





Universidade do Minho
Escola de Engenharia

Inês Sofia Paiva Bastos

**Visualização de Dados:
Fatores Críticos de Sucesso**

Dissertação de Mestrado
Mestrado em Engenharia de Sistemas

Trabalho efetuado sob a orientação de
Professora Doutora Maria Teresa Torres Monteiro
Professor Doutor Luís Paulo Peixoto dos Santos

Junho de 2023

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.



Atribuição-NãoComercial-SemDerivações
CC BY-NC-ND

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

AGRADECIMENTOS

O espaço desta secção de agradecimentos é insuficiente para agradecer, como devia, a todas as pessoas que ao longo da concretização deste projeto me acompanharam e contribuíram para a conclusão do mesmo. Uma dissertação nunca é um projeto individual, são várias as pessoas que, de uma forma ou de outra, se cruzaram e influenciaram o meu percurso académico.

O meu primeiro agradecimento é endereçado à minha orientadora, Professora Doutora Maria Teresa Torres Monteiro, por todo o apoio inegável e conhecimento partilhado neste percurso académico, pela prontidão, disponibilidade constante, entusiasmo e confiança que sempre demonstrou. Não poderia ter escolhido uma pessoa melhor para me guiar neste último passo desta jornada académica. Um agradecimento especial ao meu coorientador, Professor Doutor Luís Paulo Peixoto dos Santos, por todo o acompanhamento, apoio e conhecimentos transmitidos para a realização desta dissertação.

A todos os meus amigos que me acompanharam nesta vida académica ao longo destes 5 anos e fizeram deles os melhores anos da minha vida!

À Diana e à Sara, por estarem comigo desde o 1º ano de escolaridade, não sei se algum dia irei conseguir agradecer a tudo o que fizeram por mim.

Ao David, ao Rui, à Lucie, à Inês, à Joana e ao “*Sunsetadas*”, por estarem sempre presentes e disponíveis sempre que precisei.

À *JTA Consulting* e a todos os meus colegas, por me estarem a dar uma experiência empresarial incrível e por se mostrarem dispostos a ajudar sempre que necessário.

Ao João por ser o meu braço direito e me apoiar em todas as minhas decisões e por estar sempre lá para mim.

Por fim, o agradecimento mais importante, a toda a minha família. Aos meus pais e irmão pelo apoio incondicional em toda a minha vida, permitindo-me atingir os meus objetivos e ser a pessoa que sou hoje. Sempre acreditaram em mim e sempre me deram a mão para me ajudar a alcançar tudo o que consegui até hoje.

Sem vocês nada disto seria possível.

DECLARAÇÃO DE INTEGRIDADE

Declaro ter atuado com integridade na elaboração do presente trabalho académico e confirmo que não recorri à prática de plágio nem a qualquer forma de utilização indevida ou falsificação de informações ou resultados em nenhuma das etapas conducente à sua elaboração.

Mais declaro que conheço e que respeitei o Código de Conduta Ética da Universidade do Minho.

RESUMO

Visualização de Dados: Fatores Críticos de Sucesso

Com a chegada da era digital, tem havido a necessidade de apresentar grandes quantidades de dados de forma simples, acessível e compreensível.

Esta dissertação consiste na realização de um estudo que visa identificar fatores de sucesso e boas práticas associados à Visualização de Dados.

O presente relatório insere-se em contexto empresarial, a pedido da “*JTA - The Data Scientists*” e foi realizada no âmbito do segundo ano do Mestrado em Engenharia de Sistemas da Universidade do Minho (UM).

Com a recolha dos Fatores Críticos de Sucesso da Visualização de Dados feita neste projeto, é possível desenvolver *Reports* de dados eficientes.

Para uma melhor compreensão dos fatores que devem ser evitados na Visualização de Dados, foram apresentados várias más práticas e exemplos visuais utilizados em projetos desta natureza.

De forma a validar a solução de boas práticas, foram elaborados três *Reports* com a ferramenta *PowerBI* tendo em consideração esses mesmos Fatores. Assim, foi possível apresentar Visualizações de Dados eficazes, com a implementação de boas práticas.

A título de curiosidade, recorreu-se à ferramenta de Inteligência Artificial *ChatGPT* para colocar duas questões relacionadas com os objetivos deste trabalho.

Palavras-chave: Visualização de Dados, Inteligência Empresarial, Análise Empresarial, *Power BI*

ABSTRACT

Data Visualization: Critical Success Factors

With the arrival of the digital age, there has been a need to present large amounts of data in a simple, accessible, and understandable way.

This dissertation consists of a study that aims to identify success factors and best practices associated with Data Visualization.

This report is written in a business context, at the request of "JTA - The Data Scientists" and was performed during the second year of the MSc in Systems Engineering at the University of Minho (UM).

By collecting the Critical Success Factors of Data Visualization done in this project, it is possible to develop efficient Data Reports.

For a better understanding of what factors should be avoided in Data Visualization, several bad practices and visual examples used in projects of this nature have been presented.

In order to validate this solution, three reports were created with the PowerBI tool taking into account these same factors. Thus, it was possible to visually demonstrate effective Data Visualizations and the implementation of best practices.

For curiosity, the ChatGPT Artificial Intelligence tool was used to ask two questions related to the objectives of this work.

Keywords: *Data Visualization, Business Intelligence, Business Analytics, Power BI*

ÍNDICE

Agradecimentos	ii
Resumo	iv
Abstract	v
Índice de figuras	ix
Índice de tabelas	xii
Lista de abreviaturas, siglas e acrónimos	xiii
1. Introdução	1
1.1 Enquadramento e Motivação	1
1.2 Objetivos e Resultados Esperados	1
1.3 Metodologia de Investigação	2
1.4 Questões de investigação	4
1.5 Estrutura da dissertação	5
2. Revisão Bibliográfica	6
2.1 Estratégia de Pesquisa	6
2.2 Business Intelligence	7
2.2.1 Processo do <i>Business Intelligence</i>	7
2.2.2 Vantagens da adoção do <i>Business Intelligence</i>	8
2.3 Business Analytics	9
2.3.1 Vantagens da adoção do <i>Business Analytics</i>	10
2.4 Data Visualization	11
3. Ferramentas de Visualização de dados	13
3.1 Microsoft Power BI	13
3.1.1 <i>Dataset, Dashboard e Report</i>	13
3.1.2 Escolher os melhores Visuais no <i>Power BI</i>	15
3.2 Excel	19

3.3	<i>Google Charts</i>	20
3.4	<i>Tableau</i>	21
3.5	<i>Zoho Analytics</i>	22
3.6	<i>Datawrapper</i>	23
3.7	<i>Qlik Sense</i>	24
3.8	<i>Fusion Charts</i>	25
3.9	<i>Domo</i>	26
3.10	<i>Google Analytics</i>	27
3.11	<i>Visme</i>	28
4.	Fatores Críticos de Sucesso	30
4.1	Definição de Sucesso na área do <i>Business Intelligence</i>	30
4.3	Definição de Sucesso na Visualização de Dados	30
4.4	Fatores Críticos de Sucesso na Visualização de Dados	31
4.5	Fatores Críticos de Sucesso no <i>Business Intelligence</i>	37
5.	Más práticas de visualização de dados.....	41
5.1	Exemplos de Más Práticas de Visualização de Dados	41
5.2	Exemplos de Visualização de Dados Mal Desenvolvidas.....	47
5.3	O impacto de uma Má Visualização dos Dados	49
6.	Reports em <i>Power BI</i> tendo em consideração os Fatores de Sucesso	51
6.1	<i>Dataset</i>	51
6.2	Report Filmes Originais <i>Netflix</i>	52
6.3	Report de vendas do Supermercado <i>Giant</i>	55
6.4	Report <i>Top 50</i> Músicas do <i>Spotify</i>	60
7.	Conclusão	64
7.1	Síntese do trabalho realizado.....	64
7.2	ChatGPT	65

7.3 Perspetivas de trabalho futuro	68
Referências Bibliográficas.....	69
Anexo I – Guião de avaliação de uma Visualização de Dados.....	74
Anexo II – Datasets	75
A.1- Dataset Netflix.....	75
A.2- Dataset Spotify	76
A.3- Dataset Supermercado.....	77

ÍNDICE DE FIGURAS

Figura 1- Processo de uma Revisão Sistemática da Literatura.....	4
Figura 2- As 5 fases do Business Intelligence	8
Figura 3- As três categorias do Business Analytics.....	10
Figura 4- Exemplo de um Dashboard em Power BI.....	14
Figura 5- Exemplo de um Report em Power BI	14
Figura 6- Estrutura do Dataset, Report e Dashboard no Power BI.....	15
Figura 7- Tipos de Visualizações disponíveis por padrão no Power BI.....	15
Figura 8- Visualização da ferramenta Power BI	18
Figura 9- Visualização da ferramenta Microsoft Excel.....	19
Figura 10- Visualização da ferramenta Google Charts	20
Figura 11- Visualização da ferramenta Tableau	21
Figura 12- Visualização da ferramenta Zoho Analytics.....	22
Figura 13- Visualização da ferramenta Datawrapper	23
Figura 14- Visualização da ferramenta QilkSense.....	24
Figura 15- Visualização da ferramenta Fusion Charts	25
Figura 16- Visualização da ferramenta Domo	26
Figura 17- Visualização da ferramenta Google Analytics.....	27
Figura 18- Visualização da ferramenta Visme	28
Figura 19- Esquema de cor Sequencial	32
Figura 20- Esquema de cor Divergente.....	32
Figura 21- Esquema de cor Qualitativo.....	32
Figura 22- Abordagem visual de pequenos múltiplos	33
Figura 23- 10 princípios da Visualização de Dados	34
Figura 24- Fórmula do Data-Ink Ratio	35
Figura 25- Exemplo de um Data-Ink baixo	35
Figura 26- Exemplo de um Data-Ink elevado.....	35
Figura 27- Fatores Chave de Sucesso do Business	41
Figura 28- Exemplo do uso excessivo de variáveis num Pie Chart	42
Figura 29- Utilização de escalas inconsistentes	43
Figura 30- Utilização de valores errados no eixo dos Y	44

Figura 31- Utilização errada do eixo dos y - Exemplo real.....	44
Figura 32- Escala Linear vs Escala Logarítmica	45
Figura 33- Exemplo 1 de Escolha Errada de Cores	46
Figura 34- Exemplo 2 de Escolha Errada de Cores	47
Figura 35- Exemplo de uma DV com manipulação dos dados.....	47
Figura 36- Pie Chart com má representação dos dados.....	48
Figura 37- Gráfico de barras com má representação dos dados.....	48
Figura 38- Mapa com má representação dos dados	49
Figura 39- Cartões com variáveis relevantes do Report Netflix.....	52
Figura 40- Top 10 filmes com melhor classificação e Top 10 géneros de filmes.....	52
Figura 41- Report filmes Originais Netflix.....	53
Figura 42- Visuals do Report de Vendas do supermercado Giant.....	55
Figura 43- Visuals com o Top 3 compradores do Supermecado Giant.....	56
Figura 44- Visuals relativos ao lucro e total de vendas do Supermecado Giant.....	57
Figura 45- Top 10 Produtos mais vendidos do Supermecado Giant.....	57
Figura 46- Report de vendas do Supermecado Giant	58
Figura 47- Visualizações do artista mais popular	60
Figura 48- Secção Top 5 de várias categorias.....	60
Figura 49- Visual com informação relativa às músicas	61
Figura 50- Report Top 50 músicas Spotify	61
Figura 51- Visualização com legenda- propriedades da música	62
Figura 52- Visual com legenda do Top 5 músicas com mais Vivacidade.....	62
Figura 53- Reposta do ChatGPT à pergunta "Quais os fatores críticos de sucesso da Visualização de Dados?"	66
Figura 54- Reposta do ChatGPT à pergunta "Quais são as más práticas de Visualização de Dados?.." ..	67
Figura 55- Excerto do Dataset Netflix.....	75
Figura 56- Excerto do Dataset Spotify.....	76
Figura 57- Excerto do Dataset Supermecado Giant	77

ÍNDICE DE TABELAS

Tabela 1- Ferramentas de Visualização de Dados	29
Tabela 2- Compilação dos Fatores Críticos de Sucesso.....	36
Tabela 3- Fatores Críticos de Sucesso do Business Intelligence	38
Tabela 4- Tabela de Avaliação de uma Visualização de Dados	74

LISTA DE ABREVIATURAS, SIGLAS E ACRÓNIMOS

BA – Business Analytics

BD – Base de Dados

BI – Business Intelligence

CSF – Critical Success Factors

DA – Data Analytics

DIR – Data-Ink Ratio

DV – Data Visualization

IMDB – Internet Movie Database

KSF – Key Success Factors

PBI – Power BI

ROI – Return On Investment

RSL - Revisão Sistemática da Literatura

1. INTRODUÇÃO

Esta dissertação insere-se em contexto empresarial, a pedido da “JTA - The Data Scientists” e foi realizada no âmbito do Mestrado em Engenharia de Sistemas da Universidade do Minho.

A JTA Consulting é uma empresa de consultoria especializada em oferecer aos seus clientes uma gama de serviços integrados que abrangem Plataformas de Dados, Análise de Dados e Relatórios de Dados.

No presente capítulo é apresentado o enquadramento e motivação que levaram à realização da presente dissertação, bem como os objetivos e resultados esperados da mesma. Posteriormente, será descrita a abordagem metodológica de investigação selecionada para a elaboração da dissertação. Por fim, é apresentada a estrutura do presente documento.

1.1 Enquadramento e Motivação

Com a chegada da era digital, tem havido a necessidade de apresentar grandes quantidades de dados de forma simples, acessível e compreensível. As empresas geram dados todos os dias e é difícil para os utilizadores visualizar, explorar, e utilizar todos estes dados e, quando justificável, em tempo real (Musa et al., 2016; Muskan et al., 2022). O *Business Intelligence* permite a análise de grandes volumes de dados e é definido como um sistema que combina 3 tarefas: recolha de dados, armazenamento de dados e gestão do conhecimento. (Błazewicz, 2000). A análise destes dados permite tratar aspetos determinantes para o sucesso da empresa, tais como possíveis melhorias, previsões de mercado e otimização na tomada de decisão (AlWreikat et al., 2019). Os dados devem ser apresentados de forma clara para quem os recebe e analisa.

Esta dissertação consiste na realização de um estudo que visa identificar fatores de sucesso e boas práticas associados à Visualização de Dados.

1.2 Objetivos e Resultados Esperados

O principal objetivo desta dissertação é identificar quais os fatores críticos de sucesso e quais as boas práticas do desenvolvimento de uma Visualização de Dados.

Atualmente, existem muitas visualizações que não são bem desenvolvidas e, conseqüentemente, a análise dessas representações visuais não vão ao encontro do que era pretendido. Algumas delas podem até mesmo representar o oposto do objetivo inicial do desenvolvimento dessa mesma visualização de dados.

Deste modo, é pretendido identificar os aspetos a ter em atenção na representação visual dos dados propriamente dita.

Para isso, será necessário:

- Definir o que deve ser considerado como “sucesso” no que concerne à representação de dados;
- Definir quais os principais fatores de sucesso da Visualização de Dados;
- Catalogar ambientes e produtos lógicos de *software* mais utilizados para produzir Visualizações de Dados;
- Ilustrar visualmente más práticas encontradas em projetos desta natureza;
- Fazer a ligação das ilustrações com os respetivos fatores de sucesso.

Em suma, será importante identificar o impacto que a Visualização de Dados tem na área de *Business Intelligence* e a importância de seguir os fatores de sucesso para um bom desenvolvimento de uma Visualização de Dados (Camm et al., 2017).

1.3 Metodologia de Investigação

Foi necessário planear uma abordagem metodológica para que fosse possível o desenvolvimento desta dissertação. Como tal, o método que se pretende usar para uma melhor orientação e execução deste trabalho de dissertação é a Revisão Sistemática de Literatura.

A Revisão de Literatura resume e fornece uma avaliação crítica da literatura disponível sobre um assunto específico (Thomé et al., 2016). Nesta dissertação, vai ser feita essa mesma avaliação crítica para analisar e obter os fatores críticos de sucesso da Visualização de Dados.

Uma Revisão Sistemática pode ser explicada como um método e processo de investigação para identificar e avaliar criticamente a investigação relevante, bem como para recolher e analisar dados da referida investigação. O objetivo deste tipo de revisão é identificar todas as provas empíricas que se enquadram nos critérios de inclusão pré-especificados para responder a uma determinada questão ou hipótese de investigação (Snyder, 2019).

Ou seja, é importante rever pontos críticos do conhecimento atual sobre um determinado tema e respetivas questões de investigação e encontrar estudos mais específicos relacionados com as variáveis da questão (Carrera-Rivera et al., 2022).

Como tal, segundo (Xiao & Watson, 2019), todas as revisões podem ser conduzidas seguindo oito etapas comuns:

- 1- Definir o problema de investigação: secção 1.4 - as questões de investigação impulsionam todo o processo de revisão bibliográfica porque é a partir do problema de investigação que vai ser feita toda a pesquisa;
- 2- Desenvolver e validar o protocolo de revisão: secção 1.1, 1.2, 1.3 - o protocolo de revisão é um plano pré-definido que especifica os métodos utilizados na realização da revisão. Este protocolo deve incluir o objetivo do estudo, questões de investigação, estratégias de pesquisa e um calendário de projeto.
- 3- Fazer uma pesquisa da literatura: secção 2.1 - a pesquisa bibliográfica fornece materiais e dados para a revisão da literatura; por conseguinte, uma revisão sistemática depende de uma pesquisa sistemática da literatura. Existem três fontes principais para encontrar literatura: (1) bases de dados eletrónicas (*Web of Science, EBSCO, ProQuest, IEEE Xplore, Google Scholar*); (2) pesquisa para trás (investiga e analisa todas as referências citadas num único artigo); e (3) pesquisa para a frente (consiste em encontrar todos os artigos que citam um artigo específico);
- 4- Delimitar os critérios de inclusão: secção 2.1 - cada artigo deve ser examinado para decidir se deve ser incluído para extração e análise de dados. Este critério deve excluir todos os artigos que não estão relacionados com as perguntas de investigação ou que fornecem pouca orientação sobre o tema.
- 5- Avaliar a qualidade dos artigos: secção 2.1 - depois de restringir os artigos, surge a parte da avaliação desses artigos para preparar o conjunto de artigos para a fase seguinte: extração e síntese de dados. Os estudos com mais qualidade devem ser os que dão suporte para construir grandes argumentos de investigação antes de passar aos estudos de qualidade média. Os estudos de qualidade baixa devem ser utilizados apenas como suplemento;
- 6- Extrair os dados: secção 2.1 - este processo envolve muitas vezes a codificação, especialmente para alargar as revisões. É importante estabelecer se a codificação será indutiva ou dedutiva (ou seja, se a codificação será baseada ou não nos dados ou conceitos pré-existentes);
- 7- Analisar e sintetizar os dados: secção 2.1 - uma vez concluído o processo de extração de dados, é necessário organizar os dados de acordo com a revisão escolhida.

- 8- Relatar os resultados: secção 2.1 - para que as revisões de literatura sejam fiáveis, o processo de revisão sistemática da literatura deve ser descrito com detalhe suficiente. Isto permite que outros investigadores sigam os mesmos passos descritos e cheguem aos mesmos resultados.

Estes oito passos podem ser esquematizados na Figura 1.

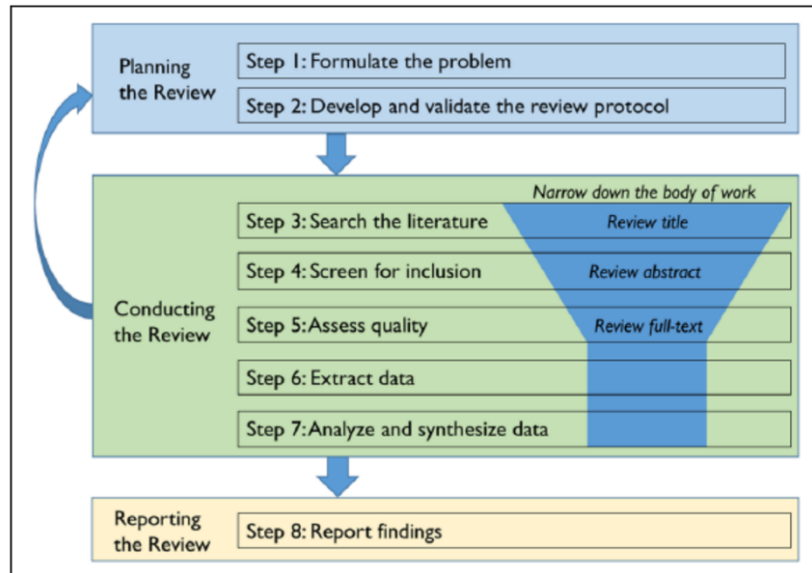


Figura 1- Processo de uma Revisão Sistemática da Literatura

(Xiao & Watson, 2019)

A partir desta abordagem dedutiva, espera-se identificar os passos fulcrais a seguir para atingir os fatores críticos de sucesso da Visualização de Dados.

É importante salientar que, a partir do ponto 4, os passos não foram tratados de forma explícita e apenas foi apresentada a abordagem e o método de desenvolvimento na secção 2.1.

1.4 Questões de investigação

Para o desenvolvimento do ponto 1 (“Definir o problema de investigação”) mencionado anteriormente, foi necessário definir quais as questões de investigação fulcrais que vão impulsionar a dissertação. É um primeiro passo essencial para que seja possível fazer a Revisão Sistemática da Literatura de acordo com os objetivos pretendidos e sem perder o foco de investigação. Sendo assim, as questões de investigação chave para a presente dissertação são:

- O que deve ser considerado como “sucesso” no que toca à visualização de dados?
- Quais os principais fatores de sucesso da Visualização de Dados?
- Quais são os principais ambientes e produtos lógicos de *software* mais utilizados para produzir Visualizações de Dados?

Além de responder extensivamente a estas perguntas de investigação, o objetivo será de ilustrar visualmente boas e más práticas encontradas em projetos desta natureza e fazer a ligação das ilustrações com os respetivos fatores de sucesso.

1.5 Estrutura da dissertação

Um documento devidamente estruturado, facilita a compreensão dos conceitos nele mencionados podendo, assim, servir de facilitador para o alcance de alguns objetivos previamente referidos. Baseado nisso, a escolha recaiu numa estrutura composta por 7 capítulos e 2 Anexos.

O documento inicia-se com uma contextualização da dissertação, seguida dos objetivos e resultados esperados. Ainda neste capítulo é abordada a metodologia de investigação adotada no desenvolvimento deste documento e as respetivas questões de investigação. Por fim, é apresentada a estrutura da dissertação.

O segundo capítulo é destinado à revisão bibliográfica. Primeiramente é descrita a estratégia de pesquisa adotada e, de seguida, o capítulo é dividido em três conceitos fulcrais do tema em questão: *Business Intelligence*, *Business Analytics* e *Data Visualization*. É feita uma breve explicação de cada um, apresentado o processo e as respetivas vantagens.

No terceiro capítulo são apresentadas várias ferramentas interessantes de Visualização de dados, tais como o *Microsoft Power BI*, *Excel*, *Google Charts*, *Tableau*, *Zoho Analytics*, *Datawrapper*, *Qlik Sense*, *Fusion Chart*, *Domo*, *Google Analytics* e *Visme*.

O quarto capítulo é referente aos fatores críticos de sucesso. Neste capítulo, o sucesso é definido no contexto do *Business Intelligence* e da Visualização dos Dados. Por fim, são apresentados os fatores críticos de sucesso na Visualização dos Dados.

No quinto capítulo são apresentados exemplos de más práticas de Visualização de Dados e o respetivo impacto.

No sexto capítulo são apresentados e explicados 3 *Reports* que foram desenvolvidos tendo em consideração os fatores críticos de sucesso.

No último capítulo são apresentadas as Conclusões do projeto e é efetuada uma análise global dos objetivos atingidos. São ainda reportadas as respostas da ferramenta ChatGPT a duas questões relacionadas com este projeto, algumas sugestões para um futuro desenvolvimento e apreciação final.

No Anexo I é apresentado um guião para a avaliação de uma Visualização de Dados baseado nos fatores críticos de sucesso. O Anexo II contém um excerto dos datasets usados no desenvolvimento dos reports apresentados no Capítulo 6.

2. REVISÃO BIBLIOGRÁFICA

Este capítulo será dedicado à revisão da literatura. Este tópico surge para aprofundar os conceitos teóricos fundamentais de forma a suportar toda a elaboração da dissertação. Irão ser abordados vários temas tais como: *Business Intelligence*, *Business Analytics* e *Data Visualization*.

2.1 Estratégia de Pesquisa

Para a realização da revisão da bibliografia foi necessário recorrer à Revisão Sistemática da Literatura. Esta estratégia consiste em reunir, a partir de uma pesquisa exaustiva, um conjunto de artigos, dissertações, teses e livros relacionados com o objeto de estudo em questão para a realização de uma análise de todos esses dados e fundamentar o problema em questão.

Como tal, antes de passar para a revisão da literatura propriamente dita, para o ponto 2 mencionado anteriormente, o desenvolvimento de um protocolo de revisão é essencial para facilitar o trabalho de pesquisa e para ser uma procura consistente. Além disso, formular o problema ajuda a alcançar as questões-chave que se pretendem alcançar com esta dissertação.

Sendo assim, a pesquisa e a recolha de documentos foram feitas a partir de conceitos chave essenciais para sustentar as questões de investigação, como por exemplo "*Data Visualization*", "*Business Intelligence*", "*Power BI*", "*Business Analytics*", entre outros.

A fonte de pesquisa consistiu em bases de dados eletrónicas, maioritariamente a *Web of Science*, *Scopus*, *Google Scholar* e RepositóriUM. Além disso, também foi feita uma pesquisa para trás que consiste em investigar e analisar todas as referências citadas num único artigo. Isto é, parte por recolher todos os artigos que serviram de suporte para um documento que tem bastante relevância para o problema em questão e, assim, recolher mais informação e detalhe acerca desse tema.

