

Universidade do Minho
Escola de Engenharia

Luís Rafael Araújo Ribeiro

Data Analytics: Abordagem para Visualização da informação



Universidade do Minho
Escola de Engenharia

Luís Rafael Araújo Ribeiro

Data Analytics: Abordagem para Visualização da informação

Dissertação de Mestrado
Mestrado Integrado em Engenharia e Gestão de Sistemas de Informação

Trabalho efetuado sob a orientação do
Professor Jorge Oliveira e Sá

DECLARAÇÃO

Nome: Luís Rafael Araújo Ribeiro

Endereço eletrónico: lr_ribeiro@hotmail.com

Telefone: 969324699

Número do Bilhete de Identidade: 13632018

Título:

Data Analytics: Abordagem para Visualização da informação

Orientador(es):

Professor Jorge Oliveira e Sá

Ano de conclusão: 2015

Designação do Mestrado: Mestrado Integrado em Engenharia e Gestão de Sistemas de Informação

DE ACORDO COM A LEGISLAÇÃO EM VIGOR, NÃO É PERMITIDA A REPRODUÇÃO DE QUALQUER PARTE DESTE TRABALHO

Universidade do Minho, ___/___/____

Assinatura: _____

AGRADECIMENTOS

Uma tese de mestrado, apesar de ser processo solitário a que qualquer investigador está destinado, reúne contributos de várias pessoas.

Logo desde o início do mestrado, senti-me privilegiado por poder contar com a confiança e o apoio de inúmeras pessoas. Sem esses contributos, esta investigação não teria sido possível.

Ao Professor Jorge Oliveira e Sá, orientador da dissertação, agradeço o apoio, a partilha do saber e a sua grande disponibilidade sempre que precisei para me apoiar no trabalho. Acima de tudo, obrigado por continuar a acompanhar-me nesta jornada e por estimular o meu interesse pelo conhecimento e incentivar-me cada vez mais numa área em que tenho grande interesse e pela qual quero seguir no mundo do trabalho.

Quero também agradecer aos meus familiares e amigos pelo incentivo recebido ao longo destes anos. Aos meus pais Manuel e Celeste, á minha irmã Diana, um obrigado especial pelo apoio, amor e atenção prestados ao longo destes anos

A todos os meus colegas de mestrado e de universidade, mas particularmente ao Joel Sousa, João Gomes, Tiago Costa, Rui Fernandes, João Cadime, Vítor Pinheiro, João Gonçalves e Henrique Carvalho por todo o companheirismo e entreaajuda durante todos estes fantásticos anos.

Por fim, o meu profundo e sentido agradecimento a todo o corpo docente pela passagem de todo o conhecimento que contribuiu para a concretização desta dissertação, estimulando-me intelectual e emocionalmente.

RESUMO

Na área de *Data Analytics*, existem poucas metodologias/abordagens reconhecidas para se obter uma correta visualização da informação. Antes de sermos capazes de visualizar toda a informação é necessário preparar e compreender os diferentes tipos de dados: recolher, caracterizar, corrigir e transformar, para que, quando utilizados pelos diferentes tipos de *softwares* que irão apresentar a informação, não surjam erros ou incoerências. A relevância deste aspeto é primordial visto que a adequação de uma técnica de visualização depende do tipo de dados a serem exibidos e do estado em que os dados se apresentam no momento de apresentarmos.

Sendo esta uma área em crescimento devido ao aumento do uso massivo de dados por parte das organizações, torna-se relevante perceber como se pode retirar o melhor partido desse enorme volume de dados, nomeadamente através da sua visualização para posterior tomada de decisão. Sendo este um trabalho em desenvolvimento, pretende-se, a partir de uma revisão sistemática de literatura, identificar o estado-da-arte relativamente ao processo de criar uma visualização de informação, permitindo identificar informações de um conjunto de estudos existentes que incluam boa qualidade metodológica. Como resultado apresenta-se uma *framework* para servir como base para uma metodologia/abordagem para se obter uma visualização de informação.

Palavras-Chave: Data Analysis, Data Analytics, Information Visualization, InfoViz, Dashboards, Reports, Infografia

ABSTRACT

In the Data Analytics area, there is a lack of methods / approaches recognized to obtain a correct visualization of information, then, the main objective of this work involves making a bibliographic review and collection of all aspects related to the VI to provide a definition of all the concepts related to it, as well as the importance and the impact that results from their application. This is a growing area and increasingly important in research, analysis and technological development, but suffers from a well-defined framework, scientific knowledge and a more targeted teaching application. It is therefore of utmost importance to try to define the problem, identify gaps and propose a possible solution.

Keywords: Data Analysis, Data Analytics, Information Visualization, InfoViz, Dashboards, Reports, Inforgrafia

ÍNDICE

Agradecimentos.....	i
Resumo.....	iii
Abstract.....	iv
Lista de Abreviaturas, Siglas e Acrónimos	xi
1 Introdução	1
1.1. Problema	1
1.2. Motivação / Enquadramento	3
1.3. Objetivos.....	4
1.4. Estrutura do Documento.....	4
2 Abordagem Metodológica	7
2.1. Método Design Science Research	9
2.2. Outputs do método Design Science Research	10
2.3. Aplicação do método de Investigação	12
3 Revisão de Literatura	21
3.1. Conceitos Importantes.....	21
3.2. Design/Visualization Theories and Principles	23
4 Data Analytics: Visualização da informação.....	33
4.1. Como resolver o Problema da má utilização da VI.....	33
4.2. Desenvolvimento	34
4.3. Desenvolvimentos futuros.....	38
4.4. Contribuição.....	38
4.5. Validação	39
5 Framework	64
5.1. Framework conceptual Visualização da Informação.....	74
6 Conclusões.....	76
6.1. Síntese do trabalho desenvolvido	76
6.2. Contribuição da dissertação	77
6.3. Limitações e perspetivas de desenvolvimentos futuros	77

Bibliografia 79

LISTA DE FIGURAS

Figura 1: Fases DSR, adaptado de (Vaishnavi & Kuechler, 2004).....	10
Figura 2: Outputs do método DSR adaptado de (Vaishnavi & Kuechler, 2004)	11
Figura 3 - Gestalt Theory: Similaridade	24
Figura 4 - Gestalt Theory: Continuação.....	24
Figura 5 - Gestalt Theory: Encerramento	25
Figura 6 - CRAP Theory : Contraste	25
Figura 7 - CRAP Theory : Repetição.....	26
Figura 8 - CRAP Theory : Alinhamento.....	26
Figura 9 - CRAP Theory : Proximidade.....	26
Figura 10 - CRAP Theory : Proximidade 2.....	27
Figura 11 - CRAP Theory : Proximidade 3.....	27
Figura 12 - CRAP Theory : Proximidade 4.....	27
Figura 13 - CRAP Theory : Forma e fundo	27
Figura 14 - CRAP Theory : Forma e fundo 2	28
Figura 15 - CRAP Theory : Forma e fundo 3	28
Figura 16 - Dieter Ram's Principles: Inovação.....	28
Figura 17 - Dieter Ram's Principles: Usabilidade.....	29
Figura 18 - Dieter Ram's Principles: Estética	29
Figura 19- Dieter Ram's Principles: Compreendido	29
Figura 20 - Dieter Ram's Principles: Discreto.....	30
Figura 21 - Dieter Ram's Principles: Honesto.....	30
Figura 22 - Dieter Ram's Principles: Duradouro	30
Figura 23 - Dieter Ram's Principles: Detalhado.....	31
Figura 24 - Dieter Ram's Principles: Amigo do ambiente.....	31
Figura 25 - Má conceção	36
Figura 26 - Boa conceção	37
Figura 27 - Questionário VI	41
Figura 28 - Questionário VI	42
Figura 29 - Questionário VI	43
Figura 30 - Questionário VI	44

Figura 31 – Indicador número 1 47

Figura 32 - Indicador número 2 48

Figura 33 - Indicador número 3 49

Figura 34 - Indicador número 4 50

Figura 35 - Indicador número 5 51

Figura 36 - Indicador número 6 52

Figura 37 - Indicador número 7 53

Figura 38 - Indicador número 8 54

Figura 39 - Indicador número 9 55

Figura 40 - Indicador número 10 56

Figura 41 - Indicador número 11 57

Figura 42 - Indicador número 12 58

Figura 43 - Indicador número 13 59

Figura 44 - Indicador número 14 60

Figura 45 - Framework 74

LISTA DE TABELAS

Tabela 1 - Características do <i>Design Science Research</i>	8
Tabela 2 - Outputs do método DSR	11
Tabela 3 - Matriz de Conceitos	14
Tabela 4 - Métodos de Avaliação	18
Tabela 5 - Soluções VI	34
Tabela 6 - Casos Considerados	46
Tabela 7 - Indicador número 1	47
Tabela 8 - Indicador número 2	48
Tabela 9 - Indicador número 3	49
Tabela 10 - Indicador número 4	50
Tabela 11 - Indicador número 5	51
Tabela 12 - Indicador número 6	52
Tabela 13 - Indicador número 7	53
Tabela 14 - Indicador número 8	54
Tabela 15 - Indicador número 9	55
Tabela 16 - Indicador número 10	56
Tabela 17 - Indicador número 11	57
Tabela 18 - Indicador número 12	58
Tabela 19 - Indicador número 13	59
Tabela 20 - Indicador número 14	60

LISTA DE ABREVIATURAS, SIGLAS E ACRÓNIMOS

VI - Visualização da Informação

BI- *Business Intelligence*

TI - Tecnologias de Informação

SI - Sistemas de Informação

DSR - *Design Science Research*

DA - *Data Analytics*

HCI - *Human Computer Interaction*

1 INTRODUÇÃO

Neste primeiro capítulo será introduzida a temática e a motivação para a realização deste trabalho assim como apresentados os objetivos, motivação e enquadramento.

1.1. Problema

Atualmente, uns dos aspetos de maior relevância relacionado com o *Business Intelligence (BI)* prendem-se com *Data Analytics* e a Visualização da Informação. Realça-se que todo o processo de BI, nomeadamente a recolha, tratamento e armazenamento de dados é importante e decisivo para o bom funcionamento, mas o que o gestor/decisor de facto vê e utiliza são as ferramentas de análise de dados. As organizações estão atualmente preocupadas em gerir informação de forma a terem acesso a informações corretas, atuais e relevantes para o negócio (Blewett, 2011).

Os dados que suportam essa informação são provenientes de múltiplas fontes, estruturadas ou mesmo não estruturadas, são dados provenientes de sistemas operacionais existentes na organização, podem também ser dados provenientes de experiências de utilizadores sobre os produtos e serviços que a organização oferece no mercado, estes dados são denominados de *Big Data*.

Manipular grandes volumes de dados de forma a obter-se uma correta informação sobre o negócio, é um dos desafios que as organizações enfrentam. Para além deste aspeto, a disponibilização dos dados de uma forma adequada permite identificar pontos positivos da organização e potenciá-los para que se consiga um melhor aproveitamento.

Uma das dificuldades que os profissionais da área sentem é saber se estão a fornecer dados da melhor forma para que o utilizador obtenha a melhor informação, isto é denominado de *Data Analytics*.

Uma das componentes importantes de *Data Analytics* é a VI, tendo por objetivo conseguir uma boa visibilidade dos dados, em tempo real, consistente e extrair o máximo da informação. É um campo que está a ser investigado, solicitado por pessoas que trabalham com dados, e utilizados por uma percentagem crescente de organizações no mercado de trabalho, especialmente as de análise e visualização de dados.

Na realidade, a VI é ainda muito mal compreendida, pois a sua utilização carece de perícia, de conhecimento científico, nomeadamente para compreender os aspetos da perceção visual, é ainda uma área que reúne conhecimentos de muitas outras áreas do conhecimento completamente díspares, nomeadamente, da arte, da estatística, da psicologia, da informática, etc.

Estas capacidades raramente são encontradas no mundo dos negócios, não porque sejam difíceis de aprender, mas porque a necessidade de as aprender raramente é reconhecida.

O desafio de exibir uma grande variedade de informação através dos diferentes modelos de visualização da informação, muitas das vezes, resulta num excesso de informação, incoerente e nada organizada. A VI utiliza imagens, gráficos para representar a informação que são ferramentas poderosas no que diz respeito à interpretação da informação, mas apesar desse potencial, os benefícios são prejudicados devido à falta de compreensão da matéria, levando muitas vezes à confusão ao invés da compreensão.

Por outro lado, está cada vez a utilizar-se o conceito de infografia, muito utilizado no jornalismo, em que se tenta contar uma “história” através dos dados, isto é, que a informação faça sentido (Blewett R., 2011).

Dessa forma, a partir de tabelas e gráficos, de fontes e cores, e de infografias para contar histórias sobre os dados, VI é uma capacidade que muitos profissionais de BI aprendem informalmente no dia-a-dia do trabalho, muitas vezes a partir de tentativas que resultam em erros, em vez de serem peritos em VI.

No entanto, VI é a linguagem que permite fazer a transição de dados puramente estatísticos para dados que fazem sentido e dessa forma melhoram a perceção e ação sobre o negócio.

1.2. Motivação / Enquadramento

Há inúmeros benefícios de análise de dados no entanto, existem alguns que se destacam, tais como: ajuda na estruturação dos resultados recolhidos de diferentes fontes, utilidade na estruturação de problemas complexos em peças independentes para uma melhor análise.

Esta análise de dados age como um filtro quando se trata de adquirir conhecimentos significativos de um conjunto de dados. Cada analista tem de trabalhar um grande volume de dados antes de chegar a uma conclusão, que neste caso passa pelo momento de Visualização da Informação (VI). A análise dos dados revela-se crucial neste processo pois fornece uma base significativa para decisões críticas, ajuda a criar uma base de confiança para a VI.

Dependendo de como queremos expor a informação, podemos visualizar a mesma fonte de dados de várias formas diferentes. É impreterível, quando preparamos qualquer tipo de informação garantir a qualidade e integridade da mesma.

Um dos pontos de maior importância passa por conseguir clarificar qual a “imagem ou história” que desejamos expor através das diferentes formas de VI. Por norma, a maior parte das ferramentas apresenta modelos generalistas de VI, o seu conteúdo deve ser o mais específico e sucinto possível – é isso que o torna eficaz. Ajudar a tomar boas decisões, informar, persuadir, ensinar, mover as pessoas à ação e resolver problemas usando a cognição visual.

A principal função dos modelos de VI não se prende em apresentar a informação de forma detalhada, mas sim, a sintetizar essa informação ao mínimo possível, sem comprometer a sua compreensão. Diferentes tipos de dados requerem diferentes tipos de abordagens e metodologias de tratamento, mas em todas elas é necessário uma colheita dos dados de maior relevância, pois seria impossível apresentar os modelos de VI com a quantidade de informação presente, ficando assim, apenas as métricas fundamentais para a compreensão da informação e as facilmente compreendidas, que dispensem qualquer explicação adicional.

1.3. Objetivos

A utilização da VI a partir de um conjunto de dados previamente trabalhados é uma área em constante expansão em que, nos últimos anos, tem-se assistido a um claro desenvolvimento da forma como a informação é trabalhada e posteriormente apresentada. Apesar deste crescimento, existem ainda diversos desafios por investigar na área da VI nomeadamente qual a melhor forma de a apresentar consoante o tipo de dados a serem apresentados.

O objetivo fundamental desta dissertação passa por detetar e caracterizar um conjunto de regras, passos a seguir no momento da análise de dados, que posteriormente resultará na VI de uma forma mais limpa e apelativa ao utilizador.

Seguindo a metodologia DSR (*Design Scienc Research*), as regras e passos a seguir devem ser identificadas, através de um revisão de literatura baseada no tema *Data Analytics*, o que resultará num conjunto de informação onde seja identificada a existência de algum tipo de informação que sirva de suporte ao desenvolvimento de modelos de VI e se essa informação é a mais adequada. Por outro lado, se essa informação não existir, e mediante a revisão de literatura realizada será então, criado um *framework*.

Os resultados esperados prendem-se com a recolha dos diferentes métodos e modelos mais promissores que servirão de suporte ao desenvolvimento de novos modelos de VI, em particular para entender os fatores principais a ter em conta no desenvolvimento e a análise da evolução que tem acontecido ao longo do tempo com a finalidade, de eventualmente propor um método mais atual e correto, potenciando todos os benefícios que apresentam para que o utilizador obtenha uma visualização dos dados clara e concisa.

1.4. Estrutura do Documento

Após a introdução, nomeadamente motivação/enquadramento e objetivos termina o ponto 1 do documento. O restante documento está dividido em mais 5 capítulos.

No capítulo 2 é definida e descritos os passos da metodologia de investigação utilizada para a realização deste trabalho de investigação.

O capítulo 3 é a revisão de literatura, capítulo esse que está dividido em 3 sub secções: Na primeira são referenciados e descritos alguns conceitos importantes relacionados com *Business Intelligence*; na segunda secção são descritas teorias importantes relacionadas com a perceção,

design e visualização; finalmente na terceira é feita a revisão de literatura dos temas desta dissertação.

De seguida no capítulo 4 é apresentado o trabalho desenvolvido com base na revisão de literatura, no capítulo 5 é apresentado resultado do trabalho, neste caso o framework conceptual, seguido pelo capítulo 6 onde estão descritas as conclusões, seguidos finalmente pelas referências bibliográficas.

2 ABORDAGEM METODOLÓGICA

A metodologia *Design Science Research* (DSR) oferece grandes benefícios no que se refere à solução de problemas da vida real das organizações, devido ao seu pragmatismo, é adequada à investigação de problemas de natureza prática, em vez de, por outro lado, focar-se na verificação de leis naturais ou teorias comportamentais (Hevner, March et al., 2004).

Não é uma tarefa fácil definir DSR. Inicialmente, é necessário posicioná-la dentro do contexto do trabalho, investigando as questões paradigmáticas, tipológicas e metodológicas. Devem-se assim também apresentar comparações com outras metodologias, com a finalidade de melhoria de forma, conceito e suas aplicações.

Pode-se ver esta metodologia como um instrumento para a estruturação de artefactos que visam solucionar problemas de cariz prático, em ambiente real, e, desta maneira, conduzir à geração de conhecimento.

Fundamentalmente, o método DSR diz que o conhecimento, compreensão e solução de um problema são obtidos através da construção e aplicação de um artefacto, constituído por um determinado número de fases, para o contexto específico.

Depois da recolha das necessidades organizacionais, bem como dos problemas de interesse do investigador, a DSR serve de suporte ao desenvolvimento e construção de artefactos, bem como no aperfeiçoamento de teorias existentes. Estes artefactos, futuramente padecem de avaliações, bem como da necessidade de justificar a sua importância. Para sustentar estas atividades de desenvolvimento, construção, justificação e avaliação a base de conhecimento presente precisa de ser consultada e utilizada. A base de conhecimento referida é estabelecida pelos fundamentos e métodos consolidados, reconhecidos pelos autores das diferentes versões da teoria reconhecida em âmbito académico. Estes métodos sustentam particularmente as atividades de justificação e avaliação dos artefactos construídos ou da teoria aprimorada (Hevner et al., 2004).

