

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

**EDUARDO RODRIGUES DE OLIVEIRA
ELLEN CRYSTINA LESTENSKY PEREIRA**

**DESENVOLVIMENTO DE UM ROTEIRO PARA A ANÁLISE DE PROJETOS E
CARGA DE TRABALHO UTILIZANDO FERRAMENTAS DE *BUSINESS*
INTELLIGENCE COM BASE EM DADOS DISPONÍVEIS NO SETOR DE BENS DE
CONSUMO**

CURITIBA

2022

**EDUARDO RODRIGUES DE OLIVEIRA
ELLEN CRYSTINA LESTENSKY PEREIRA**

**DESENVOLVIMENTO DE UM ROTEIRO PARA A ANÁLISE DE PROJETOS E
CARGA DE TRABALHO UTILIZANDO FERRAMENTAS DE *BUSINESS
INTELLIGENCE* COM BASE EM DADOS DISPONÍVEIS NO SETOR DE BENS DE
CONSUMO**

**Development of a script for the analysis of projects and workload using
business intelligence tools based on data available in the consumer goods
sector**

Trabalho de conclusão de curso de graduação
apresentado como requisito para obtenção do título de
Bacharel em Engenharia Mecânica da Universidade
Tecnológica Federal do Paraná (UTFPR).
Orientadora: Profa. Dra. Cleina Yayoe Okoshi.

**CURITIBA
2022**



Esta licença permite download e compartilhamento do trabalho desde que sejam atribuídos créditos ao(s) autor(es), sem a possibilidade de alterá-lo ou utilizá-lo para fins comerciais. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

**EDUARDO RODRIGUES DE OLIVEIRA
ELLEN CRYSTINA LESTENSKY PEREIRA**

**DESENVOLVIMENTO DE UM ROTEIRO PARA A ANÁLISE DE PROJETOS E
CARGA DE TRABALHO UTILIZANDO FERRAMENTAS DE *BUSINESS*
INTELLIGENCE COM BASE EM DADOS DISPONÍVEIS NO SETOR DE BENS DE
CONSUMO**

Trabalho de conclusão de curso de graduação
apresentado como requisito para obtenção do título de
Bacharel em Engenharia Mecânica da Universidade
Tecnológica Federal do Paraná (UTFPR).

Data de aprovação: 02/12/2022

Cleina Yayoe Okoshi
Doutora em Engenharia de Produção e Sistemas
Universidade Tecnológica Federal do Paraná

Oswaldo Verussa Junior
Mestre em Engenharia Mecânica e de Materiais
Universidade Tecnológica Federal do Paraná

Walter Luís Mikos
Doutor em Engenharia Mecânica
Universidade Tecnológica Federal do Paraná

CURITIBA

2022

AGRADECIMENTOS

Agradecemos a Universidade Tecnológica Federal do Paraná e aos docentes que nos direcionaram pelo caminho do aprendizado da engenharia, nos fornecendo todas as ferramentas e conhecimentos que foram essenciais para o nosso desenvolvimento pessoal-profissional e para a resolução deste trabalho.

Agradecemos em especial à nossa orientadora, Profa. Dra. Cleina Yayoe Okoshi, que nos deu todo o suporte, e auxiliou com a sua experiência em todos os nossos questionamentos para o desenvolvimento deste projeto e, acima de tudo, nos apoiou e se engajou em nossa ideia, demonstrando genuíno interesse e envolvimento.

Aos nossos pais agradecemos por serem apoio e sustentação durante toda a nossa trajetória acadêmica, e aos nossos amigos agradecemos por todo o suporte emocional e colaboração ao longo desta jornada.

RESUMO

A quantidade de dados a quais estamos expostos atualmente não necessariamente implica em um maior conhecimento, mas possibilita o acesso a cada vez mais informações para basear nossas decisões. Através de modelagens de dados se faz possível criar representações visuais que auxiliem na compreensão das informações obtidas e melhorem a tomada de decisões empresariais. Este trabalho tem como objetivo desenvolver um roteiro aplicável de elaboração de *Dashboards* a respeito de projetos em organizações de bens de consumo, dentro do *software Power BI*, utilizando os passos da metodologia *Crisp-DM*, que consistem em identificar, mapear e compreender esses dados. Como efeito, este roteiro se torna aplicável para ser empregado nas bases de dados que envolvam informações semelhantes as utilizadas neste projeto e possibilitam que as organizações tenham análises mais padronizadas e em menor tempo.

Palavras-chave: Análise de dados; *Power BI*; *Business Intelligence*.

ABSTRACT

The data amount to which we are currently exposed does not necessarily imply in a greater knowledge, but it allows access to increasingly more information to base our decisions on. Through data modeling it is possible to create visual representations that help in understanding the information obtained and improve business decision making. This work aims to develop an applicable script for the elaboration of Dashboards regarding projects in consumer goods organizations, within the Power BI software, using the steps of the Crisp-DM methodology, which consist of identifying, mapping and understanding this data. As an effect, this script becomes applicable to be used in databases that involve similar information to those used in this project and allow organizations to have more standardized analyzes and in less time.

Keywords: Data Analysis; Power BI; Business Intelligence.

LISTA DE ILUSTRAÇÕES

Figura 1 - Interdependência da cadeia de suprimentos no contexto da indústria 4.0	19
Figura 2 - Os 3 Vs do Big Data	20
Figura 3 - O Valor da Narrativa.....	25
Figura 4 - Ciclo da metodologia Crisp-DM	26
Figura 5 - Etapas do Crisp-DM	27
Figura 6 - Quadrantes mágicos Gartner	28
Figura 7 - Exemplo de transformação no <i>Power Query</i>	29
Figura 8 - Linguagem M	30
Figura 9 - Visuais do <i>Power BI</i>	31
Figura 10 - Relações no <i>Power BI</i>	32
Figura 11 - Exemplo de aplicação DAX	33
Figura 12 - Fluxograma de desenvolvimento do TCC.....	34
Figura 13 - Tabelas mínimas disponíveis.....	37
Figura 14 - Multidimensionalidade dos dados.....	38
Figura 15 - Tipos de visuais padrões no <i>Power BI</i>	39
Figura 16 - Esboço das relações entre as tabelas '.CSV'.....	44
Figura 17 - Importar novos arquivos para o <i>Power BI</i>	45
Figura 18 - Abrir editor do <i>Power Query</i>	45
Figura 19 - Navegação entre tabelas e etapas no <i>Power Query</i>	46
Figura 20 - Tabela de Projetos modelada.....	47
Figura 21 - Ferramenta de divisão em colunas.....	47
Figura 22 - Tabela de Demandas modelada.....	48
Figura 23 - Tabela de Capacidade modelada	48
Figura 24 - Índice de categorias.....	49
Figura 25 - Índice de Cidades.....	50
Figura 26 - Índice de Setores.....	50
Figura 27 - Relações entre as Tabelas.	51
Figura 28 - Esboço da página de Sumario	53
Figura 29 - Esboço da página Financeiro.....	54
Figura 30 - Esboço da página de Maiores Projetos.....	55
Figura 31 - Esboço da página de Gráficos de Demanda.....	55
Figura 32 - Esboço da página de Tabelas de Demanda.....	56
Figura 33 - Esboço da página de Extração de dados.....	56
Figura 34 - Modelo no <i>Power Point</i>	57
Figura 35 - <i>Templates</i> aplicados	57
Figura 36 - Aplicação dos Modelos no <i>Power BI</i>	58
Figura 37 - Adição de segmentadores de dados	59
Figura 38 - Uso das dicas de ferramentas.....	59
Figura 39 - Distribuição das áreas de conhecimento.....	60
Figura 40 - Experiência média dos entrevistados	61
Figura 41 - Avaliação da adaptabilidade do processo	61
Figura 42 - Avaliação da Aplicabilidade do processo	62
Figura 43 - Avaliação do resultado das análises	62
Figura 44 - Avaliação do impacto dos dashboards na eficiência	62
Figura 45 - Resultado geral do Procedimento	63

LISTA DE QUADROS

Quadro 1 - Técnicas de BI Atuais	24
Quadro 2 - Colunas da tabela de Projetos	42
Quadro 3 - Colunas da tabela de Demanda.....	43
Quadro 4 - Colunas da tabela de Capacidade.....	43

LISTA DE ABREVIATURAS E SIGLAS

BI	<i>Business Intelligence</i>
Crisp-DM	<i>Cross Industry Standard Process for Data Mining</i>
CSV	<i>Comma Separated Value</i>
DAX	<i>Data Analysis Expressions</i>
FTE	<i>Full Time Employee</i>
KPI	<i>Key Performance Indicator</i>

SUMÁRIO

1	INTRODUÇÃO	13
1.1	Caracterização do problema	14
1.2	Objetivos	14
1.2.1	Objetivo geral	14
1.2.2	Objetivos específicos.....	14
1.3	Justificativa	15
1.4	Etapas do trabalho	16
2	FUNDAMENTAÇÃO TEÓRICA	17
2.1	Revoluções tecnológicas industriais	17
2.2	Indústria 4.0	18
2.3	<i>Big Data</i>	20
2.3.1	CSV	21
2.4	<i>Business Intelligence</i>	21
2.4.1	OLAP	22
<u>2.4.1.1</u>	<u>Multidimensionalidade</u>	<u>22</u>
<u>2.4.1.2</u>	<u>Drill Up e Drill Down</u>	<u>22</u>
2.4.2	Advanced Analytics	23
2.4.3	Corporate Performance Management	23
2.4.4	Real Time BI.....	23
2.4.5	Data Warehouse e Data Marts	23
2.5	Técnicas de BI	24
2.6	<i>Storytelling</i>	25
2.7	<i>Crisp-DM</i>	26
2.8	Plataformas de BI	28
2.8.1	Power BI.....	29
<u>2.8.1.1</u>	<u><i>Power Query</i></u>	<u>29</u>
<u>2.8.1.2</u>	<u>Linguagem M</u>	<u>30</u>
<u>2.8.1.3</u>	<u>Tabelas</u>	<u>31</u>
<u>2.8.1.4</u>	<u>Visuais</u>	<u>31</u>
<u>2.8.1.5</u>	<u>Relações</u>	<u>32</u>
<u>2.8.1.6</u>	<u>Linguagem DAX</u>	<u>32</u>
2.9	Aplicação	33
3	PROCEDIMENTOS METODOLÓGICOS	34

3.1	Crisp-DM	35
3.2	Necessidades do processo	35
3.3	Dados disponíveis	36
3.3.1	Definição dos dados de entrada mínimos necessários	36
3.4	Definição da ferramenta a ser utilizada	37
3.5	Modelagem das relações multidimensionais	37
3.6	Definição dos dados tratados a serem utilizados	38
3.7	Definição e modelagem das métricas a serem utilizadas	38
3.8	Definição dos visuais que melhor representem os dados	39
3.9	Aplicação das ferramentas de <i>Storytelling</i>	39
3.10	Construção dos <i>dashboards</i>	40
3.11	Consolidação dos passos desenvolvidos nas etapas anteriores	40
3.12	Desenvolvimento de um questionário para validação do processo	40
3.13	Validação do processo com base nas respostas do questionário ..	41
3.14	Revisão do processo com base nos <i>feedbacks</i> recebidos	41
3.15	Avaliação da metodologia	41
4	RESULTADOS	42
4.1	Dados disponíveis	42
4.1.1	Projetos	42
4.1.2	Demanda	43
4.1.3	Capacidade	43
4.2	Relações multidimensionais	44
4.3	Modelagem dos dados no <i>Power Query</i>	44
4.3.1	Projetos	46
4.3.2	Demanda	47
4.3.3	Capacidade	48
4.4	Correlação entre as tabelas no <i>Power BI</i>	49
4.4.1	Tabelas de dimensão	49
<u>4.4.1.1</u>	<u>Calendário</u>	<u>49</u>
<u>4.4.1.2</u>	<u>Categorias</u>	<u>49</u>
<u>4.4.1.3</u>	<u>Países</u>	<u>50</u>
<u>4.4.1.4</u>	<u>Setores</u>	<u>50</u>
4.4.2	Modelagem e correlações	51
4.5	Criação de medidas com base nas tabelas calculadas	51
4.5.1	Carga de trabalho	51

4.5.2	Eficiência	52
4.5.3	Retorno dos maiores projetos	52
4.6	Desenvolvimento dos <i>dashboards</i> com base nos dados	52
4.6.1	Sumário	53
4.6.2	Financeiro.....	54
4.6.3	Maiores projetos	54
4.6.4	Gráficos de demanda	55
4.6.5	Tabelas de demanda.....	56
4.6.6	Extração de dados brutos.....	56
4.7	O uso do <i>Storytelling</i>	57
4.7.1	Desenvolvimento das telas de fundo.....	57
4.7.2	Posicionamento dos elementos gráficos	58
4.7.3	Segmentadores de dados	58
4.7.4	Dicas de ferramenta	59
4.7.5	Conclusão do desenvolvimento.....	60
4.8	Desenvolvimento do processo.....	60
4.9	Avaliação do questionário.....	60
4.10	Revisão do Processo	63
5	CONCLUSÃO	64
	REFERÊNCIAS.....	66
	APÊNDICE A - PROCEDIMENTO PARA ANÁLISE DE PROJETOS ...	69
	APÊNDICE B - QUESTIONÁRIO DE VALIDAÇÃO DO PROCEDIMENTO	73

1 INTRODUÇÃO

O número cada vez maior de dados aos quais estamos expostos não necessariamente implica em um maior conhecimento, ou uma melhora na tomada de decisões. Tão importante quanto coletar e possuir dados, é transformá-los em informação para que se possa extrair conhecimento e fazer análises relevantes deles. Esse é o resultado de uma gestão do conhecimento eficaz dentro de uma empresa (TERRA, 2005).

Ainda na década de 90, quando o capitalismo se torna mais competitivo devido ao avanço da globalização, ocasionando em um salto de crescimento das grandes indústrias, já se via a importância e o diferencial competitivo que uma gestão do conhecimento trazia para o negócio. Para Nonaka e Takeuchi (1997), a criação do conhecimento organizacional é o que tornou possível o sucesso das empresas japonesas estudadas por eles naquela época, e ela se fundamenta na habilidade que as empresas possuem em produzir conhecimento, compartilhar em seus setores e introduzi-los nas suas produções.

Com o avanço da indústria 4.0 e do *Big Data* (grandes bases de dados), ter o processo de transformação desses dados em informação e posteriormente em conhecimento, de uma forma cada vez mais rápida, e até mesmo, instantânea, é o que as empresas mais buscam para que as tomadas de decisões e a distribuição do conhecimento dentro da organização seja ágil (dados em tempo real), que é o foco da quarta revolução industrial, tornar os processos mais produtivos e eficientes.

O objetivo deste estudo é então, utilizando ferramentas de BI (*Business Intelligence*), construir e avaliar informações a respeito de times que desenvolvem projetos em indústrias de bens de consumo, a fim de que, utilizando meios de sistemáticos no tratamento de dados, desenvolver um roteiro para análise condicionada transformando-os em informação.

A implementação de um *dashboard* (painel de visualização que contém informações, métricas e indicadores) tornará o processo mais funcional e autoexplicativo, facilitando propor métricas que ajudem na tomada de decisões do negócio, prevendo o impacto de cada projeto em número de trabalhadores e impacto financeiro para as organizações, através da padronização de técnicas e maior agilidade na análise de dados, comparado com métodos manuais, ao longo do tempo.

1.1 Caracterização do problema

Organizações que estão na vanguarda da tecnologia, em busca de processos mais dinâmicos e inovações constantes, mantendo para isto times focados em desenvolver e testar novas estratégias para a empresa. Para isso, existem muitos investimentos e se faz necessário avaliar a eficácia destes grupos.

Para mensurar esta eficácia, validar hipóteses ou entender correlação de dados é necessário um processo padronizado de análise dos dados disponíveis, desta maneira, para o desenvolvimento deste trabalho, uma base de dados de projetos baseada em uma indústria de bens de consumo foi adaptada, de maneira a não apresentar informações específicas e sensíveis a respeito das empresas ou grupos representados pela mesma, contudo, mantendo as informações relevantes para uma avaliação, como dados financeiros, localizações, categorias e recursos alocados por área ao longo do tempo.

Estes dados são referentes a milhares de projetos que foram executados nos últimos 2 anos, e estão atualmente alocados em diversas tabelas.

Com os dados contidos nessa base, é possível definir as entradas necessárias, e como transformá-las em informação, a fim de facilitar futuras tomadas de decisão, que servirão de base para o desenvolvimento de um roteiro padronizado para análise de bases similares das indústrias.

1.2 Objetivos

Buscando aplicar estes conceitos caracterizados anteriormente, detalha-se aqui, os objetivos gerais e específicos que balizaram o desenvolvimento deste trabalho.

