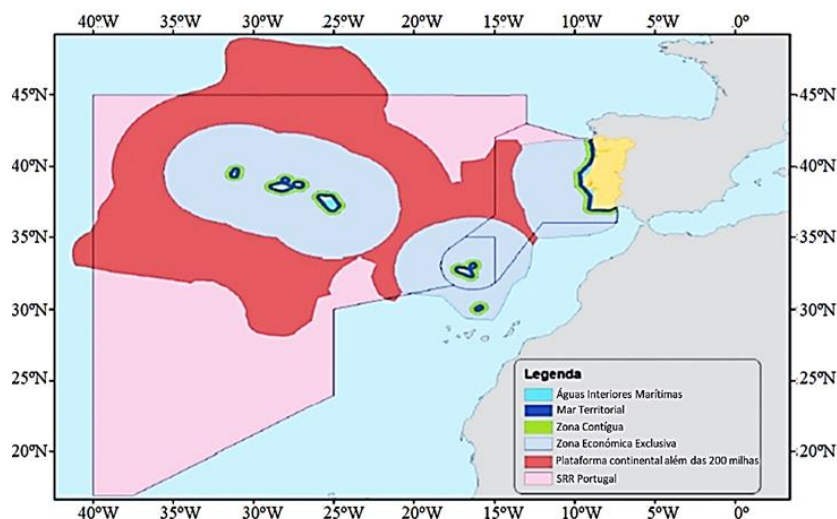


Figura 14 - Espaços marítimos sob soberania, jurisdição e responsabilidade nacional²³

Portugal, através das Organizações Regionais de Pesca para águas internacionais, usufrui de outras possibilidades de pesca, nomeadamente nas seguintes áreas: Organização da Pesca do Atlântico Noroeste (NAFO), Comissão da Pesca do Atlântico Nordeste (NEAFC), Comissão Internacional para a Conservação do Atum Atlântico (ICCAT) e Comissão dos Atuns do Oceano Índico (CTOI). Para além disso, tem estabelecido Protocolos de Pesca²⁴ anexos aos Acordos de Parceria entre a União Europeia e Países Terceiros, para águas das respetivas ZZE (INE, 2017, p. 115).

²² Fonte: Governo de Portugal, s.d., p.30.

²³ Fonte: <https://www.dgrm.mm.gov.pt/web/guest/am-ec-zonas-maritimas-sob-jurisdicao-ou-soberania-nacional>, (consultado em fevereiro de 2018).

²⁴ Em 2016, Protocolos com Guiné-Bissau, Marrocos, Cabo Verde e Madagáscar.

Uma porção dos recursos naturais marinhos são os recursos vivos, nos quais se denota uma significativa insuficiência na captura de pescado, pelo que se deve desenvolver práticas de aquicultura (Governo de Portugal, s.d., p. 46).

As pescas, incluindo a aquicultura, constituem uma “fonte vital de alimentos, emprego e bem-estar económico” a nível mundial. Para garantir a conservação dos recursos aquáticos vivos, a *Food and Agriculture Organization of the United Nations* (FAO) estabelece comportamentos responsáveis relativos à exploração dos recursos da pesca, através do Código de Conduta para uma Pesca Responsável (FAO, 1995).

Posto isto, a FAO estabeleceu zonas de pesca (Figura 15), onde se deve incidir a investigação e a verificação de capturas (Pinto, 2017, p. 65). Essas zonas estão distribuídas em áreas geográficas, a nível mundial. As áreas onde Portugal realiza o exercício de jurisdição nacional, o controlo das atividades de pesca e onde, por sua vez, tem maior incidência de capturas (INE, 2017) são as seguintes áreas: 21 (NAFO), 27 (NEAFC) e 34 (CECAF – *Central East Commission Atlantic Fishing*).

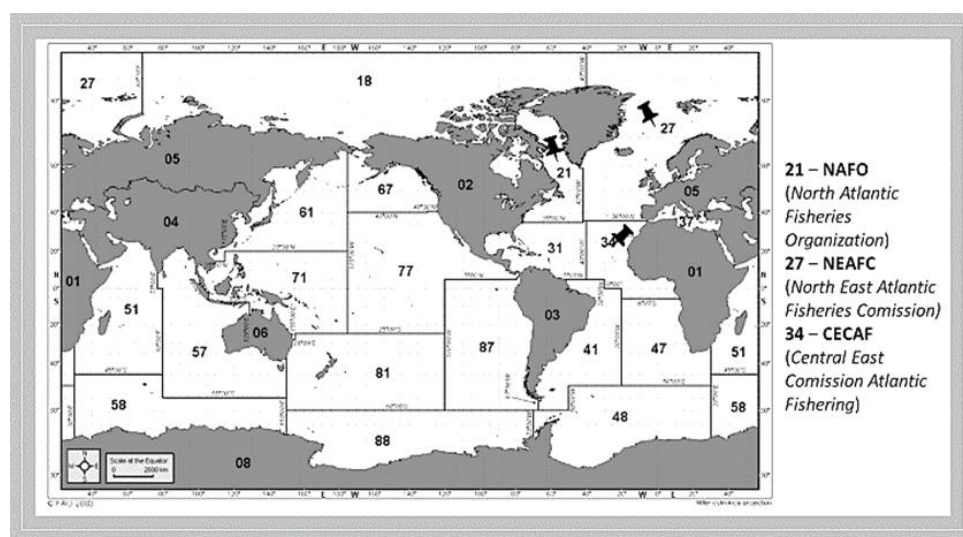


Figura 15 - Áreas de pesca (Divisão FAO)²⁵

De modo a tornar a pesca sustentável do ponto de vista ambiental, económico e social, assegurando a conservação e gestão dos recursos da pesca, definiu-se uma política comum em matéria de pescas: Política Comum das Pescas (PCP) (União Europeia, 2016).

A PCP é uma política da União Europeia (UE), instituída em 1983, através do Regulamento (CEE) n.º 170/83, de 25 de janeiro de 1983 (revogado). O processo de reforma da PCP concluiu-se com a publicação do Regulamento (UE) n.º 1380/2013, de 11 de dezembro (revoga uma revisão da Política Comum realizada em 2002). A nova PCP foi implementada, através do Regulamento (UE) 2015/812, de 20 de maio (DGRM, 2018).

²⁵ Fonte: Adaptado de INE, 2017.

Portugal, como país-membro da UE, está ligado à PCP de acordo com o DL n.º 10/2017, de 10 de janeiro, que institui um regime comunitário de controlo com a finalidade de assegurar o cumprimento das regras da PCP.

De modo a apoiar a nova PCP, a incentivar o exercício e o controlo de práticas de pesca sustentável, a UE concebeu o Fundo Europeu dos Assuntos Marítimos e das Pescas (FEAMP), nomeadamente para os mares e para as costas da Europa.

O FEAMP tem como fundamentos: tornar a pesca e a aquacultura mais sustentáveis e lucrativas; melhorar a diversidade de economias locais para um desenvolvimento sustentável nas comunidades costeiras e interiores; diminuir os impactos da atividade humana no meio marinho; entre outros (União Europeia, 2016, p. 50). Uma parte do orçamento do FEAMP (19,1%) visa a aplicação da PCP, de maneira a melhorar os conhecimentos científicos, a recolha de dados, o controlo e a fiscalização da legislação da pesca (União Europeia, 2016, p. 51).

Para além disso, segundo a União Europeia (2016), a maior parte das águas do mundo estão compostas por Organizações Regionais de Gestão das Pescas (ORGP), para um controlo e gestão sustentável dos recursos marinhos comuns. As ORGP são entidades internacionais, integradas por países com interesses na área da pesca, numa mesma região e/ou espécie (ou grupo de espécies). Assim, para atender às necessidades de conservação, gestão e sustentabilidade de recursos, estabeleceram-se Acordos de Parceria no Domínio da Pesca Sustentável²⁶ (APPS), celebrados entre a UE e países terceiros (União Europeia, 2016, p. 8).

Posto isto, um dos deveres de cada nação é cooperar e incentivar boas práticas para a conservação e gestão de uma pesca sustentável, cumprindo as regras estabelecidas e mantendo o controlo e a fiscalização sobre o exercício da pesca (Pinto, 2017, p. 67).

2.2.3. Recursos Haliêuticos

A ação da Humanidade tem provocado riscos no Oceano, o que leva a que muitas espécies se encontrem ameaçadas e, muitas das vezes, devido à pesca ilegal (Governo de Portugal, s.d., p. 13). Para além disso, no continente, por se situar numa zona de transição para ecossistemas mais quentes, verifica-se a existência de uma elevada diversidade de espécies de pescado²⁷, mas em baixa abundância (Ministério da Agricultura e do Mar, 2007, p. 9).

²⁶ Os APPS definem as normas a seguir relativamente à pesca internacional, sendo que a gestão sustentável de frotas de pesca externas, especificadamente o âmbito de um APPS, está descrita no Regulamento (UE) n.º 2017/2403, de 12 de dezembro de 2017, art.º 8º.

²⁷ Outros fatores que contribuem para a elevada diversidade de espécies de pescado: existência de fenómenos sazonais e afloramentos costeiros (costa ocidental), e a influência das águas do Mediterrâneo (costa algarvia).

As principais espécies exploradas na costa portuguesa são os pequenos pelágicos²⁸ (como a cavala, o carapau e a sardinha) e os demersais²⁹ (como o polvo, a pescada e o choco). Os dados que possibilitam uma análise das capturas nominais³⁰ de pescado (em toneladas) por Porto de descarga e Espécie, no período de interesse, estão disponíveis no *site* do Instituto Nacional de Estatística (INE).

No que concerne ao registo de embarcações³¹ da frota de pesca nacional, a 31 de dezembro de 2016, estavam registadas 7 980 embarcações (INE, 2017, p.29). Estas embarcações exercem a sua atividade, predominantemente, através das seguintes modalidades de pesca: a pesca do cerco, a pesca polivalente e o arrasto (INE, 2017, p. 50).

A pesca por arte de cerco consiste em “qualquer método de pesca que utiliza parede de rede sempre longa e alta, que é largada de modo a cercar completamente as presas e a reduzir a capacidade de fuga” (Portaria n.º 1102-G/2000, de 22 de novembro, art.º 2º). Este tipo de pesca é direcionado para a captura dos pequenos pelágicos (Portaria n.º 1102-G/2000, de 22 de novembro, art.º 7º).

A pesca polivalente caracteriza-se por utilizar diversas artes³², tais como aparelhos de anzol, armadilhas, alcatruzes, ganchorra, redes camaroeiras, xávegas, entre outras (INE, 2017, p. 125).

A pesca por arte de arrasto compreende “qualquer método de pesca que utiliza estruturas rebocadas essencialmente compostas por bolsa, em geral grande, e podendo ser prolongada para os lados por “asas” relativamente pequenas (Portaria n.º 1102-E/2000, de 22 de novembro, art.º 2º).

Atendendo às diferentes modalidades de pesca e as capturas delas resultantes, verifica-se, através do Gráfico 1, que a pesca do cerco, apesar da redução de 2015 para 2016³³, mantém-se como a arte predominante, sucedendo-se a pesca polivalente e, por último o arrasto, no qual se verificou um aumento do volume de capturas.

²⁸ A espécie pelágica é uma espécie que vive na coluna de água (ou à superfície), mas sem relação ao fundo.

²⁹ A espécie demersal é uma espécie que se caracteriza por viver no fundo, ou perto deste, mas não é permanentemente dependente do fundo.

³⁰ A captura nominal é o peso vivo correspondente aproximadamente à pesca descarregada, ou seja, representa o peso líquido no momento da descarga do peixe.

³¹ O registo de embarcações pode ser consultado através da base de dados da UE.

³² A arte de pesca define-se em função das espécies a capturar e da época do ano.

³³ As capturas não ultrapassaram as 53 270 toneladas, que revela uma diminuição de 19.8 % devido, essencialmente, às menores quantidades de peixes marinhos (cavala e carapau, nomeadamente).

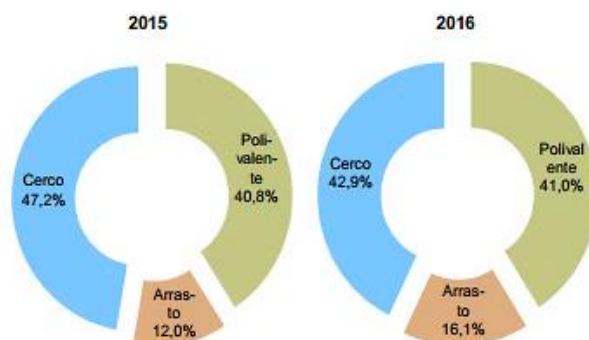


Gráfico 1 - Capturas nominais de pescado fresco ou refrigerado (ton), por arte de pesca nos períodos de 2015 e 2016³⁴

O exercício da pesca marítima e da cultura de espécies marinhas, nos quais não existe uma distribuição uniforme, requerem medidas de equilíbrio, ou seja, assegurar a gestão e o aproveitamento dos recursos da fauna e flora que residem nas águas sob soberania e jurisdição portuguesas, de forma sustentável. Posto isto, a Marinha Portuguesa elaborou Roteiros de Fiscalização de Pesca, com a finalidade de obter uma melhor visão da pesca, a nível nacional. Estes roteiros estão divididos pelas seguintes zonas marítimas: Zona Marítima do Norte (ZMN), Zona Marítima do Centro (ZMC), Zona Marítima do Sul (ZMS), Zona Marítima dos Açores (ZMA) e Zona Marítima da Madeira (ZMM) (Pinto, 2017, p. 64).

2.2.4. Controlo e Fiscalização das Atividades de Pesca

Uma ação de fiscalização a uma embarcação de pesca compreende diversos tipos de matéria a serem verificados, tais como: aspetos específicos de pesca³⁵, confirmação dos requisitos de segurança³⁶, meios de segurança e de navegabilidade, condições da tripulação da embarcação³⁷ (DGAM, 2011), meios de salvação (DL n.º 191/98, de 10 de julho, art.º 16º), documentos obrigatórios para embarque dos marítimos (DL n.º 280/2001, de 23 de outubro, n.º 4 do anexo V) e licenças relativas ao equipamento radioelétrico das embarcações (DL n.º 190/98, de 10 de julho, art.º 50º). A duração de uma ação de fiscalização não deve ser superior a 4 horas, exceto se for detetada uma infração aparente ou no caso dos agentes precisarem de informação adicional (Regulamento de Execução (UE) n.º 404/2011, da Comissão, de 08 de abril, n.º 4 do artigo 104º). Estas vistorias visam otimizar a fiscalização marítima, porque se analisam vários componentes para além da pesca, apenas numa ação (Mourinha, 2012, p. 98).

No decurso de uma inspeção, a verificação do cumprimento das normas e dos procedimentos técnicos, realiza-se através do preenchimento do relatório de inspeção. O conteúdo mínimo necessário para preencher um relatório de inspeção está descrito no

³⁴ Fonte: INE, 2017, p.50.

³⁵ Licença, tipo de artes, zonas de operação.

³⁶ Coletes, extintores, pirotécnicos.

³⁷ Lotação de segurança, habilitações dos marítimos embarcados.

Anexo XXVII do Regulamento de Execução (UE) n.º 404/2011 da Comissão, de 8 de abril.

Assim, se algo não estiver em conformidade com os parâmetros estabelecidos, poderão ser consideradas presumíveis infrações, pelo que toda a informação que for pertinente deve ser registada no relatório de inspeção (Regulamento de Execução (UE) n.º 404/2011 da Comissão, de 8 de abril, art.º 115º).

Aquando da deteção das referidas presumíveis infrações, pressupõe-se a elaboração de documentação e o levantamento de meios de prova, para posteriormente se tomar uma apropriada decisão processual (Mourinha, 2012, p. 98).

Sempre que sejam detetados ilícitos contraordenacionais por unidades navais da Marinha, em áreas sob jurisdição marítima nacional, compete ao comandante do navio levantar o auto de notícia e salvaguardar os meios de prova, tomando as medidas necessárias adequadas (DL n.º 45/2002, de 2 de março, art.º 10º).

Assim sendo, segundo dados obtidos pela Marinha, recolhidos das ações de fiscalização, verifica-se que a MP apurou diversas presumíveis infrações, entre os períodos de 2015 a 2018 (apenas o mês de janeiro) (Tabela 2).

Tabela 2 - Quantitativos de infrações averiguadas pela Marinha Portuguesa, nos períodos entre 2015 e janeiro de 2018

Infração		Quantitativos de Presumíveis Infrações averiguadas pela Marinha Portuguesa						
		2015		2016		2017		JAN 18
Maior	Menor	Diversos *	Potência motora ou arqueação excessiva	Diversos *	Diário de Pesca inexistente	Diversos *	Diário de Pesca inexistente	Diversos *
Quantitativo		1280	3	962	6	725	4	26

* Falta de documentos a bordo; falta de pirotécnicos; extintores caducados; entre outros

Das referidas presumíveis infrações, observa-se uma redução das infrações referente aos diversos, que por englobar algumas possibilidades, torna-se difícil de aferir qual a infração mais alarmante que integra os diversos. Por outro lado, a infração que se verificou com menos quantitativo nos anos de 2016 e de 2017, foi a inexistência do diário de pesca, apesar de se observar uma melhoria.

2.2.5. Ferramentas de apoio no âmbito da fiscalização da pesca

No que concerne à matéria de pescas, tendo em conta toda a envolvente explicitada, é essencial que existam sistemas que possibilitem a monitorização de embarcações de pesca. Isto leva a que a instalação de um equipamento que permita a monitorização contínua nas embarcações nacionais que atuam em águas portuguesas, seja um requisito essencial para o exercício da pesca e respetivo licenciamento (Decreto Regulamentar n.º 3/93, de 8 de fevereiro).

Como tal, é necessário o uso de ferramentas de apoio à decisão de modo a alcançar atuações nas missões de fiscalização pertencentes à Marinha de forma eficaz e eficiente, pelo controlo sobre todos os meios humanos e materiais disponíveis.

Para programar, coordenar e executar a fiscalização, a vigilância e o controlo das atividades da pesca, a DGRM usufrui do Sistema de Fiscalização e Controlo das Atividades da Pesca (SIFICAP), do Sistema de Monitorização Contínua da Atividade de Pesca (MONICAP) (DL n.º 49-A/2012, de 29 de fevereiro, art.º 2º, na sua atual redação).

O SIFICAP, primeiro sistema a ser desenvolvido por Portugal (Pinto, 2017, p. 74), suporta as ações de vigilância, fiscalização e controlo das atividades da pesca. É um sistema integrado numa rede de comunicação e tratamento informativo de dados³⁸ que visa contribuir para uma melhor defesa, conservação e gestão dos recursos piscatórios (DL n.º 79/2001, de 5 de março, arts.º 2º e 3º). Os objetivos do SIFICAP estão descritos no artigo 4º do DL n.º 79/2001, de 5 de março, e nos quais se destacam: garantir, em tempo útil, a conjugação tanto das entidades participantes como dos meios operacionais; possibilitar o acesso a informação relevante, consoante a atividade de cada uma das entidades; fornecer elementos estatísticos e de apoio à decisão.

Para atender aos objetivos, o SIFICAP é constituído por um SI³⁹, uma rede de comunicação de dados (RCD)⁴⁰, pelo MONICAP e pelos meios humanos e materiais das entidades participantes (DL n.º 79/2001, de 5 de março, art.º 5º).

O MONICAP é um sistema de Monitorização Contínua da Atividade da Pesca, baseado em tecnologias de telecomunicações e informação geográfica, que permite acompanhar a atividade das embarcações de pesca, através de representação gráfica sobre carta digitalizada (DL n.º 79/2001, de 5 de março, art.º 3º). Para além disso, este sistema apresenta as seguintes funcionalidades: aquisição, receção e transmissão de dados meteorológicos e posicionamento célere e precisa de navios e embarcações em perigo (Mota, 2010, p. 16). O MONICAP é um dos sistemas que serve como fonte de dados do Sistema de Apoio à Decisão para a Atividade de Patrulha (SADAP).

No âmbito da análise operacional da atividade de fiscalização da pesca por parte do CN, a Direção de Análise e Gestão da Informação (DAGI) desenvolveu um sistema de apoio à decisão, no ano de 2006, designado de SADAP. Este sistema serve como suporte no que diz concerne à orientação do esforço de patrulha, como uma extensão temporal e geográfica. Para além disso, é um sistema com capacidade para assegurar uma

³⁸ Dados provenientes dos seguintes órgãos: Ministérios da Defesa Nacional, das Finanças, da Administração Interna, da Economia, da Agricultura, do Desenvolvimento Rural e das Pescas e do Ambiente e do Ordenamento do Território.

³⁹ SI – Conjunto de dados que apoia a interligação entre as unidades participantes no SIFICAP e os seus sistemas envolventes.

⁴⁰ RCD – Tecnologia de informação que permite a disponibilização dos dados e da informação.

cobertura eficiente do espaço marítimo, e também para aperfeiçoar a brevidade na conceção de relatórios estatísticos da atividade operacional, através da recolha automática de dados, provenientes de diversas fontes e formatos (Marinha, s.d.b, p. 1.1).

O SADAP é composto por oito módulos⁴¹, cada um com um contributo específico, e de forma conjunta permitem o planeamento de missão, a preparação de saída diária, o apoio à vistoria, o apoio na elaboração de mensagens ligadas à fiscalização marítima e o apoio às atividades de busca e salvamento (Marinha, s.d.b, p. 1.2).

Das aptidões que o sistema possui, distribuídas pelos diversos módulos, são de destacar as seguintes: acesso e cruzamento de dados de fiscalização; mensagens formatadas elaboradas pelas UN e pelas capitánias; verificação e análise das trajetórias praticadas pelas embarcações de pesca; panorama situacional marítimo atualizado; elaboração de mensagens formatadas relativas à fiscalização (tais como FISCREP, NAVSIT, entre outras); deteção automática de infrações; averiguação da zona de vistoria, tendo em conta os dados das embarcações e respetiva atividade de pesca; criação, edição e visualização de alertas SAR.

Em suma, o SADAP é uma ferramenta de apoio operacional, que fornece um panorama através da visualização e da capacidade de filtrar apenas a informação de interesse, consoante a necessidade de uma determinada entidade.

⁴¹ Módulos do SADAP: análise de fiscalização; análise da atividade da pesca; capacidade *Automatic Identification System* (AIS); elaboração de mensagens formatadas; regras, artes, malhagens e espécies; coberturas; visualização de alertas e computação *Search and Rescue* (SAR); e busca e salvamento.

CAPÍTULO 3

RECOLHA E TRATAMENTO DE DADOS

- 3.1. Conceito e objetivo das MTF e das Mensagens Estruturadas
- 3.2. Identificação e Tratamento das Fontes de Dados
- 3.3. Recolha de Dados
- 3.4. Análise e tratamento dos dados recolhidos
- 3.5. Base de Dados

3. Capítulo 3 – Recolha e Tratamento de Dados

3.1. Conceito e objetivo das MTF e das Mensagens Estruturadas

A MP é uma organização composta por diversas entidades, que necessitam de compartilhar informação entre si. A base da informação são os dados, que neste caso constituem uma ampla dimensão, tendo em conta o volume de dados que circula entre as unidades, quer em terra, quer no mar. Para que ocorra a difusão e troca de dados, é necessário estabelecer um meio de comunicação, entre os quais se destacam as mensagens.

Entende-se por “Mensagem”, no âmbito da teoria da informação, a sequência ordenada de caracteres destinada a transportar informação (Marinha, 2005, p. II.233).

Os tipos de mensagens analisados na presente dissertação são: Mensagens Estruturadas e Mensagens de Texto Formatado, e por isso, seguem-se as respetivas definições.

Uma Mensagem Estruturada é uma mensagem cujo texto segue uma sequência específica, constituída por parágrafos padrão. Cada parágrafo contém um identificador, no qual pode ser usado texto livre. Este tipo de mensagem requiere interpretação e processamento manual. Por isso, “as mensagens neste formato serão substituídas por MTF num futuro próximo” (NATO, 2015, p. 6-1).

Uma MTF é uma mensagem cujo texto segue um formato específico, e é constituída por diversos campos. Este tipo de mensagem permite o tratamento automático dos dados (NATO, 2015, p. 1-1).

A estrutura das mensagens é orientada de acordo com o identificador de parágrafo MSGID (*Message Identification*), pois este campo corresponde ao tipo de mensagem, sendo que cada mensagem possui uma estrutura a seguir.

Essas mensagens são elaboradas nas mais diversificadas situações, através das seguintes publicações: APP11 e IONAV1010, nível NATO e nível nacional respetivamente.

O APP11 é um compêndio de mensagens (de texto formatado) para o Comando e Controlo de Forças NATO, referentes aos diversos níveis das operações (em terra, no mar, ou no ar). Esta publicação estabelece um procedimento comum de troca de informação entre os aliados, e reduz o risco de ambiguidade nas comunicações (NATO, 2015, p. 1-1).

A IONAV1010 é constituída por comunicados e relatos operacionais (mensagens estruturadas) que são empregues a nível nacional. Esta publicação foi elaborada de forma idêntica ao sistema NATO, para possibilitar uma doutrina semelhante e para que não

existam comunicados com estruturas e terminologias díspares para o mesmo fim (Marinha, 2002, pp. 1.1-1.2).

Neste momento, a IONAV1010 é composta por mensagens estruturadas. No entanto, as diretrizes são transformar essas mensagens em MTFs, num futuro próximo. Isto justifica-se pela necessidade de analisar os dados provenientes das diversas mensagens, o que se consegue alcançar através do processamento de forma automática, que este agiliza essa análise.

A redação manual das MTFs exige tempo disponível, quer para a sua codificação como para a respetiva descodificação, tendo sempre em conta a suscetibilidade do erro humano. Por isso, é adequada a utilização de um *software* focalizado tanto para a codificação (sistema já implementado na MP para a maioria das MTFs) como para a descodificação. O foco da presente dissertação é proporcionar uma visualização da informação útil para a tomada de decisão, com base nos dados contidos nessas mensagens. Perante isto, a MP está a realizar um projeto que se baseia no processamento automático de mensagens, daí a necessidade de transformar as mensagens estruturadas em mensagens formatadas, para simplificar, posteriormente, o acesso e a visualização da informação relevante e de interesse.

3.2. Identificação e Tratamento das Fontes de Dados

Os dados significativos estão incorporados nos campos que constituem as mensagens.

As mensagens contemplam a informação relevante, mas não de forma direta. Ou seja, as mensagens são constituídas por inúmeros dados que necessitam de ser descodificados e tratados para se transformarem em informação. Seguidamente, quando os dados já têm um significado atribuído, é necessária uma seleção da informação.

A seleção de informação foi feita com base nas publicações **APP11 (D – Versão 1)** e **IONAV1010 (Alteração N.º 5)**.

De seguida, explicita-se a seleção das mensagens e dos respetivos campos, e o tratamento dos dados. Este processo consistiu na identificação das mensagens, dessas referidas fontes, e da comparação dos dados que as constituem com a finalidade de obter uma versão final dos dados de interesse. Cada medida de interesse foi tratada em várias folhas do Excel, de forma individual, isto é, cada ficheiro corresponde a uma medida de interesse. No entanto, apenas a medida de interesse que servirá como prova de conceito, será analisada de forma minuciosa.

Fiscalização da Pesca

- Mensagens – a seleção das mensagens teve como critério qualquer campo que contivesse dados acerca da medida em estudo. A publicação APP11 não detém mensagens deste

âmbito. Assim, os comunicados que têm pelo menos um campo relativo à atividade da Fiscalização da Pesca, são os que se encontram na Tabela 3.

Tabela 3 - Comunicados que contêm pelo menos um campo referente à fiscalização

IONAV1010	APP11
INCIDENTREP	NIL
FISCREP	
FISCSUM	
FISHGEARSUM	
LAWREP	
MISREP	
LOGCONTREP	

Cada comunicado é composto por diversos campos, o que requiere a análise de cada um deles, de modo a se alcançar o comunicado mais completo no que diz respeito aos dados referentes à Fiscalização da Pesca. Os campos estão organizados consoante a finalidade de cada comunicado, sendo que apenas foram analisados os campos que estavam relacionados com a medida de interesse (vide Tabela 4).

Tabela 4 - Comunicados relativos à fiscalização e a respetiva finalidade⁴²

Comunicado	Finalidade	Campos observados
INCIDENTREP	Descrever um evento anômalo que requiere uma especial atenção ou controlo, nas águas sob soberania ou jurisdição nacional	
FISCREP	Relatar o resultado da inspeção realizada a uma embarcação de pesca ou de recreio, nacional ou estrangeira, em atividade nas águas sob soberania ou jurisdição nacional	Todos, exceto REF
FISCSUM	Relatar o resumo da atividade de fiscalização no decurso de uma missão, ao Comando Naval e às entidades que concorrem para a fiscalização da atividade de pesca e recreio	Todos, exceto REF
LAWREP	Explicar o resultado das ações administrativa e judicial delineadas pelo Capitão de Porto competente, na sequência de um ilícito marítimo	
FISHGEARSUM	Relatar a localização das artes de pesca caladas detetadas, no âmbito das atividades de fiscalização executadas no decurso de uma missão	
MISREP	Descrever como decorreu uma missão atribuída	VIFISC
LOGCONTREP	Conceder informação e dados estatísticos relativos à atividade operacional, assim como o custo efetivo das missões atribuídas	ACTID

Campos – o modo de seleção dos campos foi de encontro com os que abrangiam qualquer tipo de dados acerca da medida em estudo. O objetivo foi perceber que tipo de dados incluíam as mensagens, e se subsistia correlação. É possível verificar que existem muitos campos em comum, nas diversas mensagens, através da análise da *Identificação dos dados, campos e respetivas mensagens* (Apêndice B).

⁴² Fonte: Marinha, 2002.

Perante isto, o comunicado que abrange mais campos direcionados para a análise e controlo da Fiscalização da Pesca é o FISCREP⁴³ (Tabela 5).

