



LISBON
SCHOOL OF
ECONOMICS &
MANAGEMENT
UNIVERSIDADE DE LISBOA

MESTRADO

MÉTODOS QUANTITATIVOS PARA A DECISÃO ECONÓMICA E EMPRESARIAL

TRABALHO FINAL DE MESTRADO

RELATÓRIO DE ESTÁGIO

A VISUALIZAÇÃO DE DADOS NAS LOTAS DA DOCAPESCA

BERNARDO SIMONS DE ALMEIDA

OUTUBRO - 2020



MESTRADO
MÉTODOS QUANTITATIVOS PARA A
DECISÃO ECONÓMICA E EMPRESARIAL

TRABALHO FINAL DE MESTRADO

RELATÓRIO DE ESTÁGIO

A VISUALIZAÇÃO DE DADOS NAS LOTAS DA DOCAPESCA

BERNARDO SIMONS DE ALMEIDA

ORIENTAÇÃO:

PROF. CARLOS J. COSTA

FILOMENA SARAIVA

OUTUBRO - 2020

Agradecimentos

Em primeiro lugar, quero agradecer ao ISEG por toda a minha formação enquanto académico e pessoa e em especial às coordenadoras do Mestrado de MQDEE, pelo apoio e incentivo durante estes três anos.

Ao professor Carlos J. Costa agradeço toda a preocupação, as sugestões, conselhos e o apoio e compreensão demonstrados durante todo o processo da realização do Trabalho Final de Mestrado.

À Docapesca, em especial à Dr.^a Filomena Saraiva por toda a disponibilidade revelada, assim como todos os conselhos e pela confiança depositada em mim.

Agradeço também aos meus colegas que partilharam uma sala comigo durante os meses em que realizei o estágio pelos conselhos e companheirismo, assim como a integração na empresa. Ao Filipe, à Carla, ao Rui, à Conceição, à Helena e à Cláudia, os meus mais sinceros obrigados.

Quero agradecer ao Departamento de Controlo Orçamental da Docapesca pela oportunidade que me foi dada no final do meu estágio e por todo o apoio que me deram durante a elaboração deste relatório e por toda a compreensão e amizade. Obrigado Sizalda e Margarida.

Aos meus companheiros do futebol, que me acompanharam durante todo o meu percurso académico, quero agradecer todo o suporte e motivação para que conseguisse ultrapassar todos os obstáculos que surgiram na minha vida académica e pessoal.

Ao José Raposo, João Azeredo, Vasco Lopo, Diogo Oliveira, Pedro Sousa e ao Daniel Tomás por tudo o que me fizeram por mim nesta caminhada que já tem quase 20 anos de história. Sem eles, nada disto seria possível e não estaria aqui hoje.

Ao João Pedro, que é como um filho, que sempre acreditou em mim e que me motivou para que fosse além das minhas capacidades para ultrapassar todos os desafios mais importantes da minha vida.

À minha família, por tudo o que representam, especialmente aos meus pais, avós paternos e ao meu irmão por terem sido sempre aqueles que me empurraram

nos bons e nos maus momentos e que nunca desistiram de mim mesmo quando nem eu acreditava.

Finalmente, quero agradecer à Leonor por todo o apoio e por me ter posto sempre no topo das suas prioridades.

Sem vocês, nada disto seria possível.

Resumo

Na segunda década do século XXI, assistiu-se a um crescimento exponencial dos dados criados diariamente. Consequentemente, a visualização de dados é cada vez mais relevante no contexto empresarial, revelando-se cada vez mais essencial para a tomada de decisões e planeamento de estratégias de negócio.

Este trabalho foi realizado, em conjunto com a Docapesca, com a finalidade de criar um *dashboard* capaz de analisar os dados específicos relativos ao pescado transacionado nas lotas em Portugal num determinado período (Entre 2009 e 2018). Com uma abordagem metodológica orientada para a *data visualization*, foi possível elaborar um *dashboard* com as propriedades indicadas para o que foi solicitado.

Foi possível implementar a revisão da literatura, criando um *dashboard* atendendo ao que foi referido anteriormente.

Palavras-chave: *Dashboard, Business Intelligence, R, Rdata, Package, Dados, Pescado*

Abstract

The second decade of the 21st century saw an exponential growth in the data created daily. Consequently, data visualization is increasingly relevant in the business context, proving to be increasingly essential for decision making and business strategy planning.

In association with Docapesca, this work was carried out with the purpose of creating a dashboard capable of analyzing the specific data related to fish traded at auction in Portugal in a given period (Between 2009 and 2018). With a methodological approach oriented to data visualization, it was possible to elaborate a dashboard with the properties indicated for what was requested.

Despite not being implemented, the final assessment was positive and allowed conclusions to be drawn that were correct in relation to the data under study.

Índice

| | |
|---|----|
| Agradecimentos..... | 4 |
| Resumo | 6 |
| Abstract | 7 |
| Índice..... | 8 |
| Índice de Figuras | 9 |
| 1. Introdução..... | 10 |
| 1.1. Enquadramento..... | 10 |
| 1.2. Contexto Organizacional: Docapesca – Portos e Lotas, S.A | 10 |
| 1.3. Problema/Objetivo..... | 11 |
| 1.4. Abordagem Metodológica..... | 11 |
| 1.5. Estrutura do Relatório de Estágio..... | 12 |
| 2. Revisão da Literatura..... | 13 |
| 2.1. Business Intelligence e Business Analytics | 13 |
| 2.2. Processo de Business Intelligence | 14 |
| 2.3. Visualização de Dados | 15 |
| 2.4. Dashboards..... | 17 |
| 3. Proposta Conceptual | 19 |
| 3.1. Arquitetura da solução..... | 19 |
| 3.2. Modelo de processos de BI utilizado | 19 |
| 4. Trabalho empírico | 21 |
| 4.1. Perceber o negócio | 21 |
| 4.2. Preparação e processamento dos dados | 21 |
| 4.3. Desenvolvimento..... | 22 |
| 4.4. Visualização de dados..... | 26 |
| 4.5. Resultados e Interpretação dos dados | 27 |
| 5. Conclusões e trabalhos futuros..... | 36 |
| 6. Bibliografia | 38 |

Índice de Gráficos

| | |
|--|----|
| Gráfico 1 – Kgs de biqueirão pescados entre 2009 e 2018 com arte de pesca - cerco | 27 |
| Gráfico 2 – Kgs de carapau pescados entre 2009 e 2018 com arte de pesca - cerco..... | 28 |
| Gráfico 3 – Kgs de cavala pescados entre 2009 e 2018 com arte de pesca - cerco | 29 |
| Gráfico 4 – Kgs de sardinha pescados entre 2009 e 2018 com arte de pesca - cerco | 29 |
| Gráfico 5 – Kgs de pescada-branca pescados entre 2009 e 2018 com arte de pesca - polivalentes..... | 30 |
| Gráfico 6 – Kgs de polvo-vulgar pescados entre 2009 e 2018 com arte de pesca - polivalentes | 31 |
| Gráfico 7 – Kgs de choco-vulgar pescados entre 2009 e 2018 com arte de pesca - polivalentes | 32 |
| Gráfico 8 – Kgs de amêijoia-branca pescados entre 2009 e 2018 com arte de pesca - polivalentes..... | 33 |

Índice de Figuras

| | |
|---|----|
| Figura 1- Modelo genérico de processo de Business Intelligence (Fonte: Vuori,2006) | 14 |
| Figura 2- Características dos diferentes tipos de dashboards (Fonte: autor)..... | 18 |
| Figura 3 - Arquitetura do sistema e sua implementação (Fonte: autor)..... | 19 |
| Figura 4 - Modelo BI utilizado para este problema | 20 |
| Figura 5 - Ficheiro Excel com os dados sobre a pesca de Arrasto em 2009 do grupo de pesca A1 | 23 |
| Figura 6 - Ficheiro ArrastoA1 | 24 |
| Figura 7 – Função de importação de dados | 24 |
| Figura 8 - Data..... | 25 |
| Figura 9 - Rstudio | 25 |
| Figura 10- Homepage do Dashboard..... | 26 |
| Figura 11 – Localização geográfica e dados das lotas | 34 |

1. Introdução

1.1. Enquadramento

Com a informatização dos serviços, e com o aumento exponencial de dados, assistiu-se na segunda década do sec. XXI a um aumento da informação numa base diária (Cavique, 2014). Segundo a IBM, em 2017 eram criados cerca de 2.5 quintilhões de bytes de data diariamente e que 90% dos dados do mundo foram criados entre 2015 e 2017 (Canabarro, 2017). Em consonância, um artigo publicado na Forbes afirma que em 2020 será gerada cerca de 1.7 Megabytes de informação por segundo e por pessoa. É cada vez mais frequentemente ouvirmos falar de termos como Big Data e Data Science. Big Data está associado à mudança de escala no volume de dados e na sua taxa de atualização. Em 2012, um estudo conclui que o valor estimado de informação no planeta era de 2.8 zetabytes. Adicionalmente, permitiu a criação da profissão considerada pela Harvard Business Review como sendo a mais sexy do século XXI: Data Scientist – alguém melhor em estatística que um engenheiro informático e melhor em programação que um matemático.

1.2. Contexto Organizacional: Docapesca – Portos e Lotas, S.A

A 27 de Março de 1990, nos termos do Decreto-Lei nº 107/90, foi constituída uma Sociedade Anónima de capitais integralmente do Estado denominada DOCAPESCA – PORTOS E LOTAS, SA, porém, o nome Docapesca remonta a 1956, data da publicação do Decreto-Lei nº 40764, que determinou que a concessão da exploração da Doca de Pesca de Pedrouços seria realizada por uma sociedade anónima de responsabilidade limitada.