De modo a auxiliar o desenvolvimento da DSR, (Hevner, March et al., 2004) definem alguns critérios que devem ser considerados por quem efetua a pesquisa. Os critérios definidos são imprescindíveis e são apresentados de seguida:

Tabela 1 - Características do *Design Science Research*

Características do DSR	Descrição
Pragmatismo	O método DSR procura aprimorar tanto a teoria como a prática. O valor da teoria é avaliada pelo grau em que os seus princípios informam e melhoram a prática.
Relevância do Problema	O objetivo do método é desenvolver soluções baseadas em tecnologia para problemas importantes e relevantes.
Flexibilidade e interatividade	Quem realiza a pesquisa está envolvido em todos os processos do projeto trabalhando todos em conjunto. Os procedimentos de pesquisa são flexíveis. Várias técnicas para a recolha e análise dos dados podem ser aplicadas.
Avaliação do Design	A utilidade, qualidade e eficácia do artefacto devem ser, rigorosamente, demonstradas por meio de métodos de avaliação bem executados.
Contribuições do Design	O método DSR deve promover contribuições claras e verificáveis nas áreas específicas dos artefactos desenvolvidos, nas fundamentações de design e/ou nas metodologias de design.
Rigor da Pesquisa	A pesquisa baseia-se na aplicação de rigorosos métodos na construção e na avaliação do design do artefacto.
Design como um Processo de Pesquisa	A procura por um artefacto eficaz exige o uso de meios disponíveis para alcançar os fins desejados, desde que satisfaçam as leis no ambiente de problema.
Contextualização	O processo de pesquisa, os resultados da investigação e as alterações do plano inicial devem ser documentados. Os resultados da pesquisa estão relacionados com o processo de design e configuração.
Comunicação da Pesquisa	A pesquisa deve ser apresentada para o público orientado à tecnologia bem como para os orientados à gestão.

Este processo conceptual permite aos investigadores o adequado desenvolvimento de um artefacto em DSR e um modelo mental para a sua apresentação. Uma vez que o objetivo deste trabalho passa por recolher quais as melhores práticas/metodologias no que diz respeito à análise de dados e a forma como estes vão ser trabalhados e estruturados para apresentação, decidiu-se assim seguir os passos aconselhados pelo método DSR.

Depois da análise de diferentes variantes desenvolvidas sobre a metodologia DSR, verificou-se que existe uma abordagem proposta pelos autores Vaishnavi & Kuechler (2004) que suporta os objetivos fundamentais necessários ao desenvolvimento de projetos na área dos Sistemas de Informação de modo a facultar um processo devidamente comprovado para a realização da investigação. Dessa forma, nesta dissertação irá ser seguido este modelo.

2.1. Método Design Science Research

A abordagem seguida é constituída por cinco fases distintas. Estas fases são descritas de seguida:

1. Consciencialização do problema - O foco nesta etapa para quem está a efetuar o trabalho de pesquisa passa por identificar e compreender o problema de estudo e para o qual pretende apresentar uma solução, assim como deve definir qual a performance necessária para o sistema em estudo;
2. Sugestão - pretende-se que sejam propostas soluções exequíveis para o problema que está a ser estudado. Esta etapa deve ser realizada utilizando-se como base o método científico abduutivo, ou seja, quem está a sugerir uma possível solução terá a necessidade de usar a sua criatividade e os conhecimentos/experiências de modo a apresentar soluções capazes e que possam ser usadas na melhoria da solução atual
3. Desenvolvimento - consiste em utilizar um dos artefactos que foram propostos na etapa anterior e tentar solucionar esse problema. Se os desenvolvimentos se mostrarem adequados e capazes, serão futuramente avaliados (quarta etapa). Contudo, se no processo de desenvolvimento, ou mesmo na avaliação, o artefacto não se mostrar adequado como uma solução às necessidades da pesquisa, o investigador terá a opção de retornar à etapa de consciencialização, com o intuito de compreender melhor o problema e, posteriormente, dar continuidade à pesquisa
4. Avaliação – permite gerar novas aprendizagens e conhecimentos, não só para quem efetua a investigação, mas também para quem têm acesso à mesma. Na figura 1, as setas representam as interações entre as etapas, essas interações são denominadas de circunscrição. O processo de circunscrição é fundamental para aperfeiçoar a compreensão da pesquisa efetuada, pois permite que quem tenha acesso ao artefacto compreenda e aprenda com o processo apresentado. Outra vantagem passa pelo facto de que quem efetua o trabalho de pesquisa possa também aprender com o que não funcionou, fazendo um contraponto dos seus resultados com a teoria existente; e
5. Conclusão - são apresentados os resultados obtidos, podendo estar ou não em concordância com os resultados encontrados, é neste etapa onde o autor da pesquisa se consciencializa se obteve ou não sucesso no desenvolvimento do artefacto. Devido a este motivo é possível retornar ao ciclo de *design* e eventualmente gerar contribuições a respeito de lacunas existentes na teoria, e que, no momento da conscientização, por sua falta, podem resultar num artefacto inadequado para resolver o problema de estudo.

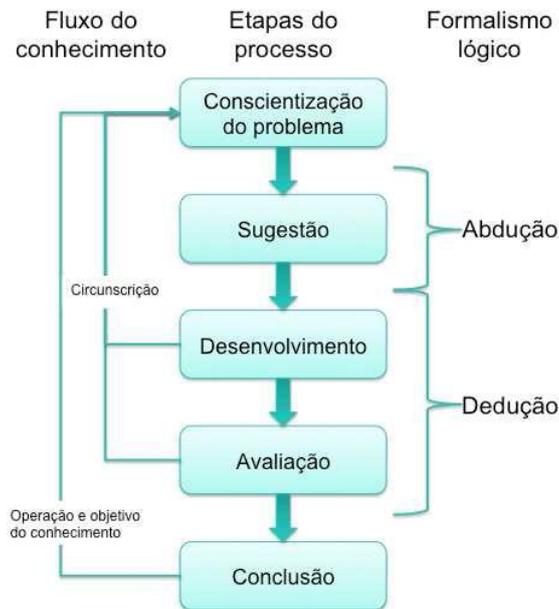


Figura 1: Fases DSR, adaptado de (Vaishnavi & Kuechler, 2004)

A abordagem adotada no desenvolvimento deste trabalho de investigação foi a abordagem proposta por (Vaishnavi & Kuechler, 2004) com uma pequena variante: surgiu a necessidade de efetuar uma segunda iteração da revisão bibliográfica devido ao facto de terem surgido algumas sugestões nos questionários que suscitaram uma nova pesquisa onde foi identificado outro conjunto de regras.

2.2. Outputs do método Design Science Research

No método DSR para apresentar os diferentes tipos de processos e resultados do artefacto de desenvolvimento utilizaremos os artefactos propostos pela teoria desenvolvida por (March & Smith, 1995):

Segundo March & Smith (1995) existem quatro artefactos ou outputs relevantes derivados do método DSR: são os constructos, modelos, métodos ou instâncias, ou uma combinação dos mesmos. Quanto à avaliação de artefactos de *design*, essa categorização pode servir como um guia com o objetivo de avaliar a utilidade ou a qualidade global do artefacto projetado (ou seja, construções (conceitos), modelos, métodos ou instâncias) para resolver o problema que foi formulado no início do processo de investigação. Na tabela seguinte aparece uma descrição geral dos quatro diferentes tipos de artefactos de *design* e suas definições por (March & Smith, 1995):

Tabela 2 - Outputs do método DSR

Artefacto Design	Definição baseada em March & Smith (1995)
Constructo (Conceitos)	Constructos ou conceitos, no contexto da <i>DSR</i> , são compreendidos como o vocabulário de um domínio. São aplicados para descrever os problemas dentro do domínio e para especificar as respetivas soluções. São ainda utilizados para descrever e pensar sobre as tarefas.
Modelos	Os modelos são compreendidos como um “conjunto de proposições ou declarações que expressam as relações entre os constructos”. Quando queremos retratar/representar/descrever uma realidade utilizamos os modelos, tanto apresentam as variáveis de determinado sistema, como também suas relações. Um modelo deve tentar captar sempre a realidade, procurando os detalhes e a precisão de modo a fornecer utilidade.
Métodos	Representam um conjunto de passos (um algoritmo) usado para executar uma tarefa. Os métodos são baseados num conjunto de construções subjacentes (linguagem) e uma representação (modelo) da solução. Podem ser representados graficamente ou em algoritmos específicos. Os artefactos deste tipo podem estar diretamente ligados aos modelos, havendo uma interação. A principal vantagem dos métodos passa por favorecer a transformação dos sistemas e conseqüente procura pela melhoria.
Instâncias	São as realizações de um artefacto no seu ambiente, ou seja, um conjunto de regras que orientam a utilização de um artefacto num ambiente real. As instâncias tratam de transmitir como implementar os artefactos desenvolvidos e os possíveis resultados esperados.

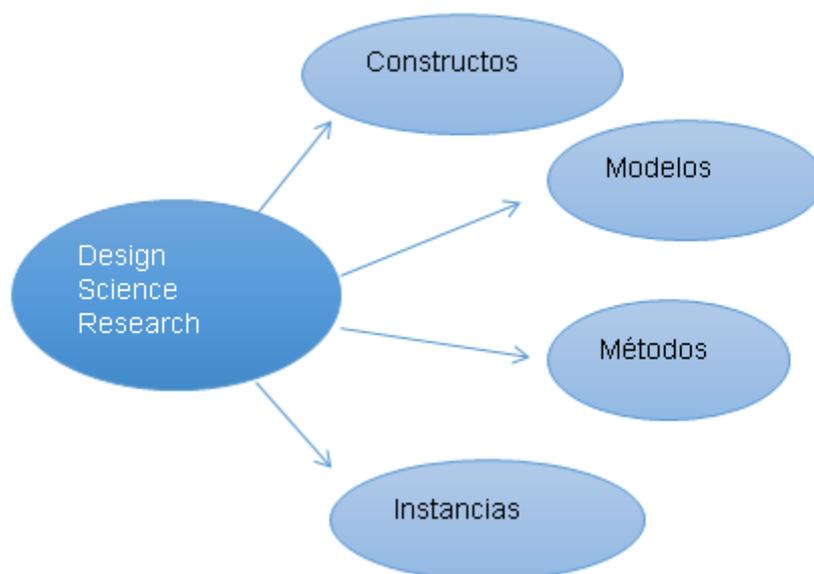


Figura 2: Outputs do método DSR adaptado de (Vaishnavi & Kuechler, 2004)

A aplicação do método DSR é essencial no âmbito deste trabalho de dissertação pois pretende distinguir os vários resultados e proporcionar uma perceção de qual será o output resultante do desenvolvimento do artefacto desta investigação, o “*Framework Conceptual Para Visualização da Informação (VI)*”.

O que se pretende da execução do Framework conceptual desta investigação, é uma instanciação porque operacionaliza construções (arquiteturas) e modelos de domínio da questão de investigação, ao nível da análise de dados e a posterior visualização da informação, fazendo a relação dos conceitos relevantes através do estudo e investigação através de uma revisão bibliográfica.

2.3. Aplicação do método de Investigação

Depois da compreensão do método DSR nomeadamente do modelo proposto pelos autores Vaishnavi & Kuechler (2004) e que será seguido neste trabalho, e da exibição dos outputs que resultam da realização deste trabalho, será apresentado o planeamento e execução das diferentes fases do método devidamente ajustadas ao contexto do problema.

2.4. Consciencialização do problema

É neste momento que se inicia a construção formal ou informal do processo de investigação. Nessa fase, é requisitada a análise de um conjunto de conceitos, teorias e relações verificadas e testadas, sendo úteis para explicar processos e resultados organizacionais. Os autores Vaishnavi & Kuechler (2004), explicam que nesta fase deve-se identificar e compreender o problema de estudo e para o qual se pretende apresentar uma solução.

Este trabalho de investigação é direccionado para a compreensão de trabalhos relevantes previamente realizados na área *BI*, nomeadamente *Data Analytics* e *VI*.

Sendo assim devido à necessidade de obter uma base suficientemente rica e estruturada, de modo a executar o desenvolvimento das diferentes tarefas com eficiência e eficácia, foi realizada uma pesquisa e revisão bibliográfica, que servirá de suporte ao desenvolvimento deste trabalho, pesquisa essa que será apresentada numa matriz com os conceitos, autores e anos das bibliografias recolhidas.

O que se pretende nesta fase é uma pesquisa e recolha de artigos bibliográficos, ou publicações, relacionadas com *Data Analytics*, VI e todos os processos implícitos que estão relacionados.

Métodos de pesquisa: Devido a complexidade da questão, a pesquisa de artigos/bibliografias foi efetuada utilizando o *Google Scholar*, *Microsoft Academic* e algumas sugestões de publicações/*papers* referenciadas por um autor credenciado na matéria, as palavras-chaves utilizadas foram *Data Analysis*, *Data Analytics*, *Information Visualization*, *Data Analysis Theories*, *Information Visualization Theories*, *InfoViz*. De modo a executar uma correta seleção de conteúdo, procedeu-se à leitura e procura das palavras-chaves no *abstract* dos documentos de modo a recolher apenas os que se enquadram no âmbito do trabalho. Para finalizar os artigos foram categorizados por autor, conceito e ano.

Tabela 3 - Matriz de Conceitos

Conceitos Autores	<i>VI</i>	<i>Data Analytics</i>	<i>Big Data</i>	<i>Dashboard Design</i>	<i>Percepção Visual</i>	<i>DSR</i>
Bezerianos, M. E. (2010)	X					
C. Ahlberg, E. W. (1995)		X				
Center, I. I. (2012)			X			
Center, M. &. (2006)		X				
Dinu Airinei, D. H. (s.d.)	X					
E.Micheline, B. (s.d.)	X			X		
Few, S. (2006)	X			X		
Few, S. (2007)	X			X		
Few, S. (2011)	X			X		
H. Lam, E. B. (2010)	X					
Herman, M. (2013)			X			
Huang, W. (2009)	X					
Hullman, J. (s.d.)	X					
J.Myatt, G. (2007)		X				
K.Robert, M. (2013)	X			X		

Marianne Jorgensen, L. J. (s.d.)	X					
Maureen, S. (2010)	X					
McIntosh, L. (2013)				X		
O'Connor, H. (s.d.)		X				
Oracle. (2013)			X			
Pamela A. Dupin-Bryant, D. H. (2014)	X	X				
paper, W. (2013)	X					
Raden, N. (2000)	x					
Blewett, R. (2011)	X					
Chang, D., Dooley, L., & Tuovinen, J. E. (2002)					X	
Geman, D., Glass, L., Murray, J. D., Kohn, R. V, Sastry, S. S., & Krishnaprasad, P. S. (2000)					X	
Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004).						X
IBM. (2012)				X		
March, S. T., & Smith, G. F. (1995)						X
Matange, S. (2013)	X					
Schoenbach, V. J. (2014)		X				

Stodder, D. (2013)	X					
Stone, M. (2009)	X					
Vaishnavi, V. K., & Kuechler, B. (2004)						X

Sugestão

Após a fase de identificação do problema, inicia-se o momento de definição de sugestão. Trata-se de uma etapa criativa onde se pretende planear o desenho de cada parte do artefacto, formular novas configurações de elementos existentes ou mesmo conceção de novas teorias ou modelos.

Nesta fase pretende-se apresentar os objetivos definidos para a realização da dissertação, bem como planear o seu desenvolvimento. Numa primeira fase foi recolhido um grande volume de informação relacionada com o tema, após a exposição da ideia de desenvolvimento ao orientador de pesquisa, o próximo passo cingiu-se na discussão da viabilidade do caminho a seguir para depois passar a um desenvolvimento sustentado. Após realizados os testes para comprovar a exequibilidade, estes eram reportados ao orientador do trabalho para validar ou não as opções tomadas.

Desenvolvimento

A etapa do desenvolvimento é a efetiva construção do artefacto, podendo ser um ou um conjunto de soluções para resolver o problema proposto, para isso, dependerá dos artefactos que foram propostos na etapa anterior e tentar solucionar esse problema.

O desenvolvimento deste trabalho passa por criar um *framework* que facilite e sirva de auxílio ao desenvolvimento de modelos de VI onde contém as abordagens mais apropriadas para a análise de dados, baseando-se em diversos tipos de dados, e posterior visualização da informação. O objetivo principal passa por mostrar como potenciar os pontos fortes para o utilizador usufruir de um serviço sintetizado, eficaz, coerente e completo

Avaliação

A avaliação é definida como um processo exigente de verificação do comportamento do artefacto no ambiente para o qual foi idealizado. Uma série de procedimentos é necessária para verificar o desempenho do artefacto.

Esta fase é essencial pois permite que quem não participou no desenvolvimento, ou quem não é um especialista na área, possa compreender na perfeição o processo apresentado.

A avaliação do método desenvolvido será contínua ao longo do processo de desenvolvimento e sofrerá os ajustes necessários pela parte do criador do trabalho, discutidos em conjunto o orientador. Em suma, será avaliado por outros peritos na área de *Business Intelligence* e *Data Analytics*.

A utilidade, qualidade e eficácia da DSR devem ser demonstradas rigorosamente por meio de métodos definidos para avaliação do resultado produzido (Hevner et al, 2004). A avaliação é uma componente essencial no método de pesquisa. A avaliação do resultado da DSR é fundamentada nas exigências empresariais que frequentemente ocorrem no contexto da utilidade, qualidade, beleza (estilo) do artefacto produzido, inclui também a integração do artefacto com a infraestrutura técnica do ambiente do negócio. De acordo com (Joahnssen, 2000) no ambiente da TI, os artefactos podem ser avaliados em termos da sua funcionalidade, da perfeição, da consistência, precisão, desempenho, confiabilidade, “*usability*”, ajuste à organização, entre outros atributos de qualidade pertinentes. Quando métricas analíticas são adequadas, os artefactos projetados podem ser avaliados matematicamente.

A avaliação de artefactos projetados é feita recorrendo a metodologias disponíveis na área científica. A tabela 4 mostra um resumo delas, segundo (Hevner et al 2004). A escolha do

método a utilizar é importante, pois o mesmo deve ser o mais adequado aos objetivos da avaliação. No caso de testes, do tipo “*black box*”, há um dispositivo ou objeto que exhibe apenas os dados de entrada e de saída, sem que se tenha conhecimento do processamento interno. Já o teste “*white box*” envolve um teste estrutural ou orientado à lógica para avaliar o comportamento interno do componente de *software*. Essa técnica trabalha diretamente sobre o código fonte para testar, por exemplo, o fluxo de dados, os caminhos lógicos e códigos específicos.