A data de publicação dos artigos e documentos foi um dos aspetos que foi tido em atenção na pesquisa de literatura. O presente tema é bastante atual e sofre constantes atualizações com a digitalização frequente da sociedade e das empresas e, portanto, recolher e analisar dados publicados com datas atuais e recentes torna esta dissertação mais atualizada. Contudo, também foram considerados artigos com datas de publicação anteriores de maneira a sustentar a parte teórica do problema e por se terem mostrado relevantes para o tema.

O método adotado para delimitar os critérios de exclusão consistiu em ler o título, palavras-chave e resumo do artigo. Caso estes três aspetos se mostrassem relevantes e acrescentassem valor ao tema, partia-se então para a leitura do artigo na íntegra.

Posteriormente, todos os documentos e artigos que foram dando respostas às questões de investigação e apresentando credibilidade da fonte, foram considerados como artigos de qualidade elevada e interessantes para o tema e, conseqüentemente, serviram de suporte para construir argumentos nesta dissertação. A pesquisa seguiu a metodologia dedutiva, ou seja, foi baseada nos dados ou conceitos já existentes.

Depois de selecionada a literatura que iria servir de suporte para esta dissertação e tendo toda a informação devidamente recolhida, foi utilizada a ferramenta Mendeley Desktop, para fazer a gestão da informação e devida referenciação.

2.2 Business Intelligence

No mundo atual altamente competitivo, a qualidade do *Business Intelligence (BI)* de uma organização pode significar não só a diferença entre lucro e prejuízo, mas também a diferença entre sobrevivência e falência (Ranjan, 2008).

Com a chegada da era digital, tem havido a necessidade de apresentar grandes quantidades de dados de forma simples, acessível e compreensível. As empresas geram dados todos os dias e, conseqüentemente, como a quantidade de dados disponíveis tem aumentado drasticamente, é bastante difícil para os utilizadores visualizar, explorar, e utilizar todos estes dados (Musa et al., 2016).

O *Business Intelligence* permite a análise de grandes volumes de dados e é definido como um sistema que combina recolha e armazenamento de dados e gestão do conhecimento (Błażewicz, 2000). O desenvolvimento de um sistema de *Business Intelligence* traz bastantes benefícios para os proprietários de empresas. Com base nos dados disponíveis e a partir da visão completa dos conhecimentos passados e atuais, permitem dar apoio na tomada de decisões e no desenvolvimento de estratégias, obtendo, assim, vantagem competitiva em relação a empresas concorrentes (AlWreikat et al., 2019)

2.2.1 Processo do Business Intelligence

Segundo (Jakhar & Krishna, 2020), *BI* é o processo de transformação de dados empresariais em informação, informação em conhecimento e esse conhecimento é transformado em inteligência.

Segundo os autores, o *Business Intelligence* divide-se em 5 fases:

1. Fontes de dados - o *Business Intelligence* consiste na extração de informação de múltiplas fontes de dados.
2. Análise de Dados - nesta fase, o *BI* está prestes a sintetizar o conhecimento útil dos dados selecionados. Consiste em estimar as tendências atuais, integrar e resumir informações

relevantes ou díspares, validar modelos de compreensão e prever informações em falta ou tendências futuras. O processo de análise de dados é também chamado *data mining* ou descoberta de conhecimento.

3. Consciência da Situação – esta fase consiste em filtrar a informação irrelevante, e colocar a restante informação útil no contexto de negócio e do ambiente envolvente.
4. Avaliação dos Riscos – consiste em descobrir possíveis tomadas de decisão. Trata-se de ajudar a ponderar o risco atual e futuro, o custo ou os benefícios de tomar uma determinada ação comparativamente a outras.
5. Apoio à Decisão – o *BI* identifica-se com a utilização correta da informação. Tem como objetivo alertar para diversas situações para que possam ser tomadas medidas preventivas.

Estas 5 fases podem ser ilustradas na Figura 2.



*Figura 2- As 5 fases do Business Intelligence
(Jakhar & Krishna, 2020)*

2.2.2 Vantagens da adoção do *Business Intelligence*

O *Business Intelligence* converte dados em conhecimento com a intenção de fornecer informação analítica aos utilizadores.

Atualmente, as empresas, através do *Business Intelligence*, podem tomar decisões empresariais estratégicas que aumentam a produtividade, melhoram as receitas e aumentam o crescimento da empresa. É uma ferramenta que ajuda as empresas a classificar, comparar e rever os dados e, posteriormente, filtrar, processar, armazenar e analisar todos esses dados para os transformar em *insights*. A partir destes insights, a empresa pode obter vantagem competitiva de mercado (Jakhar & Krishna, 2020).

O *BI* pode proporcionar muitos benefícios às empresas e faz com que seja possível eliminar muito do trabalho dentro de uma organização, melhorar a comunicação entre os diversos departamentos durante a coordenação das atividades, e permitir às empresas responder rapidamente às alterações financeiras, preferências do cliente, e operações da cadeia de abastecimento. De forma geral, o *Business Intelligence* ajuda a melhorar o desempenho global da empresa que o utiliza.

Uma outra vantagem do *BI* é que ajuda e acelera o processo de tomada de decisão, como agir de forma rápida e correta sobre a informação adquirida antes que as empresas concorrentes o façam (Ranjan, 2009).

Segundo (Jakhar & Krishna, 2020), definem as principais vantagens do *BI* como:

1. Apenas existir uma versão verdadeira - com a ajuda do *Business Intelligence*, toda a informação pode ser combinada sem ser necessário verificar a lista de fontes.
2. Persistência de dados - qualquer informação pode estar disponível e ser consultada em qualquer altura, isto significa que, se os dados forem atualizados ou porventura apagados, podem ser facilmente recuperados.
3. Desempenho - são armazenados no armazém de dados de forma atempada, antes de serem processados e analisados.
4. Flexibilidade - as possibilidades de elaboração de relatórios e análise dos mesmos são intermináveis. A flexibilidade de conceção do *Business Intelligence* permite que a solução possa ser restabelecida ou recuperada quando necessário.
5. Gestão dos Recursos Humanos Eficiente – o *BI* ajuda os recursos humanos que não são capazes de fazer análises complexas de todas as fontes de dados.

2.3 Business Analytics

Como mencionado anteriormente, tal como os dados, também a análise dos dados se tem tornado fulcral para a tomada de decisão nas organizações atuais. A crescente facilidade de recolha, armazenamento e processamento de dados é acompanhada pelo aumento do conjunto e da diversidade de dados para análise (Power et al., 2018).

(Raghupathi & Raghupathi, 2021) definem *Business Analytics* como a aplicação de modelos, métodos e ferramentas à análise de dados para obter conhecimentos para tomar decisões.

Segundo (Sun et al., 2017), essa forma alargada de análise de dados divide-se em três categorias:

- **Descritiva** - esta análise aborda factos- o que aconteceu e o que está a acontecer;
- **Preditiva** - concentra-se na previsão de tendências e fornece uma possível previsão empresarial e a razão disso acontecer;
- **Prescritiva** - responde ao que devemos fazer e por que razão o devemos fazer;



Figura 3- As três categorias do Business Analytics

(Sun et al., 2017)

2.3.1 Vantagens da adoção do *Business Analytics*

O *Business Analytics* (BA) é um conjunto de técnicas e processos que podem ser utilizados para analisar dados e ajuda a melhorar o desempenho empresarial na tomada de decisão. O *Business Analytics* pertence ao *Business Intelligence* e permite que as empresas melhorem as suas capacidades para competir no mercado (Ajah & Nweke, 2019).

O BA pode também fornecer uma visão sobre quais as estratégias que estão a funcionar e onde se deve investir no crescimento. Segundo (Liu et al., 2018), existem benefícios transacionais, informativos e estratégicos que, quando combinados, proporcionam uma vantagem competitiva à organização:

- **Benefícios transacionais** - são definidos como resultados tangíveis e diretos de um sistema de BA. Estes benefícios incluem a poupança de tempo, menor número de reuniões e redução

do número funcionários necessários. Um programa *BA* é geralmente gerido através de um *software* e por sistemas de processamento analítico online. Como consequência deste tipo de tecnologia, as métricas empresariais podem ser entregues automaticamente aos utilizadores finais de forma atempada. Esta entrega de dados em tempo real pode reduzir a necessidade de cálculos manuais e a possibilidade de erros. Como resultado, a organização pode conseguir mais com menos reuniões, menos processos manuais e, em muitos casos, menos funcionários, o que reduzirá drasticamente o custo comercial.

- **Benefícios informativos** - um programa de *BA* deve ser capaz de transformar dados em informação. Através da análise da informação adquirida, uma organização pode acompanhar os comportamentos dos clientes e os padrões, melhorar a elaboração de *websites*, analisar conteúdos gerados pelo utilizador, compreender a inteligência e competitividade de mercado e, em última análise, começar a prever o futuro.
- **Benefícios estratégicos** - uma estratégia forte é apoiada por um bom *BA*. O *Business Analytics* permite às organizações compreender as perceções empresariais, tomar decisões e resolver questões-chave que são barreiras ao sucesso. Ao extrair estes dados e ao analisá-los para tomar melhores decisões, as organizações podem concentrar-se no que é melhor para a organização. O *BA* é fulcral para ajudar uma empresa a manter-se inovadora em relação à concorrência.

2.4 Data Visualization

A Visualização de Dados é uma componente chave e fundamental para as empresas. Esta visualização tem sido amplamente utilizada para ajudar na tomada de decisões que estão intimamente ligadas às receitas chave de muitas empresas. A Visualização de Dados identifica descrições, cria hipóteses, e desenvolve ideias a partir de grandes volumes de dados que contribuem para a investigação científica e a previsão (Muskan et al., 2022). A sua essência está na transformação de dados abstratos em visões físicas e faz com que seja um meio poderoso para apresentar histórias convincentes de dados aos seres humanos (Qin et al., 2020).

O objetivo da visualização é obter uma visão, através de gráficos interativos, sobre vários aspetos relacionados com algum processo que nos interessa do mundo real. O termo "*insight*" é utilizado muito frequentemente quando o assunto é a visualização. Isto porque, a visualização pode ajudar a obter vários tipos de *insight*, respondendo a vários tipos de perguntas.

No contexto atual, utilizamos a palavra "*insight*" para descrever dois tipos de informação que obtemos de uma aplicação de visualização:

1. respostas a perguntas concretas sobre um determinado problema;
2. factos sobre um dado problema de que não estávamos cientes (Telea, 2007).

Estas representações visuais bem concebidas proporcionam acesso visual de fácil compreensão a enormes quantidades de dados e facilitam a interpretação dos resultados.

Segundo o autor (Naidoo & Campbell, 2016), as melhores práticas são categorizadas em seis áreas principais:

1. **Organização da informação e sua disposição** - os dados apresentados de forma muito complexa e confusa podem levar ao oposto do objetivo que é a Visualização de Dados. Isto pode ser ultrapassado através da redução da informação que não é relevante, e dar pistas para guiar os utilizadores aos detalhes essenciais. A memória dessa representação visual pode ser conquistada quando a disposição da página está bem planeada.
2. **Esquemas de cor** - os dados apresentados com paletas de cores discordantes e cores temáticas inconsistentes podem interferir com a retenção de memória. Ao reduzir o número de cores para gráficos complexos e ao utilizar um maior contraste entre informação relevante e o fundo permite aos utilizadores identificar mais facilmente o objetivo e a informação pretendida.
3. **Infográficos** - os dados exibidos com visualizações confusas e descoordenadas podem confundir quem analisa. Esse problema pode ser combatido através da redução da desorganização, legendas explícitas e claras, pistas de visualização, e seleções gráficas apropriadas.
4. **Visuais bem escolhidos e apresentados** - a utilização de gráficos simples e temáticos em vez de imagens e gráficos de grandes dimensões faz com haja um maior foco sem distrair os utilizadores.
5. **Legibilidade** - a utilização de tipos de letra de fácil leitura, com espaço branco entre os títulos e o texto descritivo ou imagens, bem como a distinção entre o visual e o seu fundo promove a legibilidade.
6. **Funcionalidade** - os dados apresentados sem métricas ou filtros podem dificultar a capacidade dos utilizadores de contextualizar a informação.

Um dos tópicos deste capítulo consistiu em abordar temas como: BI, BA, DV. No capítulo que se segue serão apresentadas algumas ferramentas de Visualização de Dados que agregam estes conceitos.

3. FERRAMENTAS DE VISUALIZAÇÃO DE DADOS

Neste capítulo irão ser apresentadas várias ferramentas de Visualização de Dados. Contudo, como a empresa “*JTA- The Data Scientists*” utiliza como ferramenta o *Microsoft Power BI*, irá ser feita uma abordagem de pesquisa mais profunda sobre a mesma.

3.1 Microsoft Power BI

Microsoft Power BI é uma ferramenta de *Business Intelligence* para analisar dados, partilhar conhecimentos e obter respostas rapidamente com a ajuda da visualização interativa de dados utilizando *Dashboards* (Mandava Geetha Bhargava et al., 2018).

Além disso, pode extrair, transformar e carregar os dados de várias fontes e moldar esses dados. Dentro do *software*, é utilizada a linguagem DAX para agregar os dados e criar métricas necessárias para o *Report* (Liu & Chen, 2022).

Esta ferramenta baseia-se no processamento de dados e envolve quatro etapas: limpeza dos dados, modelação dos dados, cálculo e visualização dos Dados e análise dos dados (Cheng & Gu, 2022).

3.1.1 Dataset, Dashboard e Report

Segundo (Rad, 2018), *Dataset* pode ser definido como o local onde é estabelecida a ligação à fonte de dados no *Power BI*. Não é possível criar um relatório sem um *Dataset*.

Um conjunto de dados pode ser utilizado num ou vários relatórios e as visualizações podem ser exibidas em vários *Dashboards*.

Segundo o mesmo autor, um *Dashboard* é um elemento de visualização que, normalmente, é o ponto de navegação e a visão de alto nível do conteúdo no *Power BI*. Um *Dashboard* pode ser criada em cima de um ou mais relatórios. O objetivo do *Dashboard* não é filtrar, mas apenas apresentar uma visão rápida dos relatórios existentes, com o intuito de ajudar na tomada de decisão. É importante realçar que os *Dashboards* não dispõem de título.

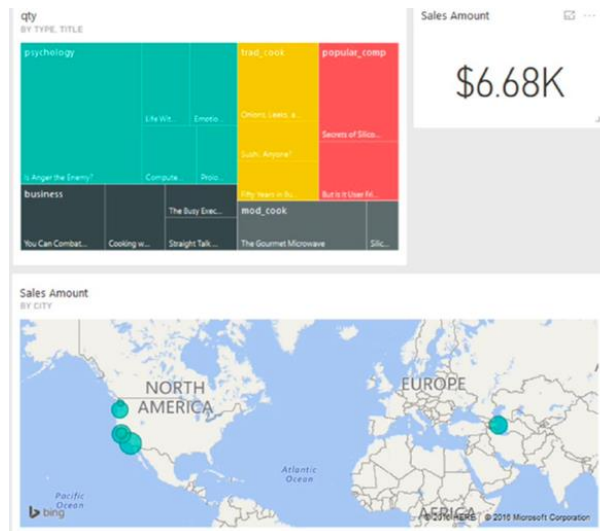


Figura 4- Exemplo de um Dashboard em Power BI
(Rad, 2018)

Contudo, *Dashboard* e *Report* são conceitos diferentes. Um *Report* é uma combinação de múltiplos elementos visuais (gráficos, textos, cartões, valores...) numa página que pode ser inter-relacionada entre si e filtrada com base em critérios.

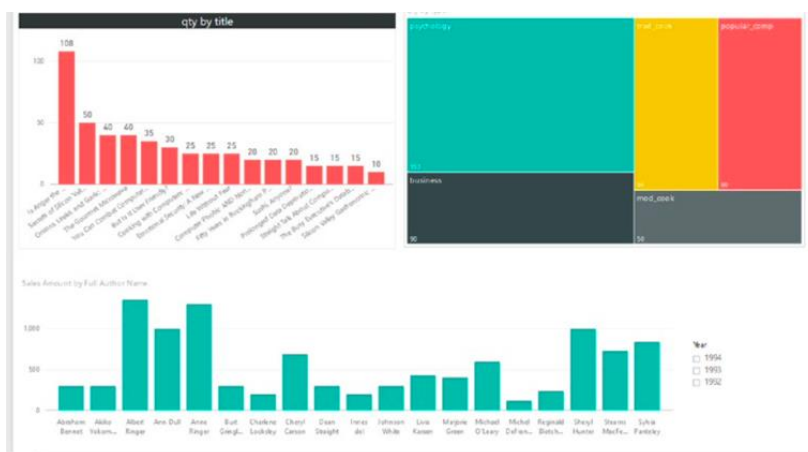


Figura 5- Exemplo de um Report em Power BI
(Rad, 2018)

Para uma melhor compreensão destes 3 conceitos, a Figura 6 define a estrutura e as diferenças entre o *Dataset*, *Dashboard* e *Report* em *Power BI*. Cada *Report* necessita de um *Dataset* e um ou vários *Reports* estão contidos num *Dashboard*.

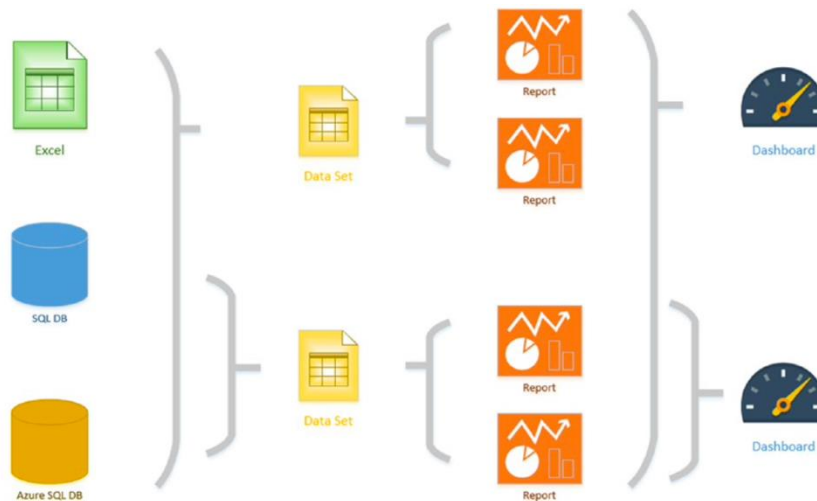


Figura 6- Estrutura do Dataset, Report e Dashboard no Power BI

(Rad, 2018)

3.1.2 Escolher os melhores Visuais no *Power BI*

O *Power BI* tem várias visualizações disponíveis por predefinição e padrão. Além disso, é possível alargar o número de visualizações a partir de *Custom Visuals* mais personalizados.

Uma visualização, ou um visual, é uma representação gráfica dos dados que permite aos utilizadores analisar aspetos dos dados. Com o *Power BI*, a representação visual dos dados é fácil e direta, tornando muito mais fácil a compreensão da informação disponível (O'Connor, 2019).

Os *Visuals* por padrão são os que constam na Figura 7.

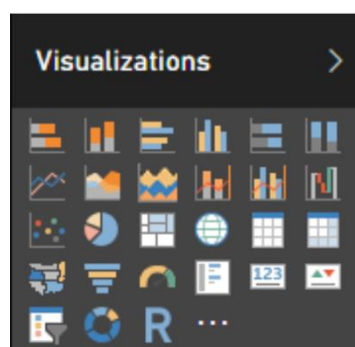


Figura 7- Tipos de Visualizações disponíveis por padrão no Power BI

(Ferrari & Russo, 2016.)

Os 27 *Visuals* disponíveis no *Power BI* são brevemente descritos pelo autor (Ferrari & Russo, 2016) a partir da seguinte lista:

- **Gráfico de barras** -este visual deve ser usado quando é necessário comparar diferentes valores da mesma medida, lado a lado, ou quando é necessário mostrar diferentes medidas que fazem parte do mesmo todo. As barras são apresentadas horizontalmente.
- **Gráfico de colunas** - este gráfico é semelhante ao Gráfico de barras, mas apresentado verticalmente.
- **Gráfico de barras agrupadas** - neste visual é possível comparar diferentes medidas lado a lado e não na mesma barra- contrariamente aos dois gráficos supracitados.
- **Gráfico de colunas agrupadas** - é semelhante ao gráfico de barras agrupadas, mas orientado verticalmente.
- **Gráfico de barras empilhado a 100%** - semelhante ao gráfico de barras, mas com cada medida a utilizar uma parte de cada barra, que corresponde sempre à largura total disponível (100%).
- **Gráfico de colunas empilhadas a 100%** - o mesmo que o gráfico de colunas empilhadas a 100%, mas apresentado verticalmente.
- **Gráfico de linhas** - este visual é normalmente utilizado para mostrar a tendência de algumas medidas ao longo do tempo.
- **Gráfico de área** - semelhante ao gráfico de linhas, mas é utilizado quando se pretende exibir dados cumulativos em vez de sequências de pontos.
- **Gráfico de área empilhada** - semelhante ao gráfico supracitado (Gráfico de área), mas cada medida é acumulada com as outras.
- **Gráfico de linhas e colunas empilhadas** - este gráfico é normalmente utilizado quando é necessário mostrar medidas com diferentes escalas (exemplo: moeda e percentagem).
- **Gráfico de linha e de colunas agrupadas** - semelhante ao gráfico de linha e de colunas empilhadas, mas utilizando colunas agrupadas em vez de colunas empilhadas.
- **Gráfico em cascata** - este visual é normalmente utilizado para exibir dados cumulativos, em que é necessário destacar para cada valor o seu valor positivo ou negativo.
- **Gráfico de dispersão** - este gráfico é normalmente utilizado para mostrar possíveis correlações entre duas medidas.

- **Pie chart/Gráfico de tarte** - é utilizado para mostrar a distribuição de valores de uma ou mais medidas. Os valores aparecem como peças da tarte, com os valores maiores a ocuparem fatias maiores.
- **Mapa em árvore** - semelhante a um *Pie Chart*, mas utiliza uma representação gráfica bastante diferente, em que os valores são representados por retângulos coloridos numa página. Pode ser uma alternativa a um gráfico de tarte, mas é quase ilegível quando se tem muitos elementos no mapa em árvore.
- **Mapa** - é utilizado para mostrar dados geográficos com formas circulares de tamanho variável em mapas.
- **Tabela** - é utilizado para exibir dados numa forma textual como uma tabela simples, onde cada atributo e cada medida é uma única coluna.
- **Matriz** - este visual estende a tabela, e faz com que seja possível agrupar as medidas por linhas e colunas.
- **Mapa preenchido** - é semelhante ao mapa, mas os dados são representados por áreas coloridas.
- **Funil** - é bastante semelhante ao gráfico de barras empilhadas, mas com uma única medida e uma representação gráfica diferente, em que as linhas são empilhadas por ordem, o que faz com que o gráfico pareça um funil.
- **Gráfico de Gauge** - é utilizado para mostrar onde se encontram os valores dentro de um intervalo/objetivo.
- **Cartão de várias filas** - este visual permite mostrar diferentes medidas e atributos para cada instância de uma entidade.
- **Cartão** - é utilizado para mostrar um único valor numérico de uma medida.
- **KPI** - este visual é utilizado para mostrar um único valor com um gráfico de linhas de tendência, em que realça o seu desempenho com cores.
- **Slicer** - é utilizado para filtrar um ou mais gráficos, em que são selecionados valores de um atributo.
- **Gráfico em Donut** - semelhante ao *Pie Chart*, mas com a visualização de um donut.
- **Visual em R** - permite gerar gráficos a partir de código em linguagem R.

Segundo o mesmo autor, apesar de termos bastantes gráficos e visuais disponíveis, não há necessidade de os utilizar todos. A presença de muitos tipos diferentes de visuais num único relatório pode torná-lo bastante confuso.

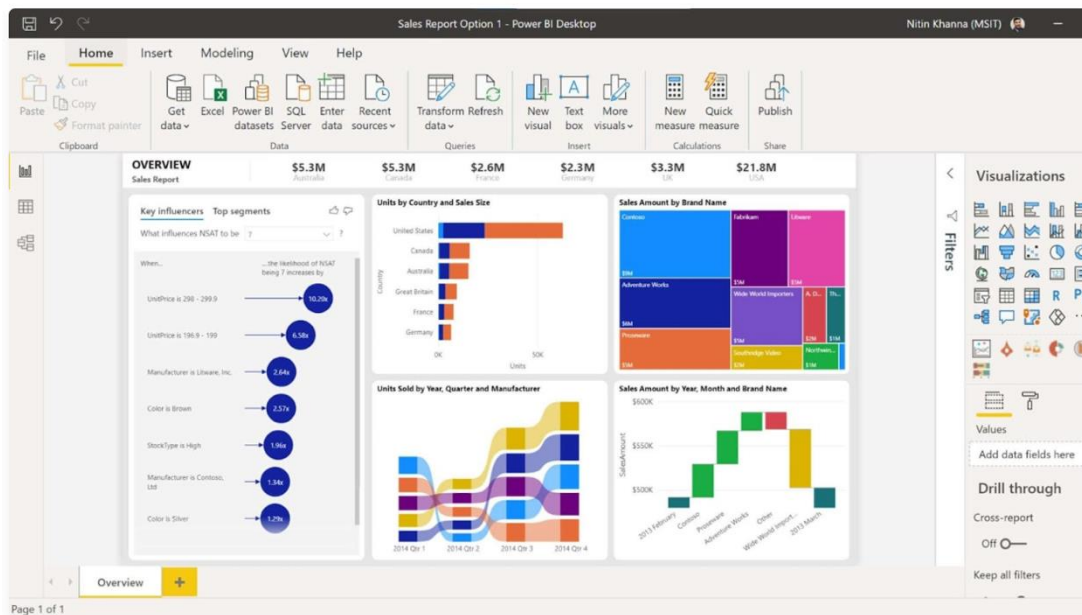


Figura 8- Visualização da ferramenta Power BI

(Duò, 2022)

O autor (Duò, 2022) define os seus prós e contras como sendo:

Prós

- Inclui muitas opções de gráficos e relatórios pré-definidos;
- Tem capacidades de Aprendizagem de Máquina;
- É possível a análise de dados a partir de várias fontes de dados tais como: *Dynamics 365, Excel, SharePoint, Salesforce, e Azure SQL DB* (entre outras fontes);
- Ideal para equipas;
- Possibilidade de criação de uma visualização de dados interativa;
- Muito seguro;
- Disponível em versão *desktop* e dispositivos móveis;

Contras

- Mais adequado para pessoas que já analisaram dados antes ou que usam o *Excel* frequentemente;
- Só pode lidar com um máximo de 2 GB de dados de cada vez (portanto, não é adequado para conjuntos de dados extensos);

3.2 Excel

O *Microsoft Excel*, de todas as ferramentas apresentadas nesta dissertação, é aquela que apresenta o programa mais antigo e, provavelmente, o mais conhecido, que permite criar tabelas, gráficos e visualizações (Duò, 2022).