Atualmente, o site oficial da Docapesca apresenta a empresa como pertencente ao setor empresarial do Estado, tutelada pelo Ministério do Mar, que tem a seu cargo no continente português o serviço público da prestação de serviços de Primeira Venda do Pescado, assim como o apoio ao Setor da Pesca e respetivos portos.

A Docapesca está territorialmente dispersa por Portugal Continental, estando sedeadada em Lisboa, e sendo composta por cinco delegações: Norte e Matosinhos, Centro Norte, Centro, Centro Sul e Sul.

A missão da Docapesca passa pela prestação do serviço público da primeira venda de pescado em lota e atividades conexas, assim como a administração dos portos de pesca e marinas de recreio e as funções de autoridade portuária nas áreas sob sua jurisdição.

A atividade geral das seis delegações passa pela promoção de venda de pescado fresco e refrigerado, bem como o gelo, o controlo higiossanitário, a gestão do mercado de segunda venda de Matosinhos e a gestão dominial das áreas sob sua jurisdição.

O cumprimento da sua missão pública depende da concretização dos seus 13 objetivos estratégicos, enquadrados por 4 eixos de intervenção: Desenvolvimento sustentável e inovação; Modernização, qualidade e segurança; Desempenho organizacional; desempenho económico e financeiro.

Os 2 primeiros eixos estão direcionados para o negócio da Docapesca, os restantes estão orientados para a componente interna da Docapesca – o alinhamento institucional e o desempenho financeiro.

1.3. Problema/Objetivo

O presente relatório de estágio surge no âmbito de um protocolo entre a Docapesca e o Instituto Superior de Economia e Gestão da Universidade de Lisboa.

O objetivo deste trabalho consiste em propor uma solução que permita a visualização de dados fornecidos pela empresa, bem como a sua análise, que permita à empresa tirar conclusões ótimas para a tomada de decisão.

A solução corresponde ao desenvolvimento de um *dashboard*, através da linguagem de programação *R*, que reúna as características mais indicadas para a informação em estudo.

1.4. Abordagem Metodológica

Em linguagem corrente, a palavra informação pode ter vários significados diferentes, nomeadamente data, conhecimento e até mesmo inteligência. A abordagem académica faz a distinção entre estes conceitos de acordo com o seu conteúdo, realçando, porém, que a interpretação destes conteúdos é ambígua e está dependente da interpretação do autor (Vuori, V. 2006).

Assim, a abordagem metodológica utilizada para a elaboração deste relatório de estágio, segmentou-se em 5 fases:

A primeira fase consistiu na compilação de dados fornecidos pela Docapesca para o estudo do pescado;

A segunda e terceira fase incidiram em 2 momentos de uma revisão bibliográfica: a primeira parte sobre os vários pontos fulcrais para a elaboração de um dashboard apropriado para os dados em estudo, nomeadamente a sua: definição, especificidades, arquitetura, design e vantagens. A segunda parte teve como foco a linguagem de programação R, mais concretamente na elaboração de um dashboard através da mesma.

A quarta fase, a introdução dos dados compilados no dashboard criado para a empresa, que reúne todos os requisitos atrás mencionados, e que permita à empresa uma otimização na tomada de decisão.

Finalmente, a quinta e última fase corresponde às conclusões do caso de estudo específico – a transação de pescado nas lotas portuguesas.

1.5. Estrutura do Relatório de Estágio

O presente relatório é composto por sete capítulos. O presente capítulo é a introdução. O segundo capítulo a revisão da literatura. No terceiro capítulo é abordado a proposta a solução conceptual, resultante fundamentalmente da revisão da literatura. No quarto capítulo é descrito o trabalho empírico, que corresponde à aplicação da proposta concetual no caso concreto. Finalmente, o quinto capítulo contém a conclusão e trabalhos futuros.

2. Revisão da Literatura

2.1. Business Intelligence e Business Analytics

A utilização de ferramentas de *Business Intelligence* (B.I) remonta à década de 60, na altura com a designação de *Decision Support System* (sistema de apoio à decisão). O termo B.I foi popularizado por Howard Dresner, do grupo Gartner – empresa de consultoria responsável pelo desenvolvimento de tecnologia e ferramentas indispensáveis para a tomada de decisão - em 1989, e atualmente são uma ferramenta essencial para as empresas (Power, 2010, Caetano & Costa, 2014). Boris Evelson designou B.I como sendo um conjunto de metodologias, processos, arquiteturas e tecnologias, responsáveis pela transformação de dados brutos em informação relevante para a tomada de decisão. (Evelson,2008). O B.I traduz-se através de *scorecards*, *dashboards* e relatórios (Evans & Lindner, 2012).

Por sua vez, Business Analytics (B.A) corresponde ao processo de transformação de dados em ações através de análise de dados, permitindo às empresas uma tomada de decisão mais eficiente e eficaz nas mais diversas áreas do contexto empresarial. A análise realizada pelas ferramentas B.A permite a automatização e consequente otimização dos seus processos produtivos, através de informações como: sazonalidade, tendências e outros padrões (Liberatore & Luo 2010, Tomás, et al., 2018). O B.A é reproduzido através de análises quantitativas, descritivas e métodos preditivos (Evans & Lindner, 2012, Marques & Costa, 2019).

2.2. Processo de Business Intelligence

Na literatura é referida a existência ao longo dos anos de diferentes processos de BI, contudo com objetivos semelhantes. A evolução dos processos de BI mostrou que a sua base assenta nos seguintes pontos: análise das necessidades de informação, aquisição de informação, armazenamento de informação, distribuição de informação e utilização de informação. Ainda é usual que a última fase do processo de business intelligence volte à origem do processo criando um ciclo de melhoria contínua (Vuori, 2006).

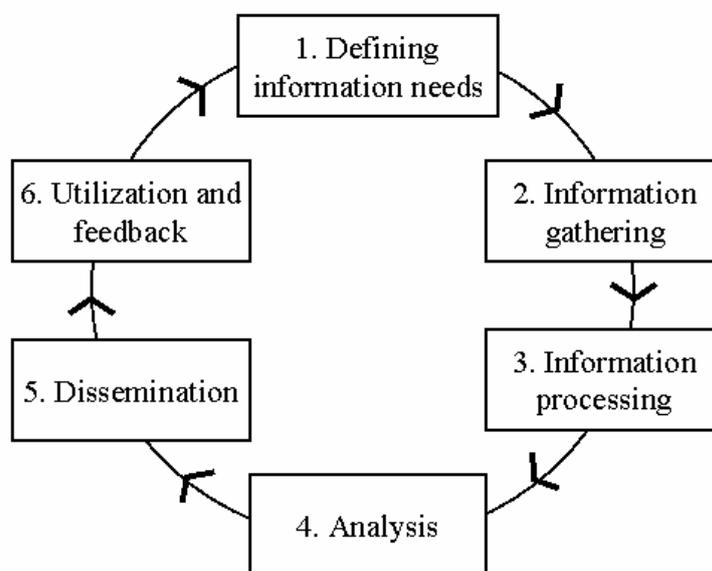


Figura 1- Modelo genérico de processo de Business Intelligence (Fonte: Vuori,2006)

A análise das necessidades de informação é o ponto crucial para a uma boa prática de BI. Com uma definição adequada da informação e o auxílio dos gestores/decisores na definição dos objetivos faz com que o processo seja ágil e efetivo para a organização (Vuori, 2006).

Os investigadores têm proposto diversos modelos de ciclo de vida para apoiar no desenvolvimento de sistemas de BI. Especificamente, no que diz respeito ao Business Analytics têm sido apresentados muitas propostas (Costa & Aparício, 2020).

2.3. Visualização de Dados

Em 2004, Stephen Few referiu que para se trabalhar com visualização de dados (*data visualization*) é necessário saber fazer a distinção entre dados categóricos e dados quantitativos, percebendo assim a melhor maneira de consolidar ambos os dados. Enquanto que os dados quantitativos correspondem ao valor dos objetos de estudo, os dados categóricos identificam aquilo a que está a ser atribuído valor (O primeiro conjunto responde à pergunta “Quanto?” e o segundo à pergunta “O quê?”). Few também realça a importância de determinados atributos na visualização de dados, nomeadamente: cor, forma, tamanho, orientação, comprimento das linhas e localização a duas dimensões. Destes todos, apenas os dois últimos atributos permitem codificar visualmente (visual encoding) valores quantitativos eficientemente.

Atendendo a estas noções introdutórias, é possível passar à “mensagem quantitativa” que queremos transmitir. Das sete existentes (comparação nominal, séries temporais, ranking, parte de um todo, desvio, distribuição de frequências e correlação), para os dados fornecidos, pelas suas várias limitações e por serem observações ao longo do tempo, foi decidido apresentá-los em formato de séries temporais, com todas as especificidades adjacentes: a utilização de gráficos de linhas, por estarmos a lidar com uma panóplia de dados (espécies diferentes de pescado, distribuídos por diversas lotas) por um período de 10 anos e colocar o tempo no eixo das abcissas.

Ciência de dados pode ser definido sendo a ciência que incide numa área de conhecimento, combinando estatística e tecnologias de Informação (Aparicio, Aparicio & Costa, 2019). Apesar de ser um termo bastante moderno, Cavique (2014) afirma que esta ciência remonta aos anos 70 sob o nome Sistemas de Apoio à Decisão (DSS), que posteriormente se passou a apelidar de Sistemas de Apoio a Executivos (EIS) nos anos 80, reinventando-se mais duas vezes: *Online Analytical Processing* (OLAP) na década de 90 e a anteriormente referida B.I na entrada do novo milénio.

Com este boom de informação, a análise do mesmo passou a ter maior importância: a visualização de dados, outrora um meio para outras áreas

consideradas relevantes, como a astronomia, economia, matemática e até mesmo a geografia, adquiriu uma luz própria no campo de data.