Tabela 4 - Métodos de Avaliação

Métodos de Avaliação	
Observação	Estudo de caso: estudo profundo do artefacto no ambiente da empresa.
	Estudo de campo: monitorar o uso do artefacto em múltiplos projetos.
Analítica	Análise estática: exame da estrutura do artefacto referente a qualidades estáticas (por exemplo: complexidade)
	Análise da arquitetura: estudo do ajuste do artefacto à arquitetura do SI
	Otimização: demonstração da otimização das propriedades do artefacto
	Análise dinâmica: estudo das qualidades dinâmicas do artefacto em uso (exemplo: performance)
Experimental	Experimento controlado: estudo do artefacto em ambiente controlado para análise das suas propriedades, como por exemplo: usabilidade
	Simulação: análise do artefacto com dados artificiais
Testes	Teste funcional (Black Box): execução do artefacto para descobrir falhas e Identificar defeitos por meio de dispositivos específicos.
	Teste estrutural (White Box): teste de desempenho em relação a métricas na implementação do artefacto (por exemplo: teste de endereços)
Argumentação	Argumentação: uso de informação com base científica para construir um argumento convincente da utilidade do artefacto

	Cenários: construção detalhada de cenários em torno do artefacto para demonstrar sua utilidade
--	--

Quanto á avaliação dos artefactos produzidos neste trabalho, seriam necessários dois tipos de avaliação: primeiro é importante avaliar os diferentes fatores a colocar no *framework*, fatores esses que foram recolhidos através de uma pesquisa bibliográfica e que foram avaliados através de métodos de avaliação analíticos, neste caso foi realizado um questionário a toda a comunidade estudante da Universidade do Minho para comprovar a importância e impacto dos diferentes fatores no desenvolvimento de métodos de VI sendo que futuramente é essencial realizar um segundo questionário a peritos na área; em segundo lugar seria importante avaliar a aplicação do *framework* em contextos reais, neste caso usando os métodos de avaliação de observação, experimental e de testes. Como o tempo de desenvolvimento é relativamente curto, seria interessante como trabalho futuro aplicar o *framework* desenvolvido e testar o resultado do mesmo.

Conclusão

A conclusão da DSR ocorre, do ponto de vista científico, para apresentar os principais resultados do estudo. É nesta fase que deverá ser apresentada a importância do trabalho realizado, utilidade, rigor e a inovação que proporciona na área de Sistemas de informação. São apresentados os resultados obtidos e onde o autor da pesquisa se consciencializa se obteve ou não sucesso no desenvolvimento do artefacto.

A conclusão será apresentada no fim do desenvolvimento da dissertação.

3 REVISÃO DE LITERATURA

Data Analytics (DA) e VI são os principais conceitos sobre os quais este trabalho é desenvolvido. Neste capítulo, são apresentadas definições de conceitos importantes relacionados com DA e VI, seguidos do desenvolvimento do tema ao qual este trabalho se propõe.

3.1. Conceitos Importantes

Nesta secção são definidos alguns conceitos importantes relacionados com VI.

Data Analytics

Pode-se definir DA como o processo usado para interagir, analisar e descobrir informação relacionada com determinada estrutura de dados, resultando em decisões apropriadas ao negócio. Existem várias técnicas de análise de dados com diferentes abordagens, mas todas elas, são direcionadas para a recolha, inspeção, tratamento e posterior apresentação da informação com o objetivo de descobrir informações úteis, sugerindo conclusões, e apoiar a tomada de decisão.

Para realizar a DA existem alguns passos que devem ser seguidos de modo a que a mesma seja feita corretamente e de uma forma estruturada (Schoenbach, 2014):

- Recolha dos dados;
- Seleção dos dados;
- Transformação dos dados (verificar erros, incoerências, etc.);
- Testar os dados;
- Análise final dos dados.

A DA tem como maiores objetivos (Schoenbach, 2014):

- Avaliar e aumentar a qualidade dos dados;
- Descrever um determinado estudo e relaciona-lo com os existentes;
- Medidas de previsão e estimativa;
- Identificar pontos críticos;
- Avaliar hipóteses;
- Avaliar graus de incerteza;
- Avaliar o impacto ou importância de determinada informação.

Big Data

Big Data é o termo usado para definir e descrever o crescimento exponencial e disponibilidade do volume de dados, quer sejam estruturados ou não estruturados. São conjuntos de dados cujo tamanho vai além da capacidade física das ferramentas comuns de recolha, armazenamento, gestão e análise da base de dados.

O autor Laney (2001) concebeu a “Teoria dos três V’s” para definir Big Data:

- **Volume:** Vários fatores contribuem para o aumento do volume de dados, transações armazenadas ao longo dos anos, transmissão de dados não estruturados, etc. Pré Big Data, o volume de dados excessivo era um problema de armazenamento, mas com a diminuição dos custos de armazenamento, outras questões emergem, incluindo a forma de determinar a relevância dentro de grandes volumes de dados e como a análise para criar valor a partir de grandes volumes de dados. É o tamanho dos dados que determina o valor e o potencial dos dados em questão e se eles realmente podem ser considerados como Big Data.
- **Velocidade:** Atualmente, os dados são transmitidos a uma grande velocidade, então, devem ser tratados o mais rapidamente possível. Reagir rapidamente o suficiente para lidar com a velocidade de dados é um desafio para a maioria das organizações. Neste contexto, "velocidade" refere-se à velocidade da geração de dados ou o quão rápido os dados são gerados e processados para atender as demandas e os desafios que temos pela frente no caminho do crescimento e desenvolvimento.
- **Variedade:** Variedade refere-se às muitas fontes e tipos de dados estruturados e não estruturados. Estamos acostumados a armazenar informação de bases de dados e folhas de Excel por exemplo, mas atualmente, os dados surgem em diferentes formatos, na forma de e-mails, fotos, vídeos, PDFs, áudio, etc. Esta variedade de dados não estruturados cria problemas para armazenamento de dados, mineração e análise. O trabalho das organizações passa por organizar esses dados e torná-los acessíveis, isso seria um processo fácil se a informação estivesse toda no mesmo formato, não sendo o caso na maioria das vezes e sendo esse o desafio a superar com o *Big Data*.

Data Lake

O conceito de *Data Lake* está relacionado com o *Apache Hadoop* e os seus projetos *open source*. Este conceito surgiu devido à necessidade capturar novos tipos de dados em grande volume. Um *Data Lake* é um enorme repositório de dados, de fácil acesso e que centraliza grandes volumes de dados, quer sejam estruturados ou não estruturados.

Características de um *Data Lake*:

- Capturar e armazenar dados brutos com baixo custo;
- Armazenar vários tipos de dados no mesmo local;
- Realizar transformações sobre os dados;
- Definir a estrutura dos dados;
- Realizar novos tipos de processamentos aos dados;
- Executar a análise sobre os dados.

A arquitetura de um *Data Lake* é uma abordagem de armazenamento orientada para *Big Data*. Os dados não são classificados quando são armazenados no repositório devido ao valor dos dados não ser claro no início. Como resultado, a preparação de dados é eliminada. Um *Data Lake* é, portanto, menos estruturado comparado com um *Data Warehouse* (repositório de dados) convencional. Quando se acede aos dados estes só então são classificados, organizados ou analisados.

3.2. Design/Visualization Theories and Principles

Gestalt Theory

Gestalt é um termo da psicologia que significa "todo unificado". Refere-se a teorias de percepção visual desenvolvidas por psicólogos alemães na década de 1920. Essas teorias tentam descrever como as pessoas tendem a organizar elementos visuais em grupos ou conjuntos unificados quando certos princípios são aplicados (Chang, Dooley, & Tuovinen, 2002; Geman et al., 2000). Estes princípios são:

- **Similaridade:** Similaridade ocorre quando os objetos parecem semelhantes um ao outro. As pessoas muitas vezes compreendem-nos como um grupo ou padrão.



Figura 3 - Gestalt Theory: Similaridade

O exemplo acima (contem 11 objetos distintos) e apresenta-se como um simples objeto pois todas as formas têm similaridade. Este processo acontece porque as formas triangulares na parte inferior do símbolo da águia são semelhante às formas que formam o “sol” que a rodeia.

- **Continuação:** Continuação ocorre quando o olho é obrigado a passar por um objeto e é “obrigado” a continuar para outro objeto.



Figura 4 - Gestalt Theory: Continuação

Nesta imagem temos um exemplo da continuação, porque o olho de quem observa, naturalmente, segue uma linha ou curva. A barra de fluxo suave do "H" conduz o olhar diretamente para a folha de plátano.

- **Encerramento:** Encerramento ocorre quando um objeto é incompleto ou um espaço não está completamente fechado. Se é apresentada uma grande percentagem da imagem preenchida, quem visualiza compreende o restante preenchimento em falta.



Figura 5 - Gestalt Theory: Encerramento

Embora, na figura 5, o panda não esteja completo, mesmo assim está o suficiente para o olho conseguir completar a forma. Quando a percepção do observador capta uma forma completa, ocorre o “encerramento”.

C.R.A.P Theory for Effective Visual Design

Os quatro princípios básicos e essenciais presentes em muitas teorias de aprendizagem são Contraste, Repetição, Alinhamento e Proximidade (CRAP). Estes princípios devem ser utilizados de modo a melhorar a leitura e compreensão no que toca ao Design.

De acordo com (Williams, 2004), os princípios são:

- Contraste - Se há elementos iguais pode-se utilizar contraste para torná-los diferentes. O contraste atrai a atenção de quem visualiza, ver Figura 6.



Figura 6 - CRAP Theory : Contraste

- Repetição - Repetir um elemento provoca uma sensação de consistência. A repetição desenvolve a organização e cria uma marca forte, ver Figura 7.

REPETITION REPETITION REPETITION

Figura 7 - CRAP Theory : Repetição

- Alinhamento - Não devem ser colocados elementos de uma forma aleatória ou simplesmente tentar preencher o espaço vazio. Cada elemento deve ter uma ligação visual com os restantes elementos. O alinhamento dos elementos transmite a sensação de estarmos na presença de uma unidade forte, coesa e com ordem, ver figura 8.

ALIGN MENT

Figura 8 - CRAP Theory : Alinhamento

- Proximidade - Itens relacionados entre si, devem ser agrupados. O agrupamento organiza as informações e reduz a desordem. Os produtos que não estão relacionados uns com os outros não devem estar na proximidade, ver figura 9 e 10. A proximidade ocorre quando os elementos são colocados juntos e tendem a ser compreendidos como um grupo, ver figura 11 e 12.

PROX IMITY PROX IMITY PROX IMITY

Figura 9 - CRAP Theory : Proximidade

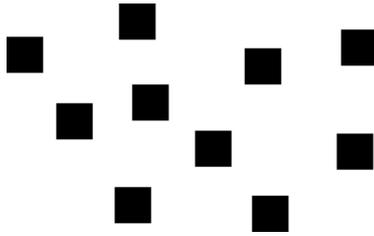


Figura 10 - CRAP Theory : Proximidade 2

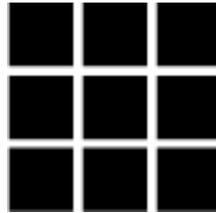


Figura 11 - CRAP Theory : Proximidade 3



Figura 12 - CRAP Theory : Proximidade 4

- Forma e fundo: O olho diferencia as formas dos objetos, a sua forma, silhueta ou sombra e possibilita a compreensão de figuras (objetos), enquanto a área circundante é definida como fundo. Balancear a figura e o fundo pode fazer com que a forma da imagem seja compreendida de forma mais clara. Na figura 13, a palavra é claramente percebida como figura devido ao espaço em branco que circunda as letras. Na figura 14, as relações forma-fundo mudam à medida que o olho percebe a forma de uma sombra ou a silhueta de um rosto. Na figura 15, são usadas relações complexas de forma-fundo, que mudam a percepção da imagem desde as folhas em cima, água em baixo e do tronco de uma árvore.

Figure

Figura 13 - CRAP Theory : Forma e fundo



Figura 14 - CRAP Theory : Forma e fundo 2



Figura 15 - CRAP Theory : Forma e fundo 3

Dieter Ram's Principles of Good Design

O que faz um bom *design*? Quais os princípios gerais que nos podem guiar para concluirmos o que faz um projeto melhor do que outro? Era o que Rams (1980) se questionava.

Em 1980, preocupado com o estado do mundo e com a sua contribuição como um *designer*, Rams (1980) desenvolveu o que ele considerava ser os 10 princípios para um bom design.

Esta teoria surgiu antes da internet e por isso antes de *web design*, sendo que hoje ainda existe quem os aplique.

De seguida apresentam-se os 10 princípios (Rams, 1980):

- Um bom *design* é inovador - o *design* deve evoluir em conjunto com a tecnologia, *design* inovador tem o elemento surpresa, que incentiva a aprendizagem e melhora a capacidade de decodificar novas informações, ver figura 16.



Figura 16 - Dieter Ram's Principles: Inovação

- Um bom *Design* torna o produto útil - no *design*, forma, de facto, tem que seguir a função. Quando o *design* se torna um obstáculo ao consumidor impedindo-o ou dificultando a utilização do produto de forma eficaz, está a falhar. Este ponto salienta que o *design* é suposto trabalhar com o produto, não contra ele, ver figura 17.



Figura 17 - Dieter Ram's Principles: Usabilidade

- Bom *Design* é estético - A qualidade estética do produto é uma parte integrante da sua utilidade. Apenas objetos bem executados podem ser bonitos. Estética e utilidade andam de mãos dadas, ver figura 18.



Figura 18 - Dieter Ram's Principles: Estética

- Um bom *Design* torna o produto compreendido - Esclarece o produto e na melhor das hipóteses um bom *design* torna-o autoexplicativo ou seja, não são necessárias instruções. Um bom *design* pode fazer isso através da clarificação estrutura e função do produto, mantendo a intuição do utilizador em mente, ver figura 19.



Figura 19- Dieter Ram's Principles: Compreendido

- Um bom *Design* é discreto - Os produtos projetados devem ser pensados como ferramentas, e não como obras de arte ou objetos decorativos. Como tal, devem ser o mais simples possível, ver figura 20.



Figura 20 - Dieter Ram's Principles: Discreto

- Um bom *Design* é honesto - O *design* não deve fazer um produto parecer mais inovador, poderoso e valioso do que realmente é. Não implicar “promessas” que não possam ser cumpridas, ver figura 21.



Figura 21 - Dieter Ram's Principles: Honesto

- Um bom *Design* é de longa duração – tentar somente acompanhar as tendências de “moda” em certa altura serão desatualizados; aqueles que se esforçam na direção da funcionalidade vão durar. Esta é uma boa razão para evitar elementos muito modernos no desenvolvimento, ver figura 22.



Figura 22 - Dieter Ram's Principles: Duradouro

- Um bom *Design* é completo até ao último detalhe - Nada deve ser arbitrário ou deixado ao acaso. Cuidados e precisão mostram respeito para com o consumidor, ver figura 23.

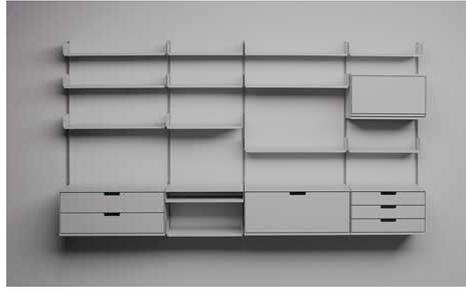


Figura 23 - Dieter Ram's Principles: Detalhado

- Um bom *Design* é amigo do ambiente - Rams pensava na sustentabilidade ambiental, bem antes de este ser amplamente divulgado. Também se preocupava com um tipo diferente de poluição: poluição visual. Este conceito diz que se deve conservar os recursos e minimizar a poluição física e visual durante todo o ciclo de vida do produto, ver figura 24.

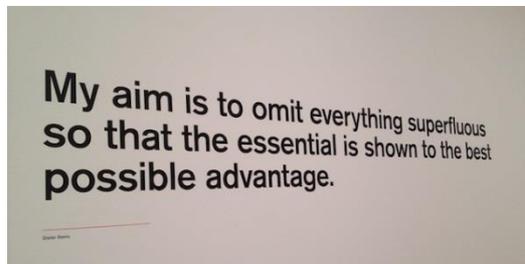


Figura 24 - Dieter Ram's Principles: Amigo do ambiente

- Um bom *Design* é o mais simples possível - Menos, mas melhor, concentrar-se apenas nos aspectos essenciais, e não sobrecarregar com não-essenciais. Voltar para a pureza, de volta à simplicidade.

Podendo ou não concordar com todos os aspectos dos princípios acima, há muita sabedoria neles. Embora esta teoria tenha sido criada pré web, estes conceitos ainda se aplicam em projetos recentes. Por exemplo, estes princípios estão no coração da filosofia de *design* atual da Apple, com os resultados comprovados e aparentemente muitas outras empresas seguem o seu exemplo para produzir um melhor design.

4 DATA ANALYTICS: VISUALIZAÇÃO DA INFORMAÇÃO

4.1. Como resolver o Problema da má utilização da VI

De modo a prevenir a má exposição da informação e consequente má interpretação, ou não obter o máximo partido da mesma, é fundamental definir um conjunto de boas práticas a seguir que, podem minimizar estes riscos. O objetivo desta definição passa por apresentar as propriedades mais relevantes, propriedades essas que devem ser rastreadas para projetar e desenvolver a infraestrutura necessária para a recolha dos dados apropriados e torná-los acessíveis:

- Definir um conjunto de normas para o correto desenvolvimento de métodos de VI;
- Compreender os custos dessas métricas;
- Ter expectativas realistas quanto aos benefícios trazidos;
- Não definir o que não se pode medir;
- Construir uma infraestrutura apropriada;
- Aplicar os princípios da percepção visual;
- Apontar os principais erros e corrigi-los.

No fundo, o objetivo passa por conceber um referencial com os melhores métodos e praticas que possibilitem que a informação seja apresentada da melhor forma possível e consequentemente o utilizador tire o máximo partido da informação apresentada. No entanto devemos adaptar-nos às diferentes situações que surgem pois nem todas são passíveis das mesmas interpretações e aplicações.

Os principais desafios no desenvolvimento dos métodos de VI passam por dar ênfase aos dados mais relevantes em relação aos menos relevantes e organizar um grande volume de dados, que por vezes parecem dispare, em algo que faça sentido. É de maior importância entender como e porque é que algo funciona, do que simplesmente ver funciona, para isso será útil compreender e aplicar os princípios da percepção visual (Stodder, 2013).

Quando se transmite informação, através dos métodos de VI, deve-se incluir apenas a informação necessária, não diminuindo o seu significado e usar mecanismos de visualização que possam facilmente ser lidos e compreendidos. Esses mecanismos necessitam de ser excepcionalmente organizados, específicos, personalizados e maximizando todo o espaço disponível (Stone, 2009).