Criado em 1985, é um programa de folha de cálculo e é muito conveniente para cálculos numéricos (Baier & Neuwirth, 2007).

O *Excel* é frequentemente utilizado para análise de dados devido às suas características de Visualização de Dados, que permitem a criação de gráficos. O programa vem com uma seleção significativa de gráficos incorporados, que podem ser usados para transmitir de forma elegante os dados (entre os quais Gráficos de barras, colunas, *Pie Chart*, gráfico de linhas, gráficos de área, entre outros) (Hati, 2023).

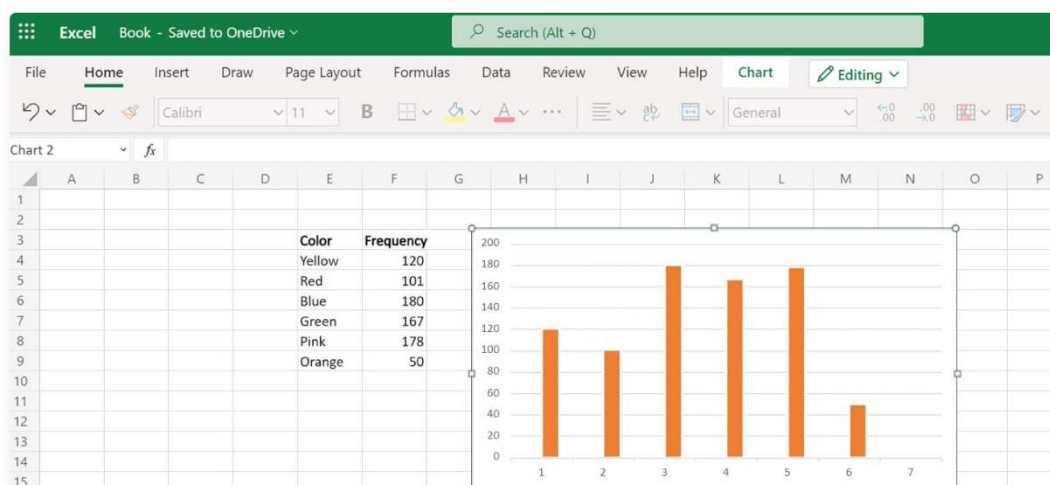


Figura 9- Visualização da ferramenta Microsoft Excel

(Duò, 2022)

O autor (Duò, 2022) define os seus prós e contras como sendo:

Prós

- Disponível em dispositivos *Windows, MacOS, Android e iOS*;
- É possível personalizar a fonte, a cor e as características da visualização;
- Permite a importação e exportação de dados;
- Inclui histórico de versões;

Contras

- Pode ser caro;
- Difícil de usar em grandes conjuntos de dados;
- Não foi projetado para equipas;

3.3 Google Charts

O *Google Charts* é uma ferramenta gratuita da Google que pode criar visualizações simples a partir de conjuntos de dados (Duò, 2022).

Com esta ferramenta, os utilizadores podem criar gráficos e incorporá-los em páginas *web*. Os gráficos são baseados em *HTML5* e tecnologia de gráficos vetoriais escaláveis e podem ser apresentados em vários navegadores e plataformas. Existem 12 tipos de gráficos: *Pie chart*, gráficos de dispersão, de *Gauge*, geográficos, tabelas, mapas de árvores, gráficos combinados, de linhas, de barras, de colunas, de área, e gráficos de castiçais.

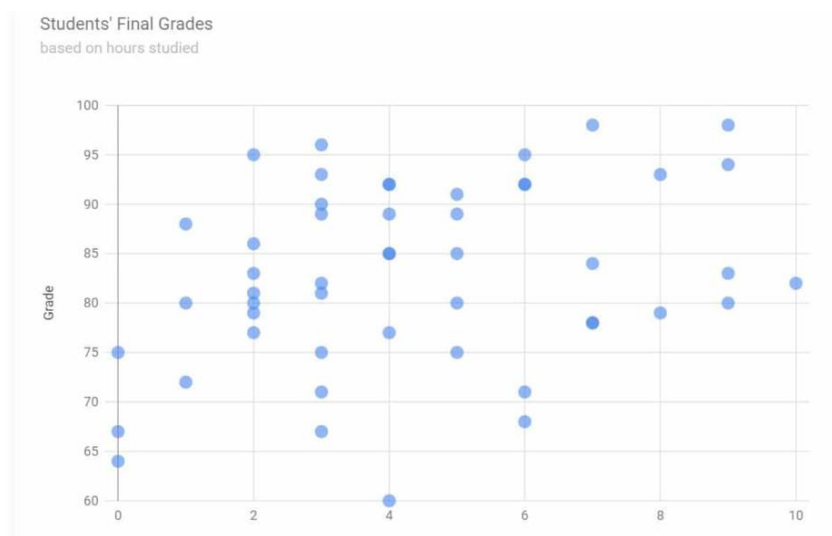


Figura 10- Visualização da ferramenta Google Charts

(Duò, 2022)

O autor (Duò, 2022) define os prós e contras do *Google Charts* como sendo:

Prós

- Possibilidade de criar vários tipos de visualizações;
- Amigável para iniciantes;
- Possibilidade de analisar dados de qualquer base de dados *SQL*;
- As visualizações podem ser personalizadas recorrendo à edição em código *CSS*;
- É totalmente gratuito;

Contras

- É necessário um conhecimento básico de *JavaScript* para publicar a visualização no site;

3.4 Tableau

Tableau é uma ferramenta poderosa de visualização de dados da *Salesforce* que permite a análise de dados de múltiplas fontes ao mesmo tempo (Duò, 2022).

É uma plataforma de análise visual que transforma a forma como os dados são utilizados para resolver problemas – fazendo com que as pessoas e organizações possam tirar o máximo partido dos seus dados. Esta plataforma analítica facilita a exploração e gestão de dados por parte das pessoas, e torna mais rápida a compreensão dos dados (Tableau, 2022).

O *Tableau* fornece uma interface de “arrastar e largar” os dados para uma exploração interativa. Utiliza a linguagem *VizQL* para descrever uma *query* analítica e disposição gráfica associada. As *queries* são causadas pelas interações dos utilizadores. Os resultados recebidos da base de dados são posteriormente processados e apresentados para obtenção de uma visualização (Sellis et al., 2011).

O autor (Duò, 2022) define os prós e contras do *Tableau* como sendo:

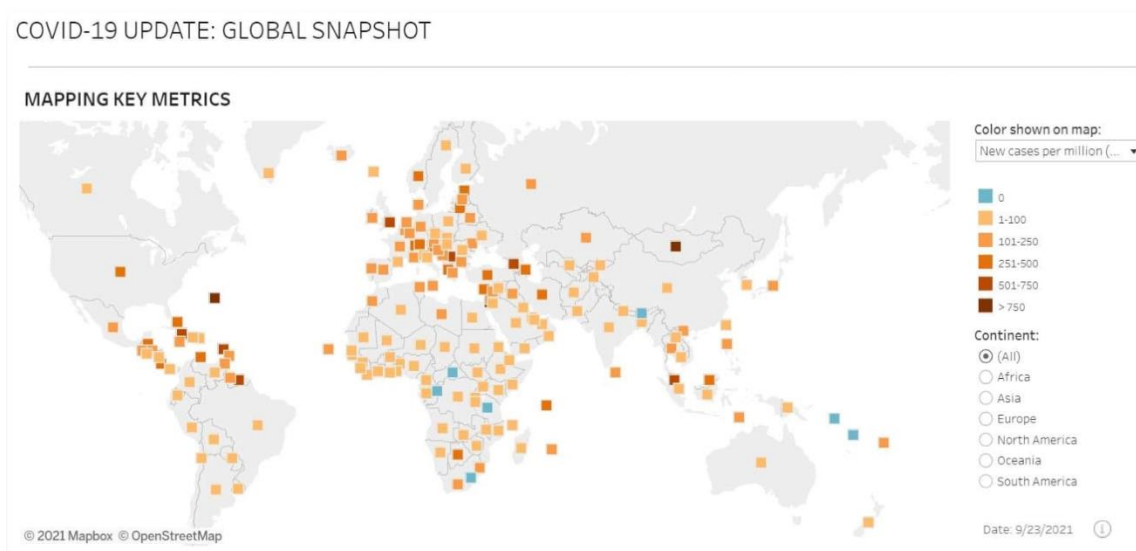


Figura 11- Visualização da ferramenta Tableau

(Duò, 2022)

Prós

- Inclui uma ampla gama de modelos de gráficos;
- Possibilidade de lidar com grandes volumes de dados;
- Pode importar dados de vários sítios;

Contras

- A aprendizagem de utilização pode ser demorada;
- O *Tableau Public* torna os dados visíveis para outros utilizadores;

3.5 Zoho Analytics

Zoho Analytics é um *software* de análise de dados que permite a análise dos dados e a criação de visualizações de dados (Zoho, 2022).

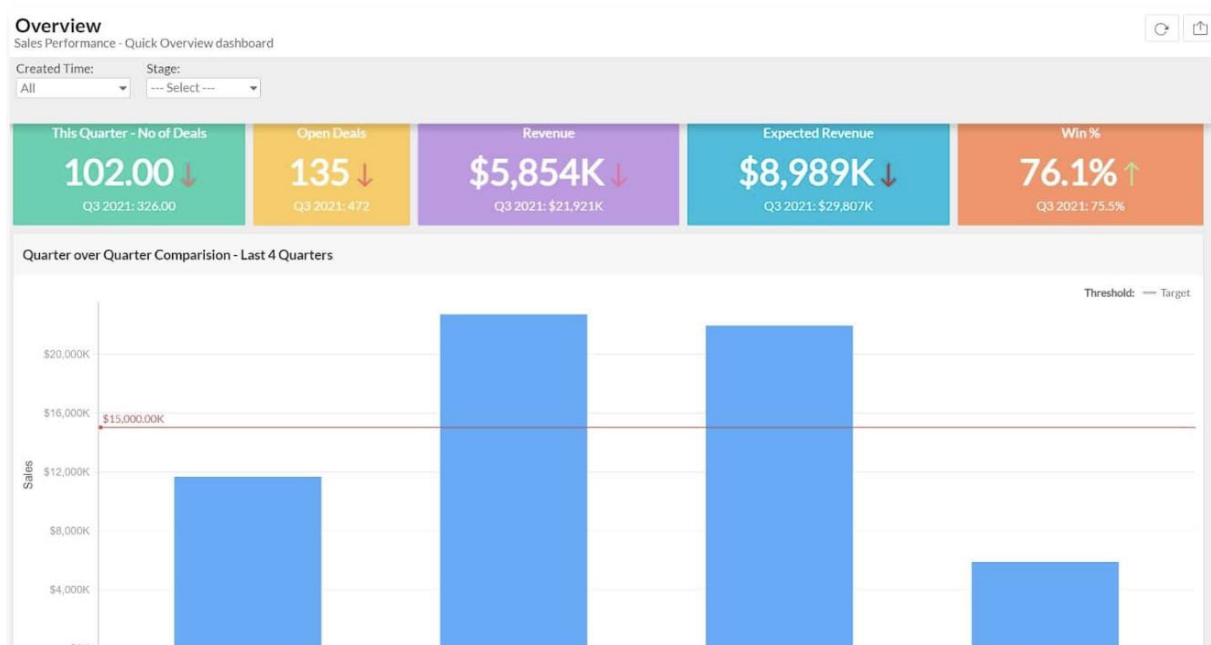


Figura 12- Visualização da ferramenta Zoho Analytics

(Duò, 2022)

O autor (Duò, 2022) define os prós e contras do *Zoho Analytics* como sendo:

Prós

- Inclui uma ampla gama de modelos pré-definidos;
- É possível incorporar visualizações do *Zoho* num site com pouco ou nenhum código;
- Integra-se em mais de 500 aplicações (incluindo *Google Ads*, *Salesforce*, e muitas outras plataformas);
- É fácil de aprender devido aos tutoriais fornecidos pela *Zoho*;

Contras

- *Zoho Analytics* é mais adequado para pessoas com uma compreensão básica de análise e análise de dados;

3.6 Datawrapper

Datawrapper é uma plataforma online gratuita que ajuda os utilizadores a criar gráficos, mapas, ou tabelas interativas facilmente. A ferramenta não requer qualquer código ou capacidades de desenho, mas fornece aos utilizadores uma interface acessível para visualizar os seus conjuntos de dados. Além disso, o *Datawrapper* não requer a criação de uma conta. Os gráficos permitem analisar conjuntos de dados complexos, revelar relações nos dados, e tornar a complexidade mais compreensível para um público mais vasto (Princeton, 2023).

Esta ferramenta tem um verificador de daltonismo integrado, o que faz com que as visualizações sejam amigáveis para pessoas com daltonismo (Duò, 2022).

Migration to the US by world region, 1820-2009

The numbers are recorded by decade. For example, the numbers recorded for 1905 tell us the number of immigrants between 1900 and 1910.

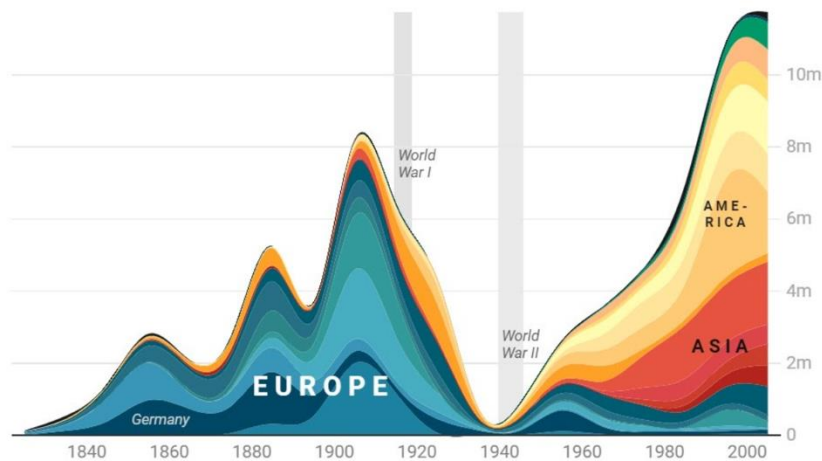


Chart: Mirko Lorenz • Source: Department of Homeland Security • Get the data • Created with Datawrapper

Figura 13- Visualização da ferramenta Datawrapper

O autor (Duò, 2022) define os prós e contras do *Datawrapper* como sendo:

Prós

- Possibilidade de criar um número ilimitado de visualizações (na conta gratuita);
- Fácil de aprender (*DataWrapper* tem vários manuais e vídeos que ensinam a trabalhar com a ferramenta);

Contras

- Só é possível importar dados de algumas fontes;
- Os planos premium do *Datawrapper* são mais caros que as ferramentas concorrentes de mercado;

3.7 Qlik Sense

Qlik Sense é uma ferramenta de Visualização de Dados desenhada para grandes empresas e para quem deseja utilizar uma análise ampliada dos dados. *Qlik Sense* é o sucessor de “*QlikView*” – uma ferramenta de análise visual similar (mas menor) (Duò, 2022).

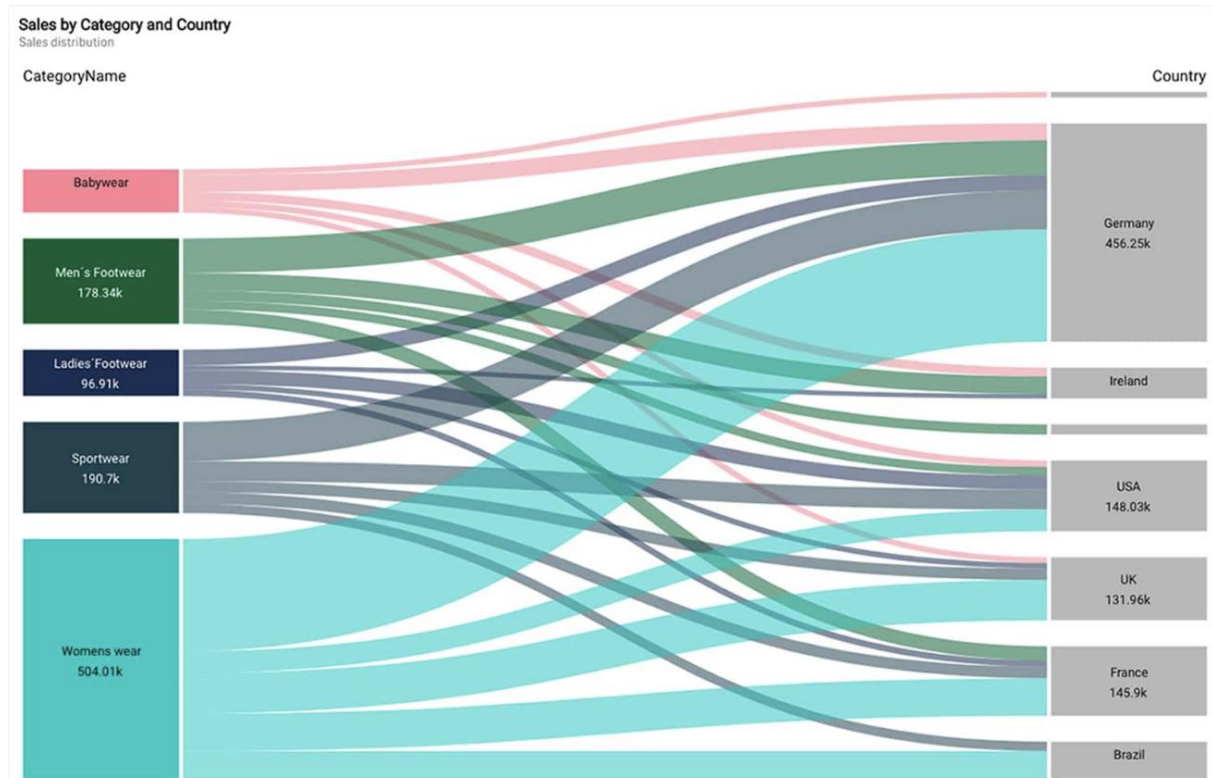


Figura 14- Visualização da ferramenta QlikSense

(Duò, 2022)

O autor (Duò, 2022) define os prós e contras do *Qlik Sense* como sendo:

Prós

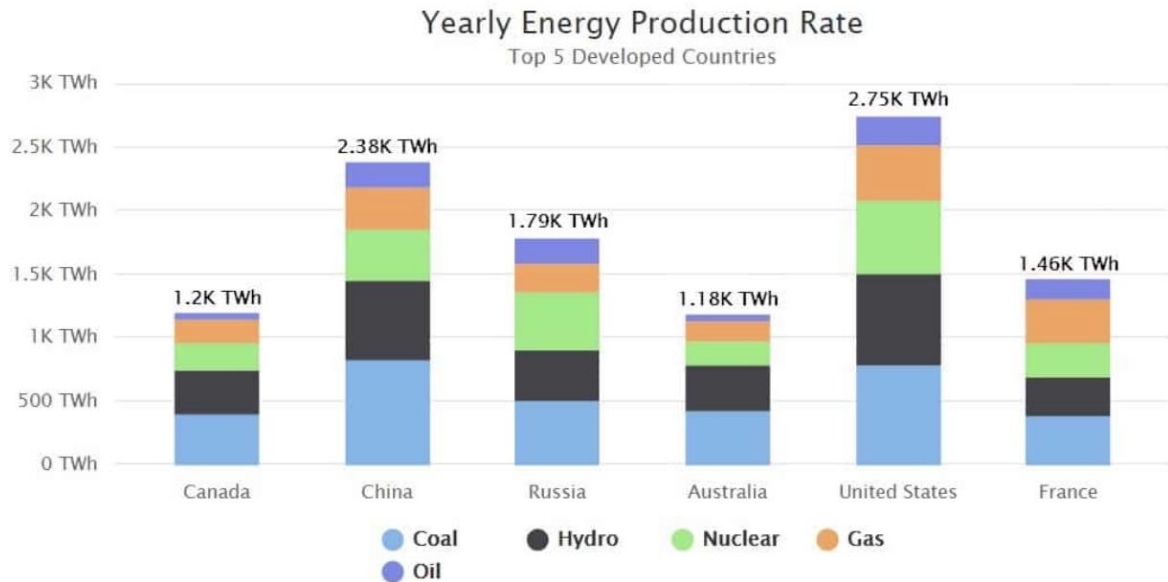
- Funciona tanto online como offline em dispositivos móveis;
- Ideal para equipas;
- Escalável para grandes empresas;

Contras

- Mais adequado para pessoas com experiência em análise de dados;

3.8 Fusion Charts

O *Fusion Charts* é uma alternativa ao *Google Charts*. Esta ferramenta permite a criação de visualizações para embutir em páginas *web* (Duò, 2022).



FusionCharts Trial

Figura 15- Visualização da ferramenta Fusion Charts

(Duò, 2022)

O autor (Duò, 2022) define os prós e contras do *Fusion Charts* como sendo:

Prós

- Oferece mais de 100 tipos diferentes de gráficos e mais de 2.000 mapas;
- É amigável para iniciantes;
- Tem uma boa interface de utilizador (IU);
- É possível exportar gráficos de forma rápida e fácil;
- Lida com grandes conjuntos de dados;

Contras

- Mais caro do que muitas outras opções de *software* de Visualização de Dados;
- Não está disponível em dispositivos *iOS* ou *Android*;

3.9 Domo

Domo é uma ferramenta de BI voltada para empresas com conjuntos de dados complicados ou grandes. Organizações como *eBay*, *ESPN*, *Cisco*, e *Emerson* utilizam a ferramenta Domo (Duò, 2022);



Figura 16- Visualização da ferramenta Domo

(Duò, 2022)

O autor (Duò, 2022) define os prós e contras do *Domo* como sendo:

Prós

- Biblioteca extensiva de modelos de relatórios e gráficos;
- *Domo* tem uma função de consulta em linguagem natural que responde a perguntas básicas sobre os dados;
- É possível analisar dados de mais de 1.000 fontes de dados;
- Oferece alertas que irão disparar quando for atingido um objetivo pré-determinado;

Contras

- Não é amigável para iniciantes;

3.10 Google Analytics

O *Google Analytics* é uma ferramenta de BI que é adequada para pequenas, médias e grandes empresas que procuram analisar a atividade de seus sites. Esta ferramenta permite rastrear números críticos como taxa de recuperação, duração média da sessão e páginas por sessão em tempo real (Duò, 2022).

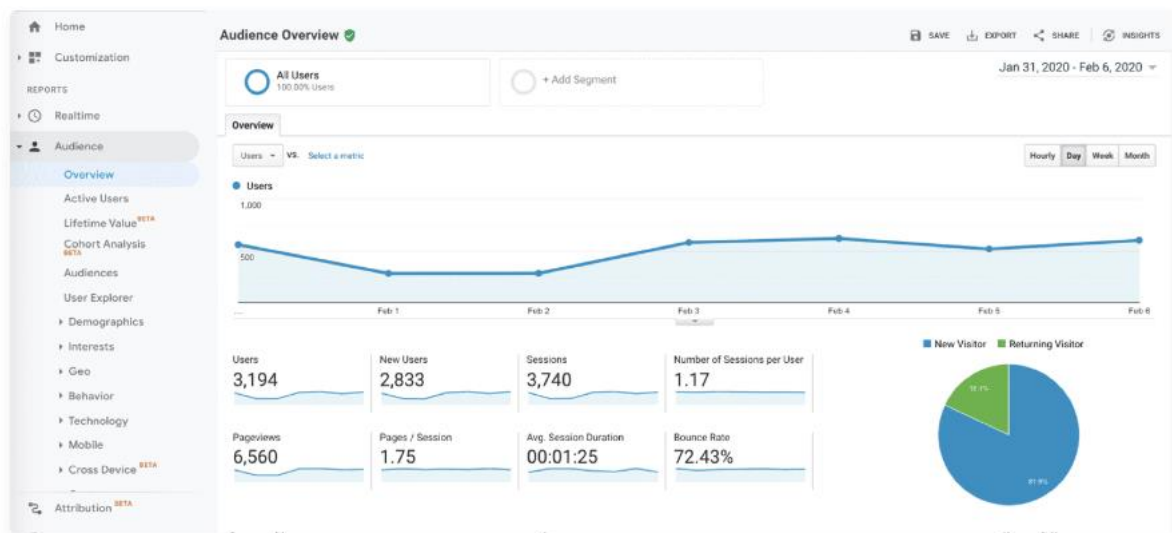


Figura 17- Visualização da ferramenta Google Analytics

(Duò, 2022)

O autor (Duò, 2022) define os prós e contras do *Google Analytics* como sendo:

Prós

- Fácil de configurar;
- *Google Analytics* recolhe e analisa dados sobre a audiência, conversões, aquisição de audiência, e comportamento automaticamente;
- É possível exportar os dados para outros programas (para uma análise posterior);

Contras

- Só é adequado para análise de dados *web*;
- Apenas oferece algumas visualizações pré-definidas;

3.11 Visme

Visme é um programa intuitivo que permite criar visualizações de dados e infográficos. Esta ferramenta é adequada para quem procura fazer gráficos simples com desenhos elegantes e esteticamente agradáveis (Duò, 2022).