Primeiramente, e atendendo ao crescimento acentuado de informação que se vem observando recentemente, é preciso ter uma noção atual do que significa a visualização de dados. Visualização de dados corresponde à representação gráfica de data e informação gerados por ferramentas informáticas designada de *data visualization tools* (Bikakis, 2018, Aparicio & Costa, 2015). Essas ferramentas que auxiliam os utilizadores na sua análise, através de meios intuitivos para a exploração e análise de dados, permitindo-lhes identificar padrões eficientemente interessantes, assim como inferir causalidades e correlações. A visualização de dados pode ser definida como sendo a representação visual de dados através de diagramas e gráficos englobando desde gráficos simples e estáticos até interações complexas, mas sempre com a finalidade de visualizar dados (Fry, B. 2015, Aparicio & Costa, 2015).

Num estudo social acerca das diferentes apreciações por parte de diversos utilizadores, cujo feedback não foi consensual: enquanto alguns elogiavam a simplicidade e facilidade na apreensão dos dados, outros consideravam os gráficos e diagramas demasiado confusos e coloridos para que se toma uma decisão de qualquer cariz (Bikakis, 2018).

Seguidamente, é necessário fazer uma contextualização histórica da visualização de dados. Friendly (2008) elaborou um cronograma com os diferentes períodos da visualização de dados, remontando aos meados do século XVI com o aparecimento dos primeiros mapas e diagramas. Desde 1975 que se está a viver no período da visualização de dados dinâmica e interativa, muito devido ao surgimento da Big Data. Os sistemas de informação sofreram desenvolvimento (e aperfeiçoamento) permitindo lidar com a volume, variedade e velocidade de produção e processamento de dados (Costa & Aparicio, 2019).

Apesar da qualidade dos dados não ser impeditivo na elaboração de um Dashboard, é um pré-requisito para o sucesso dos mesmos e uma condicionante para a sua utilização (Yigitbasioglu & Velcu, 2012).

Relativamente à apresentação da informação, a mesma consiste numa conjugação destas 3 premissas (Costa & Aparicio, 2019):

- Evitar saturar o utilizador com demasiada informação;
- Destacar informação fulcral;
- Apresentação simples e concisa.

Também é importante identificar os gráficos mais adequados. Primeiro, é essencial esclarecer o objetivo e identificar o público. O objetivo que se pretende atingir ao mostrar os dados poderá ser evidenciar comparação, relacionamento, composição ou distribuição. Depois, é essencial identificar o número de variáveis e a representação do tempo (Costa & Aparicio, 2019).

2.4. Dashboards

Para proceder à elaboração de um *dashboard*, é necessário primeiro ter uma noção do mesmo, assim como dos elementos que o constituem. definem um *dashboard* como sendo uma aplicação que, através de gráficos visuais simples como tabelas, gráficos e medidores num web browser. Os *dashboard* destacam-se pelo vasto número de métricas diferentes incorporadas, pelos seus sumários bastante detalhados e precisos, tal como a facilidade de compreensão da informação que é apresentada. Stephen Few aborda o tópico de uma maneira mais generalizada, defendendo que os *dashboards* se caracterizam como sendo uma forma de exibição ou um estilo de apresentação (de dados) e não um tipo específico de informação ou tecnologia. Num automóvel, o *dashboard* serve o propósito de monitorizar o seu funcionamento, a informação disponibilizada serve um determinado objetivo, o que deve ser o foco da aplicação em si (Few, 2004, Aparicio & Costa, 2015).

Os *dashboards* têm diversas utilidades, nomeadamente: Definir metas e categorizar informação; incentivar a realização de determinadas actividades de uma maneira cronológica; alertar para potenciais riscos e anomalias; definir objetivos para os diferentes sectores de atividade de uma empresa e providenciar uma User Interface (U.I) estandardizada com vista à interação e análise de data (Costa & Aparicio, 2019).

Para que seja possível criar um dashboard capaz de satisfazer estas utilidades, segundo Pauwels et al, (2009), o mesmo poderá ser direcionado para um dos

seguintes 4 propósitos: Monitorização, consistência, planeamento e comunicação.

Considerada como uma das funções mais importantes de um *dashboard*, a monitorização de performance corresponde à avaliação de métricas que poderá resultar ou não em ações corretivas. Um *dashboard* também poderá incluir planeamento se entre as suas características estiver incluído a análise de cenários. O processo de comunicação de um *dashboard* refere-se à informação que é transmitida aos *stakeholders* de uma empresa na forma de performance e valores, através das métricas indicadas para tal (Pauwels et al, 2009).

Existem três tipos de *Dashboards*: Estratégicos, Operacionais e Táticos. Cada um dos tipos destina-se para os diferentes tipos de informação. Cada tipo enfatiza as camadas superior, intermédia e inferior de informação, assim como a diferente aplicação das mesmas. A superior exhibe condições esperadas, graficamente; a intermédia permite aos utilizadores conjugarem a informação de diferentes sectores de uma organização, e a inferior permite a análise de transacções individuais, assim como de relatórios operacionais (Ikechukwu et al., 2012).

| | Operacional | Tático | Estratégico |
|--------------|-----------------------|----------------------------------|---------------------|
| Objetivo | Monitorizar Operações | Otimizar processos/procedimentos | Estratégias |
| Enfâse | Monitorização | Análise | Gestão |
| Utilizadores | Supervisores | Gestores | Executivos |
| Área/Sector | Operacional | Departamental | Empreendedorismo |
| Informação | Detalhada | Detalhada/Resumida | Resumida |
| Atualizações | Diárias | Diárias/Semanais | Mensais/Trimestrais |

Figura 2- Características dos diferentes tipos de dashboards (Fonte: autor)

Uma abordagem possível é ainda agrupar toda esta informação de forma a fazer sentido globalmente. É o que acontece com o Balance Score Card. Aqui com frequência também existem desafios a ser ultrapassados (E.g. Costa & Aparicio, 2005).

3. Proposta Conceptual

3.1. Arquitetura da solução

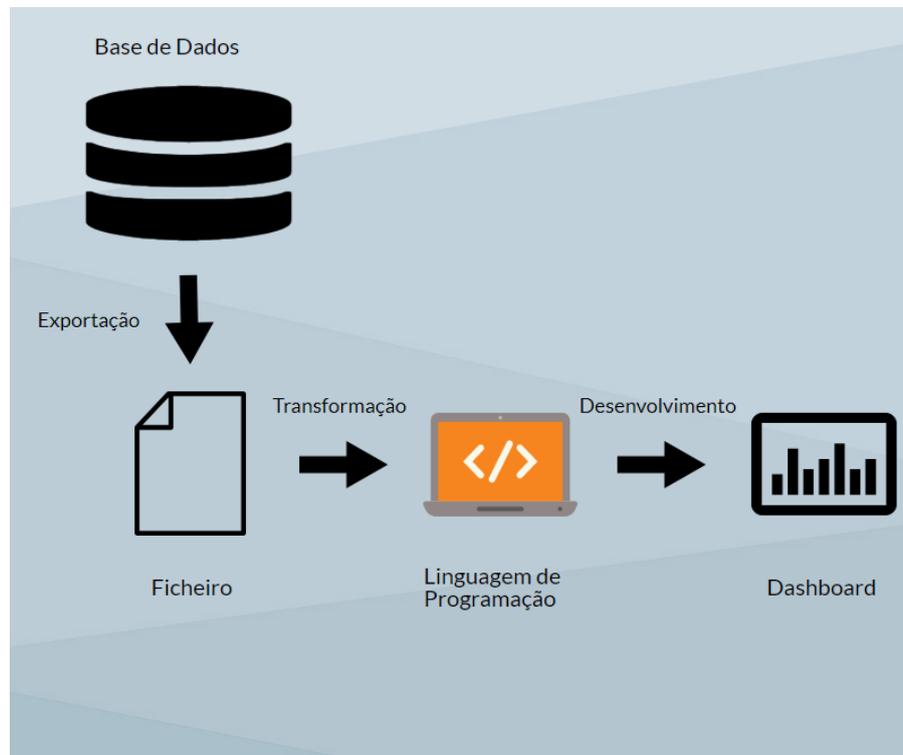


Figura 3 - Arquitetura do sistema e sua implementação (Fonte: autor)

A proposta de solução passa pela extração dos dados diretamente da base de dados para um ficheiro em qualquer formato (neste caso xls), seguido de uma transformação do mesmo para ser trabalhado numa linguagem de programação (neste caso R) e finalmente desenvolvido num dashboard.

3.2. Modelo de processos de BI utilizado

Foi desenvolvido um modelo de processos BI que tem por base o ciclo de melhoria contínua e que contempla os seguintes passos:

- Perceber o negócio e os dados – fase inicial, de cariz mais introdutória, que se foca numa assimilação dos valores, missão e estratégia de determinada empresa e de que maneira funciona o seu *core-business*;
- Preparação e processamento dos dados - analisada o tipo de informação disponibilizada e planeia-se a melhor abordagem para a respetiva análise e as ferramentas a utilizar para se atingir esse fim. É nesta fase que se

decide, por exemplo, se se usa uma linguagem de programação e qual a mais indicada ou o tipo de ferramentas para análise;

- Desenvolvimento - considerada a mais importante no processo, consiste no desenvolvimento e consequente modelação daquilo que foi realizado ao longo das restantes fases;
- Visualização dos dados;
- Resultados e interpretação dos dados - Nas duas últimas fases, procede-se à visualização dos dados e as respetivas conclusões. Assim, cabe aos responsáveis pela análise dos dados tomarem decisões estratégicas que visem o desenvolvimento e o progresso.

Existe uma repetição periódica deste processo, essencial neste ciclo de melhoria contínua.