1.2.1 Objetivo geral

Propor um processo para a análise de dados que permita uma avaliação rápida e consistente a respeito de indicadores utilizados no desenvolvimento de projetos, utilizando ferramentas de *Business Intelligence* e *Big Data*, sendo balizado pela metodologia Crisp-DM.

1.2.2 Objetivos específicos

- Selecionar uma base de dados sobre projetos como exemplo;

- Entender e aplicar as etapas do processo metodológico Crisp-DM;
- Definir os principais dados de entrada necessários;
- Modelar e correlacionar os dados de entrada;
- Desenvolver *dashboards* que permitam uma análise gerencial destes dados;
- Construir um roteiro detalhando as etapas de entrada, modelagem e construção de *dashboards* que se aplique a bases semelhantes do segmento;
- Desenvolver um questionário qualitativo que busque entender as melhorias e vantagens que o roteiro desenvolvido desempenha nas organizações;
- Avaliar, através das respostas do questionário, a aplicabilidade do processo desenvolvido para análise de dados construída com base na metodologia Crisp-DM para avaliar projetos em organizações.

1.3 Justificativa

Proporcionar maneiras de facilitar o tratamento e interpretação de dados, a fim de melhorar a tomada de decisões estratégicas no segmento de bens de consumo, pois geralmente, em grandes empresas do setor, observa-se que parte do trabalho de análise de desempenho é realizado de maneira manual, consolidando as informações e dados em planilhas, sem informações de fácil visualização.

Grandes indústrias possuem inúmeras bases de arquivos que contém diversas informações. Estes dados são, em geral, organizados e concatenados em um mesmo registro, que após passar por um tratamento de dados e análise, demonstra os resultados daquela organização, porém, para chegar nesse fim, geralmente os responsáveis executam estas análises de maneira manual.

Este procedimento é complexo e exige alto nível de conhecimento das pessoas que o executam, assim como aloca tempo de trabalho em atividades repetitivas sem valor agregado, que pode contar com diversas falhas de execução e padronização. Desta maneira, este trabalho visa, com o auxílio de ferramentas de BI, desenvolver um roteiro padronizado para automatizar o tratamento destes dados, com a construção de *dashboards* para avaliação destes indicadores ao longo do tempo, o que pode impactar no retorno financeiro de empresas deste segmento, possibilitando

que as informações possam ser disponibilizadas em tempo hábil e permitam que as melhores decisões estratégicas sejam tomadas.

1.4 Etapas do trabalho

A introdução demonstrou o contexto e caracterização do problema, que geraram o ambiente que este trabalho irá atuar, bem como descreveu a proposta de objetivos gerais e específicos em que este modelo de BI estará focado.

Na fundamentação teórica, contextualiza-se onde estamos na revolução industrial, conceituando *Big Data* e *Business Intelligence*, bem como define as técnicas de BI e *Storytelling*, utilizando a metodologia Crisp-DM, uma plataforma de BI e bases comuns às indústrias de bens de consumo.

Na etapa de procedimentos metodológicos, define-se a sequência de ações necessárias para o desenvolvimento deste trabalho, trazendo mais detalhes das ferramentas e metodologias utilizadas.

Com base nos procedimentos metodológicos definidos, nos resultados se explicita-se a execução do fluxograma de desenvolvimento proposto, que culmina no roteiro de desenvolvimento, em conjunto com a elaboração e avaliação do questionário de validação do roteiro.

Na etapa de conclusão, serão comentados a respeito dos resultados alcançados, bem como propostas para continuidade do referido trabalho.

2 FUNDAMENTAÇÃO TEÓRICA

Para se iniciar o processo de elaboração deste trabalho, é necessário, previamente, entender os conceitos e métodos que norteiam o desenvolvimento científico da pesquisa e embasamento seguido, de maneira a explicitar todas as suas esferas. Para isso, a princípio é apresentado o momento atual das revoluções tecnológicas e industriais, assim como o conceito de *Big Data*, devido a possibilidade de avaliação de quantidades massivas de dados, juntamente com *Business Intelligence*, suas técnicas e ferramentas.

O conceito de *Storytelling* é uma parte muito significativa, pois é a partir dos seus princípios que é possível construir uma estrutura lógica e que permita fácil compreensão das informações que serão apresentadas no *dashboard*.

É necessário explicitar a metodologia Crisp-DM utilizada para o processamento dos dados, norteando a construção do procedimento padronizado a ser gerado.

2.1 Revoluções tecnológicas industriais

Segundo Stearns (2020), em sua obra *the industrial revolution in world history*, a revolução industrial foi o mais importante passo de desenvolvimento da história humana dos últimos três séculos, tão importante, que não pode ser considerado em uma data ou episódio único. Mohajan (2019) relata que a revolução iniciada em 1750, continua a moldar o mundo em que vivemos atualmente, alterando relações econômicas globais, relações trabalhistas e maneira de viver.

A primeira revolução industrial, que se iniciou na Inglaterra no setor têxtil no século XVIII e depois se espalhou para os demais países. No qual, as relações “homem *versus* campo” se transformaram e o homem passa a ocupar mais espaços nas cidades. Com isso, como aponta Cavalcante e Silva (2011), o capitalismo comercial, já iniciado, começa a requerer maiores demandas de produtos e de uma maneira rápida, motivando assim o surgimento do capitalismo industrial. É nesse momento em que as máquinas à vapor aprimorado por James Watt, através da geração de energia oriunda do carvão, tomam destaque dentro das indústrias, possibilitando maquinizar os processos.

A segunda revolução industrial se inicia no século XIX, a partir do ano de 1850, em um cenário, de acordo com Lopes, Garcia e Assumpção (2020), em que sai do

período de criação de máquinas e motores simples, que marcaram a primeira revolução industrial, e entra em projetos que apresentam uma relação maior entre ciência e tecnologia, proporcionando a criação de trens, desenvolvimento de motores à combustão interna com a utilização de petróleo e, até mesmo, o surgimento das indústrias de base, conhecidas como siderúrgicas e indústrias químicas, que substituíram a utilização do ferro pelo aço. Foi também durante esse período em que a eletricidade começou a ser utilizada como fonte de energia nas indústrias, como conta Dathein (2003), que além de possibilitar a criação dos motores elétricos, ainda mudou o cotidiano das pessoas e facilitou a inserção de pequenas empresas no mercado.

Com o consumo em massa e o patamar tecnológico que essas empresas alcançaram com a inserção dessas novas tecnologias mostrou-se a necessidade de grandes produções e, segundo Wood Jr. (1992), foi nesse período que surgiu Henry Ford com seus conceitos de linha contínua de produção, que promoveram produtividade e redução de custos.

Com o fim da Segunda Guerra Mundial, e por volta de 1950, surge a terceira revolução industrial. Nela se apresentam mudanças profundas envolvendo a robótica e a eletrônica que proporcionaram grandes avanços no meio industrial, como caracteriza Lopes, Garcia e Assumpção (2020), modernizando e facilitando os processos de fabricação. O que mais se destaca no período é a criação de computadores e *softwares*, além das evoluções envolvendo transporte e telecomunicações. Todos esses avanços foram essenciais para chegar ao processo de globalização.

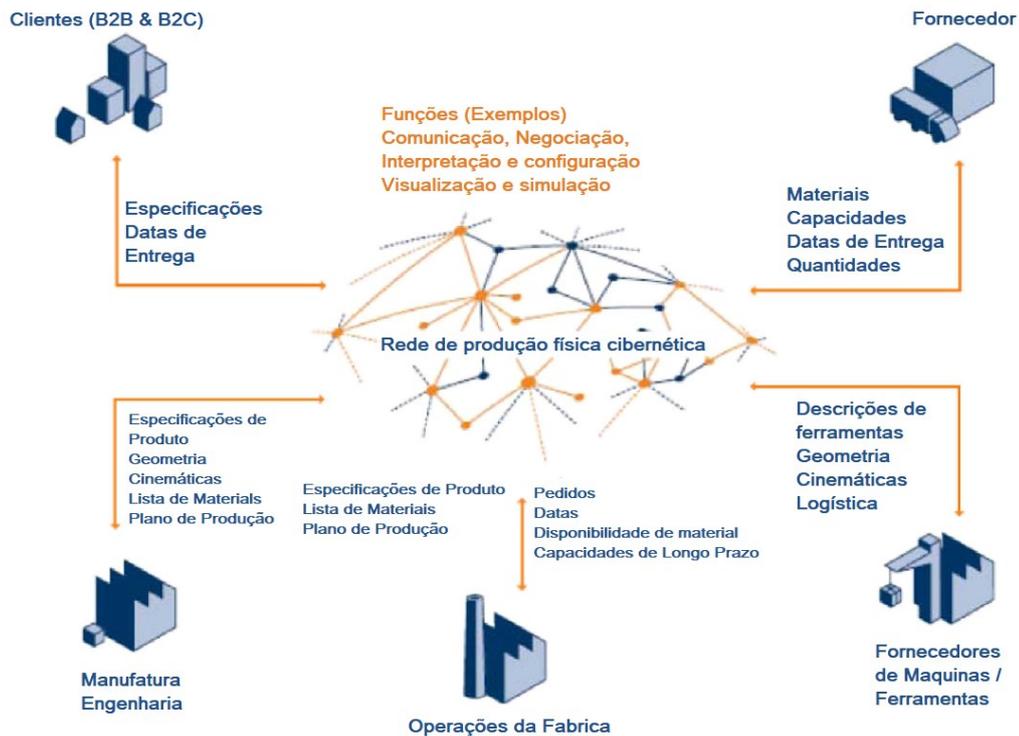
2.2 Indústria 4.0

A quarta revolução Industrial é a mais recente das revoluções, comumente conhecida por Indústria 4.0. Segundo Culot *et al.* (2020), o termo foi definido em 2011, por um time de cientistas durante pesquisas para o Ministério de Educação e Pesquisa do governo alemão, que buscava promover a computadorização da manufatura. De acordo com Sacomano *et al.* (2018), o governo estabeleceu o projeto Platform Industrie 4.0 durante a feira de Hannover, no qual, tinha como propósito desenvolver a tecnologia necessária para que as máquinas e equipamentos na indústria pudessem se comunicar entre si e trocar informações, de maneira a otimizar e automatizar processos.

Para Schwab (2017), o que torna a quarta revolução diferente das outras é a maneira como existe e interage a união entre a tecnologia e as funções físicas, biológicas e digitais. Com a globalização, a técnica se difundiu mais rapidamente, e a inovação se tornou uma cultura, intitulada pelas indústrias e adotadas pelo homem.

Lasi *et al.* (2014), determinaram os conceitos fundamentais da Indústria 4.0, diferente de iniciativas isoladas e bem estabelecidas em revoluções anteriores, a quarta revolução industrial pode se referir a uma ampla gama de conceitos distintos, porém com comportamentos chaves padrões, como fábricas inteligentes e autônomas, sistemas de inteligência artificial, mudanças nos sistemas de distribuição, adaptação às necessidades humanas e responsabilidade corporativa social. A Figura 1 demonstra as consequências desta integração dos diferentes componentes da cadeia de suprimentos.

Figura 1 - Interdependência da cadeia de suprimentos no contexto da indústria 4.0



Fonte: Adaptado de Lasi et al. (2014)

A Figura 1 é um demonstrativo das oportunidades que a indústria 4.0 como tecnologia proporciona a indústria. Todos esses tópicos relacionados a imagem geram dados possíveis de serem processados, lidos e analisados, dentro da necessidade exigida pelo homem e pelo negócio. É possível observar pelas redes, como a digitalização da informação viabiliza conexões entre setores e possibilidades em diferentes locais.

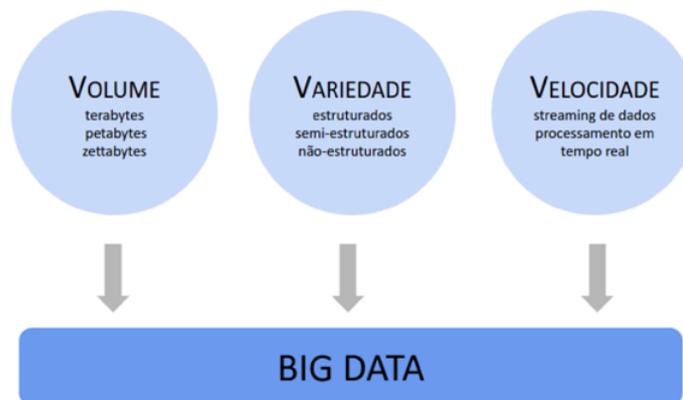
A indústria 4.0 incentiva a troca de dados que, após passar por tratamento em *softwares* e análise do homem, se tornam informações. Porém, para que isso aconteça, a quantidade de dados armazenada no sistema modular descentralizado (ou simplesmente, *cloud*) é muito densa, e então surge o termo *Big Data*.

2.3 Big Data

Big Data é um fenômeno em que há uma geração ampla de dados, pois os dispositivos e equipamentos possuem um nível de processamento que permite produzir rapidamente esses dados em diversos formatos e armazená-los, de maneira estruturada ou não (AMARAL, 2016).

Na opinião de Marquesone (2016), o que fez com que ocorresse esse volume de dados significativos é uma maior utilização de dispositivos móveis com poder de processamento, além da utilização da internet das coisas, em contrapartida de baixos custos para o armazenamento desses dados. Ela ainda define a composição de *Big Data* a partir de 3 Vs, como demonstrando na Figura 2.

Figura 2 - Os 3 Vs do Big Data



Fonte: Adaptado de Marquesone (2016)

Para Marquesone (2016), Volume é a principal característica do *Big Data*, representando a quantidade. A Variedade é determinada conforme as condições que eles estão apresentados em uma tabela, eles podem estar definidos através de esquemas rígidos, pré-definidos ou os não estruturados. Se tratando da Velocidade, se refere a rapidez para gerar, coletar, fazer as corretas correlações e analisar esses dados, e de maneira eficaz. Todos esses Vs compõem o *Big Data*.

Na indústria, os métodos que envolvem o *Big Data* fornecem novas oportunidades e um campo antes inexplorado de inovação e novas tecnologias (SENA, 2019). A combinação de múltiplas informações admite análises de modo a direcionar a produção para um modelo enxuto e flexível ao mercado, permitindo uma melhor adaptação na interação homem máquina. Com isso, o conhecimento que é proporcionado pode potencializar uma maximização dos ganhos e diminuição de riscos.

Para conseguir absorver toda a demanda pela inteligência de negócios e análise de dados que acompanham a quarta revolução industrial, o *Big Data*, termo descrito por Schermann *et al.* (2014) resume todo o desenvolvimento tecnológico na área de armazenamento e processamento de dados, que possam oferecer a possibilidade de tratar o aumento exponencial no volume de dados apresentados, de qualquer tipo, em um intervalo de tempo cada vez menor.

Um ponto importante quando se trata de *Big Data*, são os formatos para compartilhamento de dados entre aplicações. Para contexto e uso neste trabalho, comentaremos sobre arquivos CSV - *Comma-Separated Values*.

2.3.1 CSV

Segundo a IETF - *Internet Engineering Task Force* (2005), instituição normatizadora do formato, através da instrução RFC 4180, o CSV é um protocolo para transferência e conversão de dados entre programas que permite uma ampla gama de usos, salvando arquivos em formatos padronizados através de linhas e caracteres separadores, por padrão, a vírgula.

2.4 *Business Intelligence*

Negash e Gray (2008) citam, em sua obra sobre definições atuais a respeito do significado do termo *Business Intelligence*, ou sua sigla, BI, que se pode classificar como um mecanismo de suporte a tomada de decisões com base em dados, onde são reunidos a coleta de dados, o armazenamento e todo o gerenciamento do conhecimento com bases nestas entradas de dados, que possibilitaram decisões de negócio. Este termo, que começou a ser utilizado em 1989, permeia cada vez mais no dia a dia das organizações e, a longo prazo, novas técnicas e metodologias de BI serão incorporadas em todos os processos e negócios.

Ranjan (2009), define os componentes básicos para o desenvolvimento e implementação de um projeto utilizando BI. São eles, o OLAP (*Online Analytical Processing*), *Advanced Analytics*, *Corporate Performance Management*, *Real Time BI* e os *Data Warehouses* e *Data Marts*.

2.4.1 OLAP

OLAP é um conceito tecnológico utilizado para organizar bancos de dados de grandes empresas e dar suporte a trabalhos de BI, significa *Online Analytical Processing* (Processamento Analítico Conectado) e é atribuído ao modo como o usuário consegue dividir em dimensões os dados, como hierarquia ou tempo, utilizando a ajuda de ferramentas. Ele é empregue em análises e relatórios, e modelado para que o sistema possa fornecer visualizações resumidas e multidimensionais (RANJAN, 2009).