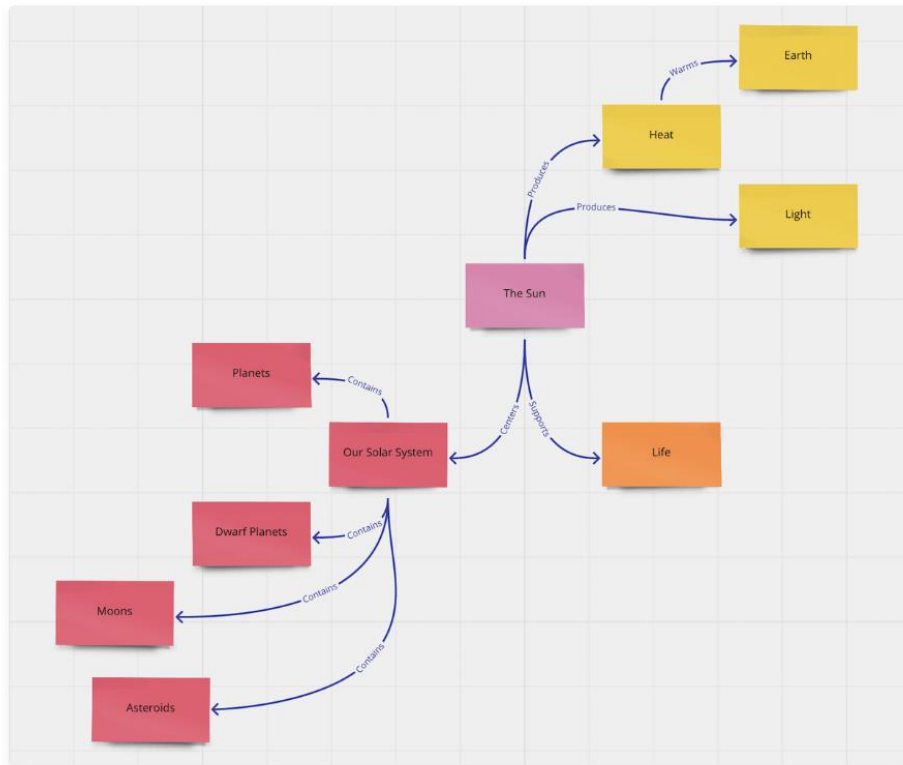


Figura 18- Visualização da ferramenta Visme

(Duò, 2022)

O autor (Duò, 2022) define os prós e contras do *Visme* como sendo:

Prós

































- Oferece muitos modelos de gráficos e infográficos;
- É possível obter os gráficos como múltiplos tipos de arquivo;
- Funciona offline e online;
- Permite a personalização das fontes, cores, fundo e design do gráfico;

Contras

- *Visme* é mais adequado para empresas com conjuntos de dados de pequeno e médio porte;
- Todos os projetos são públicos no plano básico da *Visme*;
- Não está disponível em dispositivos *Android* ou *iOS*;

Os prós e contras das ferramentas são apresentadas na Tabela 1.

Tabela 1- Ferramentas de Visualização de Dados

	Fácil aprendizagem e utilização;	Lida com grandes conjuntos de dados;	Várias opções de gráficos;	Ideal para equipas;	Barato/Gratuito;
Power BI					
Excel					
Google Charts					
Tableau					
Zoho Analytics					
Datawrapper					
Qlik Sense					
Fusion Chart					
Domo					
Google Analytics					
Visme					

* Todas as células apresentadas em branco representam ausência de informação desse parâmetro.

Após a apresentação das várias ferramentas de Visualização de Dados e de comparação das suas funcionalidades, está-se em condições de passar para o cerne deste trabalho- fatores críticos de sucesso. O próximo capítulo é dedicado a este assunto.

4. FATORES CRÍTICOS DE SUCESSO

O presente capítulo é destinado à apresentação da definição de sucesso na área do *Business Intelligence* e, mais especificamente, na Visualização de Dados. São também apresentados detalhadamente os Fatores de Sucesso.

4.1 Definição de Sucesso na área do *Business Intelligence*

O sucesso do *BI* está relacionado com o valor positivo que uma organização obtém a partir do seu investimento em *BI* (Işik et al., 2013). Contudo, o sucesso é definido de forma diferente pelas diferentes organizações e depende dos benefícios organizacionais que eram pretendidos quando foi implementado o *BI*.

O sucesso do *BI* pode representar a obtenção de benefícios tais como maior rentabilidade, diminuição dos custos, e obtenção de uma maior eficiência (Ujang Subagja et al., 2020).

Algumas organizações procuram medir o sucesso do *BI* a partir de benefícios tangíveis e utilizam medidas explícitas, tais como a melhoria do retorno do investimento (*ROI*) na eficiência operacional ou o aumento da rentabilidade da organização. Outras concluem que o seu *BI* é bem-sucedido se os custos forem razoáveis em relação aos benefícios acumulados. As medidas específicas de sucesso do *BI* diferem entre organizações e mesmo dentro de uma organização. Por exemplo, uma empresa pode implementar o *BI* para alcançar uma melhor gestão da cadeia de fornecimento, enquanto outra pode adotá-la para alcançar um melhor serviço ao cliente (Işik et al., 2013).

Portanto, o sucesso do *BI* pode ser definido como todos os benefícios positivos que as organizações obtêm através da sua utilização do *BI*.

4.3 Definição de Sucesso na Visualização de Dados

Segundo o autor (Münster, 2017), as Visualizações de Dados podem ter objetivos e funções muito diferentes, dependendo da área de aplicação. Por esse mesmo motivo, o sucesso na Visualização de Dados deve ser medido em função de diferentes critérios de qualidade, dependendo da tarefa.

O autor defende que, quando a Visualização de Dados é utilizada:

- Na análise de dados, sucesso significa que um cientista de dados pode identificar claramente as estruturas e padrões que necessita para o passo seguinte.
- Na elaboração de relatórios, o sucesso significa que um funcionário consegue ler e compreender facilmente as informações que lhe dizem respeito.
- Na comunicação, o sucesso consiste em contar uma história ou aprofundar os detalhes de uma história.

De forma geral, uma Visualização de Dados é bem-sucedida quando atinge o seu objetivo principal-transmitir a informação de forma clara e eficaz, deve também apresentar dados corretos e precisos, sem grande complexidade de interpretação e que alerte o leitor para situações chave. Além disso, o visual deve ser esteticamente agradável e atrativo.

4.4 Fatores Críticos de Sucesso na Visualização de Dados

Para alcançar o sucesso, o autor (Midway, 2020), defende que existem 10 princípios que devem ser considerados:

- **Princípio 1: Escolha do diagrama certo** - é necessário primeiro pensar na informação e na mensagem que se pretende transmitir com a Visualização de Dados.
- **Princípio 2: Utilizar o *software* correto** - visuais eficazes requerem, normalmente, um bom *software* para serem desenvolvidos. Atualmente, a maioria dos *softwares* está prontamente disponível e muitos deles gratuitos.
- **Princípio 3: Selecionar um Gráfico eficaz** - gráficos são representações dos dados em diferentes formas, e muitas vezes pode haver mais do que um gráfico a considerar consoante o tipo de dados que queremos apresentar. Por exemplo, quantidades ou comparações são frequentemente apresentadas com um gráfico de barras.
- **Princípio 4: As cores têm sempre significado** - o uso da cor na Visualização de Dados é bastante poderoso para melhor representar a informação pretendida.

A cor representa informação, quer de forma direta e óbvia, quer de forma indireta e subtil. Existem três esquemas de utilização da cor:

- a. **Sequencial**- variam desde a tonalidade da cor mais clara até à mais escura (Figura 19);

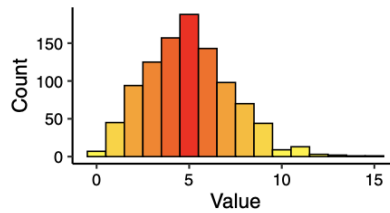


Figura 19- Esquema de cor Sequencial

(Midway, 2020)

- b. **Divergente**- são aqueles que têm dois esquemas sequenciais que representam dois extremos (Figura 20);

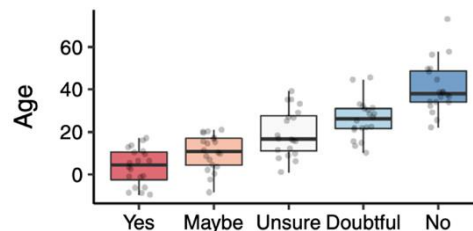


Figura 20- Esquema de cor Divergente

(Midway, 2020)

- c. **Qualitativo**- são utilizados quando a intensidade da cor não é de grande importância, mas sim o objetivo é utilizar cores diferentes para transmitir diferenças qualitativas (Figura 21);

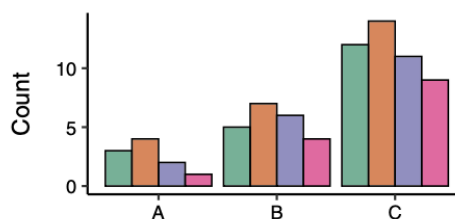


Figura 21- Esquema de cor Qualitativo

(Midway, 2020)

- **Princípio 5: Incluir a incerteza** - existem dois desafios primários com a inclusão da incerteza na Visualização de Dados: falhar ao incluir a incerteza e a má interpretação da mesma. Representar a incerteza é muitas vezes um desafio. O desvio padrão, erro padrão e os intervalos de confiança são métricas comuns de incerteza. A expressão da incerteza exige que os leitores

estejam familiarizados com as métricas de incerteza e a sua interpretação; contudo, é também da responsabilidade de quem desenvolve o visual adotar a medida de incerteza mais apropriada.

- **Princípio 6: Utilizar forma Painel, quando possível (Pequenos Múltiplos)** - uma abordagem visual particularmente eficaz consiste em repetir uma representação visual para realçar as diferenças. Esta abordagem é frequentemente chamada de pequenos múltiplos e o objetivo é tornar os dados mais facilmente comparáveis e as diferenças nos dados são mais fáceis de mostrar (Figura 22).

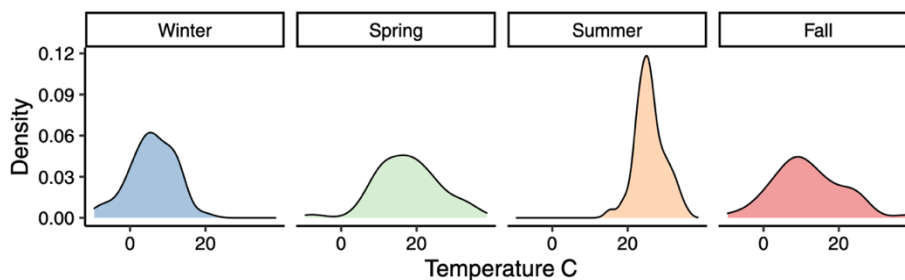


Figura 22- Abordagem visual de pequenos múltiplos

(Midway, 2020)

- **Princípio 7: Dados e modelos são coisas diferentes** - os dados brutos são muitas vezes relativamente simples, contudo, um modelo gráfico pode exigir mais explicações para quem o analisa. Certamente qualquer modelo de um estudo deve ser relatado de uma forma completa que assegure a reprodutibilidade. Contudo, qualquer modelo gráfico deve ser explicado na legenda da figura ou referenciado noutra local do documento para que o leitor possa encontrar os detalhes completos sobre o que o modelo visual está a representar.
- **Princípio 8: Visuais Simples, Legendas Detalhadas** - é bastante importante ter legendas detalhadas que expliquem ao detalhe a visualização e as representações utilizadas. As legendas devem ser autónomas, o que significa que se a figura e a legenda fossem analisadas independentemente do resto do estudo, os pontos principais ainda poderiam ser compreendidos.
- **Princípio 9: Considerar um Infográfico** - as figuras visuais e gráficos tendem a concentrar-se em dados e modelos representativos propriamente ditos, enquanto os infográficos incorporam texto e imagens. Um infográfico pode ser mais eficaz e mais facilmente memorável.
- **Princípio 10: Obter uma opinião** - embora possa haver princípios e teorias sobre a visualização eficaz de dados, a realidade é que os visuais mais eficazes são aqueles que se salientam e ficam na memória dos leitores. Por isso, muitas vezes, é necessário ter uma revisão

externa objetiva e obter *feedback* útil sobre o que os leitores percebem e absorvem dessa Visualização de Dados e, portanto, perceber se é eficaz ou não em termos visuais.

Estes 10 princípios podem ser divididos em 3 etapas: desenho da representação visual (*Figure design*), executar a representação propriamente dita (*Figure making*) e a revisão da representação visual final (*Figure review*) e são representadas na Figura 23.

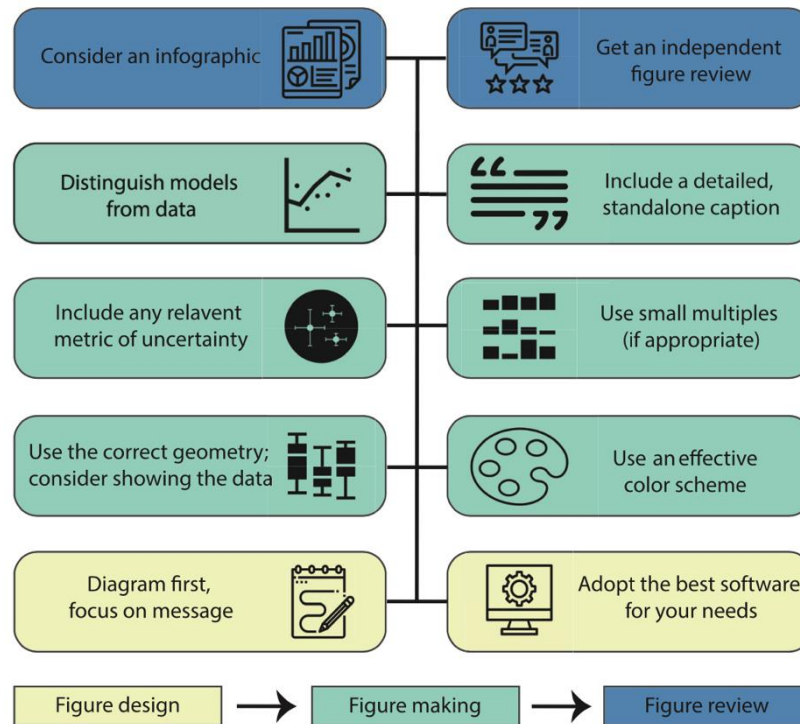


Figura 23- 10 princípios da Visualização de Dados

(Midway, 2020)

Os dois princípios a amarelo são os que ocorrem primeiro, durante a fase de desenho da representação visual. Os seis princípios a verde são considerados durante a realização da representação visual. Os dois princípios a azul são passos finais considerados depois da *DV* ter sido elaborada. Embora o fluxo dos princípios seja geralmente adotado de baixo para cima, não há ordem específica ou exigida.

Complementarmente, o autor (Metwalli, 2020), defende que o seguinte fator também contribui para uma visualização de sucesso:

- **Simplicidade nas visualizações** - ter uma visualização complexa e com bastante informação é algo que deve ser evitado.

O autor refere que sempre que se está a criar uma Visualização de Dados, é necessário prestar atenção ao *Data-Ink Ratio*. *Data-Ink Ratio* é um termo utilizado para se referir à quantidade de

dados vs. “tinta” que foi necessária para criar o gráfico (efeitos/cores de fundo e representação 3D dos dados). É possível observar a fórmula do *Data-Ink Ratio* na Figura 24.

$$\text{Data-Ink Ratio} = \frac{\text{Data-Ink}}{\text{Total "ink" used to print the graphic}}$$

= Proportion of a graphic's "ink" devoted to the non-redundant display of data-information

= 1.0 – proportion of a graphic that can be erased without loss of data-information

Figura 24- Fórmula do *Data-Ink Ratio*

(Blaxell, 2019)

Para que a DV seja simples e eficaz, o *Data-Ink Ratio* tem de ser elevado. Quando o contrário acontece, é sinal de que a visualização está a apresentar demasiada informação visual tornando-se, assim, bastante confuso visualmente.

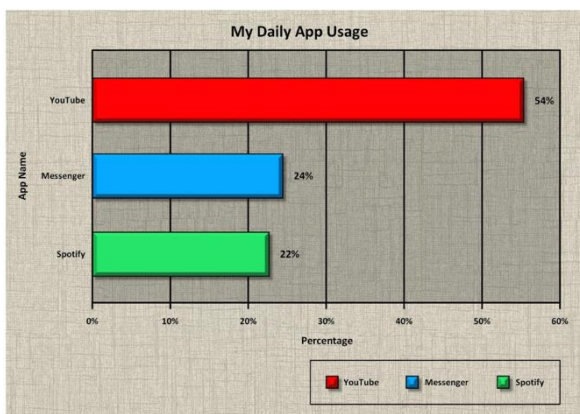


Figura 25- Exemplo de um *Data-Ink* baixo

(Blaxell, 2019)

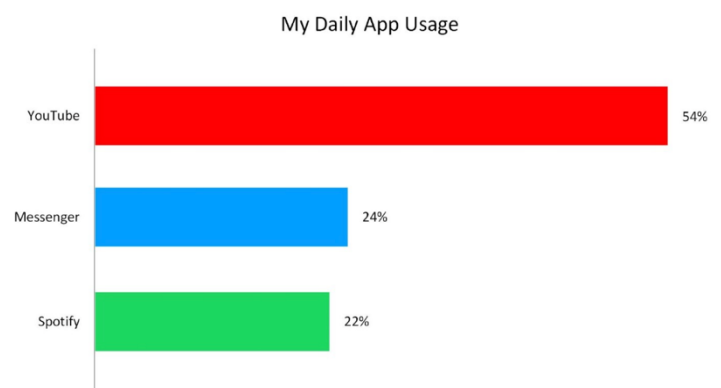


Figura 26- Exemplo de um *Data-Ink* elevado

(Blaxell, 2019)

Após terem sido detalhados os vários fatores de sucesso, a Tabela 2 faz a compilação dos mesmos.

Tabela 2- Compilação dos Fatores Críticos de Sucesso

Objetivo da Visualização de Dados bem definido
Escolha de uma ferramenta adequada para a Visualização de Dados
Selecionar o melhor gráfico consoante o tipo de dados a apresentar
Utilização correta das cores
Inclusão de incerteza
Utilização de Pequenos Múltiplos
Legendas e texto detalhado
Apresentação dos dados de forma clara e correta
Simplicidade na Visualização de Dados (Data-Ink Elevado)

Está presente no Anexo I uma tabela de ajuda para determinar se a Visualização de Dados foi bem desenvolvida de acordo com os Fatores Críticos de Sucesso.

Além destes fatores mencionados anteriormente, para que o projeto de Visualização de Dados seja bem conseguido, há inevitavelmente muitos outros fatores que terão uma influência na implementação do mesmo. É especialmente importante identificar e reconhecer o impacto das condições contextuais, dentro e à volta do projeto que irão afetar o seu desenvolvimento (Kirk, 2012):

- **O objetivo** - existem diferentes maneiras que podem originar um projeto. Um cliente pode não definir detalhadamente o que pretende. Nestas situações, a responsabilidade passa por apresentar um trabalho que vá ao encontro do pedido e da satisfação do cliente, mesmo que os parâmetros não sejam muito especificados. Noutras situações, pode ser pedido algo que vá contra a prática do autor (por exemplo o assunto em questão ou o estilo solicitado) ou até pode ser simplesmente impossível de entregar (devido aos recursos disponíveis ou à complexidade do pedido).

- **Pressão de tempo** - a maioria dos projetos tem prazos de entrega claros e o objetivo é planear o trabalho e utilizar o tempo de forma eficaz para evitar tempo desnecessariamente em tarefas sem relevância ou com menor importância.

- **Custos** - os custos influenciam bastante o tempo que se pode dedicar a um projeto, a possibilidade de incluir mais colaboradores e as ferramentas e os recursos que se pode utilizar.

- **Pressão dos clientes** - para além da pressão de tempo, é necessário antecipar e reduzir o impacto de potenciais pressões originadas pelo cliente. O cliente pode simplesmente mudar os requisitos anteriormente dados ou pedir novas exigências. O objetivo é manter o cliente sempre atualizado do

progresso do projeto e, quando possível, envolvê-lo em momentos de decisão ao longo do processo para que sejam feitas pequenas alterações ao longo do projeto em vez de grandes alterações quando o projeto já estiver finalizado.

- **Formato** - a representação visual pode ser estática ou interativa e interfere bastante no tipo de conceção, devendo ser definida o mais cedo possível. Se for uma representação visual interativa, é necessário escolher a ferramenta que melhor se adequa para cada tipo de dispositivo (telemóvel, sites, computador). Caso seja estático, é necessário definir se vai ser um pequeno gráfico numa publicação ou, por exemplo, *flyers*.

- **Capacidades técnicas** - é necessário analisar quais as ferramentas que estão disponíveis e, além disso, é importante definir se apenas existe acesso a ferramentas gratuitas ou se é possível aceder a *softwares* premium com mais opções e funcionalidades.

4.5 Fatores Críticos de Sucesso no *Business Intelligence*

Os fatores críticos de Sucesso, do Inglês CSF (*Critical Success Factors*), são vistos como um conjunto de tarefas e procedimentos que devem ser abordados e cumpridos para assegurar as realizações dos sistemas *BI* durante a sua implementação (Olszak & Ziemba, 2012).

A teoria dos CSF fornece uma boa base para identificar os critérios que devem ser seguidos durante a implementação das aplicações *BI* (Olszak, 2016).

Uma compreensão correta dos CSF permite aos participantes de um projeto de *BI* que otimizem os seus recursos e esforços, concentrando-se em fatores significativos que apresentam maior probabilidade de ajudar a alcançar o sucesso na implementação do projeto (Yeoh & Koronios, 2010).

Segundo (Williams et al., 2022), os CSF dividem-se em 3 perspetivas:

1. Perspetiva Técnica;
2. Perspetiva Organizacional;
3. Perspetiva do Processo;

A Tabela 3 ilustra as categorias que estão inseridas dentro de cada uma dessas 3 perspetivas.

Tabela 3- Fatores Críticos de Sucesso do Business Intelligence

(Williams et al., 2022)

Perspetiva	Critério	Fatores
1. Perspetiva Técnica	<p>A. Qualidade do sistema</p> <p>B. Qualidade da informação</p> <p>C. Satisfação do utilizador</p>	<p>A.1. Flexibilidade e escalabilidade</p> <p>A.2. Integrações</p> <p>A.3. Acessibilidade</p> <p>B.1. Fiabilidade e integridade dos dados</p> <p>B.2. Gestão de dados eficaz</p> <p>C.1. Facilidade de utilização</p> <p>C.2. Vantagem relativa</p>
2. Perspetiva Organizacional	<p>A. Apoio à gestão</p> <p>B. Cultura organizacional</p> <p>C. Disponibilidade de recursos adequados</p>	<p>A.1. Patrocínio executivo</p> <p>A.2. Apoio dos principais intervenientes</p> <p>B.1. Uma visão estratégica bem definida</p> <p>B.2. Cultura de aprendizagem organizacional</p> <p>B.3. Cultura empresarial aberta e analítica</p> <p>C.1. Recursos de apoio funcional e de sistema</p> <p>C.2. Competências de equipa</p> <p>C.3. Centro de competência <i>BI</i></p>

D. Perspetiva do Processo	A. Gestão de projetos	A.1. Abordagem de gestão de projetos A.2. Processos de gestão de projetos
	B. Interação do utilizador	B.2. Participação e envolvimento dos utilizadores B.3. Desenvolvimento iterativo impulsionado pelo utilizador final B.4. Requisitos impulsionados pelo negócio
	C. Gestão da mudança	C.1. Gestão da mudança orientada para o utilizador C.2. Formação e capacitação dos utilizadores C.3. Resistência do utilizador

Segundo (García & Pinzón, 2017) existem diferentes critérios-chave para assegurar o sucesso da solução BI:

1. **Gestão e Direção de topo:** Um elevado nível de apoio da gestão de topo está associado a um elevado nível de sucesso do projeto de BI.
2. **Planeamento e visão de negócio** - Um sistema de BI que não é orientado para os negócios, é um sistema falhado. Um modelo empresarial sólido deve incorporar todas as propostas estratégicas que o projeto irá abordar, recursos de trabalho necessários, possíveis riscos, custos a assumir e prazos para executar o projeto. Muitas das vezes, a principal causa do fracasso das soluções BI não é por parte tecnológica, mas devido a um mau alinhamento com o negócio, a sua visão e os seus objetivos.
3. **Eleição de um líder do projeto de BI** - Esta pessoa deve ter conhecimentos técnicos e funcionais suficientes e, ao mesmo tempo, deve ter excelentes capacidades interpessoais para resolver conflitos organizacionais.
4. **Estratégia** - Uma estratégia a longo prazo resulta numa melhoria contínua a nível organizacional. O projeto deve ter também uma visão do BI, e fornecer os recursos necessários

para a sua realização. É igualmente importante e fundamental ter, de forma clara, os objetivos que se pretende alcançar.

5. **Mudança de gestão** – Ter um programa de mudança de gestão é importante uma vez que ajuda a reduzir a resistência de novas implementações e há uma maior disposição a aceitar mudanças.
6. **Projeto** – O projeto BI é um dos fatores mais importantes. O principal objetivo da sua formulação, prazos, avanços e resultados pode ser alcançado, enquadrado e alinhado com os propósitos comerciais estrategicamente afixados numa fase anterior pela gestão de topo.
7. **Recursos humanos qualificados** - Os membros do projeto devem ter experiência, conhecimentos e competências adequadas. As competências das pessoas envolvidas no projeto influenciam o sucesso da sua implementação.
8. **Aprendizagem e competências** – É importante que as organizações forneçam aos trabalhadores as competências e conhecimentos apropriados para uma melhor utilização das soluções do BI. Tendo em conta que este tipo de projeto exige um elevado empenho técnico, deve dispor de uma equipa que dê apoio e formação mais precisa.
9. **Informação e tecnologia** – Este fator crítico de sucesso está centrado na arquitetura, desenvolvimento de *software* e ferramentas e é um dos mais utilizados na investigação *BI*.
10. **Recursos** – De maneira geral, este fator é visto de uma perspetiva económica, principalmente para os gestores que avaliam este tipo de projeto consoante as relações de custo-benefício.
11. **Métricas** - As métricas são sempre importantes para conhecer os resultados dos projetos. O sucesso de uma solução de BI pode ser influenciado pelos benefícios, utilidade, qualidade e satisfação por parte dos utilizadores.
12. **Ambiente** - As organizações precisam de sistemas de informação que possam melhorar essa situação e dar-lhes uma melhor compreensão sobre as forças ambientais com as quais podem melhorar o seu desempenho.

Estes 12 critérios-chave definidos por (García & Pinzón, 2017) podem ser visualizados na Figura 27.