Tabela 5 - Formato do comunicado FISCREP⁴⁴

C.O. (1)	I.P. (2)	TÍTULO	OBSERVAÇÕES (3)
M	MSGID	Identificador de Mensagem	FISCREP/Unidade Naval/Nº de ordem anual do comunicado/GDH da inspeção//
C	REF	Referência	Identificador do documento/Originador/GDH/Mês/Nº sequencial/SIC//
M	IDEMB	Identificador da Embarcação	Tipo de embarcação (Tabela 1 Anexo D)/Subtipo (Tabelas 2-A, 2-B, 2-C e 2-D, do Anexo D)/Matricula ou Registo/Nome/Nacionalidade (Anexo A)//
M	AREA	Posição Inspeção / Área	Posição Geográfica/Área (apenas para embarcações de pesca – Tabela 4 Anexo D)//
M	RES	Resultado	LEGAL OU PRESUMÍVEL INFRATORA
C	ACT	Tipo de Atividade	Referir a atividade em que a embarcação está empenhada (Tabela 5 Anexo D)/Arte de Pesca em uso (para embarcações de pesca – Tabela 3 Anexo D)/Classe de Malhagem (para embarcações de pesca, caso aplicável)//
C	LAW	Presumíveis Infrações	Nº sequencial (mesmo que seja apenas uma infração)/Código da infração (Tabela 7 Anexo D)/Código de comportamento grave (Tabela 6 Anexo D), para embarcações de pesca caso aplicável)/Presumível infração por extenso//
C	FISH	Capturas de Pescado	Registo de todas as capturas existentes a bordo da seguinte forma: Nº sequencial (mesmo que seja apenas uma linha)/Quantidade (em kg) – Espécie (Tabela 8 Anexo D; esta coluna repete-se tantas vezes quantas as necessárias) / Nº da página do diário de pesca do dia (quando aplicável)/Arte de Pesca utilizada (Tabela 3 Anexo D)/Classe de Malhagem utilizada (caso aplicável)//
M	ENVCON	Condições Ambientais	Condições de tempo/Direção e Força do Vento/Estado do Mar/Altura e Direção da Ondulação/Visibilidade//
M	DIV	Diversos	Nome do Mestre ou Patrão da Embarcação/Nº e emissão da cédula ou carta de recreio/Nº contribuinte (se Presumível Infrator)/Morada (se Presumível Infrator)
C	RMKS	Observações	Ação Tomada (embarcação apresada, mandada seguir para o porto de ..., artes ou pescado selados e/ou cautelarmente apreendidos, auto notícia a ser enviado à Capitania do Porto de ..., etc); Informação Complementar como o Nº dos selos utilizados e a sua localização, etc.

(1) C.O. – Critério de Obrigatoriedade (M – Mandatório; C – Condicional; O – Opcional)

(2) I.P. – Identificador de Parágrafo

(3) Observações:

-Anexo A: Códigos Designadores de Bandeiras e Entidades Geográficas

-Anexo D: Tabelas relativas à Fiscalização das Atividades de Pesca

- Tabela 1: Tipos de Navios/Embarcações

- Tabela 2-A: Subtipos de Embarcações de Pesca

- Tabela 2-B: Subtipos de Embarcações de Recreio

- Tabela 2-C: Subtipos de Embarcações Marítimo-Turísticas

- Tabela 2-D: Subtipos de Artes Caladas (Alagem)

- Tabela 3: Artes de Pesca (Código SIFICAP)

- Tabela 4: Áreas / Águas de Pesca

- Tabela 5: Atividades

- Tabela 6: Tipos de Comportamento que infringem gravemente as regras de Políticas Comum das Pescas

- Tabela 7: Códigos de Infração

- Tabela 8: Espécies de Pescado

De forma a obter uma análise simples, sintetizada e objetiva relativa ao conteúdo do FISCREP, sucedeu-se a organização dos dados. Tendo em conta que existe um tema central, que é o assunto em estudo, designado de “fato”, e que existe um conjunto de

⁴³ Os resultados das ações de fiscalização, realizadas pela MP, são descritos no comunicado FISCREP.

⁴⁴ Fonte: Marinha, 2002, p.7.5.

dados competentes para ser instituída uma ligação entre eles, obtém-se oito dimensões e os respetivos atributos (Tabela 6).

Tabela 6 - Organização dos dados por dimensão e respetivos atributos, referente ao fato

Fato: FISCALIZAÇÃO DA PESCA			
Dimensão	Atributos	Dimensão	Atributos
Unidade	Unidade Naval		
Zona Marítima	Posição de Inspeção	Infrações	Código Infração
	Área		Comportamento Grave – Código
			Presumível Infração
Condições Ambientais	Condições Tempo	Pescado	Quantidade (kg) – Espécie
	Direção e Força do Vento		Número da Página do Diário
	Estado do Mar		Arte de Pesca Utilizada
	Altura e Direção da Ondulação		Classe de Malhagem
	Visibilidade		
Embarcação	Tipo de Embarcação	Resultados Vistoria	Resultado
	Subtipo de Embarcação		Observações (Ação tomada, entre outros)
	Matrícula ou Registo		
	Nome		
	Nacionalidade		
Atividade	Tipo de Atividade	Tempo	Dia
	Arte de Pesca em Uso		Mês
	Classe de Malhagem		Ano
			Hora

Os dados que constituem o comunicado FISCREP estão organizados conforme os diversos assuntos, o que proporciona uma melhor perceção dos dados. Por isso, foi criada uma folha em Excel que contém o fato, as dimensões, os atributos e o local de interesse de cada um dos atributos (vide Compilação dos dados de interesse).

Este processo foi aplicado em todas as medidas de interesse, mas de uma forma mais sintética. Tendo em conta que a explicação que se seguiu teve meramente foco na fiscalização da pesca (por ser a prova de conceito), foi elaborado o Apêndice C para esclarecer o procedimento referente às restantes medidas de interesse.

Posto isto, para que se aplique este tipo de organização e visualização dos dados, é necessário obter esses referidos dados no âmbito da fiscalização da pesca, e de seguida proceder ao seu tratamento.

3.3. Recolha de Dados

O comunicado apurado com mais dados acerca da medida de interesse foi o FISCREP. O ideal seria prosseguir com o projeto referido anteriormente, que neste momento está a ser realizado pela Direção de Tecnologias de Informação e Comunicações (DITIC). Ou seja, a recolha de dados seria diretamente de comunicados FISCREP. No

entanto, atendendo ao fato do projeto não estar concluído, que tem como consequência esse comunicado ainda ser uma mensagem estruturada, surge a necessidade de adquirir dados que sejam o mais semelhante possíveis aos campos que este comunicado engloba. Perante isto, procede-se a uma manipulação do que será o futuro das MTFs.

Os dados idênticos aos que o FISCREP contém, são concedidos pela estrutura do SADAP. É importante salientar que esse sistema possibilita averiguar se a referida organização dos dados auxilia no processo de tomada de decisão, isto porque esse sistema ao fornecer uma mensagem formatada, possibilita uma aproximação do futuro próximo. Os dados recolhidos (Tabela 7), disponibilizados pela DAGI (em formato Excel), são referentes à Fiscalização da Pesca e estão distribuídos numa tabela principal⁴⁵, relativos ao período de janeiro de 1999 a janeiro de 2018.

Tabela 7 - Dados recolhidos com a respetiva descrição

Dados	Descrição
Latitude	Latitude onde se encontrava a embarcação de pesca aquando da fiscalização
Longitude	Longitude onde se encontrava a embarcação de pesca aquando da fiscalização
Ano	Ano no qual ocorreu a fiscalização
Mês	Mês no qual ocorreu a fiscalização
Dia do Mês	Dia do Mês no qual ocorreu a fiscalização
Dia da Semana	Dia da Semana no qual ocorreu a fiscalização
Hora	Hora à qual ocorreu a fiscalização
Período do Dia	Período do Dia em que ocorreu a fiscalização, onde o dia está dividido em 4 período de 6 horas
Unidade	Unidade da Marinha Portuguesa que efetuou a fiscalização
Tipo de Embarcação	Tipo de embarcação a que pertence
Subtipo de Embarcação	Subtipo de embarcação a que pertence
Artes	Arte de pesca fiscalizada, a ser utilizada pela embarcação de pesca
Resultado	Resultado legal da fiscalização
Número FISCREP	Número do FISCREP associado ao registo
Área de Operação	Área relativa ao tipo de pesca (local, costeira, do largo)
Nacionalidade	Nacionalidade da embarcação fiscalizada
Número de Registo	Número de Registo associado à embarcação de pesca
GDH (Grupo-Data-Hora)	Identificador de data e hora
Área Oceânica	Área Oceânica por distância à costa, aquando da fiscalização
Número Infrações Tipo 1, 2, 3, ..., 14	Quantidade de infração verificadas
Local de Inspeção	Local onde ocorreu a fiscalização

Infrações (IONAV1010 ALT5 – Tabela 7, Anexo D)

- Tipo 1: Diário de Pesca inexistente
- Tipo 2: Diário de Pesca preenchido incorretamente
- Tipo 3: Artes Proibidas
- Tipo 4: Pesca em zona proibida ou interdita
- Tipo 5: Pesca proibida por potência motora ou arqueação excessiva
- Tipo 6: Capturas indevidas por pesca direta
- Tipo 7: Capturas indevidas por captura acessória
- Tipo 8: Capturas indevidas por pescado de tamanho inferior ao mínimo legal
- Tipo 9: Atividade exercida sem licença ou autorização
- Tipo 10: Sinalização e/ou identificação indevidas das artes de pesca
- Tipo 11: Sinalização e/ou identificação indevidas das embarcações
- Tipo 12: Diversos – Certificados inválidos
- Tipo 13: Inscrição Marítima inexistente/inválida
- Tipo 14: Diversos (Por exemplo: falta de documentos a bordo, falta de pirotécnicos, extintores caducados, entre outros)

Os dados não serão alvo de qualquer tipo de análise relativa à fiscalização da pesca, ou seja, o objetivo é ter dados que possibilitem a construção de um panorama

⁴⁵ A tabela principal contém 183 547 registos (linhas) de dados.

clarificado para concluir se da forma que será edificado preenche os requisitos para auxílio à tomada de decisão.

Pela observação dos dados recolhidos, pode-se constatar a carência de alguns dados, constituintes dos campos do FISCREP, tais como as condições ambientais e o pescado⁴⁶.

No que concerne à interpretação dos dados, muitas das colunas estão codificadas, com o objetivo de facilitar o tratamento dos dados. A essas referidas colunas correspondem outros ficheiros (também em Excel), designados de tabelas auxiliares, que contêm os dados descodificados, como segue o exemplo através da Figura 16.

C	D	E	F
Ano	Mês	Dia do Mês	Dia da semana
2	1	29	6
2	1	19	3
2	1	19	3
2	1	19	3

MESES		
Janeiro	January	1
Fevereiro	February	2
Março	March	3

DIAS DA SEMANA	
Domingo	1
Segunda	2
Terça	3

Figura 16 - Exemplo de interpretação dos códigos ao respetivo significado

Ou seja, os dados relativos à Fiscalização da Pesca estão distribuídos numa única folha do Excel, sendo que as colunas que estão codificadas correspondem a tabelas auxiliares que esclarecem o significado dos códigos.

As tabelas auxiliares recolhidas dizem respeito aos seguintes dados: ano, mês, dia da semana, período do dia, unidade, tipo de embarcação, subtipo de embarcação, artes, resultado, área de operação, nacionalidade, área oceânica e local de inspeção. Nessas tabelas, algumas dispõem de dados complementares (Tabela 8), essencialmente quando existem possibilidades de referir o mesmo dado (como por abreviaturas, por exemplo).

Tabela 8 - Tabelas auxiliares que dispõem de dados complementares

Tabela Auxiliar	Dados complementares
Área Oceânica	Área Oceânica Abreviada
Área de Operação	Área de Operação Abreviada
Artes	Nome Incompleto
	Abreviatura
	Ficheiros
Local de Inspeção	Local em Maiúsculas
	Local Abreviado
Mês	Mês em inglês
Nacionalidade	Nacionalidade Abreviada
Tipo de Embarcação	Tipo de Embarcação Abreviado
Unidade Naval	Indicativo de Chamada
	Tipo de Navio
	Classe do Navio

⁴⁶ O sistema SADAP permite a leitura do pescado, mas através de uma funcionalidade independente da que lê os FISCREP, daí a ausência destes dados na tabela principal.

Os dados recolhidos estão em concordância com a análise realizada na fonte de dados, mais precisamente, no FISCREP. Ou seja, o tipo de embarcação, o subtipo de embarcação e as artes de pesca estão em conformidade com as tabelas da IONAV1010 (referidas na secção anterior, Tabela 7).

Para além da folha em Excel que possui os dados da medida em estudo, foram recolhidos dados relativos ao tempo (incorporados noutra folha em Excel), respeitantes ao período de 1990 a 2099. Esta folha de dados funciona como um modelo, que contém dados acerca do tempo (Tabela 9). Este modelo é importante porque compreende uma grande diversidade de dados que poderão vir a ser úteis na fase da criação de relatórios.

Tabela 9 - Dados recolhidos relativos ao tempo

Dados - Tempo	Dados - Tempo
Número do Ano	Nome do Semestre do Ano
Número do Mês	Indicação se é Fim-de-Semana
Número do Dia do Ano	Quantidade de Dias do Mês
Número do Dia do Mês	Data (Dia-Mês-Ano)
Número do Dia da Semana	Semana
Número da Semana do Ano	Data com Hora (Ano/Mês/Dia Hora)
Nome do Dia da Semana	Nome da Semana
Nome no Mês	Número da Semana do Mês
Número do Trimestre	Nome da Semana
Nome do Trimestre	Dia da Semana ordenado
Nome do Trimestre do Ano	Número do Ano ordenado
Número do Semestre	
Nome do Semestre	

3.4. Análise e tratamento dos dados recolhidos

A fase de análise e tratamento de dados realiza-se para verificar e resolver eventuais disparidades nos dados recolhidos, para que estes se tornem precisos e coerentes, e para possam ser alvos de uma posterior análise.

Pelo fato de não terem sido recolhidos dados relativos ao pescado, procedeu-se à introdução de duas colunas, na folha do Excel que diz respeito aos dados referentes à Fiscalização da Pesca, que são: espécie do pescado e a respetiva quantidade. A elaboração destas colunas foi feita através da função de números aleatórios, ou seja, são dados fictícios, o que não se torna crítico pelo fato de que nenhum dado será alvo de análise. A inserção destas colunas enriquece a variedade de dados, mesmo que irreais, para a elaboração de um relatório referente a este tipo de dados.

Os dados estão organizados em várias tabelas. As tabelas auxiliares, referidas na secção anterior, caracterizam-se por conter valores exclusivos, isto é, a cada registo

corresponde um número que não existe noutro registo, e que por sua vez também não pode ser nulo⁴⁷.

Sendo assim, estamos perante o conceito da *primary key*⁴⁸ (PK), que efetua a ligação entre as tabelas auxiliares e a tabela principal. Ou seja, as tabelas auxiliares são as tabelas da PK. Perante isto, através da Tabela 10, é possível observar os dados que possuem essa chave, sendo que as duas primeiras tabelas são pertencentes aos dados relativos à Fiscalização da Pesca, e as duas restantes dizem respeito aos dados do tempo.

Tabela 10 - Dados recolhidos com a verificação da *primary key* nas tabelas auxiliares

Dados	Tabela PK	Dados	Tabela PK
Latitude		Resultado	✓
Longitude		Número FISCREP	
Ano	✓	Área de Operação	✓
Mês	✓	Nacionalidade	✓
Dia do Mês		Número de Registo	
Dia da Semana	✓	GDH (Grupo-Data-Hora)	
Hora		Área Oceânica	✓
Período do Dia	✓	Número Infrações	
Unidade	✓	Tipo 1, 2, 3, ..., 14	
Tipo de Embarcação	✓	Local de Inspeção	✓
Subtipo de Embarcação	✓	Espécie de Pescado	✓
Artes	✓	Quantidade de Pescado	

Dados - Tempo	Tabela PK	Dados - Tempo	Tabela PK	Dados - Tempo	Tabela PK
Data	✓	Nome do Trimestre do Ano		Número da Semana do Mês	
Número do Ano		Número do Semestre		Nome da Semana	
Número do Mês		Nome do Semestre		Ano	✓
Número do Dia do Ano		Nome do Semestre do Ano		Mês	✓
Número do Dia do Mês		Indicação se é Fim-de-Semana		Trimestre	✓
Número do Dia da Semana		Quantidade de Dias do Mês		Semestre	✓
Número da Semana do Ano		Data (Dia-Mês-Ano)		Dia da Semana ordenado	
Nome do Dia da Semana		Semana	✓	Número do Ano ordenado	
Nome no Mês		Data com Hora (Ano/Mês/Dia Hora)			
Número do Trimestre		Nome da Semana			
Nome do Trimestre					

Considerar "PK" – Primary Key

Posto isto, sucede-se uma análise pormenorizada dos dados. É preciso ter em consideração que o tipo de embarcação e o subtipo de embarcação são duas colunas distintas, mas interligadas. Ou seja, ambos correspondem a chaves diferentes, no entanto ao tipo de embarcação 1 apenas estão associados os subtipos de embarcação 1.

A análise serve também para verificar a consistência dos dados, mesmo que o valor dos dados não seja alvo de análise, é importante ter uma perceção de algumas incoerências que surgem nas MTFs. Foram detetadas algumas incongruências, pelo que serão especificadas em seguida.

⁴⁷ Um valor nulo simboliza um valor desconhecido ou que não está disponível. É diferente de zero.

⁴⁸ Uma ou mais colunas com a finalidade de identificar exclusivamente cada linha de uma tabela, sendo que não admite valores nulos.

A existência de números decimais em determinados dados que, por isso se tornam ilógicos, nomeadamente: ano, mês, dia do mês, unidade, resultado e área oceânica. Estes dados, exceto o dia do mês, são chaves primárias, e nenhuma das tabelas auxiliares contém chaves com decimais por isso, não possuindo uma correspondência leva à carência de um significado atribuído. Apesar do dia do mês não ser uma chave, é notório que não existem dias com decimais. Um outro dado que compreende um valor irreal é a hora, por ter um valor negativo.

Os dados pertencentes ao Grupo Data-Hora (GDH) são tratados no *software* MATLAB, que faz com que o formato relativo ao GDH, presente na tabela principal, seja inadequado. Esta informação é pertinente porque é de interesse tratar todos os dados da mesma forma, ou seja, deve-se considerar dispor um formato que o sistema a utilizar possibilite um tratamento imediato.

A tabela referente às unidades que realizaram vistorias possui unidades em terra, nomeadamente capitánias que têm uma chave atribuída e outras unidades ligadas a este meio envolvente (como a DGAM), que não correspondem a nenhuma chave. A coluna referente à unidade naval contém apenas as unidades navais, ou seja, a partir do término destas, estão valores em branco que na coluna “indicativo de chamada” correspondem às capitánias. Os relatos de vistoria são elaborados pelas unidades navais e também pelas capitánias⁴⁹, após cada ação de fiscalização. Na presença desta situação, uma solução para tornar os dados mais precisos e perceptíveis é diferenciar unidades navais de capitánias. Nestas tabelas verifica-se que é necessária uma atualização que visa retirar as entidades que não correspondem a chaves.

Em todas as tabelas auxiliares, foi verificada a correspondência entre a chave “0” e o respetivo significado, “TODOS”, ou seja, sempre que o valor na tabela principal (no caso de corresponder a uma tabela auxiliar) for zero, aparece essa expressão. Perante isto, presume-se que é uma chave atribuída quando algo não é coerente. Por exemplo, não seria lógico surgir que uma embarcação era de todos os tipos possíveis de embarcação, ou que tinha obtido todos os resultados aquando de uma vistoria em todos os dias da semana. Apesar de ser importante conter uma chave atribuída a uma expressão para que se perceba, de imediato, que esse dado não está disponível, a expressão atual não é de percepção direta, se por exemplo, for visualizada por um indivíduo que não tem uma visão integrada do assunto.

No que concerne aos dados, incorporados na tabela do Tempo, foi detetado que o mês de dezembro era escrito de duas formas: “Dezembro” e “Decembro”.

Perante a percepção da falta de consistência de alguns dados, como foi dito anteriormente, surgem algumas alterações a serem efetuadas. Relativamente às unidades, não foram consideradas as unidades em terra que não possuíam uma chave, e por isso,

⁴⁹ As capitánias relatam uma inspeção realizada a uma embarcação, estação ou indivíduo.

não tinham dados na tabela principal. No que diz respeito às restantes modificações, nomeadamente: a diferenciação entre unidades navais e capitánias; o ajuste da chave “0” e da respetiva expressão, em situações que os dados não estejam disponíveis; a correção dos dados que apresentam valores decimais; e a incorreta redação do mês de Dezembro, serão tratados posteriormente, durante todo o processo realizado através de um *software*, inicializando esse processo com a inserção dos dados numa base de dados.

3.5. Base de Dados

As bases de dados compreendem ficheiros otimizados com a finalidade de armazenar dados de um modo estruturado, cuja estrutura assegura a integridade e a exatidão desses dados. A gestão e preservação das bases de dados são realizadas através de SGBDs (Magalhães, 2015).

O SGBD aplicado na presente dissertação é o *Structured Query Language (SQL) Server* (2017), cuja linguagem é SQL.

O SQL Server é uma base de dados relacional pois usa tabelas para armazenar dados (Magalhães, 2015). Uma tabela é composta por colunas e linhas, as quais se designam de campos e registos, respetivamente. Este modelo relacional permite a comparação e a verificação de possíveis relações entre os dados que constam nas várias tabelas.

Após a recolha e a análise dos dados, é essencial que sejam dispostos numa forma estruturada para proceder à sua organização. Esta estrutura pode ser alcançada através da inserção dos dados numa base de dados. O SQL Server é constituído por diversas instâncias⁵⁰, utilizadas consoante a finalidade de cada usuário. Tendo em conta que, neste momento, o objetivo é a criação de uma base de dados, a instância a explorar é *Database Engine*.

A base de dados criada designa-se de “FonteDadosMTF”, sendo que corresponde ao sistema OLTP, isto é, um sistema transacional que processa os dados concebidos no dia-a-dia de uma organização.

Primeiramente, foi introduzida a tabela principal, denominada de “Historico”, e de seguida, as tabelas auxiliares. Aquando dessa introdução, os diversos campos correspondem a um tipo de dados e há também a possibilidade de aceitar ou rejeitar valores nulos. Na base de dados “FonteDados MTF” todos os registos permitem nulos.

⁵⁰ Uma instância é um dos componentes do SQL Server, que se caracteriza pela combinação de vários processos. Um utilizador pode ter várias instâncias, e que cada uma possui um conjunto de características.

No que concerne à tabela “Historico”, devido à ausência de chaves primárias, procedeu-se ao incremento automático da chave primária⁵¹.

É importante referir que a tabela referente ao tempo não foi introduzida na base de dados, o que se justifica pelo fato de não ter chaves de ligação com nenhuma outra tabela e, por isso, não subsistia essa necessidade. O processo da não inserção dos dados na base de dados e, de serem tratados posteriormente (como veremos na secção referente ao processo ETL), podia ter sido aplicado em todas as tabelas, mas visto que existia uma ligação entre elas (através das chaves), optou-se por criar uma base de dados de forma a organizar os dados.

Posteriormente, foram modificados os nomes dos campos de modo a retirar os espaços em branco e a acentuação das palavras. É importante referir que, na alteração dos nomes dos campos, especificamente nas tabelas auxiliares, o valor “zero” e a expressão “TODOS” foram substituídos por nomes identificadores de cada coluna, um que revela a PK e o outro o tipo de assunto a que se refere (Figura 17).



Figura 17 – Primeiro passo para evitar confusão na interpretação de dados

Assim, a confusão que poderia ocorrer por “0” corresponder à expressão “TODOS”, dá-se por finalizada. Para além disso, aos campos que não correspondem a quantitativos, aplicou-se a substituição de zero por “-1” (Figura 18), para ser um valor inconsistente, exceto na hora, que foi atribuído o valor “9999”.

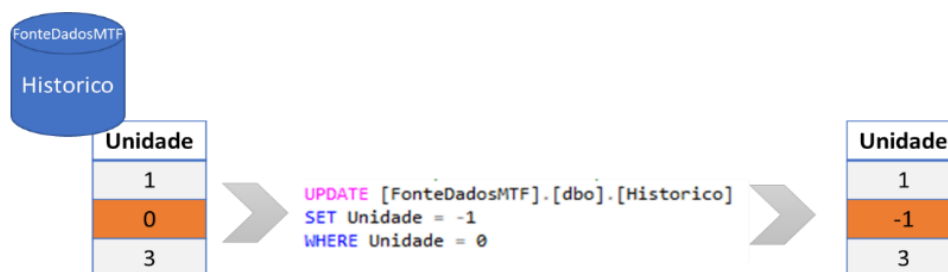


Figura 18 - Atualização da tabela principal, nomeadamente o campo referente à unidade naval

Tendo em conta que “-1” não era uma chave, passaram a existir valores na tabela principal, que carecem de uma decodificação. Por isso, inseriu-se pela ordem dos vários atributos de cada campo, essa nova chave a corresponder à expressão “N/D”. (Não Definido) (Figura 19).

⁵¹ O incremento automático da chave primária permite a criação dessa chave e atribuição do valor de incremento, neste caso foi estabelecido um incremento de valor um.



Figura 19 – Inserção de valores na tabela auxiliar que diz respeito à unidade naval

Posto isto, o ajuste da chave e da respetiva expressão, ou seja, “-1” e “N/D” respetivamente, está concretizado para todas as tabelas – ver *Query* para tratamento de dados com chave “0”.

Sendo assim, o tratamento dos valores decimais pode ser tratado através desta nova chave. Ou seja, os dados que possuíam essa falta de lógica, foram direcionados para a chave que corresponde aos valores com falta de coerência, que neste momento é “-1” (vide *Query* para tratamento de dados decimais).

De seguida, procedeu-se à inserção do tipo de dados desde a tabela principal às tabelas auxiliares. Esta etapa não foi realizada logo inicialmente com a finalidade de realizar primeiramente algumas correções, como foi visto anteriormente. Isto porque os dados ao serem inseridos neste *software*, assumem um determinado tipo de dados, e era importante que os decimais não fossem classificados como inteiros, pois aconteceria que o valor “1.235” passava a “1”, e não a uma expressão que revela que esse dado não é possível de analisar.

Nesta etapa, é importante ter uma perceção minuciosa do tipo de dados, ou seja, saber especificamente do que se tratam. A classificação dos dados foi determinada, da seguinte forma:

- Caracteres⁵²: *varchar(n)*, em que *n* corresponde ao comprimento da cadeia de caracteres e varia consoante a dimensão de cada dado. A definição deste tipo deu-se porque armazena apenas o tamanho do dado inserido⁵³, e por isso, não origina espaços em branco;

⁵² Uma cadeia de caracteres ou *string* é, neste caso, uma sequência de letras ou símbolos (como espaços em branco e acentuação).

⁵³ Ao ser selecionado outro tipo de dados, como o char, o tamanho corresponderia ao valor inserido. Por exemplo, a palavra “unidade” possui sete caracteres, se fosse selecionado char (10) (por exemplo), o campo passava a armazenar dez espaços, pelo que se o dado apenas contém sete caracteres, criará um espaço em branco (que são os restantes três que não estão a ser utilizados), e é algo que se deve procurar evitar.

- Numéricos exatos⁵⁴: *smallint*, *int*, *bigint* e *decimal(p,s)*, em que *p* e *s* representam a precisão⁵⁵ e a escala⁵⁶, respetivamente. O que diferencia os três primeiros tipos de dados é o intervalo de valores aceitáveis e a quantidade de espaço que ocupa na base de dados. A escolha destes tipos de dados ocorreu de acordo com o intervalo de valores admissíveis dos dados. A seleção do último tipo deveu-se à necessidade de ter os valores exatos, ou seja, sem arredondamentos;
- Numéricos aproximados⁵⁷: *float*. Este tipo de classificação foi escolhido pela sua simples precisão, ou seja, nem todos os valores são representados na forma exata, pelo que poderão suportar arredondamentos;
- Hora: *time(s)*, em que *s* corresponde à escala de frações de segundos. A seleção deste tipo de dados sucedeu-se porque engloba apenas a hora, e porque corresponde a um período de 24 horas, que é o pretendido.

Perante esta visão geral do tipo de dados escolhidos, obtém-se uma visão específica, através do Apêndice E, referente aos dados da tabela principal e das tabelas auxiliares.

Um dos aspetos referidos aquando da análise detalhada dos dados, foi a relação entre os campos: tipo de embarcação e subtipo de embarcação. Para alcançar a correspondência correta, foi elaborada uma *query*⁵⁸ (Figura 20). Desta forma, consoante o tipo de embarcação selecionado, subsistem os respetivos subtipos.

⁵⁴ Os numéricos exatos representam-se por um número finito de valores que, neste caso, são utilizados para armazenar valores inteiros (compreendem valores negativos, positivos e zero, com a restrição de não englobar números decimais).

⁵⁵ A precisão traduz-se no número total de dígitos a conservar, tanto à esquerda como à direita da vírgula.

⁵⁶ A escala consiste no número de casas decimais a armazenar, à direita do ponto decimal.

⁵⁷ Os numéricos aproximados não armazenam os valores exatos para muitos números, mas sim uma aproximação muito próxima do valor.

⁵⁸ *Query* é uma consulta que pode ser realizada através de vários comandos (como “SELECT”, “UPDATE”) para obtenção do que se pretende.

```

SELECT 'Pesca comercial' as Tipo_Embarcacao,
'P' as Tipo_Abreviado, 1 as PK_Tipo_Embarcacao,
Subtipo_capslock as Subtipo_Capslock,
Código as SubtipoCodigo, Ficheiros as
FicheirosSubtipo, PK_Subtipo1 as PK_Subtipo,
Subtipo
from [FonteDadosMTF].[dbo].[Subtipo1Pesca]

union all
SELECT 'Recreio' as Tipo_Embarcacao,
'R' as Tipo_Abreviado, 2 as PK_Tipo_Embarcacao,
'N/D' as Subtipo_Capslock,
'N/D' as SubtipoCodigo, 'N/D' as
FicheirosSubtipo, PK_Subtipo2 as PK_Subtipo,
Subtipo
from [FonteDadosMTF].[dbo].[Subtipo2Recreio]

union all
SELECT 'Maritimo turisticas' as
Tipo_Embarcacao,
'MT' as Tipo_Abreviado, 3 as
PK_Tipo_Embarcacao, 'N/D' as Subtipo_Capslock,
Codigo as SubtipoCodigo, 'N/D' as
FicheirosSubtipo, PK_Subtipo3 as PK_Subtipo,
Subtipo
from [FonteDadosMTF].[dbo].[Subtipo3MarTur]

union all
SELECT 'Artes caladas' as Tipo_Embarcacao,
'AC' as Tipo_Abreviado, 4 as
PK_Tipo_Embarcacao, 'N/D' as Subtipo_Capslock,
Codigo as SubtipoCodigo, 'N/D' as
FicheirosSubtipo, PK as PK_Subtipo,
Subtipo
from
[FonteDadosMTF].[dbo].[Subtipo4ArtesCaladas]

union all
SELECT 'Outras' as Tipo_Embarcacao,
'O' as Tipo_Abreviado, 5 as PK_Tipo_Embarcacao,
'N/D' as Subtipo_Capslock,
Codigo as SubtipoCodigo, 'N/D' as
FicheirosSubtipo, PK as PK_Subtipo,
Subtipo
from [FonteDadosMTF].[dbo].[Subtipo5Outras]

```

Figura 20 - Query que estabelece a relação entre o tipo de embarcação e o subtipo de embarcação

Após a inserção dos dados recolhidos na base de dados e do tratamento de algumas das incoerências, a próxima etapa é construir um modelo que cumpra com a organização estabelecida *a priori* para os dados, descrito no próximo capítulo.