Propondo um método de resolução do problema de VI, conseguiremos transmitir de uma forma mais clara o que está a acontecer. Dando o exemplo de um painel de um automóvel que nos fornece

medidas como velocidade, combustível entre outras, um *Dashboard*, por exemplo, fornece uma visão geral dos dados que podem ser assimilados rapidamente e que necessitam de uma maior atenção.

De seguida são apresentados alguns exemplos de problemas aos quais os modelos de visualização da informação ajudam a responder:

Tabela 5 - Soluções VI

Problema	Solução
As métricas chave do negócio são abstratas e difíceis para os funcionários interiorizarem.	Visualizações simples e padronizadas que todos conseguem entender.
Quem realiza a análise precisa de orientação.	Definir e comunica metas de negócio.
Descrever as principais métricas é um processo lento.	Construir modelos de visualização simples que com alteração de dados atualizem automaticamente.
Aceder a informação através de múltiplas plataformas através de um login é um risco de segurança.	Reunir os dados mais importantes e uteis derivados de diferentes serviços em um único local.
As tendências não podem ser identificadas facilmente entre sistemas.	Recolher dados de múltiplas fontes dentro de um único local para de imediato obtermos um melhor contexto.

4.2. Desenvolvimento

O tipo de comunicação gráfica que normalmente é necessária no negócio, não é difícil de aprender a fazer, no entanto não surge naturalmente, projetar gráficos individuais é uma tarefa simples em comparação com a conceção de modelos de visualização da informação complexos, mas a conceção de ambos, deve seguir regras de modo a que a exposição da informação possa ser aproveitada da melhor forma (IBM, 2012).

Um modelo eficaz permite ao utilizador visualizar com rapidez, precisão e sem distrações um conjunto de tarefas relevantes e atuais, por exemplo os custos de gestão ou qualquer outra informação crítica cujos efeitos sejam necessários para atingir os objetivos de negócio.

Os modelos são otimizados para ajudar a acelerar a avaliação das tendências e estatísticas atuais, fazendo com que a informação seja acessível para os colegas e clientes (Few, 2007).

Associadas a esta definição, está um conjunto de características e boas práticas que devem ser seguidas na conceção dos métodos de VI (Few, 2007; Matange, 2013):

- **Selecionar o que visualizar** - selecionar quais as medidas, indicadores devem ser incluídos nos modelos. Deve-se ser altamente seletivo para determinar o que é importante ser visualizado.

- **Optar pela simplicidade** - os modelos de VI são concebidos com o intuito de serem fáceis e rápidos de interpretar, um exemplo de uma má implementação deste ponto são os relatórios e tabelas baseadas sobretudo em texto.
- **Ser interativo** - um modelo deve ser concebido para ser interativo, o que permitirá que diferentes utilizadores com diferentes objetivos de análise possam, na mesma página, através de uma personalização obter a informação de que necessitam.
- **Estar atualizado** - uma das principais preocupações na conceção dos modelos de VI passa por apresentar apenas informação útil e atualizada de modo a refletir os desafios atuais dos negócios.
- **Tornar o acesso e utilização simples** - tornar os modelos acessíveis é fundamental, essencialmente com distribuição web pois deste modo podem extrair constantemente informação atual e seguir os protocolos IT e padrões de segurança.

Existem também um conjunto de erros comuns que devem ser evitados (Few, 2006):

- **Começar com muita complexidade;**
- **Organizar mal os dados;**
- **Usar medidas e indicadores que não estão devidamente compreendidas;**
- **Subestimar o tempo ou recursos necessários para criar e manter os modelos de VI;**
- **Incoerências na definição das medidas, indicadores e objetivos;**
- **Usar gráficos e diagramas mal desenhados e ineficazes;**
- **Não destacar o que é importante de analisar;**
- **Mau uso das cores;**
- **Evitar o “Chart Junk”.**

Posto isto, podemos visualizar de seguida dois exemplos, que fundamentando no que foi descrito até então, apresentam uma conceção mais e menos adequada de dois modelos de VI, neste caso aplicados aos Dashboards:

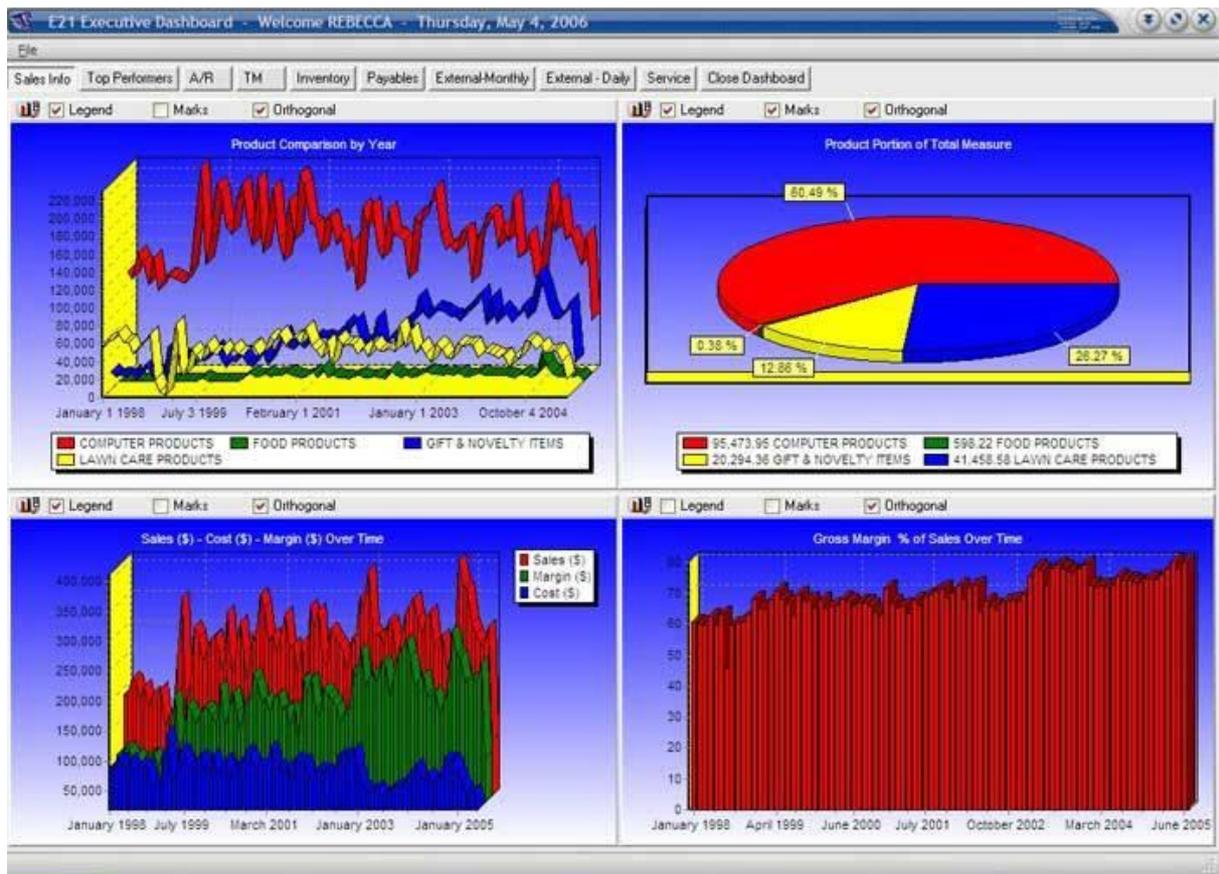


Figura 25 - Má conceção

A Figura 26 é um exemplo de uma conceção de um *Dashboard* menos adequada. Depois de o analisar, verificamos que não obedece às normas que foram descritas anteriormente, nomeadamente uma má seleção das medidas e indicadores, falta de simplicidade pois são usados vários tipos de gráficos que não se complementam entre si o que torna a análise bastante confusa e má organização dos dados. Tem uma apresentação bastante desorganizada, descontextualizado, mau *design*, sendo quase impossível ler a informação. Para um analista trata-se de um *Dashboard* inútil pois falha no momento da comunicação, não é possível identificar rápido e claramente o foco de estudo, apresenta pouca informação e a que apresenta não esta exposta da melhor forma.

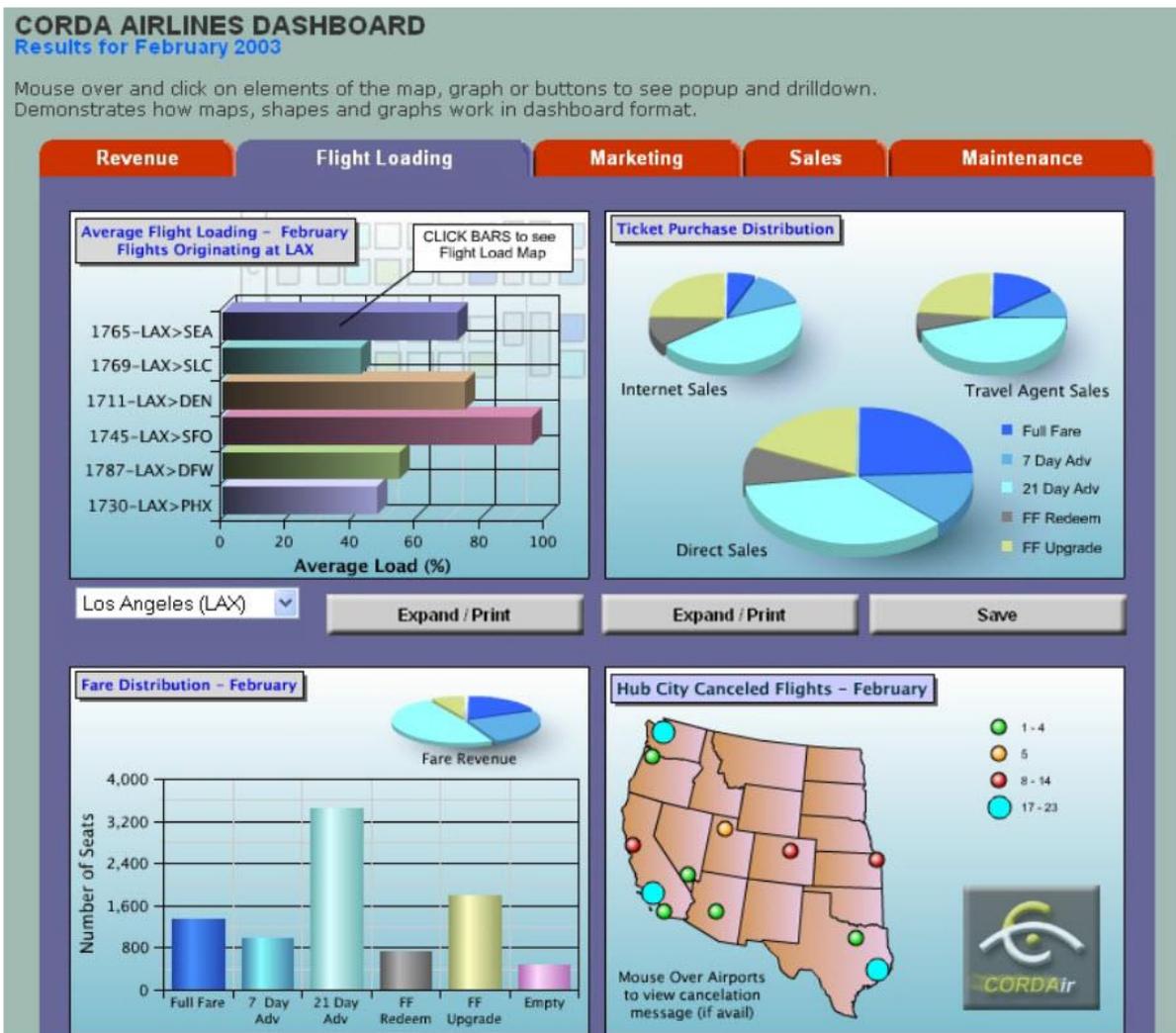


Figura 26 - Boa concepção

Nesta Figura 27 temos um exemplo de um *Dashboard* mais adequado para análise, conseguimos visualizar um contexto e um fio condutor bem delineado. No exemplo em concreto temos a informação importante relacionada com os voos de uma companhia aérea, nomeadamente a percentagem de ocupação de um avião, quais os meios pelos quais se adquirem os bilhetes e a percentagem relacionada, lugares livres e os voos cancelados por cidade.

Tem um ótimo *design*, apresenta vários tipos de gráficos com a possibilidade de se expandirem em outros selecionando as medidas desejadas, apresenta as medidas e indicadores que são importantes para um analista visualizar, é possível identificar rapidamente o problema exposto, simples e bem conseguido, é interativo, permite que diferentes utilizadores com diferentes objetivos de análise, na mesma página, possam personalizar a análise de modo a obter a informação de que necessitam, é coerente, utiliza gráficos bem desenhados e sobretudo destaca o que é fundamental á análise. Esta concepção evita erros como exibir muita complexidade, incoerências e mau uso das cores.

Cada vez mais, os *Dashboards* são vistos como parte integrante da análise e estudo numa organização. Esses *Dashboards* permitem-nos converter dados brutos em informação e traduzir essa informação em conhecimento que resultarão em decisões e ações mais apropriadas. É essencial que os gráficos criados sejam eficazes na comunicação de informação. Um gráfico é eficaz se a informação no gráfico pode ser decodificado com precisão pelo utilizador (Few, 2007, 2008).

Para servir de suporte nestas tarefas, os principais investigadores nesta área formularam as regras apresentadas anteriormente que podem ajudar na criação de *Dashboards* bons e eficazes. Estas regras são concebidas com base na ciência da perceção visual e em estudos sobre a perceção humana. Podem-se aplicar essas regras e diretrizes para criar gráficos que transmitem a informação com precisão, tentando diminuir ao máximo a má interpretação (IBM, 2012).

4.3. Desenvolvimentos futuros

O trabalho aqui apresentado pode servir de suporte para que futuramente se realize um desenvolvimento mais aprofundado acerca da emergente e cada vez mais importante área da VI, podendo também servir de apoio ao desenvolvimento de ferramentas mais sofisticadas e eficientes na produção de formas de VI.

Por outro lado serve de apoio ao desenvolvimento de uma revisão bibliográfica e ao desenvolvimento de um *Framework* que permite estabelecer regras, identificar os pontos fortes e principais lacunas no que diz respeito ao desenvolvimento de metodologias de VI. Então, torna-se fundamental efetuar um levantamento, classificação e registo do histórico da matéria relacionada com VI. A visualização envolve muito mais do que apenas a interpretação dos dados, envolve a perceção humana, *design* e até arte, sendo assim poderá também servir de suporte ao desenvolvimento de trabalhos relacionados com outras áreas e não só a área de *BI*.

4.4. Contribuição

De modo a que VI forneça um suporte a tomada de decisão, é essencial que esta seja composta por vários níveis, onde esses níveis devem comunicar entre eles.

No primeiro nível, aquele que entra em contacto com o utilizador, a VI requer um ambiente interativo, onde é possível configurar várias visualizações, sincronizadas e configuradas rapidamente, de forma a potenciar a rica exploração de dados.

O segundo nível pode ser descrito como o nível de infraestrutura. O fundamental deste nível passa por garantir que dados que são provenientes de diferentes fontes e diferentes formatos podem ser lidos.

O terceiro nível, mas não menos importante passa por construir o modelo de exibição da informação. Estes modelos devem ser construídos baseando-se nas técnicas de visualização apropriadas e nas regras que foram descritas anteriormente. É importante acompanhar o presente, com a evolução, vão surgindo novas técnicas de análise/visualização, logo é fundamental realizar um estudo contínuo e manter-se informado sobre as tendências atuais na área.

Combinar os processos de acesso, processamento e visualização da informação deve ser um procedimento fácil, visual e interativo (Blewett, 2011). A infraestrutura deve ser o mais flexível possível de forma a ser possível a integração com aplicações já existentes, sem grande dificuldade, além de garantir vários níveis de acesso, dependendo do perfil do utilizador final.

Concluindo, uma boa visualização deve potenciar os aspetos de maior importância para quem esta a efetuar a análise, não só uma clara compreensão da informação, mas também ampliar a quantidade de informação que é “compreendida” e “apreendida” através da visualização. Identificar valores máximos e mínimos, desvios, erros, tendências ou intervalos, são algumas das principais metas que estão presentes na VI, pois só deste modo, é possível proporcionar um melhor entendimento de sistemas complexos, descobrir informação “escondida” ou ainda dar apoio a decisões.

4.5. Validação

Introdução

Com o intuito de avaliar as regras apresentadas anteriormente, que influenciam o desenvolvimento de métodos de Visualização de informação, considerou-se fundamental desenvolver um questionário englobando todos os aspetos relacionados com esses modelos, presentes na dissertação.

O referido Questionário foi aplicado de 6 de Julho a 25 de Julho de 2015 e, versou um conjunto de perguntas que permitem identificar a importância desses componentes, avaliando o grau de importância assim como sugerir novos aspetos relevantes.

O questionário classifica a importância de subtópicos relacionados com: um conjunto de características e boas práticas que devem ser seguidas na conceção de métodos de visualização da informação; um conjunto de erros comuns que devem ser evitados na conceção de modelos de visualização da informação; sugestão de características, boas práticas ou erros a evitar que considerem importantes na conceção de métodos de visualização da informação.

Aquando a elaboração das questões e o formato do questionário utilizado foi necessário ter atenção ao tipo de população a inquirir, nomeadamente, a idade, o nível educacional e a motivação.

O questionário contemplou um conjunto de 2 questões e a escala aplicada foi de 1 a 5, sendo o nível 1 correspondente ao nível Pouco importante e 5 Muito Importante.

O questionário foi seguido por um campo, de carácter facultativo, onde podiam ser colocadas sugestões para melhoria ou observações.

Questionário

Questionário VI

Caro/a participante,

este inquérito enquadra-se num estudo relacionado com a **Visualização da Informação** que tem como principal objetivo apresentar uma definição de **Visualização de Informação (VI)** e de todos os conceitos relacionados com a mesma, assim como a importância e impacto que resulta da sua aplicação, realizado no âmbito da Dissertação de Mestrado Integrado em Engenharia e Gestão de Sistemas de Informação na Universidade do Minho, realizada por Luís Rafael Araújo Ribeiro, orientado por Jorge Oliveira e Sá.

O seu preenchimento é voluntário. Não existem respostas certas ou erradas. O questionário é **anónimo** e é garantida toda a **confidencialidade** dos dados.

Este é um estudo conduzido pela Escola de Engenharia da Universidade do Minho, Guimarães, Portugal.

A sua colaboração é fundamental para o estudo da Visualização da Informação.

Desde já agradecemos a sua colaboração.

Pode obter mais informações através do e-mail: a58133@alunos.uminho.pt

Figura 27 - Questionário VI

Nesta secção é apresentado um conjunto de características e boas práticas que devem ser seguidas na conceção de métodos de visualização da informação. Segue uma breve descrição dos indicadores, para uma melhor compreensão do inquirido.

Selecionar o que visualizar - selecionar quais as medidas, indicadores devem ser incluídos nos modelos de visualização da informação.