Segue abaixo o modelo proposto, inspirado na revisão bibliográfica e que contempla as seguintes fases:

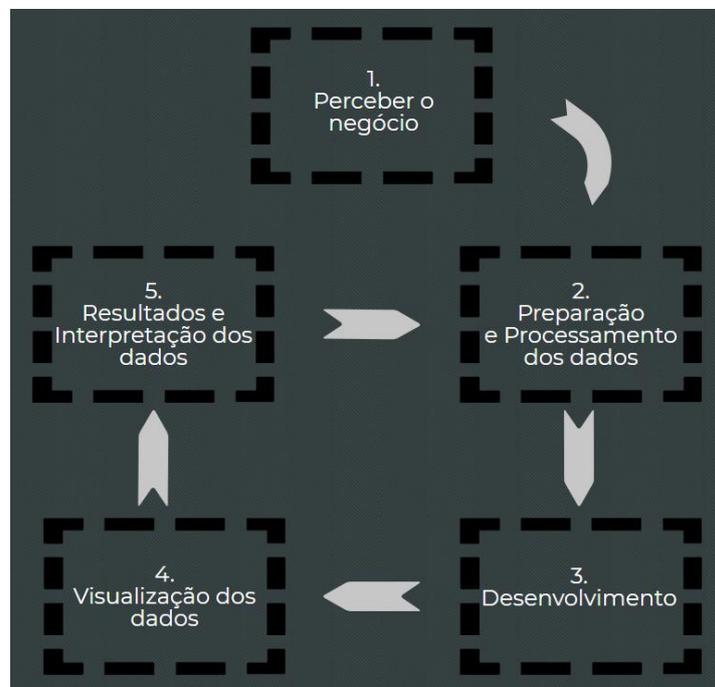


Figura 4 - Modelo BI utilizado para este problema

4. Trabalho empírico

4.1. Perceber o negócio

Primeiramente, é importante fazer uma contextualização do problema antes de o resolver. Para isso foi necessário fazer uma investigação relativamente à Docapesca e ao seu negócio de transações de pescado em lota.

4.2. Preparação e processamento dos dados

Em segundo lugar, foram analisados dois grupos de pescado – A1 e A2 (no primeiro grupo, tem-se em conta o tamanho e a frescura do pescado, apesar do segundo dado ser irrelevante para a análise dos mesmos), através de 3 artes de pesca diferentes (Cercos, Arrasto e Polivalente) no período entre 2009 e 2018 nas 22 lotas do país com número de controlo veterinário. Esta foi a resolução escolhida para lidar com o problema pois era uma solução que agradava a ambas as partes: à empresa, porque os dashboards são cada vez mais indispensáveis na avaliação de performance empresarial e à faculdade porque apesar de ser um tema complexo e pouco fundamentado no mestrado, permite disponibilizar dados de uma maneira clara, bem estruturada e apelativa que, conjugado com o R e a sua panóplia de funções, terá os resultados desejados. Nota: foi utilizado o *Rstudio*, uma versão mais *user-friendly* e avançada do R (*window docking* – todas as janelas relevantes estão disponibilizadas para o utilizador); a possibilidade de “navegar” entre gráficos feitos anteriormente; possui um editor de texto que facilita a parte da sintaxe de uma linguagem de programação; o software é compatível com todas as plataformas e funciona da mesma maneira, sendo assim possível passar ficheiros de *Windows* para *Mac* e vice-versa e carregando no *Tab*, o *Rstudio* irá completar o argumento por nós, o que poupará bastante tempo e erros de sintaxe).

Para os ficheiros Excel com mais de um separador foram uniformizadas primeiras linhas dos respetivos separadores, que correspondiam aos nomes das colunas, para ser possível extrair os dados para R através do *package readyXL*. Como são 30 ficheiros (10 anos e 3 artes de pesca), foi criada uma função “importar” e depois aplicada para os mesmos. Em R, as funções correspondem a um conjunto de instruções que queremos utilizar repetidamente. O R não só

contém uma data de funções pré-definidas, como também é possível criar as nossas próprias funções.

Seguidamente, aglutinaram-se todos os ficheiros extraídos num só através da função *rbind*, que combina funções, matrizes ou data frames por linhas, e removeram-se todos os ficheiros “separados” pois a partir daquele momento de nada nos serviriam utilizando a função *rm*.

Alterou-se também o nome da primeira coluna, de “DD-MM-AAAA” para “Data” apenas por uma questão de preferência, criou-se uma coluna extra com os preços médios dividindo os euros pelos quilogramas de cada entrada de peixe nas lotas, fazendo um arredondamento de 2 casas decimais.

Finalmente, criaram-se as colunas ano e mês para posteriormente tornar os dados mensais. Anteriormente com os dados diários, havia uma maior ocorrência de “entradas vazias” pois o peixe não entra em lota todos os dias. Analisando mensalmente, este fenómeno tem naturalmente uma taxa de ocorrência bastante menor.

Tudo isto foi guardado em formato *RData*, que são ficheiros que guardam *workspaces*, nomeadamente funções e objetos de valor. Quer isto dizer que, quando passarmos ao passo de elaboração do *Dashboard*, não teremos de estar sempre a correr o que foi feito anteriormente, tornando-o o processo mais eficiente e menos demorado.

4.3. Desenvolvimento

Relativamente ao *Dashboard*, foi criado um *script* final. Utiliza-se a função *load* para carregar tudo o que foi utilizado no *Rdata*. Foram criados botões com os objetos de estudo considerados relevantes (euros, quilogramas e, no caso do grupo A1, tamanho), que por sua vez estão inseridos nos botões A1 e A2).

Como o trabalho é realizado totalmente em R, software que não é utilizado na empresa, foi dada autonomia total no manejo dos dados. Por não estar a lidar diretamente com a base de dados (SQL Server), os mesmos terão de ser importados directamente da respetiva directoria, apenas mudando a “barra” (exemplo: de “C:\Users\Desktop\DOCAPESCA”, para “C:/Users/Desktop/DOCAPESCA”). O primeiro mês de trabalho correspondeu à

fase de “brincar com os dados” onde se agruparam os anos todos por arte de pesca (CercoA1, ArrastoA1, PolivalentesA1, CercoA2, ArrastoA2 e PolivalentesA2) e dividiram-se em 2 scripts – A1 e A2; criou-se uma nova coluna com o preço médio, com duas casas decimais, calculado através da divisão dos euros pelos quilogramas de cada entrada de peixe em lota e mudou-se o nome da primeira coluna de “DD-MM-AAAA” para “Data” pois quando se trabalha com R, convém evitar os sinais e o acentos de pontuação de forma a evitar alguns erros.

A partir do segundo mês, já com objetivo de criar um *dashboard* definido, teve de ser escolhido um *package* de R que me permitisse a elaboração do mesmo. Como tal, foi escolhido o *Shinydashboard* pois, no âmbito da visualização de dados, não requer conhecimentos sobre *css*, *javascript* ou *html* para desenvolver uma aplicação.

| DD-MM-AAAA | Nome Lota | Cod Esp | Nome Esp | Tam | Fresc | kg | euros |
|---------------------|------------|---------|----------------|-----|-------|--------|--------|
| 02-01-2009 00:00:00 | Aveiro | HOM | Carapau | T5 | A | 13,20 | 50,16 |
| 02-01-2009 00:00:00 | Aveiro | HOM | Carapau | T5 | A | 19,30 | 74,31 |
| 02-01-2009 00:00:00 | Aveiro | HOM | Carapau | T5 | A | 100,00 | 385,00 |
| 02-01-2009 00:00:00 | Matosinhos | HOM | Carapau | T4 | A | 9,60 | 26,88 |
| 02-01-2009 00:00:00 | Matosinhos | HOM | Carapau | T3 | A | 12,10 | 36,30 |
| 02-01-2009 00:00:00 | Aveiro | HOM | Carapau | T5 | A | 8,20 | 56,58 |
| 02-01-2009 00:00:00 | Aveiro | HOM | Carapau | T4 | A | 6,00 | 23,10 |
| 02-01-2009 00:00:00 | Aveiro | HOM | Carapau | T5 | A | 10,00 | 38,50 |
| 02-01-2009 00:00:00 | Matosinhos | HOM | Carapau | T3 | A | 12,60 | 37,80 |
| 02-01-2009 00:00:00 | Matosinhos | HOM | Carapau | T4 | A | 300,00 | 960,00 |
| 02-01-2009 00:00:00 | Aveiro | HOM | Carapau | T5 | A | 62,60 | 244,14 |
| 02-01-2009 00:00:00 | Aveiro | HOM | Carapau | T4 | A | 22,40 | 101,93 |
| 02-01-2009 00:00:00 | Aveiro | HKE | Pescada-branca | T5 | A | 8,00 | 36,80 |
| 02-01-2009 00:00:00 | Peniche | HKE | Pescada-branca | T5 | A | 25,70 | 103,06 |
| 02-01-2009 00:00:00 | Peniche | HKE | Pescada-branca | T5 | A | 11,10 | 48,06 |
| 02-01-2009 00:00:00 | Peniche | HKE | Pescada-branca | T5 | A | 34,00 | 155,72 |
| 02-01-2009 00:00:00 | Peniche | HKE | Pescada-branca | T4 | A | 11,80 | 48,62 |
| 02-01-2009 00:00:00 | Peniche | HKE | Pescada-branca | T5 | A | 12,20 | 53,92 |
| 02-01-2009 00:00:00 | Matosinhos | PIL | Sardinha | T2 | A | 24,00 | 8,16 |
| 02-01-2009 00:00:00 | Aveiro | HKE | Pescada-branca | T4 | A | 10,90 | 57,77 |
| 02-01-2009 00:00:00 | Aveiro | PIL | Sardinha | T3 | A | 24,00 | 14,40 |
| 02-01-2009 00:00:00 | Peniche | HKE | Pescada-branca | T5 | A | 12,10 | 53,72 |
| 02-01-2009 00:00:00 | Peniche | HKE | Pescada-branca | T5 | A | 24,30 | 95,26 |
| 02-01-2009 00:00:00 | Aveiro | HKE | Pescada-branca | T3 | A | 5,40 | 49,14 |
| 02-01-2009 00:00:00 | Peniche | HKE | Pescada-branca | T3 | A | 15,20 | 119,62 |
| 02-01-2009 00:00:00 | Matosinhos | HKE | Pescada-branca | T1 | A | 1,40 | 10,50 |
| 02-01-2009 00:00:00 | Aveiro | HKE | Pescada-branca | T4 | A | 7,20 | 59,76 |