CAPÍTULO 4

CONCEÇÃO DA SOLUÇÃO

- 4.1. Modelação Dimensional
- 4.2. Processo Extração, Transformação
e Carregamento
- 4.3. Cubo Multidimensional
- 4.4. Visualização da Informação

4. Capítulo 4 – Conceção da Solução

4.1. Modelação Dimensional

A fase da Modelação Dimensional tem como objetivo: obter uma organização dos dados. Essa organização consiste na distribuição dos dados pelas dimensões e respetivos atributos referentes a um determinado fato. Ou seja, por outras palavras, pretende-se edificar um modelo dimensional, que será o DW. Para isso, foi criada outra base de dados, dentro do mesmo projeto, designada de *Data Warehouse*.

O Modelo Dimensional é uma forma de moldar o tipo de relações entre os dados. Este modelo pode assumir vários esquemas, dos quais se destaca o esquema em estrela, que será o esquema a utilizar. A seleção deste tipo de esquema deu-se porque, apesar da falta de normalização, possibilita consultas simples, uma fácil interpretação, flexibilidade para alterações e também pela carência de necessidade de relacionar dimensões.

Este esquema tem no centro o fato, que neste caso é a Fiscalização da Pesca, e está circundado de dimensões. Enquanto que o fato dispõe a componente quantitativa dos dados, as respetivas dimensões são formadas por dados qualificativos. Neste caso, o tipo de fato é não aditivo porque nenhuma das suas medidas possui a possibilidade de realizar operações (soma, por exemplo) com os atributos das dimensões. Relativamente ao tipo de relação entre a tabela de fato e as tabelas de dimensão, é uma relação denominada de “um para muitos”, porque a tabela de fato está associada aos vários atributos das tabelas de dimensões, mas o invés não, ou seja, as tabelas de dimensões apenas correspondem a uma medida da tabela de fatos.

A tabela do fato é constituída por *foreign keys*⁵⁹ e por medidas⁶⁰. Como se viu anteriormente, cada tabela auxiliar está associada a uma PK. Deste modo, as tabelas auxiliares serão os atributos das dimensões. Ou seja, a *primary key* de cada dimensão ligar-se-á à chave correspondente na tabela de fatos – a designada *foreign key*.

Posto isto, o modelo em estrela que se obteve foi o seguinte (Figura 21).

⁵⁹ As *foreign keys* são chaves constituintes da tabela de fato que estabelecem ligação com as tabelas de dimensão.

⁶⁰ Medidas são valores numéricos que representam o fato.

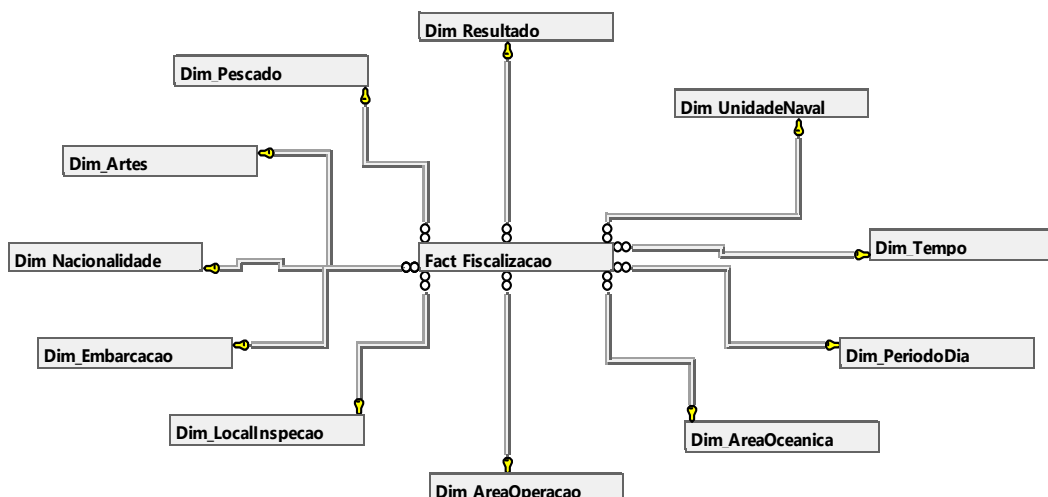


Figura 21 - Esquema em estrela, com as respectivas relações às dimensões

Para uma observação discriminada, nomeadamente com as medidas (tabela do fato) e os atributos (tabelas das dimensões), foi elaborado o Apêndice F. Através da análise do esquema detalhado, verifica-se que existem dados da tabela principal que não foram colocados nesta nova base de dados, nomeadamente: GDH, número de registo, latitude e longitude. Esta não colocação deve-se ao fato de serem dados que não são considerados de interesse⁶¹ para o que se pretende: visualização da informação relevante que possa influenciar na tomada de decisão. Por outro lado, foi criado um campo, na tabela do facto, que se designa de “DataETL” para aquisição das datas dos carregamentos dos dados (processo ETL).

A modelação dimensional foi também aplicada nas restantes medidas de interesse, com base na compilação dos dados de interesse e das respetivas conclusões. Esta modelação tem como finalidade ser alvo de uma futura utilização, daí a elaboração do Apêndice G – Modelação dimensional aplicada nas restantes medidas de interesse.

O modelo em estrela está criado, mas vazio em termos de dados. Para se efetuar o carregamento dos dados nas diversas dimensões, será necessário percorrer o processo ETL.

4.2. Processo Extração, Transformação e Carregamento

O processo ETL visa tornar os dados úteis para serem submetidos a uma análise. Os referidos dados para se tornarem úteis, têm de percorrer os processos de tratamento e eliminação de incoerências.

Para a realização desta fase, a instância a explorar é *SQL Server Integration Services (SSIS)*, através do *software Microsoft Visual Studio (2017)*. Esta plataforma

⁶¹ A latitude e a longitude poderiam ser dados de interesse, se o sistema a ser concebido abrangesse informação geográfica.

possibilita a importação de dados de outras fontes, a manipulação desses dados (Magalhães, 2015, p.77), que se traduz no processo ETL, efetuando o carregamento para outras fontes de natureza *Data Warehousing*.

Primeiramente, foi criado um projeto, designado de *VerifyDW*. De seguida, foram estabelecidas as conexões às bases de dados de origem e de destino, *data source* e *data destination*. Concretamente, a origem corresponde à base de dados *FonteDadosMTF* e o destino à base de dados *DataWarehouse*.

Posto isto, foram criados (primeiro as dimensões e, por último o fato), os seguintes pacotes SSIS⁶²: *Dim_AreaOceanica*, *Dim_AreaOperacao*, *Dim_Artes*, *Dim_LocalInspecao*, *Dim_Nacionalidade*, *Dim_PeriodoDia*, *Dim_Pescado*, *Dim_Resultado*, *Dim_Tempo*, *Dim_TipoEmbarcacao*, *Dim_UnidadeNaval* e *Fato*.

Seguidamente, em todas as dimensões, nos respetivos pacotes, foi criado no *Control Flow*, uma *Data Flow Task*. Esta tarefa agrega os dados entre a origem e o destino, através do fluxo de dados, onde são determinadas as ferramentas a utilizar para a limpeza, transformação ou alteração dos dados.

O fluxo de dados, aplicado nas dimensões (exceto a dimensão referente ao Tempo), consistiu na aplicação das seguintes ferramentas: ADO NET *source*, *data conversion*, *derived column*, *lookup* e ADO NET *destination* (Figura 22). Estas ferramentas têm as seguintes potencialidades:

- ADO NET *Source*⁶³ (*FonteDadosMTF*) – conexão e extração dos dados da origem;
- *Data Conversion* – Conversão do tipo de dados entre a coluna de entrada e a coluna de saída;
- *Derived Column* (*NULL* e *WhiteSpaces*) – aplica expressões, através de funções, para criar ou substituir valores existentes, originando uma nova linha de saída;
- *Lookup* – pesquisa das chaves nas tabelas de destino efetuando a correspondência entre a coluna de entrada disponível e a coluna de destino disponível;
- ADO NET *Destination*⁶⁴ (*DataWarehouse*) – carregamento, atualização e eliminação dos dados no destino, neste caso, nas respetivas dimensões.

⁶² Os pacotes SSIS armazenam o fato e as respetivas dimensões.

⁶³ ADO NET *Source* permite a leitura de um fornecedor .Net, e dispõe esses dados no fluxo de dados.

⁶⁴ ADO NET *Destination* possibilita o carregamento dos dados numa base de dados compatível com ADO.NET.

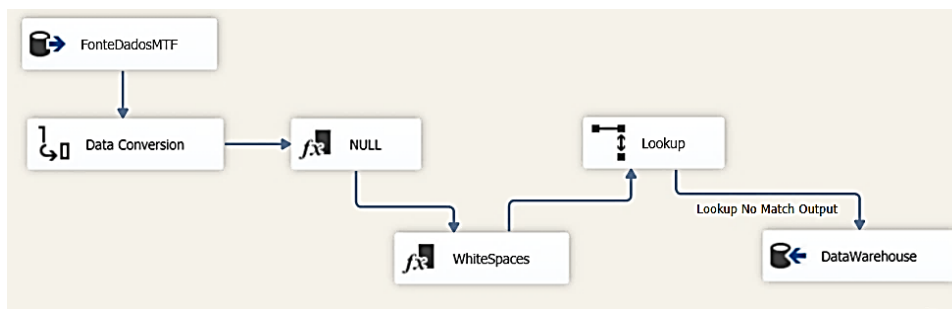


Figura 22 - Ferramentas utilizadas para o carregamento dos dados nas respectivas dimensões

No que diz respeito à *derived column*, foram introduzidas duas: uma consiste na detecção de espaços vazios, também designados de “NULL”, sendo que ao serem detetados a coluna derivada substituirá por “N/A” (Não Aplicável); a segunda (apenas aplicável em caracteres) visa eliminar os espaços em branco, que se encontram à direita e à esquerda de uma sequência de caracteres, através das funções “RTRIM” e “LTRIM”, para que retorne a sequência num formato de saída que seja o mais preciso possível.

Na dimensão que diz respeito à Unidade Naval, nomeadamente o campo que contém a unidade naval, os valores “NULL” foram substituídos por “Capitania”, porque como já foi referido, a partir do término das unidades navais, os campos vazios correspondem a capitánias (na coluna “indicativo de chamada”). Posto isto, as unidades navais e as capitánias estão diferenciadas (Figura 23).

Indicativo de chamada	Unidade Naval
IndicativoX	UnidadeX
IndicativoY	UnidadeY
...	...
CapitaniaA	NULL
CapitaniaB	NULL

➤

Indicativo de chamada	Unidade Naval
IndicativoX	UnidadeX
IndicativoY	UnidadeY
...	...
CapitaniaA	Capitania
CapitaniaB	Capitania

Figura 23 - Tratamento da incoerência referente à diferenciação entre unidades navais e capitánias

Relativamente à dimensão Tempo, as alterações no fluxo de dados em comparação às restantes dimensões, foram as seguintes: a conexão da origem dos dados é referente a uma tabela do Excel (modelo recolhido), e possui uma outra coluna derivada, para além das que foram referidas acima, que tem como função substituir o caractere “Dezembro” por “Dezembro”, através da função “Replace”.

Perante isto, as dimensões estão carregadas no esquema em estrela. Assim, segue-se o carregamento da tabela do fato.

No que concerne à tabela do fato (*Fact_Fiscalizacao*), foram criadas, no *control flow*, as seguintes tarefas: *Execute SQL Task*⁶⁵ (*EliminarDados*), *Data Flow Task* (*CarregarDados*) e *Execute SQL Task1* (*NULL*) (Figura 24).

⁶⁵ A *Execute SQL Task* executa uma ou mais instruções, de forma sequencial.

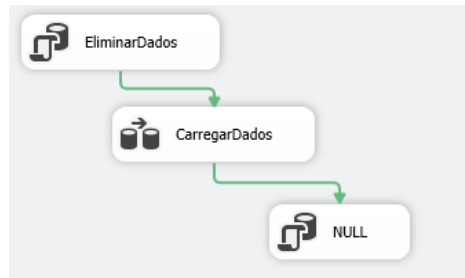


Figura 24 - Control flow referente à tabela do fato

A primeira tarefa serve para eliminar quaisquer dados existentes na tabela do fato, isto porque sempre que existir necessidade de carregar novamente o fato, todos os dados que o constituem são eliminados para que possam ser armazenados os novos dados (atualização ou correção de dados, por exemplo) (Figura 25).

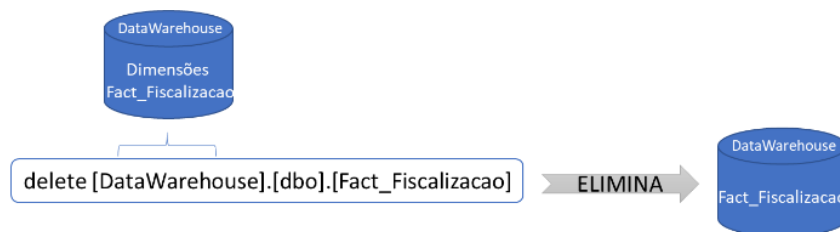


Figura 25 - Comando SQL para eliminar os dados existentes na tabela do fato

A razão pela qual se optou por eliminar a tabela do fato foi, porque neste caso específico, a estrutura do SADAP apresenta constitui um histórico, ou seja, possibilita eliminar visto que os novos dados a serem carregados permitem uma análise do que se passou anteriormente.

A segunda tarefa é semelhante ao fluxo de dados realizado nas dimensões, mas com algumas diferenças. O carregamento dos dados na tabela do fato, procedeu-se da seguinte forma (os nomes que constam entre parêntesis dizem respeito ao nome inserido nas diversas ferramentas):

- ADO NET *Source (QueryFonteDadosMTF)* – diz respeito a uma *query*, em que através da tabela principal (ou “Historico”) são especificados os dados que se pretendem extrair para o Fato. Os dados selecionados apenas precisam de ter uma seleção de cada tabela, por isso, foram selecionados os campos que envolviam menos caracteres ou valores, para facilitar o tratamento dos dados (ver *Query* para extração de dados da *FonteDadosMTF*). Com esta *query*, os dados que estão codificados, passam a visualizar-se decodificados;
- *Data conversion* – conversão do tipo de dados, para estarem em concordância com o tipo de dados estabelecidos na tabela do Fato;
- *Derived Column (WhiteSpaces)* – abrange a função de eliminação dos espaços em branco, tanto à direita como à esquerda dos caracteres;
- *Lookup, Derived Column1 e Union All* – aplicados em conjunto, com a finalidade da pesquisa de chaves com a união do valor “-1” para os dados inexistentes. Este

conjunto foi aplicado a todas as dimensões, exceto à dimensão tempo por esta englobar todos os dados.

- ADO NET *Destination (DataWarehouse)* – carregamento na tabela de fato, referente à base de dados *DataWarehouse*.

É importante salientar que a última *derived column*, designada de “DataETL”, contém a função *GETDATE*. Esta função retorna a data atual do sistema, o que possibilita saber a respetiva data e hora em que foi executado o processo ETL para o carregamento de dados na tabela de fato.

Para uma visão particular do processo executado para o carregamento dos dados na tabela do fato, ver *Processo de Carregamento de Dados*.

A terceira tarefa, realizada através de uma *query*, traduz-se em verificar todas as medidas que incluem dados “NULL”, para serem substituídas por “0”. Sendo que possuem um caráter quantitativo, o valor zero é alvo de interpretação (vide *Query* para dados com campos *NULL*).

Em suma, o processo ETL foi concluído com sucesso, o que significa que o fato e as respetivas dimensões estão conservados com os dados limpos, tratados e integrados.

Posto isto, os dados são precisos e podem ser analisados. A análise dos dados, estando perante um DW, é realizada através de cubos multidimensionais.

4.3. Cubo Multidimensional

A estrutura das relações estabelecidas na modelação dimensional dá origem à formação do cubo⁶⁶.

O cubo tem como base a tecnologia OLAP⁶⁷, isto é, uma tecnologia que visa fornecer ferramentas e operações que possibilitam o manuseamento dos dados.

O cubo tem diversas funcionalidades, entre as quais o cruzamento de dados e a criação de hierarquias entre os vários atributos de uma determinada dimensão. Isto possibilita a obtenção de variadas perspetivas da informação de interesse.

Para a execução desta fase, a instância a explorar foi *SQL Server Analysis Services (SSAS)*, através do *software Microsoft Visual Studio (2017)*. Esta plataforma proporciona uma visualização integrada de todos os dados (análises OLAP, relatórios, entre outros⁶⁸), o que possibilita a conceção e a manutenção de cubos constituintes de dados agregados

⁶⁶ Também conhecido por “cubo multidimensional” ou “estrutura multidimensional”.

⁶⁷ A tecnologia OLAP permite o desenho, a criação e a manutenção dos cubos multidimensionais.

⁶⁸ *Data Mining*, documentos do Office, entre outros.

provenientes de outras fontes (como a base de dados relacional, por exemplo) (Magalhães, 2015, p. 8).

Primeiramente, foi criado um projeto, designado de *CuboOLAP*, onde foram criadas fontes de dados e de visualização de dados, denominadas de *DataWarehouse*. A fonte de dados tem como conexão a base de dados de destino (também designada de *DataWarehouse*) que abrange o fato e as dimensões, enquanto que a fonte de visualização de dados está conectada à fonte de dados concebida.

Posto isto, procedeu-se à formação do cubo *DataMart_Fiscalizacao*. Para o cubo foram selecionados o fato e as respetivas dimensões. Para cada dimensão, foram definidos os atributos que se pretendiam para uma posterior análise e visualização. Ou seja, cada dimensão é constituída por vários atributos, no entanto nesta fase há a possibilidade de os particularizar para que cada dimensão fique apenas com os atributos que serão alvo de análise, consoante a necessidade do utilizador. Neste caso, em todas as dimensões foram consideradas os respetivos atributos e as respetivas chaves primárias, exceto na dimensão referente ao tempo.

No que diz respeito à dimensão *Tempo*, foram apenas considerados os atributos para a conceção de uma hierarquia. A hierarquia foi construída para possibilitar uma perspetiva lógica dos dados, ou seja, ao serem distribuídos hierarquicamente e coerentes torna-se mais fácil para o utilizador analisar os dados organizados. Por isso, a hierarquia contém os seguintes atributos: ano > semestre > trimestre > mês > número do dia (Figura 26).

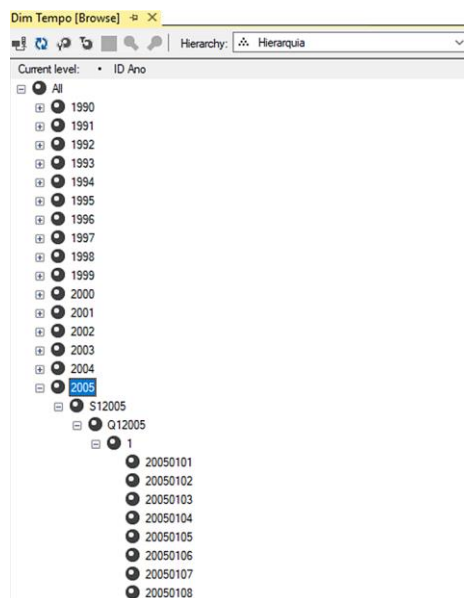


Figura 26 – Hierarquia respeitante à dimensão do tempo

Relativamente às medidas do fato, particularmente a medida “DataETL”, foi aplicado o máximo, ou seja, retorna a data mais recente.

Assim, segue-se o processamento e a edificação do cubo para verificar se os dados que estão distribuídos pelo fato e pelas dimensões estão em conformidade (Figura 27).

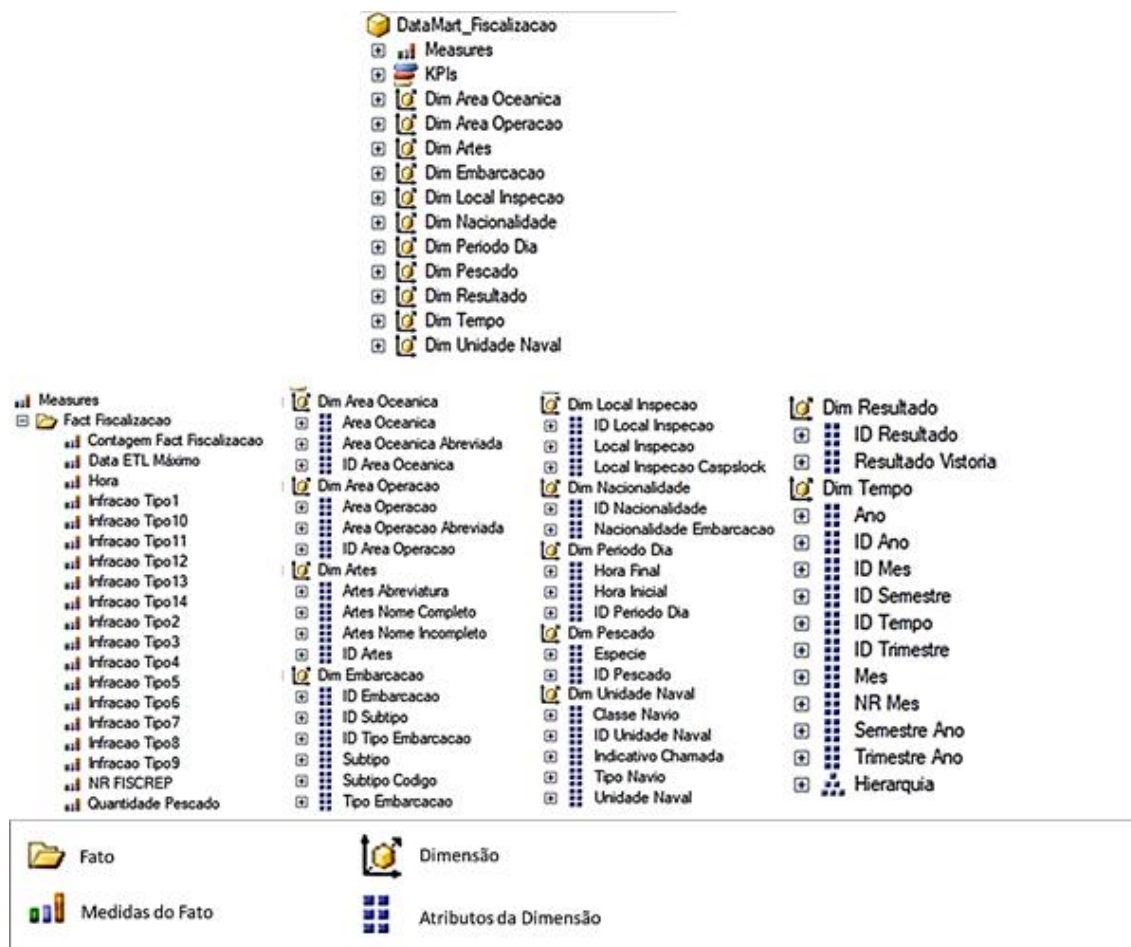


Figura 27 - Informação organizada no cubo multidimensional, através do fato e respetivas dimensões

Deste modo, o cubo foi concluído com sucesso e assume-se que os dados estão disponíveis num formato organizado e coerente, através do fato e das respetivas dimensões para proporcionar uma visualização o mais precisa e útil possível.

4.4. Visualização da Informação

Perante todo o procedimento efetuado, nomeadamente a criação das bases de dados, a modelação dimensional, o processo ETL e a edificação do cubo multidimensional, dar-se-á início a uma das grandes potencialidades do *Business Intelligence*: cruzamento e visualização da informação de interesse num *dashboard*⁶⁹. É

⁶⁹ Um *dashboard* é um painel que abrange a informação relevante distribuída por diversos diagramas (designados de relatórios – gráficos, tabelas, entre outros).

importante referir que os dados se transformam em informação a partir do momento que são disponibilizados para o consumidor final.

Para se dar início à conceção da referida visualização, foi definida a informação de interesse, ou seja, a que se vai dispor nos diversos relatórios⁷⁰. Tendo em conta a informação disponível, a que se considerou relevante para apresentar foi a seguinte: número de fiscalizações ao longo do ano, entidade que realizou as vistorias, tipo e nacionalidade das embarcações vistoriadas, artes de pesca, presumíveis infratores por área e tipo de pescado verificado nos presumíveis infratores.

Para a realização desta fase, foi necessária a instalação da instância referente ao *SQL Server Reporting Services (SSRS)*. Esta instância dispõe de um servidor de relatórios, que dá acesso à ferramenta *Report Builder*. Esta ferramenta possibilita a criação de relatórios e, por isso foi utilizada para realizar esta etapa.

Primeiramente, no servidor de relatórios, foi criada uma fonte de dados conectada ao *CuboOLAP*, designada de *Cubo*.

De seguida, através do *Report Builder*, procedeu-se à criação dos relatórios. Os relatórios foram elaborados de forma individual, isto é, foram produzidos diversos relatórios para cada tipo de informação, até se obter toda a informação consolidada. Após isso, foi criado um *dashboard* (conjunto de relatórios) com toda essa informação.

Para a construção de cada relatório, foi estabelecida uma conexão ao *Cubo*. Seguidamente, a informação de interesse foi organizada e agrupada de forma lógica (Figura 28).

Número total de Fiscalizações	Número total de Embarcações Portuguesas Vistoriadas	Número total de Presumíveis Infratores
Fiscalizações ao longo do ano	Tipo e Nacionalidade da Embarcação Vistoriada	Presumíveis Infratores por Área Oceânica
Entidades (unidade naval ou capitania) com mais vistorias realizadas	Artes de pesca mais verificadas	Tipo de Pescado dos Presumíveis Infratores

Figura 28 - Estrutura do *dashboard*, com a respetiva informação relevante distribuída

Para que a estrutura do *dashboard* seja alcançada, é fundamental obter a informação necessária. Essa obtenção ocorre através da criação de *datasets*. Um *dataset* é a informação ou o conjunto de informação selecionada de entre as opções (medidas do fato e atributos das dimensões) para ser analisada.

⁷⁰ Conjunto de gráficos que constituem o *dashboard*.

Foram criados dez *datasets* para a estrutura definida, onde se estabeleceu o cruzamento de informação, através dos atributos e das medidas disponíveis. É importante referir que foram criados dois *datasets* para a secção referente ao tipo de embarcação e à nacionalidade da embarcação vistoriada, respetivamente. Relativamente aos três valores totais, que se encontram na primeira “linha” da estrutura do *dashboard*, foram aplicadas funções de agregação, de forma a proporcionar um panorama referente a três tipos de informação importantes. A única medida calculada⁷¹ visa determinar o período homólogo, como se pode observar na Figura 29.

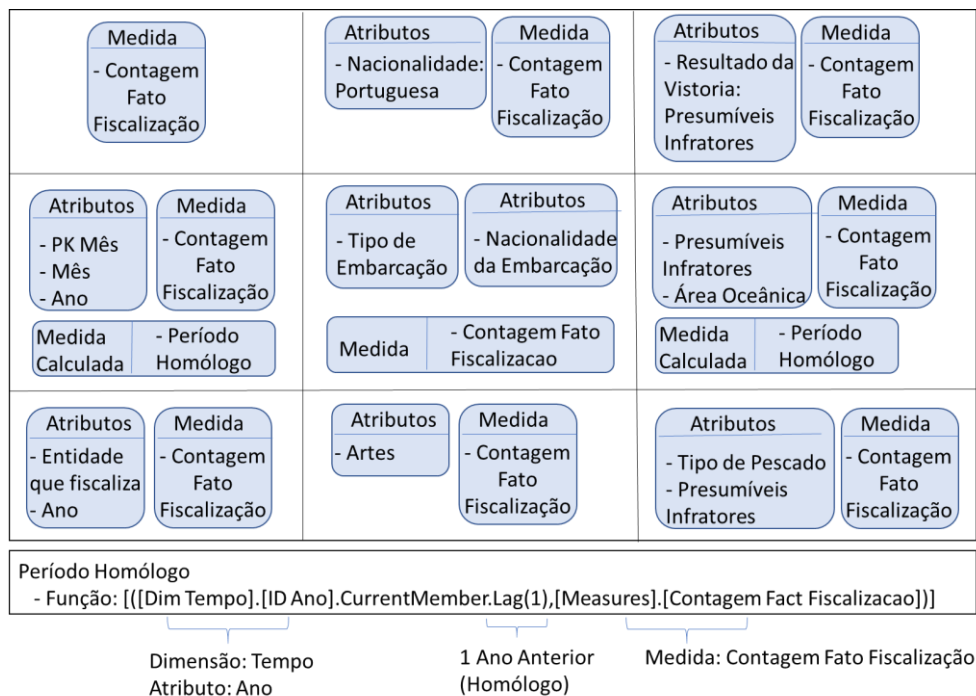


Figura 29 - Estrutura do dashboard com a informação selecionada para os *datasets*

É importante constatar que nem toda a informação foi considerada para o dashboard, isto porque este painel deve ter informação relevante, fácil de analisar e sumariada. Por isso, foram criadas algumas hiperligações, onde se pode ter uma visão completa de toda a informação. Estas hiperligações estão conectadas a relatórios individuais com informação complementar, nomeadamente: subtipo de embarcação, unidade naval, capitania, resultados das vistorias, artes de pesca e tipo e quantidade de pescado. Como tal, para cada um destes relatórios foram também criados *datasets* com a respetiva informação (Figura 30).

⁷¹ Medidas calculadas são operações possíveis de realizar através de funções (como máximo, mínimo, soma, entre outras). Estas medidas podem se efetuadas na fase do cubo multidimensional.

Relatório	Atributos	Medida
Capitania	Indicativo de Chamada	Contagem Fato Fiscalização
Unidade Naval	Unidade Naval	
Subtipo de Embarcação	Subtipo	
Artes de Pesca	Artes – Nome Completo	
Resultados das Vistorias	Resultado da Vistoria	
Tipo e Quantidade de Pescado	Espécie	
	Quantidade de Pescado	

Figura 30 - Informação selecionada para os *datasets* referente aos relatórios complementares ao dashboard

Foi também criado um outro *dataset*, tanto no dashboard como nos restantes relatórios, para se obter a data de atualização, que é uma informação pertinente para o utilizador. Este *dataset* é composto pela medida “DataETL”.

Nos *datasets* do *dashboard* e dos restantes relatórios, exceto no *dataset* referente à data de atualização, foram aplicados parâmetros⁷² com base na hierarquia contruída (ano, semestre, trimestre e mês). Esta hierarquia foi aplicada na dimensão “Tempo”, como já foi descrito anteriormente, o que possibilita uma escolha de entre os valores existentes. Posto isto, no *dashboard* e nos restantes relatórios foi definido o ano de 2017 (Figura 31) para possibilitar uma vista mais completa da informação, visto que do ano atual apenas se possui informação relativa ao mês de janeiro.

Ano Semestre Trimestre Mês

Figura 31 - Parâmetros estabelecidos para o *dashboard* e para os relatórios complementares

Somente nos relatórios individuais foram criados parâmetros referentes a esse tipo de informação, ou seja, no relatório que diz respeito às artes de pesca é possível selecionar, se existir essa necessidade, apenas as artes que se pretendem analisar (Figura 32).