Optar pela simplicidade - modelos baseados sobretudo em informação gráfica ao invés da informação textual.

Ser interativo - um modelo de visualização da informação interativo permitirá que diferentes utilizadores, com diferentes objetivos de análise, possam personalizar o modelo de modo a obterem a informação de que necessitam.

Estar atualizado - informação útil e atualizada de modo a refletir os desafios atuais do negócio.

Tornar o acesso e utilização simples - tornar os modelos de visualização da informação acessíveis a qualquer tipo de utilizador, essencialmente com distribuição web pois deste modo podem extrair constantemente informação atual e seguir os protocolos e padrões de segurança.

Classifique as seguintes características quanto à importância:

	Pouco Importante			Muito Importante	
	1	2	3	4	5
Selecionar o que visualizar	<input type="radio"/>				
Simplicidade	<input type="radio"/>				
Interatividade	<input type="radio"/>				
Atualizado	<input type="radio"/>				
Acesso e utilização simples	<input type="radio"/>				

Figura 28 - Questionário VI

Existem também um conjunto de erros comuns que devem ser evitados, classifique quanto ao impacto nos modelos de Visualização de Informação:

	Pouco Importante			Muito Importante	
	1	2	3	4	5
Muita complexidade	<input type="radio"/>				
Má organização dos dados	<input type="radio"/>				
Medidas e indicadores incompreendidas	<input type="radio"/>				
Subestimar o tempo ou recursos necessários para criar e manter os métodos de visualização da informação	<input type="radio"/>				
Incoerências na definição das medidas, indicadores e objetivos	<input type="radio"/>				
Usar gráficos e diagramas mal desenhados e ineficazes	<input type="radio"/>				
Não destacar o que é importante de analisar	<input type="radio"/>				
Mau uso das cores	<input type="radio"/>				
'Chart Junk' (elementos que não são fundamentais à compreensão da informação)	<input type="radio"/>				

Figura 29 - Questionário VI

Tem alguma sugestão de características, boas práticas ou erros a evitar que considere importante na conceção de métodos de visualização da informação?

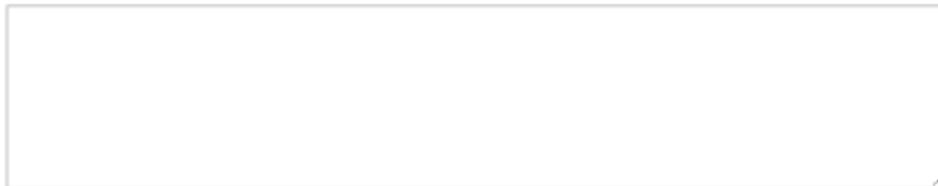
A large, empty rectangular box with a thin black border, intended for the user to provide their suggestions or answers to the question above. The box is currently blank.

Figura 30 - Questionário VI

Método de recolha e tratamento dos dados

O questionário teve natureza confidencial, 16 questionários a alunos da disciplina de VI do 4º ano do curso de Engenharia e Gestão de Sistemas de Informação da Universidade do Minho, na prática estes alunos são os que têm mais conhecimento na área, logo servirão de termos de comparação, a população alvo desta análise inclui a totalidade dos alunos presentes na sala de aula aquando da entrega dos questionários; o outro conjunto de questionários foi enviado por email para toda a comunidade estudante, onde o feedback foi claramente positivo, foram contabilizados 168 questionários iniciados mas apenas 95 totalmente respondidos, o que faz uma taxa de conclusão de 57%, podendo concluir que se obteve uma amostra aceitável representativa daquilo que pensa a comunidade, assim como algumas sugestões.

Perfil dos Inqueridos

As maiorias dos inqueridos que preencherem os questionários caracterizam-se por um tipo de população muito heterogénea, relacionada com diferentes faixas etárias que constituem a diferente comunidade estudante. A sua maioria são jovens com idade igual ou superior a 18 anos, e também alguns professores devido ao acesso ao email institucional

Validação

A validação é um processo de grande importância na construção de um questionário, mas devido ao facto de este ser um procedimento de grande complexidade e duração, requeria uma grande percentagem do tempo existente para a realização da dissertação. Sendo assim, não foi realizada uma validação, o que não interfere nas conclusões e resultados tirados deste inquérito, sendo interessante num futuro realizar a mesma.

Amostra e população

Para calcular o tamanho da amostra que é suposto considerar, de modo a esta ser representativa da População foi utilizada a fórmula:

$$n = \frac{N \cdot Z^2 \cdot p \cdot (1 - p)}{Z^2 \cdot p \cdot (1 - p) + e^2 \cdot (N - 1)}$$

Onde:

n - amostra calculada;

N - população;

Z - variável normal padronizada associada ao nível de confiança;

p - verdadeira probabilidade do evento;

e - erro amostral.

Proporção: É a proporção que esperamos encontrar. Como não tenho nenhuma informação ou perspectiva sobre o valor que quero encontrar, usarei $p=50\%$. Caso tivesse, usaria o valor aproximado que esperava (ajustando para 50% por via das dúvidas).

Erro amostral: é a diferença entre o valor estimado pela pesquisa e o verdadeiro valor. Frequentemente o valor definido é 5%, mas serão estudados vários casos.

Nível de confiança: é a probabilidade de que o erro amostral efetivo seja menor do que o erro amostral admitido pela pesquisa. Definiu-se um erro amostral de 5%, o nível de confiança indica a probabilidade de que o erro cometido pela pesquisa não exceda 5%.

População: é o número de elementos existentes no universo da pesquisa.

Depois de concluída a fase de inquérito, a amostra obtida foi de 170. Perante essa amostra, foi necessário calcular as várias margens de erro dos diferentes níveis de confiança (Tabela 4). Para isso, foi utilizada a fórmula 1.

TABELA COM OS VÁRIOS CASOS CONSIDERADOS

Tabela 6 - Casos Considerados

Margem de erro	Nível de Confiança	Tamanho da população	Proporção	Dimensão recomendada da amostra
6.28%	90%	18000	50	170
7.5%	95%	18000	50	170
7.85%	96%	18000	50	170
8.3%	97%	18000	50	170
8.9%	98%	18000	50	170
9.86%	99%	18000	50	170

Como seria de esperar, a variação de margem de erro é diretamente proporcional ao nível de confiança, ou seja, quanto maior o nível de confiança que quisermos para a amostra, maior é a margem de erro.

Resultados obtidos

- Classifique as seguintes características quanto à importância.

Selecionar o que visualizar:

Tabela 7 - Indicador número 1

Grau de Importância (1-5)	Total de respostas	Percentual
1 – Pouco Importante		
5 – Muito Importante		
1	2	2%
2	0	0%
3	23	19%
4	53	45%
5	40	34%
Total	118	100%

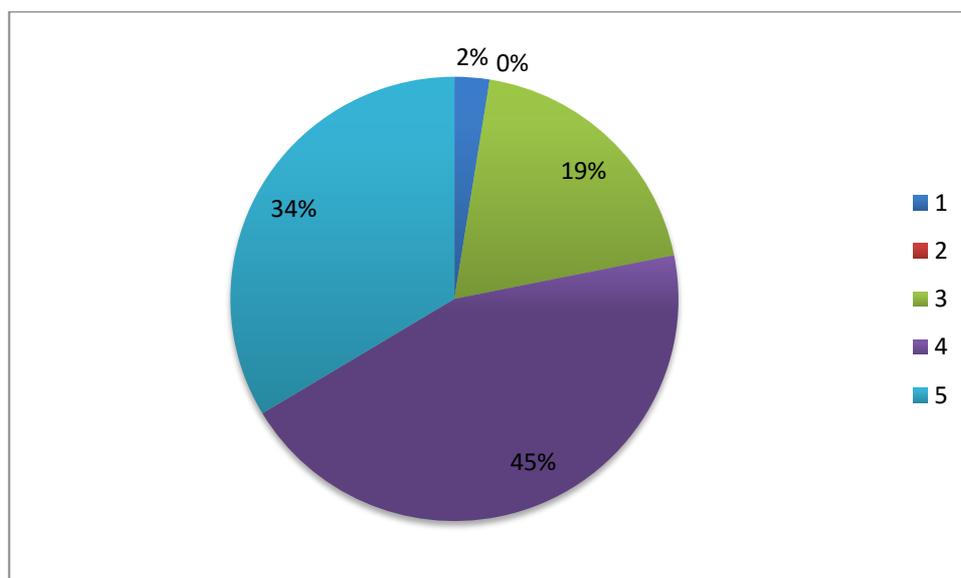


Figura 31 – Indicador número 1

Simplicidade:

Tabela 8 - Indicador número 2

Grau de Importância (1-5)	Total de respostas	Percentual
1 – Pouco Importante		
5 – Muito Importante		
1	0	0%
2	1	1%
3	20	17%
4	41	35%
5	56	47%
Total	118	100%

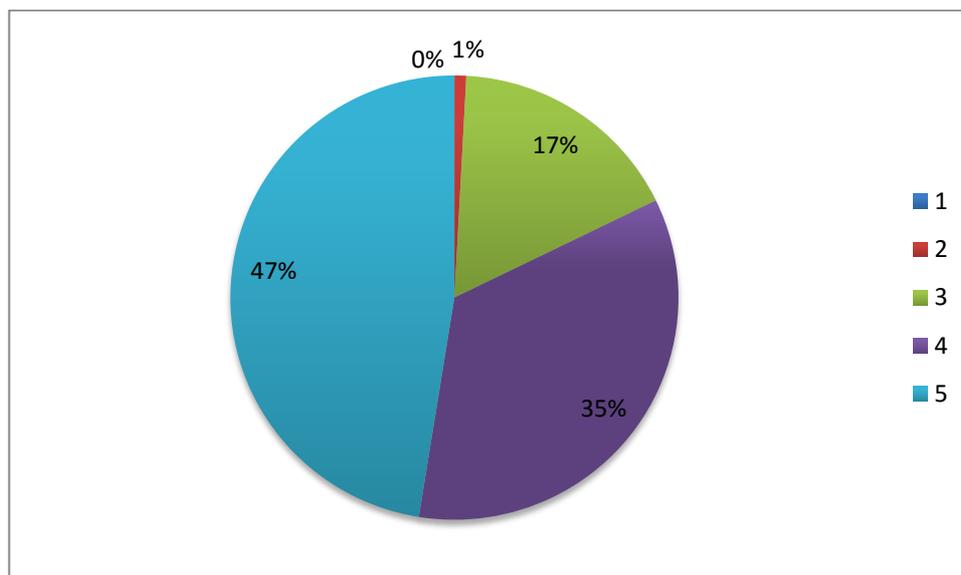


Figura 32 - Indicador número 2

Interatividade:

Tabela 9 - Indicador número 3

Grau de Importância (1-5)	Total de respostas	Percentual
1 – Pouco Importante		
5 – Muito Importante		
1	0	0%
2	4	3%
3	21	18%
4	57	48%
5	36	31%
Total	118	100%

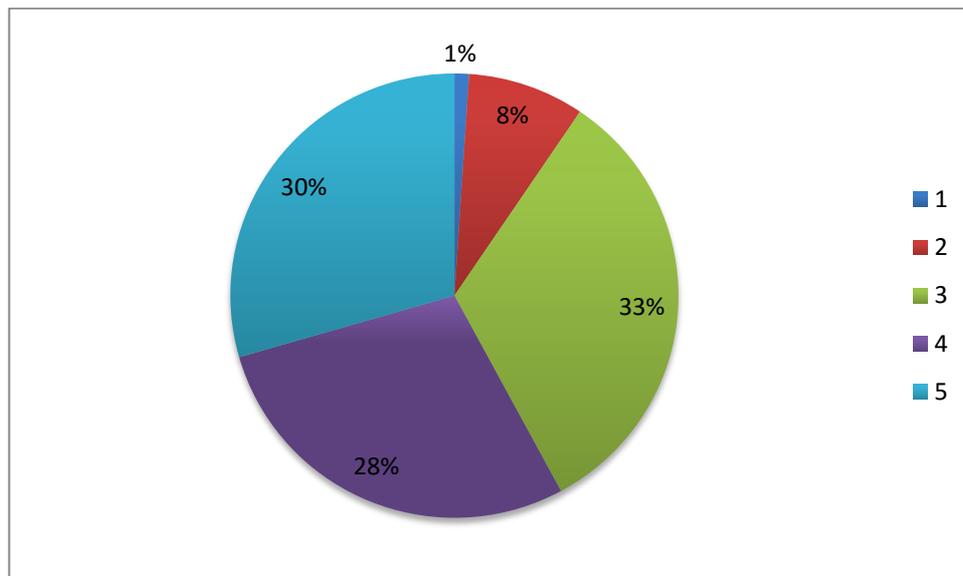


Figura 33 - Indicador número 3

Atualizado:

Tabela 10 - Indicador número 4

Grau de Importância (1-5)	Total de respostas	Percentual
1 – Pouco Importante		
5 – Muito Importante		
1	0	0%
2	0	0%
3	3	3%
4	33	28%
5	82	69%
Total	118	100%

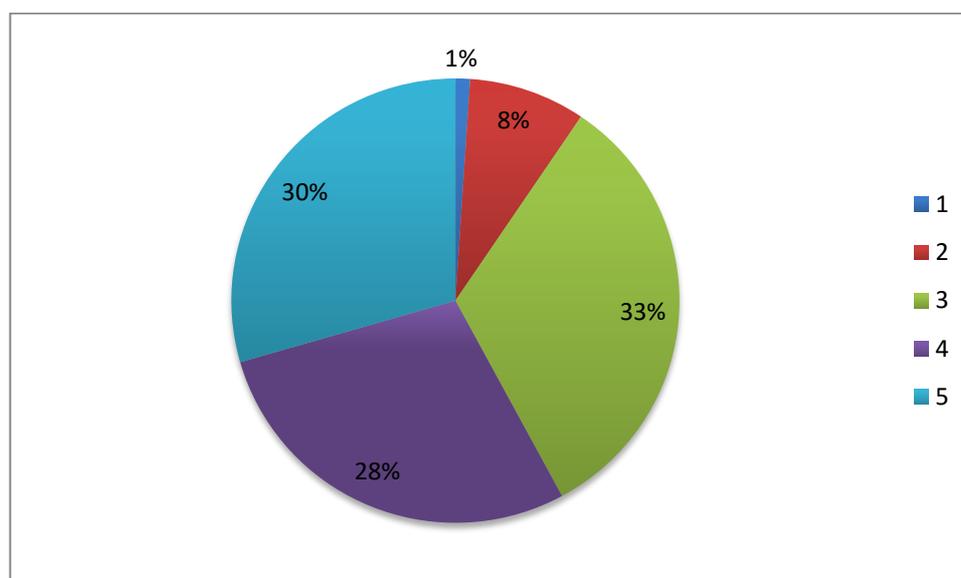


Figura 34 - Indicador número 4

Acesso e utilização simples:

Tabela 11 - Indicador número 5

Grau de Importância (1-5)	Total de respostas	Percentual
1 – Pouco Importante		
5 – Muito Importante		
1	0	0%
2	0	0%
3	5	4%
4	35	30%
5	78	66%
Total	118	100%

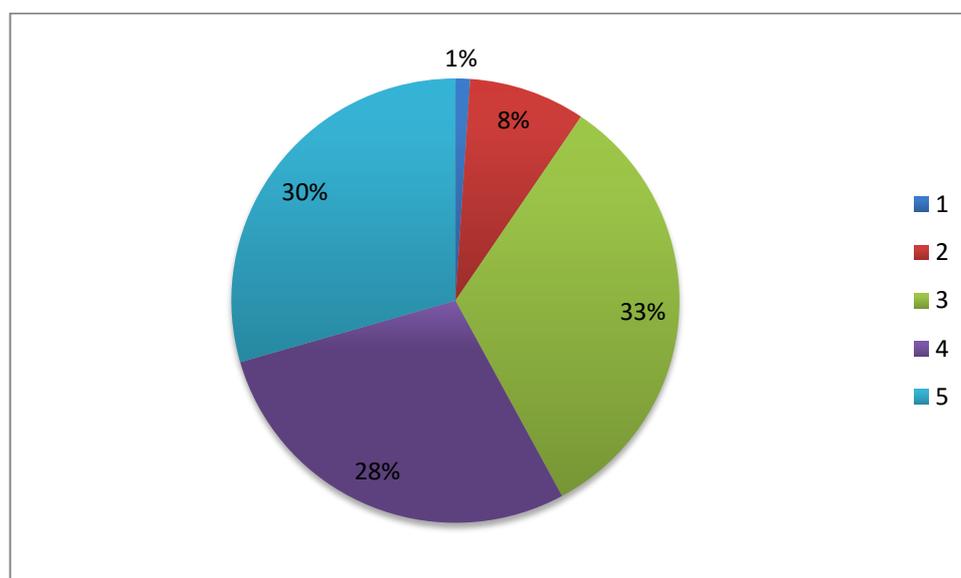


Figura 35 - Indicador número 5

- Existem também um conjunto de erros comuns que devem ser evitados, classifique quanto ao impacto nos modelos de Visualização de Informação:

Muita complexidade:

Tabela 12 - Indicador número 6

Grau de Importância (1-5)	Total de respostas	Percentual
1 – Pouco Importante		
5 – Muito Importante		
1	0	0%
2	2	2%
3	16	17%
4	40	42%
5	37	39%
Total	95	100%

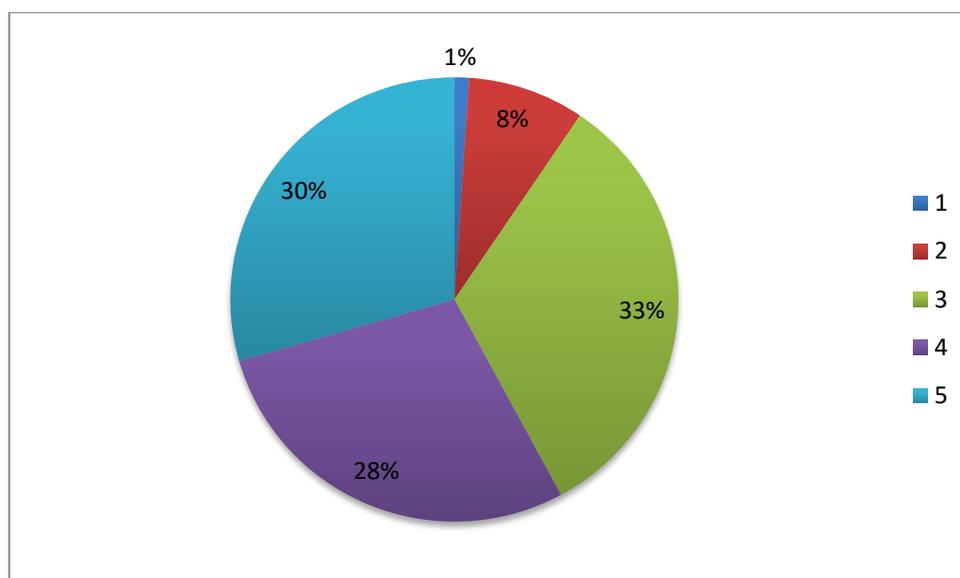


Figura 36 - Indicador número 6

Má organização dos dados:

Tabela 13 - Indicador número 7

Grau de Importância (1-5)	Total de respostas	Percentual
1 – Pouco Importante		
5 – Muito Importante		
1	0	0%
2	1	1%
3	4	4%
4	40	42%
5	50	53%
Total	95	100%

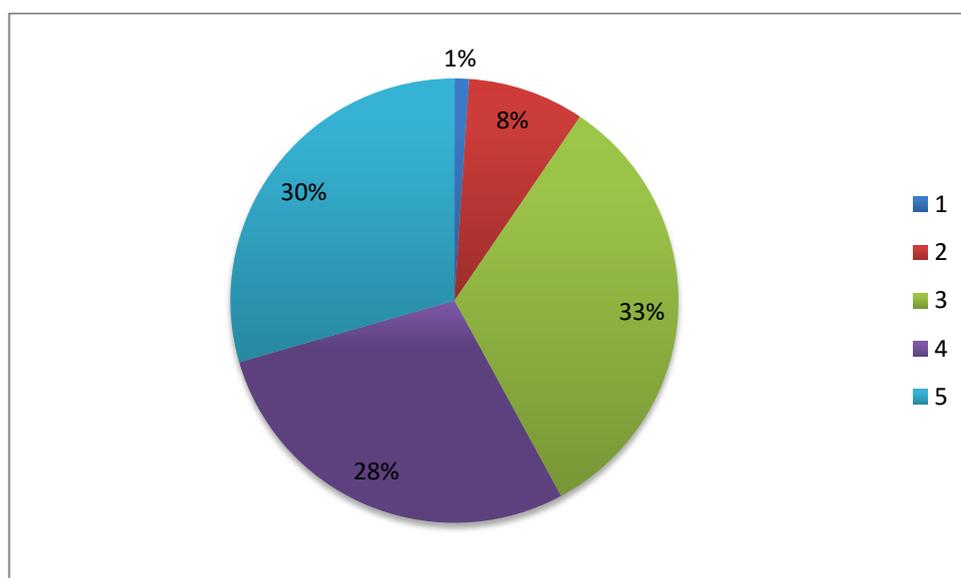


Figura 37 - Indicador número 7

Medidas e indicadores incompreendidos:

Tabela 14 - Indicador número 8

Grau de Importância (1-5)	Total de respostas	Percentual
1 – Pouco Importante		
5 – Muito Importante		
1	0	0%
2	1	1%
3	16	17%
4	39	41%
5	39	41%
Total	95	100%

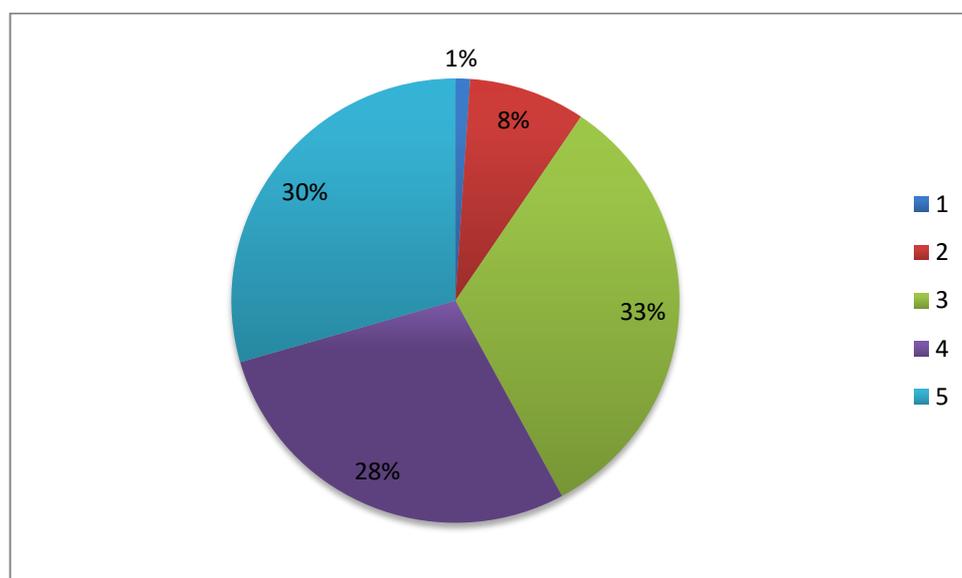


Figura 38 - Indicador número 8

Subestimar o tempo ou recursos necessários para criar e manter os métodos de VI:

Tabela 15 - Indicador número 9

Grau de Importância (1-5)	Total de respostas	Percentual
1 – Pouco Importante		
5 – Muito Importante		
1	0	0%
2	6	6%
3	39	41%
4	35	37%
5	15	16%
Total	95	100%

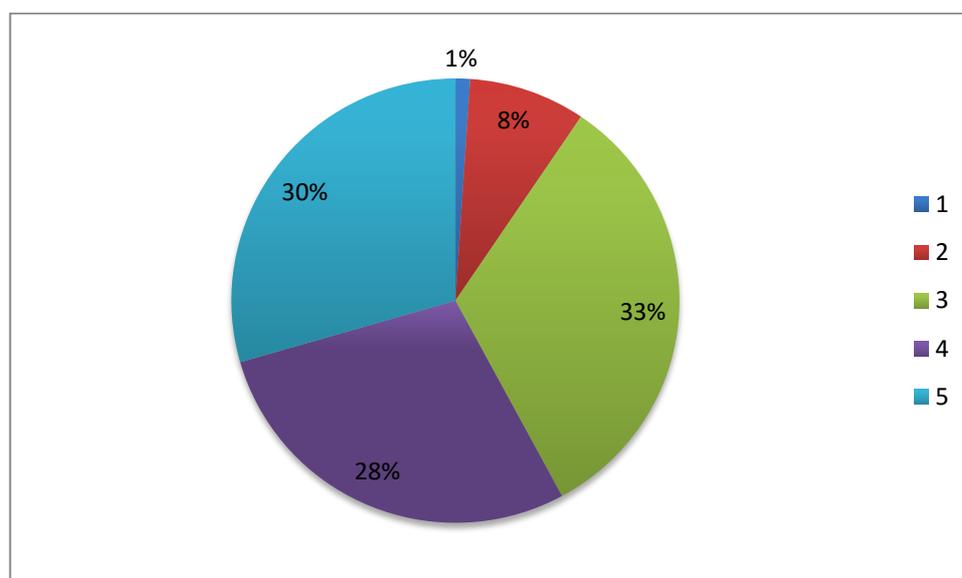


Figura 39 - Indicador número 9

Incoerências na definição das medidas, indicadores e objetivos:

Tabela 16 - Indicador número 10

Grau de Importância (1-5)	Total de respostas	Percentual
1 – Pouco Importante		
5 – Muito Importante		
1	0	0%
2	3	3%
3	29	31%
4	32	34%
5	31	33%
Total	95	100%

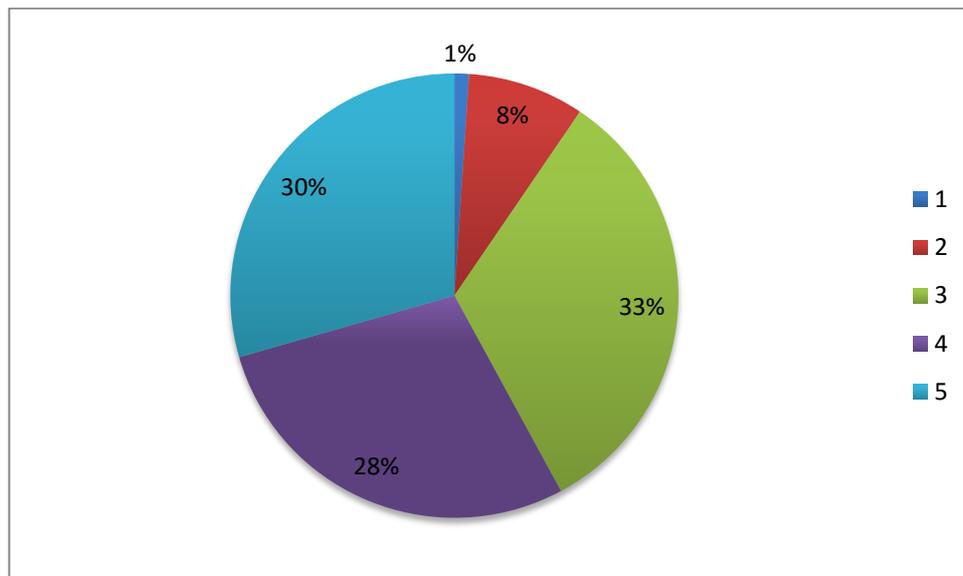


Figura 40 - Indicador número 10

Usar gráficos e diagramas mal desenhados e ineficazes:

Tabela 17 - Indicador número 11

Grau de Importância (1-5)	Total de respostas	Percentual
1 – Pouco Importante		
5 – Muito Importante		
1	0	0%
2	1	1%
3	18	19%
4	40	42%
5	36	38%
Total	95	100%

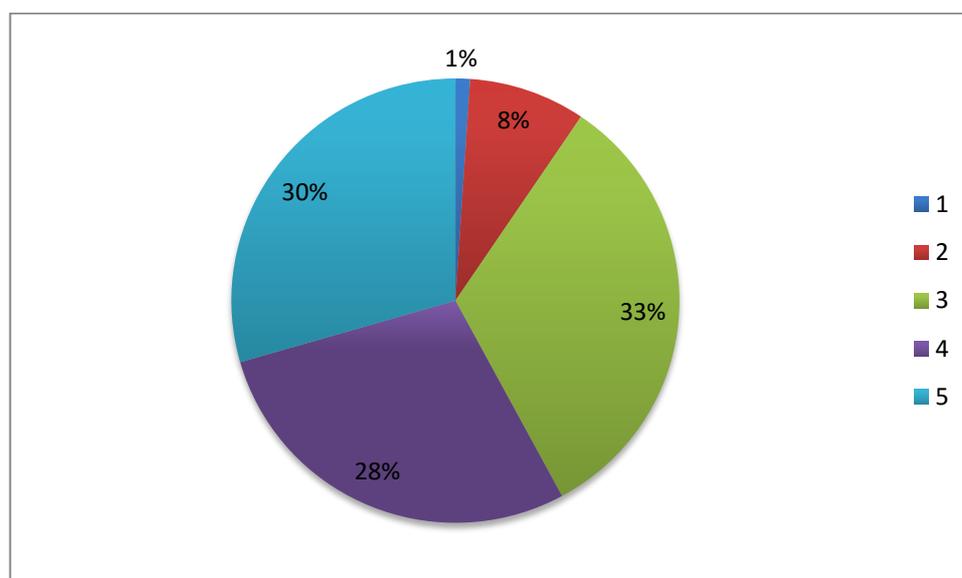


Figura 41 - Indicador número 11

Não destacar o que é importante de analisar:

Tabela 18 - Indicador número 12

Grau de Importância (1-5)	Total de respostas	Percentual
1 – Pouco Importante		
5 – Muito Importante		
1	0	0%
2	4	4%
3	19	20%
4	48	51%
5	24	25%
Total	95	100%

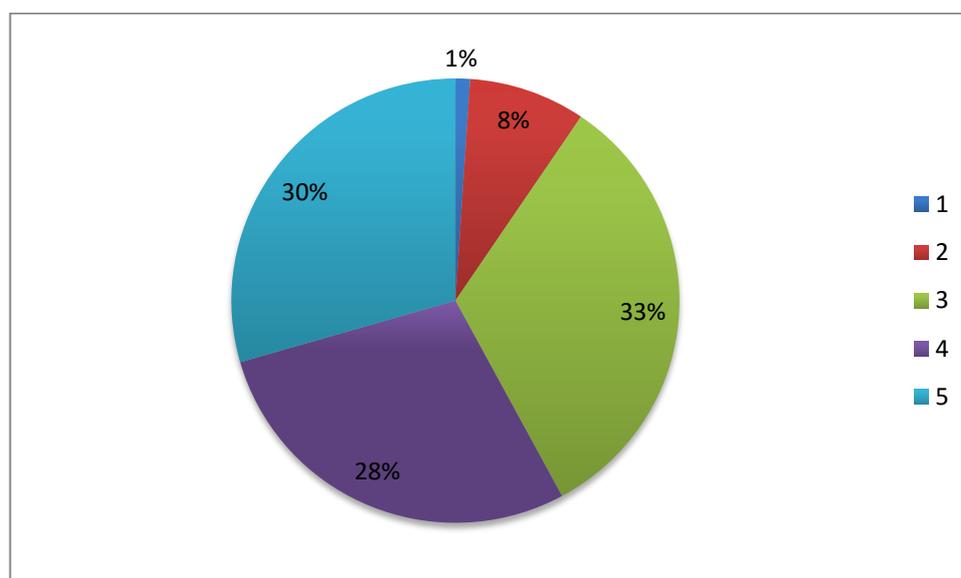


Figura 42 - Indicador número 12

Mau uso das cores:

Tabela 19 - Indicador número 13

Grau de Importância (1-5)	Total de respostas	Percentual
1 – Pouco Importante		
5 – Muito Importante		
1	3	3%
2	15	16%
3	33	35%
4	31	33%
5	13	14%
Total	95	100%

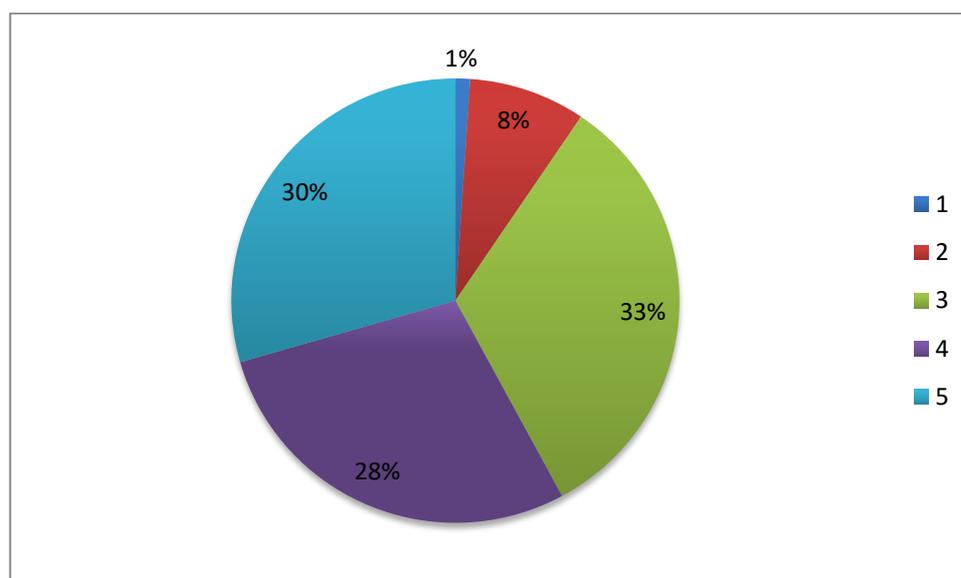


Figura 43 - Indicador número 13

“Chart Junk” (elementos que não são fundamentais à compreensão da informação):

Tabela 20 - Indicador número 14

Grau de Importância (1-5)	Total de respostas	Percentual
1 – Pouco Importante		
5 – Muito Importante		
1	1	1%
2	8	8%
3	31	33%
4	27	28%
5	28	29%
Total	95	100%

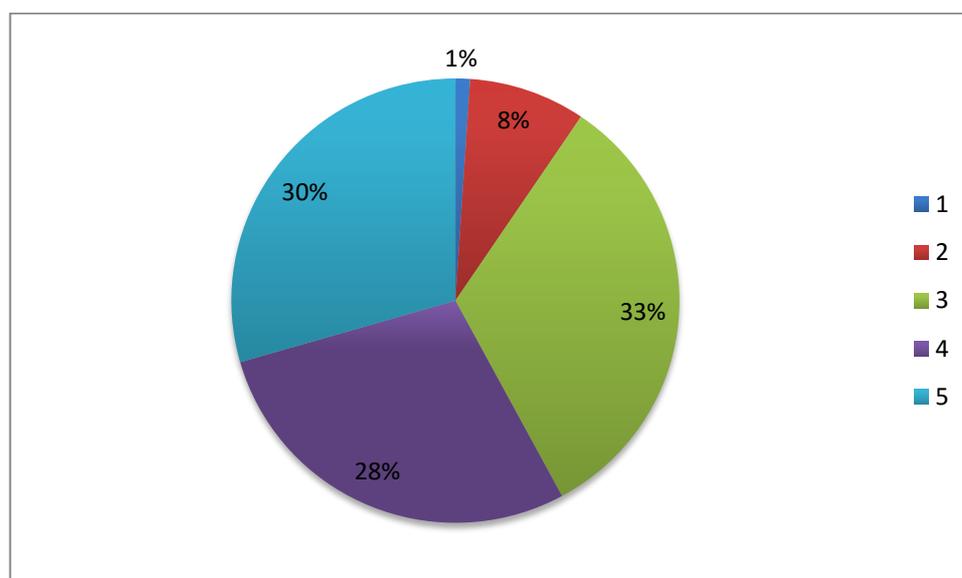


Figura 44 - Indicador número 14

Conclusões

Com o intuito de percebermos a opinião da nossa comunidade escolar, desenvolvemos um questionário acerca da importância de variáveis relacionadas com Visualização da Informação. Restringimo-nos, contudo, a alunos de uma disciplina de VI e a todo o universo estudantil da Universidade do Minho. O feedback foi claramente positivo, pelo que obtivemos uma excelente amostra representativa daquilo que pensa a comunidade.