Figura 5 - Ficheiro Excel com os dados sobre a pesca de Arrasto em 2009 do grupo de pesca A1

| Ano | Mes | Data | Nome.Lota | Cod.Esp | Nome.Esp | Tam | Fresc | kg | euros | prmed |
|------|-----|------------|------------|---------|----------------|-----|-------|-------|--------|-------|
| 2009 | 1 | 2009-01-02 | Aveiro | HOM | Carapau | T5 | A | 13.2 | 50.16 | 3.80 |
| 2009 | 1 | 2009-01-02 | Aveiro | HOM | Carapau | T5 | A | 19.3 | 74.31 | 3.85 |
| 2009 | 1 | 2009-01-02 | Aveiro | HOM | Carapau | T5 | A | 100.0 | 385.00 | 3.85 |
| 2009 | 1 | 2009-01-02 | Matosinhos | HOM | Carapau | T4 | A | 9.6 | 26.88 | 2.80 |
| 2009 | 1 | 2009-01-02 | Matosinhos | HOM | Carapau | T3 | A | 12.1 | 36.30 | 3.00 |
| 2009 | 1 | 2009-01-02 | Aveiro | HOM | Carapau | T5 | A | 8.2 | 56.58 | 6.90 |
| 2009 | 1 | 2009-01-02 | Aveiro | HOM | Carapau | T4 | A | 6.0 | 23.10 | 3.85 |
| 2009 | 1 | 2009-01-02 | Aveiro | HOM | Carapau | T5 | A | 10.0 | 38.50 | 3.85 |
| 2009 | 1 | 2009-01-02 | Matosinhos | HOM | Carapau | T3 | A | 12.6 | 37.80 | 3.00 |
| 2009 | 1 | 2009-01-02 | Matosinhos | HOM | Carapau | T4 | A | 300.0 | 960.00 | 3.20 |
| 2009 | 1 | 2009-01-02 | Aveiro | HOM | Carapau | T5 | A | 62.6 | 244.14 | 3.90 |
| 2009 | 1 | 2009-01-02 | Aveiro | HOM | Carapau | T4 | A | 22.4 | 101.93 | 4.55 |
| 2009 | 1 | 2009-01-02 | Aveiro | HKE | Pescada-branca | T5 | A | 8.0 | 36.80 | 4.60 |
| 2009 | 1 | 2009-01-02 | Peniche | HKE | Pescada-branca | T5 | A | 25.7 | 103.06 | 4.01 |
| 2009 | 1 | 2009-01-02 | Peniche | HKE | Pescada-branca | T5 | A | 11.1 | 48.06 | 4.33 |
| 2009 | 1 | 2009-01-02 | Peniche | HKE | Pescada-branca | T5 | A | 34.0 | 155.72 | 4.58 |

Figura 6 - Ficheiro ArrastoA1

Ficheiro ArrastoA1, com todas as modificações feitas para a visualização de dados através do *Dashboard*

```

#Função para importar Bases de Dados
importar = function(path) {
  sheetnames = excel_sheets(path)
  mylist = lapply(excel_sheets(path), read_excel, path = path)

  # name the dataframes
  names(mylist) <- sheetnames
  do.call(rbind, mylist)
}

#Importar Bases de Dados
#2009
"Cerco"
CercoA12009 = importar("C:/Users/ASUS/Desktop/DOCAPESCA/A1_Cerco_T&F Pr _22NCV_2009.xlsx")

"Arrasto "
ArrastoA12009 = importar("C:/Users/ASUS/Desktop/DOCAPESCA/A1_Arrasto_T&F Pr _22NCV_2009.xlsx")

```

Figura 7 – Função de importação de dados

A função criada para importar os dados dos ficheiros *Excel*, com dois exemplos práticos. Desta forma garantimos a eficiência do código

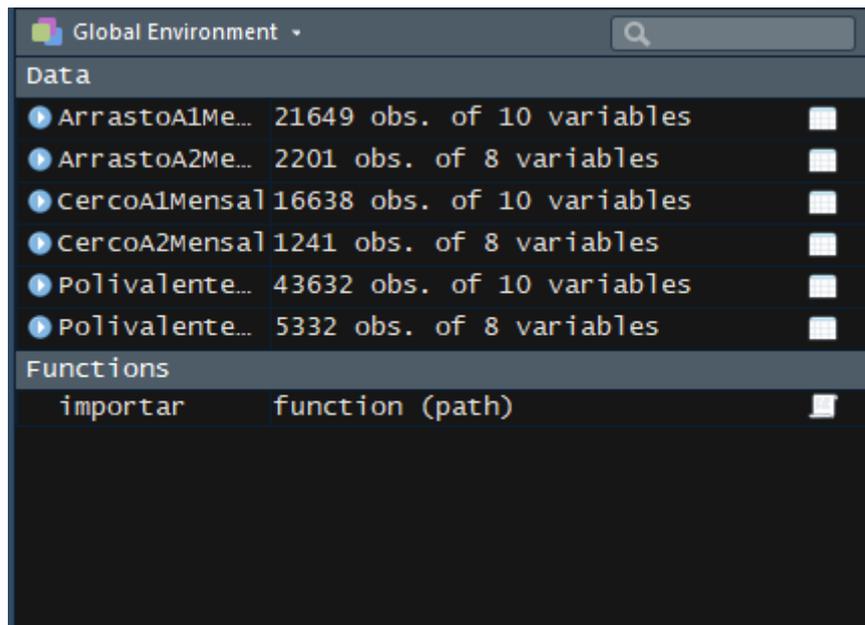


Figura 8 - Data

Data – abrindo isto com o *Rstudio*, fica já gravado, não sendo necessário incorporar isto no *script* do *dashboard*, tornando o processo mais eficiente e poupando na quantidade de linhas de código necessárias.

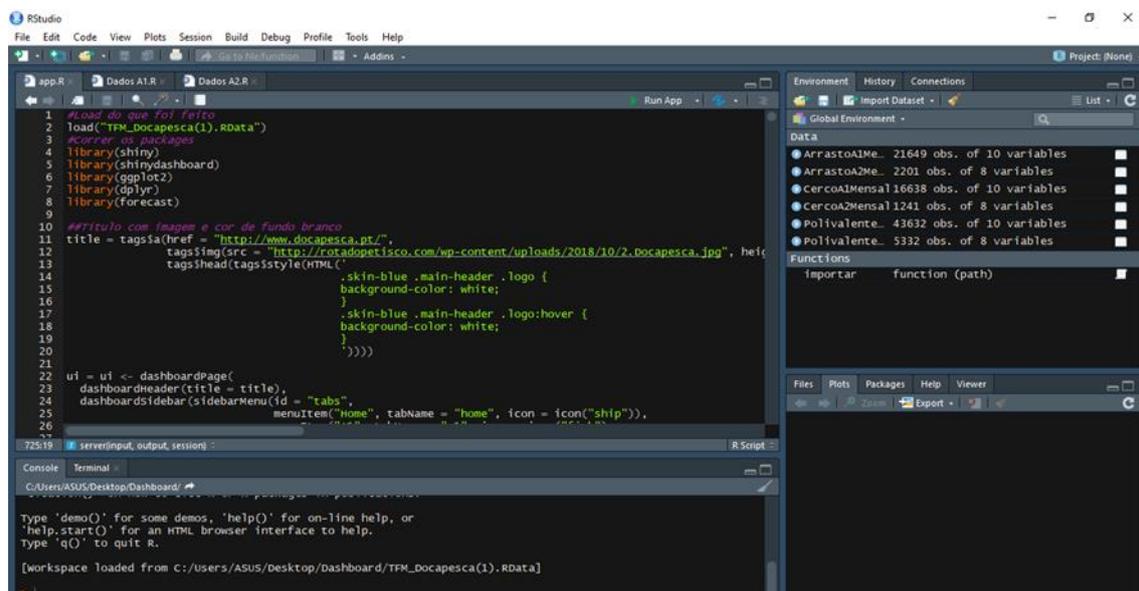


Figura 9 - Rstudio

Na parte superior esquerda temos os 3 *scripts* criados, divididos em separadores como nos *browsers*; no canto inferior esquerdo temos a consola onde introduzimos os comandos; no canto superior direito temos o *global environment* que conserva aquilo que introduzimos na consola. Neste caso específico, estão

visíveis os objetos do *R.Data* – os dados dos dois grupos de espécies de pescado, juntamente com a função “importar”; finalmente, no canto inferior direito temos o espaço reservado a (entre outros) os gráficos, que não vai ser utilizado pois está tudo a ser direcionado para o *dashboard*). A figura que se segue apresenta o modelo final do *dashboard* criado.

4.4. Visualização de dados

Para o produto final, devido ao propósito e à natureza do *dashboard*, foram incorporados maioritariamente gráficos de linhas para a consequente análise das espécies de pescado, através dos *packages shinydashboard* e *shiny*. Adicionalmente, foram introduzidos um gráfico de barras e um mapa (este via o *package leaflet*) como instrumentos de análise complementares, simultaneamente evidenciando outras funcionalidades possíveis de incorporar na solução.