Artes Ano Semestre Trimestre
Mês

Figura 32 - Parâmetros estabelecidos para o relatório das artes de pesca

De seguida, procede-se a escolha dos diagramas a inserir em cada uma das partes da estrutura definida. A informação a inserir nos diagramas é retirada a partir dos *datasets* criados. Quando se trata de gráficos, a informação pode ser colocada nos seguintes campos: *values*, *category groups* e *series groups*, ou seja, o que é referente a valores – contagem do fato fiscalização e período homólogo – são inseridos nos *values*, e consoante ao que diz respeito o gráfico é inserido em *category* ou *series groups*. Como por exemplo, na parte relativa às fiscalizações ao longo do ano, foram considerados: *values* – Contagem do Fato Fiscalização e Período Homólogo, e *category groups* – Mês.

⁷² Parâmetros são filtros que se definem para possibilitar a escolha de determinados atributos. Por exemplo, se aplicar um parâmetro referente ao ano é possível escolher o ano que se pretende analisar a informação.

Sendo assim, os *datasets* foram distribuídos pelos diversos gráficos, apresentados de forma ordenada a partir dos valores mais elevados, o que deu origem ao *dashboard* e aos relatórios com a informação completa.

No que diz respeito aos relatórios, provenientes das hiperligações, englobam o máximo de informação. O subtipo de embarcação e as artes de pesca possuem dois gráficos circulares, um com os valores mais elevados e outro com os valores mais baixos, e uma lista com todos os subtipos e artes suscetíveis de selecionar apenas os que se pretendem observar, através dos parâmetros. As unidades navais, as capitánias e o tipo e quantidade de pescado apresentam uma lista ordenada de modo descendente. O resultado das vistorias para além dessa lista, engloba um quadro onde estão distribuídos os vários resultados pelas diversas áreas oceánicas.

Para uma observação geral da distribuição da informação pelos diversos diagramas, o *dashboard* e os relatórios complementares estão representados no Apêndice I.

Posto isto, sucede-se a fase de avaliação do artefacto produzido de forma a averiguar se os objetivos e os requisitos são adequados às necessidades.

CAPÍTULO 5

ANÁLISE DE RESULTADOS

5.1. Avaliação do Artefacto

5.2. Análise dos Resultados Obtidos

5. Capítulo 5 – Análise de Resultados

5.1. Avaliação do Artefacto

A avaliação de um sistema produzido consiste em verificar se os objetivos e os requisitos do artefacto correspondem às necessidades. Uma das formas de avaliar a funcionalidade do artefacto é verificar se corresponde às necessidades dos “utilizadores-alvo”. Deste modo, para a obtenção de dados no que diz respeito à solução edificada e, tendo em conta que há um público-alvo identificado, sucede-se a seleção do instrumento de investigação – questionários.

Albarello, em 1997, mencionou que “o questionário é tanto um ponto de chegada de uma reflexão como o ponto de partida para análises ulteriores” (como referido em Cerqueira & Teixeira, 2015, p. 157). O questionário é utilizado num determinado meio de investigação, e serve como suporte para a recolha de dados, devendo apresentar algumas características como ser claro, objetivo e de fácil interpretação (Lopes, 2006, p. 241).

O presente questionário tem como objetivo investigar se a solução edificada auxilia no processo de tomada de decisão, nomeadamente, se o tipo de informação e a forma como se encontra disponível permite uma tomada de decisão correta, num determinado panorama.

O questionário está dividido da seguinte forma: agradecimento, objetivo, instruções, dados demográficos, caracterização do respondente, leitura de quadros e análise da solução edificada.

Quanto à forma, o questionário compreende questões de resposta fechada, nomeadamente respostas de escolha múltipla, de modo a evitar ambiguidade.

Quanto aos inquiridos, o presente questionário tem como público-alvo: Aspirantes do 5º ano da Escola Naval, Oficiais do Comando Naval ligados à fiscalização da pesca, Oficiais da DAGI, Oficiais de Patrulhas, Oficiais de Lanchas e Chefes de Serviço de Navegação das Corvetas.

A plataforma utilizada para elaboração dos questionários foi *Google Forms*. Essa ferramenta para além de possibilitar a elaboração de questionários, permite a obtenção dos dados num ficheiro em Excel, o que simplifica a análise dos questionários. A aplicação do questionário foi realizada através de mail, onde consta o link concedido pela referida plataforma, o que possibilita um acesso direto ao questionário.

Em suma, trata-se de um questionário *on-line* constituído por um conjunto de perguntas. As respostas às referidas perguntas foram solicitadas a uma amostra não aleatória de indivíduos. Em termos de requisitos dos respondentes, foi estabelecido uma grande porção relativamente a indivíduos com conhecimentos ligados à fiscalização da

pesca (em unidades navais e em unidades em terra) e, numa menor fração, indivíduos sem esses saberes.

Para auxiliar a averiguar se o objetivo foi alcançado, este inquérito tem também como finalidade dar resposta às questões de investigação elaboradas. O Apêndice J foi elaborado para a visualização do questionário. Perante isto, sucede-se a análise dos resultados obtidos.

5.2. Análise dos Resultados Obtidos

O presente capítulo tem como finalidade analisar os dados obtidos, de forma a se retirar conclusões acerca do sistema edificado, nomeadamente no que diz respeito a limitações, qualidade e quantidade de informação disponível, qualidade dos diagramas apresentados, se atende às necessidades e quais as melhorias que devem ser implementadas.

O questionário foi dividido em quatro secções:

- A. Dados Demográficos – concebe a recolha de dados demográficos gerais relativamente ao género, ao estatuto militar e à categoria;
- B. Caracterização do Respondente – possibilita a aferição do tipo de conhecimento e influência que pode ter no âmbito da fiscalização da pesca;
- C. Leitura de Quadros – analisar a capacidade de análise e de tomada de decisão, através do sistema edificado;
- D. Análise da Solução Edificada – proporciona a validação deste tipo de sistema de informação de gestão para auxílio à tomada de decisão, quer em termos de informação, panorama, extensão deste sistema a outras medidas, entre outros.

No que concerne ao conjunto de questões realizadas, algumas permitem dar resposta às questões derivadas, levantadas no Capítulo Um:

- QD 1 – secção D, questões 6 e 7;
- QD 2 – secção D, questões 2 e 4;
- QD 3 – secção D, questão 8.

A recolha da amostra foi definida com base no Teorema do Limite Central, ou seja, se a dimensão (n) da amostra for suficientemente grande, em que n assume valores maiores ou iguais a 30, então a distribuição de amostragem pode ser aproximada de uma distribuição normal (Martins, 2006, p. 30). Perante isto, a amostra obtida teve um total de 45 inquiridos.

5.2.1. Características da Amostra

As características da amostra são disponibilizadas nas secções um e dois, através da análise de gráficos circulares, nos quais constam os respetivos quantitativos.

A primeira secção possibilita uma análise dos inquiridos em termos de categoria. Através do Gráfico 2, é possível observar que mais de metade dos inquiridos são oficiais, sendo que possuem uma ligação à fiscalização da pesca ou ao tratamento e gestão da informação, como é possível constatar pelo público-alvo.

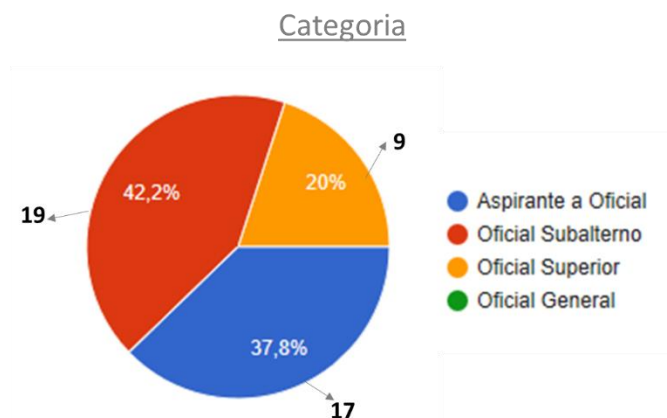


Gráfico 2 - Categoria dos inquiridos (Pergunta 3, Secção A)

A segunda secção permite observar o perfil do respondente, em termos de experiência e conhecimentos relativos à fiscalização da pesca. Através do Gráfico 3, verifica-se que mais de 50% dos inquiridos detém ligação com a medida de interesse.



Gráfico 3 - Ligação à atividade da fiscalização da pesca (Pergunta 1, Secção B)

A informação relativa ao contato da atividade da fiscalização da pesca, é completada através dos gráficos que se seguem na Figura 33. Estes gráficos possibilitam concluir que, mais de metade dos inquiridos estabelecem contato com esta área e também que possuem conhecimentos relativos à envolvente da medida em estudo (legislação, tipos de artes de pesca, tipos de embarcações de pesca, entre outros).

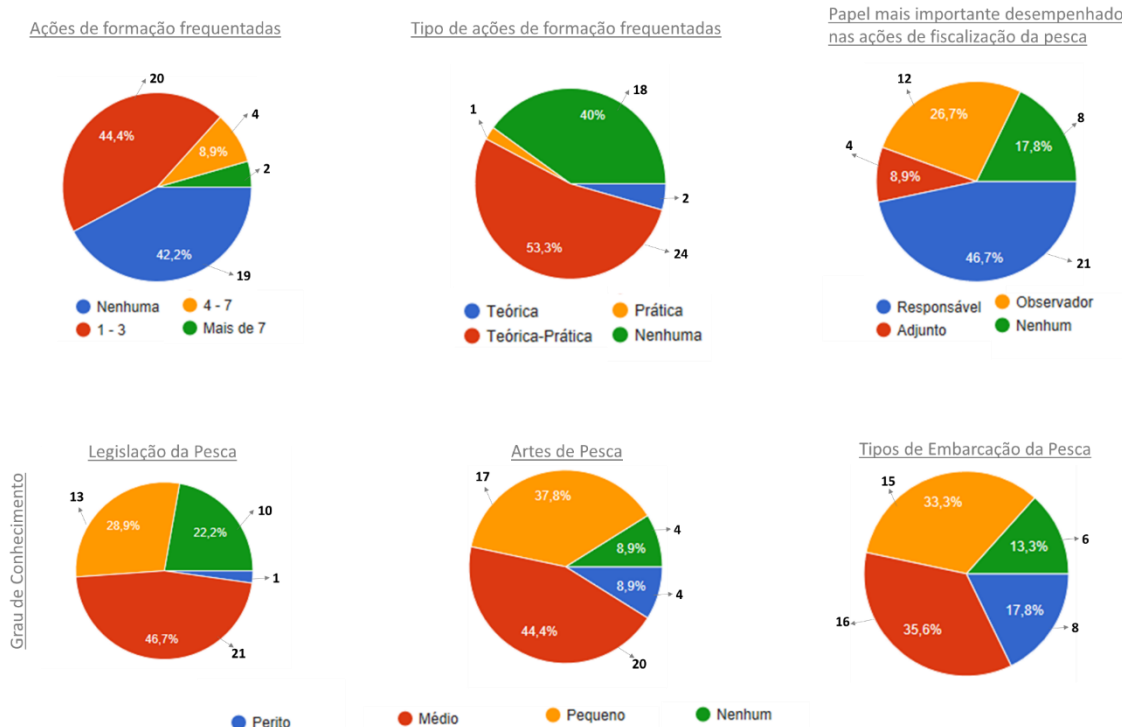


Figura 33 - Conjunto de gráficos complementares à ligação relativa à atividade da pesca (Perguntas 2 a 7, Secção B)

5.2.2. Análise detalhada dos resultados obtidos

A **secção 3** permite retirar conclusões relativamente aos diversos quadros apresentados, consoante a leitura que o inquirido fez.

Secção 3 – Conjunto de questões relativas à leitura de quadros.

Questão 1 – Em que meses, do ano corrente, se verificam o maior número de fiscalizações?

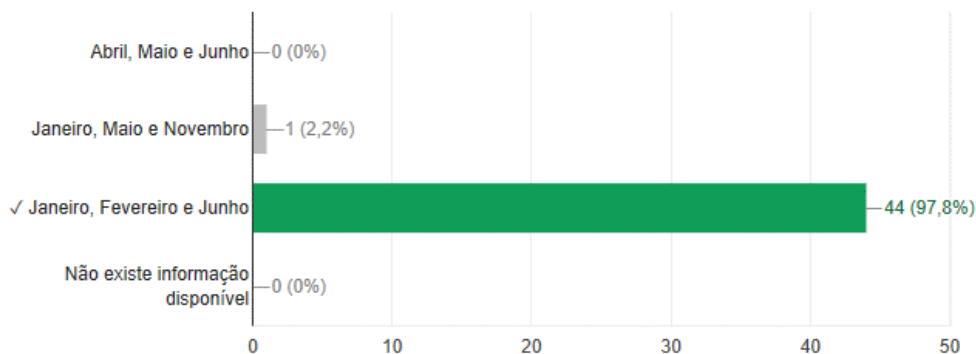


Gráfico 4 - Verificação dos meses com maior número de fiscalizações

Análise – Apenas uma pessoa examinou o quadro de forma errada, pelo que se pode concluir que o quadro apresenta as características suficientes para análise.

Questão 2 – Quais os tipos de embarcação mais vistoriados?

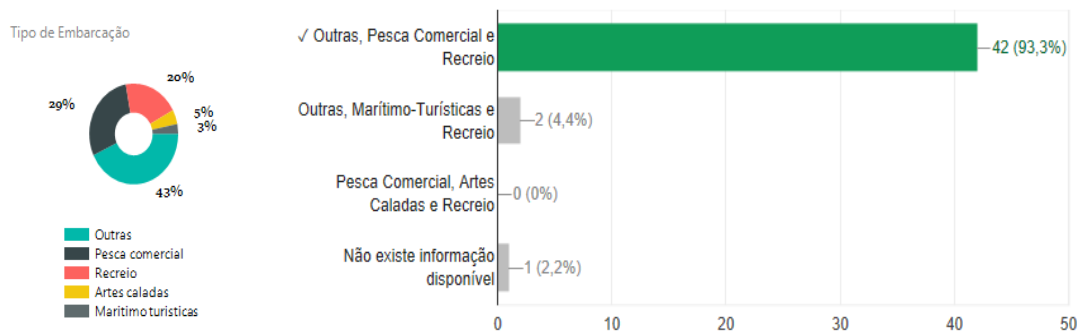


Figura 34 - Resultados da análise do gráfico referente aos tipos de embarcação

Análise – A resposta a ter em conta em termos de visualização do quadro, excluindo o conteúdo referente aos tipos de embarcação, foi a que teve a maior percentagem de respostas. No entanto, em termos de conteúdo e noção da informação, é necessário que se tenha a perceção de que a componente “outras” é constituída por partes menores que 43%. Ou seja, apesar da percentagem de “outras” ser a maior, leva a concluir que este elemento deverá ser colocado em separado deste gráfico, ou colocar-se o devido tipo de embarcação para clarificar esta grande percentagem. A nomenclatura está descrita na Tabela 1 da IONAV1010, e poderá ser essencial colocar todos os tipos de embarcação, em vez de “outras”. Por isso, considerar que não existe informação disponível para responder à pergunta é lógico, apesar de se conseguir observar, mesmo assim, que dos tipos de embarcação mais vistoriados são pesca comercial e recreio. É também importante salientar que se obtiveram duas respostas em “outras, marítimo-turísticas e recreio”, potencialmente pela dificuldade em distinguir as cores, nomeadamente, o cinzento escuro (pesca comercial) do cinzento claro (marítimo-turísticas).

Questão 3 – Por ordem crescente (da menor percentagem para a maior), quais as nacionalidades inspeccionadas?

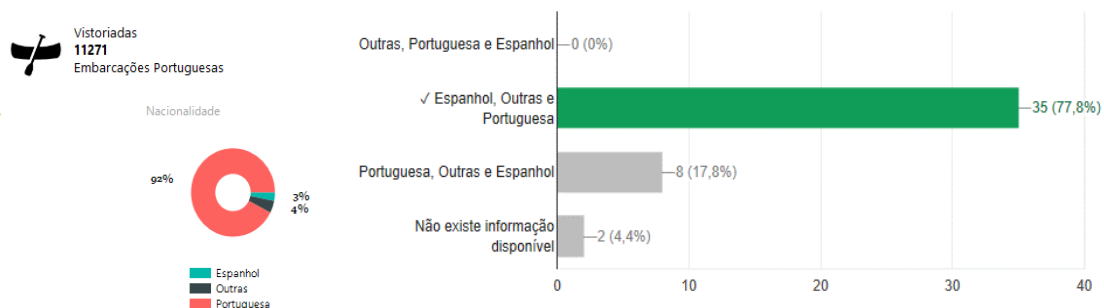


Figura 35 - Verificação das respostas no que diz respeito às nacionalidades inspeccionadas

Análise – A análise realizada na questão anterior revê-se nesta questão também. Isto porque apesar do gráfico transmitir que a resposta à pergunta é a que mais de metade

dos inquiridos concederam, surge novamente a incoerência de “outras”, uma vez que a ordem correta, analisando o conteúdo e não o que o gráfico demonstra, é: outras – espanhol – portuguesa (do menor para o maior). A conclusão que se tira é a mesma que na questão anterior, ou se faz um quadro separadamente com a discriminação das “outras”, ou uma espécie de alerta para que qualquer pessoa que observe o gráfico, consiga analisar de forma correta, mesmo não tendo noção do conteúdo. Neste caso, não se aplica uma alteração na IONAV1010, porque a Tabela 1 abrange todas as nacionalidades.

Questão 4 – Quantas embarcações de nacionalidade portuguesa foram vistoriadas?

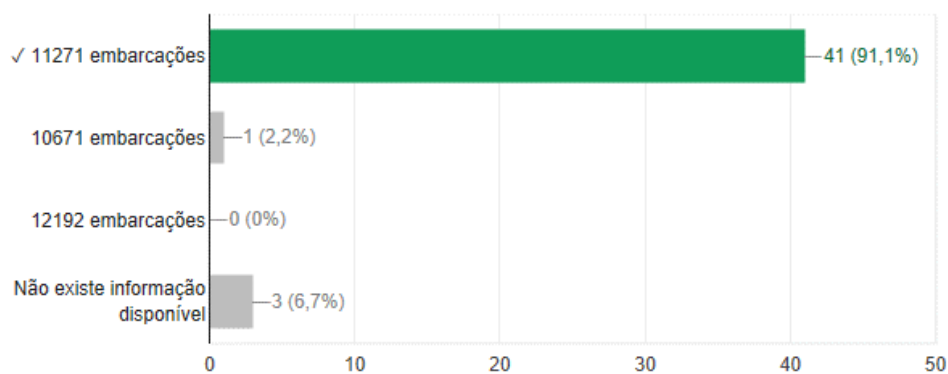


Figura 36 - Análise de respostas obtidas relativamente às embarcações portuguesas vistoriadas

Análise – Apesar de uma pequena parte considerar que não existe informação disponível para dar resposta à questão, mais de 90% dos inquiridos, deram a resposta que era possível observar no quadro, o que leva a concluir que é uma informação fácil de observar e de entender.

Questão 5 – Face à informação apresentada, que decisão tomaria no sentido de melhorar a eficácia da fiscalização?

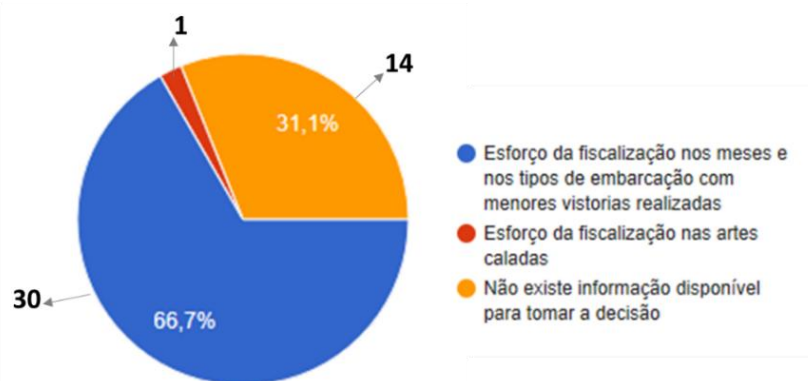


Gráfico 5 - Respostas obtidas relativamente à eficácia da fiscalização

Análise – Apesar da maior percentagem incidir sobre a resposta que se aproxima mais da decisão que prevê a melhoria da eficácia da fiscalização, é importante assinalar que há uma grande quantidade de inquiridos que revelam falta de informação para tomar uma decisão. Por isso, conclui-se que o sistema devia abranger mais dimensões de interesse, essencialmente para se ter mais bases que apoiem o processo de tomada de decisão.

Questão 6 – Quantos presumíveis infratores foram detetados ao longo do ano de 2017?

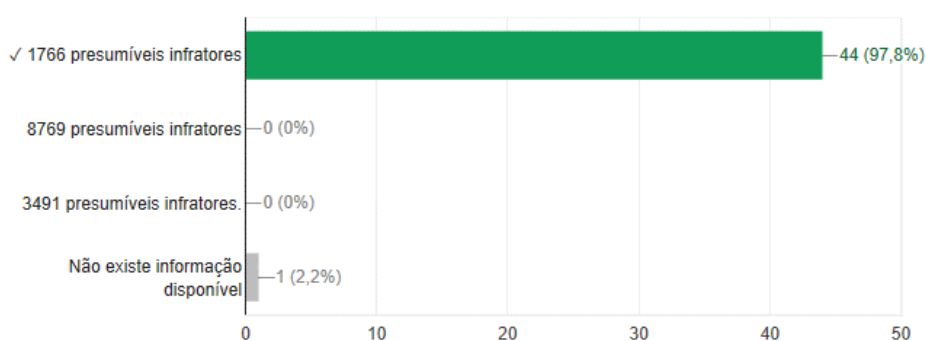


Gráfico 6 - Análise de respostas obtidas acerca das áreas que compreendem presumíveis infratores

Análise – A percentagem obtida por 44 inquiridos revela que a informação está clara e esclarecedora.

Questão 7 – No ano corrente (2017), em que áreas se verifica um maior número de presumíveis infratores?

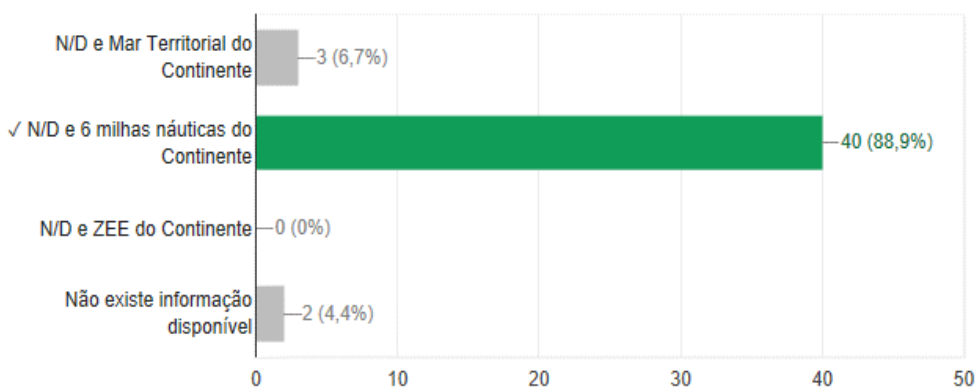


Gráfico 7 - Respostas obtidas referentes aos quantitativos dos presumíveis infratores

Análise – O conjunto de respostas disponíveis revela que uma das áreas é “N/D – Não Definido”, e, no entanto, isso não deve ser considerado uma área. Isto revela falta de

informação, que apenas pode ser deportado através de um maior cuidado na elaboração das mensagens.

Questão 8 – Qual o resultado de maior ocorrência nas vistorias?

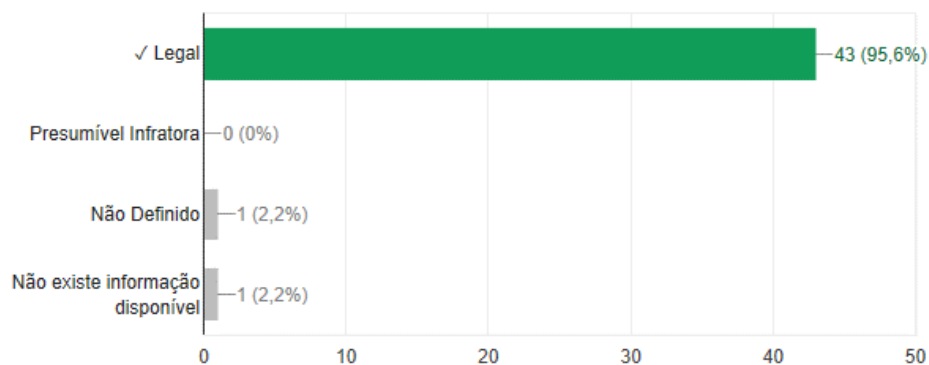


Gráfico 8 - Resultados obtidos no que concerne aos resultados das vistorias

Análise – A percentagem obtida permite verificar que o quadro possui informação suficiente para apurar o resultado mais aplicado nas vistorias. Relativamente ao inquirido que considera que não existe informação disponível, pode dever-se ao facto de um dos resultados disponíveis ser “Não Definido”, pelo que poderá ser pertinente a justificação da escolha deste resultado.

Questão 9 – Se tivesse de definir uma área a ser vistoriada, seleccionaria como critério o número de presumíveis infratores detetados numa determinada área?

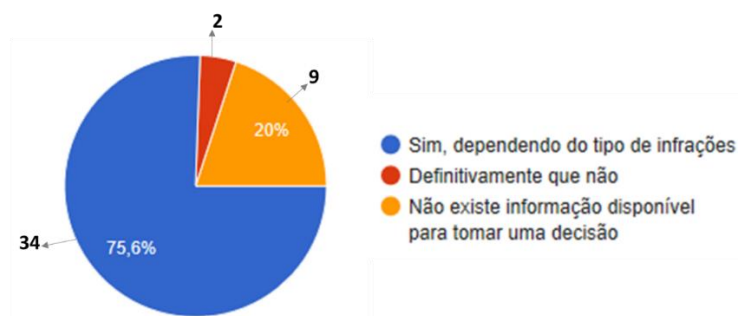


Gráfico 9 - Resultados obtidos numa tomada de decisão

Análise – A maior quantidade de respostas revela que é importante o acrescento de um quadro com o tipo e quantidade de infrações.

A **secção 4** diz respeito à opinião dos inquiridos relativamente ao sistema edificado, em termos de visualização e interpretação dos quadros, da qualidade e quantidade de informação, e da possível extensão deste sistema a outras medidas de

interesse. Deste modo, sucede-se uma análise dos resultados obtidos (através de diversas figuras) a fim de se avaliar o sistema.

A Figura 37 permite observar que a solução edificada é adequada para o planeamento e a análise no âmbito da fiscalização da pesca, e por sua vez, a informação disponibilizada é de fácil interpretação.

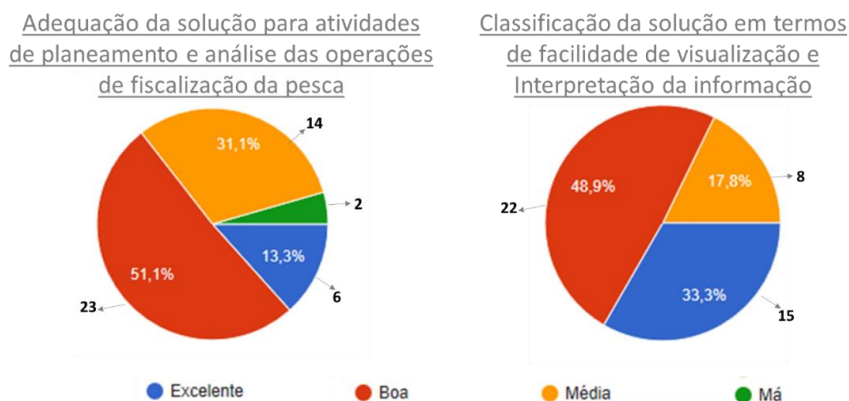


Figura 37 - Respostas obtidas referentes à solução edificada (Perguntas 1 e 2, Secção D)

No que concerne à organização do *dashboard* (Figura 38), apesar de permitir um adequado acesso à informação durante o processo de tomada de decisão (percentagem de 51,1% no segundo gráfico abaixo), observa-se que deve subsistir um maior cruzamento de informação, o que dá origem a algumas melhorias a realizar no sistema.

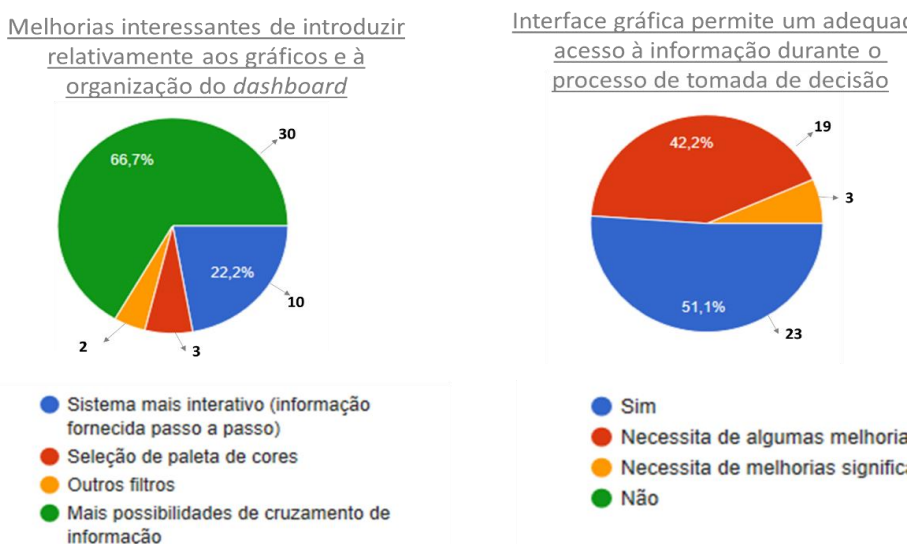


Figura 38 - Análise de questões relativas à organização do *dashboard* e respetiva interface gráfica (Perguntas 3 e 4, Secção D)

Em termos gerais do sistema edificado, através da Figura 39 é possível verificar que é um sistema que apresenta potencial no que concerne ao apoio à tomada de decisão, no âmbito das atividades de fiscalização da pesca, com limitações pouco significativas.

Isso vai de encontro à falta de algumas dimensões de análise. Essa referida carência leva a que uma das dificuldades encontradas tenha sido a tomada de decisão com a informação existente.