O conceito OLAP pode ser usado com *Data Warehouse* e *Data Marts*, além de outras ferramentas de BI que são utilizadas para fazer análises e armazenamento de dados. Empregadas no meio empresarial para definir tendências e determinar fatores críticos, seus relatórios fornecem uma base de dados confiável para uso em tomadas de decisões em negócios.

2.4.1.1 Multidimensionalidade

A análise de multidimensionalidade de dados de um *Data Warehouse*, se refere, segundo Tetila (2014), ao modelo de relações baseado em tabelas fato e dimensão. No qual, dentro dos dados existem associações que podem ser interligadas de maneira simples ou complexas, para que a informação chegue completa na representação visual, e elas podem estar resumidas conforme o contexto temporal, geográfico ou em outro nível de detalhamento.

2.4.1.2 Drill Up e Drill Down

Drill Up e *Drill Down*, de acordo com Funes (2017), é a forma como os dados estão detalhados dentro do *Data Warehouse*, sendo que podem estar menos detalhados (*Drill Up*) ou mais (*Drill Down*), fazendo com que a varredura de dados ocorra em dimensões mais profundas ou não.

2.4.2 Advanced Analytics

Definido por Ranjan (2009), *Advanced Analytics* é simplesmente uma análise avançada de mineração de dados, no qual, utiliza técnicas de estatística para uma previsão de negócio, fornecendo graus de certeza sobre possíveis cenários futuros.

Segundo Camilo e Silva (2009), a mineração de dados é um assunto multidisciplinar, mas que possui como seus principais fatores: método estatístico, aprendizado de máquina e banco de dados. Sendo, que o objetivo da mineração é encontrar relações, que podem ser inesperadas, a fim de resumir de uma forma compreensível aos *stakeholders* as informações contidas na base.

Para o autor Camilo e Silva (2009), o processo de mineração de dados é necessário para se ter um caminho na direção da busca do conhecimento dentro do negócio, e uma maneira em que se pode fazer isso é utilizando a metodologia CRISP-DM.

2.4.3 Corporate Performance Management

Ranjan (2009) afirma ser o local onde os dados podem ser agrupados e apresentados através de padrões visuais relacionados, como: *dashboards*, *scorecards* e *portals*. Essas estruturas permitem um melhor entendimento e leitura de dados, a partir do que é montado e associado pelo analista responsável.

Suas representações visuais em *dashboards*, por exemplo, determinam para os *stakeholders* as informações importantes para a análise e conhecimento atual do negócio.

2.4.4 Real Time BI

É associado a análise e distribuição de métricas em tempo real dentro do sistema *corporate performance management*, como descreve Ranjan (2009). Sistemas de *business intelligence* permitem que aconteça essa atualização dos dados em *real time*, desde que sejam alimentadas de maneira correta pela nuvem de dados, permitindo que isso aconteça.

2.4.5 Data Warehouse e Data Marts

Data Warehouse é o sistema dentro de BI que contém esses dados, vindos de diversas fontes, de modo que podem ser traduzíveis em informações em sua

operação e lidando com a propagação desses muito registros atualizados, de acordo com Ranjan (2009). Ele ainda contém muitos dados ativos, históricos e, até mesmo, os não instantâneos. Podendo estar inclusos dados de operações e externos. Ele suporta vários tipos de multimídias, no qual podem estar contidos em tabelas, textos simples ou imagens.

Já o *Data Marts* seria a divisão em setores ou departamentos dessas mesmas composições de dados. É como se o *Data Warehouse* seja o todo dentro de uma empresa e o *Data Marts* a porção no qual se deseja um maior foco, ou naquele momento é mais interessante que seja feita a análise de acordo com o futuro da companhia. Segundo Ranjan (2009), eles são específicos e, até mesmo, peculiares, pois cada parte dentro da empresa contém as suas necessidades.

A diferença entre as duas definições é que o *Data Marts* é determinado conforme a necessidade do meio, e é predefinida em um determinado agrupamento de dados dentro do departamento característico, fornecendo suporte a uma determinada função. Enquanto o *Data Warehouse* não existe essa distinção de especificações.

2.5 Técnicas de BI

Ranjan (2009) define as técnicas de BI comumente utilizadas, demonstradas no Quadro 1.

Quadro 1 - Técnicas de BI Atuais

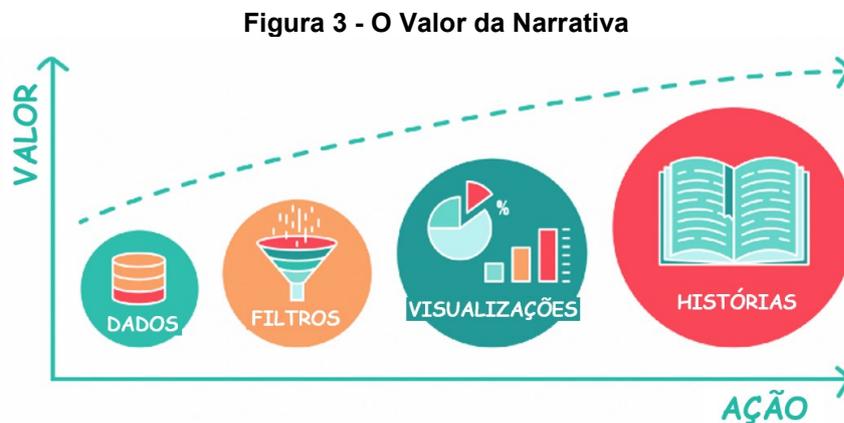
Técnica	Descrição
Modelagem Preditiva	Predição de valores para um atributo de dado específico
Caracterização e mineração descritiva de dados	Distribuição, dispersão e exceção dos dados
Associação, Correlação, Análise de Causalidade	Identificar relações entre atributos
Classificação	Determinar a que classe um item pertence
Clusterização e análise de <i>Outliers</i>	Particionar um conjunto em classes, onde itens com características similares são agrupados juntos
Análise de Padrões temporais e sequenciais	Tendência e desvios, padrões sequenciais, periodicidade
OLAP	As ferramentas OLAP permitem que os usuários analisem diferentes dimensões de dados multidimensionais. Por exemplo, fornecendo séries temporais e visualizações de análise de tendência
Visualização de Modelos	Tornar o conhecimento descoberto facilmente compreendido usando gráficos, histogramas e outros meios visuais
Análise Exploratória de Dados	Explorar um conjunto de dados sem uma forte dependência de suposições ou modelos; O objetivo é identificar padrões de forma exploratória

Fonte: Adaptado de Ranjan (2009)

O Quadro 1 descreve métodos estatísticos de mineração de dados, sendo elas: predição, caracterização dos dados minerados, as associações e relações formadas, os *clusters*, as análises temporais e sequenciais e, finalizando com a representação visual da análise exploratória apresentada, assim como, com a descrição de cada técnica.

2.6 Storytelling

Ramm *et al.* (2021) detalham, em sua pesquisa sobre o *Storytelling* aplicado ao BI, a importância das histórias e das narrativas, como demonstrado na Figura 3. Ao longo de toda nossa existência, fomos balizados pela transmissão de informações, histórias, experiências e ideias por meio da narrativa, e nada mais comum do que entender a necessidade desta narrativa atualmente, na análise de dados e compreensão do mundo moderno, com uma ferramenta que foi aprendida por nossos subconscientes durante o fluxo evolutivo humano. Ao contar histórias com os dados, geramos informação e entregamos o contexto necessário para análise dos dados, garantindo assim, maior assertividade na tomada de decisão.



Fonte: Adaptado de Paixão (2017)

Gessert (2016) argumenta que o *Storytelling* é uma habilidade essencial para o BI. Ela pondera que os *insights* encontrados nos dados de um analista podem parecer evidentes e óbvios para o analista, mas eles não devem presumir que seu público os compreenderá. Em vez disso, ela acredita que o profissional precisa colocar seus *insights* em um contexto de narrativa de dados, onde ele pode explicar o valor de sua lógica analítica e evidências para um público menos técnico, ao mesmo tempo em que o enquadra em um contexto que ajuda o público a entender sua relevância para o problema que estão tentando resolver.

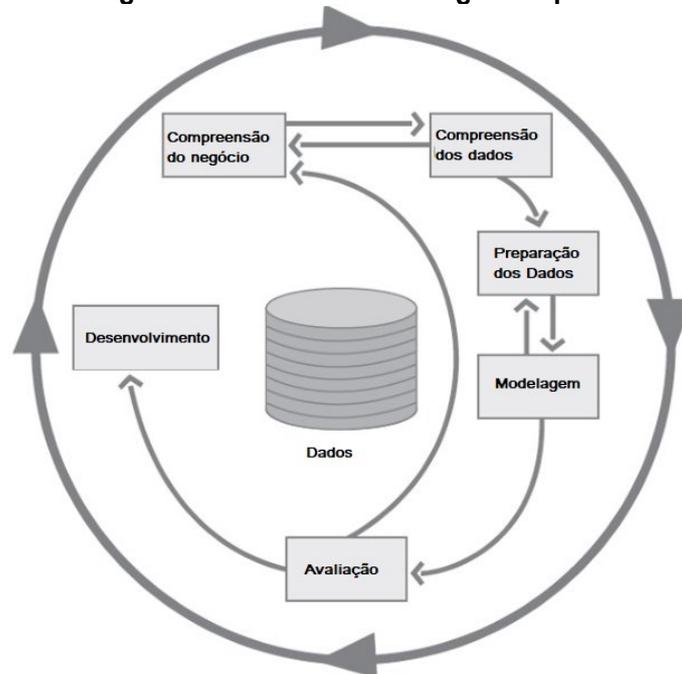
Paixão (2017) cita, como demonstrado na Figura 3, o valor desta narrativa, que visa demonstrar o que o contexto necessário trás, para que um dado se transforme em uma História, e bons dados contam boas histórias.

Storytelling é estruturado em técnicas que visam destacar pontos que, pensando no público-alvo de apresentação, irão agregar mais valor. É o procedimento que fundado em estratégia de visualização, é utilizado pelo analista a fim de contar a melhor história, a história que ele quer narrar e que sua plateia quer ouvir.

2.7 Crisp-DM

Para o desenvolvimento deste projeto, utilizou-se uma metodologia conhecida no ramo da análise de dados, chamada Crisp DM (*Cross Industry Standard Process for Data Mining*), ela visa transformar dados de uma empresa em conhecimentos e informações de gerenciamento. O sistema, que tem como base, um fluxo de melhoria contínua, foi criada a pouco mais de 20 anos, e cresce junto com a necessidade cada vez maior no tratamento de uma quantidade massiva de dados. Sua estrutura pode ser visualizada na Figura 4.

Figura 4 - Ciclo da metodologia Crisp-DM



Fonte: Adaptado de Lima (2020)

O princípio básico do Crisp-DM, como demonstrado na Figura 4, se inicia com a compreensão do negócio, conhecendo e entendendo todas as métricas associadas a empresa, bem como todas as regras de negócio, identificando assim os

especialistas, levantando todas as expectativas, *hardwares* e *softwares* disponíveis, inventariando bases de dados existentes e verificando a existência de bases de dados multidimensionais.

A seguir, para a segunda etapa, inicia-se a compreensão dos dados, etapa que normalmente é executada junto à primeira, levantando e compreendendo todos os dados que a organização necessita para prover o negócio com informações. Qual a quantidade, qualidade e granularidade das informações, onde elas estão disponíveis, qual a taxa de atualização e volume total a disposição.

Com estes dados mapeados, inicia-se a preparação, com tratamento e construção de modelos que possam adequar os dados nas informações finais necessárias. Nesta etapa, eles são organizados em estruturas conhecidas e definidas, selecionadas as colunas de interesse, realizado a limpeza em colunas e dados inconsistentes, adequado os formatos e cálculo de novos atributos com base nos dados originais.

Figura 5 - Etapas do Crisp-DM

Entendimento do Negócio	Entendimento dos Dados	Preparação dos Dados	Modelagem	Avaliação	Desenvolvimento
Determinação dos Objetivos	Coleta Inicial dos dados	Seleção dos dados	Seleção da técnica de Modelagem	Avaliação dos Resultados	Aplicação do Projeto
Situação a ser avaliada	Descrição dos Dados	Limpeza dos dados	Geração do projeto de teste	Revisão do Processo	Planejar Entrega
Determinar metas de mineração de dados	Exploração dos dados	Construção dos dados	Construção do Modelo	Determinar Proximos Passos	Plano de monitoramento e manutenção
Produzir o Projeto	Verificação da Qualidade dos dados	Integração dos dados	Modelo a ser avaliado		Produção do Relatório Final
		Formatação dos dados	Ajustes no Modelo		Revisão do Projeto
		Descrição das bases			Lições Aprendidas

Fonte: Adaptado de Lima (2020)

Com as bases tratadas e prontas para uso, inicia-se a etapa de modelagem, onde, por meio de diversas ferramentas, determinadas de modo empírico, são utilizadas para definição e calibração de parâmetros, que resultam em resultados estatísticos a partir dos dados inseridos. Com estes dados é realizada uma avaliação dos modelos e acuracidade, em um ciclo com a etapa quatro, de Modelagem, até atingir os resultados esperados, e por fim o sistema é implantado.

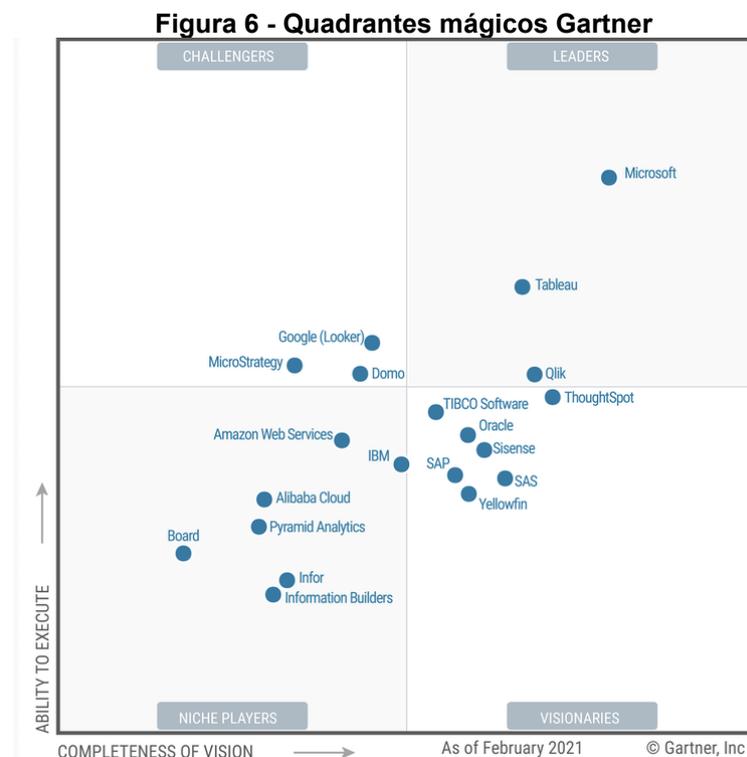
Um resumo destas etapas descritas pode ser visualizado na Figura 5, que define passos sugeridos para cada etapa. Em um primeiro momento, conhecer e

entender o negócio é essencial para então, determinar o que e como serão minerados os dados em busca do que se deseja apresentar ou relacionar. A preparação e construção das associações e modelos é crucial para se ter bons resumos dos dados. E na etapa de Avaliação, revisar o que foi feito nas etapas anteriores, finalizando com as representação visuais, de modo a expor os resultados encontrados.

2.8 Plataformas de BI

Quando se pensa em iniciar um projeto de BI, é importante entender qual ou quais plataformas estão disponíveis para planejamento e execução do ciclo de tratamento de dados. Para ajudar com esta busca, a Gartner (2021) analisa, em um relatório publicado anualmente, por meio da Metodologia *Magic Quadrant*, uma análise gráfica com o posicionamento competitivo dos quatro tipos de tecnologias dos mercados de rápido crescimento, sendo eles, Líderes, Visionários, Jogadores de Nicho e Desafiadores.

Na seção de *Business Intelligence Platforms*, o estudo mais recente é de 2021, e pode ser visualizado na Figura 6, que dimensiona a efetividade de uma plataforma de BI pela habilidade em executar e em ver o quadro geral.



Fonte: Adaptado de Gartner (2021)

Verificamos neste estudo que as empresas Microsoft, Tableau e Qlik são as empresas que lideram estas soluções na execução de BI. Dentre elas, a Microsoft lidera o *Ranking* a 14 anos consecutivos.