*Figura 27- Fatores Chave de Sucesso do Business
(García & Pinzón, 2017)*

Tendo sido apresentados os fatores críticos de sucesso numa VD, no sentido de melhor os concretizar, serão apresentados no capítulo seguinte exemplos de más práticas de Visualização de Dados.

5. MÁS PRÁTICAS DE VISUALIZAÇÃO DE DADOS

Este capítulo é destinado à apresentação e descrição de más práticas de Visualização de Dados.

5.1 Exemplos de Más Práticas de Visualização de Dados

Como referido e explicado anteriormente, a Visualização de Dados tem um papel essencial nos dias de hoje e facilita bastante a análise dos dados. Contudo, quando estes dados não são bem apresentados, podem ter o efeito oposto do pretendido ou até mesmo transmitir uma certa desinformação.

Segundo o autor (Nguyen et al., 2021), o tipo de desinformação pode ser dividido em duas categorias:

1. **Desinformação intencional** - utilização de gráficos para desvirtuar, ocultar ou modificar os dados numa tentativa de enganar os utilizadores;
2. **Desinformação não intencional** - utilização de informações não fidedignas, falsas e não exatas devido à falta de cuidado na conceção/seleção dos dados e, conseqüentemente, fornecer informações erradas ao utilizador;

O autor (Stobierski, 2021), apresenta 5 exemplos comuns de más visualizações de dados:

1. **Utilizar o tipo de gráfico errado** - Existem vários tipos de gráficos para representar os dados visualmente. Contudo, a escolha do gráfico pode não ser a mais apropriada para o objetivo que se pretende ilustrar. A natureza dos dados determina o formato da visualização e a característica mais importante a ter em atenção é se os dados são qualitativos (descreve ou categoriza) ou quantitativos (mensuráveis). Os dados qualitativos tendem a ter uma melhor aplicação em gráficos de barras e *Pie Charts*, enquanto os dados quantitativos são melhor representados em formato de gráficos e histogramas.
2. **Incluir demasiadas variáveis** - O objetivo de criar uma visualização é criar uma história a partir dos dados. Por esse mesmo motivo, o objetivo é filtrar e incluir a informação mais relevante possível e excluir informação irrelevante e detalhes desnecessários. Antes de desenvolver a visualização propriamente dita, é necessário delinear qual o objetivo dessa Visualização de Dados e quais as variáveis com maior importância e relevância. Assim, a transmissão correta dos dados é garantida e a audiência presta atenção à parte pretendida e importante dos dados. A Figura 28 demonstra o uso excessivo de variáveis no gráfico *Pie Chart*.

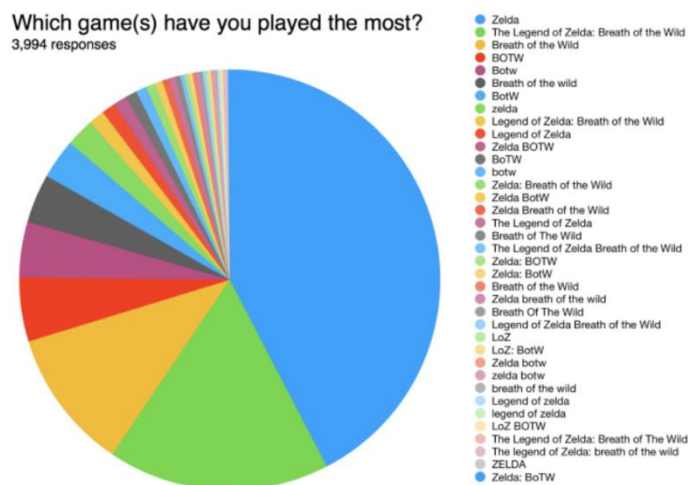


Figura 28- Exemplo do uso excessivo de variáveis num *Pie Chart*

(Violet, 2021)

Como é possível observar pela Figura 28, um *Pie Chart* que compara muitas variáveis, para além de ser difícil observar as diferenças entre os valores devido à complexidade do gráfico, também distrai e dispersa quem analisa do ponto de vista do que se pretende transmitir. Além disso, como são consideradas muitas variáveis, as cores escolhidas acabam por se tornar semelhantes e confusas.

3. **Utilizar escalas inconsistentes ou manipulação do eixo** - Caso o objetivo do gráfico seja a demonstração de diferentes pontos de vista, a escala deve ser consistente. Se a escala escolhida for inconsistente, pode causar confusão a quem analisa os dados. Por exemplo, na Figura 29, deveria ter sido utilizada a mesma imagem com o mesmo tamanho em ambas as colunas para que a representação visual fosse consistente. Só desta forma é possível observar e comparar o caso A ao caso B. No segundo exemplo foi utilizada uma imagem que, apesar de ser notório que o caso B apresenta valores maiores, não é claro nem perceptível quais são os valores exatos do caso B. Este exemplo descreve uma situação em que foi utilizada uma escala inconsistente e que pode levar a uma maior dificuldade de compreensão de quem analisa.

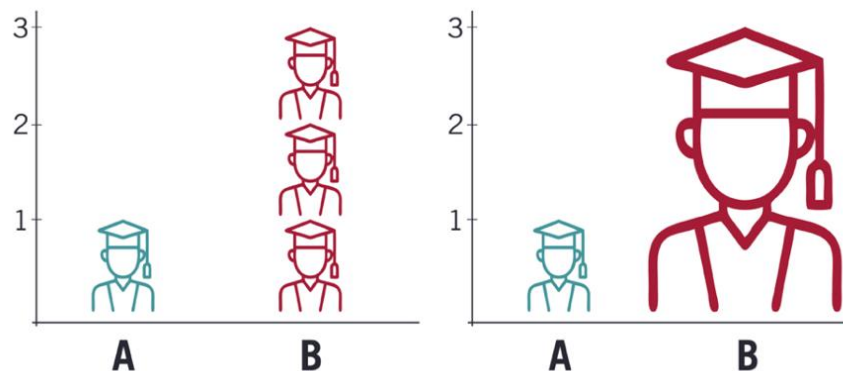


Figura 29- Utilização de escalas inconsistentes

(Stobierski, 2021)

No que toca ao eixo dos gráficos, outra situação que pode acontecer e fazer com leve à incompreensão e confusão do leitor é a manipulação do eixo dos y. O valor onde o eixo dos y começa pode alterar bastante a compreensão de um gráfico. O eixo deve começar sempre no valor 0 para que seja consistente. Por exemplo, na Figura 30, estão apresentados os mesmos valores nos dois gráficos, contudo, como a escala do primeiro gráfico começa no valor 10 no

eixo dos y, dá uma perceção que há uma maior diferença de valores do que no segundo, que efetivamente está bem representado e com o começo do eixo correto. Este exemplo mostra que, ao começar o eixo dos y com um valor que não o zero, pode levar o leitor a ter uma interpretação mais exagerada dos valores.

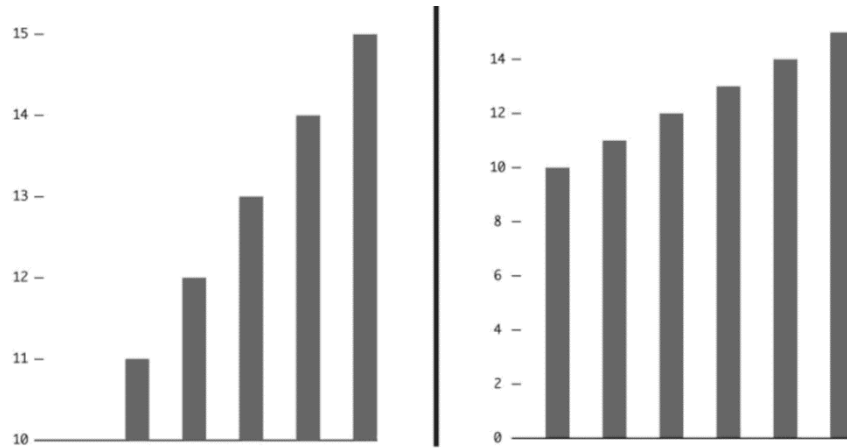


Figura 30- Utilização de valores errados no eixo dos Y

(Berry, 2021)

Recorrendo a um exemplo real, os gráficos da Figura 31 consistem numa votação para apurar a opinião da população com a saída do Reino Unido da União Europeia. Como é possível observar pelo primeiro gráfico, a utilização de um valor diferente do 0 no eixo dos y (neste caso 37%) faz com que a perceção da votação seja muito mais dispersa do que realmente é. No segundo gráfico, utilizando o valor correto no eixo dos y, mostra que efetivamente 43% e 39% são valores bastante próximos e que não apresentam tanta disparidade de valores como é percecionado no primeiro. Este é um exemplo real de como a utilização de valores errados no eixo dos y consegue alterar a perceção de um leitor quando analisa um gráfico.

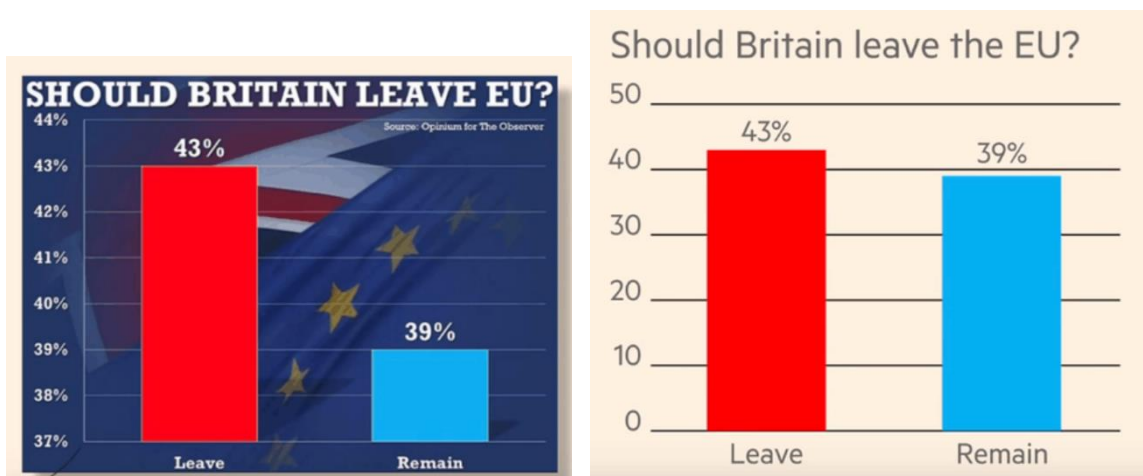


Figura 31- Utilização errada do eixo dos y - Exemplo real

(Berry, 2021)

4. Escala Linear ou Escala Logarítmica

A forma mais fácil de interpretar a diferença entre uma escala linear e uma escala logarítmica é a partir da observação dos eixos de cada uma delas. Quando um gráfico é construído numa escala linear, o valor entre quaisquer dois pontos ao longo de qualquer um dos eixos é sempre igual e imutável. Quando um gráfico é construído numa escala logarítmica, o valor entre quaisquer dois pontos ao longo de qualquer um dos eixos muda de acordo com um padrão em particular.

Embora a escala logarítmica possa ser um meio eficaz de comunicação de dados, o gráfico deve clarificar que é essa a escala que está a ser utilizada. Quando não estiver claro, o visualizador pode assumir que está a analisar um gráfico com uma escala linear (o que é mais comum) e gerar alguma confusão.

Na Figura 32, podemos observar dois gráficos que transmitem os mesmos dados. A diferença entre eles é a escala utilizada. No gráfico da esquerda é utilizada uma escala linear e no da direita uma escala logarítmica. Não é uma má prática utilizar escalas logarítmicas, simplesmente devem estar bem identificadas e clarificadas para quem analisa.

5. Escolha errada de cores

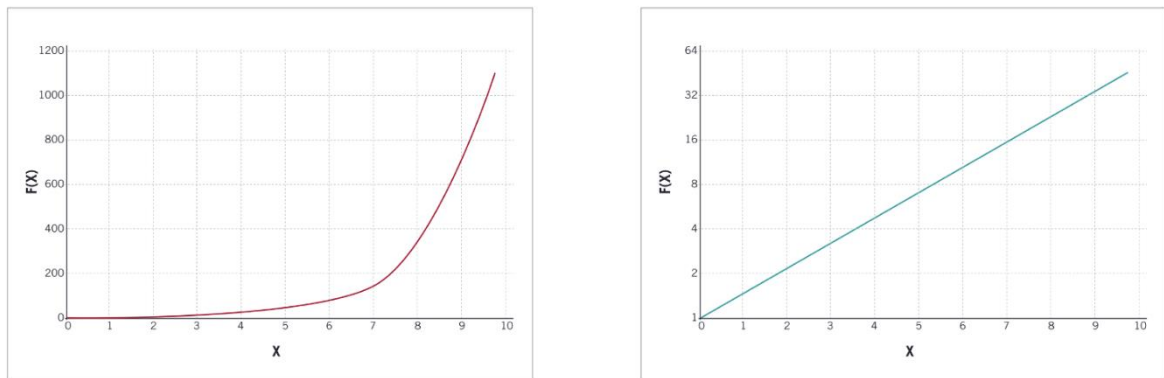


Figura 32- Escala Linear vs Escala Logarítmica

(Stobierski, 2021)

A utilização de cor facilita a compreensão dos dados quando são usadas e escolhidas com cuidado e corretamente. Contudo, quando usadas incorretamente, podem causar bastante confusão e até incompreensão do gráfico. É bastante importante perceber a informação que se pretende transmitir a partir da DV e fazer a escolha correta de cores.

Existem problemas que surgem quando se incorpora as cores erradas nas visualizações:

- Utilizar demasiadas cores, dificultando a compreensão;
- Utilizar cores habituais (exemplo: vermelho e verde) contra a utilização normal- utilizar a cor verde associada ao decréscimo e o vermelho ao aumento pode causar uma certa confusão quando essas cores, por norma, estão associadas ao aumento e ao decréscimo, respetivamente;
- Utilizar cores com pouco contraste;

Na Figura 33 está apresentado um mapa dos Estados Unidos da América (EUA) que mostra as taxas de infeção de uma doença de cada estado. Normalmente, os mapas cartográficos utilizam diferentes tonalidades dentro da mesma família de cores. Quanto mais clara a tonalidade, menos casos nesse estado; quanto mais escura a tonalidade, mais casos existem. Se for contra esta suposição, como acontece neste exemplo, a utilização de uma cor mais escura para indicar menos casos pode confundir o observador.

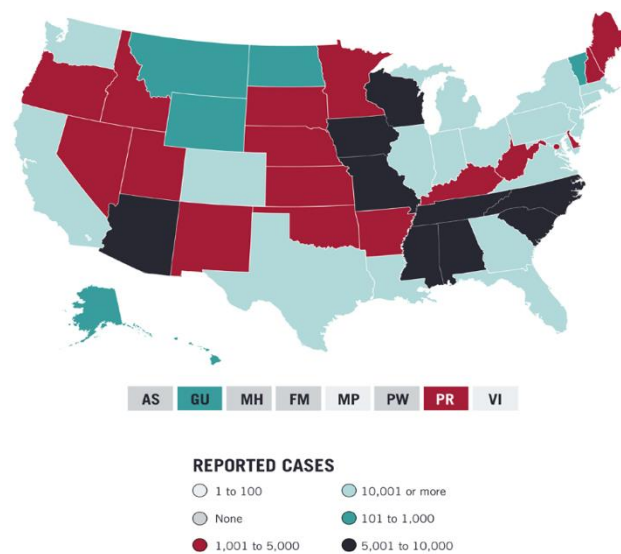


Figura 33- Exemplo 1 de Escolha Errada de Cores

(Stobierski, 2021)

O caso seguinte, figura 34, representa outro exemplo da má utilização de cores. Como se pode observar na figura, ambas as entidades foram representadas por diferentes tonalidades de vermelho e, como estão incorporadas na mesma barra, torna-se difícil diferenciá-las.

MOST WICKETS IN DEATH OVERTS IN ODIS

SINCE THE START OF JANUARY 2017

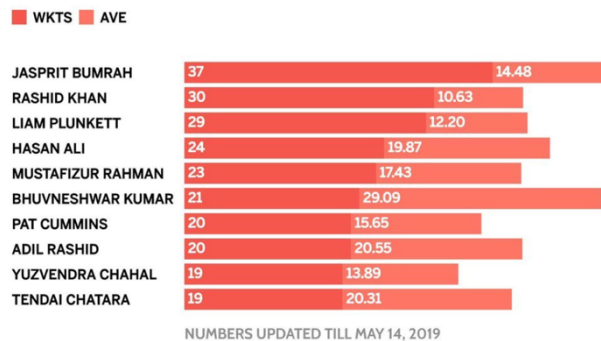


Figura 34- Exemplo 2 de Escolha Errada de Cores

(Violet, 2021)

5.2 Exemplos de Visualização de Dados Mal Desenvolvidas

Além dos 5 exemplos de más práticas de visualizações de dados apresentados anteriormente, existem exemplos de visualizações que estão mal desenvolvidas. No caso seguinte, Figura 35, a percentagem de Joe Biden, apesar de apresentar uma percentagem de 54% encontra-se atrás de líderes com 32% (Emmanuel Macron). Este é um exemplo de uma manipulação intencional dos dados.

Approval rate of some world leaders

April 2021



Figura 35- Exemplo de uma DV com manipulação dos dados

(Berry, 2021)

Outro exemplo que demonstra um mau desenvolvimento é a Figura 36. Podemos observar que a menor fatia apresenta uma maior percentagem (56%) e a maior fatia a menor percentagem (42%). Isto faz com que, apesar dos valores estarem apresentados explicitamente no gráfico, haja uma confusão por parte de quem analisa e, por vezes, pode ficar com a perceção contrária dos valores devido à dimensão das fatias. Para além disso, um segundo erro que se pode encontrar neste visual é o facto de a soma de ambas as fatias dar um total de 98% (e não de 100% como expectável).

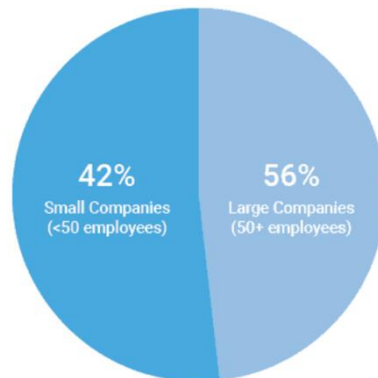


Figura 36- Pie Chart com má representação dos dados

(Berry, 2021)

O exemplo seguinte, Figura 37, revela um gráfico de barras que não apresenta qualquer consistência no que toca ao tamanho das barras apresentadas. Apesar de estar ordenado cronologicamente, o tamanho das barras não está consistente e de acordo com os valores apresentados- uma barra com 2,926 casos é maior do que uma barra com 4,049 casos.

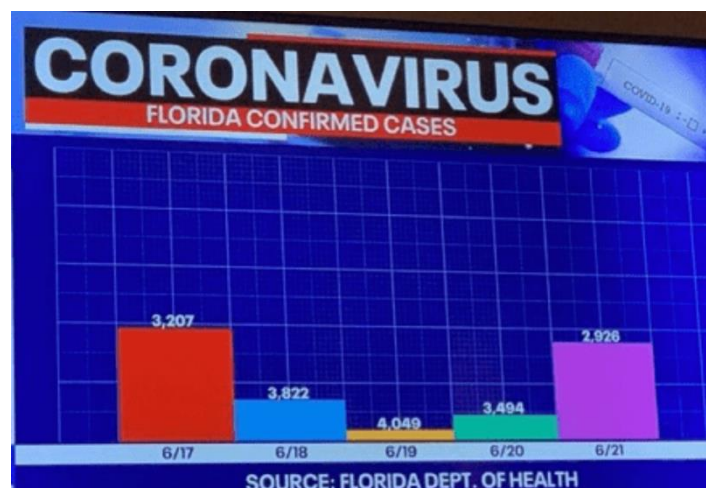


Figura 37- Gráfico de barras com má representação dos dados

(Berry, 2021)

É necessário ter bastante atenção a como apresentar dados geográficos. Na Figura 38, o autor tentou mostrar um valor linear ao longo do tempo, mas usando um mapa. Esta adoção torna a comparação e análise bastante difícil. Além disso, a relação entre o tamanho atribuído a cada cor no mapa e a percentagem associada não é coerente quando é apresentado um valor percentual. Adicionalmente, não é explícito se as etnias que não aparecem no mapa equivalem a 0%.

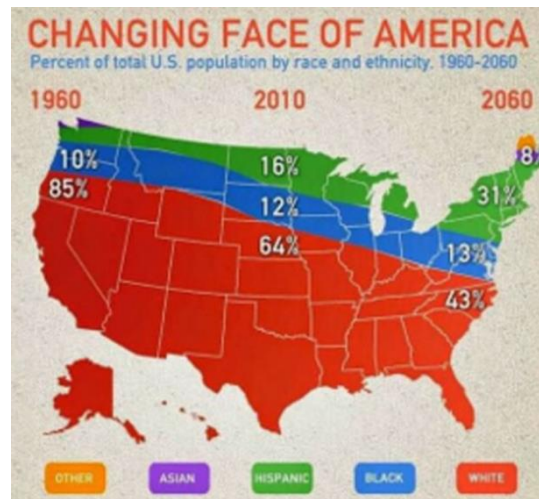


Figura 38- Mapa com má representação dos dados

(Berry, 2021)

5.3 O impacto de uma Má Visualização dos Dados

Depois de apresentados alguns exemplos comuns de más práticas que se tornam confusas e até mesmo enganadoras, é necessário perceber qual o impacto de trabalhar com informações pouco éticas ou fazer más escolhas na demonstração dos dados.

Um gráfico mal desenvolvido pode parecer um pequeno erro da organização, contudo, pode ter repercussões graves. Visualizações de dados erradas levam a uma compreensão errada dos dados apresentados. Este mal-entendido pode levar a percepções erradas e, conseqüentemente, tomada de más decisões comerciais – (sob o pretexto de que são apoiadas por dados). Levado até ao extremo, um quadro ou gráfico mal desenvolvido pode levar a questões legais ou regulamentares. Por exemplo, uma visualização enganadora dos dados incluídos num relatório financeiro pode levar os investidores a comprar ou vender ações de empresas (Stobierski, 2021).

O autor (Tinted, 2015) explicou o impacto dividindo em 3 más práticas:

1. **Narrações fragmentadas** - como mencionado anteriormente, uma narrativa informativa ou de infográficos é um dos maiores componentes de uma estratégia de BI de sucesso. Um dos

maiores erros que as pessoas cometem quando trabalham com informação visual é trabalhar com uma figura, ou imagens que apenas apresentam fragmentos de um visual inteiro. Quando isto acontece, faltam os antecedentes, um certo contexto e as principais percepções e descobertas. Para além de tornar os dados menos persuasivos, quebra-se uma narrativa em partes ilógicas. Quando isto acontece, é possível que a organização sofra um retorno do investimento (ROI) fraco, criando simultaneamente ineficiência em toda a organização.

2. **Má tomada de decisões** - quer seja escala errada, más escolhas de gráficos, ou má conceção de DV, trabalhar com métricas enganosas resulta muitas vezes em más decisões. Quando os conhecimentos obtidos a partir da visualização não vão ao encontro da verdade, podem ser tomadas más decisões:

- Reduzir o tempo necessário para o projeto e orçamento;
- Conduzir campanhas de vendas e marketing ineficazes;
- Colocar uma sobrecarga nos processos internos;
- Acelerar o crescimento organizacional;

3. **Ineficiência operacional e financeira** - Trabalhar com conhecimentos imprecisos ou errados pode resultar numa má eficiência operacional e ou financeira.

Sem percepções visuais concretas que transmitam exatamente o que está a dar bons resultados ou não, é improvável que se chegue ao centro das questões que desperdiçam os orçamentos ou que diminuem a produtividade. Como resultado, os processos operacionais e fiscais não irão atingir os objetivos da organização.

Para contrapor os exemplos de más práticas apresentados neste capítulo, serão reportados no capítulo que se segue, 3 exemplos de boas práticas de Visualização de Dados.

6. REPORTS EM *POWER BI* TENDO EM CONSIDERAÇÃO OS FATORES DE SUCESSO

Neste capítulo serão apresentados três exemplos de *Reports* desenvolvidos em *Power BI*.

Estes *Reports* foram desenvolvidos tendo em consideração todos os fatores e boas práticas de desenvolvimento de Visualização de Dados apresentadas em capítulos anteriores para uma exemplificação prática das mesmas.

6.1 *Dataset*

Para o desenvolvimento destes *Reports* em *Power BI*, foram selecionados três *Datasets* do *site Kaggle*. O *Kaggle* permite aos utilizadores encontrar conjuntos de dados que pretendam utilizar na construção de modelos de Inteligência Artificial, publicar conjuntos de dados, trabalhar com outros cientistas de dados e engenheiros de aprendizagem automática, e participar em competições para resolver desafios de ciência de dados (Hassan, 2022).

Um excerto de cada um destes *Datasets* está apresentado nos Anexos (A.1, A.2 e A.3).

Sendo assim, os *Datasets* selecionados para o desenvolvimento de *Reports* em *Power BI* foram:

- **Filmes Originais da *Netflix* e respetivas classificações *IMDB (Internet Movie Database)*** - Este conjunto de dados consiste em todos os filmes originais da *Netflix* lançados até 2021. Além disso, também inclui todos os documentários e especiais da *Netflix*. Os dados foram *integrados* com todas as pontuações *IMDB* correspondentes. O *IMDB* é uma base de dados com informação sobre cinema, música, jogos e televisão. As pontuações do *IMDB* são votadas pelos membros da comunidade, e a maioria dos filmes tem mais de 1.000 críticas (Corter, 2021);
- **Vendas do Supermercado *Giant*** – Conjunto de dados que contém informações relacionadas com as vendas, os lucros e outros aspetos interessantes do Supermercado *Giant* de 2014 a 2017 (Chowdhury, 2022);
- **Top 50 músicas *Spotify*** – Conjunto de dados que contém as 50 músicas mais ouvidas no mundo pelo *Spotify* em 2019. Este conjunto de dados contém várias variáveis sobre as músicas e os respetivo músicos (Henrique, 2019);

6.2 Report Filmes Originais Netflix

O objetivo com o desenvolvimento deste *Report*, é mostrar as informações mais relevantes presentes no *Dataset*. Sendo assim, primeiramente, estão apresentados quatro cartões com variáveis relevantes (número total de filmes originais da *Netflix*, número de línguas disponíveis nos filmes, número total de gêneros dos filmes e tempo médio de duração dos filmes). Com estas variáveis é possível ter uma noção da dimensão do *Dataset* e do *Report* que estamos a analisar.