Tendo em conta não só os questionários como também as nossas interações e discussões acerca do tópico, concluí que:

- Quanto às características de um modelo de VI:

- A maior parte dos inqueridos acha que selecionar o que visualizar é Importante ou Muito importante;
- A maior parte dos inqueridos acha que simplicidade é Importante ou Muito importante;
- A maior parte dos inqueridos acha que interatividade é Importante ou Muito importante;
- A maior parte dos inqueridos acha que estar atualizado é Muito importante;
- A maior parte dos inqueridos acha que acesso e utilização simples é Muito importante;

- Quanto ao impacto de um conjunto de erros que normalmente existem no desenvolvimento de métodos de VI:

- A maior parte dos inqueridos acha que Muita Complexidade é um erro importante ou muito importante de ser evitado;
- A maior parte dos inqueridos acha que má organização dos dados é um erro importante ou muito importante de ser evitado;
- A maior parte dos inqueridos acha que medidas e indicadores incompreendidos é um erro importante ou muito importante de ser evitado;
- A maior parte dos inqueridos acha que Subestimar o tempo ou recursos necessários para criar e manter os métodos de visualização da informação é um erro de média importância ou importante de ser evitado;
- A maior parte dos inqueridos acha que Subestimar o tempo ou recursos necessários para criar e manter os métodos de visualização da informação é um erro de média importância ou importante de ser evitado;
- Quanto Às Incoerências na definição das medidas, indicadores e objetivos os inqueridos acham em percentagens idênticas que é um erro de média importância, importante ou muito importante a ser evitado;

- A maior parte dos inqueridos acha que usar gráficos e diagramas mal desenhados e ineficazes é um erro importante ou muito importante de ser evitado;
- A maior parte dos inqueridos acha que não destacar o que é importante de analisar é um erro importante a ser evitado;
- A maior parte dos inqueridos acha que mau uso das cores é um erro de média importância ou importante a ser evitado;
- Quanto ao “Chart Junk” os inqueridos acham em percentagens idênticas que é um erro de média importância, importante ou muito importante a ser evitado;

As conclusões apresentadas, que nos permitiram compreender as posições que a nossa comunidade escolar tem relação a estas variáveis, foram retiradas com base numa análise dos dados recolhidos nos questionários. Deste modo, ao elaborarmos tabelas e gráficos representativos de cada pergunta/resposta, foi possível determinar com alguma precisão aquilo que, em geral, a comunidade pensa sobre a importância das diferentes variáveis, chegando à conclusão que todas elas têm relativamente grande impacto no desenvolvimento dos métodos de VI.

5 FRAMEWORK

Este trabalho pretende estabelecer um conjunto de regras (devidamente validadas) a serem utilizadas na definição de sistemas de VI. Pretende-se mostrar que a aplicação destas regras, resultará numa contribuição para a sistematização da criação de métodos de análise de dados mais eficazes para a gestão.

Com base na revisão de literatura efetuada, chegou a necessidade de desenvolver um *Framework* que descrevesse e sintetizasse o contributo da matéria recolhida. Este *Framework* serve de suporte, como ferramenta conceptual, na orientação dos decisores e no desenvolvimento dos modelos de VI.

A VI, como ferramenta de análise de dados, fornece vários outputs aos utilizadores, por exemplo: *Dashboards*, *Reports* e Infografias, estes são alguns dos principais outputs abordados neste trabalho, podendo posteriormente adaptar-se a diferentes métodos de análise e visualização. Estes outputs servem de apoio à tomada de decisão por parte dos analistas/utilizadores, permitindo aos mesmos atingir os seus objetivos com excelência.

Existem vários fatores que são decisivos, determinando a forma e o tipo de informação que será apresentada aos analistas/utilizadores:

Uma das disciplinas estudadas de onde foi retirada informação importante e necessária para o desenvolvimento deste *Framework* foi a HCI (*Human Computer Interaction*). HCI pode ser definida como uma disciplina voltada para o estudo, conceção, construção e implementação de sistemas de computação interativos. Uma interface de utilizador, assim como uma interface gráfica, são os meios como um ser humano interage com uma máquina, a HCI vai mais além e estuda o raciocínio por detrás da construção das interfaces integrando conceitos e métodos de ciência da computação, *design* e psicologia para construir interfaces que são acessíveis, fáceis de usar e eficientes. Todos esses conceitos serão direcionados e adaptados para utilizar na conceção de métodos de VI.

Tanto visualização, como análise visual, dependem fortemente do conhecimento transmitido pelas teorias estudadas em HCI, então, a qualidade da informação derivada da visualização e análise visual, depende em grande parte das capacidades dos utilizadores nesse processo, assim como o tipo de utilizador e as tarefas são igualmente importantes para os métodos e algoritmos utilizados nos processos de análise e visualização. Foley (2011) diz mesmo que a “análise visual” é um domínio especializado da HCI.

A chave para relacionar HCI com VI é a interação e a experiência do utilizador, ou seja, é necessário traduzir um grande volume de informação em conjuntos de dados e visualizações interativas que

atendam às necessidades específicas do utilizador ou seja, pretende-se uma interação simples e intuitiva, de design simples, que seja compreensível para os utilizadores peritos e não peritos na área. Migrar estas teorias para visualização é, portanto, a conceção de um sistema que seja suficientemente simples e intuitivo para não dificultar a criatividade dos utilizadores, mas também flexível o suficiente para permitir explorar plenamente os dados - talvez de forma ainda que o utilizador do sistema não tinha antecipado.

Posto isto, foram selecionados alguns princípios fundamentais de HCI que podem ser utilizados na conceção de métodos de visualização da informação:

- Perceção, em termos de perceção, cor, luminosidade, padrão, efeito e objeto:
 - Perceção: a perceção é a forma como algo pode ser visto ser visto, dependendo do utilizador. Assim, quando se trata de projetar algo, será necessário projetar de tal maneira, que qualquer utilizador, pode usar perfeitamente sem confusão. Na conceção, pode haver mais de uma pessoa envolvida com a ideia, então é importante que todos tenham o mesmo entendimento e mesma visão.
 - Cor: o uso da cor é muito importante no *design*. Por isso, é importante que os *designers* utilizam cores adequadas que não terão efeito sobre o utilizador. Isto significa que a cor é algo que o *designer* precisa levar em conta na no momento da conceção. Cores diferentes, têm efeito diferentes, em diferentes utilizadores.
 - Luminosidade: Um sistema de cores é usado para certificar de que o utilizador final tem os melhores efeitos visuais possíveis. O tipo mais comum de sistema utilizado é o chamado sistema *Trichromatic*, que usa três tipos de cores para exibir, que são: Vermelho, verde, e azul (em inglês *Red, Green and Blue* - RGB). RGB é a melhor maneira de exibir cores devido à forma como os nossos olhos funcionam, e como nossos olhos vêem estas cores na tela. Os *designers* precisam levar em conta, o brilho de telas de utilizadores, por exemplo.
 - Padrão (*Gestalt*): o padrão é essencialmente o modelo utilizado para a criação de uma interface. Permite ao utilizador navegar facilmente numa interface A lei *Gestalt* é usado muitas das vezes no desenvolvimento para descrever a perceção de padrão e formar as regras para projetos.
 - Efeito "*Pop Out*": este é um efeito que é usado para fazer com que os nossos olhos se concentrem num determinado item ou objeto. Fazem isso usando cores mais brilhantes contra as menos brilhantes, etc. Esta técnica é usada para assegurar que os utilizadores se concentram num determinado elemento e em tudo o que o rodeia, por exemplo, numa página *web* o *designer* vai querer ter certeza de que os botões e barra de navegação se destacam na sua página, então vai fazer o texto maior, usou cores diferentes, etc., para se certificar de que o utilizador vê isso facilmente.
 - Objeto: os objetos são essencialmente as partes que compõem a interface, por exemplo, os botões, menus, etc. É necessário ter em conta os tamanhos e formas dos objetos, de modo a que eles se encaixem e sejam adequados para o utilizador final.
- Modelos de comportamento através de *predictive model, throughput*.

- *predictive model*: é o processo em que se tenta prever o resultado de alguma coisa, por exemplo, na criação de um novo jogo, a partir do que está em jogo, o esquema de cores, se houver qualquer tipo de violência, etc., é possível prever o público-alvo e se vai vender bem ou não. Em relevância para HCI, quanto mais complexo o interface, menor mercado haverá para o produto.
- *throughput*: é usado para estimar quanto tempo vai demorar para transferir uma quantidade específica de dados de uma localização para outra e / ou a quantidade de trabalho que o computador faz num determinado período de tempo. Um exemplo disso seria o tempo que leva para transferir dados do seu computador para um cartão de memória USB.
- Modelos descritivos, nomeadamente KAM, ...
- KAM (*Key, Action, Model*): é um modelo que dá informação detalhada sobre cada “pedaço ” de tecnologia. Um exemplo de KAM poderia ser um teclado QWERTY normal, que tem os botões com funções únicas. É possível também usar uma combinação de botões, por exemplo, pode-se usar o botão CTRL + C para copiar algo, etc.
- *Buxton’s model*: É um modelo que é utilizado durante a criação de uma interface; decide qual o tipo de técnicas de entrada que podem ser utilizadas. É importante pensar como o utilizador pode querer interagir e ter todas as opções disponíveis para eles.
- *Guiard’s model*: É um modelo que é usado para se certificar se a interface é fácil de usar para pessoas destros ou esquerdas.
- Processamento da informação: O processamento da informação é sobre como a informação é processada ou para ser processada e quão rápido é feito.
- “*Human as a component*”: O Humano é o utilizador final e uma das variáveis-chave. *Human as a component* basicamente significa que o ser humano é o principal componente no projeto. Ao projetar a interface o designer precisa de pensar fora da caixa, pensar sobre outros utilizadores, pois vários utilizadores podem ter problemas diferentes. Por isso, é vital que seja projetado para atender a qualquer utilizador e não apenas para suas próprias necessidades. Há muita pesquisa que precisa de ser feita ao projetar qualquer tipo de modelo, por exemplo, as necessidades do utilizador, como o utilizador irá interagir com a interface e outras informações, são algumas das coisas que o designer precisa pensar.
- Processamento Humano da informação (HIP): É a forma como o ser humano processa a informação. Este princípio é utilizado para analisar como nós, os seres humanos, processamos certos tipos de informação, e como nós, em seguida, a introduzimos num modelo de alguma forma. Muitas pessoas acreditam que os seres humanos e computadores, ambos têm forma semelhante de armazenamento de dados, a longo prazo e de curto prazo.
- Computadores, por outro lado, são capazes de realizar cálculos muito rápidos do que os seres humanos, e também continuar por muito mais tempo sem falhar, então em algum aspeto os computadores são mais eficientes.
- GOMS (*Goals, Operators, Methods, Section Rules*), que inclui, goals, operators, methods e rules :
 - *Goals* – São os objetivos finais do utilizador. Ao trabalhar numa tarefa, existem metas de curto prazo e metas de longo prazo.

- *Operators* - estas são as ações que precisam de ser tomadas para que a tarefa destinada seja feita. Por exemplo, a fim de ir para outra página da web, é necessário o rato clicar, e o teclado para digitar a página web.
- *Methods* - este é o plano que vai ser utilizado, a fim de fazer alguma coisa. Por exemplo, para mover um atalho para a reciclagem, primeiro tem que se pensar em mover o rato para pegar no atalho e depois muda-lo para a reciclagem.
- *Rules* - Este é o método que se usa a fim de completar uma tarefa específica. Primeiro de tudo é necessário identificar qual é a tarefa que se quer fazer, então só depois se deve pensar sobre o qual o método mais eficiente para completar a tarefa.

Existe também um conjunto de regras chamadas - Heurísticas de usabilidade desenvolvidas por (Nielsen, 1993), que podem ser usados no *design* de interfaces de modelos de VI:

- Evitar informação irrelevante: Esforços devem ser feitos para evitar informação irrelevante. (Nielsen, 1993) diz que cada unidade extra de informação compete com unidades de informações relevantes o que diminui a sua visibilidade.
- “Falar” a linguagem do utilizador: Todas as informações devem ser expressas em conceitos que são familiares para o utilizador, em vez de familiares para o operador ou para o sistema.
- Minimizar a carga de memória ao utilizador: É importante que o utilizador não tenha que se lembrar de informações a partir de uma parte de um diálogo para outro. A ajuda deve estar disponível em pontos facilmente recuperáveis no sistema.
- Consistência: Palavras situações e ações devem sempre dizer a mesma coisa, independentemente de onde ocorrem no sistema.
- Feedback: Os utilizadores devem sempre ser informados sobre o que acontece no sistema de uma forma oportuna e relevante.
- Clarificar as saídas: Os erros muitas das vezes ocorrem na escolha de funções que não são necessárias e que não precisam de ter uma rápida saída de emergência para retornar ao estado anterior.
- Atalhos: Requerida pelo utilizador experiente (e invisível para o utilizador básico) para acelerar a interação com o sistema.
- Boas mensagens de erro: Estes precisam de ser expressos em linguagens simples que o utilizador entenda e que sejam específicas o suficiente para identificar o problema e sugerir uma solução.
- Precaver erros: Um *design* cuidadoso irá impedir que ocorram menos problemas.
- Ajuda e documentação: Os melhores sistemas podem ser utilizados sem documentação. No entanto, quando tal ajuda é necessária deve estar facilmente disponível, focada nas tarefas dos utilizadores e lista de passos específicos para soluções.

Depois da revisão de literatura e do estudo das diferentes teorias e necessidades dos utilizadores, chegaram-se aos seguintes objetivos a atingir com o apoio da VI:

- Explorativos: Através dos objetivos explorativos descobrimos a história e os elementos. Aqui, a visualização da informação ajuda a descobrir ideias, padrões e factos significativos para ajudar a formar uma opinião sobre o assunto, nesta secção enquadram-se algumas infografias assim como *reports*. Devem ser interativos, utilizar visualizações e ferramentas que facilitem a exploração da informação, logo é importante compreender as necessidades e as questões do público-alvo.
- Explicativos/analíticos: Este é provavelmente o objetivo que mais frequentemente se pensa quando se trata de visualização de informação. Explicação é sobre a compreensão, não apenas o que aconteceu, mas porquê e como isso aconteceu. Neste aspecto a visualização é importante de modo a auxiliar na procura de detalhes e pontos de prova para depois os explicar. Por exemplo, estas visualizações podem apresentar contextos de histórico, tendências ou ser incorporados em diferentes meios de comunicação de modo a incluir texto para contar uma história convincente, neste caso todos os modelos de VI vão interessar.
- Decisivos: Este objetivo é sobre o uso de visualizações de dados para dar suporte à tomada de decisão. *Scorecards* e *dashboards* são os exemplos mais comuns que ajudam no apoio à tomada de decisão, permitem aos utilizadores efetuar uma rápida visualização e triagem dos dados de modo a dar ênfase à informação que requer uma maior atenção.
- Estratégicos: Na maioria das vezes a informação é necessária para fins estratégicos. Fornecem uma visão geral e rápida da "saúde" da organização, visões gerais e comparações rápidas dos valores-alvo, bem como avaliações simples acerca do desempenho (por exemplo, bom ou mau), neste caso *Scorecards* e *reports* serão os modelos de VI adequados.
- Operacionais: Os objetivos operacionais são para dados de monitorização, principalmente em tempo real, não tendo funcionalidades estatísticas ou de análise. Servem também para fornecer alertas e apresentação de respostas adequadas a eventos de emergência. São geralmente usados a nível dos departamentos e não a nível executivo. Um pobre fluxo de informação para os empregados que são diretamente responsáveis para a eficiência e otimização de processos é o calcanhar de Aquiles de muitas organizações. Muitas das vezes os funcionários de uma organização são obrigados a adiar a execução de uma determinada operação pois são procurar informação de um determinado cliente em varias interfaces, os *Dashboards* e os *reports* podem ajudar na medida em que podem consolidar informações para relatórios e análises de fácil utilização contribuindo para a eficiência operacional.

Tipos de dados:

O estado em que os dados se encontram no momento da concepção dos modelos é um dos aspectos de maior relevância, podendo ser estruturados ou não estruturados.

- Estruturados: Quando os dados estão coerentes, sem erros e prontos a serem utilizados.
- Não estruturados: Quando é necessário efetuar transformações aos dados para que possam ser utilizados posteriormente. Por exemplo, corrigir eventuais erros ou categorizar os dados.

Tipo de utilizador

Outro aspecto de grande relevância na VI prende-se com a existência ou ausência de conhecimento das teorias da percepção visual, por exemplo: design, contrastes, alinhamentos, etc. Utilizadores com estes conhecimentos são mais exigentes com os modelos de VI apresentados, sendo este um aspeto importante no desenvolvimento dos mesmos.

Devido aos aspetos apresentados anteriormente, existem dois tipos de utilizadores para os modelos de visualização de informação: básicos e avançados.

- Os utilizadores básicos não possuem conhecimentos das teorias da percepção visual nem necessidades profundas de análise de dados, apenas necessitam da visualização rápida e não muito detalhada da informação. Assim sendo, a estes utilizadores apenas interessa apresentar a informação em grande volume, sem grande nível de detalhe, mas onde os mesmos possam ter uma visão que abranja todas as áreas de análise.
- Quanto aos utilizadores avançados (gestores, analistas, etc) interessa a apresentação da informação de forma detalhada e trabalhada, pois necessitam desta para suportar da melhor forma as suas decisões. Sendo estes, teoricamente, utilizadores conhecedores das diferentes teorias da percepção visual e design, torna-se assim relevante a apresentação da informação nas diferentes formas e com diferentes níveis de detalhe. Todos os aspetos interessam, desde as teorias da percepção (que vão conjugar com as cores, contrastes, etc) até às características, boas praticas e erros a evitar na concepção de métodos de VI.

Para melhor descrever esta comparação entre utilizador básico e avançado, podemos usar uma analogia – um cliente irá escolher jantar num restaurante com requinte ou um restaurante pequeno e familiar. Quem frequenta restaurantes requintados (estes seriam os utilizadores avançados) pretende todos os serviços e todos os detalhes possíveis, por exemplo, o nível da luz, a atenção do servente, a qualidade da comida, a qualidade da carta de vinhos, a recomendação do melhor vinho para acompanhar os vários pratos, a música ambiente em termos de qualidade e nível, etc. Quem frequenta restaurantes pequenos e familiares (utilizadores básicos), procura comer e beber bem e eventualmente pagar pouco, mas não se importa com o ruído existente, se é música ou televisão ou se na mesa ao lado estão a falar alto, tal como não se importa se os pratos são colocados em cima de toalhas de papel descartáveis, os guardanapos são de papel, se os talheres estão colocados corretamente, etc.

Regras a seguir e erros a evitar no desenvolvimento de métodos de VI:

Associados a estes fatores estão ainda um conjunto de regras e de erros a ser evitados que são de extrema importância no desenvolvimento dos modelos de VI. Estas regras e erros foram recolhidos na revisão bibliográfica realizada tendo sido posteriormente avaliada a sua importância e impacto através da realização de um questionário feito á comunidade estudante.

As Características e boas práticas que devem ser seguidas na conceção dos métodos de VI são:

- Selecionar o que visualizar - selecionar quais as medidas, indicadores devem ser incluídos nos modelos. Deve-se ser altamente seletivo para determinar o que é importante ser visualizado.
- Optar pela simplicidade - os modelos de VI são concebidos com o intuito de serem fáceis e rápidos de interpretar, um exemplo de uma má implementação deste ponto são os relatórios e tabelas baseadas sobretudo em texto.
- Ser interativo - um modelo deve ser concebido para ser interativo, o que permitirá que diferentes utilizadores com diferentes objetivos de análise possam, na mesma página, através de uma personalização obter a informação de que necessitam.
- Estar atualizado - uma das principais preocupações na conceção dos modelos de VI passa por apresentar apenas informação útil e atualizada de modo a refletir os desafios atuais dos negócios.
- Tornar o acesso e utilização simples - tornar os modelos acessíveis é fundamental, essencialmente com distribuição web pois deste modo podem extrair constantemente informação atual e seguir os protocolos IT e padrões de segurança.