2.8.1 Power BI

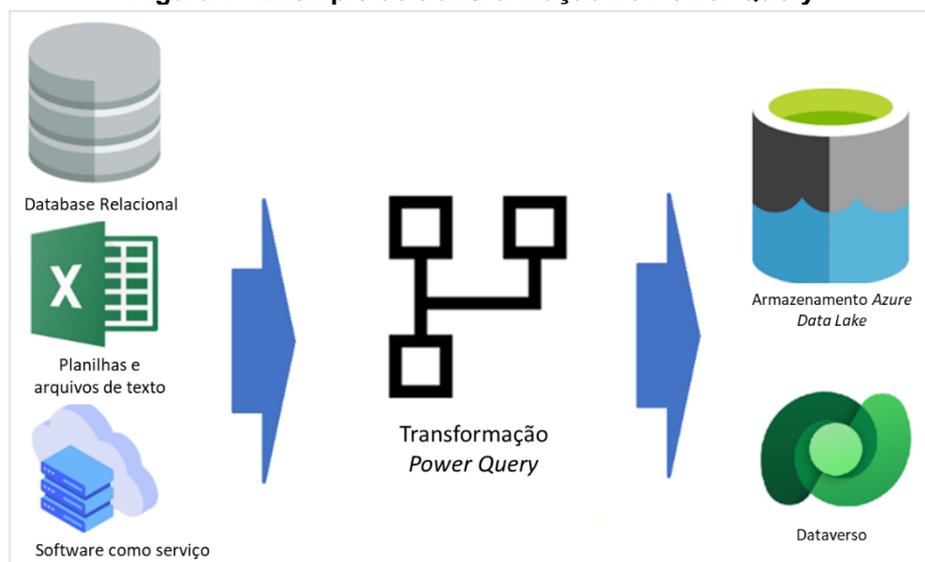
Segundo a Microsoft (2021), *Power BI* é uma coleção de serviços de *software*, aplicativos e conectores que funcionam juntos para tornar suas fontes de dados não relacionadas em informações coerentes, visualmente imersivas e interativas, permitindo que você se conecte facilmente às suas fontes de dados, limpe e modele seus dados sem afetar a fonte subjacente, visualize o que é importante e compartilhe isso com quem você quiser.

A Microsoft (2021), pontua que os três principais blocos de construção do *Power BI* são: conjuntos de dados, relatórios e *dashboards* que são então todos organizados em pastas compartilhadas.

2.8.1.1 Power Query

Power Query é o mecanismo de transformação de dados usado no fluxo de dados, como visualizado na Figura 7 e definido por Microsoft (2021).

Figura 7 - Exemplo de transformação no *Power Query*



Fonte: Adaptado de Microsoft (2021)

Esse mecanismo é capaz de dar suporte a muitas transformações avançadas nos dados, utilizando uma interface simples, mas poderosa, chamada *Power Query*.

Nesta interface, é possível desenvolver fluxos de dados para desenvolver soluções de integração de dados com mais rapidez e facilidade.

Na Figura 7 podemos entender que o *Power Query* é o sistema responsável por gerenciar um fluxo de informações, ou seja, tratar, de maneira ordenada e sistêmica, uma série de dados, transformando-os em outros. As instruções são atribuídas a uma coluna contendo o passo a passo, e podem ser aplicadas a qualquer outro dado posterior, automaticamente, além de funcionar automaticamente na nuvem, interpretando quaisquer novos dados sem necessitar de configurações manuais.

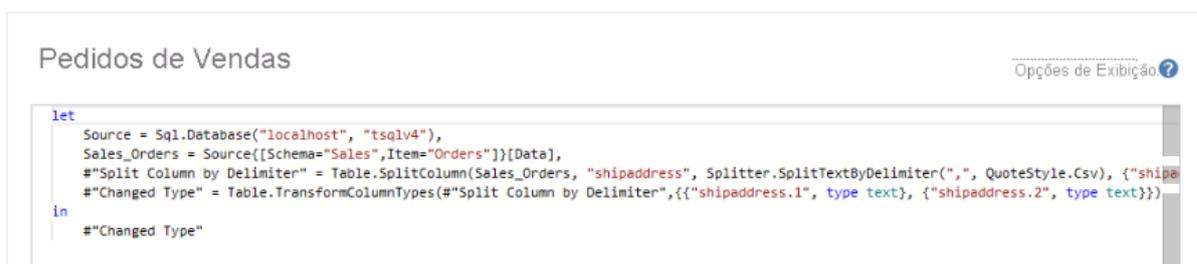
2.8.1.2 Linguagem M

Segundo explicação na documentação da Microsoft (2021), sempre que você formatar dados no *Power Query*, criará uma etapa no processo do *Power Query*, como demonstrado na Figura 8.

Figura 8 - Linguagem M



A tela a seguir deverá ser exibida.



Fonte: Adaptado de Microsoft (2021)

Essas etapas podem ser reordenadas, excluídas e modificadas, quando apropriado. Cada etapa de limpeza gerada pode ser criada a partir da interface gráfica, mas o *Power Query* usa a linguagem M para realizar estas transformações. As etapas combinadas estão disponíveis para leitura por meio do Editor Avançado do *Power Query*. A linguagem M sempre está disponível para ser lida e modificada diretamente. Como cada etapa no *Power Query* é escrita em código M, mesmo que o passo tenha sido criado utilizando a interface gráfica, a mesma pode ser acessada em

sua versão codificada neste menu avançado, e personalizado de acordo com as necessidades do projeto, como demonstrado na Figura 8.

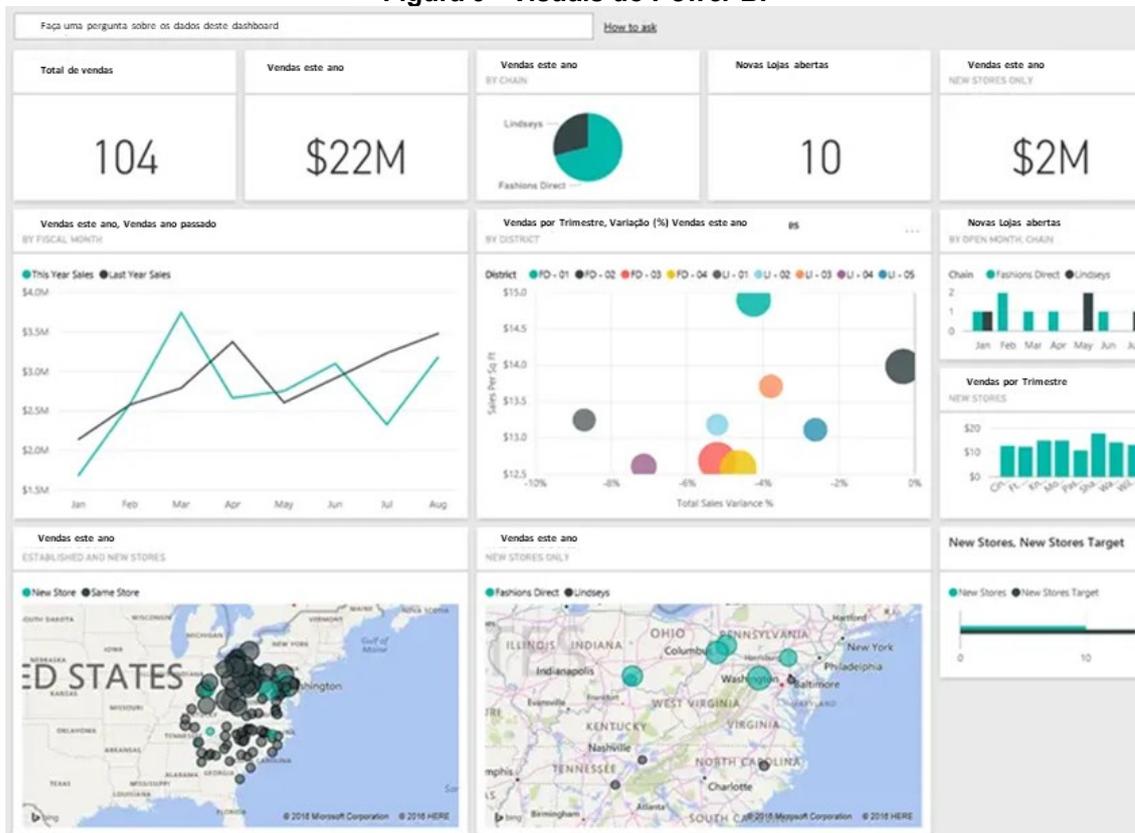
2.8.1.3 Tabelas

As tabelas são os elementos bases que serão utilizadas no *Power BI* e são definidas por Microsoft (2021) como sendo o resultado das transformações realizadas no *Power Query*, e contém os dados tratados e organizados em colunas com o formato que se adeque aos dados.

2.8.1.4 Visuais

Os Visuais são os elementos gráficos que irão demonstrar os dados das tabelas, e por onde o usuário irá interagir com o *dashboard*. Gráficos, Filtros, Imagens, Mapas e *Cards* são exemplos de visuais, e parte fundamental do programa. Alguns exemplos são demonstrados na Figura 9, adaptada de Microsoft (2021).

Figura 9 - Visuais do Power BI

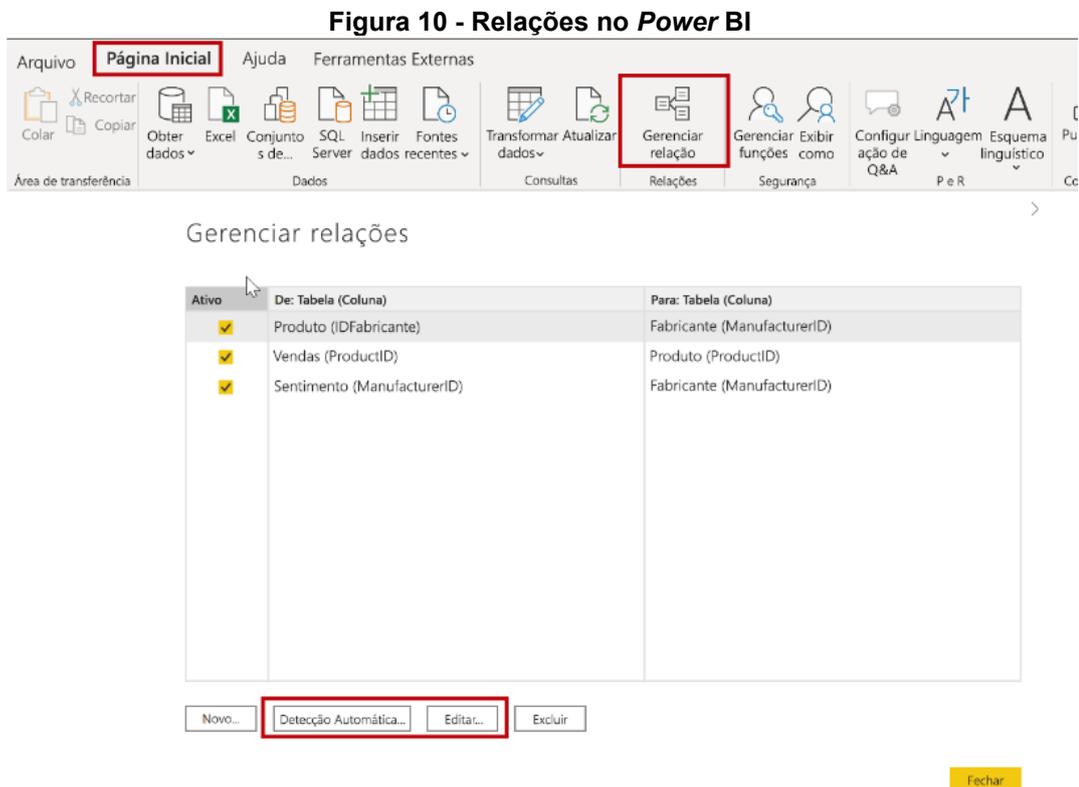


Fonte: Adaptado de Microsoft (2021)

2.8.1.5 Relações

A exibição de Modelo do *Power BI Desktop* permite definir visualmente o relacionamento entre tabelas ou elementos. Segundo a Microsoft (2021), uma relação é quando duas ou mais tabelas são vinculadas porque contêm dados relacionados. Isso permite que os usuários executem consultas de dados relacionados em várias tabelas. Dentro do *Power BI*, a seção de exibição de Modelo permite ver uma exibição diagramática dos dados.

A adição e remoção de relações pode ser realizada como demonstrado na Figura 10, e pode conter múltiplas opções de cardinalidade, além de permitir relacionamentos direcionais ou bidirecionais.

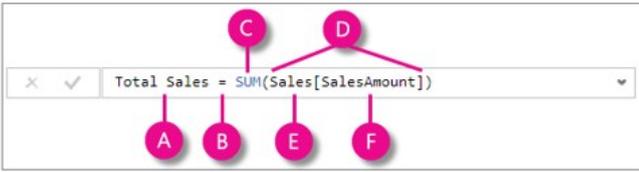


Fonte: Adaptado de Microsoft (2021)

2.8.1.6 Linguagem DAX

A linguagem DAX (*Data Analysis Expressions*) é uma coleção de funções, operadores e constantes que podem ser usados em uma fórmula, ou expressão, para calcular e retornar um ou mais valores. Resumindo, o DAX ajuda a criar informações de dados já presentes em seu modelo. A Figura 11 traz um exemplo de como utilizar, com uma função de soma.

Figura 11 - Exemplo de aplicação DAX



Esta fórmula inclui os seguintes elementos de sintaxe:

- R.: O nome da medida, **Total Sales**.
- B. O operador de sinal de igual (=), que indica o início da fórmula. Quando calculada, ela retornará um resultado.
- C. A função DAX **SUM**, que soma todos os números da coluna **Sales[SalesAmount]**. Você aprenderá mais sobre as funções mais tarde.
- D. Os parênteses (), que envolvem uma expressão que contém um ou mais argumentos. A maioria das funções exige pelo menos um argumento. Um argumento transmite um valor para uma função.
- E. A tabela referenciada, **Sales**.
- F. A coluna referenciada, **[SalesAmount]**, na tabela **Sales**. Com este argumento, a função **SUM** sabe em que coluna deve agregar uma **SUM**.

Fonte: Adaptado de Microsoft (2021)

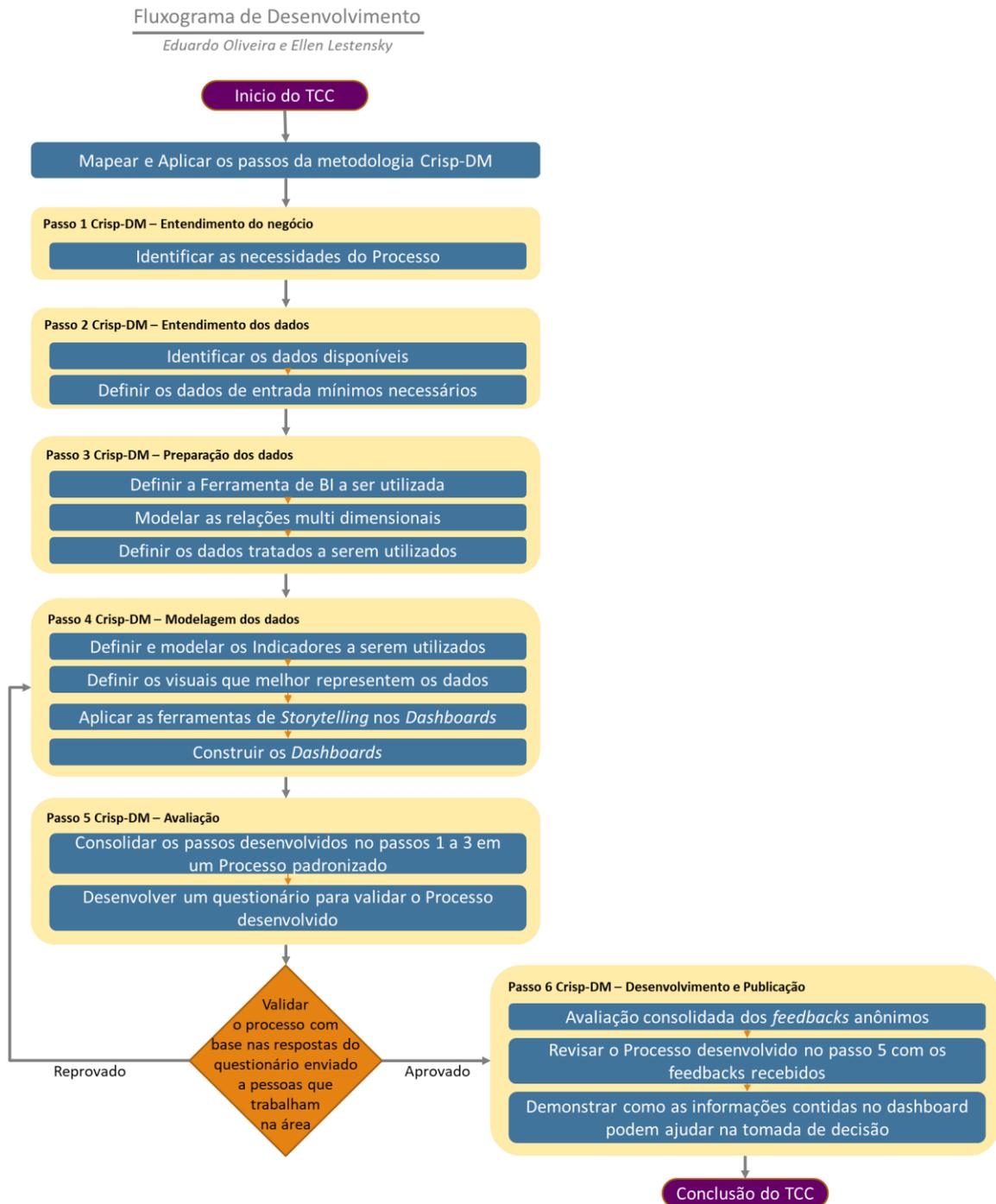
2.9 Aplicação

A partir da base teórica desenvolvida, relacionamos aos assuntos que permeiam o trabalho e que forneceram suporte ao desenvolvimento do projeto, com a aplicação da metodologia, elaboração do *dashboard* e criação do roteiro, que se demonstra na seção 3.