Figura 10- Homepage do Dashboard

4.5. Resultados e Interpretação dos dados

Através da informação apresentada no *dashboard*, foi possível concluir que:

Espécies A1:

Biqueirão:

O período do ano onde se verificam as maiores quantidades de biqueirão vendido nas lotas é entre julho e outubro, no entanto, no resto ano as quantidades chegam a atingir valores muito significativos.

A arte de pesca predominante na sua captura é o Cerco. O valor mais elevado registou-se em setembro de 2016, cerca de 4 mil toneladas.

Apenas como curiosidade, o valor mais elevado capturado com arte de arrasto, foi de 10 toneladas em 2012 e de 45 toneladas em 2018 com artes polivalentes.

A lota que registou a maior quantidade vendida de biqueirão foi Matosinhos, com mais de 3,5 mil toneladas de kg no ano de 2016.

Monetariamente, isto traduz-se em mais de 7 milhões de euros transacionados da espécie biqueirão, no período de maior captura, como acima referido, no mês de setembro de 2016.

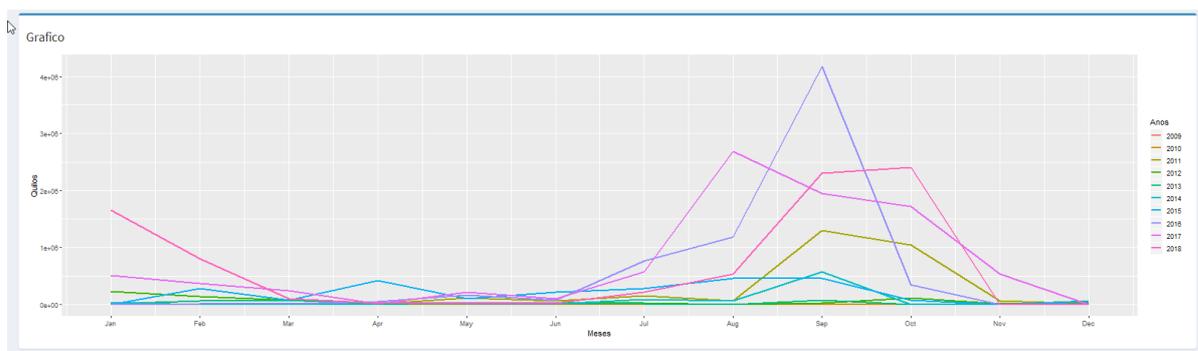


Gráfico 1 – Kgs de biqueirão pescados entre 2009 e 2018 com arte de pesca - cerco

Carapau:

Contrariamente ao que se verificou com o biqueirão, a pesca do carapau tem uma atividade mais duradoura e também mais linear ao longo do ano, com um período de menor intensidade no inverno.

Apesar de cerco registar picos mais elevados, é no arrasto que se verificam as maiores quantidades de carapau pescado ao longo do ano.

No mês de maio de 2016, registaram-se cerca de 1,75 milhões de kg de carapau transacionado no conjunto das lotas do continente.

As principais lotas responsáveis pelas maiores transações de carapau são: Matosinhos, Nazaré, Peniche e Portimão.

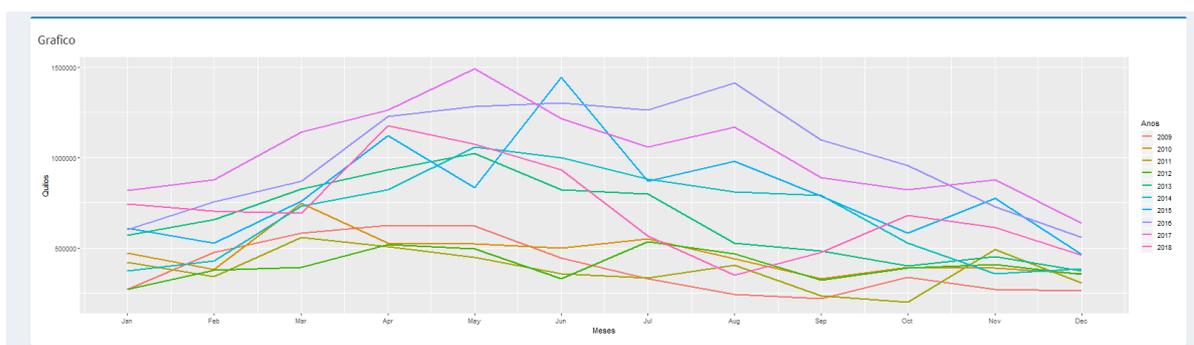


Gráfico 2 – Kgs de carapau pescados entre 2009 e 2018 com arte de pesca - cerco

Cavala:

Apesar de o padrão de atividade piscatória da cavala ser um pouco mais irregular que a do carapau, é possível observar que há maior quantidade de carapau transacionado em lota, na segunda metade do ano.

A “Campanha da Cavala” – que visou promover o aumento do consumo desta espécie entre os consumidores nacionais. A “Campanha da Cavala” foi lançada em 2012 e contribuiu para um aumento significativo das receitas relativas à venda desta espécie, a partir deste ano e atingiu um pico de mais de 1,6M de euros em 2018.

A mesma tendência se verificou com as quantidades transacionadas (kg) da cavala.

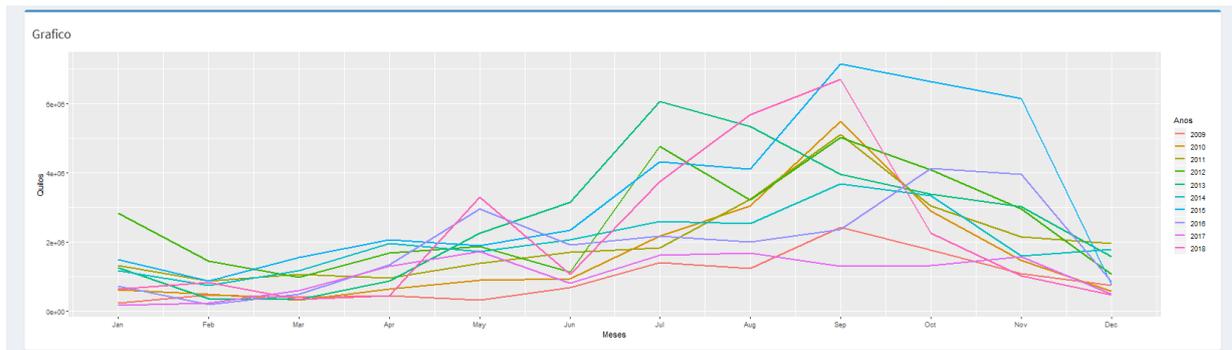


Gráfico 3 – Kgs de cavala pescados entre 2009 e 2018 com arte de pesca - cerco

Sardinha:

A Sardinha foi durante décadas, denominada como “A rainha”.

No período em análise 2009 a 2018, o maior pico nas quantidades vendidas foi registado no mês de novembro de 2011, com cerca de 8 mil toneladas, provenientes da arte de cerco, em novembro de 2011.

Devido à imposição de quotas ibéricas restritivas à pesca da sardinha, as quantidades transacionadas desta espécie reduziram-se acentuadamente nos últimos anos.

O período mais intenso de atividade piscatória corresponde ao verão e outono, que na última década são as estações que registam as vagas de calor mais elevadas e duradouras.

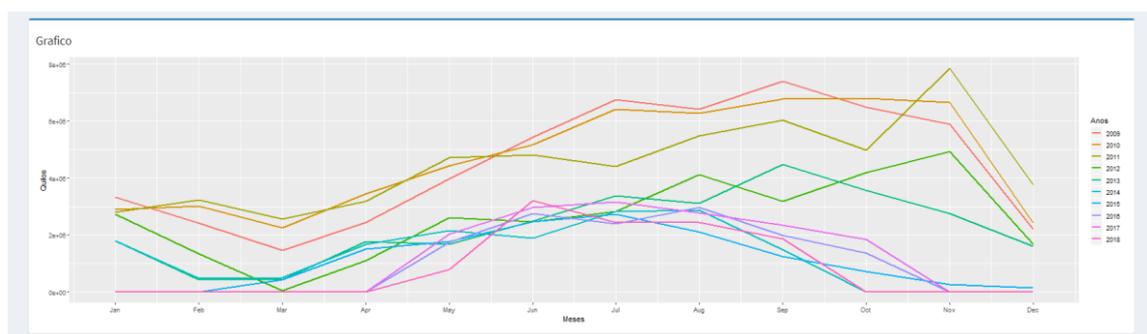


Gráfico 4 – Kgs de sardinha pescados entre 2009 e 2018 com arte de pesca - cerco

Pescada-Branca:

Das 5 espécies, é a que regista os números mais baixos de pesca, com maior ênfase na primavera e no verão;

O pico mais alto foi em julho de 2012 com quase 300 toneladas, proveniente das artes de pesca polivalentes.

As lotas onde se registaram as maiores quantidades de pescada-branca, foram nas delegações do Centro Norte e Norte, como Nazaré, Figueira da Foz, Matosinhos e Póvoa de Varzim.