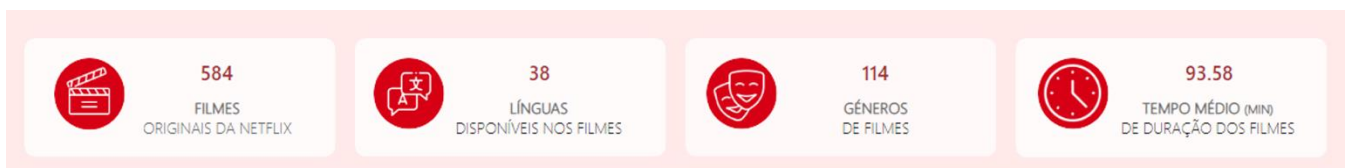


Figura 39- Cartões com variáveis relevantes do Report Netflix

Logo de seguida é apresentado o “*Top 10 filmes com melhor classificação*” e o “*Top 10 gêneros de filmes*”. O pretendido com esta tabela e gráfico, respetivamente, é salientar os melhores filmes e gêneros de filmes para quem analisa.

É importante referir que, para uma maior facilidade de análise, além da classificação *IMDB* dos filmes, foi adicionada a coluna “Gênero” à tabela para servir de ligação ao gráfico de barras. Sendo, assim, é possível analisar que, por exemplo, o gênero de filme original da *Netflix* com melhor classificação corresponde ao gênero de filme mais comum (159 filmes).

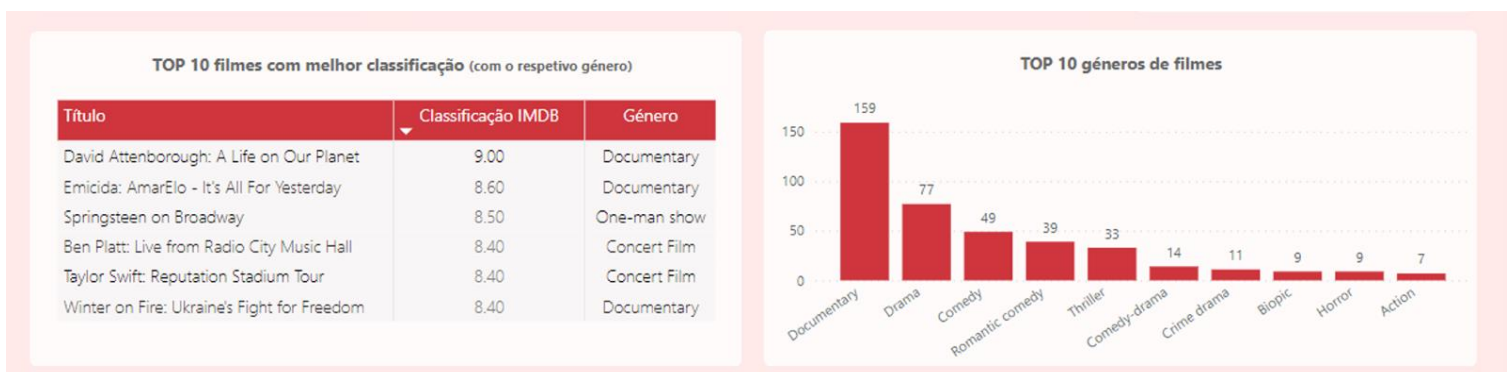


Figura 40- Top 10 filmes com melhor classificação e Top 10 gêneros de filmes

Posteriormente, e fazendo ligação com a tabela anterior foram apresentados com maior detalhe as características principais (duração, classificação, gênero e data de estreia) do filme com melhor classificação e do filme mais antigo. Com isto, é possível ter uma maior noção da janela temporal dos filmes originais da *Netflix*.

O resultado do *Report* é o seguinte:

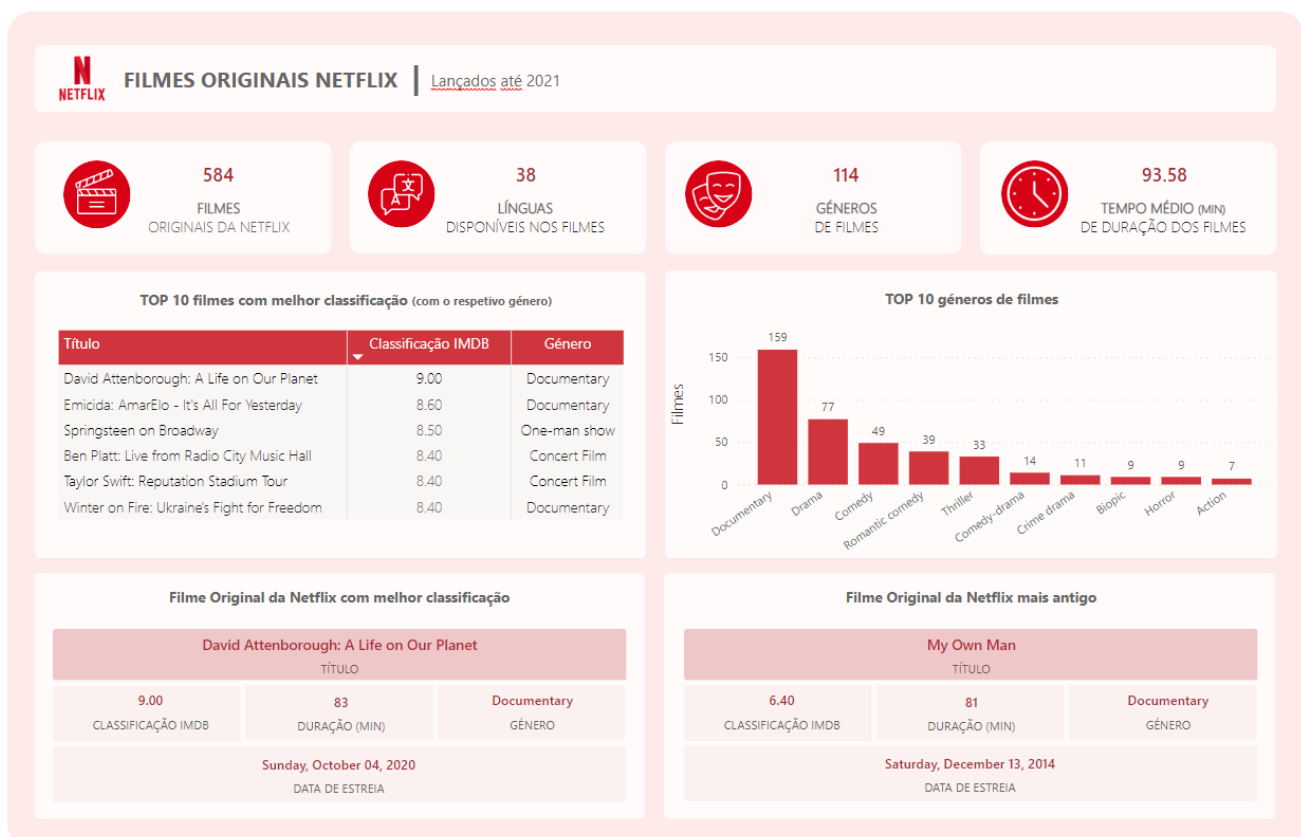


Figura 41- Report filmes Originais Netflix

Neste caso, para atingir o sucesso do *Report*, como mencionado em capítulos anteriores, era necessário que as informações presentes no mesmo fossem claras e de fácil compreensão. Como é possível observar pela Figura 41, houve um certo cuidado em apresentar toda a informação e gráficos de forma legível, sem muita complexidade e sem que fosse necessário muito esforço para compreender os dados que estão a ser transmitidos. A apresentação dos dados de forma clara, legível e simples é um dos aspetos chave a ter em consideração no desenvolvimento de *Reports*.

Além disso, foram escolhidos os *visuals* do *Power BI* que mais faziam sentido para a apresentação dos dados:

- Nas métricas da parte superior e inferior do *Report*, foram utilizados os cartões para apresentar os valores das métricas;
- Foi utilizada uma tabela para a apresentação do *Top 10* filmes com melhor classificação. Foi feita a escolha da tabela para que, ao apresentar o título do filme, a informação da classificação e o género não se tornassem muito confusos e difíceis de interpretar. Como era bastante informação a ser transmitida, a tabela foi a melhor opção para esse efeito;
- No *Top 10* géneros de filmes foi utilizado um gráfico de barras para se ter uma perceção visual da diferença entre o género com mais filmes daquele que está em décimo lugar (150 e 7, respetivamente). Como se trata de quantidade de filmes com o objetivo de comparar valores, o gráfico de barras foi a melhor escolha;

Para o desenvolvimento deste *Report* foi selecionada uma paleta de cores do logotipo da empresa Netflix. O objetivo com a utilização de cores da mesma tonalidade, era de demonstrar que, apesar da cor fazer toda a diferença para quem analisa, não é necessário a utilização de muitas e variadas cores para apresentar distintivamente e claramente a informação.

Além disso, foi considerado e tido em conta o *Report* como um infográfico, em que a junção de *visuals* e texto se pode tornar bem-sucedida.

Para finalizar, o *Report* foi desenvolvido tendo em conta um *Data-Ink elevado* não foram utilizados efeitos nem gráficos 3D e foi adotada a estratégia da seleção de uma única cor (simplificando o relatório).

6.3 Report de vendas do Supermercado *Giant*

O *Dataset* relativo ao Supermercado *Giant* é bastante mais complexo que o *Dataset* que foi usado no *Report* anteriormente apresentado. Como tal, foi necessário ter em consideração que seria um *Report* que iria ter mais informação a ser apresentada.

Primeiramente, foi apresentado o título do *Report* “Vendas do Supermercado *Giant*” e utilizado um ícone que se adequasse ao mesmo. Além disso, como o relatório em questão é relativo às vendas efetuadas, o primeiro *visual* a ser apresentado foi o de número de vendas por zona dos Estados Unidos da América (EUA) (o *Dataset* apenas refere vendas dos EUA- por essa mesma razão, a moeda utilizada no *Report* foi o *Dollar* (\$)). Assim, logo de início é possível ter a perceção de quais foram as zonas dos EUA em que foram efetuadas mais vendas nesse supermercado. Como tal, foi utilizado um gráfico de barras para ter uma melhor comparação entre as zonas (Oeste apresentou o número máximo das vendas e Sul apresentou os menores valores de vendas). Adicionalmente, para uma análise temporal das vendas, foram apresentadas as 10 melhores vendas por data. Para apresentar estes dados, foi selecionado um gráfico de linhas para uma melhor perceção das vendas a nível temporal. Assim, é possível ver em que alturas foram realizadas as melhores vendas.

Para também haver informação sobre a forma como as vendas foram entregues, foi adicionado um gráfico de barras para demonstrar o modo de envio utilizado para as vendas do supermercado. Em que é possível analisar que o “*Standard class*” foi o modo de envio mais utilizado.



Figura 42- Visuals do Report de Vendas do supermercado *Giant*

Foram utilizados cartões para demonstrar o nome dos *Top 3* compradores e as vendas totais de cada-ordenados decrescentemente. Para adicionar informação sobre as vendas dos *Top 3* compradores, foi adicionado um gráfico de linhas para cada um dos clientes para demonstrar o histórico de compras entre 2014 e 2017 e, assim, conseguir analisar as compras dos mesmos ao longo do tempo.

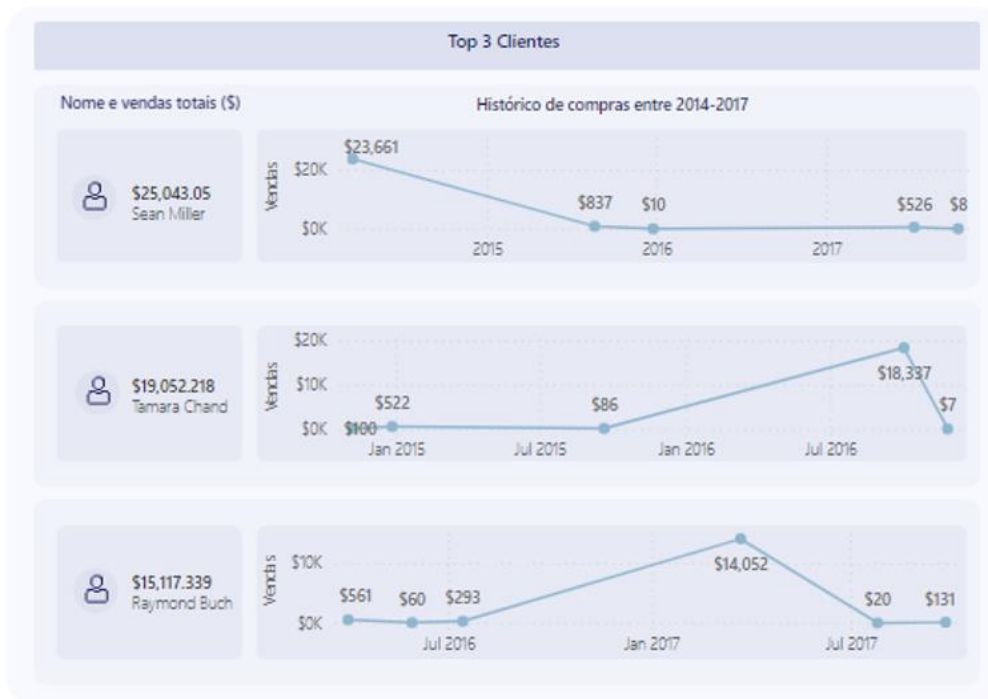


Figura 43- Visuals com o Top 3 compradores do Supermercado Giant

Para uma percepção dos valores de lucro do supermercado, foram adicionados vários cartões com o lucro total de cada ano (2014, 2015, 2016, 2017) e outro com o valor do lucro total (de 2014 a 2017). Além disso, também foi adicionado um cartão com o total de vendas entre 2014 e 2017 (Figura 44).

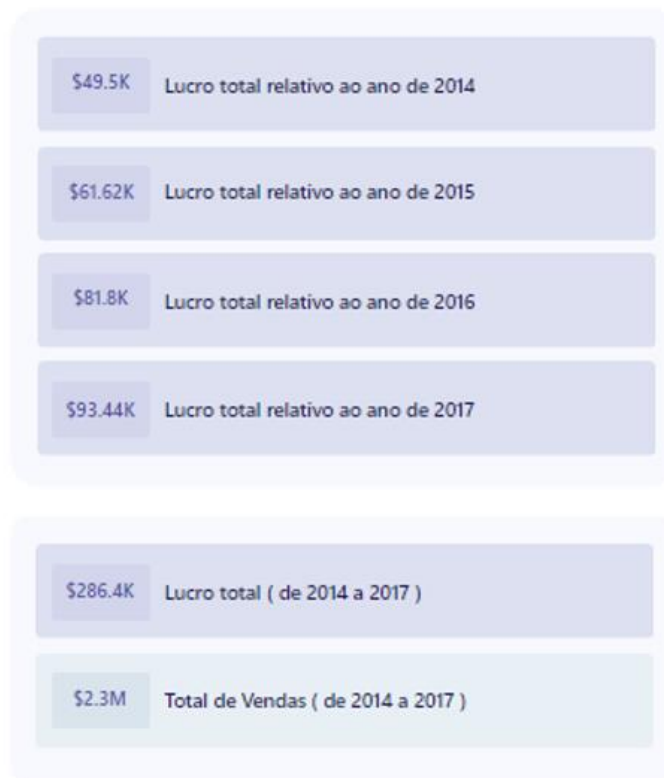


Figura 44- Visuals relativos ao lucro e total de vendas do Supermercado Giant

Para finalizar, foi utilizado um gráfico de barras para apresentar o Top 10 produtos mais vendidos e o valor de vendas dos mesmos.

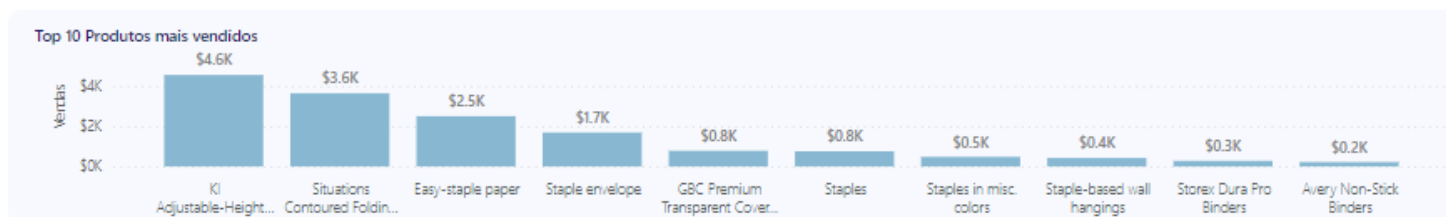


Figura 45- Top 10 Produtos mais vendidos do Supermercado Giant

Assim, após explicação de cada secção do *Report*, o resultado do *Report* das vendas do Supermercado *Giant* é o seguinte:

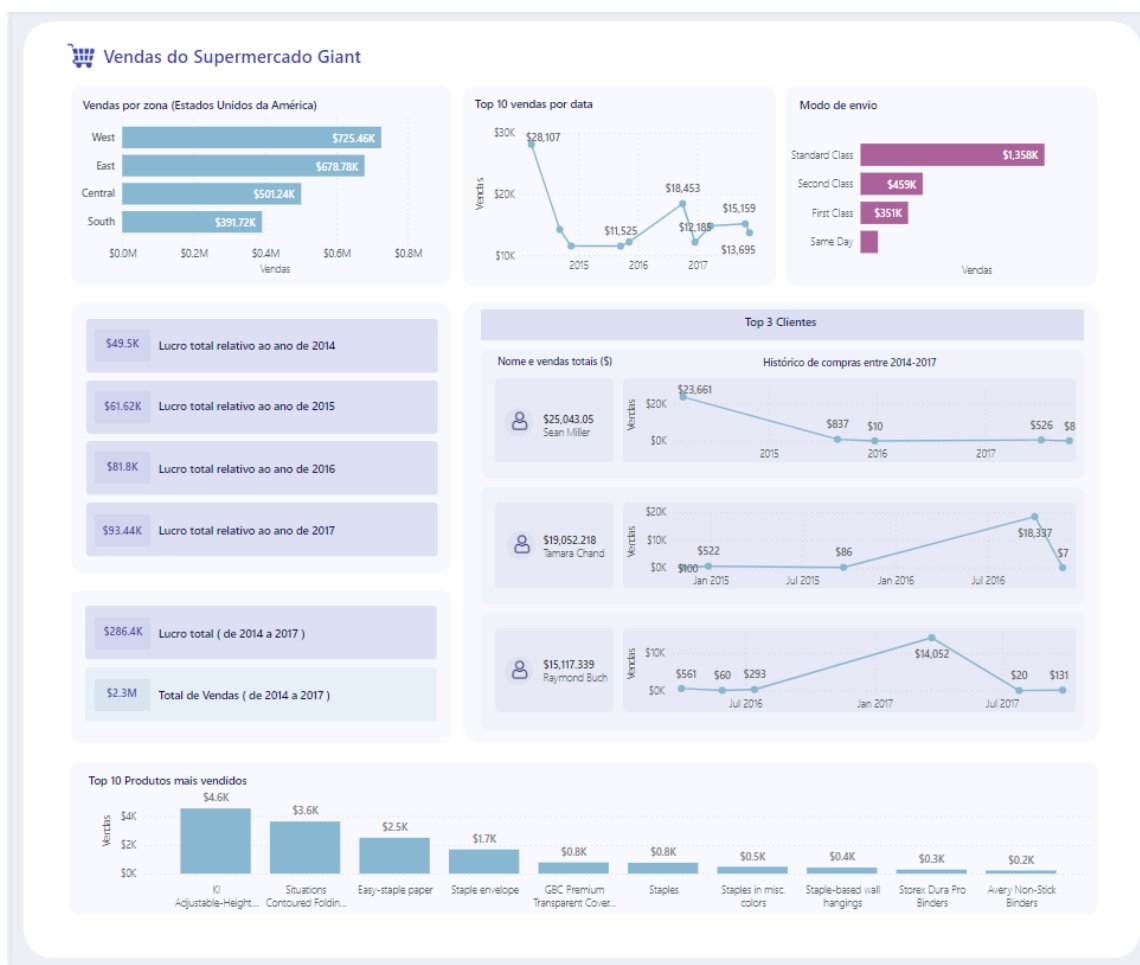


Figura 46- Report de vendas do Supermercado Giant

Neste caso, para atingir o sucesso do *Report*, como mencionado nos outros capítulos, era necessário que as informações presentes no mesmo fossem claras e de fácil compreensão. Como é possível observar pela Figura 46, apesar de ter bastante mais informação e vários dados a serem apresentados, houve um certo cuidado em apresentar os gráficos de forma legível, sem muita complexidade e sem que fosse necessário muito esforço de compreensão. Além disso, a organização dos gráficos foi essencial no *Report*. primeiro foi dada informação acerca das vendas (vendas por zona, vendas por data e modo de transporte), posteriormente informação relativa aos clientes (localização e top 3 clientes), os valores de lucros gerados com as vendas ao longo dos anos e, por fim, a apresentação dos produtos mais vendidos. Esta ligação: vendas, clientes, lucro e produtos, resulta numa melhor e mais clara interpretação do *Report*.

Foram escolhidos os *visuals* do *Power BI* que mais faziam sentido para a apresentação dos dados:

- Foi selecionado o gráfico de linhas para apresentar dados temporais, como é o exemplo do gráfico “*Top 10 vendas por data*” e o “*Histórico de compras entre 2014-2017*”. Esta escolha foi feita porque, a partir destes gráficos, é possível analisar com mais clareza todos os dados de vendas que ocorrem num período contínuo, neste caso, de 2014 a 2017;
- Em métricas simples, como é o caso da apresentação do valor do lucro, do total de dólares feito nas vendas e do total gasto pelos *Top 3* clientes, foram selecionados os cartões. Estes *visuals* são a melhor opção quando se pretende apresentar valores totais;
- Para os *visuals* “*Vendas por zona*” e “*Modo de envio utilizado*”, foi selecionado o gráfico de barras horizontais para ser possível efetuar uma melhor comparação dos dados. Assim, é possível analisar de forma clara quais as zonas em que mais vendas foram feitas e quais os modos de envio mais utilizados;
- No “*Top 10 produtos mais vendidos*” foi selecionado o gráfico de barras vertical para que seja possível fazer uma melhor comparação dos produtos que mais foram vendidos e o respetivo valor total de vendas;

Para o desenvolvimento deste *Report* foi selecionada uma paleta de cores pastel. A intenção foi de apenas utilizar 3 cores: a cor azul-clara para vendas, o lilás para o modo de transporte utilizado e o roxo para os lucros. Assim, é mais fácil para quem analisa associar automaticamente os dados às cores utilizadas.

Adicionalmente, o *Report* apresenta vários *visuals* e texto porque foi considerado um infográfico para se tornar uma Visualização de Dados bem-sucedida.

Como fator chave, o *Data-Ink* do *Report* é *elevado* uma vez que não foram utilizados efeitos nem gráficos 3D e foram utilizadas cores da mesma paleta de cores. Tornando-se, assim, um *Report* simples e sem grande complexidade visual (apesar de apresentar bastante informação).

6.4 Report *Top 50* Músicas do *Spotify*

O objetivo com o *Report* do *Spotify* é apresentar a informação relativamente às *Top 50* músicas em 2019. Para tal, para uma primeira abordagem, foram apresentados vários *visuals* com informações do artista mais ouvido:

- Nome;
- Género de música do artista;
- Popularidade do artista (este é o valor mais elevado de popularidade dos artistas- uma vez que é o artista mais popular);
- Música mais ouvida do artista;

Assim, com estas informações, é possível ter dados relativamente ao artista mais ouvido.



Figura 47- Visualizações do artista mais popular

Seguidamente, foram apresentados os *Top 5* de várias categorias. Como o *Dataset* em si já se tratava do *Top 50* músicas mais ouvidas no *Spotify*, era interessante filtrar a informação e apresentar os dados com menos complexidade. Sendo assim, foram selecionados:

- *Top 5* músicas com mais Oralidade;
- *Top 5* músicas com mais Vivacidade;
- *Top 5* músicas mais ouvidas;
- *Top 5* músicas com mais Energia;



Figura 48- Secção *Top 5* de várias categorias

Posteriormente, foi apresentada a informação relativa às músicas:

- Média de batidas por minuto;
- Média de duração das músicas;
- Género de música mais comum;

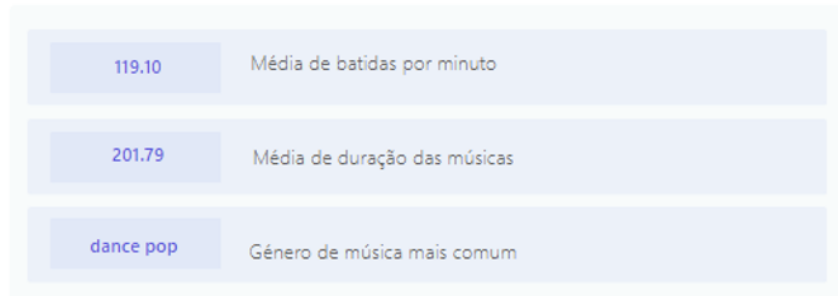


Figura 49- Visual com informação relativa às músicas

Sendo assim, o resultado do *Report* foi o seguinte:



Figura 50- Report Top 50 músicas Spotify

A ideia do *Report* foi de, primeiramente, apresentar o artista mais ouvido com as respetivas informações para revelar qual o artista que mais popularidade teve. Depois, a ideia foi a de mostrar o *Top 5* de diversas características das músicas (Oralidade, Vivacidade, Energia e mais ouvidas) para ter a perceção de quais foram os valores máximos alcançados nestas características e quais as músicas que alcançaram o *Top*.