3 PROCEDIMENTOS METODOLÓGICOS

Neste capítulo, serão definidos os procedimentos necessários para execução deste trabalho. O fluxograma deste desenvolvimento pode ser verificado na Figura 12.

Figura 12 - Fluxograma de desenvolvimento do TCC



Fonte: Autoria própria (2022)

Inicialmente foram mapeados quais dados estavam presentes na base de dados e, seguindo a metodologia Crisp-DM no qual se baseia nos passos de: compreensão das necessidades do negócio, compreensão dos dados presentes,

preparação dos dados, modelagem, avaliação e desenvolvimento, foram fundamentadas as etapas metodológicas de elaboração do trabalho de conclusão de curso.

Sendo assim, no passo 1 foi identificado as necessidades que o processo deveria atender no final. No passo 2, além da identificação das bases de dados disponíveis, em alinhamento com os objetivos que foram constatados como importantes, também foi definido os dados de entrada mínimos necessários. No passo 3, na fase de preparação dos dados, foi estabelecido a ferramenta de *Business Intelligence* que seria utilizada, modelado as relações de multidimensionalidade e definindo os dados tratados que seriam aproveitados.

Na etapa de modelagem dos dados, descrita no passo 4, foram estipulados os indicadores, os visuais, e com o auxílio das ferramentas de *storytelling*, criados os *dashboards*. No passo 5, foi realizado uma avaliação do que foi construído consolidando os passos 1 a 3 e utilizando um questionário desenvolvido pelos autores a fim de validar o processo.

Com o resultado do questionário foi avaliado se o resultado era o desejado ou se havia alguma lacuna que não havia sido mapeada anteriormente. Se houvesse, o indicado seria retornar ao passo 4 e remodelar, se não, segue para o passo 6.

No passo 6, foi consolidado o resultado dos pareceres anônimos e revisado o passo 5. Com essas informações, foi demonstrado analiticamente como esses dados poderiam auxiliar na tomada de decisão, finalizando assim o desenvolvimento do projeto.

3.1 Crisp-DM

Utilizando as etapas destacadas na metodologia Crisp-DM, definidas no capítulo anterior, é possível definir o fluxograma de desenvolvimento nas etapas do Crisp-DM. Permitindo seguir todos os passos de maneira lógica e ordenada.

3.2 Necessidades do processo

Como descrito no objetivo principal deste trabalho, as necessidades principais são em relação ao desenvolvimento de um procedimento que permita às organizações realizarem, de maneira mais rápida, automatizada e confiável, análises de indicadores relacionadas ao desenvolvimento de projetos, utilizando bases de dados comuns ao

desenvolvimento de projetos em empresas de bens de consumo, permitindo assim que decisões gerenciais sejam tomadas com base em dados.

3.3 Dados disponíveis

O procedimento a ser desenvolvido precisa ser apto a funcionar em diversas bases, desta maneira, não serão explicitados aqui todos os dados que podem estar disponíveis, e sim uma lógica base que pode ser adaptada a diversas realidades, utilizando para isto, uma base comum, tratada.

Estes dados base são relativos a projetos com uma grande quantidade de dimensões. Estão disponíveis, para diversas combinações de Categoria, Países e Grupo de Recursos, os dados de Demanda, medidos em FTE (*Full Time Employee*), onde 1 FTE corresponde a demanda completa de um trabalhador em um intervalo de tempo, por Projeto, bem como as entregas financeiras e detalhes específicos de cada marca.

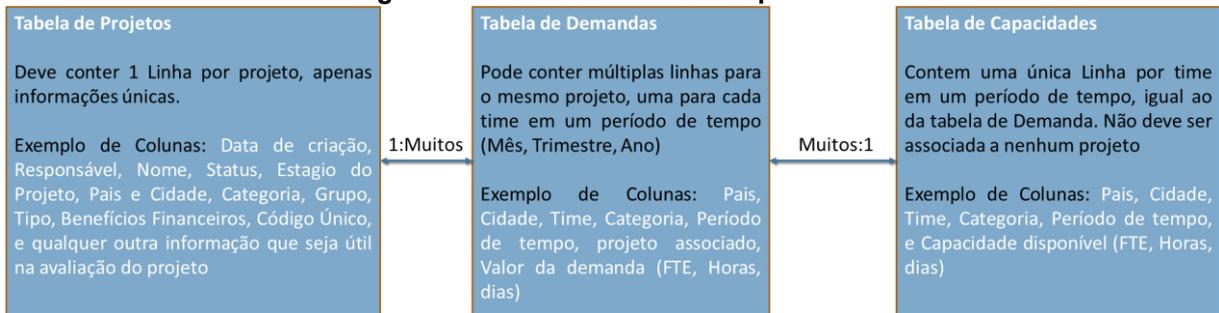
Estes recursos estão divididos em três tabelas, que devem trazer:

- Os Projetos, com todas as informações únicas associadas, como: dados financeiros, detalhamento de segmento, localidade geográfica, dentre outros;
- A demanda de um grupo, categoria e dados geográficos em um determinado projeto e trimestre, sendo uma informação avaliada em muitos para um, contra a tabela de projetos, ou seja, poderá conter múltiplas linhas para o mesmo projeto;
- A Capacidade total deste time para execução destes projetos, por sub-grupo, categoria e Cidade em um intervalo de tempo, de maneira igual a demanda.

3.3.1 Definição dos dados de entrada mínimos necessários

Desta maneira, de forma a demonstrar os dados mínimos disponíveis, foi desenvolvido, com base na avaliação da base comum, e disponibilizado na Figura 13, um resumo gráfico com as tabelas e dados mínimos necessários para a execução do procedimento proposto.

Figura 13 - Tabelas mínimas disponíveis



Fonte: Autoria própria (2022)

Ressalta-se que esta é apenas a estrutura base, e qualquer dado, em tabela ou coluna adicional, que possa enriquecer a avaliação, pode e deve ser adicionada e relacionada à modelagem base.

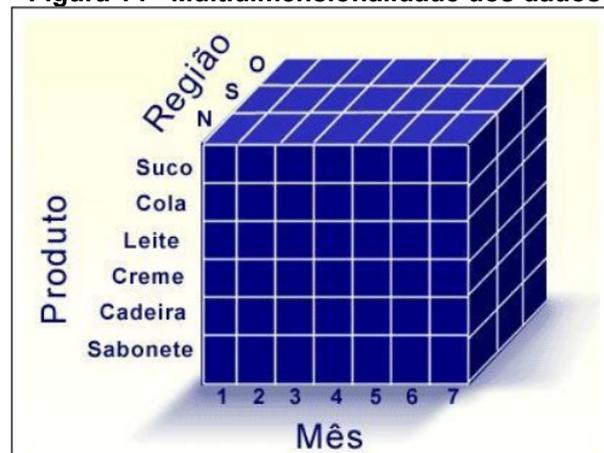
3.4 Definição da ferramenta a ser utilizada

A partir de uma avaliação das opções disponíveis em ferramentas de BI, do estudo disponibilizado pela Gartner (2021), em relação a custo, usabilidade, facilidade em escalar e implementação na organização, optou-se pelo programa da Microsoft, conhecido como *Power BI*. Suas principais vantagens são, além do preço competitivo, o fato de ser um ecossistema completo, onde é possível coletar os dados de maneira automática por meio de sua plataforma de automação, chamada *Power Automate*, guarda-los de maneira segura e com acesso apenas pelos colaboradores dentro da organização na nuvem interna da organização, utilizando *Onedrive* ou *Sharepoint*, e após isso realizar coletas de dados nativas usando o Microsoft *Power BI*, que realiza a modelagem dos dados automaticamente e em tempo real, salvando os resultados também na nuvem da empresa, garantindo assim, a segurança da informação fornecida internamente em todas as etapas e sem gastos adicionais.

3.5 Modelagem das relações multidimensionais

Utilizando um banco de dados do tipo OLAP, disponível no *Power BI*, pode utilizar a técnicas de modelagem multidimensional, que possibilitara ser avaliado diversas dimensões diferentes do banco de dados utilizado neste trabalho ao mesmo tempo. Date (2004) demonstra na Figura 14 um exemplo de relação multidimensional que avalia para um determinado dado, o mês, produto e região, com diversas combinações entre os dados.

Figura 14 - Multidimensionalidade dos dados



Fonte: Adaptado de Date (2004)

Avaliando o banco de dados, será utilizado principalmente as 8 dimensões abaixo, contudo, outras dimensões presentes poderão ser aplicadas no futuro também.

- Categoria do Projeto (Com *Drill Down* para sub-categoria);
- Categoria do Time da demanda;
- País do Projeto (Com *Drill Down* para Cidade);
- País do Grupo da Demanda (Com *Drill Down* para Cidade);
- Data de finalização do Projeto;
- Data da análise;
- Grupo de Projeto (Com *Drill Down* para sub-grupo).

3.6 Definição dos dados tratados a serem utilizados

Após a finalização das etapas de modelagem dos dados e correlação entre as tabelas, é relevante realizar uma limpeza da base, de maneira a remover informações que não são úteis, que podem estar repetidas ou serem um ruído nas avaliações.

3.7 Definição e modelagem das métricas a serem utilizadas

Os KPIs (*Key Performance Indicators*) serão escolhidos com base nas melhores combinações entre os dados modelados no Projeto, que impactem a demanda e o aspecto financeiro da organização.

Quando aplicado a algum caso ou empresa específica, os KPIs podem e devem ser adaptados aos objetivos que a situação está necessitando no momento,

neste momento deve ser avaliado criticamente quanto a demanda exigida de aplicação.

3.8 Definição dos visuais que melhor representem os dados

Para representação dos dados no *Power BI*, serão selecionados os gráficos padrões disponibilizados pela ferramenta que mais se adequem aos dados. Numerro (2021) classifica os principais tipos de visuais disponíveis na ferramenta, classificando-os em 6 categorias (Figura 15). Não necessariamente usaremos todos, mas buscaremos, utilizando a metodologia e técnicas do *storytelling*, entender os que melhores se enquadram em cada análise.

Figura 15 - Tipos de visuais padrões no *Power BI*



Fonte: Adaptado de Numerro (2021)

3.9 Aplicação das ferramentas de *Storytelling*

Para Auxiliar na construção dos *dashboards*, utiliza-se as técnicas definidas de *storytelling*, que contém a história por trás dos dados, utilizando cores que destaquem os dados que são mais representativos, fontes personalizadas a necessidade e os visuais que melhor demonstrem o conteúdo dos dados.

O *storytelling* passa por critérios como: evitar saturações desnecessárias, destacar o que o público-alvo tem interesse de saber ou conhecer e o que os autores

e modeladores dos dados entendem como importantes pensando nos objetivos de transmissão de informação que impactaram na razão em que o *dashboard* foi criado.

3.10 Construção dos *dashboards*

Utilizando o *Power BI*, com os dados modelados, é possível construir painéis personalizados e dinâmicos, que podem ser publicados em uma pasta compartilhada e ficar disponível para todos os membros da organização que tenham relação e interesse profissional com os dados.

3.11 Consolidação dos passos desenvolvidos nas etapas anteriores

Consolidando os tópicos 3.2 ao 3.10, objetiva-se construir, em um documento disponibilizado no Apêndice I, um roteiro com um procedimento passo a passo que permita que bases similares sejam analisadas utilizando o mesmo método da Pesquisa.

3.12 Desenvolvimento de um questionário para validação do processo

Como mecanismo para validar o Apêndice I e de acordo com o passo 5 do procedimento metodológico Crisp-DM descrito no fluxograma como avaliação, parte do trabalho foi o desenvolvimento de um questionário, a ser respondido de maneira anônima por pessoas que trabalhem na área de análise de dados ou tomada de decisão de empresas de bens de consumo, avaliando três pontos principais.

- O roteiro disponibilizado se adapta a sua base de dados?
- O *dashboard* desenvolvido melhora a eficiência e qualidade das análises atuais?
- O roteiro é aplicável a necessidades da sua função?
- Para classificar e agrupar os feedbacks recebidos no questionário, os seguintes parâmetros serão utilizados
- Tempo (em anos) na função;
- Experiência em Análise de dados ou Tomada de decisão (múltipla escolha);
- Em uma escala, a facilidade de adaptação da base atual do entrevistado ao seguir os procedimentos de criação do *dashboard*;
- Em uma escala, o a aplicabilidade desse processo na visualização dos dados;

- Em uma escala, o impacto destes *dashboards* na qualidade das análises atuais;
- Em uma escala, a eficácia das análises atuais que o procedimento proporcionou;
- E se processo está aprovado para substituir o procedimento de análise atual, com duas respostas, nas quais, se positivo finaliza o questionário e, se negativo, é necessário justificar antes da finalização.

O Formulário completo com estas questões se encontra disponível no Apêndice II.

3.13 Validação do processo com base nas respostas do questionário

Para toda e qualquer resposta negativa à validação de uso no procedimento de análise atual, os comentários serão avaliados e utilizados como nova interação no desenvolvimento do passo a passo. E assim como estruturado na metodologia do Crisp-DM, é necessário retornar ao passo 4 de modelagem e revisar os tópicos de estruturação do *dashboard*, a fim de solucionar as lacunas identificadas pelo público-alvo. Além de uma revisão dos passos 1 a 3 do desenvolvimento, aprofundando-se nos objetivos que se deseja alcançar.

3.14 Revisão do processo com base nos *feedbacks* recebidos

Ao utilizar como base os comentários feitos pelo entrevistado no questionário, se faz necessário uma revisão de indicadores e visuais, para que sejam mais bem direcionados, com base em ferramentas de modelagem e readequação do *storytelling*.

Ao final desse processo de readaptação, é fundamental a reaplicação do questionário, e a reavaliação por parte do público-alvo de todo o método e, por fim, pelos autores do procedimento, repetindo esse fluxo até a aplicabilidade do sistema satisfazer os critérios criados e esperados pelos autores.

3.15 Avaliação da metodologia

Utilizando a metodologia Crisp-DM, foi possível definir a base e passos necessários para a elaboração deste projeto, e seguindo suas etapas, foi possível criar o procedimento validado através do questionário. Os resultados obtidos estão detalhados no capítulo 4.

4 RESULTADOS

O fluxograma de desenvolvimento proposto no trabalho (apresentado na Figura 12) contém 6 passos. Sendo que o primeiro passo (identificação das necessidades do processo) foi definido na etapa teórica deste trabalho. Nesta seção foram relatados os 5 passos posteriores do desenvolvimento proposto. Onde identificou-se os dados disponíveis, a avaliação dos dados essenciais para o objetivo analisado e os resultados obtidos.

4.1 Dados disponíveis

Os dados estão disponíveis em 3 arquivos de extensão ‘.csv’ (projetos, capacidade e demanda). Nas seções seguintes foram detalhadas as informações de cada um deles.

4.1.1 Projetos

É a tabela principal a ser utilizada nas análises posteriores, e contém todos os detalhes únicos a respeito de um projeto, como o nome, tipo, categoria, datas relevantes, informações financeiras, e qualquer outro tipo de dado. O link entre estas tabelas e todas as demais que referenciam um projeto, é um atributo chamado “Identificador do Projeto”, este identificador é um código único que representa a linha. No Quadro 2 temos um resumo de todas as colunas, bem como um descritivo do dado associado.