Os Erros comuns que devem ser evitados são:

- Começar com muita complexidade;
- Organizar mal os dados;
- Usar medidas e indicadores que não estão devidamente compreendidas;
- Subestimar o tempo ou recursos necessários para criar e manter os modelos de VI;
- Incoerências na definição das medidas, indicadores e objetivos;
- Usar gráficos e diagramas mal desenhados e ineficazes;
- Não destacar o que é importante de analisar;
- Mau uso das cores; Evitar o “*Chart Junk*”.

Foi identificada uma necessidade de efetuar uma segunda iteração da revisão bibliográfica devido ao facto de terem surgido algumas sugestões nos questionários que suscitaram uma nova pesquisa onde foi identificado outro conjunto de regras que todos os modelos de VI devem seguir, seguidos de um conjunto particular de regras que cada modelo de VI deve seguir:

- Começar com a questão - sem saber o que se procura, nenhum tipo dados ou gráficos chamativos vão ajudar. Começar com uma pergunta e construir o projeto em torno disso.
- Pensar nas cores: aplica-se a qualquer tipo de *design*, existem cores que funcionam bem juntas, outras nem tanto, logo é fundamental pensar cuidadosamente no esquema de cores.
- Usar filtros: as visualizações são muitas das vezes grandes e complexas para os utilizadores identificarem claramente alguns padrões, logo é importante adicionar filtros de modo a que os utilizadores possam decidir que tipos de informação se querem centrar.
- Escolha adequada da ferramenta e do tipo de gráfico: a escolha da ferramenta é um dos pontos mais importantes, é fundamental usar uma ferramenta com que se esteja familiarizado e á vontade a trabalhar, podendo então adequar os tipos de gráficos existentes á informação.
- Verificar os dados: é importante verificar a consistência dos dados, certificarmo-nos que fazem sentido. Começar com gráficos simples para verificar se existem valores atípicos e daí em diante aumentar a complexidade.
- Explicar as codificações: devem-se explicar as codificações existentes para guiar e não assumir que o analista sabe o significado. Para isso podem-se utilizar, por exemplo, escalas de cores, rótulos, usar um parágrafo para descrever, etc.
- Considerar o público-alvo: devemos ter em conta o que são e para quem são os modelos de visualização desenvolvidos. De nada vale apresentarmos um gráfico com muitos detalhes para um conjunto de pessoas que nada percebem de VI.

Apresentam-se agora as regras no desenvolvimento das diferentes técnicas de VI que respondem aos objetivos a atingir com o apoio da VI:

1. *Dashboards / Reports*

- a. Manter o *dashboard* em linha com os objetivos estratégicos;
- b. Equilibrar cuidadosamente as diferentes categorias de indicadores nos *dashboards*;
- c. Usar *dashboards* inteligentes e dinâmicos, simples e de fácil utilização;
- d. Verificar sempre os resultados, comparar com dados históricos e com os objetivos a atingir;
- e. Certificar que a mesma pessoa é responsável pela edição, preenchimento e a partilha do *dashboard*;
- f. Partilhar o *dashboard* numa ferramenta multiplataforma para que seja possível que vários utilizadores acedam á informação;
- g. Selecionar o tipo certo de gráficos;
- h. Agrupar a informação de forma lógica
- i. Tornar os dados relevantes para a audiência;
- j. Apresentar apenas as métricas mais importantes;

2. Scorecards

- a. Determinar os objetivos do projeto;
- b. Desenvolver um plano estratégico: desenvolver a visão e a missão, analise SWOT e determinar a estratégia final;
- c. Identificar os indicadores de desempenho:
- d. Preparar o mapa estratégico: Descreve a relação entres os objetivos, desde o ponto de vista do cliente ao ponto de vista financeiro.
- e. Identificar os objetivos do público-alvo:
- f. Preparar um manual dos indicadores de performance: o manual contém informações detalhadas sobre os indicadores de desempenho, oferecendo informações aos responsáveis pela avaliação dos indicadores.
- g. Preparar os *Balanced Scorecards*: O processo de preparação dos *balanced scorecards* envolve o preenchimento de dados que foram recolhidos e alocados ao longo do projeto, a fim de demonstrar a interdependência de tudo o que foi preparado durante todo o projeto.
- h. Implementar os *Balanced Scorecards*.

3. Infografias

- a. Recolha de dados: tudo começa com a recolha de dados. Normalmente os dados surgem desorganizados e com erros, então é necessário trabalhar os dados de modo a corrigir essas incoerências erros.
- b. Ler toda a informação: normalmente, quem visualiza a informação tem tendência a ler apenas a informação destacada e não dar importância ao resto, então o desafio passa por conceber infográficas hábeis que façam com que nenhuma informação seja perdida.
- c. Encontrar a narrativa: quando a informação a apresentar não é a mais interessante, então a infografia desenvolvida pode-se tornar desinteressante também, então, o objetivo passa por clarificar e tornar interessante um conjunto de dados e encontrar uma narrativa que ajude nesse processo.
- d. Conceber bons gráficos: o design deve ajudar á compreensão e cativar quem observa;
- e. Identificar o objetivo: qual o objetivo a atingir com a infografia. Assim que identificamos o objetivo podemos decidir qual o tipo de informação e formato que funciona melhor.
- f. Identificar o público-alvo: importante para saber que tipo de informação apresentar, saber o que lhes interessa e que tipo de linguagem usar.
- g. Ser relevante: apresentar algo novo, curto e simples.
- h. Realizar uma boa pesquisa: recolher informação relevante e de diferentes fontes e perspectivas.
- i. Usar um bom título: um título cativante ajuda a cativar a atenção para o resto do trabalho.

5.1. Framework conceptual Visualização da Informação

Nesta secção é apresentado o esquema do Framework referente às descrições anteriores:

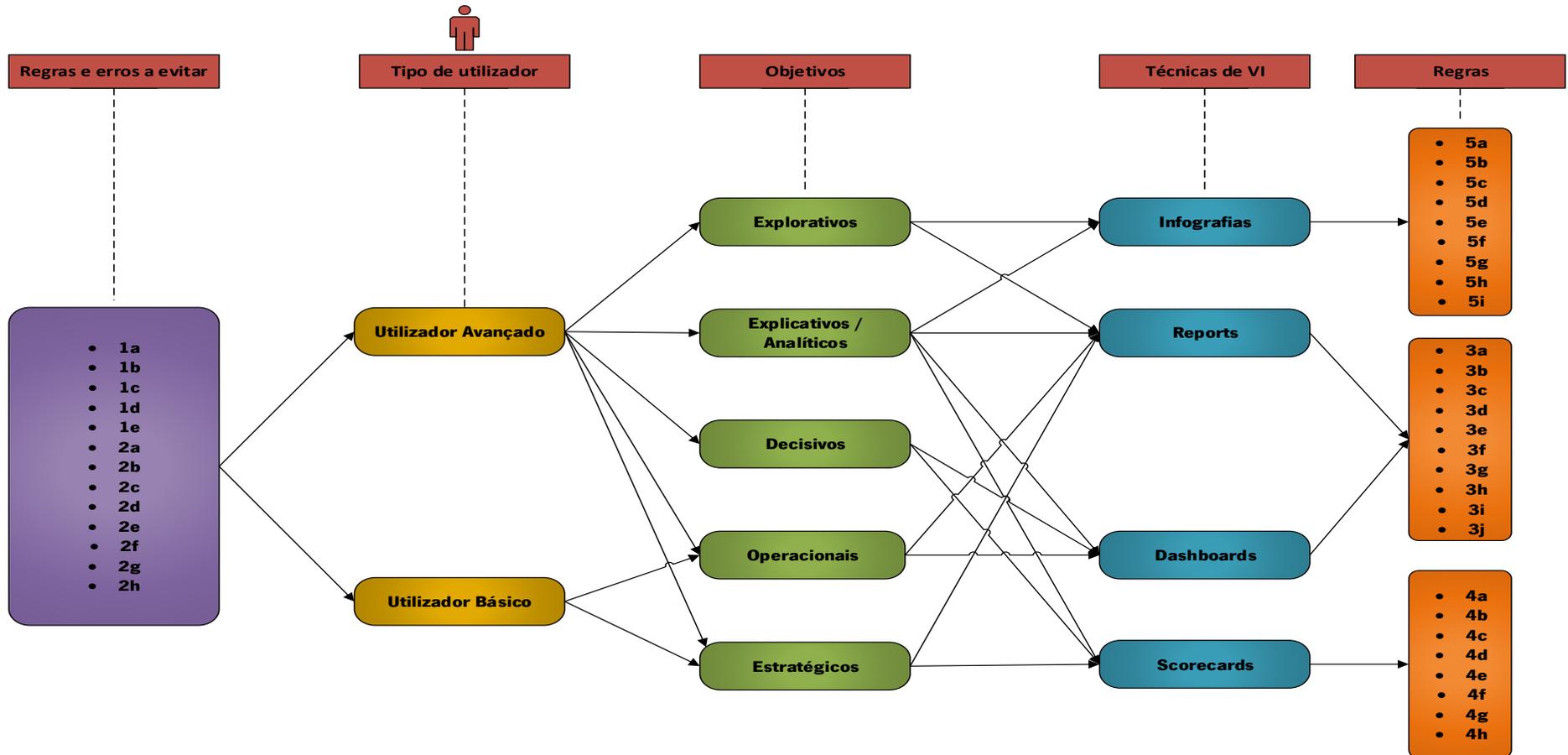


Figura 45 - Framework

O objetivo do *framework* passa apenas por servir de apoio a um correto desenvolvimento de modelos de VI, não tendo como intenção dizer qual a forma correta de desenvolver os mesmos, pois a forma como se visualiza a informação não é linear, dois analistas podem retirar conclusões diferentes da mesma visualização, com o mesmo tipo de informação podem-se desenvolver Visualizações completamente diferentes, mas, é possível facultar uma ajuda com procedimentos básicos á tomada de decisão e diminuir o mau desenvolvimento através das normas que aqui foram descritas.

O *framework* deve ser visto como uma espécie de fluxograma, inicialmente deve-se proceder a leitura e compreensão das regras a seguir e erros a evitar de modo a que os mesmos estejam sempre subentendidos durante todo o processo; de seguida devemos identificar qual o tipo de utilizador e os objetivos associados ao mesmo; relacionados com os objetivos estão relacionadas determinadas técnicas de VI que serão mais adequadas para responder aos diferentes objetivos; identificados os objetivos e as técnicas de VI será necessário então proceder a uma leitura e compreensão das regras específicas que o desenvolvimento desses modelos de VI deve seguir.

6 CONCLUSÕES

6.1. Síntese do trabalho desenvolvido

Este trabalho inicialmente foi concebido com dois objetivos: explicar importância crescente da VI nas organizações e conceber um *framework* que definisse um conjunto de normas para o correto desenvolvimento de modelos de VI.

Usar a VI para exibir grandes volumes de informação não é nada de novo, no entanto, o valor da VI na tomada de decisão é muitas vezes esquecido ou mal implementado. Então, o objetivo deste trabalho passou por descrever a importância da VI e como a mesma deve ser objetiva e coerente. Os pontos-chave para garantir o sucesso do uso da VI passam por garantir que os melhores e mais adequados tipos de visualizações são usados, os dados são sempre expostos em perspectiva com a informação contextual permitindo que a informação seja compreendida universalmente e principalmente que os dados a serem visualizados permitam aos analistas tomar as melhores decisões baseando-se no que veem.

Por outro lado, no que toca ao desenvolvimento dos modelos de VI, é cada vez mais uma grande necessidade estabelecer um conjunto de normas que permita que o erro no desenvolvimento seja cada vez menor.

Então, com base na investigação de teorias relacionadas com visualização, percepção e Design, chegou-se à conclusão que a forma como se visualiza a informação não é linear, dois analistas podem retirar conclusões diferentes do mesmo gráfico, com o mesmo tipo de informação podem-se desenvolver Visualizações completamente diferentes, mas, é possível facultar uma ajuda com procedimentos básicos à tomada de decisão e diminuir o mau desenvolvimento através das normas que aqui foram descritas.

No entanto, o desenvolvimento de métodos de VI padece ainda de estudos empíricos que ajudem a cimentar o trabalho nesta área, sendo assim, a dissertação em desenvolvimento está preparada para continuar com a fase de Desenvolvimento (DSR), para que, futuramente possa ter uma contribuição de relevância e ser uma referência nesta área.

6.2. Contribuição da dissertação

As principais contribuições deste trabalho estão incluídas na Framework em desenvolvimento e da investigação de uma área ainda não muito explorada, das quais se podem salientar:

- Framework adaptado á modelação de diferentes modelos de VI: para um contexto específico, existe um tipo de forma de Visualização específico;
- Possibilidade de gerar visualizações mais simples e intuitivas: seguindo um conjunto de regras e minimizando alguns erros é possível criar algo simples e fácil de interpretar;
- Economia de tempo e recursos: proporcionar uma visão vasta sobre toda a execução de um determinado projeto e ajudar a economizar tanto o tempo de trabalho como os recursos que são despendidos.
- Decisões mais rápidas e conscientes: a informação apresentada adequadamente nos modelos de VI dá um panorama geral de toda a execução do projeto e permite ainda o aprofundamento em determinados pontos para que se possa compreender todo o contexto apresentado. Com uma rápida visualização, os gestores conseguem ter uma visão mais consciente e rápida, otimizando todo o ciclo de vida do projeto.

6.3. Limitações e perspetivas de desenvolvimentos futuros

A conceção do trabalho e o trabalho em si, no estado em que se encontram, têm algumas limitações, mas são uma importante base de trabalho para desenvolvimentos futuros. Nos parágrafos seguintes, são assinalados, sem nenhuma classificação de prioridade ou importância, algumas limitações e alguns dos possíveis trabalhos futuros.

As principais limitações que surgiram:

- Uma das grandes limitações deste estudo encontra-se na sua reduzida amostra na resposta ao questionário, neste caso alunos da Universidade do Minho. Desta forma, as conclusões deste trabalho aplicam-se apenas a essa mesma população, tratando-se por isso de um estudo exploratório;
- Este trabalho abarcou apenas duas fases do processo, o início e o desenvolvimento do mesmo, faltando, por isso, a análise da terceira e última fase, a finalização do processo de aprendizagem, neste caso a aplicação do framework em contexto real;
- Constrangimentos temporais e logísticos decorrentes do facto de se tratar de um trabalho no âmbito do mestrado;

Como desenvolvimentos futuros:

- Servir de suporte para que futuramente se realize um desenvolvimento mais aprofundado acerca da emergente e cada vez mais importante área da VI;
- Servir de apoio ao desenvolvimento de ferramentas mais sofisticadas e eficientes na produção de formas de VI;
- Apoio ao desenvolvimento mais aprofundado de uma revisão bibliográfica e ao desenvolvimento de um framework que permita estabelecer regras, identificar os pontos fortes e principais lacunas no que diz respeito ao desenvolvimento de metodologias de VI;
- Criar um mecanismo que facilite o desenvolvimento de modelos de VI;
- Melhorar e facilitar a análise realizada sobre os dados, nomeadamente de quem não tem grande conhecimento e à vontade a trabalhar com métodos de VI;

Certamente que muitas questões de investigação continuam sem resposta, mas espera-se que este trabalho possa motivar a realização de desenvolvimentos futuros, dando assim continuidade ao trabalho aqui iniciado.

BIBLIOGRAFIA

- Bezerianos, M. E. (2010). Exploration Views: Understanding Dashboard Creation and Customization for Visualization Novices.
- C. Ahlberg, E. W. (1995). *Information Visualization and*.
- Center, I. I. (2012). Big Data Analytics.
- Center, M. &. (2006). Introduction to Data Analysis Handbook.
- Dinu Airinei, D. H. (s.d.). Data Visualization in Business Intelligence.
- E.Micheline, B. (s.d.). *Exploration Views: Understanding Dashboard Creation and Customization for Visualization Novices*.
- Few, S. (2006). Information Dashboard Design.
- Few, S. (2007). *Perceptual Edge*.
- Few, S. (2007). Show me the numbers.
- Few, S. (2011). Infographics and the brain.
- Few, S. (2011.). Dashboard Design for at-a-glance monitoring.
- Few, S. (2011). Now You See It.
- Foley, J. (2011). *Visual Analytics Education*.
- H. Lam, E. B. (2010). *Seven Guiding Scenarios for Information Visualization Evaluation*.
- Herman, M. (2013). The Data Lake: Taking Big Data.
- Huang, W. (2009). *Measuring effectiveness of graph visualizations: A cognitive load perspective*.
- Hullman, J. (s.d.). Benefitting InfoVis with Visual Difficulties .
- Informação, D. d. (2013). *Guião de Avaliação de Pré-Dissertações e Pré-Projetos do Mestrado Integrado em Engenharia e Gestão de Sistemas de Informação*.
- J.Myatt, G. (2007). *Making sense of data*.
- Joahnsen, J. (2000). *On the impact of network latency on distributed systems design*.
- K.Robert, M. (2013). *Storytelling: The Next Step for Visualization*.
- Ken Peffers, T. T. (2008). A Design Science Research Methodology for Information Systems Research. In *Published in Journal of Management Information Systems*.
- Marianne Jorgensen, L. J. (s.d.). *Discourse Analysis as Theory and Method*.
- Maureen, S. (2010). *Information Visualization: Challenge for the Humanities*.
- McIntosh, L. (2013). Grammar of Graphics:Visualizing Data. *Machine Learning & Data Science Meetup*.
- Nielsen, J. (1993). *Usability Engineering, Academic Press*. London.

O'Connor, H. (s.d.). *A Step-by-Step Guide to Qualitative Data Analysis*.

Oracle. (2013). Big Data for the Enterprise.

Pamela A. Dupin-Bryant, D. H. (2014). Business Intelligence, Analytics And Data Visualization: A Heat Map Project Tutorial.

paper, W. (2013). Principles of Data Visualization - What We See in a Visual.

Raden, N. (2000). The Foundations of Analytics: Visualization, Interactivity and Utility.

Rebeckah, B. (2011). *The Importance of Data Visualization to Business Decision Making*.

S.Phil. (2014). *The Visual Organization*.

Schoenbach, V. J. (2014). Data analysis and interpretation.

Schoenbach, V. J. (2014). Data analysis and interpretation: Concepts and techniques for managing, editing, analyzing and interpreting data from epidemiologic studies.

Stodder, D. (2013). *Data Visualization and Discovery for Better Business Decisions*.

ULARU, E. G. (2012). Perspectives on Big Data and Big Data Analytics.

Ware, C. (2004). *Information Visualization: Perception for Design*.

Ware, C. (s.d.). *Information Visualization: Perception for design*.