Posteriormente, após terem sido dadas noções acerca do artista mais ouvido e das características das músicas, foi também adicionada informação acerca das propriedades das músicas (média de batidas e média de duração) e género de música mais popular.

Além disso, este *Report* salienta-se pela presença de legendas bastante descritivas e detalhadas. Para uma melhor perceção e noção do significado de algumas propriedades musicais, foi necessário reservar uma parte do *Report* para apresentar uma breve explicação de cada uma delas. Além do mais, assim, quem analisa consegue fazer a ligação com os *visuals* apresentados anteriormente e perceber que o valor máximo de cada propriedade musical corresponde ao *Top 1* apresentado.

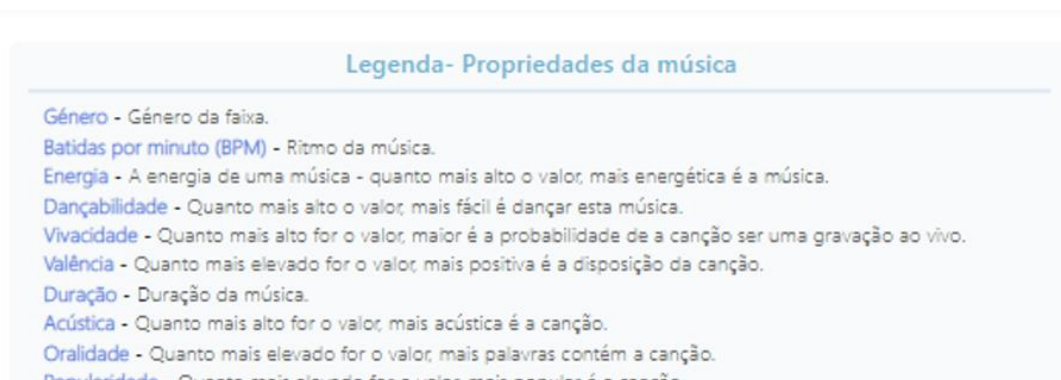


Figura 51- Visualização com legenda- propriedades da música

No visual que apresenta dados relativos ao *Top 5* músicas com mais Vivacidade, foi utilizado um *Pie Chart* para apresentar as percentagens. Neste caso, foi necessário explicitar que as percentagens apresentadas apenas eram calculadas a partir das 5 músicas com mais vivacidade e não no total.

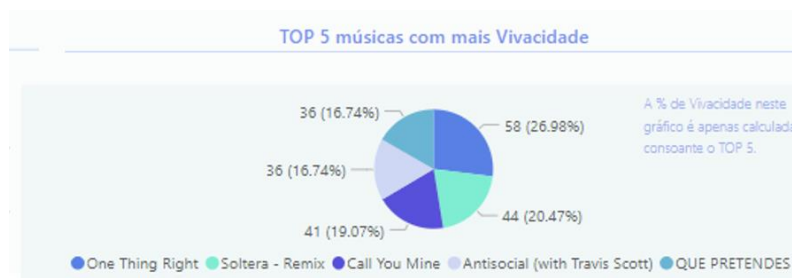


Figura 52- Visual com legenda do Top 5 músicas com mais Vivacidade

Para uma visualização de sucesso, foram selecionados os *visuals* que mais se enquadravam para a apresentação dos dados:

- Cartões para valores simples, como por exemplo, o nome e informações do artista mais ouvido e média de batidas por minuto, média de duração das músicas e género de música mais comum. Como são dados que não requerem grande complexidade, este *visual* foi a melhor opção para os revelar.
- Foi selecionada uma tabela para apresentar o *Top 5* músicas mais ouvidas porque, para além do nome da música e da popularidade (que define o *Top 5*), também fazia sentido apresentar informações e propriedades relevantes (género da música, nome do artista, dançabilidade e energia) para um maior perceção e visão geral dos dados apresentados;
- O gráfico de barras horizontal foi o visual escolhido para apresentar os dados do *Top 5* músicas com mais energia. Assim, a partir deste gráfico, é possível analisar claramente quais as músicas com mais energia desde o maior valor até ao menor.
- Para apresentar o *Top 5* músicas com mais vivacidade, foi selecionado o *Pie chart*. Apesar deste *visual*, normalmente, ter a possibilidade de ser enganador na partilha dos dados, nesta situação em específico recorreu-se a uma legenda para não transmitir informação errada ou enganadora. Neste caso, como explicado anteriormente, foi necessário explicitar que as percentagens apresentadas no *Pie Chart* eram calculadas apenas tendo em consideração as 5 músicas presentes e não todas as músicas do *Dataset* a soma de todas as percentagens dá um total de 100%, o que revela consistência na apresentação dos dados.
- Inicialmente, para apresentar o *Top 5* músicas com mais Oralidade, foi selecionado o gráfico de barras. Contudo, como também seria interessante analisar essas mesmas músicas consoante a Acústica e a Valência, foi selecionado um gráfico de barras agrupadas verticalmente para ser possível analisar cada música tendo em conta essas três propriedades.

Neste *Report* foram utilizados tons pastéis da mesma paleta de cores do símbolo da empresa “*Spotify*” para uma harmonização *visual*.

No desenvolvimento deste *Report*, à semelhança dos outros apresentados anteriormente, para ser bem-sucedido, foi considerado como um infográfico no sentido em que apresenta vários *visuals* e várias legendas detalhadas.

7. CONCLUSÃO

Neste capítulo é apresentada uma síntese do trabalho desenvolvido que originou esta dissertação de mestrado. Por curiosidade, foram colocadas duas questões relacionadas com os objetivos deste trabalho à ferramenta de Inteligência Artificial *ChatGPT*. Além disso, são também apontadas propostas de trabalho futuro.

7.1 Síntese do trabalho realizado

Esta dissertação foi inserida em contexto empresarial, a pedido da “*JTA - The Data Scientists*” e foi realizada no âmbito do Mestrado em Engenharia de Sistemas da Universidade do Minho.

O objetivo principal desta dissertação foi definir quais os principais fatores de sucesso da Visualização de Dados. Além disso, definir o que deve ser considerado como “sucesso” no que toca à representação de dados, catalogar ambientes e produtos lógicos de *software* mais utilizados para produzir Visualizações de Dados, ilustrar visualmente más práticas encontradas em projetos desta natureza e o desenvolvimento de *Reports* com ligação aos respetivos fatores de sucesso eram também objetivos desta dissertação.

Com o intuito de avaliar o sucesso dos objetivos propostos, é necessário efetuar uma análise do cumprimento de cada um. Sendo assim, o principal objetivo foi abordado no Capítulo 4 no qual, para um maior contexto, começou-se por definir o que era o sucesso no *Business Intelligence* e na Visualização dos Dados e, por fim, foram detalhados e explicados os Fatores Críticos de Sucesso.

A metodologia de investigação da presente dissertação foi a Revisão Sistemática da Literatura e, por esse mesmo motivo, ao longo da dissertação foi contextualizado o problema e foram identificadas e estudadas diversas fontes de literatura. Com estas fontes, foi possível esclarecer e aprender vários conceitos chave nesta dissertação tais como *Data Visualization*, *Business Intelligence* e *Business Analytics*.

Para concretizar outro objetivo pretendido, foram estudadas algumas ferramentas de Visualização de Dados presentes atualmente no mercado. A ferramenta *Microsoft Power BI* foi estudada com maior detalhe uma vez que é a ferramenta utilizada pela empresa em que se insere esta dissertação. Foram abordadas outras ferramentas tais como: *Excel*, *Google Charts*, *Tableau*, *Zoho Analytics*, *Datawrapper*, *Qlik Sense*, *Fusion Chart*, *Domo*, *Google Analytics* e *Visme*.

Outro objetivo foi concretizado a partir da ilustração visual de más práticas de Visualização de Dados e respetiva explicação. Foram detalhadas todas as más práticas que devem ser evitadas no desenvolvimento de qualquer ilustração de dados.

Por último, no Capítulo 6, para fazer a ligação com os respetivos fatores de sucesso, foram desenvolvidos *três Reports* utilizando a ferramenta *Power BI* para demonstrar visualmente o sucesso proveniente do bom uso desses fatores.

7.2 ChatGPT

A Inteligência Artificial tem vindo, cada vez mais, a fazer parte do dia-a-dia do ser humano. O *ChatGPT* é uma ferramenta de processamento de linguagem natural impulsionada pela tecnologia de Inteligência Artificial que lhe permite ter conversas semelhantes às humanas. O modelo de linguagem pode responder a várias perguntas e até ajudar a realizar tarefas, tais como escrever e-mails, textos e código (Ortiz, 2023).

Com o desenvolvimento e aparecimento do *ChatGPT*, várias questões são respondidas em tempo real e muitas vezes de forma clara. É importante referir que esta ferramenta de Inteligência Artificial não está isenta de erros e a sua fiabilidade não é total. A sua utilização pressupõe um forte sentido crítico e conhecimento sobre os temas a investigar. Como tal, tendo esta ferramenta de Inteligência Artificial bastantes dados e sendo capaz de fornecer informações em tempo real das perguntas que lhe são feitas recorreu-se, a título de curiosidade, ao *ChatGPT* para efetuar algumas questões-chave.

O primeiro tema a ser colocado ao *ChatGPT* relaciona-se com a questão fulcral deste documento: “Quais são os fatores críticos de sucesso da Visualização de Dados?” e, de seguida, outra pergunta essencial para o projeto: “Quais são as más práticas de Visualização de Dados?”.

As respostas dadas pelo *ChatGPT* são apresentadas nas imagens 53 e 54.



O que é considerado sucesso na Visualização dos dados?



O sucesso na visualização de dados pode ser medido com base em vários critérios. Aqui estão alguns aspectos que são frequentemente considerados indicadores de sucesso na visualização de dados:

1. **Compreensão dos dados:** Uma visualização de dados é considerada bem-sucedida quando transmite de forma clara e eficaz as informações contidas nos dados. Os usuários devem ser capazes de entender facilmente os padrões, tendências e insights apresentados na visualização.
2. **Comunicação eficaz:** O sucesso na visualização de dados está relacionado à capacidade de comunicar a mensagem desejada de maneira clara e impactante. A visualização deve contar uma história coerente e envolvente, destacando os principais insights e facilitando a compreensão dos dados por parte do público-alvo.
3. **Relevância e contexto:** Uma visualização de dados é bem-sucedida quando apresenta informações relevantes e significativas para o público-alvo. Ela deve ser capaz de responder às perguntas-chave e fornecer insights valiosos que ajudem na tomada de decisões ou no entendimento de um determinado problema ou fenômeno.
4. **Escolha adequada de gráficos:** O sucesso na visualização de dados envolve selecionar o tipo de gráfico correto para representar os dados de maneira mais eficaz. Diferentes tipos de gráficos são adequados para diferentes tipos de dados e propósitos, e a escolha adequada pode facilitar a compreensão e a interpretação dos dados.
5. **Estética e apelo visual:** Uma visualização de dados bem-sucedida é esteticamente agradável e visualmente atraente. A combinação adequada de cores, o layout equilibrado, o uso de elementos visuais adequados e a legibilidade das informações são fatores importantes para criar uma visualização atraente e envolvente.
6. **Interatividade e exploração dos dados:** O sucesso na visualização de dados pode ser medido pela capacidade de interatividade e exploração dos dados. Recursos interativos, como filtros, zoom, tooltips e animações, permitem que os usuários explorem os dados em maior detalhe, descubram padrões ocultos e personalizem a experiência de visualização.
7. **Impacto e tomada de decisões:** Uma visualização de dados é considerada bem-sucedida quando tem um impacto real nas decisões e ações tomadas com base nos insights fornecidos. Se a visualização leva a uma melhor compreensão dos dados e impulsiona a tomada de decisões informadas, ela é considerada eficaz.
8. **Feedback positivo dos usuários:** O sucesso na visualização de dados também pode ser medido pelo feedback positivo dos usuários. Se os usuários consideram a visualização útil, clara, intuitiva e valiosa para seu trabalho ou processo de tomada de decisão, isso indica que ela atingiu seu objetivo.

Lembrando que o sucesso na visualização de dados pode variar dependendo do contexto específico, do público-alvo e dos objetivos da visualização. É importante adaptar a visualização às necessidades e preferências dos usuários e buscar constantemente melhorias com base no feedback recebido.

Figura 53- Reposta do ChatGPT à pergunta "Quais os fatores críticos de sucesso da Visualização de Dados?"



Quais são as más práticas de Visualização de Dados?



Existem algumas más práticas comuns na visualização de dados que podem comprometer a eficácia e a compreensão das informações apresentadas. Aqui estão algumas delas:

1. **Sobrecarga de informações:** Incluir muitos elementos visuais, dados excessivos ou detalhes desnecessários em uma única visualização pode sobrecarregar o público e dificultar a compreensão dos dados. É importante manter a visualização limpa e concisa, destacando apenas as informações relevantes.
2. **Uso inadequado de gráficos:** Utilizar o tipo de gráfico errado para representar os dados pode levar a interpretações equivocadas. É importante escolher o gráfico apropriado com base na natureza dos dados e nos insights que se deseja comunicar. Evite gráficos inadequados ou distorcidos que possam induzir a conclusões errôneas.
3. **Falta de contexto:** Apresentar dados sem contexto adequado pode levar a interpretações equivocadas. É essencial fornecer informações explicativas, rótulos claros nos eixos, legendas informativas e qualquer outra informação que ajude o público a entender o que está sendo mostrado e por quê.
4. **Excesso de cores e estilos:** O uso excessivo de cores, estilos e elementos visuais pode distrair o público e dificultar a compreensão dos dados. É importante manter a visualização visualmente coerente, usando uma paleta de cores adequada, evitando o uso excessivo de gráficos em 3D ou sombras desnecessárias.
5. **Falta de acessibilidade:** Ignorar a acessibilidade pode excluir uma parte significativa do público. É importante garantir que a visualização seja acessível a pessoas com deficiências visuais ou outras necessidades especiais. Isso pode incluir o uso de contraste adequado, legendas alternativas para elementos visuais e a capacidade de navegar pela visualização usando apenas o teclado.
6. **Ausência de título e contexto:** A falta de um título claro e de informações contextuais pode dificultar a compreensão da visualização. É importante fornecer um título descritivo que resuma o objetivo da visualização e forneça orientações claras sobre o que está sendo mostrado.
7. **Falta de atualização:** Apresentar dados desatualizados ou não fornecer informações em tempo real pode comprometer a relevância da visualização. É importante garantir que os dados sejam atualizados regularmente para refletir a situação mais recente e manter a relevância da visualização.
8. **Ausência de testes e iteração:** Não testar a visualização com o público-alvo e não buscar feedback pode resultar em problemas de usabilidade e compreensão. É essencial realizar testes com usuários reais, obter feedback e iterar com base nas sugestões recebidas para melhorar a eficácia da visualização.

Evitar essas más práticas pode ajudar a criar visualizações de dados mais eficazes, claras e compreensíveis, permitindo que o público tire o máximo proveito das informações apresentadas.

Figura 54- Resposta do ChatGPT à pergunta "Quais são as más práticas de Visualização de Dados?"

A resposta apresentada pelo *ChatGPT* relativa aos fatores de sucesso na DV refere tópicos mencionados ao longo desta dissertação a partir de uma Revisão Sistemática da Literatura. As más práticas apresentadas pelo *ChatGPT* coincidem com as práticas apresentadas no Capítulo 5 deste documento. Conforme foi referido, esta consulta ao *ChatGPT* foi efetuada apenas a título de curiosidade uma vez que esta ferramenta tem sido muito publicitada e discutida gerando polémica nos meios de comunicação social, nas academias e na sociedade em geral em todo o mundo.

7.3 Perspetivas de trabalho futuro

Com o desenvolvimento deste projeto foi possível alcançar os fatores necessários para um desenvolvimento de Visualização de Dados de sucesso e algumas más práticas que devem ser evitadas. Neste projeto, foram desenvolvidos três Reports de acordo com esses fatores. Uma meta proposta para o futuro é a disponibilização de um formulário para obter feedback de cada um destes relatórios de dados. Assim, seria possível efetuar possíveis melhorias de acordo com a opinião do público-alvo.

O autor do documento considera que o projeto atingiu com sucesso os objetivos inicialmente propostos, tanto pelo cumprimento dos requisitos como pela experiência e conhecimento adquirido durante o período do seu desenvolvimento.

REFERÊNCIAS BIBLIOGRÁFICAS

- Ajah, I. A., & Nweke, H. F. (2019). Big data and business analytics: Trends, platforms, success factors and applications. In *Big Data and Cognitive Computing* (Vol. 3, Issue 2, pp. 1–30). MDPI. <https://doi.org/10.3390/bdcc3020032>
- AlWreikat, AlShawa, A., Al-Rimawi, D., & Fraihat, S. (2019, December 2). Business intelligence and data analytics system for mobile money. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3368691.3368729>
- Baier, T., & Neuwirth, E. (2007). Excel :: COM :: R. *Computational Statistics*, 22(1), 91–108. <https://doi.org/10.1007/s00180-007-0023-6>
- Berry, C. (2021). *When Data Visualization Really Isn't Useful (and When It Is)*. <https://www.oldstreetsolutions.com/good-and-bad-data-visualization>
- Blaxell, R. (2019). *The Manifesto of the Data-Ink Ratio*. <https://speedwell.com.au/insights/2019/the-manifesto-of-the-data-ink-ratio>
- Błazewicz, Jacek. (2000). *Handbook on parallel and distributed processing*. Springer.
- Camm, J. D., Fry, M. J., & Shaffer, J. (2017). A practitioner's guide to best practices in data visualization. *Interfaces*, 47(6), 473–488. <https://doi.org/10.1287/inte.2017.0916>
- Carrera-Rivera, A., Larrinaga, F., & Lasa, G. (2022). Context-awareness for the design of Smart-product service systems: Literature review. In *Computers in Industry* (Vol. 142). Elsevier B.V. <https://doi.org/10.1016/j.compind.2022.103730>
- Cheng, J., & Gu, X. (2022). Design and Application of Real-Time Analysis Model for Accounts Receivable Based on PowerBI. *Proceedings - 2022 Asia-Pacific Computer Technologies Conference, APCT 2022*, 16–20. <https://doi.org/10.1109/APCT55107.2022.00011>
- Chowdhury, V. (2022). *Superstore Dataset*. <https://www.kaggle.com/datasets/vivek468/superstore-dataset-final?resource=download>
- Corter, L. (2021). *Netflix Original Films & IMDB Scores*. <https://www.kaggle.com/datasets/luisortor/netflix-original-films-imdb-scores>
- Duò, M. (2022). *O Guia para as Melhores Ferramentas de Visualização de Dados em 2023*. <https://kinsta.com/pt/blog/ferramentas-visualizacao-de-dados/>
- Ferrari, A., & Russo, M. (n.d.). *Introducing Microsoft Power BI*.

- García, J. M. V., & Pinzón, B. H. D. (2017). Key success factors to business intelligence solution implementation. *Journal of Intelligence Studies in Business*, 7(1), 48–69. <https://doi.org/10.37380/jisib.v7i1.215>
- Hassan, M. (2022). *What is a kaggle? | Data Science and Machine Learning*. <https://www.kaggle.com/general/328265>
- Hati, S. (2023). *Building Data Visualization in Excel*. <https://www.knowledgehut.com/blog/business-intelligence-and-visualization/data-visualization-excel#data-visualization-tools-in-excel>
- Henrique, L. (2019). *Top 50 Spotify Songs - 2019*. <https://www.kaggle.com/datasets/leonardopena/top50spotify2019>
- Işik, Ö., Jones, M. C., & Sidorova, A. (2013). Business intelligence success: The roles of BI capabilities and decision environments. *Information and Management*, 50(1), 13–23. <https://doi.org/10.1016/j.im.2012.12.001>
- Jakhar, R., & Krishna, C. (2020). *Business Intelligence: As a Strategic Tool for Organization Development (A Literature Review)*. <https://www.researchgate.net/publication/348621052>
- Kirk, Andy. (2012). *Data visualization : a successful design process : a structured design approach to equip you with the knowledge of how to successfully accomplish any data visualization challenge efficiently and effectively*. Packt Pub.
- Liu, Y., & Chen, X. (2022). Application of Big Data Analysis Based on Power BI in Sales Forecasts. *ACM International Conference Proceeding Series*, 722–726. <https://doi.org/10.1145/3569966.3571272>
- Liu, Y., John, S., Han, H., & DeBello, J. E. (n.d.). *The Challenges of Business Analytics: Successes and Failures*. <http://hdl.handle.net/10125/49992>
- Mandava Geetha Bhargava, B., Tara Phani Surya Kiran, K., Rajeswara Rao, D., Geetha Bhargava α, M., Tara Phani Surya Kiran σ, K., & Rajeswara Rao ρ, D. (2018). Analysis and Design of Visualization of Educational Institution Database using Power BI Tool. In *Type: Double Blind Peer Reviewed International Research Journal Software & Data Engineering Global Journal of Computer Science and Technology: C* (Vol. 1).
- Metwalli, S. A. (2020). *Data Visualization 101: 7 Steps for Effective Visualizations*. <https://towardsdatascience.com/data-visualization-101-7-steps-for-effective-visualizations-491a17d974de>
- Midway, S. R. (2020). Principles of Effective Data Visualization. In *Patterns* (Vol. 1, Issue 9). Cell Press. <https://doi.org/10.1016/j.patter.2020.100141>

- Münster, E. (2017). *What makes a data visualization successful?*
<https://www.kdnuggets.com/2017/09/what-makes-data-visualization-successful.html>
- Musa, S. M., Akujuobi, C., Sadiku, M. N. O., Shadare, A. E., Akujuobi, C. M., & Perry, R. G. (2016). Publication Impact Factor (PIF): 1.02 www.sretechjournal.org DATA VISUALIZATION. *International Journal of Engineering Research And Advanced Technology*.
<https://www.researchgate.net/publication/311597028>
- Muskan, Singh, G., Singh, J., & Prabha, C. (2022). Data Visualization and its Key Fundamentals: A Comprehensive Survey. *7th International Conference on Communication and Electronics Systems, ICCES 2022 - Proceedings*, 1710–1714. <https://doi.org/10.1109/ICCES54183.2022.9835803>
- Naidoo, J., & Campbell, K. (2016). Extended abstract: Best practices for data visualization. *IEEE International Professional Communication Conference, 2016-November*.
<https://doi.org/10.1109/IPCC.2016.7740509>
- Nguyen, V. T., Jung, K., & Gupta, V. (2021). Examining data visualization pitfalls in scientific publications. *Visual Computing for Industry, Biomedicine, and Art*, 4(1). <https://doi.org/10.1186/s42492-021-00092-y>
- O'Connor, E. (n.d.). *Microsoft Power BI dashboards : step by step*.
- Olszak, C. M. (2016). Toward Better Understanding and Use of Business Intelligence in Organizations. *Information Systems Management*, 33(2), 105–123.
<https://doi.org/10.1080/10580530.2016.1155946>
- Olszak, C. M., & Ziemia, E. (2012). Critical Success Factors for Implementing Business Intelligence Systems in Small and Medium Enterprises on the Example of Upper Silesia, Poland. In *Interdisciplinary Journal of Information, Knowledge, and Management* (Vol. 7).
- Ortiz, S. (2023). *What is ChatGPT and why does it matter? Here's everything you need to know*.
<https://www.zdnet.com/article/what-is-chatgpt-and-why-does-it-matter-heres-everything-you-need-to-know/>
- Power, D. J., Heavin, C., McDermott, J., & Daly, M. (2018). Defining business analytics: an empirical approach. *Journal of Business Analytics*, 1(1), 40–53.
<https://doi.org/10.1080/2573234X.2018.1507605>
- Princeton, U. (2023). *Interactive Charts in Datawrapper – Remote Ethnography Workshop*.
<https://commons.princeton.edu/remote-ethnography/interactive-charts-in-datawrapper/>
- Qin, X., Luo, Y., Tang, N., & Li, G. (2020). Making data visualization more efficient and effective: a survey. *VLDB Journal*, 29(1), 93–117. <https://doi.org/10.1007/s00778-019-00588-3>

- Rad, R. (2018). Pro Power BI Architecture. In *Pro Power BI Architecture*. Apress. <https://doi.org/10.1007/978-1-4842-4015-1>
- Raghupathi, W., & Raghupathi, V. (2021). Contemporary business analytics: An overview. In *Data* (Vol. 6, Issue 8). MDPI. <https://doi.org/10.3390/data6080086>
- Ranjan, J. (2008). Business justification with business intelligence. *VINE*, 38(4), 461–475. <https://doi.org/10.1108/03055720810917714>
- Ranjan, J. (2009). Business intelligence: Concepts, components, techniques and benefits. *Journal of Theoretical and Applied Information Technology*, 9, 60–70.
- Sellis, Timos., ACM Digital Library., & Association for Computing Machinery. Special Interest Group on Management of Data. (2011). *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*. ACM.
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- Stobierski, T. (2021). *Bad Data Visualization: 5 Examples of Misleading Data*. <https://online.hbs.edu/blog/post/bad-data-visualization>
- Sun, Z., Strang, K., & Firmin, S. (2017). Business analytics-based enterprise information systems. *Journal of Computer Information Systems*, 57(2), 169–178. <https://doi.org/10.1080/08874417.2016.1183977>
- Tableau. (2022). *What is Tableau?* <https://www.tableau.com/why-tableau/what-is-tableau>
- Telea, A. C. (2007). Data Visualization: Principles and practice. In *Data Visualization: Principles and Practice*. <https://doi.org/10.1201/b10679>
- Thomé, A. M. T., Scavarda, L. F., & Scavarda, A. J. (2016). Conducting systematic literature review in operations management. In *Production Planning and Control* (Vol. 27, Issue 5, pp. 408–420). Taylor and Francis Ltd. <https://doi.org/10.1080/09537287.2015.1129464>
- Tinted, R. (2015). *BI Blog | Data Visualization & Analytics Blog | datapine*. <https://www.datapine.com/blog/misleading-data-visualization-examples/>
- Ujang Subagia, Angela Shierly Hermanto, & Gunawan Wang. (2020). Measuring Success Factors for Business Intelligence: Management Perspective. *International Journal of Science and Applied Information Technology*, 9(6), 31–34. <https://doi.org/10.30534/ijisait/2020/03962020>
- Violet, N. (2021). *Data Visualization Examples: Good, Bad and Misleading*. <https://www.syntaxtechs.com/blog/data-visualization-examples>