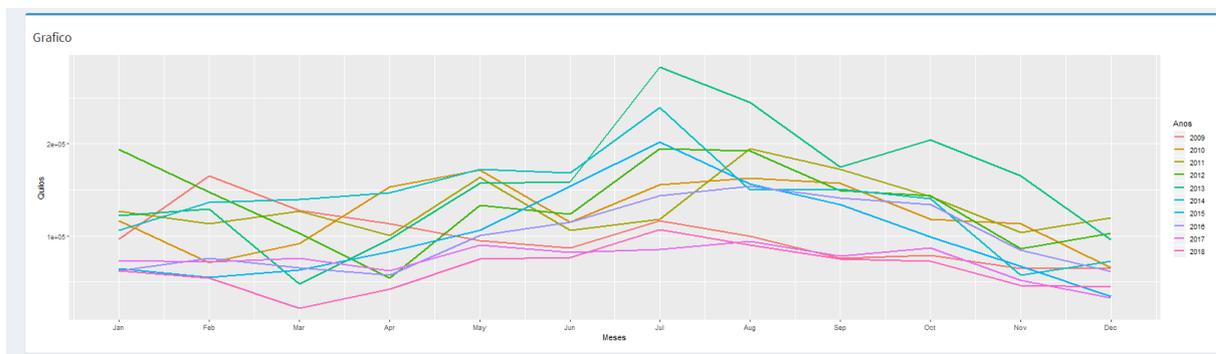


Gráfico 5 – Kgs de pescada-branca pescados entre 2009 e 2018 com arte de pesca - polivalentes

Espécies Grupo A2:

Polvo-vulgar:

As artes de pesca mais recorrentes para pescar Polvo-vulgar são as polivalentes, Foi em outubro de 2014 que se registou o pico dos kgs de polvo pescado (cerca de 1,25 milhões de kgs). Em abril de 2013 verificou-se outro pico na ordem dos 900.000 kgs.

Em Cerco, verificou-se um pico semelhante em julho de 2010, de cerca de 975.000 kgs. A arte de pesca de arrasto não apresenta qualquer registo relevante para a estatística.

Monetariamente, isto traduz-se em receitas bastante elevadas. Nas artes de pesca polivalentes, excluindo o ano atípico de 2009, só por uma vez é que se verificaram receitas abaixo do 1M de euros (setembro de 2017). Relativamente aos outros anos, com picos em março e em outubro, os valores são bastante elevados, com períodos a ultrapassarem os 3M de euros em receitas, destacando-se outubro de 2014 e 2018 com valores a rondar os 3,5M de euros.

A pesca de polvo-vulgar segue uma tendência de maior captura nas estações de primavera e outono, porém isso difere de lota para lota. É possível observar que, apesar de não existirem números destacáveis numa determinada lota, é possível ver que, enquanto há lotas a registar números mais elevados só no outono

(Aveiro e Figueira da Foz, por exemplo), casos com comportamentos bastante erráticos (Lagos) e ainda lotas que o pescam mais na primavera (Olhão).

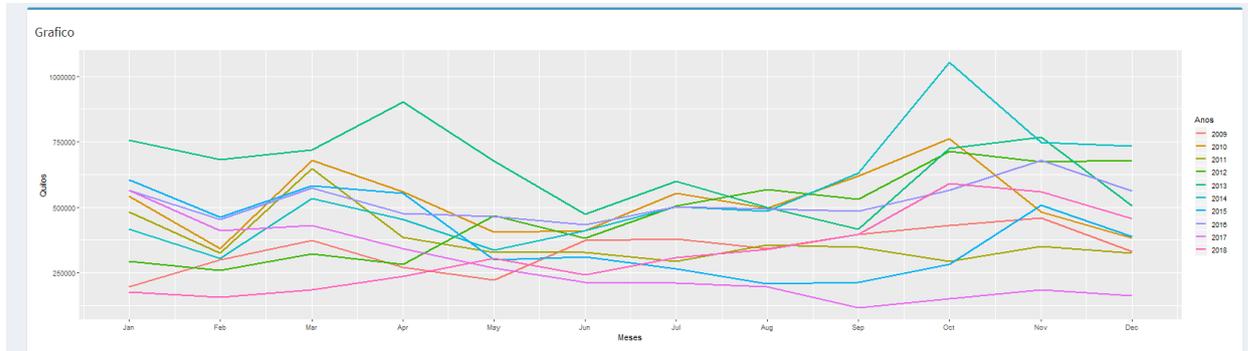


Gráfico 6 – Kgs de polvo-vulgar pescados entre 2009 e 2018 com arte de pesca - polivalentes

Choco-vulgar:

Com valores consideravelmente inferiores aos do polvo-vulgar, o choco apresenta uma época alta entre fevereiro e abril, com destaque para março e abril de 2010, registando valores acima dos 250.000 kgs.

Traduzindo a pesca de choco em receitas, tendo apenas em conta as artes de pesca polivalentes visto que é a arte predominante na pesca do choco-vulgar, o período de maior receita corresponde ao período entre o final do inverno e o início do verão, com os maiores picos a ocorrerem entre março e maio em diversos anos, registando valores acima dos 600.000 euros.

A principal lota na pesca do choco corresponde à lota de Aveiro.

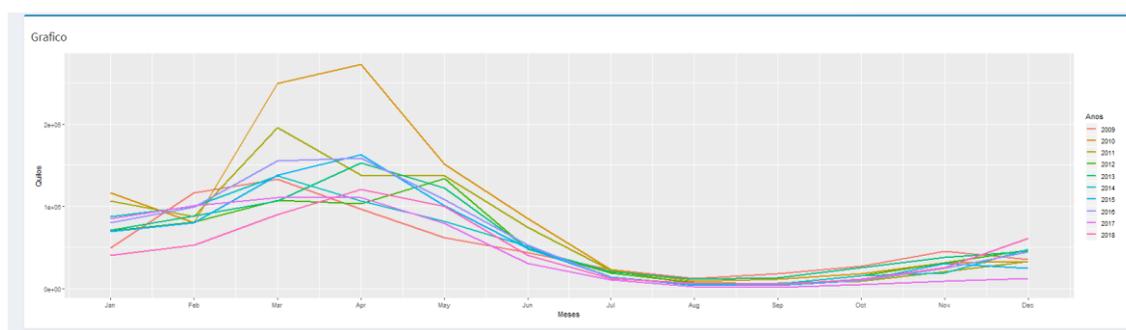


Gráfico 7 – Kgs de choco-vulgar pescados entre 2009 e 2018 com arte de pesca - polivalentes

Amêijoa-branca:

De todas as espécies em estudo, a amêijoa-branca é a única que regista valores apenas num tipo de pesca.

O período mais intenso da apanha da amêijoa-branca corresponde ao verão, com o pico da mesma a coincidir com agosto. O registo mais elevado foi em agosto de 2018, registando quase 200.000 kgs de amêijoa transacionada. (Existe um pico mais alto registado em dezembro de 2018, mas acredita-se que seja um erro no registo do valor).

Traduzindo isto em receita, o comportamento do gráfico é idêntico: época alta nos verões e período mais fraco no 1.º semestre de cada ano. Em números, o pico mais alto corresponde a quase meio milhão de euros.

O porto onde se verifica maior entrada de amêijoa-branca é o de Matosinhos, só por si, regista cerca de metade dos valores totais registados no conjunto das lotas do continente.

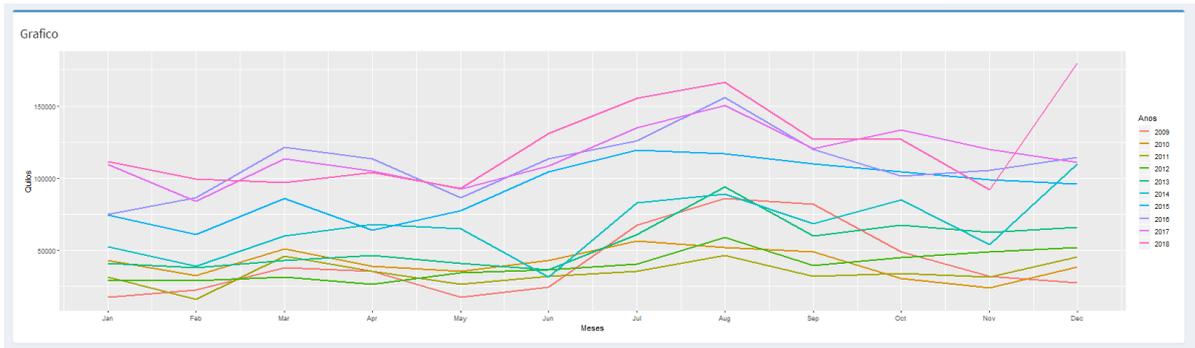


Gráfico 8 – Kgs de amêijo-a-branca pescados entre 2009 e 2018 com arte de pesca - polivalentes

Mapa

Finalmente, foi introduzido um mapa de Portugal Continental onde estão identificadas as lotas geograficamente. Ao selecionar uma lota, surge um balão com dados informativos acerca da respetiva lota, nomeadamente, nome da lota, qual o tipo de pescado (grupo A1 e grupo A2) mais transacionado, respetivas quantidades e valores transacionados.

Implementando este mapa, foi possível sintetizar a atividade de cada lota, de uma maneira apelativa para o utilizador.