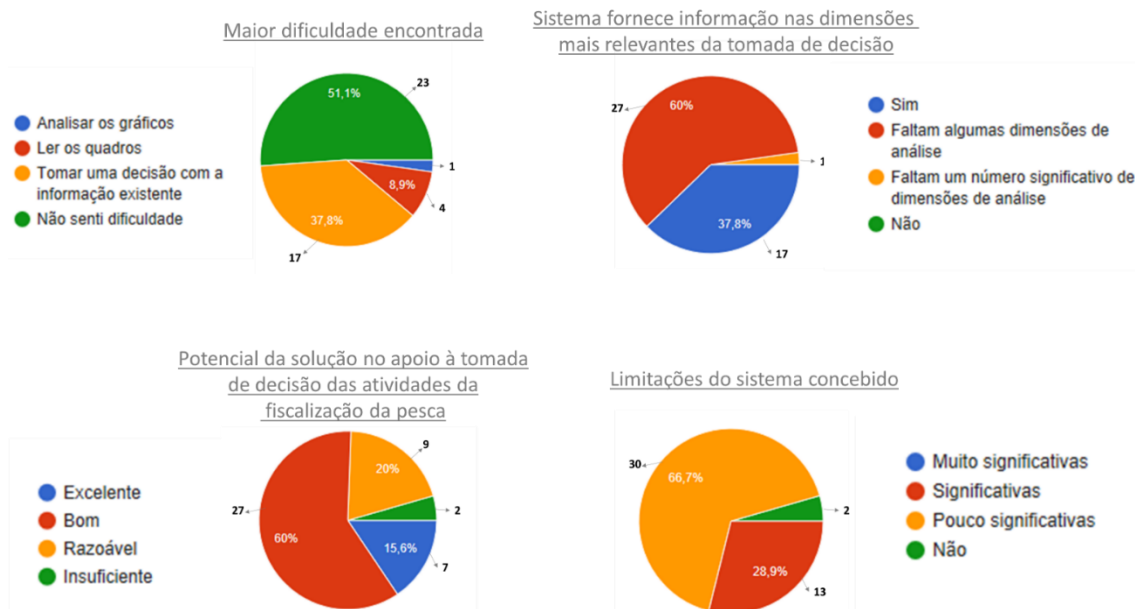


Figura 39 - Resultados obtidos no que diz respeito à envolvente do sistema (Perguntas 5,6,7 e 8, Secção D)

No que diz respeito à aplicação desta solução noutras áreas, que na presente dissertação algumas foram descritas como medidas de interesse, verifica-se que os inquiridos consideram relevante essa aplicação de modo que os dados recolhidos não tenham erros. Para além disso, em termos globais, a utilização de soluções automatizadas, tendo como instrumento de apoio à decisão um *dashboard*, é classificado como uma utilização boa (vide Figura 40).

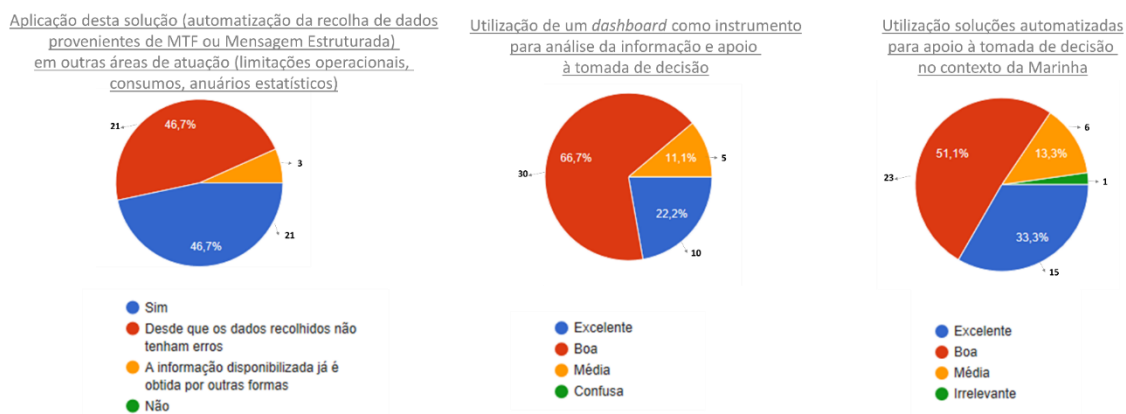


Figura 40 - Resultados obtidos acerca da aplicação de soluções automatizadas (Perguntas 9, 10 e 11, Secção D)

Em suma, a solução edificada apresenta algumas limitações nomeadamente em termos de cruzamento de informação e pela falta de algumas dimensões de análise, que

têm impacto no processo de tomada de decisão. Perante isto, surgem algumas melhorias que ao serem implantadas poderão reduzir as limitações verificadas:

- Cruzamento de informação:
 - Tipos de artes de pesca utilizadas com a área de operação;
 - Tipos de infração com o tipo de embarcação;
 - Tipos de infração e quantidade de presumíveis infratores com a área de operação.
- Dimensão de análise:
 - Áreas com menor esforço de fiscalização.
- Visualização:
 - Mapa/carta com as posições geográficas das ações de fiscalização, complementado com informação relativa a presumíveis infratores e respectivas infrações, em determinadas áreas.
- Redação de Mensagens:
 - Importante elaborar mensagens de forma cuidada e precisa, de modo a evitar áreas e resultados classificados como “Não Definido”, ou caso seja pertinente, uma justificação da razão dessa escolha.

5.2.3. Análise da eficácia do sistema edificado

A análise dos resultados obtidos, para além de permitir retirar conclusões muito relevantes acerca do sistema, pretende dar resposta à questão principal e às questões derivadas formuladas no Capítulo Um.

Primeiramente, é necessário avaliar se as respostas dos inquiridos estavam sujeitas à leitura que foi realizada, através da análise estatística. Para isso, foram formuladas hipóteses⁷³, sendo que a resposta dos inquiridos é representada pela variável ordinal X, que é constituída por valores da escala de Likert (valores de 1 a 4). Essa escala permite avaliar o grau de conformidade e o quanto o inquirido está satisfeito com a solução apresentada, neste caso. Os valores da escala traduzem-se em vários níveis (por exemplo concordância, frequência, satisfação) que dependem do tipo de respostas.

De modo a se alcançar uma resposta às questões derivadas, estabeleceu-se uma ligação entre estas e as questões do inquérito, sendo que a cada uma das questões está associado um grau, que servirá de base para o teste de hipóteses (vide Tabela 11).

⁷³ As hipóteses definem-se em H_0 e em H_1 , hipótese nula e hipótese alternativa (a que se pretende validar), respetivamente.

Tabela 11 - Ligação entre questões derivadas e questões do questionário, com o respetivo grau

QD	Questões Inquérito	Grau (X)
QD1	Q6 - O sistema fornece informação nas dimensões mais relevantes da tomada de decisão?	4- Sim 3- Faltam algumas dimensões de análise 2- Faltam um número significativo de dimensões de análise 1- Não
	Q7 - Como classifica o potencial da solução apresentada, no apoio à tomada de decisão das atividades de fiscalização da pesca?	4- Excelente 3- Bom 2- Razoável 1- Insuficiente
QD2	Q2 - Como classifica esta solução, em termos da facilidade de visualização e interpretação da informação?	4- Excelente 3- Boa 2- Média 1- Má
	Q4 - A interface gráfica do sistema permite um adequado acesso à informação durante o processo de tomada de decisão?	4- Sim 3- Necessita de algumas melhorias 2- Necessita de melhorias significativas 1- Não
QD3	Q8 - Considera que o sistema concebido tem limitações?	4- Não 3- Pouco significativas 2- Significativas 1- Muito significativas
Considerar: QD - Questão Derivada; Q - Questão		

De seguida, sucedem-se os testes de hipóteses unidirecionais para um nível de significância *alfa* de 0.05 (probabilidade de erro de 5%). A cada um desses testes está associada uma variável ordinal X, que foi definida tendo em conta o rigor que se pretende atingir, ou seja, pretende-se testar se o sistema preenche os requisitos sem qualquer lacuna ou limitação. A análise dos testes de hipóteses foi feita através da Tabela 12, que possui o quantitativo de respostas pela parte dos inquiridos, a respetiva percentagem, e os valores que nos permitem dar resposta (*valor t* e *t crítico*). Estes referidos valores foram obtidos através do Teste t^{74} , com variância desconhecida (uma vez que se desconhece a variância da população). Os testes possuem uma região crítica, que viabiliza rejeitar a hipótese nula. Perante isto, a hipótese nula é rejeitada quando se encontra na região crítica, ou seja, quando o *valor t* assume valores superiores ou iguais a *t crítico*.

QD1:

- H_0 : O sistema de informação de gestão concebido e implementado, **não fornece** informação necessária à tomada de decisão (Q6, Q7: $X \leq 3$)
- H_1 : O sistema de informação de gestão concebido e implementado, **fornece** informação necessária à tomada de decisão (Q6, Q7: $X > 3$)

Análise: No que diz respeito aos valores da questão 6, como o *valor t* é superior ao *t crítico* ($4.509 > 1.680$) e *p-value* é inferior ao valor *alfa* ($0.000023 < 0.05$), é suficiente para apurar que a hipótese nula pode ser rejeitada. No entanto, os valores obtidos na

⁷⁴ Atendendo à dimensão da amostra (em que *n* assume valores acima de 30), o teste t aproxima-se ao teste z, adotando uma distribuição normal.

questão 7 demonstram o contrário. Logo, os valores obtidos respeitantes às questões 6 e 7, não permitem retirar conclusões de que o sistema de informação de gestão concebido fornece informação necessária à tomada de decisão.

QD2:

- H_0 : A organização, visualização e interpretação da informação contida no *dashboard* **não é a adequada** durante o processo de tomada de decisão (Q2: $X < 3$; Q4: $X \leq 3$);
- H_1 : A organização, visualização e interpretação da informação contida no *dashboard* **é a adequada** durante o processo de tomada de decisão (Q2: $X \geq 3$; Q4: $X > 3$)

Análise: Relativamente à questão 2, uma vez que o *valor t* é inferior ao *t crítico* ($1.478 < 1.680$), não é possível rejeitar H_0 . Por seu turno, na questão 4, o *valor t* é superior ao *t crítico* ($4.780 > 1.680$) e *p-value* é inferior ao valor alfa, existem condições para rejeitar H_0 . No entanto, tendo em conta as diferenças que existem entre ambas as questões, não é possível concluir que a organização, visualização e interpretação da informação contida no *dashboard* é a adequada durante o processo de tomada de decisão.

QD3:

- H_0 : O sistema de informação de gestão **possui limitações** que afetem o potencial do sistema no apoio à tomada de decisão das atividades de fiscalização da pesca (Q8: $X \leq 3$)
- H_1 : O sistema de informação de gestão **não possui limitações** que afetem o potencial do sistema no apoio à tomada de decisão das atividades de fiscalização da pesca (Q8: $X > 3$)

Análise: Os valores obtidos para este teste de hipóteses revelam que não é possível rejeitar a hipótese nula, uma vez que o *valor t* é inferior ao *t crítico* ($-3.100 < 1.680$), logo o sistema de informação de gestão possui limitações que afetam o potencial do sistema no apoio à tomada de decisão, no âmbito da fiscalização da pesca.

Tabela 12 - Análise das respostas dos inquiridos em função das questões derivadas

	Questão Derivada 1				Questão Derivada 2				Questão Derivada 3	
	Questão 6		Questão 7		Questão 2		Questão 4		Questão 8	
	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
	17	37.8	7	15.6	15	33.3	23	51.1	2	4.4
	27	60	27	60	22	48.9	19	42.2	30	66.7
	1	2.2	9	20	8	17.8	3	6.7	13	28.9
	0	0	2	4.4	0	0	0	0	0	0
Total	45	100	45	100	45	100	45	100	45	100
Média	3.355555556		2.866666667		3.155555556		3.444444444		2.755555556	
Desvio padrão	0.528959337		0.726135474		0.705676837		0.623609564		0.528959337	
Erro médio amostral	0.078852602		0.108245885		0.105196092		0.092962225		0.078852602	
valor <i>t</i>	4.509116317		-1.231763524		1.478719911		4.780914437		-3.100017468	
<i>t crítico</i>	1.680229977		1.680229977		1.680229977		1.680229977		1.680229977	
<i>p-value</i>	0.000023925		0.112290012		0.073169535		0.000009904		0.001684945	
alfa	0.05									

A análise das questões derivadas permite observar que não existe uma concordância absoluta, ou seja, que o sistema é excelente. Por isso, o sistema concebido apesar de apresentar informação útil e de interesse, ainda carece de alguns dos requisitos

necessários para o apoio à tomada de decisão. Isso revela que a análise estatística vai de encontro às conclusões que se retiraram na análise detalhada de cada gráfico, feita na secção anterior. Para além disso, essas referidas análises constituem uma forma de dar resposta à questão principal:

Como efetuar a criação de um sistema de informação de gestão, que permita a tomada de decisão, nas várias áreas de atuação da Marinha?

Um sistema de informação de gestão que permita a visualização da informação que o utilizador necessita é uma mais-valia, em qualquer organização. No entanto, uma das práticas que se deve ter em conta antes da sua conceção é o levantamento dos requisitos necessários, nomeadamente do tipo de informação e, conseqüentemente do tipo de visualização que melhor se poderá adequar (dependendo do assunto). Isto porque o levantamento de requisitos deve ser a primeira etapa a realizar para a conceção de um sistema que pretende auxiliar a organização, através da sua influência no processo de tomada de decisão.

A tomada de decisão é com base na informação que o sistema apresenta. Perante isto, é impreterível que os dados recolhidos para serem tratados e transformados em informação, sejam confiáveis e precisos. De modo que, toda a organização, independentemente do nível hierárquico, deve ser sabedora de que os dados são a base de todo o processo e, como tal, devem ser minuciosos.

A Marinha é composta por diversas unidades, de entre as quais circula um grande volume dados. É essencialmente devido a este referido volume que é tão importante a criação de sistemas destes nas várias áreas de atuação desta organização.

Na presente dissertação, o sistema concebido foi no âmbito da fiscalização da pesca, verificando-se de que necessita de algumas melhorias. E, por isso, é um sistema que está sujeito a várias iterações, até se alcançar o panorama que preencha os requisitos necessários à tomada de decisão.

Um sistema de informação de gestão possui características que podem fazer um grande avanço na organização e, como tal, a criação de um sistema destes na Marinha, entidade que mantém um papel ativo e predominante, é um desafio que deve ser realizado.

Conclusões e Recomendações

Conclusões

A presente dissertação permitiu obter conhecimento muito variado, por ter abordado vários assuntos e conceitos, nomeadamente nas áreas do *Business Intelligence*, da fiscalização da pesca, das mensagens de texto formatado e das mensagens estruturadas referentes às várias áreas de atuação da Marinha.

O começo da presente dissertação surge com o estado da arte, que se baseou na investigação realizada para perceber como surgiu o BI, em que consiste, quais os seus componentes, quais as ferramentas que predominam no mercado, e, como finalidade principal obter os conceitos e procedimentos a realizar para edificar um sistema BI.

O sistema BI foi implementado no âmbito da fiscalização da pesca e, por isso, sucedeu-se um estudo acerca desta medida de interesse. Esta medida permitiu obter uma visão mais clarificada dos conceitos, das competências, da legislação, das infrações e dos sistemas implementados na Marinha.

O volume de dados que circula entre as unidades da Marinha é amplo, e como tal, surgiu a necessidade de analisar um dos meios mais utilizados para difusão de informação – as mensagens. Essa análise, aplicada em várias medidas de interesse, permitiu adquirir um conhecimento detalhado acerca das várias mensagens, das diferentes estruturas, dos campos e dos dados que as compõem. As mensagens estruturadas não são vantajosas para o tratamento automático dos dados e, como tal, foram levantadas algumas recomendações e trabalhos futuros que vão de encontro à transformação dessas mensagens.

O objetivo da presente dissertação foi alcançado, nomeadamente a conceção de um sistema de informação de gestão que contribui para que as unidades da Marinha obtenham uma visão unificada da informação relevante associada ao processo de tomada de decisão. Através da análise de dados, foi possível concluir que o sistema necessita de algumas melhorias (referidas na presente dissertação) para que possa atender aos requisitos dos utilizadores e, essencialmente, para que possa ser uma ferramenta de apoio à tomada de decisão.

Em suma, o presente trabalho permitiu concluir que um sistema de informação de gestão apresenta distintamente benefícios e que pode ter uma boa influência no processo de tomada de decisão, tendo em conta todos os aspetos que foram referidos para a sua devida conceção. Para além disso, é uma ferramenta que deve ser implementada, a longo prazo, nas restantes áreas de atuação da Marinha Portuguesa.

Recomendações e Trabalhos Futuros

Na presente dissertação, foi implementado um artefacto cuja prova de conceito foi referente à fiscalização da pesca. Tendo em conta todo o processo efetuado e os resultados obtidos, a extensão deste sistema nas outras medidas de interesse apresentadas poderá ser benéfica. Nas restantes medidas de interesse, foram analisadas as mensagens e os respetivos dados, de onde se retiraram conclusões importantes que poderão fazer a diferença ao serem realizadas.

As propostas para trabalhos futuros são as seguintes:

- Integração de alterações e efetuar na IONAV1010:
 - Transformação de mensagens estruturadas, presentes na IONAV1010, em MTFs, para subsistir apenas uma estrutura de mensagens cuja vantagem é o tratamento automático dos dados;
 - Colocar o comunicado de consumo de munições na IONAV1010, tendo em conta que neste momento não existe nenhum comunicado nessa publicação que aborde apenas o consumo de munições;
 - Substituição do comunicado “OPSTAT DEFECT” pelo “NAVOPDEF”, por este já ser uma MTF;
- Aperfeiçoar o sistema implementado, aquando da mensagem analisada (FISCREP) estiver no formato de MTF, para retirar toda a informação e principalmente no que diz respeito ao tipo e quantidade de pescado, tendo em conta que foram dados inseridos de forma aleatória;
- Análise de outras possíveis medidas de interesse, tais como atividade operacional e atividades de treino e aprontamento;
- Integração das medidas de interesse analisadas num sistema idêntico ao edificado.

Bibliografia

- AGARWAL, Ritu, & DHAR, Vasant (2014), "Editorial – Big Data, Data Science, and Analytics: The Opportunity and Challenge for IS Research", *Information System Research*, vol. 25, n.º 3, pp. 443–448, <http://dx.doi.org/10.1287/isre.2014.0546>, acessado em 26 de dezembro de 2017.
- AKITA, Fabio (2006), *Repensando a Web com Rails*, 1ª edição, Rio de Janeiro, Brasport Livros e Multimídia Ltda.
- ALI, Shaheb., KHAN, Shahadat, & MIAH, Shah. J. (2017), "Analysis of Interaction between Business Intelligence and SMEs: Learn from Each Other", *Journal of Information Systems and Technology Management*, vol. 14, n.º 2, pp. 151–168. doi: 10.4301/S1807-17752017000200002
- ALVARO, Troche Clavijo (2014), "Aplicacion de La Minería de Datos sobre Bases de Datos Transaccionales - Application of Data Mining on Transactional Data Bases", *Fides et Radio - Revista de Difusión cultural y científica de la Universidad La Salle*, vol. 7, pp. 58–66, http://www.scielo.org.bo/scielo.php?script=sci_arttext&pid=S2071-081X2014000100005&lang=pt, acessado em 26 de dezembro de 2017.
- APPUSWAMY, Raja *et al.* (2017), "Analyzing the Impact of System Architecture on the Scalability of OLTP Engines for High-Contention Workloads", *Very Large Data Bases*, vol. 11, n.º 2, pp. 121–134. doi: 10.14778/3149193.3149194
- ARAÚJO, Erika Maria Teixeira, BATISTA, Mônica de Lourdes Souza, & MOREIRA DE MAGALHÃES, Teresinha (2007), *OLAP: Características, Arquitetura e Ferramentas*, Instituto Vianna Júnior.
- AUAD, Arnaldo. (s.d.), *Conceitos de Business Intelligence - Guia Definitivo*, Direção e Sentido Coaching e Treinamentos.
- BALLARD, Chuck, FARRELL, Daniel M., GUPTA, Amit, MAZUELA, Carlos, & VOHNIK, Stanislav (2006), *Dimensional Modeling: In a Business Dimensional Intelligence Environment*, <https://www.redbooks.ibm.com/redbooks/pdfs/sg247138.pdf>, acessado em 20 de dezembro de 2017.
- BALTZAN, Paige, & PHILLIPS, Amy (2009), *Essentials of Business Driven Information Systems*, 2ª edição, Nova Iorque, McGraw-Hill Irwin.
- BARREIROS DOS SANTOS, Lúcio Agostinho *et al.* (2016), *Orientações Metodológicas para a Elaboração de Trabalhos de Investigação*, Lisboa, Instituto

Universitário Militar.

- BAX, Marcello Peixoto (2013), "Design Science: filosofia da pesquisa em ciência da informação e tecnologia", *Ciência da Informação*, vol. 42, n.º 2, pp. 298–312.
- BOG, Anja (2014), *Benchmarking Transaction and Analytical Processing Systems: The Creation of a Mixed Workload Benchmark and its Application*, Alemanha, Springer Science & Business Media.
- BOULEKROUCHE, Boubaker, JABEUR, Nafaâ., & ALIMAZIGHI, Zaia (2015), "An Intelligent ETL Grid-Based Solution to Enable Spatial Data Warehouse Deployment in Cyber Physical System Context", *Procedia Computer Science*, vol. 56, n.º 1, pp. 111–118. doi: 10.1016/j.procs.2015.07.176.
- BOTELHO, Fernando Rigo, & FILHO, Edelvino Razzolini (2014), "Conceituando o termo Business Intelligence: origem e principais objetivos", *Sistemas, Cibernética e Informática*, vol. 11, n.º 11, pp. 55–60.
- BRAGHITTONI, Ronaldo (2017), *Business Intelligence: Implementar do jeito certo e a custo zero*, Casa do Código.
- CALDAS, Marcos Paulo Kohler, SCANDELARI, Luciano, & KOVALESKI, João Luiz (2006), "Aplicações sobre uma Data Warehouse no ambiente das organizações e suas vantagens", Em *XII Simpósio de Engenharia de Produto*.
- CARVALHO, João Álvaro (2010), "Tecnologias e Sistemas de Informação: uma área científica orientada às necessidades de conhecimento dos profissionais envolvidos na contínua transformação das organizações através das tecnologias da informação", *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, pp. 1–25.
- CERQUEIRA, Fábio Vergara, & TEIXEIRA, Rute (2015), "Os visitantes do Santuário de Panóias: um estudo patrimonial de um sítio arqueológico português", *Revista de História e Geografia-ágora*, vol. 17, n.º 2, pp. 96–109, <https://online.unisc.br/seer/index.php/agora/article/view/6690/4845>, acessado em 12 de fevereiro de 2018.
- CODY, W. F. *et al.* (2002), "The integration of business intelligence and knowledge management", *IBM Systems Journal*, vol. 41, n.º 4, pp. 697–713. doi: 10.1147/sj.414.0697.
- COELHO, Danilo *et al.* (2016), "Towards of a Business Intelligence Platform to Portuguese Misericórdias", *Procedia Computer Science*, vol. 100, pp. 762–767. doi: 10.1016/j.procs.2016.09.222.

- CUNHA, Juscelino Cássio Vieira, Business intelligence: conceitos, técnicas, sistemas e ferramentas, Pós-graduação em Banco de Dados com ênfase em BI, Institut Brasileiro de Tecnologia, Brasil, 2008.
- DALFOVO, Oscar, & TAMBORLIN, Norberto (2017), "Business Intelligence – Estudos e casos: Gestão de Tecnologia da Informação como Inteligência nos Negócios", 1.ª edição, Blumenau, *Clube de Autores*.
- DENILSON DA SILVA *et al.* (2016), "Business Intelligence", *Revista Maiêutica*, vol. 1, n.º 1, pp. 73–90, Obtido em 25 de novembro de 2017, de <https://doi.org/10.1007/978-3-642-36318-4>.
- DEVENS, Richard Millar (1865), *Cyclopaedia of Commercial and Business Anecdotes*, Nova Iorque, D. Appleton.
- DGAM (2011), *A fiscalização da pesca*, [www.amn.pt/DGAM/Documents/Fiscalização da pesca.pdf](http://www.amn.pt/DGAM/Documents/Fiscalização%20da%20pesca.pdf), acessado em fevereiro de 2018.
- DGRM (2018), *Política Comum das Pescas*, <https://www.dgrm.mm.gov.pt/web/guest/pesca-pp-politica-comum-das-pescas>, acessado em fevereiro de 2018.
- ECKERSON, Wayne W. (2002), *The Rise of Analytical Applications: Build or Buy?*, Obtido em 3 de dezembro de 2017, de http://download.101com.com/tdwi/research_report/0402aareport.pdf
- (2006) *Deploying Dashboards and Scorecards*, Obtido em 18 de dezembro de 2017, de <http://sophitech.mx/files/1113/7718/8544/TDWI-Best-Practices-Report-Deploying-Dashboards-and-Scorecards.pdf>
- EL-SAPPAGH, Shaker H. Ali, HENDAWI, Abdeltawab M. Ahmed, & EL BASTAWISSY, Ali Hamed (2011), "A proposed model for data warehouse ETL processes", *Journal of King Saud University – Computer and Information Sciences*, vol. 23, n.º 2, pp. 91–104. doi: 10.1016/j.jksuci.2011.05.005
- FAO (1995), *Code of Conduct for Responsible Fisheries: Food and Agriculture Organization of the United Nations*, www.fao.org/3/a-v9878e.pdf, acessado em fevereiro de 2018.
- FERNANDES, M. D. S., & KANTORSKI, G. Z. (2001), "Neuro BI : Uma Ferramenta Web open source para apoio à Tomada de Decisão", *Hifen, Uruguaiana*, vol. 32, n.º 62, pp. 181–188, Obtido em 15 de dezembro de 2017, de <http://revistaseletronicas.pucrs.br/ojs/index.php/hifen/article/view/4597/3484>.
- FERREIRA, João *et al* (2010), "O Processo ETL em Sistemas Data Warehouse",

- INForum 2010 - II Simpósio de Informática*, pp. 757–765, Obtido em 22 de novembro de 2017, de <http://inforum.org.pt/INForum2010/papers/sistemas-inteligentes/Paper080.pdf>.
- GILES, Lionel (2013), *Sun Tzu on The Art of War*, Abingdon, Routledge.
- GOLFARELLI, Matteo, RIZZI, Stefano, & CELLA, Iuris (2004), "Beyond Data Warehousing: What's Next in Business Intelligence?", Em *7th ACM International Workshop on Data Warehousing and OLAP*, pp. 1–6. doi: 10.1145/1031763.1031765.
- GOMES, José, & QUINTELA, Hélder (2016), "Uma abordagem arquitetural - B-Health Analytics", Em *16ª Conferência da Associação Portuguesa de Sistemas de Informação (CAPSI)*, vol. 16, pp. 300–307. doi: <https://dx.doi.org/10.18803/capsi.v16.300-307>.
- GOVERNO DE PORTUGAL (s.d.), *Estratégia Nacional para o MAR 2013-2020*, https://issuu.com/dgpm_portugal/docs/enm2013_2020, acessado em 24 de dezembro de 2017.
- GASQUE, Kelley Gonçalves Dias, & TESCAROLO, Ricardo (2004), "Sociedade da aprendizagem: informação, reflexão e ética", *Ciência da Informação*, vol. 33, n.º 3, pp. 35–40. Obtido em 21 de dezembro de 2017, de <http://www.scielo.br/pdf/ci/v33n3/a05v33n3.pdf>.
- GUARDA, Teresa *et al.* (2013), "Business Intelligence as a Competitive Advantage for SMEs", *International Journal of Trade, Economics and Finance*, vol. 4, n.º 4, pp. 187–190. doi: 10.7763/IJTEF.2013.V4.283.
- HAJ-BOLOURI, Amir (2015), "The Notion of Users in Design Science Research", Em *Proceedings of the 38th Information Systems Research Seminar in Scandinavia (IRIS 38)*, pp. 1–15, Obtido em 7 de dezembro de 2017, de <http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A853537&dswid=1522%5Cnhttps://www.diva-portal.org/smash/get/diva2:853537/FULLTEXT01.pdf>.
- HEVNER, Alan R. *et al.* (2004), "Design Science in Information Systems Research", *MIS Quarterly*, vol. 28, n.º 1, pp. 75–105. doi: 10.2307/25148625.
- INE (2017), *Estatísticas da Pesca 2016*, Lisboa, Portugal.
- INMON, W. H. (2002), *Building the Data Warehouse*, 3ª edição, Nova Iorque, John Wiley & Sons.
- INMON, W. H., & KRISHNAN, Krish (2011), *Building the Unstructured Data Warehouse: Architecture, Analysis, and Design*, 1ª edição, New Jersey, Technics

Publications.

- IVAN, Mihaela-Laura (2014), "Characteristics of In-Memory Business Intelligence", *Informatica Economică*, vol. 18, n.º 3, pp 17–25. doi: 10.12948/issn14531305/18.3.2014.02.
- KIMBALL, Ralph, & CASERTA, Joe (2004), *The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data*, Indianapolis, Wiley Publishing.
- KIMBALL, Ralph, & ROSS, Margy (2002), *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*, 2ª edição, Nova Iorque, John Wiley & Sons.
- (2013) *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*, 3ª edição, Indianapolis, John Wiley & Sons.
- KOROBKO, A. V., & PENKOVA, T. G. (2012), "On-line analytical processing based on formal concept analysis", *Procedia Computer Science*, vol. 1, pp. 2311–2317. doi: 10.1016/j.procs.2010.04.259.
- LACERDA, Daniel Pacheco, DRESCH, Aline, PROENÇA, Adriano & JÚNIOR, José (2013), "Design Science Research: Método de pesquisa para a engenharia de produção", *Gestão & Produção*, vol. 20, n.º 4, pp. 741-761. doi: 10.1590/S0104-530X2013005000014.
- LARSON, Brian (2009), *Delivering Business Intelligence with Microsoft SQL Server 2008*, Estados Unidos, Mc Graw-Hill Companies.
- LAUDON, Kenneth C. & LAUDON, Jane P. (2011), *Management Information Systems: Managing the Digital Firm*, 12ª edição, Nova Jersey, Prentice Hall.
- LOH, Stanley (2014), *BI na Era do Big Data para Cientistas de Dados - indo além de cubos e dashboards na busca pelos porquês, explicações e padrões*, 1ª edição, Porto Alegre.
- LOPES, Jorge (2006), *O Fazer Do Trabalho Científico Em Ciências Sociais Aplicadas*, Recife, Editora Universitária.
- LOPES, Saldanha (2011), *A Marinha ao Serviço de Portugal*, www.marinhasplp.org/PT/asmarinhas/doutrinas/Documents/A%20Marinha%20ao%20Serviço%20de%20Portugal_23MAR.pdf, acessado em 23 de dezembro de 2017.
- LUCAS, Alexandre, CAFÉ, Ligia Maria Arruda, & VIERA, Angel Freddy Godoy

- (2016), "Inteligência de negócios e inteligência competitiva na ciência da informação brasileira: contribuições para uma análise terminológica", *Perspectivas em Ciência da Informação*, vol. 21, n.º 2, pp. 168–187. doi: 10.1590/1981-5344/2568.
- LUHN, H. P. (1958), "A Business Intelligence System", *IBM Journal of Research and Development*, pp. 314–319. doi: 10.1147/rd.24.0314
- MAGALHÃES, Alberto (2015), *SQL Server 2014 - Curso Completo*, 1ª edição, Lisboa, FCA - Editora de Informática.
- MARINHA (2002), *IONAV 1010 - Relatos e Comunicados Operacionais*, Oeiras, Comando Naval.
- (2005), *PDA 2 - Glossário de Sistemas e Tecnologias de Informação e Comunicação (GlosSTIC)*, Lisboa, Estado-Maior da Armada.
- (2017), *Diretiva Setorial da Marinha*, http://www.marinha.pt/conteudos_externos/dpm2017/files/basic-html/page25.html, acessado em novembro de 2017.
- (2018), *A Missão*, <https://www.marinha.pt/pt/a-marinha/Paginas/missao.aspx>, acessado em junho de 2018.
- (2018), *A Marinha é uma organização otimizada*, <https://www.marinha.pt/pt/a-marinha/Paginas/estrutura.aspx>, acessado em junho de 2018.
- (s.d.a), "A Marinha ao Serviço de Portugal", Marinha Portuguesa, pp. 1–16.
- (s.d.b), *Manual do Operador do Sistema de Apoio à Decisão para a Atividade de Patrulha (SADAP)*, Lisboa, Direção de Análise e Gestão da Informação.
- MARTINS, Maria Eugénia Graça (2006), *Introdução à Inferência Estatística*, <http://homepage.ufp.pt/cmanso/ALEA/introInfEstat.pdf>, acessado em junho de 2018.
- MINISTÉRIO DA AGRICULTURA E DO MAR (2007), *Plano Estratégico Nacional para a Pesca 2007 – 2013*, http://www.promar.gov.pt/Download/PROMAR/PEN_Pesca.pdf, acessado em fevereiro de 2018.
- MONTEIRO, Nuno Sardinha, & MOURINHA, António Anjinho (2011), "Marinha De Duplo Uso: Um Conceito Pós-Moderno De Utilização Do Poder Marítimo", *Anais do Clube Militar Naval*, vol. III, pp. 415–447, http://database.jornaldefesa.pt/politicas_de_defesa/portugal/Marinha%20de%20Du

pl%20Uso%20Um%20conceito%20pósmoderno%20de%20utilização%20do%20poder%20marítimo.pdf, acessado em fevereiro de 2018.