- Williams, R. A., Sheikh, N. J., Duman, G. M., & Kongar, E. (2022). Critical Success Factors of Business Intelligence Systems Implementation. *IEEE Engineering Management Review*.
<https://doi.org/10.1109/EMR.2022.3197096>
- Xiao, Y., & Watson, M. (2019). Guidance on Conducting a Systematic Literature Review. In *Journal of Planning Education and Research* (Vol. 39, Issue 1, pp. 93–112). SAGE Publications Inc.
<https://doi.org/10.1177/0739456X17723971>
- Yeoh, W., & Koronios, A. (2010). Critical Success Factors for Business Intelligence Systems. *Journal of Computer Information Systems*, *50*(3), 23–32.
<https://doi.org/10.1080/08874417.2010.11645404>
- Zoho. (2022). *Zoho Analytics - Online Help*. <https://www.zoho.com/analytics/help/basicconcepts.html>

ANEXO I – GUIÃO DE AVALIAÇÃO DE UMA VISUALIZAÇÃO DE DADOS

Tabela 4- Tabela de Avaliação de uma Visualização de Dados

	Sim	Não
Simplicidade na Visualização de Dados		
Seleção correta do gráfico para o tipo de dados apresentados		
Seleção de uma ferramenta adequada para a Visualização de Dados		
Data-Ink elevado		
Utilização correta de cores		
Utilização de legendas		
Dados apresentados de forma clara e correta		
Inclusão da incerteza		
Objetivo da Visualização alcançado		

ANEXO II – DATASETS

A.1- Dataset Netflix

	A	B	C	D	E	F	G	H
1	Title,Genre,Premiere,Runtime,IMDB Score,Language							
2	Enter the Anime,Documentary,"August 5, 2019",58,2.5,English/Japanese							
3	Dark Forces,Thriller,"August 21, 2020",81,2.6,Spanish							
4	The App,Science fiction/Drama,"December 26, 2019",79,2.6,Italian							
5	The Open House,Horror thriller,"January 19, 2018",94,3.2,English							
6	Kaali Khuhi,Mystery,"October 30, 2020",90,3.4,Hindi							
7	Drive>Action,"November 1, 2019",147,3.5,Hindi							
8	Leyla Everlasting,Comedy,"December 4, 2020",112,3.7,Turkish							
9	The Last Days of American Crime,Heist film/Thriller,"June 5, 2020",149,3.7,English							
10	Paradox,Musical/Western/Fantasy,"March 23, 2018",73,3.9,English							
11	Sardar Ka Grandson,Comedy,"May 18, 2021",139,4.1,Hindi							
12	Searching for Sheela,Documentary,"April 22, 2021",58,4.1,English							
13	The Call,Drama,"November 27, 2020",112,4.1,Korean							
14	Whipped,Romantic comedy,"September 18, 2020",97,4.1,Indonesian							
15	All Because of You>Action comedy,"October 1, 2020",101,4.2,Malay							
16	Mercy,Thriller,"November 22, 2016",90,4.2,English							
17	After the Raid,Documentary,"December 19, 2019",25,4.3,Spanish							
18	Ghost Stories,Horror anthology,"January 1, 2020",144,4.3,Hindi							
19	The Last Thing He Wanted,Political thriller,"February 21, 2020",115,4.3,English							
20	What Happened to Mr. Cha?,Comedy,"January 1, 2021",102,4.3,Korean							
21	Death Note,Horror thriller,"August 25, 2017",100,4.4,English							
22	Hello Privilege. It's Me, Chelsea,Documentary,"September 13, 2019",64,4.4,English							
23	Secret Obsession,Thriller,"July 18, 2019",97,4.4,English							
24	Sextuplets,Comedy,"August 16, 2019",99,4.4,English							
25	The Girl on the Train,Thriller,"February 26, 2021",120,4.4,Hindi							
26	Thunder Force,Superhero-Comedy,"April 9, 2021",105,4.4,English							
27	Fatal Affair,Thriller,"July 16, 2020",89,4.5,English							
28	Just Say Yes,Romantic comedy,"April 2, 2021",97,4.5,Dutch							
29	Seriously Single,Comedy,"July 31, 2020",107,4.5,English							
30	The Misadventures of Hedi and Cokeman,Comedy,"February 10, 2021",99,4.5,French							
31	5 Star Christmas,Comedy,"December 7, 2018",95,4.6,Italian							
32	After Maria,Documentary,"May 24, 2019",37,4.6,English/Spanish							
33	I Am the Pretty Thing That Lives in the House,Horror,"October 28, 2016",89,4.6,English							
34	Paris Is Us,Romance drama,"February 22, 2019",83,4.6,French							
35	Porta dos Fundos: The First Temptation of Christ,Comedy,"December 3, 2019",46,4.6,Portuguese							
36	Rattlesnake,Horror,"October 25, 2019",85,4.6,English							
37	The Players,Comedy,"July 15, 2020",88,4.6,Italian							
38	We Are One,Documentary,"July 14, 2020",86,4.6,French							
39	Finding Agnes,Drama,"November 30, 2020",105,4.7,Filipino							
40	IO,Science fiction/Drama,"January 18, 2019",95,4.7,English							
41	Sentinelles>Action,"March 5, 2021",80,4.7,French							
42	Sol Levante,Anime / Short,"April 2, 2020",4,4.7,English							
43	The Binding,Drama,"October 2, 2020",93,4.7,Italian							
44	We Can Be Heroes,Superhero,"December 25, 2020",100,4.7,English							
45	Christmas Crossfire,Thriller,"December 4, 2020",106,4.8,German							
46	Coin Heist,Heist,"January 6, 2017",97,4.8,English							
47	Mrs. Serial Killer,Thriller,"May 1, 2020",106,4.8,Hindi							
48	Nobody Sleeps in the Woods Tonight,Horror,"October 28, 2020",103,4.8,Polish							
49	Take the 10,Comedy,"January 20, 2017",80,4.8,English							
50	The Main Event,Comedy,"April 10, 2020",101,4.8,English							

Figura 55- Excerto do Dataset Netflix

A.2- Dataset Spotify

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	,"Track.Name", "Artist.Name", "Genre", "Beats.Per.Minute", "Energy", "Danceability", "Loudness.dB.", "Liveness", "Valence", "Length.", "Acousticness.", "Speechiness.", "Popularity"													
2	1,"SeOorita", "Shawn Mendes", "canadian pop", 117,55,76,-6,8,75,191,4,3,79													
3	2,"China", "Anuel AA", "reggaeton flow", 105,81,79,-4,8,61,302,8,9,92													
4	3,"boyfriend (with Social House)", "Ariana Grande", "dance pop", 190,80,40,-4,16,70,186,12,46,85													
5	4,"Beautiful People (feat. Khalid)", "Ed Sheeran", "pop", 93,65,64,-8,8,55,198,12,19,86													
6	5,"Goodbyes (Feat. Young Thug)", "Post Malone", "dfw rap", 150,65,58,-4,11,18,175,45,7,94													
7	6,"I Don't Care (with Justin Bieber)", "Ed Sheeran", "pop", 102,68,80,-5,9,84,220,9,4,84													
8	7,"Ransom", "Lil Tecca", "trap music", 180,64,75,-6,7,23,131,2,29,92													
9	8,"How Do You Sleep?", "Sam Smith", "pop", 111,68,48,-5,8,35,202,15,9,90													
10	9,"Old Town Road - Remix", "Lil Nas X", "country rap", 136,62,88,-6,11,64,157,5,10,87													
11	10,"bad guy", "Billie Eilish", "electropop", 135,43,70,-11,10,56,194,33,38,95													
12	11,"Callaita", "Bad Bunny", "reggaeton", 176,62,61,-5,24,24,251,60,31,93													
13	12,"Loco Contigo (feat. J. Balvin & Tyga)", "DJ Snake", "dance pop", 96,71,82,-4,15,38,185,28,7,86													
14	13,"Someone You Loved", "Lewis Capaldi", "pop", 110,41,50,-6,11,45,182,75,3,88													
15	14,"Otro Trago - Remix", "Sech", "panamanian pop", 176,79,73,-2,6,76,288,7,20,87													
16	15,"Money In The Grave (Drake ft. Rick Ross)", "Drake", "canadian hip hop", 101,50,83,-4,12,10,205,10,5,92													
17	16,"No Guidance (feat. Drake)", "Chris Brown", "dance pop", 93,45,70,-7,16,14,261,12,15,82													
18	17,"LA CANCI"Ñ", "J Balvin", "latin", 176,65,75,-6,11,43,243,15,32,90													
19	18,"Sunflower - Spider-Man: Into the Spider-Verse", "Post Malone", "dfw rap", 90,48,76,-6,7,91,158,56,5,91													
20	19,"Lalala", "Y2K", "canadian hip hop", 130,39,84,-8,14,50,161,18,8,88													
21	20,"Truth Hurts", "Lizzo", "escape room", 158,62,72,-3,12,41,173,11,11,91													
22	21,"Piece Of Your Heart", "MEDUZA", "pop house", 124,74,68,-7,7,63,153,4,3,91													
23	22,"Panini", "Lil Nas X", "country rap", 154,59,70,-6,12,48,115,34,8,91													
24	23,"No Me Conoce - Remix", "Jhay Cortez", "reggaeton flow", 92,79,81,-4,9,58,309,14,7,83													
25	24,"Soltera - Remix", "Lunay", "latin", 92,78,80,-4,44,80,266,36,4,91													
26	25,"bad guy (with Justin Bieber)", "Billie Eilish", "electropop", 135,45,67,-11,12,68,195,25,30,89													
27	26,"If I Can't Have You", "Shawn Mendes", "canadian pop", 124,82,69,-4,13,87,191,49,6,70													
28	27,"Dance Monkey", "Tones and I", "australian pop", 98,59,82,-6,18,54,210,69,10,83													
29	28,"It's You", "Ali Gatie", "canadian hip hop", 96,46,73,-7,19,40,213,37,3,89													
30	29,"Con Calma", "Daddy Yankee", "latin", 94,86,74,-3,6,66,193,11,6,91													
31	30,"QUE PRETENDES", "J Balvin", "latin", 93,79,64,-4,36,94,222,3,25,89													
32	31,"Takeaway", "The Chainsmokers", "edm", 85,51,29,-8,10,36,210,12,4,84													
33	32,"7 rings", "Ariana Grande", "dance pop", 140,32,78,-11,9,33,179,59,33,89													
34	33,"0.9583333333333333", "Maluma", "reggaeton", 96,71,78,-5,9,68,176,22,28,89													
35	34,"The London (feat. J. Cole & Travis Scott)", "Young Thug", "atl hip hop", 98,59,80,-7,13,18,200,2,15,89													
36	35,"Never Really Over", "Katy Perry", "dance pop", 100,88,77,-5,32,39,224,19,6,89													
37	36,"Summer Days (feat. Macklemore & Patrick Stump of Fall Out Boy)", "Martin Garrix", "big room", 114,72,66,-7,14,32,164,18,6,89													
38	37,"Otro Trago", "Sech", "panamanian pop", 176,70,75,-5,11,62,226,14,34,91													
39	38,"Antisocial (with Travis Scott)", "Ed Sheeran", "pop", 152,82,72,-5,36,91,162,13,5,87													
40	39,"Sucker", "Jonas Brothers", "boy band", 138,73,84,-5,11,95,181,4,6,80													
41	40,"fuck, i'm lonely (with Anne-Marie) - from 13 Reasons Why: Season 3", "Lauv", "dance pop", 95,56,81,-6,6,68,199,48,7,78													
42	41,"Higher Love", "Kygo", "edm", 104,68,69,-7,10,40,228,2,3,88													
43	42,"You Need To Calm Down", "Taylor Swift", "dance pop", 85,68,77,-6,7,73,171,1,5,90													
44	43,"Shallow", "Lady Gaga", "dance pop", 96,39,57,-6,23,32,216,37,3,87													
45	44,"Talk", "Khalid", "pop", 136,40,90,-9,6,35,198,5,13,84													
46	45,"Con Altura", "ROSALÍA", "r&b en español", 98,69,88,-4,5,75,162,39,12,88													
47	46,"One Thing Right", "Marshmello", "brostep", 88,62,66,-2,58,44,182,7,5,88													
48	47,"Te Robaré", "Nicky Jam", "latin", 176,75,67,-4,8,80,202,24,6,88													
49	48,"Happier", "Marshmello", "brostep", 100,79,69,-3,17,67,214,19,5,88													
50	49,"Call You Mine", "The Chainsmokers", "edm", 104,70,59,-6,41,50,218,23,3,88													
51	50,"Cross Me (feat. Chance the Rapper & PnB Rock)", "Ed Sheeran", "pop", 95,79,75,-6,7,61,206,21,12,82													

Figura 56- Excerto do Dataset Spotify

A.3- Dataset Supermecado

Row ID	Order ID	Order Date	Ship Date	Mode	Customer ID	Customer Name	Segment	City	State	Postal Code	Region	Product ID	Category	Sub-Category	Product Name	Sales	Quantity	Discount	Profit
1	CA-2016-152156	11/8/2016	11/11/2016	Second Class	CG-12520	Claire Guts	Consumer	United States	Henderson	Kentucky	42420	South	FUR-BO-10001798	Furniture	Bookcases	Bush Somerset Collection Bookcase	261.96	2,0	41.9136
2	CA-2016-152156	11/8/2016	11/11/2016	Second Class	CG-12520	Claire Guts	Consumer	United States	Henderson	Kentucky	42420	South	FUR-CH-10000454	Furniture	Chairs	Hon Deluxe Fabric Upholstered Stacking Chairs, Rounded Back	731.94	3,0	219.582
3	CA-2016-138688	6/12/2016	6/16/2016	Second Class	DV-13045	Darin V Huff	Corporate	United States	Los Angeles	California	90036	West	OFF-LA-10000240	Office Supplies	Labels	Self-Adhesive Address Labels for Typewriters by Universal	14.62	2,0	6.8714
4	US-2015-108966	10/11/2015	10/18/2015	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	Florida	33311	South	FUR-TA-10000577	Furniture	Tables	Bretford CR4500 Series Slim Rectangular Table	957.5775	5,0	45.383031
5	US-2015-108966	10/11/2015	10/18/2015	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	Florida	33311	South	OFF-ST-10000760	Office Supplies	Storage	Eldon Fold 'N Roll Cart System	22.368	2,0	2.5164
6	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	FUR-FU-10001487	Furniture	Furnishings	Eldon Expressions Wood and Plastic Desk Accessories, Cherry Wood	48.86	7,0	14.1694
7	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	OFF-AR-10002833	Office Supplies	Art	Newell 322.7.28.4.0.1.9656			
8	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	PH-10002275	Technology	Phones	Mitel 5320 IP Phone VoIP phone	907.152	6,0	2,90.7152
9	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	OFF-BI-10003910	Office Supplies	Binders	DXL Angle-View Binders with Locking Rings by Samsill	18.504	3,0	2,5.7825
10	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	OFF-AP-10002892	Office Supplies	Appliances	Belkin F5C206VTEL 6 Outlet Surge	114.9	5,0	34.47
11	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	FUR-TA-10001539	Furniture	Tables	Chromcraft Rectangular Conference Tables	1706.184	9,0	2,85.3092
12	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	TEC-PH-10002033	Technology	Phones	Konftel 250 Conference Phone	11.424	4,0	2,68.3568
13	CA-2017-114412	4/15/2017	4/20/2017	Standard Class	AA-10480	Andrew Allen	Consumer	United States	Concord	North Carolina	28027	South	OFF-PA-10002365	Office Supplies	Paper	Xerox 1967.15.552.3.0.2.5.4432			
14	CA-2016-161389	12/5/2016	12/10/2016	Standard Class	IM-15070	Irene Maddox	Consumer	United States	Seattle	Washington	98103	West	OFF-BI-10003656	Office Supplies	Binders	Fellowes PB200 Plastic Comb Binding Machine	407.976	3,0	2,132.5922
15	US-2015-118983	11/22/2015	11/26/2015	Standard Class	HP-14815	Harold Pawlan	Home Office	United States	Fort Worth	Texas	76106	Central	OFF-AP-10002311	Office Supplies	Appliances	Holmes Replacement Filter for HEPA Air Cleaner, Very Large Room, HEPA Filter	68.81	5,0	8.123.858
16	US-2015-118983	11/22/2015	11/26/2015	Standard Class	HP-14815	Harold Pawlan	Home Office	United States	Fort Worth	Texas	76106	Central	OFF-BI-10000756	Office Supplies	Binders	Storex DuraTech Recycled Plastic Frosted Binders	2.544	3,0	8.316
17	CA-2014-105899	11/11/2014	11/18/2014	Standard Class	PK-19075	Pete Kruz	Consumer	United States	Madison	Wisconsin	53711	Central	OFF-ST-10004186	Office Supplies	Storage	Stur-D-Stor Shelving, Vertical 5-Shelf: 72"H x 36"W x 18 1/2"D	665.88	6,0	13.3176
18	CA-2014-167164	5/13/2014	5/15/2014	Second Class	AG-10270	Alejandro Grove	Consumer	United States	West Jordan	Utah	84084	West	OFF-ST-10000107	Office Supplies	Storage	Fellowes Super Drawer	55.5	2,0	9.99
19	CA-2014-143336	8/27/2014	9/1/2014	Second Class	ZD-21925	Zuschuss Donatelli	Consumer	United States	San Francisco	California	94109	West	OFF-AR-10003056	Office Supplies	Art	Newell 341.8.56.2.0.2.4824			
20	CA-2014-143336	8/27/2014	9/1/2014	Second Class	ZD-21925	Zuschuss Donatelli	Consumer	United States	San Francisco	California	94109	West	TEC-PH-10001949	Technology	Phones	Cisco SPA 501G IP Phone	213.48	3,0	2,16.011
21	CA-2014-143336	8/27/2014	9/1/2014	Second Class	ZD-21925	Zuschuss Donatelli	Consumer	United States	San Francisco	California	94109	West	OFF-BI-10002215	Office Supplies	Binders	Wilson Jones Hanging View Binder, White, 1"	22.72	4,0	2,7.384
22	CA-2016-137330	12/9/2016	12/13/2016	Standard Class	KB-16585	Ken Black	Corporate	United States	Fremont	Nebraska	68025	Central	OFF-AR-10000246	Office Supplies	Art	Newell 318.19.46.7.0.5.0596			
23	CA-2016-137330	12/9/2016	12/13/2016	Standard Class	KB-16585	Ken Black	Corporate	United States	Fremont	Nebraska	68025	Central	OFF-AP-10001492	Office Supplies	Appliances	Acco Six-Outlet Power Strip, 4' Cord Length	60.34	7,0	15.6884
24	US-2017-106909	7/16/2017	7/18/2017	Second Class	SF-20065	Sandra Flanagan	Consumer	United States	Philadelphia	Pennsylvania	19140	East	FUR-CH-10002774	Furniture	Chairs	Global Deluxe Stacking Chair, Gray	71.372	2,0	3.10196
25	CA-2015-106320	9/25/2015	9/30/2015	Standard Class	EB-13870	Emily Burns	Consumer	United States	Orem	Utah	84057	West	FUR-TA-10000577	Furniture	Tables	Bretford CR4500 Series Slim Rectangular Table	1044.63	3,0	2,40.2649
26	CA-2016-121755	1/16/2016	1/20/2016	Second Class	EH-13945	Eric Hoffmann	Consumer	United States	Los Angeles	California	90049	West	OFF-BI-10001634	Office Supplies	Binders	Wilson Jones Active Use Binders	11.648	2,0	2,4.2224
27	CA-2016-121755	1/16/2016	1/20/2016	Second Class	EH-13945	Eric Hoffmann	Consumer	United States	Los Angeles	California	90049	West	TEC-AC-10003027	Technology	Accessories	Imation 18GB Mini TravelDrive USB 2.0 Flash Drive	90.57	3,0	11.7741
28	US-2015-106030	9/17/2015	9/21/2015	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	FUR-BO-10004834	Furniture	Bookcases	Riverside Palais Royal Lawyers Bookcase, Royale Cherry Finish	3083.43	7,0	5.1665.0522
29	US-2015-106030	9/17/2015	9/21/2015	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	OFF-BI-10000474	Office Supplies	Binders	Avery Recycled Flexi-View Covers for Binding Systems	9.618	2,0	7.70532
30	US-2015-106030	9/17/2015	9/21/2015	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	FUR-FU-10004848	Furniture	Furnishings	Howard Miller 13-3/4" Diameter Brushed Chrome Round Wall Clock	124.2	3,0	2,15.525
31	US-2015-106030	9/17/2015	9/21/2015	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	OFF-EN-10001509	Office Supplies	Envelopes	Poly String Tie Envelopes	3.264	2,0	2,1.1016
32	US-2015-106030	9/17/2015	9/21/2015	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	OFF-AR-10004042	Office Supplies	Art	BOSTON Model 1800 Electric Pencil Sharpeners, Putty/Woodgrain	86.304	6,0	2,9.7092
33	US-2015-106030	9/17/2015	9/21/2015	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	OFF-BI-10001525	Office Supplies	Binders	Acco Pressboard Covers with Storage Hooks, 14 7/8" x 11", Executive Red	6.858	6,0	7.5715
34	US-2015-106030	9/17/2015	9/21/2015	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	OFF-AR-10001683	Office Supplies	Art	Lumber Crayons	15.76	2,0	2,3.546
35	CA-2017-107727	10/19/2017	10/23/2017	Second Class	MA-17560	Matt Abelman	Home Office	United States	Houston	Texas	77095	Central	OFF-PA-10000249	Office Supplies	Paper	Easy-staple paper	29.472	3,0	2,9.9468
36	CA-2016-117590	12/8/2016	12/10/2016	First Class	GH-14485	Gene Hale	Corporate	United States	Richardson	Texas	75080	Central	TEC-PH-10004977	Technology	Phones	GE 30524EE4.1097.544.7.0.2.123.4737			
37	CA-2016-117590	12/8/2016	12/10/2016	First Class	GH-14485	Gene Hale	Corporate	United States	Richardson	Texas	75080	Central	FUR-FU-10003664	Furniture	Furnishings	Electric Architect's Clamp-On Swing Arm Lamp, Black	190.92	5,0	6.147.963
38	CA-2015-117415	12/27/2015	12/31/2015	Standard Class	SN-20710	Steve Nguyen	Home Office	United States	Houston	Texas	77041	Central	OFF-EN-10002986	Office Supplies	Envelopes	110-4 1/8" x 9 1/2" Premium Diagonal Seam Envelopes	113.328	9,0	2,35.415
39	CA-2015-117415	12/27/2015	12/31/2015	Standard Class	SN-20710	Steve Nguyen	Home Office	United States	Houston	Texas	77041	Central	FUR-BO-10002545	Furniture	Bookcases	Atlantic Metals Mobile 3-Shelf Bookcases, Custom Colors	532.3992	3,0	32.46.9764
40	CA-2015-117415	12/27/2015	12/31/2015	Standard Class	SN-20710	Steve Nguyen	Home Office	United States	Houston	Texas	77041	Central	FUR-CH-10004218	Furniture	Chairs	Global Fabric Manager's Chair, Dark Gray	212.058	3,0	13.147
41	CA-2015-117415	12/27/2015	12/31/2015	Standard Class	SN-20710	Steve Nguyen	Home Office	United States	Houston	Texas	77041	Central	PH-10000486	Technology	Phones	Plantronics HL10 Handset Lifter	371.168	4,0	21.47564
42	CA-2017-120999	9/10/2017	9/15/2017	Standard Class	LC-16930	Linda Cazamian	Corporate	United States	Naperville	Illinois	60540	Central	TEC-PH-10004093	Technology	Phones	Panasonic Kx-TSS50.147.168.4.0.2.16.5564			
43	CA-2016-101343	7/17/2016	7/22/2016	Standard Class	RA-19885	Ruben Ausman	Corporate	United States	Los Angeles	California	90049	West	OFF-ST-10003479	Office Supplies	Storage	Eldon Base for stackable storage shelf, platinum	77.88	2,0	3.894
44	CA-2017-139619	9/19/2017	9/23/2017	Standard Class	ES-14080	Erin Smith	Corporate	United States	Melbourne	Florida	32935	South	OFF-ST-10003282	Office Supplies	Storage	Advantus 10-Drawer Portable Organizer, Chrome Metal Frame, Smoke Drawers	95.616	2,0	2,9.5616
45	CA-2016-118255	3/11/2016	3/13/2016	First Class	ON-18715	Odella Nelson	Corporate	United States	Eagan	Minnesota	55122	Central	TEC-AC-10000171	Technology	Accessories	Verbatim 25 GB 6x Blu-ray Single Layer Recordable Disc, 25/Pack	45.98	2,0	19.7714
46	CA-2016-118255	3/11/2016	3/13/2016	First Class	ON-18715	Odella Nelson	Corporate	United States	Eagan	Minnesota	55122	Central	OFF-BI-10003291	Office Supplies	Binders	Wilson Jones Leather-Like Binders with Duplicor Round Rings	17.46	2,0	8.2062
47	CA-2014-146703	10/20/2014	10/25/2014	Second Class	PO-18865	Patrick O'Donnell	Consumer	United States	Westland	Michigan	48185	Central	OFF-ST-10001713	Office Supplies	Storage	Gould Plastics 9-Pocket Panel Bin, 18-3/8w x 5-1/4d x 20-1/2h, Black	211.96	4,0	8.4784
48	CA-2016-169194	6/20/2016	6/25/2016	Standard Class	LH-16900	Lena Hernandez	Consumer	United States	Dover	Delaware	19901	East	TEC-AC-10002167	Technology	Accessories	Imation 18gb Micro Traveldrive Usb 2.0 Flash Drive	45.3	4,0	9.5
49	CA-2016-169194	6/20/2016	6/25/2016	Standard Class	LH-16900	Lena Hernandez	Consumer	United States	Dover	Delaware	19901	East	TEC-PH-10003988	Technology	Phones	LF Elite 3D Dazzle Designer Hard Case Cover, LF Stylus Pen and Wiper For Apple Iphone 5c Mini Lite	21.8	2,0	6.104

Figura 57- Excerto do Dataset Supermecado Giant