Quadro 2 - Colunas da tabela de Projetos

Coluna	Descrição	Obrigatorio
Tipo	Detalha diferentes tipos de projeto, de acordo com a metodologia	
Criado Em	Data (DD/MM/AAAA)	X
Criado Por	Nome de quem criou o projeto	X
Identificação do Projeto	Identificador Único (XX.X-00000)	X
Status do Projeto	Cancelado; Ativo; Novo; Congelado; Finalizado	X
Estagio do Projeto	Estagio de desenvolvimento do projeto (%)	
Pais	Pais do Projeto	X
Categoria	Categoria do Projeto	X
Grupo	Grupo em que o projeto se encontra	X
Tipo	Tipo do projeto dentro do grupo	
Impacto Financeiro	Impacto Financeiro percentual do projeto	X
Data de Finalização Esperada	Data (DD/MM/AAAA)	X
Cidade	Cidade do Projeto	X
Cadeado	Sim; Não (Projeto ainda pode ser alterado)	X

Fonte: Autoria própria (2022)

4.1.2 Demanda

A segunda tabela visa mapear todas as demandas de um projeto, para cada time envolvido em um trimestre em específico, sendo assim, diversas linhas pertencem ao mesmo projeto, e se conecta com os projetos a partir da coluna chave de identificação do Projeto. No Quadro 3 pode ser visualizado um resumo da tabela de Demandas.

Quadro 3 - Colunas da tabela de Demanda

Coluna	Descrição	Obrigatorio
Time	Pais_Cidade_Time_Categoria	X
Demanda	Valor em FTE da Demanda no Período de tempo	X
Trimestre	Data inicial do trimestre (DD/MM/AAAA)	X
 Identificação do Projeto	Identificador Único (XX.X-00000)	X

Fonte: A autoria própria (2022)

A coluna de time mapeia diversas informações em um mesmo campo, utilizando ferramentas de modelagens, será necessário expandi-la nas quatro diferentes colunas independentes demonstradas no Quadro 3 (Pais, Cidade, Time e Categoria), enquanto na coluna de Demanda, o Valor numérico em FTE é o percentual de uma pessoa necessário para execução de um projeto em um trimestre em específico.

Neste Quadro de demanda, as quatro colunas destacadas são mandatórias, contudo, colunas adicionais, se existentes, podem ser utilizadas para filtros diferentes dos mapeados neste estudo.

4.1.3 Capacidade

De maneira similar ao quadro de demandas, o Quadro 4, de capacidade, contém o valor em FTEs, ou seja, quantas pessoas estão disponíveis em um time específico dentro de um trimestre.

Quadro 4 - Colunas da tabela de Capacidade

Coluna	Descrição	Obrigatorio
Time	Pais_Cidade_Time_Categoria	X
Demanda	Valor em FTE da Capacidade total no Período de ten	X
Trimestre	Data inicial do trimestre (DD/MM/AAAA)	X

Fonte: A autoria própria (2022)

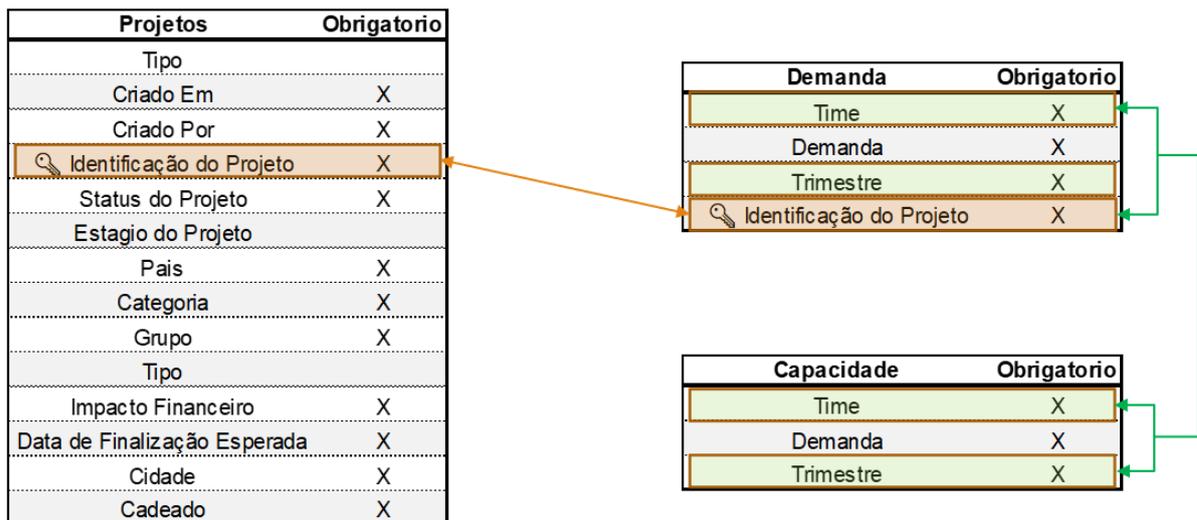
Os times definidos aqui necessitam ser os mesmos que os utilizados no Quadro de demanda, para que se construa uma relação de uma linha no quadro de capacidade para muitos no quadro de demanda.

Neste quadro de capacidade, as três colunas destacadas são mandatórias, contudo, colunas adicionais, se existentes, podem ser utilizadas para filtros diferentes dos mapeados neste estudo.

4.2 Relações multidimensionais

Baseando-se na Figura 13, é possível mapear quais as relações entre as tabelas são relevantes no modelo, e como conectar os diferentes dados. Um esboço desta relação pode ser identificado na Figura 16.

Figura 16 - Esboço das relações entre as tabelas '.CSV'



Fonte: Autoria própria (2022)

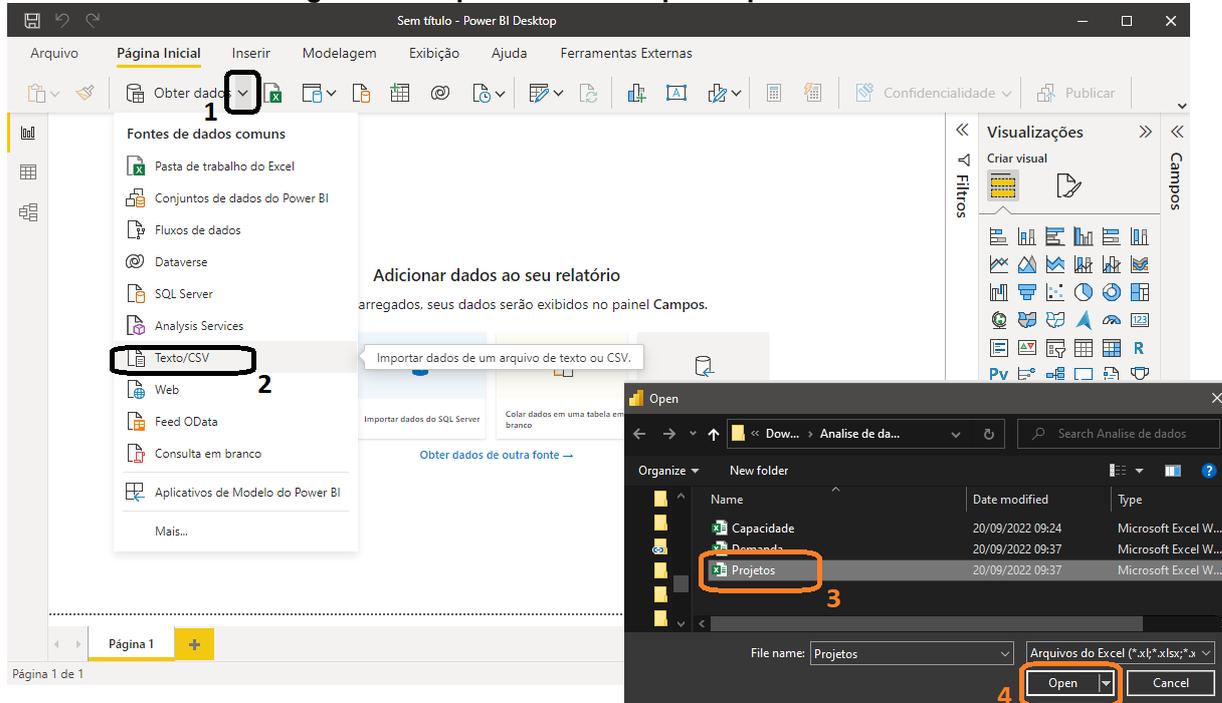
A Figura 16 apresenta a relação entre um projeto único na tabela de projetos, com múltiplas linhas de demanda, para diferentes times, na tabela de demandas. Estas demandas são associadas, pelo conjunto de pais, cidade, time e categoria, com a capacidade total de cada um destes times, desta maneira, estabelece-se a relação necessária entre as três tabelas.

4.3 Modelagem dos dados no Power Query

Com o esboço lógico das tabelas finalizado, pode-se iniciar o processo de modelagem dentro do editor avançado do *Power BI*. O primeiro passo é abrir um novo

documento no *Power BI*, em branco, e importar os arquivos através do passo a passo disponível na Figura 17, repetindo-o para cada uma das 3 bases.

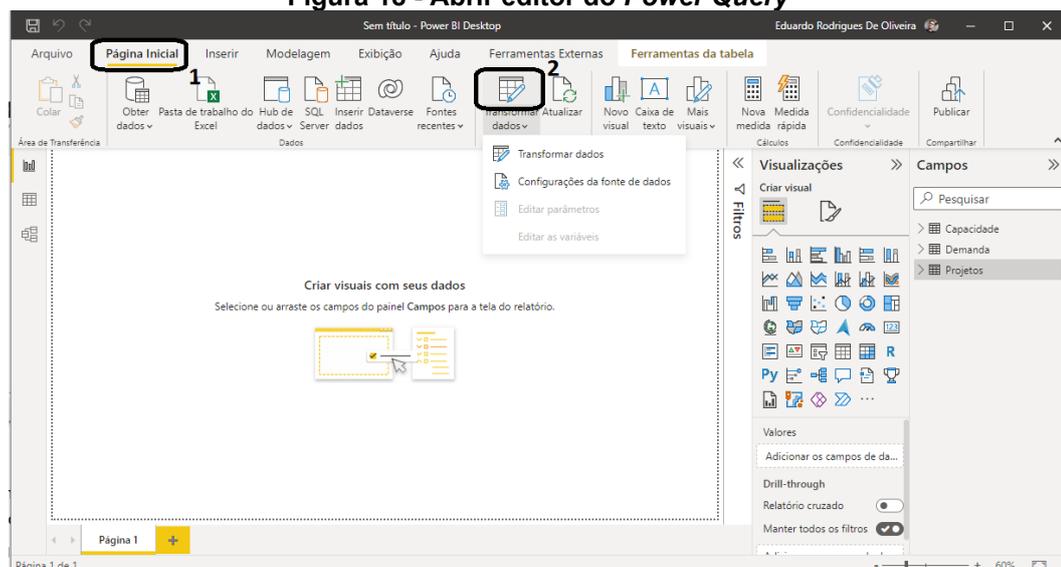
Figura 17 - Importar novos arquivos para o *Power BI*



Fonte: Autoria própria (2022)

Com os arquivos bases importados, podemos, através do menu superior do *Power BI*, selecionar “Página Inicial -> Transformar Dados”, como demonstrado na Figura 18.

Figura 18 - Abrir editor do *Power Query*



Fonte: Autoria própria (2022)

Dentro do Editor *Power Query*, podemos visualizar, conforme a Figura 19, na lateral esquerda as tabelas importadas, e na lateral direita todas as etapas de

transformação aplicadas a esta tabela. Quando um arquivo é importado, o *Power BI* automaticamente tenta fazer alguns passos, como transformar a primeira linha em cabeçalho, e alterar as colunas para seus devidos tipos de dado, como números inteiros, decimais, textos e datas. Você pode desfazer qualquer uma destas etapas a selecionando com o botão direito e a deletando.

Figura 19 - Navegação entre tabelas e etapas no Power Query

The screenshot displays the Power Query Editor window. The main area shows a table with columns: 'Identificação do Projeto', 'Capacidade', and 'Trimestre'. The 'Identificação do Projeto' column contains various project category names. The 'Capacidade' column contains numerical values, and the 'Trimestre' column contains dates. The interface includes a ribbon with options like 'Transformar', 'Adicionar Coluna', and 'Exibição'. A left-hand pane shows a list of queries, and a right-hand pane shows the configuration for the selected query, including 'ETAPAS APLICADAS' such as 'Tipo Alterado'.

Fonte: Autoria própria (2022)

Para as três tabelas importadas, vamos agora destacar as transformações efetuadas para permitir que as mesmas possam ser utilizadas para as análises buscadas.

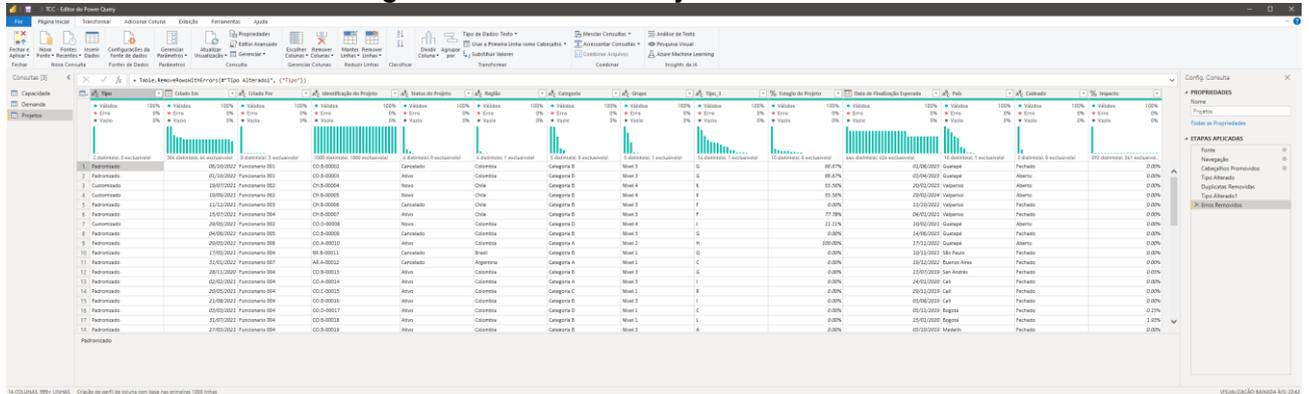
4.3.1 Projetos

A tabela de projetos precisa de poucas alterações, e estão citadas aqui:

- 1) Promover a primeira linha para cabeçalhos das colunas;
- 2) Alterar o tipo das colunas de data, texto e números;
- 3) Remover valores duplicados da coluna "Identificação do Projeto", pois este deve ser um identificador único e exclusivo;
- 4) Definir as colunas numéricas de Percentual (Estágio do Projeto; Impacto) para percentuais.

O resultado após as transformações pode ser visualizado na Figura 20, que contém também *clusters* demonstrando a distribuição dos dados na tabela.

Figura 20 - Tabela de Projetos modelada



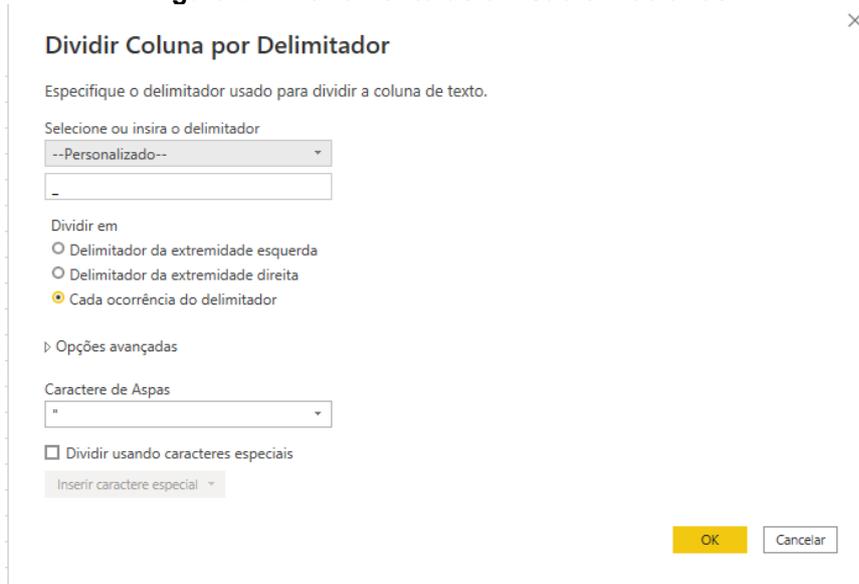
Fonte: Autoria própria (2022)

4.3.2 Demanda

Dentro da tabela de demanda, o mais relevante é expandir a Coluna de Time, transformando-a em 4 colunas, desta maneira, as etapas são as seguintes:

- 1) Promover a primeira linha para cabeçalhos das colunas;
- 2) Como demonstrado na Figura 21, utilizar a ferramenta de dividir colunas na Coluna de Time, utilizando o delimitador adequado;
- 3) Alterar o tipo das colunas de data, texto e números;
- 4) Renomear as quatro novas colunas para País, Cidade, Setor e Categoria.