MOODY, Daniel L., & KORTINK, Mark A. R. (2000), "From Enterprise Models to Dimensional Models: A Methodology for Data Warehouse and Data Mart Design", Em *Proceedings of the International Workshop on Design and Management of Data Warehouses (DMDW'2000)*, vol.28, pp. 5-1-12, <http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-28/paper5.pdf>, acessado em 25 de dezembro de 2017.

— (2003) "From ER Models to Dimensional Models: Bridging the Gap between OLTP and OLAP Design, Part I", *Business Intelligence Journal*, pp. 7–24, https://www.researchgate.net/publication/238070575_From_ER_Models_to_Dimensional_Models_Bridging_the_Gap_between_OLTP_and_OLAP_Design_Part_I, acessado em 23 de dezembro de 2017.

MORESI, Eduardo Amadeu Dutra (2000), "Delineando o valor do sistema de informação de uma organização", *Ciência da Informação*, vol. 29, n.º 1, pp. 14–24. doi: 10.1590/S0100-19652000000100002.

MOTA, Fernando Manuel Silva, *Segurança Marítima - O Caso Nacional e Perspectivas de Futuro*, Trabalho de Investigação Individual do CPOS, Instituto de Estudos Superiores Militares, Lisboa, 2010.

MOURINHA, António Mateus Anjinho (2012), "Monitorização, Controlo e Fiscalização da Pesca", *Revista Científica Eletrónica - Maria Scientia*, edição n.º 2, pp. 83–122.

NATO (2015), *APP-11 – NATO Message Catalogue*, Edição D, Versão 1.

NEGASH, Solomon (2004), "Business intelligence", *Communications of the Association for Information Systems*, vol. 13, pp. 177–195. doi: 10.1007/978-3-540-48716-6_9.

NELSON, Greg (2007), "Introduction to the SAS 9 Business Intelligence Platform: A Tutorial", *ThotWave Technologies*, Carolina do Norte, pp. 1–12, <https://www.researchgate.net/publication/237750763>, acessado em novembro de 2017.

OLIVEIRA, Miguel, CARDOSO, Elsa, & SANTANA, Marina (2015), "Medição de Usabilidade e Eficiência de Dashboards: Framework Desenvolvida usando os Princípios de Desenho de Business Intelligence e Interação Homem-Máquina", Em *15ª Conferência da Associação Portuguesa de Sistemas de Informação (CAPSI)*, pp. 338–362. doi: <http://dx.doi.org/10.18803/capsi.v15.338-362>

- PARRA, Octavio J. Salcedo, GALEANO, Rita Milena, & B., Luis G. Rodriguez (2010), "Metodología crisp para la implementación Data Warehouse", *Revista Tecnura*, vol. 14, n.º 26, pp. 35–48, <http://www.scielo.org.co/pdf/tecn/v14n26/v14n26a05.pdf>, acedido em 3 de dezembro de 2017.
- PASCOAL, Sérgio *et al.* (2009), "Using Business Intelligence techniques to increase the safety of citizens – The Tilburg case", *Revista do Departamento de Inovação, Ciência e Tecnologia da Universidade Portucalense*, pp. 1–13, <http://hdl.handle.net/11328/452>, acedido em 30 de novembro de 2017.
- PINTO, Raquel do Carmo Monteiro, *Análise de dados da fiscalização da pesca – Estudo da evolução e exploração dos recursos piscícolas no ecossistema marinho português*, Dissertação de Mestrado em Ciências Militares Navais - especialidade de Marinha, Escola Naval, Marinha Portuguesa, 2017.
- POWELL, Gavin (2005), *Oracle Data Warehouse Tuning for 10g*, Oxford, Elsevier Digital Press.
- POWER, Daniel J. (2006), *A Brief History of Decision Support Systems*, <http://dssresources.com/history/dsshhistory.html>, acedido em 22 de novembro de 2017.
- RAINARDI, Vicent (2008), *Building a Data Warehouse: With Examples in SQL Server*, 1ª edição, Estados Unidos da América, Apress.
- ROWLEY, Jennifer (2007), "The wisdom hierarchy: representations of the DIKW hierarchy", *Journal of Information Science*, vol. 33, n.º 2, pp. 163–180. doi: <https://doi.org/10.1177/0165551506070706>
- SANDU, Daniela Ioana (2010), "Operational and real-time Business Intelligence", *Informatica Economică*, vol. 47, n.º 3, pp. 33–36.
- SANTOS, Maribel Yasmina., & RAMOS, Isabel (2009), *Business Intelligence: Tecnologias da Informação na Gestão de Conhecimento*, 2ª edição, Lisboa, FCA - Editora de Informática.
- SCHEPS, Swain (2008), *Business Intelligence For Dummies*, 1ª edição, Indianapolis, Wiley Publishing.
- SCHWABER, Ken, & SUTHERLAND, Jeff (2013), *Guia do SCRUM*, Boston, Harvard Business Review, https://www.scrum.org/Portals/0/Documents/Scrum Guides/Scrum_Guide.pdf, acedido em 19 de dezembro de 2017.
- SEZÕES, Carlos, OLIVEIRA, José, & BAPTISTA, Miguel (2006), *Business*

Intelligence, Porto, Sociedade Portuguesa de Inovação,
<https://doi.org/10.1002/9780470753866>, acessado em 5 de novembro de 2017.

SHEN, Liangzhong, *et al.* (2012), "The Application Research of OLAP in Police Intelligence Decision System", *Procedia Engineering*, vol. 29, pp. 397–402. doi: <https://doi.org/10.1016/j.proeng.2011.12.730>

SINGH, Hardeep, & SINGH, Bikram Pal (2013), "Business Intelligence: Effective machine learning for business administration", *International Journal of IT, Engineering and Applied Sciences Research (IJIEASR)*, vol. 2, n.º 1, pp. 13–19.

SVERRISDOTTIR, Hrafnhildur Sif, INGASON, Helgi Thor, & JONASSON, Haukur Ingi (2014), "The role of the product owner in scrum-comparison between theory and practices", *Procedia - Social and Behavioral Sciences*, vol. 119, pp. 257–267. doi: 10.1016/j.sbspro.2014.03.030

TEOREY, Toby *et al.* (2013), *Projeto e Modelagem de Banco de Dados: Tradução da 5ª Edição*, 2ª edição, Rio de Janeiro, Elsevier.

TERESO, Marco, & BERNARDINO, Jorge (2011), "Open Source Business Intelligence tools for SMEs", Em *6th Iberian Conference on Information Systems and Technologies (CISTI)*, pp. 1–4,
<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5974187&isnumber=5974162>, acessado em fevereiro de 2018.

TURBAN, Efraim, ARONSON, Jay E., & LIANG, Ting-Peng (2007), *Decision Support and Business Intelligence Systems*, 7ª edição, New Delhi, Prentice-Hall.

UNIÃO EUROPEIA (2016), *Factos e números sobre a política comum das pescas*, https://ec.europa.eu/fisheries/sites/fisheries/files/docs/body/pcp_pt.pdf, acessado em fevereiro de 2018.

VALACICH, Joseph., & SCHNEIDER, Christoph (2017), *Information Systems Today: Managing in the Digital World*, 8ª edição, Reino Unido, Pearson Education Limited.

VASCONCELOS, André *et al.* (2002), *Arquitetura de Sistemas de Informação: A Ferramenta de Alinhamento Negócio/Sistemas de Informação*, Centro de Engenharia Organizacional, INES Inovação, Lisboa.

VERCELLIS, Carlo (2009), *Business Intelligence: Data Mining and Optimization for Decision Making*, 1ª edição, Reino Unido, John Wiley & Sons.

WIERINGA, Roel (2009), *Design Science as Nested Problem Solving*, Universidade de Twente, Holanda.

— (2014) *Design Science Methodology for Information Systems and Software Engineering*, 1ª edição, Nova Iorque, Springer.

WILLIAMS, Steve, & WILLIAMS, Nancy (2006), *The Profit Impact of Business Intelligence*, 1ª edição, San Francisco, Elsevier.

WIXOM, Barbara, & WATSON, Hugh (2010), "The BI-Based Organization", *International Journal of Business Intelligence Research*, vol. 1, n.º 1, pp. 13–28. doi: 10.4018/jbir.2010071702

XAVIER, Cristiano, & MOREIRA, Fernando (2013), "Agile ETL", *Procedia Technology*, vol. 9, pp. 381–387. doi: 10.1016/j.protcy.2013.12.042

Z Aidan, Fernando Hadad, BAX, Marcello Peixoto, & PARREIRAS, Fernando Silva (2016), "Design Science Research: Application in a Research and Development Project", Em *3th International Conference on Information Systems & Technology Management - CONTECSI*. doi: 10.5748/9788599693124-13CONTECSI/PS-4163.

Legislação Consultada

CONSELHO UE, Regulamento UE 404/2011, Jornal Oficial da União Europeia, Série L-112/1, 8 de abril de 2011.

— Regulamento UE 1380/2013, Jornal Oficial da União Europeia, Série L-354/22, 11 de dezembro de 2013.

— Regulamento UE 2015/812, Jornal Oficial da União Europeia, Série L-133/1, 20 de maio de 2015.

— Regulamento UE 2017/2403, Jornal Oficial da União Europeia, Série L-347/81, 12 de dezembro de 2017.

REPÚBLICA PORTUGUESA, Ministério do Mar, Decreto Regulamentar 3/93, Diário da República, I Série – B, 8 de fevereiro de 1993.

— Assembleia da República, Diário da República 60-B/97, 03 de abril de 1997.

— Ministério do Equipamento, do Planeamento e da Administração do Território, Decreto-Lei 190/98, Diário da República, I Série – A, 10 de julho de 1998.

— Ministério do Equipamento, do Planeamento e da Administração do Território, Decreto-Lei 191/98, Diário da República, I Série – A, 10 de julho de 1998.

— Ministério da Agricultura, do Desenvolvimento Rural e das Pescas, Decreto-Lei

- 383/98, Diário da República, I Série – A, 27 de novembro de 1998.
- Ministério da Agricultura, do Desenvolvimento Rural e das Pescas, Portaria 1102-E/2000, Diário da República, I Série – B, 22 de novembro de 2000.
- Ministério da Agricultura, do Desenvolvimento Rural e das Pescas, Portaria 1102-G/2000, Diário da República, I Série – B, 22 de novembro de 2000.
- Ministério da Agricultura, do Desenvolvimento Rural e das Pescas, Decreto-Lei 79/2001, Diário da República, I Série – A, 5 de março de 2001.
- Ministério do Equipamento Social, Decreto-Lei 280/2001, Diário da República, I Série – A, 23 de outubro de 2001.
- Ministério da Defesa Nacional, Decreto-Lei 45/2002, Diário da República, I Série – A, 02 de março de 2002.
- Ministério da Agricultura, do Desenvolvimento Rural e das Pescas, Decreto-Lei 246/2002, Diário da República, I Série – A, 8 de novembro de 2002.
- Ministério da Educação, Decreto-Lei 34/2006, Diário da República, I Série – A, 28 de julho de 2006.
- Ministério da Defesa Nacional, Decreto Regulamentar 86/2007, Diário da República, I Série, 12 de dezembro de 2007.
- Ministério da Agricultura, do Mar, do Ambiente e do Ordenamento do Território, Decreto-Lei 49-A/2012, Diário da República, I Série, 29 de fevereiro de 2012.
- Ministério da Defesa Nacional, Decreto-Lei 185/2014, Diário da República, I Série, 29 de dezembro de 2014.
- Ministério da Defesa Nacional, Decreto Regulamentar 10/2015, Diário da República, I Série, 31 de julho de 2015.
- MAR, Decreto-Lei 10/2017, Diário da República, I Série – N.º 7, 10 de janeiro de 2017.

Glossário

Artefacto – objeto que produz conhecimento e a sua construção segue métodos científicos (modelos, algoritmos, métodos). Produzido pelo ser humano para abordar um problema prático.

Base de dados – coleção de dados organizada de uma determinada forma sistematizada. Mantém informação sobre vários tipos de objetos, eventos, pessoas e locais.

Cubo Multidimensional – modelo organizado que permite o relacionamento cruzado de dados. São denominados de “Cubos OLAP”.

Dados – representação de informação.

Dashboard – ferramenta de visualização da informação (conjunto de relatórios - diagramas).

Data Mining – processo de procurar informação empresarial valiosas numa base de dados volumosa.

Dimensão – proporciona uma diferente visão de negócio, por conter vários atributos organizados consoante o tipo de dados. Armazenada no DW sob a forma de tabela.

Data Mart – subconjunto específico de informação do DW. Criado, em norma, quando o âmbito em estudo é mais restrito.

Data Warehouse – repositório de dados (provenientes de várias fontes) que consolida toda a informação da organização. Possibilita aos seus utilizadores uma análise de dados de forma seletiva.

Extract, Transform and Load – processos que permitem efetuar a extração de dados de diversas fontes, transformam esses dados num formato específico (limpo e coerente) e efetua o carregamento dos dados tratados em bases de dados de destino, para análise.

Fato – contexto que gera o valor de um determinado indicador. Armazenado no DW sob a forma de tabela.

Informação – resultado do tratamento de dados.

Negócio – atividades das organizações (não apenas associado a atividades com fins lucrativos)

Online Analytical Processing – sistema analítico que importa dados dos sistemas transacionais (OLTP). Permite criar cubos para analisar a informação sob diversas perspetivas.

Online Transaction Processing – sistema que processa a informação de acordo com regras definidas, armazena a informação e atualiza a informação existente para refletir a nova informação.

Sistemas de Apoio à Decisão – sistemas que combinam modelos e dados na tentativa de resolver problemas semiestruturados e não estruturados.

Sistemas de Informação – conjunto de componentes inter-relacionados que coletam, processam, armazenam e distribuem informação para dar suporte à tomada de decisão.

Apêndices

- A – Conceitos fundamentais sobre *Business Intelligence*
- B – Identificação dos dados, campos e respetivas mensagens
- C – Identificação das mensagens e dos dados referentes às medidas de interesse
- D – Tratamento de incongruências detetadas nos dados recolhidos
- E – Classificação do tipo de dados
- F – Modelação dimensional: esquema em estrela
- G – Modelação dimensional aplicada nas restantes medidas de interesse
- H – Processo ETL referente ao fato
- I – *Dashboard* e relatórios complementares referentes à fiscalização da pesca
- J – Inquérito para avaliação do artefacto produzido

Apêndice A – Conceitos fundamentais sobre *Business Intelligence*

Na sociedade atual assiste-se à conjugação da “produção de quantidades gigantescas de informação” e à “utilização intensiva de tecnologias” (Gasque & Tescarolo, 2004, p. 35).

A informação é um dos recursos mais importantes das organizações. A análise do seu valor intrínseco, torna necessário referir os elementos que para ela contribuem (Moresi, 2000, p. 14).

A organização dos elementos relacionados com a informação é denominada de “Pirâmide do Conhecimento”⁷⁵. Esta pirâmide é constituída por dados, informação, conhecimento e sabedoria, daí ter a designação de “Hierarquia DIKW”⁷⁶ (Rowley, 2007, p.163) (Figura 41).



Figura 41 – Pirâmide do conhecimento⁷⁷

Ackoff, definiu dados como símbolos e produtos da observação, não sendo úteis até que estejam na forma relevante (Rowley, 2007, p. 166). Ou seja, os têm sentido quando se transformam em informação. A partir do momento que se atribui um significado aos dados, estes transformam-se em informação (Loh, 2014, p.13).

A **informação** é usada nas atividades diárias. O conhecimento surge a partir da seleção de informação que se retém, isto é, da informação com credibilidade e relevância (Loh, 2014; Moresi, 2000), sendo que se altera como resultado da interação com o ambiente (Moresi, 2000, p. 19). Por isso, o **conhecimento** consiste na informação organizada e processada, posteriormente transmitida como aprendizagem ou experiência (Turban, Aronson, & Liang, 2007, p. 215). O modo como cada indivíduo utiliza o seu conhecimento designa-se por “**sabedoria**” ou “**inteligência**”, isto é, a capacidade de

⁷⁵ Também denominada por: “hierarquia da informação”; “hierarquia do conhecimento”.

⁷⁶ Hierarquia DIKW: *Data-Information-Knowledge-Wisdom*.

⁷⁷ Fonte: Baseado em Rowley, 2007, p. 163.

aplicar o conhecimento adquirido da informação útil para conseguir aplicar a qualquer tipo de problema. O próprio conceito de “sabedoria/inteligência” remete para o BI (Loh, 2014, p. 13).

Resenha Histórica

Em 1958, Hans Peter Luhn referiu que “a comunicação eficiente é a chave para o progresso em todos os campos do esforço humano” (Luhn, 1958, p. 314). Na década de 50, o conceito de BI ganhou sentido prático quando Luhn, pesquisador da IBM, definiu⁷⁸, os passos necessários para o funcionamento do sistema. Surge assim, a primeira definição de BI – “um sistema automático e flexível em desenvolvimento para divulgar informação às diversas secções de qualquer organização” (Luhn, 1958).

No final da década de 1980 ocorrem melhorias das capacidades de visualização gráfica e de armazenamento de dados. Desse fato resultou a exploração de bases de dados e o desenvolvimento de modelos de simulação dando origem ao conceito “Sistemas de Apoio à Decisão” (Vercellis, 2009, p. 35).

Em 1989, o termo BI é introduzido pela empresa Gartner Group por Howard Dresner, que o descreve como sendo um conjunto de conceitos e métodos que utilizam sistemas de apoio assentes em fatos e dimensões, tendo a finalidade de melhorar a tomada de decisão (Braghittoni, 2017; Power, 2006). A partir daí, usufruindo dos sistemas de apoio à decisão e dos sistemas de informação, torna-se necessário fazer uma separação dos dados que se destinam a esses sistemas. Isto deu origem aos conceitos: *Data Warehouse* e *Data Mart* (Vercellis, 2009, p. 36).

De acordo com Larson (2009), uma organização deve praticar decisões efetivas⁷⁹, de forma a obter a entrega de informação relevante e precisa aos decisores, dentro de um prazo fundamental para o apoio dessa tomada de decisão. Esses objetivos são a definição de BI.

Definições de Business Intelligence

Tabela 13 - Definições de *Business Intelligence*

Conceito	Autor
“é uma consequência natural de uma série de sistemas anteriores projetados para apoiar a tomada de decisão”	Negash (2004)
“reunião de diversos recursos usados para extrair, transformar e analisar grandes volumes de dados, produzindo conhecimento capaz de auxiliar a empresa	Segundo Leme Filho (2004), referido em Dalfovo & Tamborlim, 2017, p. 192

⁷⁸ Através da redação do artigo “A Business Intelligence System”.

⁷⁹ Escolhas que levam a organização a atingir as metas estabelecidas, em tempo eficiente.

a tomar decisões de negócio com mais garantia de sucesso”	
“processo de transformar dados em informação e de seguida em conhecimento”	Golfarelli, Rizzi, & Cella (2004)
“processo produtivo cuja matéria-prima é a informação e o produto final o conhecimento”	Sezões <i>et al.</i> (2006)
“qualquer atividade, ferramenta, ou processo usado para obter a melhor informação para suportar o processo de tomada de decisão”	Scheps (2008)
“conjunto de modelos matemáticos e metodologias de análise que exploram os dados para produzir informação e conhecimento útil para processos de tomada de decisão complexos”	Vercellis (2009)
“informação que as pessoas utilizam para suportar os esforços da tomada de decisão”	Baltzan & Philips (2009)
““umbrella term” que é geralmente usado para descrever tecnologias, aplicações, e processos para compilar, armazenar, conectar a analisar dados para ajudar os utilizadores a tomarem melhores decisões”	Wixom & Watson (2010)
“reúne produtos, tecnologia e métodos para organizar informação chave que a direção precisa para melhorar o lucro e desempenho”	Williams & Williams (2006)
“ativo estratégico que pode dar às organizações uma vantagem competitiva, capacitando funcionários de qualquer nível a tomar decisões rápidas. Tem como foco determinar as principais medidas de desempenho, integrar dados de sistemas diferentes de toda a empresa num data warehouse, planear, prever com visão histórica, atual e futura”	Singh & Singh (2013)
“BI é um processo que envolve métodos, técnicas, tecnologias, pessoas,	Loh (2014)

informação, fontes de informação, métricas, ferramentas”	
“um conjunto de soluções tecnológicas que envolve um processo de extração, transformação, análise e distribuição de dados para a tomada de decisão”	Denilson da Silva <i>et al.</i> (2016)
“BI é uma metodologia, não uma ferramenta.”	Braghittoni (2017)
“processo automatizado para coletar dados brutos de fontes heterogêneas e organizá-las de forma sistemática”	Vo <i>et al.</i> (2017)

Em suma, tendo em conta a diversidade e quantidade de definições existentes acerca do termo BI, pode-se verificar que é um termo que surge pelas diversas necessidades que as organizações apresentam, como a “adequação às várias realidades da vida das organizações” (Sezões *et al.*, 2006, p. 10). Ao longo dos anos, ocorre a inserção dos componentes de BI, que serão abordados mais adiante. Um termo que está sempre presente em todas as definições, independentemente do ano em que surgiu, é a tomada de decisão, sendo assim, BI é um instrumento que tem fortes potencialidades para desencadear uma melhor tomada de decisão.

Sistemas Transacionais

Os sistemas transacionais sustentam as operações diárias de uma organização, cujas transações são geradas e atualizadas. Estes sistemas são analisados como um modelo relacional de base de dados e também como o “suporte principal dos sistemas de informação” (Bog, 2014; Braghittoni, 2017; Santos & Ramos, 2009). Estes sistemas foram criados para assegurar um determinado desempenho em que: as operações⁸⁰ a efetuar relativamente aos dados devem ser rápidas; todas as ações⁸¹ do utilizador devem ser inseridas na base de dados, devendo por isso ser um sistema capaz de suportar um grande volume de dados; e deve atender às ações do utilizador em tempo real (Scheps, 2008, p. 85).

As tarefas diárias, numa organização, consistem em obter os dados mais atuais para conseguir dar resposta a clientes internos e externos⁸², através do acesso à base de dados transacional. No entanto, esse acesso, neste tipo de sistema, surge como uma dificuldade (Powell, 2005, p. 13). Este tipo de sistema apresenta algumas limitações, tais como: redução da velocidade de operação do sistema devido a excesso de transações; período de armazenamento de transações limitado, ocorrendo a eliminação dos dados

⁸⁰ Ler, escrever, remoção de dados.

⁸¹ Adição, exclusão, alteração.

⁸² Consumidores, empresas, lojas.

mais remotos; tempo excedente e possibilidade de ocorrerem bloqueios ao tentar associar uma grande quantidade de transações (Larson, 2009, pp. 27–29). Por isso, uma base de dados OLTP é “intensiva em processamento e necessidade de memória, e pouco exigente em operações de acesso a disco” (Powell, 2005, pt. Prefácio).

Apêndice B – Identificação dos dados contidos nas mensagens, relativos à fiscalização da pesca

Identificação dos dados, campos e respectivas mensagens

A identificação dos dados que englobam as mensagens que têm pelo menos um campo relativo à fiscalização encontram-se descritos na Tabela 14.

Tabela 14 - Identificação de dados e respetivos campos e mensagens, relativos à fiscalização

Dados	Campo (1)	Mensagem	Dados	Campo (1)	Mensagem	Dados	Campo (1)	Mensagem	
Nome da Operação	OPER	INCIDENTREP	Artes pesca em uso	ACT	FISCREP	Nr FISCREP	SUM	FISCSUM	
Nr Sequencial – Operação			Classe de malhagem			Nr Relatório de alvo SIFICAP			
Unidade Naval	MSGID	INCIDENTREP FISCREP FISCSUM	Tipo de atividade			LAW	FISCSUM	Ação tomada	RMKS
Nr Ordem Anual do Comunicado			Nr Sequencial – Presumível Infração	Código - Presumível Infração	SUM			INCIDENTREP	
Mês			INCIDENTREP FISCSUM	Código - Comportamento Grave	LAW	FISCREP	Informação Complementar	RMKS	FISCREP FISCSUM
Descrição do Evento			DESC	Presumível Infração por extenso			Nr Sequencial da Captura de Pescado	FISH	RES
Posição Geográfica	MSGID	FISCREP	Grupo-Data-Hora	SUM	FISCSUM	Resultado da Vistoria/ Inspeção	SUM	FISCSUM	
Posição Inspeção			AREA			FISCREP	Quantidade de Pescado	FISH	VIFISC
Área	ENVCOM	INCIDENTREP FISCREP	Condições Tempo	DIV	FISCREP	Duração - Fiscalização das Atividades de Pesca	ACTID / FPESCA	LOGCONTREP	
Direção e Forção do Vento			Nr Página Diário de Pesca do dia			FISCSUM			Início - Fiscalização das Atividades de Pesca
Estado do Mar			Arte de Pesca Utilizada						Fim - Fiscalização das Atividades de Pesca
Altura e Direção da Ondulação			Classe de Malhagem Utilizada			FISCSUM	Total de Horas de Fiscalização	TEMPOS	FISCSUM
Visibilidade			Nome do Mestre/Patrão da Embarcação				Subtotais nas Áreas – Águas de Pesca		
Tipo de Embarcação			IDEMB			FISCREP	Nr Emissão da Cédula/ Carta de Recreio do Mestre da Embarcação	FISCSUM	
Nr Relatório - SIFICAP	SUM	FISCSUM	Nr Contribuinte do Mestre da Embarcação						
Subtipo de Embarcação	IDEMB	FISCREP	Morada do Mestre da Embarcação	FISCSUM					
Nome Embarcação			Matricula ou registo		SUM	FISCREP FISCSUM			

Considerar "Nr" - Número

(1) Campos:

ACT – Tipo de Atividade

ACTID / FPESCA - Identificação da Atividade / Navio em Fiscalização das Atividades de Pesca

AREA – Posição inspeção / Área

DESC - Descrição

DIV – Diversos

ENVCON – Condições Ambientais

FISH – Capturas de Pescado

IDEMB – Identificador da Embarcação

LAW – Presumíveis Infrações

MSGID – Identificador de Mensagem

OPER – Identificação da Operação

REF – Referência

RES – Resultado

RMKS - Observações

SUM – Sumário da Atividade de Fiscalização Desenvolvida

TEMPOS – "Tempos" de Missão

VIFISC – Vigilância e Fiscalização

Compilação dos dados de interesse

Tabela 15 - Compilação dos dados de interesse relativos ao fato

Fato										
FISCALIZAÇÃO DA PESCA										
Dimensão	Atributos	Contempla o campo		IONAV1010 - Mensagem "FISCREP"			APP11			
		IONAV1010	APP11	Mensagem	Campo	Local interesse	Mensagem	Campo	Local interesse	
Unidade	Unidade Naval	x			MSGID (1)	FISCREP/unidade naval/-/-//				
Zona Marítima	Posição Inspeção	x			AREA (4)	AREA/pos/-//				
	Área	x				AREA/-/area// (Tab 4 Anx D)				
Condições ambientais	Condições Tempo	x			ENVCOM (9)	ENVCOM/tempo/-/-/-//				
	Direção e Força do Vento	x				ENVCOM/-/vento/-/-/-//				
	Estado do Mar	x				ENVCOM/-/-/estadomar/-/-//				
	Altura e Direção da ondulação	x				ENVCOM/-/-/ondulacao/-//				
	Visibilidade	x				ENVCOM/-/-/-/visibilidade//				
ID Embarcação	Tipo de Embarcação	x		FISCREP	IDEMB (3)	IDEMB/tipo/-/-/-// (Tab 1 Anx D)				
	Subtipo de Embarcação	x				IDEMB/-/subtipo/-/-/-// (Tabs 2-A, 2-B, 2-C e 2-D Anx D)				
	Matrícula ou Registro	x				IDEMB/-/-/matriculaouregistro/-/-//				
	Nome	x				IDEMB/-/-/-/nome/-//				
	Nacionalidade	x				IDEMB/-/-/-/-/nacionalidade// (Anexo A)				
Tipo de Atividade	Atividade	x			ACT (6)	ACT/atividade/-/-// (Tab 5 Anx D)				
	Arte de Pesca em uso	x				ACT/-/artepesca/-// (Tab 3 Anx D)				
	Classe de Malhagem	x				ACT/-/-/classemalhagem//				
Infrações	Código	x			LAW (7)	LAW/-/codigo/-/-// (Tab 7 Anx D)				
	Comportamento grave-código	x				LAW/-/-/codigocomprave/-// (Tab 6 Anx D)				
	Presumível infração	x				LAW/-/-/-/infracao//				
Tipo de Pesca	Quantidade (kg) - Espécie	x			FISH (8)	FISH/-/quantidade - especie/-/-// (Tab 8 Anx D)				
	Nr página diário	x				FISH/-/-/nrpagina/-/-//				
	Arte de Pesca utilizada	x				FISH/-/-/-/artepesca/-// (Tab 3 Anx D)				
	Classe de Malhagem	x				FISH/-/-/-/-/classemalhagem//				
Resultados Vistoria	Resultado	x			RES (5)					
	Observações (Ação tomada, etc)	x			RMKS (10)	Todo o campo				
Tempo	Dia	x			MSGID (1)	FISCREP/-/-/GDH//				
	Mês	x								
	Ano	x								
	Hora	x								

Considerar:

“Nr” – Número; “ID” – Identificação; “Tab” – Tabela;

“Anx” – Anexo; “GDH” – Grupo-Data-Hora

Campos (números inseridos nesta coluna correspondem ao número sequencial dos campos do FISCREP):

- (1) MSGID – Identificador de Mensagem
- (2) REF – Referência
- (3) IDEMB – Identificador da Embarcação
- (4) AREA – Posição inspeção / Área
- (5) RES – Resultado
- (6) ACT – Tipo de Atividade
- (7) LAW – Presumíveis Infrações
- (8) FISH – Capturas de Pescado
- (9) ENVCON – Condições Ambientais
- (10) RMKS – Observações

Apêndice C – Identificação das mensagens e dos dados referentes às medidas de interesse

O processo que diz respeito a esta etapa decorreu da seguinte forma: identificação das mensagens que integravam mais dados acerca da medida de interesse, tanto da IONAV1010 como do APP11; extração de dados desses referidos campos; e obtenção dos dados de interesse, possibilitando a perceção dos dados em comum, ou não, das referidas publicações. A importância de se realizar a comparação de mensagens de diferentes fontes de dados, ocorre pela necessidade de atender às diretrizes: transformação de mensagens estruturadas em MTF, de modo a facilitar o tratamento e a análise de dados.