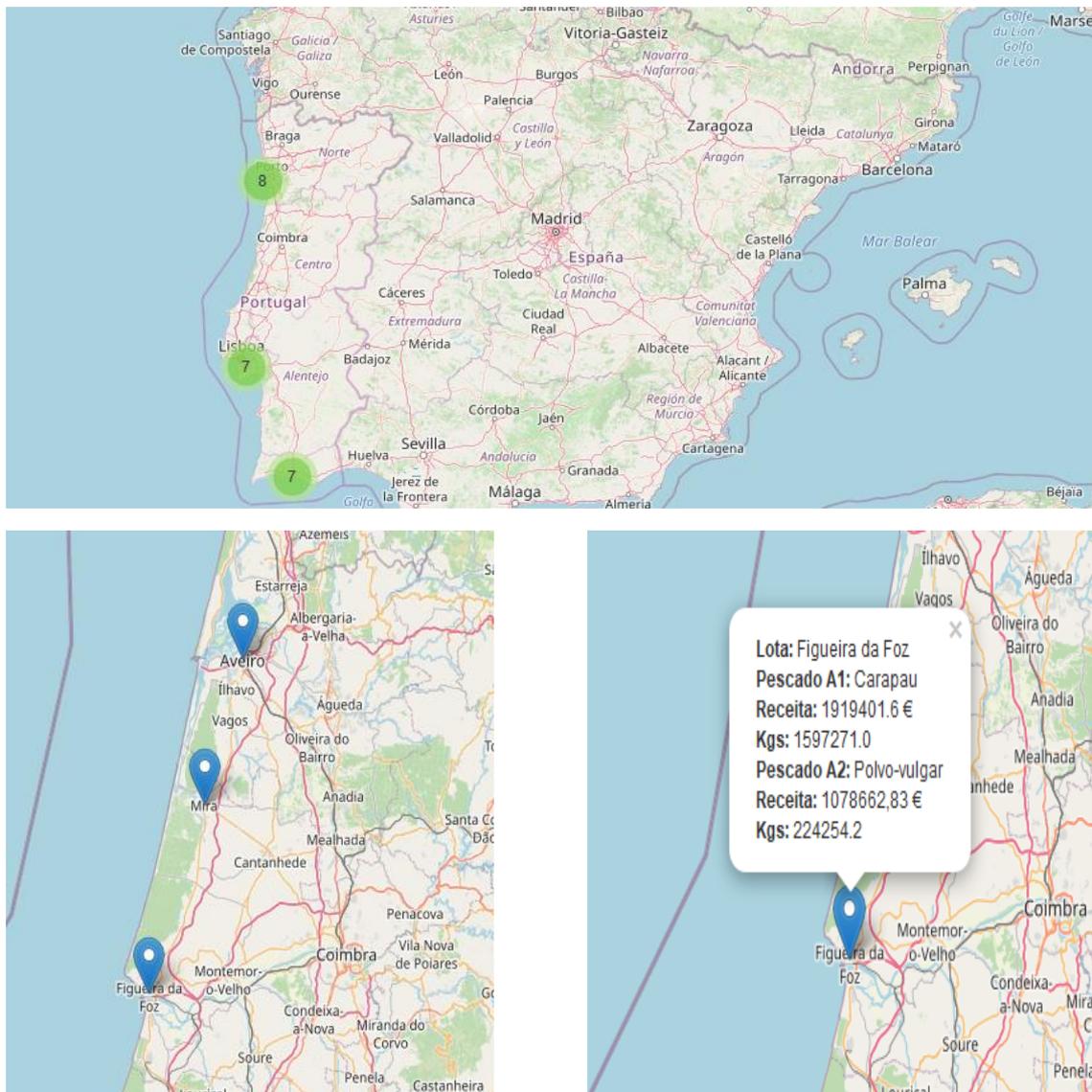


Figura 11 – Localização geográfica e dados das lotas

Como referido anteriormente, o mapa tem como maior propósito servir de complemento para a informação disponibilizada nos diversos gráficos explicados anteriormente.

5. Conclusões e trabalhos futuros

O número de dados disponíveis tem vindo a crescer exponencialmente a um ritmo diário. Como tal, é necessário para as empresas terem ao seu dispor ferramentas capazes de assimilar qualquer tipo de informação para poderem adaptar o seu modelo de negócio adequadamente.

Neste contexto, este trabalho procurou, através de uma investigação da ferramenta *R*, elaborar um *dashboard*, tendo em conta as especificidades dos dados fornecidos pela Docapesca – Portos e Lotas, S.A – dois grupos de pescado contendo uma panóplia de espécies e as suas transações pelas diversas lotas do país.

Atendendo ao que foi investigado ao longo deste trabalho, o *dashboard* mais indicado seria um estratégico, pelas seguintes razões: os dados visualizados são significativos enquanto agregado, tornando como principal foco de análise as lotas em estudo; a informação procurou ser resumida e foi orientada para a tomada de decisão de um ponto vista mais estratégico; finalmente, as atualizações deveriam ser trimestrais, pois estes dados são mais relevantes em maiores dimensões.

Contextualizando o problema, foi desenvolvida uma arquitetura para a solução do problema: os dados são exportados da base de dados em formato *xls* (*Excel*). Após a organização e formatação dos mesmos, são posteriormente exportados para *R*, onde através da instalação dos *packages*, foi possível criar um script que quando corrido nos dá um *dashboard* com as especificidades mencionadas anteriormente.

Uma temática a explorar em trabalhos futuros seria estabelecer uma ligação entre a base de dados e o *dashboard*, com vista a uma otimização mais eficiente da mesma.

Outra sugestão para trabalhos futuros seria um maior foco em séries temporais, algo que nesta solução se tornou insignificante devido à natureza dos dados estudados.

Finalmente, seria interessante explorar uma solução que utilizasse linguagens de programação com base em estatística, como *excel* ou *tableau*. Isto poderia

implicar um tratamento mais especializado dos dados (como referido anteriormente) e que para a visualização dos mesmos haveria a possibilidade de ser utilizado uma ferramenta de report, a linguagem de programação em si ou até mesmo páginas HTML.

6. Bibliografia

- Aparicio, S., Aparicio, J. T., & Costa, C. J. (2019). Data Science and AI: trends analysis. In 2019 14th Iberian Conference on Information Systems and Technologies (CISTI) pp. 1-6. IEEE. DOI: 10.23919/CISTI.2019.8760820
- Aparicio, M., & Costa, C. J. (2015). Data visualization. *Communication design quarterly review*, 3(1), 7-11.
- Bikakis, N. (2018). Big Data Visualization Tools. In *Encyclopedia of Big Data Technologies*, Springer.
- Caetano, T. V., & Costa, C. J. (2014). Data Warehousing num contexto de Sistemas Integrados. In *Atas da Conferência da Associação Portuguesa de Sistemas de Informação Vol. 12*, pp. 186-199.
- Canabarro, S. (2017). Referencing Big Data: três pequenas formas de utilizar em seu negócio. Disponível em:
<https://canaltech.com.br/big-data/big-data-tres-pequenas-formas-de-utilizar-em-seu-negocio/>
- Cavique, Luís (2014) Big data e data science. "Boletim da APDIO" [Em linha]. Nº 51, p. 11-14
- Costa, C. J., & Aparicio, M. (2005). Visualization of balanced scorecard on PDAs. In *Proceedings of the 23rd annual international conference on Design of communication: documenting & designing for pervasive information* pp. 103-107. <https://doi.org/10.1145/1085313.1085339>
- Costa, C. J., & Aparício, M. (2019). Supporting the decision on dashboard design charts. In *Proceedings of 254th The IIER International Conference 2019* pp. 10-15.
- Costa, C. J., & Aparício, J. T. (2020). POST-DS: A Methodology to Boost Data Science. 2020 15th Iberian Conference on Information Systems and Technologies (CISTI), Sevilla, Spain, 2020, pp. 1-6, doi: 10.23919/CISTI49556.2020.9140932.
- Evans, J. R., & Lindner, C. H. (2012). Business analytics: the next frontier for decision sciences. *Decision Line*, 43(2), 4-6.
- Evelson, B., & Norman, N. (2008). Topic overview: Business intelligence. *Forrester research*, 61.

- Few, S. (2004). Dashboard Confusion, 4.
https://www.perceptualedge.com/articles/ie/dashboard_confusion.pdf
- Few, S. (2004). Selecting the Right Graph for Your Message, 8.
http://www.perceptualedge.com/articles/misc/Graph_Selection_Matrix.pdf
- Friendly, M. (2008). A Brief History of Data Visualization. Handbook of Data Visualization (pp. 15–56).
- Fry, B. (2007) Visualizing Data: Exploring and Explaining Data with the Processing Environment. O'Reilly Media, Inc.
- Ikechukwu, A., Edwinah, A. & Monday, O. (2012). Use-of-Dashboard: A vital moderator of sales force competence management and marketing performance relationship. Information and knowledge management, 2(5), 30-39
- Liberatore, M. J., & Luo, W. (2010). The analytics movement: Implications for operations research. Interfaces, 40(4), 313-324
- Marques, M. A., & Costa, C. J. (2019). Social CRM Analytics Challenges. Journal of Information Systems Engineering & Management, 4(4).
- Nofal, M. I., & Yusof, Z. M. (2013). Integration of business intelligence and enterprise resource planning within organizations. Procedia Technology, 11, 658-665.
- Pauwels, K., Ambler, T., Clark, B., LaPointe, P., Reibstein, D., Skiera, B., Wierenga, B. & Wiesel, Th. (2009). Dashboards as a Service: Why, What, How, and What Research Is Needed? Journal of Service Research. 12. 175-189.
- Pires, J, P. & Costa, S (2011). Docapesca. Docapesca – Portos e Lotas, S.A, Lisboa
- Power, D. J. (2007). A brief history of decision support systems. DSSResources. COM, World Wide Web, <http://DSSResources.com/history/dsshhistory.html>, version, 4.
- Tegarden, D. P. (1999). Business Information Visualization. Communications of the Association for Information Systems, 1(1). <https://doi.org/10.17705/1CAIS.00104>
- Tomás, D.; Costa, C. J., Gaivão, J. P, & Carvalho, J.P., (2018) "Time series for incidences, orders and invoicingforecast" .CAPSI 2018 Proceedings. 36.<https://aisel.aisnet.org/capsi2018/36>
- Vuori, V. (2006). Methods of Defining Business Information Needs. Proceedings of the International Conference on Electronic Business (ICEB).

Yigitbasioglu, Ogan & Velcu, Oana. (2012). A review of dashboards in performance management: Implications for design and research. *International Journal of Accounting Information Systems*. 13. 41-59. 10.1016/j.accinf.2011.08.002