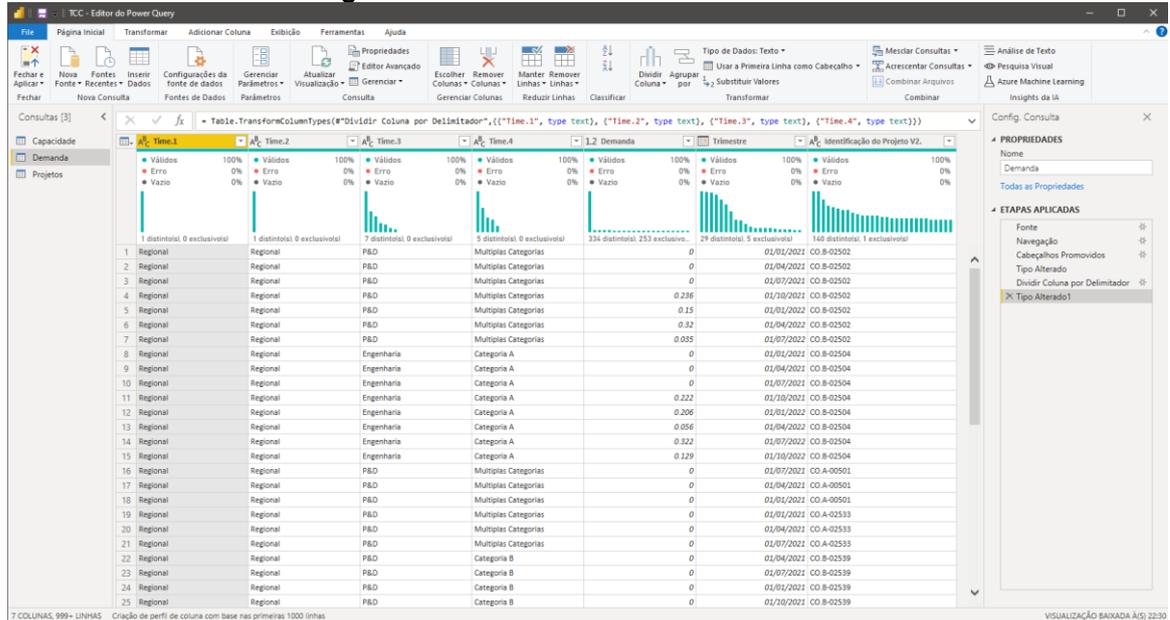
Figura 21 - Ferramenta de divisão em colunas



Fonte: Autoria própria (2022)

Desta maneira, o resultado da tabela pode ser visualizado na Figura 22.

Figura 22 - Tabela de Demandas modelada

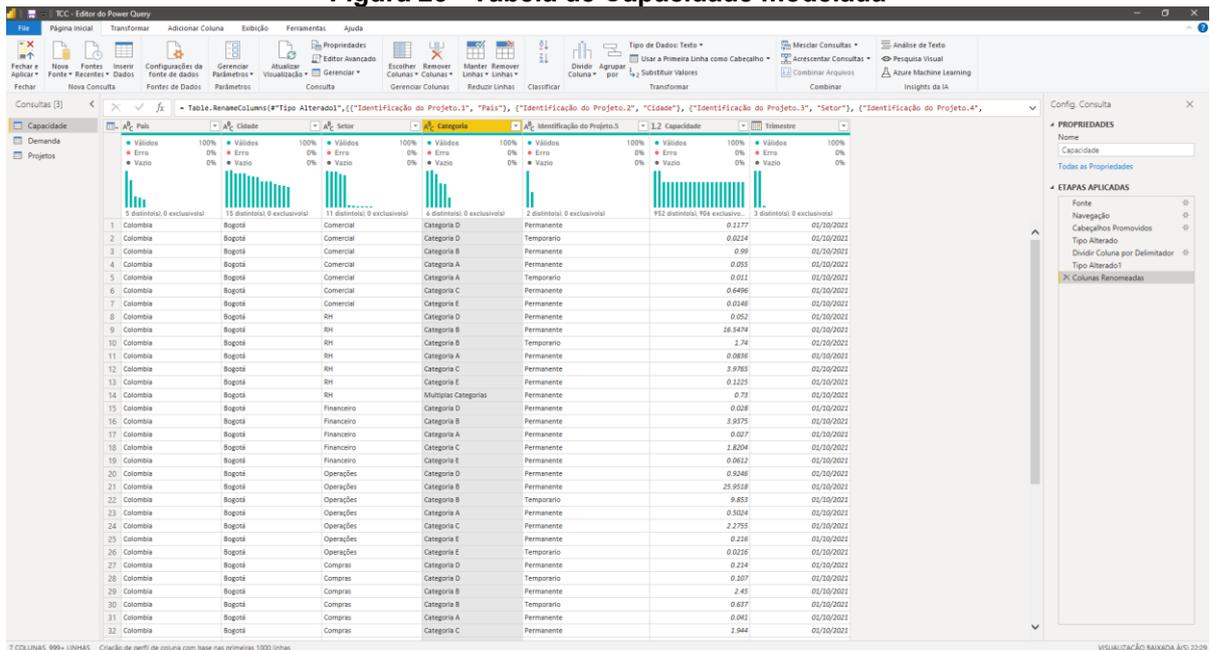


Fonte: Autoria própria (2022)

4.3.3 Capacidade

Para a tabela de capacidade, o único ajuste necessário é o mesmo realizado na tabela de Demanda, transformando a Coluna de Time em quatro distintas colunas (país, cidade, time e categoria), com base no recurso demonstrado na Figura 21, resultando na tabela modelada demonstrada na Figura 23.

Figura 23 - Tabela de Capacidade modelada



Fonte: Autoria própria (2022)

Aqui termina o tratamento de dados utilizando o *Power Query*, ou seja, esta etapa realiza a modelagem padrão dos dados, e disponibiliza os resultados das tabelas de projetos, demanda e capacidade para o *Power BI*. Quaisquer dados anteriores não são acessíveis ao usuário final, pois a modelagem os suprime. Quaisquer novos dados brutos serão tratados utilizando esta mesma lógica.

4.4 Correlação entre as tabelas no *Power BI*

Para evitar situações de referências circulares entre as tabelas, não é aconselhável que se conectem duas colunas contendo múltiplas linhas com o mesmo dado, por isto, se faz necessário a criação de tabelas virtuais, a serem utilizadas como índices, ou tabelas de dimensões, que filtrem as tabelas de fatos.

4.4.1 Tabelas de dimensão

Estas tabelas virtuais são as tabelas de dimensão, que sumarizam os dados em tempo de execução, conforme definido no capítulo 2, sobre a aplicação do OLAP.

4.4.1.1 Calendário

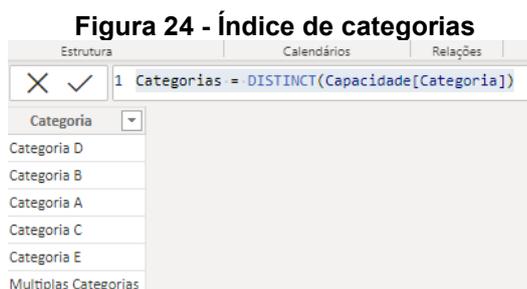
A tabela de calendário visa mapear os trimestres existentes na base, tornando possível realizar filtros com estas datas, e permitindo também visualizações simplificadas, utilizando as seguintes expressões DAX.

```
Calendario = DISTINCT(Demanda[Trimestre])
Descrição = FORMAT(Calendario[Trimestre], "mmm/yy")
```

4.4.1.2 Categorias

A tabela de Categorias visa criar um índice com todas as categorias disponíveis, utilizando uma expressão DAX similar à da tabela de calendário, resultando na tabela demonstrada na Figura 24.

```
Categorias = DISTINCT(Capacidade[Categoria])
```



Fonte: Autoria própria (2022)

4.4.1.3 Países

A tabela de países visa criar um índice a partir de todos os valores distintos de cidades existentes nas tabelas de Capacidade e Demanda, resumizando respectivamente, com base nas expressões DAX a seguir, que resultam na Figura 25.

```
Países = DISTINCT(union(DISTINCT(Capacidade[Pais]),DISTINCT(Demanda[Pais])))
Pais = LOOKUPVALUE(Capacidade[Região],Capacidade[Pais],Países[Cidade],"Regional")
```

Figura 25 - Índice de Cidades

Cidade	País
Bogotá	Colombia
Medelin	Colombia
Cali	Colombia
San Andrés	Colombia
Guatapé	Colombia
Santa Marta	Colombia
Florianópolis	Brasil
Rio de Janeiro	Brasil
Salvador	Brasil
São Paulo	Brasil
Puerto Iguazú	Argentina
Buenos Aires	Argentina
Santiago	Chile
Valparaiso	Chile
Regional	Regional
Medelin	Regional

Fonte: Autoria própria (2022)

4.4.1.4 Setores

De maneira similar as demais, o índice de setores, demonstrado na Figura 26, criado a partir da expressão DAX abaixo, define os setores da empresa.

```
Grupos de Trabalho = DISTINCT(Capacidade[Grupo de Trabalho])
```

Figura 26 - Índice de Setores

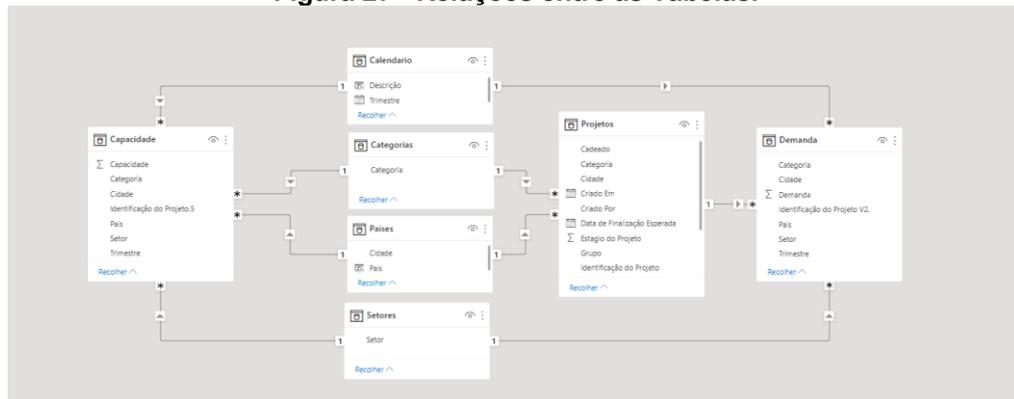
Setores
Comercial
RH
Financeiro
Operações
Compras
G10
G8
G9
P&D
G11
Engenharia

Fonte: Autoria própria (2022)

4.4.2 Modelagem e correlações

Com base nas tabelas de índice criadas, utilizando a página de Modelos dentro do *Power BI*, foram desenvolvidas as relações entre as três tabelas de fatos importadas e as quatro tabelas de índice criadas, que pode ser demonstrada na Figura 27.

Figura 27 - Relações entre as Tabelas.



Fonte: Autoria própria (2022)

A Figura 27 demonstra a lógica final, já modelada dentro do *Power BI*, construída a partir do esboço da Figura 16, demonstrando como as diferentes tabelas de fato se conectam entre si, com a adição das tabelas de dimensão, criadas como índices virtuais.

4.5 Criação de medidas com base nas tabelas calculadas

Como definido no referencial teórico, medidas são lógicas escritas que realizam um cálculo de acordo com os dados selecionados no contexto de filtro. Desta maneira pode-se ter, o que pode ser chamado de cálculo dinâmico, a partir de uma fórmula que se adequa sempre aos dados filtrados atuais.

4.5.1 Carga de trabalho

O percentual de carga de trabalho, conforme adaptado da equação apresentada por PontoTel (2021), pode ser definido pela Demanda total de um grupo, sobre a capacidade deste grupo, desta maneira, definimos, na Equação 1, em FTEs (Full Time Employee), desta maneira, dividindo, em percentual de funcionários, a demanda pela quantidade de funcionários disponíveis.

$$Carga\ de\ trabalho\ (\%) = \frac{\sum Demanda\ (FTE)}{\sum Capacidade\ (FTE)} \quad (1)$$

Este percentual tem o resultado de 100% quando a Capacidade se iguala a Demanda, e teoricamente não deve estar acima de 100%, pois significa um excesso de carga de trabalho, contudo, na prática, os valores limites devem ser definidos pelas organizações, avaliando a confiabilidade do indicador de demanda.

4.5.2 Eficiência

A eficiência, como discutido com base nas equações de PontoTel (2021), é a demanda, em FTE, aplicada ao impacto financeiro de um projeto, desta maneira, visa mapear a demanda pelo retorno financeiro, avaliando, como demonstrado na Equação 2, o valor financeiro que cada FTE entrega, dentro dos projetos.

$$Eficiencia (R\$/FTE) = \frac{\sum Impacto Financeiro (R\$)}{\sum Demanda (FTE)} \quad (2)$$

Quanto maior o resultado deste indicador, mais eficiente é a empresa, permitindo um maior retorno financeiro com menos demanda. Valores de referência não podem ser definidos, pois variam de acordo com o ramo de trabalho da organização.

4.5.3 Retorno dos maiores projetos

Uma relevante avaliação para a organização, é o valor entregue pelos maiores projetos considerados, desta maneira, a Equação 3, que é um cálculo de percentual, visa mapear o percentual monetário entregue pelos maiores 20 projetos da seleção.

$$Impacto top 20 (\%) = \frac{\sum_{20 \text{ Maiores Projos}} Impacto Financeiro}{\sum_{\text{Todos os projetos}} Impacto Financeiro} \quad (3)$$

Esta métrica irá variar entre 0% e 100%, e quanto mais próximo do limite superior, menos projetos são necessários para que a organização atinja suas métricas financeiras. O objetivo deste indicador deve estar atrelado aos objetivos da empresa.

4.6 Desenvolvimento dos *dashboards* com base nos dados

Com as medidas e dados disponíveis, o primeiro passo é desenhar um esboço dos visuais que se deseja representar. Nesta etapa, não se busca o visual final, apenas rascunhos que nos permitam entender quais dados representar e quais visuais utilizar, com base em testes e experiência.

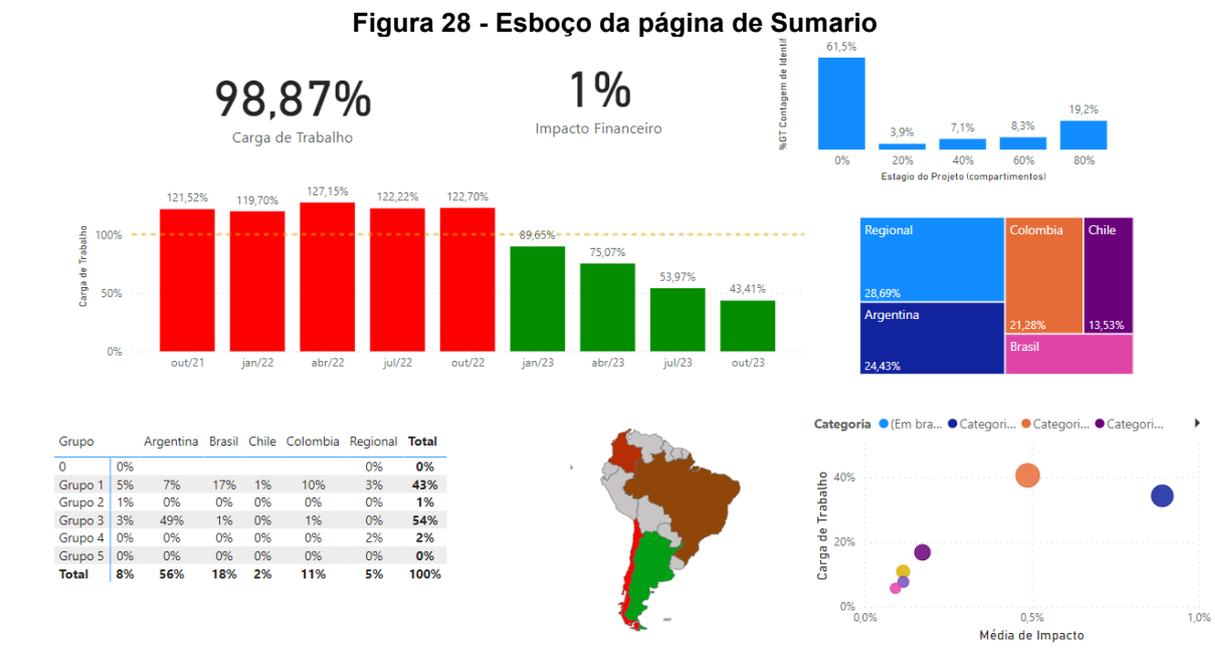
Avaliando as tabelas de dados modeladas, busca-se representar os principais dados contidos neste modelo em 6 páginas, sendo elas:

1. Sumário;
2. Financeiro;
3. Maiores Projetos;
4. Gráficos de Demanda;
5. Tabelas de Demanda;
6. Extração de dados brutos.