Sendo assim, essa comparação auxiliará a perceber as principais diferenças entre o tipo de dados que compõem os diversos comunicados.

Os valores entre parêntesis na coluna que diz respeito ao campo, representam o número sequencial desse campo numa determinada mensagem.

Consumo de Combustível

Para esta medida de interesse, existem duas mensagens a nível nacional: LOGCONTREP e LOGBASEREP (Figura 42), que se distinguem no que diz respeito ao consumo de unidades navais que tenham missão atribuída, ou não, respetivamente. Relativamente ao APP11, não há nenhuma mensagem de carácter exclusivo relativo ao consumo de combustíveis.

Fato			
CONSUMO DE COMBUSTÍVEL			
Dimensão	Atributos	IONAV1010	
		Mensagem	Campo
Unidade	Unidade Naval	LOGBASEREP LOGCONTREP	MSGID (1)
Consumo a navegar	Número sequencial de consumo	LOGBASEREP LOGCONTREP	COMBS (2) COMBS (7)
	Tipo de combustível		
	Total LT de combustível consumido		
	Preço LT (EUR)		
	Custo total (EUR)		
Consumo apoio à Formação e Treino	Número sequencial de consumo	LOGBASEREP	FORM (3)
	Tipo de equipamento		
	Total horas de funcionamento		
	Total combustível consumido		
	Total óleo consumido		
	Média de potência fornecida		
	Observações		
Consumo para Operações de Manutenção	Número sequencial de consumo	LOGBASEREP	MANUT (4)
	Tipo de equipamento		
	Total horas de funcionamento		
	Total combustível consumido		
	Total óleo consumido		
	Média de potência fornecida		
	Observações		
Informação	Info complementar	LOGBASEREP	AMPN (5)

Figura 42 - Dados de interesse referentes ao consumo de combustível, a nível nacional

Consumo de Munições

A mensagem que contém mais dados acerca do consumo de munições, a nível nacional é o LOGCONTREP. No entanto, existe um comunicado de consumo que é elaborado aquando do uso de munições, sendo que incorpora todos os dados que possibilitam uma análise do consumo. Esse comunicado de consumo tem como fonte de dados: ILDINAV709. Uma vez que esse referido comunicado não está incorporado em nenhuma das fontes de dados, seria vantajoso considerar a sua colocação na IONAV1010.

No que diz respeito às restantes medidas de interesse, há uma grande variedade de dados e na maioria, estão em ambas as fontes de dados, pelo que um panorama clarificado auxiliará a obter uma visão geral deste assunto e poderá influenciar no processo de tomada de decisão. Os campos que não estão em comum, nas diversas mensagens analisadas são de carácter importante, como já foi referido, visto possibilitarem dados estatísticos mais completos, por isso devem ser contemplados aquando da seleção de dados para posterior análise.

Limitações Operacionais

As mensagens com mais dados relativos às limitações operacionais são: OPSTAT DEFECT e NAVOPDEF, nível nacional e NATO, respetivamente (Figura 43) Essas mensagens revelam limitações operacionais quer a nível material ou pessoal.

Fato								
LIMITAÇÕES OPERACIONAIS								
Dimensão	Atributos	Contempla o campo		IONAV1010		APP11		
		IONAV1010	APP11	Mensagem	Campo	Mensagem	Campo	Subcampo
Missão	Operação	x	x	OPSTAT DEFECT	OPER (1)	NAVOPDEF	OPER (2)	1A
	Exercício	x	x		EXER (2)		EXER (1)	2A
Unidade	Unidade Naval	x	x		MSGID (3)		MSGID (3)	4A
Descrição do Incidente	Equipamento (Pessoal/Material)	x	x		EQUIP (5)		DEFEQUIP (9)	1A
	Número funcional (se for Material)	x	x					3A
	Número da deficiência (inclui área dos serviços)	x	x		DEF (7)		OPDEF (5)	1A
	Descrição da Avaria	x	x				DEFEQUIP (9)	2A
	GDH início do incidente	x			RPTYP (6)			
	Tempo estimado para reparação - reposição do sistema	x	x		ETR/ETBOL (9)		REPAIR (11)	1B e 1C
Limitação Operacional	Critério de limitação	x	x					
	Limitação Operacional	x	x	OPSLIMIT (8)	OPDEFDET (6)	3A		
	Quantidade de itens limitados		x			1A		
	Percentagem estimada de capacidade perdida	x			DEFPART (1)	4A		
	GDH retificação	x		OPSLIMIT (8)				
Assistência	Confirmação de Assistência (SIM ou NÃO)	x	x					
	Descrição do tipo de apoio necessário	x		ASSIST (10)				
	Local de assistência		x			2A		
	Data de início do pedido		x			3A		
	Data de assistência no local		x			4A		
Info Complementar	Ações tomadas a bordo	x	x		ACTION (11)		NARR	1A

Figura 43 - Dados de interesse respeitantes às limitações operacionais

No que concerne a esta medida de interesse, pode ser pertinente considerar a colocação da mensagem “NAVOPDEF” na IONAV1010, de forma a alcançar um aproveitamento da mesma informação, mas no formato correto e útil para a posterior análise de dados. Posto isto, a colocação dessa mensagem na futura IONAV1010 pode

envolver a inserção de alguns campos que constam na mensagem nacional e que podem ser úteis (como a porcentagem estimada de capacidade perdida).

Operações de Busca e Salvamento

No que diz respeito a esta medida de interesse, as mensagens analisadas (SARREP – nível nacional – e SARIR – nível NATO) (Figura 44) focam-se em qualquer situação que solicite uma operação de busca e salvamento, cujo envio sucede assim que essa ação termina.

Fato										
OPERAÇÕES DE BUSCA E SALVAMENTO										
Dimensão	Atributos	Contempla o campo		IONAV1010		APP11				
		IONAV1010	APP11	Mensagem	Campo	Mensagem	Campo	Subcampo		
Unidade	Unidade Naval	x	x	SARREP	MSGID (1)	SARIR	MSGID (3)	4A		
Identificação do Assistido	Nome	x	x		IDASS (3)			SEAINCIDT (8)	1A	
	Indicativo de Chamada	x	x						6A	
	Nacionalidade	x								
Área	Área da ação SAR	x	x		AREA (4)		SEAINCIDT (8)	3C		
Condições Meteorológicas	Vento	x	x		ENVCON (7)			MET (13)	2A	
	Mar	x	x						14A	
	Visibilidade	x	x						4A	
Assistência	Aspectos importantes da assistência	x	x		ASS (5)		SARAR	1A		
	Indicação de cooperações por outras entidades	x			COOP (6)					
Tempos	GDH do início da ação	x	x		ESTAT (8)			AVAILSAR	2A	
	GDH do fim da ação	x	x						3A	
	Número de horas da ação	x								
Encargos	Custos envolvidos com pessoal, material (combustível)	x			ENC (9)					

Figura 44 - Dados de interesse relativos às operações de busca e salvamento

A ação sugerida na medida de interesse anterior, aplica-se também neste âmbito. Isto porque a mensagem “SARIR” contempla a maioria dos campos de interesse para uma posterior análise, pelo que há a possibilidade de considerar a sua colocação na IONAV1010, em detrimento da transformação da mensagem SARREP.

Navegação

Relativamente à Navegação, a mensagem selecionada para recolha de dados foi: MISREP (nível nacional). Esta mensagem fornece uma síntese de diversas atividades relativas a uma determinada missão (Figura 45). Após uma análise às diversas mensagens que o APP11 compreende, observa-se que não há uma mensagem equivalente que possibilite a comparação de dados. Sendo assim e dada a importância que pertence a esta medida de interesse, principalmente pela posterior análise de dados, surge a ponderação de transformar essa referida mensagem em MTF e, de seguida, a sua promulgação na IONAV1010.

Fato									
Navegação									
Dimensão	Atributos	Contempla o campo		IONAV1010		APP11			
		IONAV1010	APP11	Mensagem	Campo	Mensagem	Campo	Subcampo	
Unidade	Unidade Naval	x							
Missão	Operação	x		MISREP	MSGID (3)				
	Exercício	x			OPER (1)				
	Descrição da Missão	x			EXER (2)				
	Descrição da Tarefa	x			TASK (5)				
	Indicação se os objetivos foram atingidos	x			RESULT (6)				
Tempos Navegação	GDH Início da Missão	x			TEMPOS (7)				
	GDH Fim da Missão	x							
	Duração da Tarefa	x							
	Duração da Missão	x							
	Tempo navegação	x							
	Taxa navegação	x							
	Taxa imobilização (caso de avaria)	x							
Distâncias	Taxa imersão/vela	x							
	Distância Percorrida (MI ou KM)	x			DIST (8)				
Portos	SOA	x							
	ATA	x			SAIL (10)				
	Local	x							
ATD	x								

Figura 45 - Dados de interesse respeitantes à navegação

Apêndice D – Tratamento de incongruências detetadas nos dados recolhidos

Query para tratamento de dados com chave “0”

O tratamento de dados que compreendiam a chave “0” que passaram a ser “-1”, e a correspondência desta nova chave à expressão “N/D”, foi realizado pela *query* abaixo Figura 46.

```
UPDATE Historico
SET Ano = -1
WHERE Ano = 0
go
UPDATE Historico
SET Mes = -1
WHERE Mes= 0
go
UPDATE Historico
SET Dia_Mes = -1
WHERE Dia_Mes= 0
go
UPDATE Historico
SET Dia_Semana = -1
WHERE Dia_Semana = 0
go
UPDATE Historico
SET Hora = -1
WHERE Hora= 0
go
UPDATE Historico
SET Período_Dia = -1
WHERE Período_Dia= 0
go
UPDATE Historico
SET Unidade = -1
WHERE Unidade= 0
go
UPDATE Historico
SET Tipo_embarcacao = -1
WHERE Tipo_embarcacao= 0
go
UPDATE Historico
SET Subtipo_embarcacao = -1
WHERE Subtipo_embarcacao= 0
go
UPDATE Historico
SET Arte = -1
WHERE Arte = 0
go
UPDATE Historico
SET Resultado = -1
WHERE Resultado= 0
go
UPDATE Historico
SET Area_operacao = -1
WHERE Area_operacao= 0
go
UPDATE Historico
SET Nacionalidade = -1
WHERE Nacionalidade = 0
go
UPDATE Historico
SET Area_oceanica_dist_costa = -1
WHERE Area_oceanica_dist_costa= 0
go
UPDATE Historico
SET Local_Inspecao = -1
WHERE Local_Inspecao= 0
go
UPDATE Historico
SET Tipo_Pescado = -1
WHERE Tipo_Pescado= 0
go
UPDATE Historico
SET Hora = 9999
WHERE Hora = 0

INSERT INTO Ano
VALUES (9999,-1)
go
INSERT INTO Area_Oceanica_dist_costa
VALUES ('N/D','N/D',-1)
go
INSERT INTO Area_Operacao
VALUES ('N/D','N/D',-1)
go
INSERT INTO Artes
VALUES ('N/D','N/D','N/D',-1,'N/D')
go
INSERT INTO Dia_Semana
VALUES ('N/D',-1)
go
INSERT INTO Especie
VALUES (-1,'N/D','N/D')
go
INSERT INTO Local_Inspecao
VALUES ('N/D','N/D',-1,'N/D')
go
INSERT INTO Mes
VALUES ('N/D','N/D',-1)
go
INSERT INTO Nacionalidade
VALUES ('N/D','N/D',-1)
go
INSERT INTO Período_Dia
VALUES ('9999','9999',-1)
go
INSERT INTO Resultado
VALUES ('N/D',-1)
go
INSERT INTO Subtipo1Pesca
VALUES ('N/D','N/D','N/D',-1,'N/D')
go
INSERT INTO Subtipo2Recreio
VALUES ('N/D',-1)
go
INSERT INTO Subtipo3MarTur
VALUES ('N/D','N/D',-1)
go
INSERT INTO Subtipo4ArtesCaladas
VALUES ('N/D','N/D',-1)
go
INSERT INTO Subtipo5Outras
VALUES ('N/D',-1,'N/D')
go
INSERT INTO Tipo_Embarcacao
VALUES ('N/D','N/D',-1)
go
INSERT INTO UnidadeNaval
VALUES ('N/D','N/D','N/D','N/D',-1)
```

Figura 46 - Tratamento de dados relativos a incongruências entre chave de ligação e respetiva expressão

Query para tratamento de dados decimais

A resolução dos dados decimais que, por isso tornam determinados dados irrealis, foi solucionada pela *query* apresentada abaixo (Figura 47).

```
update Historico
set Resultado = -1
where Resultado = 1.000976563
go
update Historico
set Unidade = -1
where Unidade = 37.5
go
update Historico
set Unidade = -1
where Unidade = 36.03125
go
update Historico
set Area_oceanica_dist_costa= -1
where Area_oceanica_dist_costa = 6.00390625
go
update Historico
set Area_oceanica_dist_costa= -1
where Area_oceanica_dist_costa = 8.0078125
go
update Historico
set Hora= 9999
where Hora = -615
go
update Historico
set Dia_Mes= -1
where Dia_Mes = 3.03125
go
update Historico
set Dia_Mes= -1
where Dia_Mes = 5.0625
go
update Historico
set Dia_Mes= -1
where Dia_Mes = 10.0078125
go
update Historico
set Dia_Mes= -1
where Dia_Mes = 11.125
go
update Historico
set Dia_Mes= -1
where Dia_Mes = 13.0078125
go
update Historico
set Dia_Mes= -1
where Dia_Mes = 18.25
go
update Historico
set Dia_Mes= -1
where Dia_Mes = 19.015625
go
update Historico
set Dia_Mes= -1
where Dia_Mes = 20.015625
go
update Historico
set Dia_Mes= -1
where Dia_Mes = 99
go
update Historico
set Mes= -1
where Mes = 12.0078125
go
update Historico
set Ano= -1
where Ano = 4.00390625
go
update Historico
set Ano= -1
where Ano = 4.0625
go
```

Figura 47 - Tratamento dos dados referente à presença de dados decimais

Apêndice E – Classificação do tipo de dados

Tabela 16 - Classificação do tipo de dados relativos à tabela principal e às tabelas auxiliares

TABELA PRINCIPAL	Dados	Tipo de Dados	TABELA PRINCIPAL	Dados	Tipo de Dados
	ID_Historico	int		Area_oceanica_dist_costa	smallint
Latitude	float	Nr_infracoes_tipo1			
Longitude		Nr_infracoes_tipo2			
Ano	smallint	Nr_infracoes_tipo3			
Mes		Nr_infracoes_tipo4			
Dia_Mes		Nr_infracoes_tipo5			
Dia_Semana		Nr_infracoes_tipo6			
Hora		Nr_infracoes_tipo7			
Periodo_Dia		Nr_infracoes_tipo8			
Unidade		Nr_infracoes_tipo9			
Tipo_embarcacao		Nr_infracoes_tipo10			
Subtipo_embarcacao		Nr_infracoes_tipo11			
Arte		Nr_infracoes_tipo12			
Resultado		Nr_infracoes_tipo13			
Nr_FISCREP		Nr_infracoes_tipo14			
Area_operacao		Local_inspecao			
Nacionalidade		Tipo_Pescado			
Nr_Registo	Quantidade_Pescado	decimal(6,2)			
GDH (Grupo Data-Hora)	float				

Tabela Auxiliar	Dados	Tipo de Dados
Ano	Ano	smallint
	PK_Ano	int
Area_Oceanica_dist_costa	AreaOc_abrev	varchar(30)
	AreaOc	varchar(40)
	PK_AreaOc	int
Area_Operacao	AreaOp_abrev	varchar(10)
	AreaOp	varchar(20)
	PK_AreaOp	int
Artes	Artes_nome_incompleto	varchar(50)
	Abreviatura	varchar(10)
	Ficheiros	varchar(50)
	Artes_nome_completo	varchar(60)
	PK_Artes	int
Dia_Semana	Dia	varchar(10)
	PK_Dia_Semana	int
Especie	Código	varchar(10)
	Nome	varchar(100)
	PK_Especie	int
Local_Inspecao	Local_casplock	varchar(50)
	Abreviatura	varchar(10)
	Local	varchar(50)
	PK_Local_Inspecao	int
Mes	Mes_portuges	varchar(20)
	Mes_ingles	
	PK_Mes	int
Nacionalidade	Nac_abrev	varchar(10)
	Nacionalidade	varchar(50)
	PK_Nacionalidade	int

Tabela Auxiliar	Dados	Tipo de Dados
Periodo_Dia	Hora_Inicio	time(7)
	Hora_Fim	time(7)
Resultado	PK_Periodo_Dia	int
	Resultado	varchar(30)
Subtipo1Pesca	PK_Resultado	int
	Subtipo_capslock	varchar(30)
	Código	varchar(10)
	Ficheiros	varchar(40)
	Subtipo	varchar(30)
Subtipos2Recreio	PK_Subtipo1	smallint
	Subtipo	varchar(20)
Subtipo3MarTur	PK_Subtipo2	smallint
	Subtipo	varchar(30)
	Codigo	varchar(10)
Subtipo4ArtesCaladas	PK_Subtipo3	smallint
	Subtipo	varchar(40)
	Codigo	varchar(10)
Subtipo5Outras	PK	smallint
	Subtipo	varchar(30)
Tipo_Embarcacao	Codigo	varchar(10)
	PK	smallint
	Subtipo	varchar(30)
UnidadeNaval	Tipo	varchar(20)
	Tipo_abrev	varchar(10)
Unidade	PK_Tipo_Embarcacao	int
	Indicativo_chamada	varchar(30)
	Tipo	varchar(50)
	Classe	varchar(30)
	Unidade	varchar(50)
	PK_UnidadeNaval	int

Apêndice F – Modelação dimensional: esquema em estrela

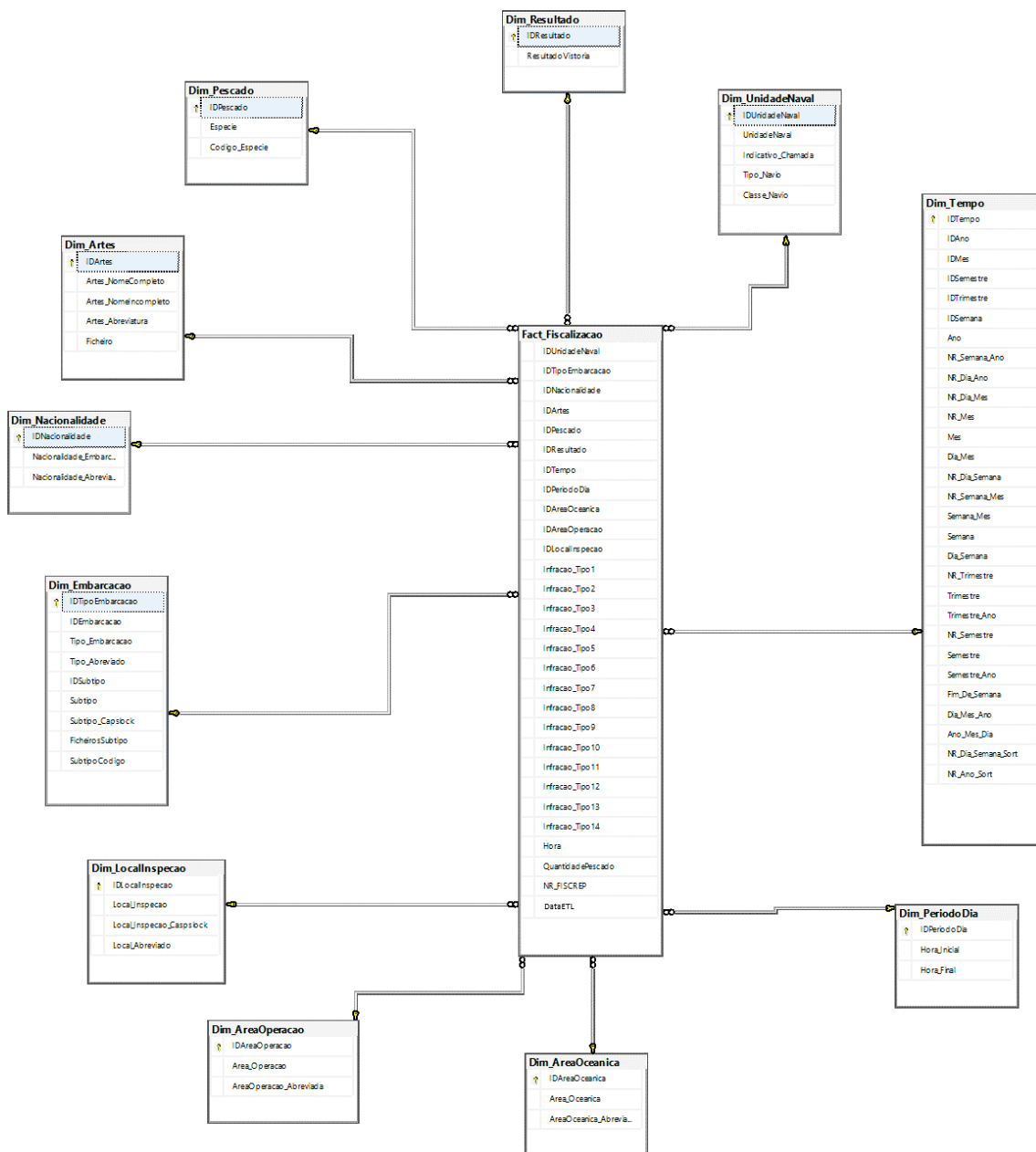


Figura 48 - Esquema em estrela, pertencente ao fato da fiscalização da pesca

Apêndice G – Modelação dimensional aplicada nas restantes medidas de interesse

Limitações Operacionais

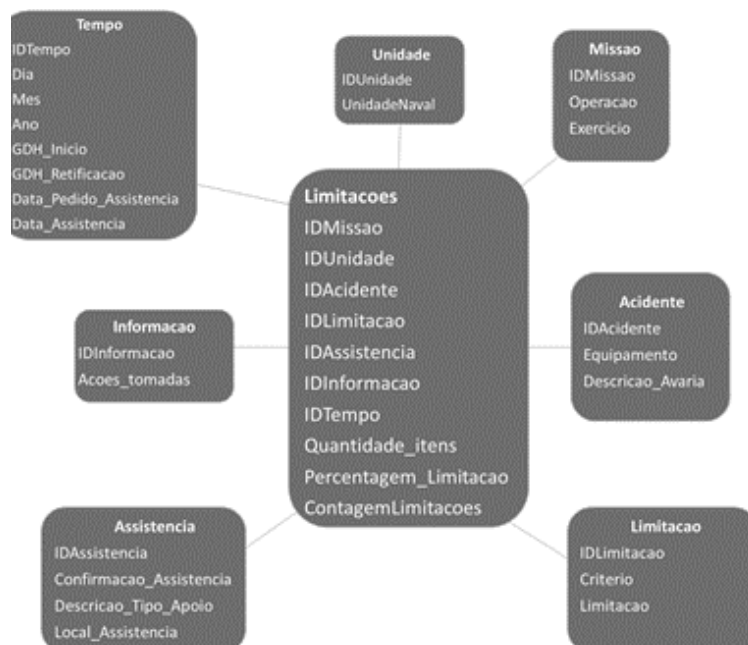


Figura 49 - Modelação dimensional referente às limitações operacionais

Operações de Busca e Salvamento

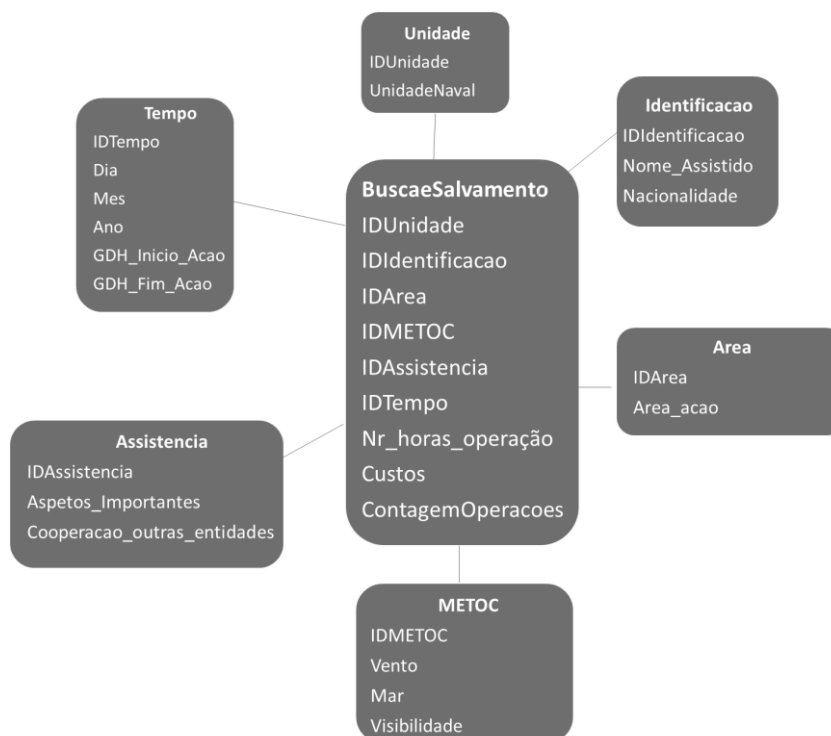


Figura 50 - Modelação dimensional aplicada nas operações de busca e salvamento

Navegação

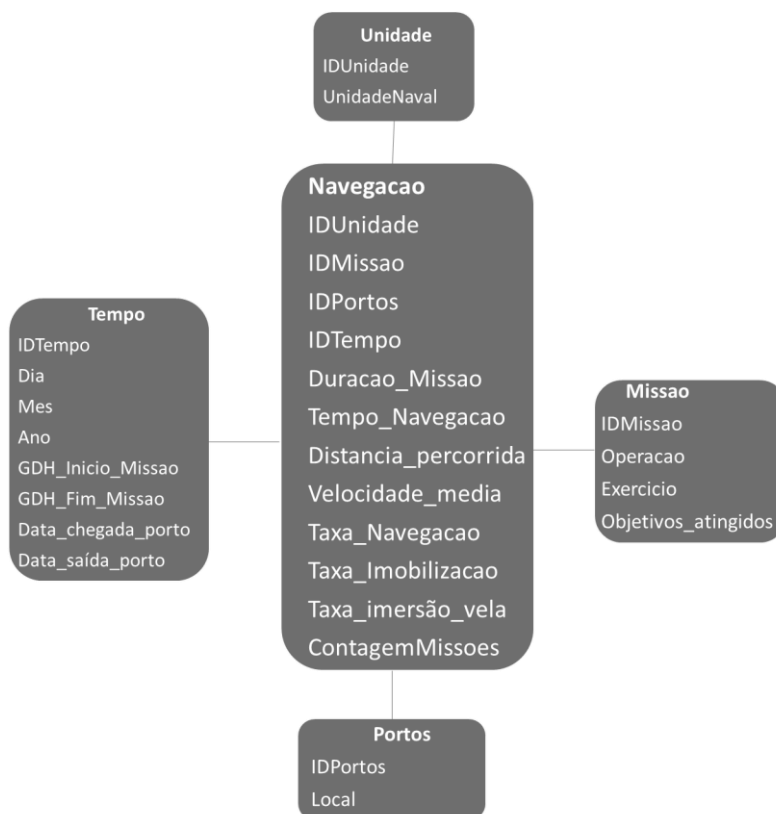


Figura 51 - Modelação Dimensional relativa à navegação

Apêndice H – Processo ETL referente ao fato

Query para extração de dados da FonteDadosMTF:

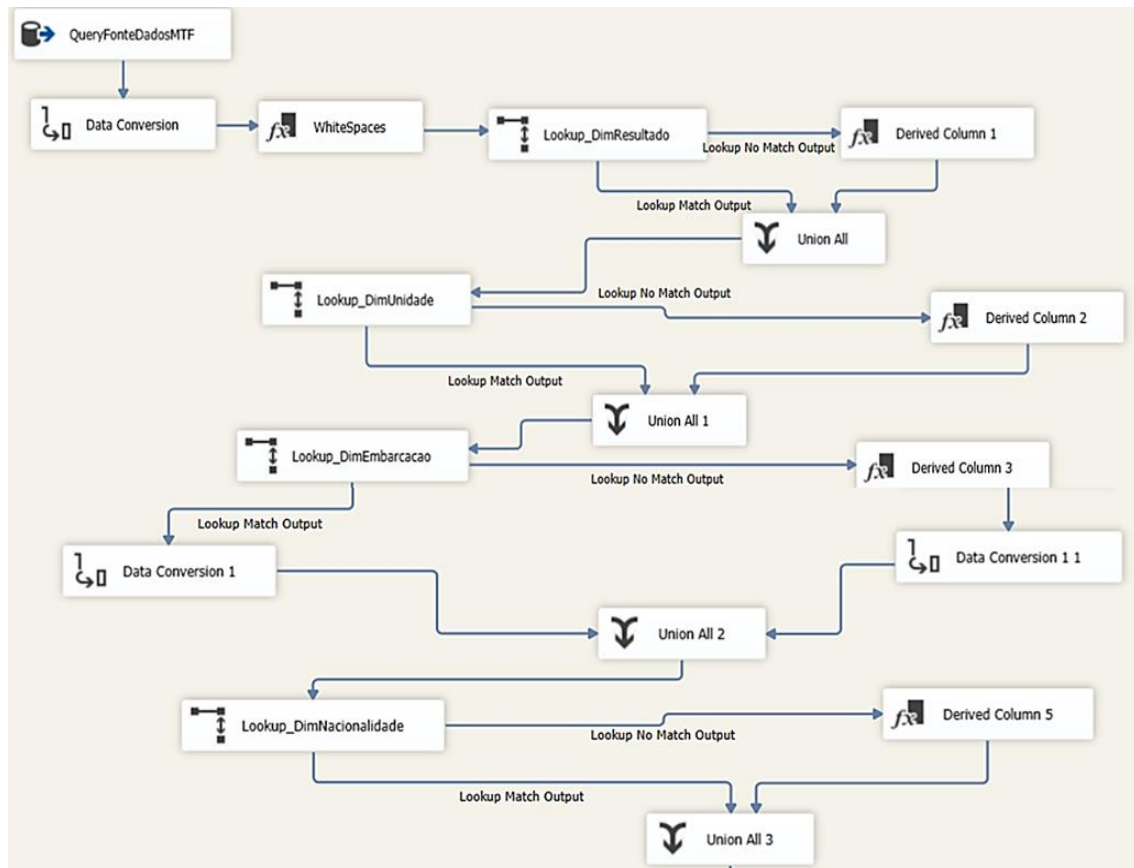
```

SELECT DBO.Ano.ANO, ltrim(rtrim(Mes_portugues)) as Mes_portugues, Dia_Mes, ltrim(rtrim(Dia)) as Dia,
HORA, Hora_Inicio, Hora_Fim, ltrim(rtrim(DBO.UnidadeNaval.Indicativo_chamada)) as Indicativo_chamada,
ltrim(rtrim(DBO.Tipo_Embarcacao.Tipo_Abreviado)) as Tipo_abrev, ltrim(rtrim(DBO.Historico.Subtipo_embarcacao)) as Código,
ltrim(rtrim(dbo.Artes.Abreviatura)) as ArtesAbreviatura, ltrim(rtrim(dbo.Resultado.Resultado)) as Resultado,
Nr_FISCREP, ltrim(rtrim(AreaOp_abrev)) as AreaOp_abrev, ltrim(rtrim(Nac_abreviada)) as Nac_abreviada,
ltrim(rtrim(AreaOc_abrev)) as AreaOc_abrev, Nr_infracoes_tipo1, Nr_infracoes_tipo2, Nr_infracoes_tipo3,
Nr_infracoes_tipo4, Nr_infracoes_tipo5, Nr_infracoes_tipo6, Nr_infracoes_tipo7, Nr_infracoes_tipo8,
Nr_infracoes_tipo9, Nr_infracoes_tipo10, Nr_infracoes_tipo11, Nr_infracoes_tipo12, Nr_infracoes_tipo13,
Nr_infracoes_tipo14, ltrim(rtrim(dbo.Local_Inspecao.Abreviatura)) as InspecaoAbreviatura,
ltrim(rtrim(dbo.Especie.Código)) as Especie, Quantidade_Pescado
FROM DBO.Historico
FULL OUTER JOIN DBO.Ano ON DBO.Historico.Ano = DBO.Ano.PK_Ano
FULL OUTER JOIN DBO.Mes ON DBO.Historico.Mes = DBO.Mes.PK_Mes
FULL OUTER JOIN DBO.Dia_Semana ON DBO.Historico.Dia_Semana = DBO.Dia_Semana.PK_Dia_Semana
FULL OUTER JOIN DBO.Periodo_Dia ON DBO.Historico.Periodo_Dia = DBO.Periodo_Dia.PK_Periodo_Dia
FULL OUTER JOIN DBO.UnidadeNaval ON DBO.Historico.Unidade = DBO.UnidadeNaval.PK_UnidadeNaval
FULL OUTER JOIN DBO.Tipo_Embarcacao ON DBO.Historico.Tipo_embarcacao = DBO.Tipo_Embarcacao.PK_Tipo_Embarcacao
FULL OUTER JOIN DBO.Artes ON DBO.Historico.Arte= DBO.Artes.PK_Artes
FULL OUTER JOIN DBO.Resultado ON DBO.Historico.Resultado= DBO.Resultado.PK_Resultado
FULL OUTER JOIN DBO.Area_Operacao ON DBO.Historico.Area_operacao= DBO.Area_Operacao.PK_AreaOp
FULL OUTER JOIN DBO.Nacionalidade ON DBO.Historico.Nacionalidade= DBO.Nacionalidade.PK_Nacionalidade
FULL OUTER JOIN DBO.Area_Oceanica_dist_costa ON DBO.Historico.Area_oceanica_dist_costa= DBO.Area_Oceanica_dist_costa.PK_AreaOc
FULL OUTER JOIN DBO.Local_Inspecao ON DBO.Historico.Local_inspecao= DBO.Local_Inspecao.PK_Local_Inspecao
FULL OUTER JOIN DBO.Especie ON DBO.Historico.Tipo_Pescado= DBO.Especie.PK_Especie

```

Figura 52 - Query realizada para extrair os dados da fonte de dados para o fato

Processo de Carregamento de Dados:



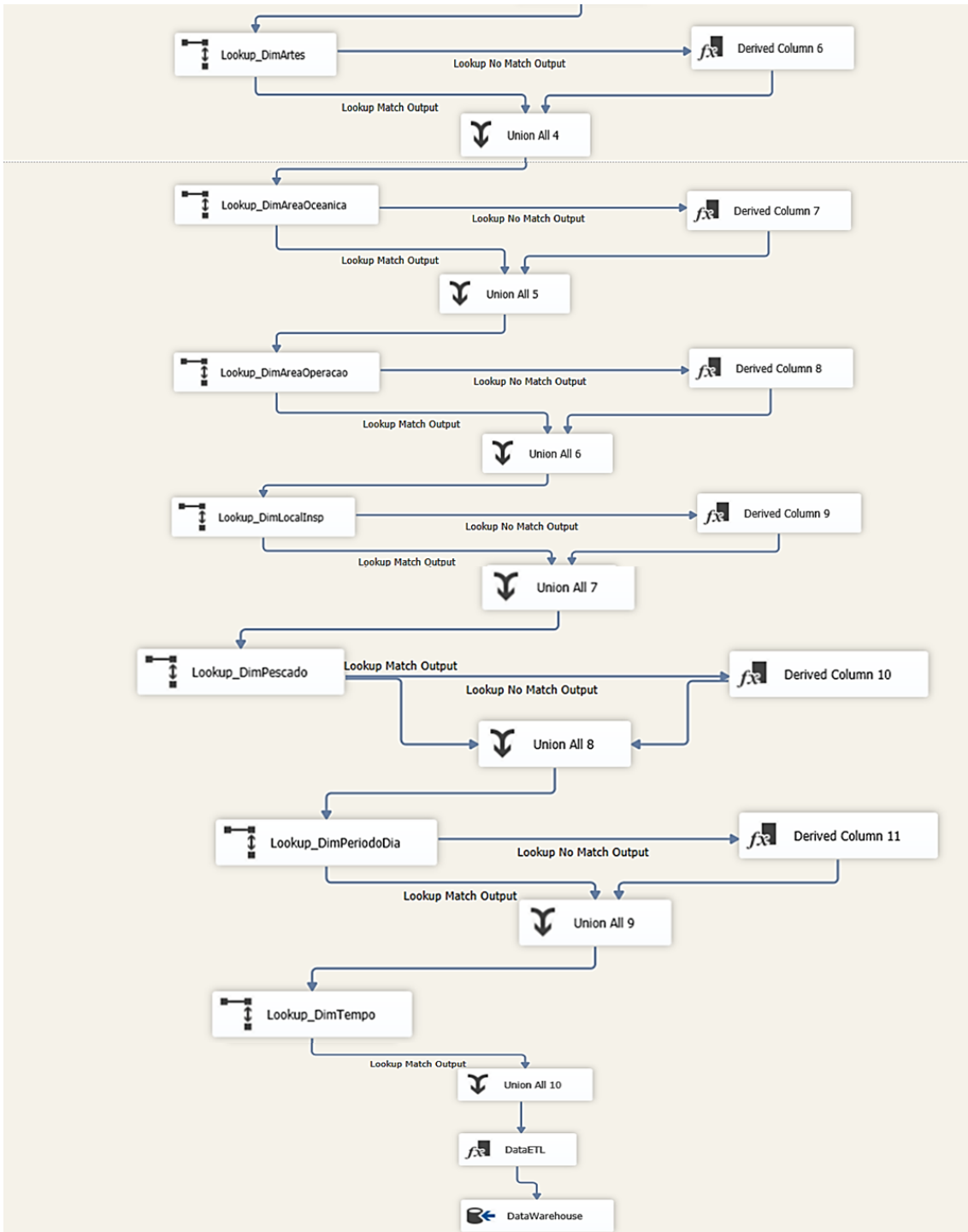


Figura 53 - Processo com as respectivas ferramentas utilizadas para o carregamento de dados na tabela do fato

Query para dados com campos *NULL*:

```
update FACT_FISCALIZACAO
SET QuantidadePescado = 0
where QuantidadePescado is NULL
go

update FACT_FISCALIZACAO
SET Infracao_Tipo1 = 0
where Infracao_Tipo1 is NULL
go

update FACT_FISCALIZACAO
SET Infracao_Tipo2 = 0
where Infracao_Tipo2 is NULL
go

update FACT_FISCALIZACAO
SET Infracao_Tipo3 = 0
where Infracao_Tipo3 is NULL
go

update FACT_FISCALIZACAO
SET Infracao_Tipo4 = 0
where Infracao_Tipo4 is NULL
go

update FACT_FISCALIZACAO
SET Infracao_Tipo5 = 0
where Infracao_Tipo5 is NULL
go

update FACT_FISCALIZACAO
SET Infracao_Tipo6 = 0
where Infracao_Tipo6 is NULL
go

update FACT_FISCALIZACAO
SET Infracao_Tipo7 = 0
where Infracao_Tipo7 is NULL
go

update FACT_FISCALIZACAO
SET Infracao_Tipo8 = 0
where Infracao_Tipo8 is NULL
go

update FACT_FISCALIZACAO
SET Infracao_Tipo9 = 0
where Infracao_Tipo9 is NULL
go

update FACT_FISCALIZACAO
SET Infracao_Tipo10 = 0
where Infracao_Tipo10 is NULL
go

update FACT_FISCALIZACAO
SET Infracao_Tipo11 = 0
where Infracao_Tipo11 is NULL
go

update FACT_FISCALIZACAO
SET Infracao_Tipo12 = 0
where Infracao_Tipo12 is NULL
go

update FACT_FISCALIZACAO
SET Infracao_Tipo13 = 0
where Infracao_Tipo13 is NULL
go

update FACT_FISCALIZACAO
SET Infracao_Tipo14 = 0
where Infracao_Tipo14 is NULL
go

update FACT_FISCALIZACAO
SET Hora = 0
where Hora is NULL
go

update FACT_FISCALIZACAO
SET NR_FISCREP= 0
where NR_FISCREP is NULL
go

update FACT_FISCALIZACAO
SET DataETL = 0
where DataETL is NULL
go
```

Figura 54 - Query realizada para transformar valores nulos em zero, no que diz respeito às medidas do fato

Apêndice I – *Dashboard* e relatórios complementares referentes à fiscalização da pesca

Dashboard com a informação relevante e sumarizada

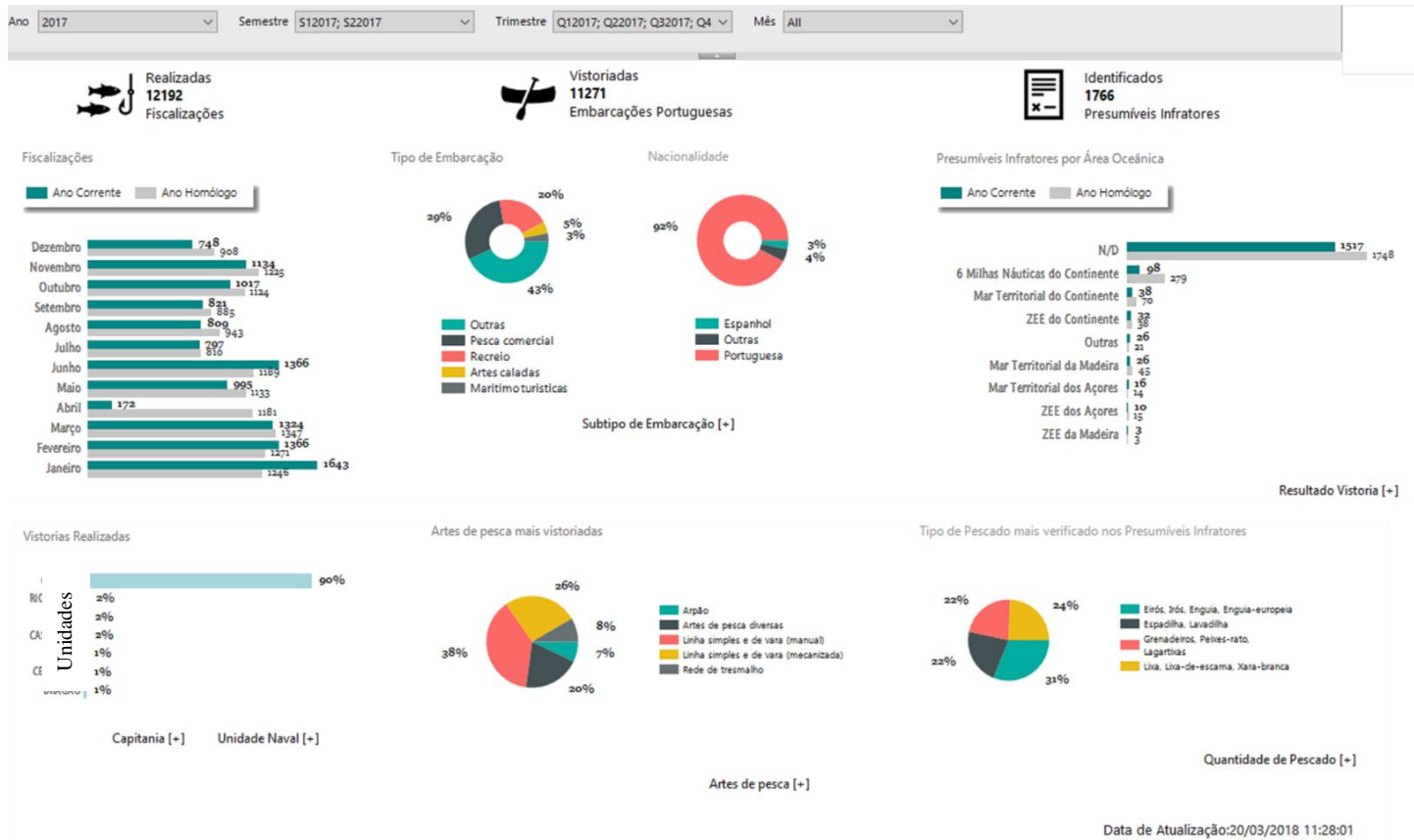
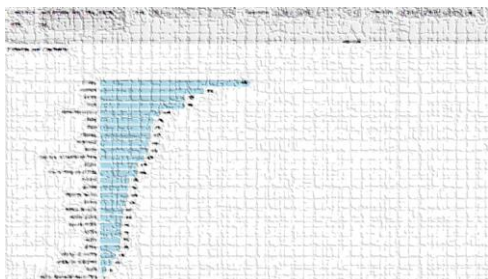


Figura 55 - *Dashboard* referente à fiscalização da pesca

Relatórios Complementares:

Os relatórios complementares estão apresentados em modo rasurado, com a finalidade de reserva estatística.

- Capitania



- Unidade Naval

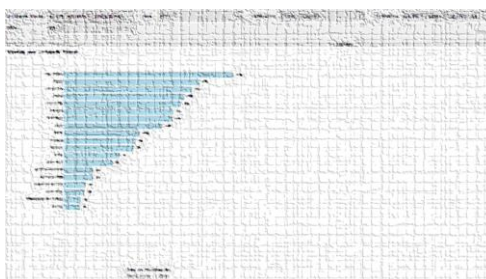
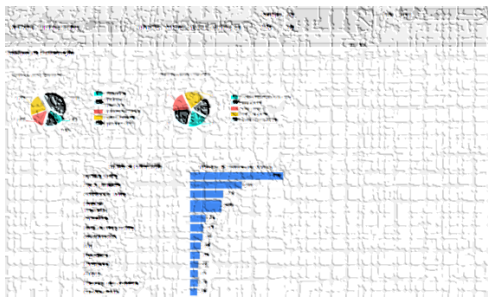


Figura 56 - Relatórios referente a fiscalizações realizadas por capitancias e por unidades navais

- Subtipo de Embarcação



- Artes de Pesca

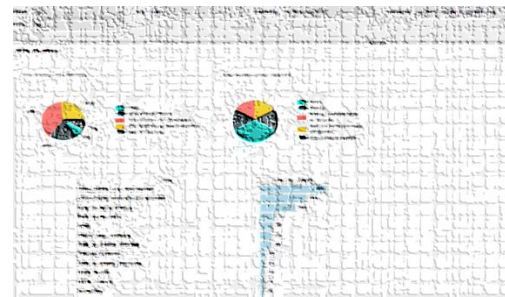


Figura 57 - Relatórios com informação complementar do subtipo de embarcação e das artes de pesca

- Resultados das Vistorias

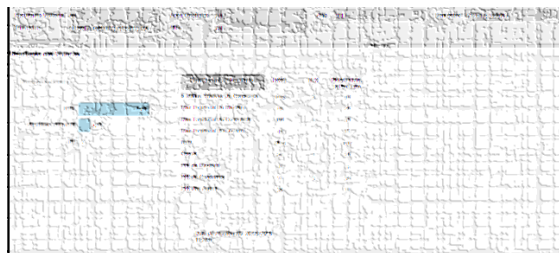


Figura 58 - Relatório referente ao resultado das vistorias

- Tipo e Quantidade de Pescado

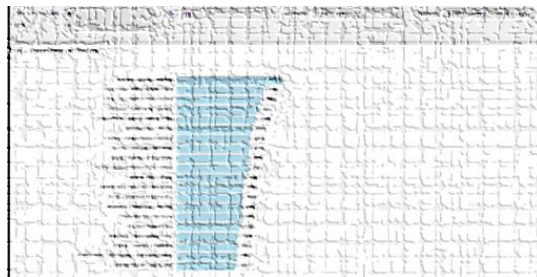


Figura 59 - Relatório com informação complementar do tipo e quantidade de pescado

Apêndice J – Inquérito para avaliação do artefacto produzido

Inquérito para avaliação de solução Business Intelligence para apoio à decisão na Atividade Operacional da Marinha Portuguesa

O presente questionário insere-se no âmbito da Dissertação de Mestrado da Aspirante Teles Machado, do Curso CMG Henrique Quirino da Fonseca (ano letivo 2017/2018).

Os dados fornecidos serão utilizados exclusivamente para fins académicos. É garantido o carácter anónimo dos dados recolhidos.

Agradeço a sua colaboração.

Objetivo do inquérito

A finalidade do questionário é a de validar a eficácia de uma solução informática construída, para suporte à tomada de decisão por parte das entidades responsáveis pelo planeamento e gestão operacional da Marinha, no âmbito da Fiscalização da Pesca.

Pretende-se reunir dados acerca da usabilidade (facilidade de utilização) e qualidade da informação disponibilizada pela solução, na tomada de decisão relativamente às ações de fiscalização da pesca da Marinha.

Instruções de preenchimento

O questionário inicia-se com um conjunto de perguntas que visam caracterizar o perfil dos respondentes e a sua familiaridade com a fiscalização da pesca.

Seguidamente, para avaliação da solução edificada, descreve-se um cenário para enquadrar as repostas ao questionário, num suposto processo de tomada de decisão. Ao responder, é importante que se imagine como responsável pela tomada de decisões, tendo por base a informação de enquadramento, que lhe será apresentada.

Note que, para atender às necessidades de informação segundo diferentes critérios, o sistema permite a seleção de vários filtros (por ano, semestre, trimestre, mês, unidade naval, artes de pesca, subtipo de embarcação, entre outros).

A. Dados Demográficos

Os seguintes itens destinam-se à recolha de dados demográficos gerais.

1. Género

- Feminino

- Masculino

2. Estatuto Militar

- Militar
- Militarizado
- Civil da Marinha
- Outro

3. Categoria

- Aspirante a Oficial
- Oficial Subalterno
- Oficial Superior
- Oficial General

B. Caracterização do Respondente

As seguintes questões destinam-se à recolha do perfil do respondente relativamente ao grau de conhecimento e experiência nas atividades de fiscalização da pesca.

1. Como classifica a sua ligação à atividade de fiscalização da pesca?

- Exerço funções de fiscalização da pesca
- Já exerci funções de fiscalização da pesca
- Fui objeto de ações de fiscalização da pesca
- Tenho formação para exercer funções, mas nunca exerci
- Nenhum conhecimento ou ligação à fiscalização da pesca

2. Quantas ações de formação em fiscalização já frequentou?

- Nenhuma
- 1 – 3
- 4 – 7
- Mais de 7

3. Como classifica o tipo das ações de formação em que maioritariamente participou?

- Teórica
- Teórica-Prática
- Prática
- Nenhuma

4. Qual o papel mais importante que desempenhou, nas ações de fiscalização em que participou?

- Responsável
- Adjunto

- Observador
- Nenhum

5. Como classifica o seu grau de conhecimento da legislação da pesca?

- Perito
- Médio
- Pequeno
- Nenhum

6. Como classifica o seu grau de conhecimento dos tipos de artes da pesca?

- Perito
- Médio
- Pequeno
- Nenhum

7. Como classifica o seu grau de conhecimento referente dos tipos de embarcações da pesca?

- Perito
- Médio
- Pequeno
- Nenhum

8. Como classifica o seu grau de conhecimento dos tipos de pescado nas águas nacionais?

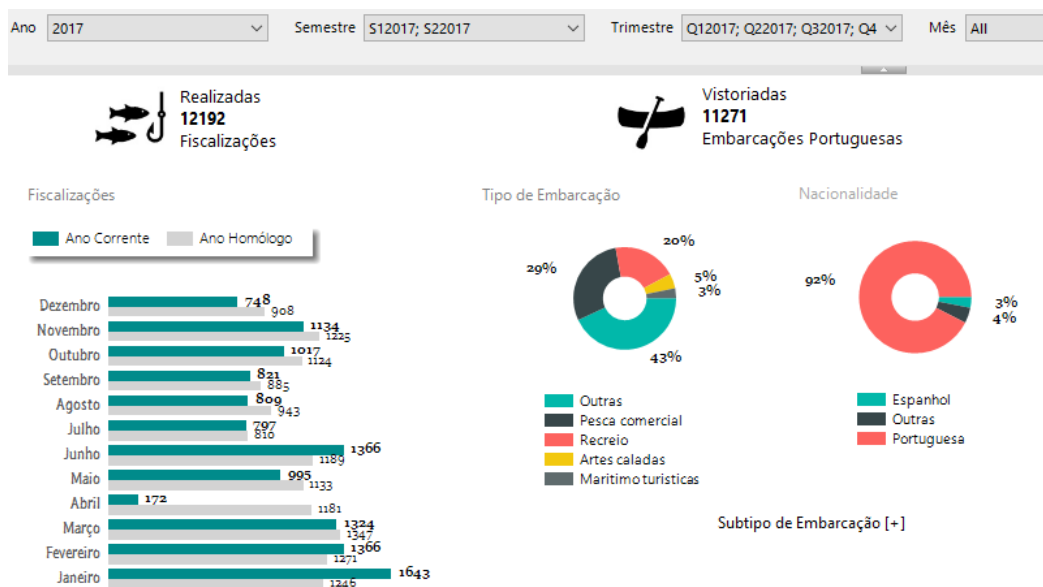
- Perito
- Médio
- Pequeno
- Nenhum

C. Leitura de Quadros

Esta secção é composta por quadros e gráficos produzidos pelo sistema edificado. O objetivo da secção é o de apurar se o sistema disponibiliza ao utilizador, informação de fácil leitura e interpretação.

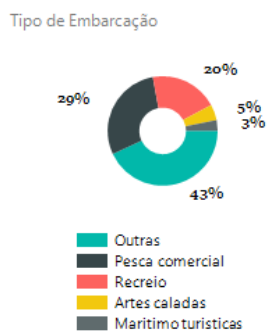
O quadro que se segue, sintetiza a atividade de fiscalização da pesca da Marinha, durante o ano de 2017. Tendo em conta a informação contida no quadro e supondo que o ano corrente é 2017, selecione a resposta que melhor responda à questão colocada.

Para cada questão, coloca-se uma parte da imagem abaixo para auxiliar a visualização entre o diagrama e a questão.




1. Em que meses, do ano corrente, se verificam o maior número de fiscalizações?

- Abril, Maio e Junho
- Janeiro, Maio e Novembro
- Janeiro, Fevereiro e Junho
- Não existe informação disponível

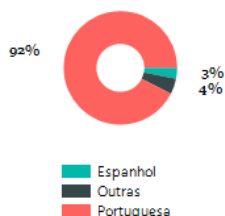


2. Quais os tipos de embarcação mais vistoriados?

- Outras, Pesca Comercial e Recreio
- Outras, Marítimo-Turísticas e Recreio
- Pesca Comercial, Artes Caladas e Recreio
- Não existe informação disponível

 Vistoriadas
11271
Embarcações Portuguesas

Nacionalidade



3. Por ordem crescente (da menor percentagem para a maior), quais as nacionalidades inspecionadas?

- Outras, Portuguesa e Espanhol
- Espanhol, Outras e Portuguesa
- Portuguesa, Outras e Espanhol
- Não existe informação disponível

4. Quantas embarcações de nacionalidade portuguesa foram vistoriadas?

- 11271 embarcações
- 10671 embarcações
- 12192 embarcações
- Não existe informação disponível

5. Face à informação apresentada, que decisão tomaria no sentido de melhorar a eficácia da fiscalização?

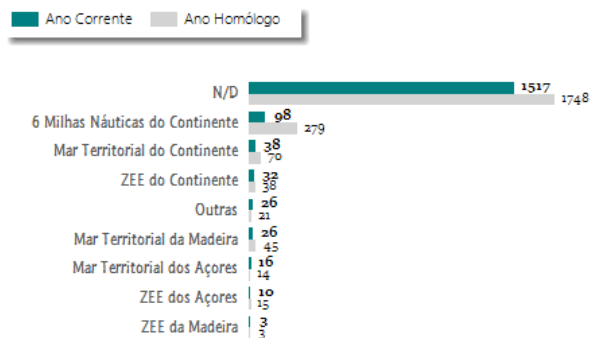
- Esforço da fiscalização nos meses e nos tipos de embarcação com menores vistorias realizadas
- Esforço da fiscalização nas artes caladas
- Não existe informação disponível para tomar a decisão

As figuras que se apresentam em seguida, mostram gráficos com informação também referente ao ano de 2017 (siglas utilizadas: N/D - Não Definido; ZEE - Zona Económica Exclusiva).



Identificados
1766
Presumíveis Infratores

Presumíveis Infratores por Área Oceânica



6. Quantos presumíveis infratores foram detetados ao longo do ano de 2017?

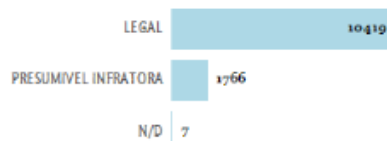
- 1766 presumíveis infratores
- 8769 presumíveis infratores
- 3491 presumíveis infratores
- Não existe informação disponível

7. No ano corrente (2017), em que áreas se verifica um maior número de presumíveis infratores?

- N/D e Mar Territorial do Continente
- N/D e 6 milhas náuticas do Continente
- N/D e ZEE do Continente
- Não existe informação disponível

Resultado Vistoria [+]

Resultado das Vistorias



Área Oceânica / Resultado	LEGAL	N/D	PRESUMIVEL INFRATORA
6 Milhas Náuticas do Continente	1095	2	98
Mar Territorial da Madeira	89		26
Mar Territorial do Continente	196	1	38
Mar Territorial dos Açores	43		16
N/D	8632		1517
Outras	251	1	26
ZEE da Madeira	8		3
ZEE do Continente	72	3	32
ZEE dos Açores	33		10

8. Qual o resultado de maior ocorrência nas vistorias?

- Legal
- Presumível Infratora
- Não Definido
- Não existe informação disponível

9. Se tivesse de definir uma área a ser vistoriada, selecionaria como critério o número de presumíveis infratores detetados numa determinada área?

- Sim, dependendo do tipo de infrações
- Definitivamente que não
- Não existe informação disponível para tomar uma decisão

D. Avaliação do Sistema Edificado

Em função do grau de facilidade que sentiu, na leitura dos gráficos e quadros anteriormente apresentados, avalie o sistema de Business Intelligence construído.

1. Como classifica genericamente, a adequação desta solução, para as atividades de planeamento e análise das operações de fiscalização da pesca?

- Excelente
- Boa
- Média
- Má

2. Como classifica esta solução, em termos da facilidade de visualização e interpretação da informação?

- Excelente
- Boa
- Média
- Má

3. No que diz respeito aos gráficos e à organização do painel (dashboard), que melhoria seria mais interessante introduzir?

- Sistema mais interativo (informação fornecida passo a passo)
- Seleção de paleta de cores
- Outros filtros
- Mais possibilidades de cruzamento de informação

4. A interface gráfica do sistema permite um adequado acesso à informação durante o processo de tomada de decisão?

- Sim
- Necessita de algumas melhorias
- Necessita de melhorias significativas
- Não

5. Qual a maior dificuldade com que se deparou?

- Analisar os gráficos
- Ler os quadros
- Tomar uma decisão com a informação existente

- Não senti dificuldade

6. O sistema fornece informação nas dimensões mais relevantes da tomada de decisão?

- Sim
- Faltam algumas dimensões de análise
- Faltam um número significativo de dimensões de análise
- Não

7. Como classifica o potencial da solução apresentada, no apoio à tomada de decisão das atividades de fiscalização da pesca?

- Excelente
- Bom
- Razoável
- Insuficiente

8. Considera que o sistema concebido tem limitações?

- Muito significativas
- Significativas
- Pouco significativas
- Não

9. Considera relevante a aplicação desta solução (automatização da recolha de dados provenientes de mensagens de texto formatado ou estruturado, o seu tratamento, análise e disponibilização de informação) em outras áreas de atuação (como por exemplo consumos, limitações operacionais, operações de busca e salvamento, informação para relatório de atividade operacional e anuários estatísticos, entre outras)?

- Sim
- Desde que os dados recolhidos não tenham erros
- A informação disponibilizada já é obtida por outras formas
- Não

10. Como classifica a utilização de um dashboard (painel que contém vários diagramas com a informação), como instrumento para a análise da informação e apoio à tomada de decisão?

- Excelente
- Boa
- Média
- Confusa

11. Como classifica a utilização de soluções automatizadas para apoio à tomada de decisão no contexto da Marinha?

- Excelente

- Boa
- Média
- Irrelevante

Agradecimento

Muito obrigada pela sua colaboração!