Sendo assim, se iniciou o desenvolvimento das telas, com os visuais disponíveis descritos na Figura 15, comentada no capítulo 3.

4.6.1 Sumário

Utilizando visuais compreensíveis, a Figura 28 se propõe a demonstrar, de maneira simples, os maiores impactos dos indicadores, como carga de trabalho e distribuição financeira ao longo do tempo e localização geográfica.



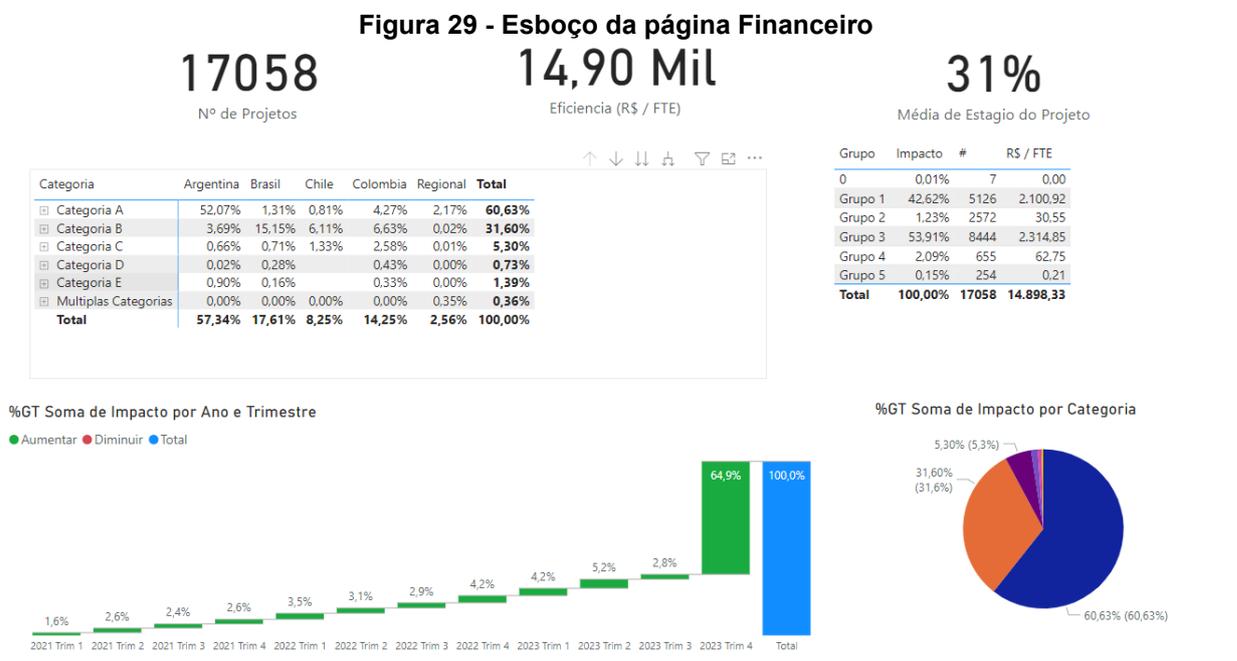
Fonte: Autoria própria (2022)

No exemplo específico da Figura 28, verifica-se visualmente que entre outubro de 2021 e outubro de 2022, a carga de trabalho excede os 100%, indicando riscos a qualidade de vida dos trabalhadores, e pode indicar oportunidades de transferir parte

da demanda para os próximos meses. Indica ainda que esta demanda é mais alta no Chile, em projetos que ainda não chegaram em 20% de execução.

4.6.2 Financeiro

Dentro de página de indicadores financeiros, busca-se sumarizar o ganho financeiro e eficiência média dos projetos ao longo de categorias, grupos e zonas geográficas, como demonstrado na Figura 29, possibilitando visualizar por exemplo um aumento constante na entrada de caixa da empresa ao longo do tempo, no intervalo de 2021 a 2023, e destacando ainda que 52% de toda a entrada é entregue na Argentina pela categoria de produtos A.



Fonte: Autoria própria (2022)

4.6.3 Maiores projetos

Como demonstrado na Figura 30, o objetivo desta página é demonstrar o impacto dos 20 maiores projetos e o status destes mesmos, de acordo com a relevância para a organização. Isto nos permite entender o risco associado aos maiores projetos da empresa, visualizando por exemplo, um risco no maior projeto, com entrega mapeada para out/2023, qual ainda não foi iniciada a execução.

Figura 30 - Esboço da página de Maiores Projetos

17058 N° de Projetos **63,95%** Impacto Financeiro Top 20 **0,34%** Impacto Demanda Top 20

Tipo	ID	Impacto	Eficiencia (RS / FTE)	Demanda	Categoria	Grupo	SubGrupo	Pais	Cidade	Finalização	Status	Estagio
Padronizado	ARA-14107	2865%	0,66	2,29	Categoria A	Grupo 3	F	Argentina	Buenos Aires	02/out/23	Ativo	0%
Padronizado	BR-B-15608	1224%	0,00		Categoria B	Grupo 1	L	Brasil	São Paulo	08/out/23	Congelado	0%
Padronizado	ARA-14802	1223%	0,00		Categoria A	Grupo 3	F	Argentina	Buenos Aires	19/mar/24	Novo	0%
Padronizado	BR-B-15135	100%	0,01	1,29	Categoria B	Grupo 1	Q	Brasil	Salvador	30/dez/23	Ativo	100%
Padronizado	ARA-14113	47%	0,00	0,30	Categoria A	Grupo 1	R	Argentina	Buenos Aires	21/mar/23	Ativo	0%
Customizado	G.A-01841	41%	0,06	13,51	Categoria A	Grupo 4	J	Regional	Regional	15/dez/24	Novo	0%
Padronizado	ARE-16631	36%	0,00		Categoria E	Grupo 1	C	Argentina	Buenos Aires	10/fev/23	Cancelado	0%
Padronizado	CH-B-04052	33%	0,01	3,51	Categoria B	Grupo 3	A	Chile	Valpariso	14/ago/24	Ativo	56%
Padronizado	CO.C-16076	27%	0,00		Categoria C	Grupo 1	L	Colombia	Santa Marta	15/mar/22	Cancelado	0%
Padronizado	CO.C-16077	27%	0,00		Categoria C	Grupo 1	L	Colombia	Santa Marta	18/mar/22	Cancelado	0%
Padronizado	CO.C-16078	27%	0,00	0,25	Categoria C	Grupo 1	L	Colombia	Santa Marta	13/dez/21	Cancelado	33%
Padronizado	AR-B-07491	27%	0,00		Categoria B	Grupo 3	A	Argentina	Puerto Iguazú	29/mar/21	Ativo	0%
Padronizado	AR.C-04976	23%	0,00		Categoria C	Grupo 3	I	Argentina	Buenos Aires	27/nov/21	Ativo	0%
Padronizado	AR.B-11973	23%	0,03	11,23	Categoria B	Grupo 1	L	Argentina	Argentina	16/jul/23	Ativo	33%
Padronizado	CO.B-04515	22%	0,00		Categoria B	Grupo 2	H	Colombia	Colombia	29/fev/24	Novo	0%
Padronizado	ARA-13945	21%	0,00		Categoria A	Grupo 3	I	Argentina	Buenos Aires	14/jun/23	Novo	0%
Customizado	G.A-09293	20%	0,02	10,08	Categoria A	Grupo 4	J	Regional	Regional	07/nov/24	Novo	0%
Customizado	G.A-15343	20%	0,02	12,00	Categoria A	Grupo 4	J	Regional	Regional	25/fev/25	Novo	22%
Padronizado	CH-B-04365	17%	0,00		Categoria B	Grupo 2	B	Chile	Valpariso	18/jan/21	Ativo	0%
Padronizado	ARA-11668	17%	0,00	1,63	Categoria A	Grupo 3	A	Argentina	Argentina	01/mar/24	Ativo	0%
Total		5843%		32,77								12%

Fonte: Autoria própria (2022)

4.6.4 Gráficos de demanda

Outro fator de relevância durante a análise destes dados, é entender se existem recursos disponíveis, e se sim, onde estão alocados, para execução das atividades ao longo do tempo, desta maneira, a Figura 31 busca demonstrar detalhes desta demanda sobre a Capacidade disponível.

Figura 31 - Esboço da página de Gráficos de Demanda

98,87% Carga de Trabalho **16,31 Mil** Soma de Demanda **16,49 Mil** Soma de Capacidade



Fonte: Autoria própria (2022)

Um exemplo de avaliação possível com os dados sumarizados, é entender que embora na média, a carga de trabalho não esteja superior a 100%, grupos específicos como engenharia e RH trabalham acima do recomendado, principalmente nas cidades de Santiago, Guatapé e Florianópolis. Demonstrando, como no sumario, as possibilidades de transferir parte desta demanda para a partir de abril de 2023.

4.6.5 Tabelas de demanda

De maneira similar aos gráficos de demanda, objetiva-se nestes visuais, demonstrados na Figura 32, validar as curvas de capacidade pela demanda em setores, categorias e posições geográficas específicas.

Figura 32 - Esboço da página de Tabelas de Demanda

122,96% Carga de Trabalho **9,57 Mil** Soma de Demanda **7,78 Mil** Soma de Capacidade

Categoria	jan/22	abr/22	jul/22	out/22
Categoria A				
Demanda	738,0	762,5	806,8	652,6
Capacidade	554,4	527,2	603,2	487,7
Carga Trabalho	133%	145%	134%	134%
Categoria B				
Demanda	770,2	929,4	831,5	826,3
Capacidade	594,8	697,3	592,9	605,3
Carga Trabalho	129%	133%	140%	136%
Categoria C				
Demanda	342,9	357,2	346,3	352,3
Capacidade	300,2	285,5	262,9	277,9
Carga Trabalho	114%	125%	132%	127%
Categoria D				
Demanda	113,8	122,6	124,3	99,2
Capacidade	89,3	103,6	103,3	84,9
Carga Trabalho	127%	118%	120%	117%
Categoria E				
Demanda	2.323,8	2.534,4	2.484,2	2.213,5
Capacidade	1.941,7	1.994,1	2.035,2	1.808,7
Carga Trabalho	120%	127%	122%	122%

Pais	Comercial	Compras	Engenharia	Financeiro	Operações	P&D	RH
Argentina							
Demanda	110,1	223,4	32,1	201,0	1.005,6	8,1	3
Capacidade	117,2	212,8			193,7	1.075,8	3
Carga Trabalho	94%	105%	Infinity	104%	93%	Infinity	1
Brasil							
Demanda	102,2	157,7	0,0	174,9	544,9	6,2	2
Capacidade	75,1	130,7		120,5	403,7		1
Carga Trabalho	136%	121%	NaN	145%	135%	Infinity	1
Chile							
Demanda	15,7	58,9	0,2	17,7	96,9	0,0	6
Capacidade	60,5	131,5		124,7	300,8		1
Carga Trabalho	26%	45%	Infinity	14%	25%	NaN	3
Colombia							
Demanda	141,6	263,3	16,3	163,7	812,0	9,8	3
Capacidade	127,9	221,6		143,7	753,9		3
Carga Trabalho	111%	119%	Infinity	114%	108%	Infinity	1
Total							
Demanda	525,3	757,6	250,2	1.042,6	3.182,1	56,2	1
Capacidade	490,8	696,7	209,5	1.092,4	3.107,7	278,6	1
Carga Trabalho	107%	109%	119%	95%	102%	20%	11

Fonte: Autoria própria (2022)

4.6.6 Extração de dados brutos

Parte importante deste trabalho de modelagem consiste em estruturar as relações entre as diferentes tabelas, de maneira que os dados possam ser analisados da melhor maneira. Sendo assim, a Figura 33 busca disponibilizar uma tabela que contenha todos os dados disponíveis a partir de todos os dados de entrada, permitindo ao usuário que extraia os mesmos, caso necessite de avaliações distintas das disponíveis dentro das páginas do *Power BI*, mesmo sem nenhum conhecimento no programa, pois pode baixar como uma planilha e avaliar em qualquer outro contexto.

Figura 33 - Esboço da página de Extração de dados

Identificação do Projeto	Categoria	Pais	Cidade	Grupo	SubGrupo	Tipo	Status do Projeto	Soma de Estágio do Projeto	Data de Finalização Esperada	Setor	Trimestre	Soma de Demanda
ARA-00443	Categoria A	Argentina	Buenos Aires	Grupo 3	F	Padronizado	Ativo	56%	24/set/22	Operações	2022	0,21
ARA-00443	Categoria A	Argentina	Buenos Aires	Grupo 3	F	Padronizado	Ativo	56%	24/set/22	Operações	sábado, 1 de outubro de 2022	0,12
ARA-00443	Categoria A	Argentina	Buenos Aires	Grupo 3	F	Padronizado	Ativo	56%	24/set/22	RH	sexta-feira, 1 de outubro de 2021	0,04
ARA-00443	Categoria A	Argentina	Buenos Aires	Grupo 3	F	Padronizado	Ativo	56%	24/set/22	RH	sábado, 1 de janeiro de 2022	0,11
ARA-00443	Categoria A	Argentina	Buenos Aires	Grupo 3	F	Padronizado	Ativo	56%	24/set/22	RH	sexta-feira, 1 de abril de 2022	0,12
ARA-00443	Categoria A	Argentina	Buenos Aires	Grupo 3	F	Padronizado	Ativo	56%	24/set/22	RH	sexta-feira, 1 de julho de 2022	0,02
ARA-00443	Categoria A	Argentina	Buenos Aires	Grupo 3	F	Padronizado	Ativo	56%	24/set/22	RH	sábado, 1 de outubro de 2022	0,03
ARA-01183	Categoria A	Argentina	Argentina	Grupo 2	H	Padronizado	Ativo	56%	09/jul/23	Comercial	quinta-feira, 1 de abril de 2021	0,00
ARA-01183	Categoria A	Argentina	Argentina	Grupo 2	H	Padronizado	Ativo	56%	09/jul/23	Comercial	quinta-feira, 1 de julho de 2021	0,00
ARA-01183	Categoria A	Argentina	Argentina	Grupo 2	H	Padronizado	Ativo	56%	09/jul/23	Comercial	sexta-feira, 1 de outubro de 2021	0,42
ARA-01183	Categoria A	Argentina	Argentina	Grupo 2	H	Padronizado	Ativo	56%	09/jul/23	Comercial	sábado, 1 de janeiro de 2022	0,46
ARA-01183	Categoria A	Argentina	Argentina	Grupo 2	H	Padronizado	Ativo	56%	09/jul/23	Comercial	sexta-feira, 1 de abril de	0,20

Fonte: Autoria própria (2022)

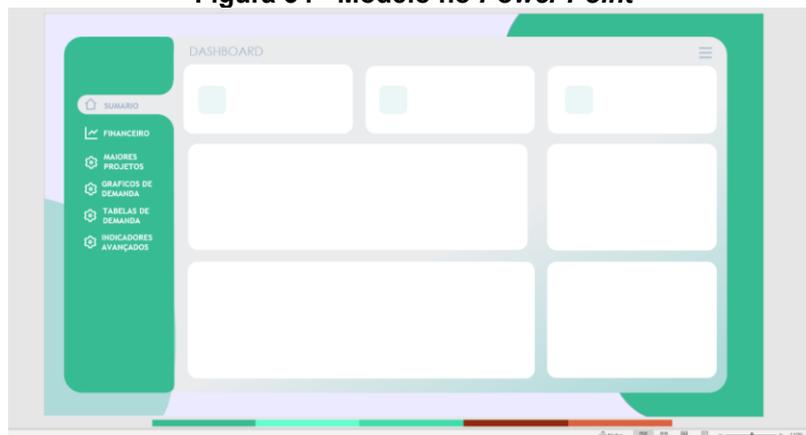
4.7 O uso do *Storytelling*

Com os visuais esboçados, o próximo passo é entender a melhor maneira de demonstrá-los, e para isto, com o entendimento de qual é o objetivo do público, é delimitado como é a configuração da página com gráficos, cores, formatos, destaques, saturações, se faz necessário desenvolver uma interface intuitiva e de simples entendimento.

4.7.1 Desenvolvimento das telas de fundo

Utilizando o programa *Microsoft Power Point*, criamos em uma tela em branco, como demonstrado na Figura 34, um modelo que nos permita demonstrar todos os dados disponíveis e defina o posicionamento dos dados.

Figura 34 - Modelo no *Power Point*



Fonte: Autoria própria (2022)

Utilizando como base este modelo, foram construídos dentro da apresentação, 6 variações para as telas disponíveis na Figura 35, e os arquivos foram salvos como imagem para serem importados pelo *Power BI*.

Figura 35 - *Templates* aplicados

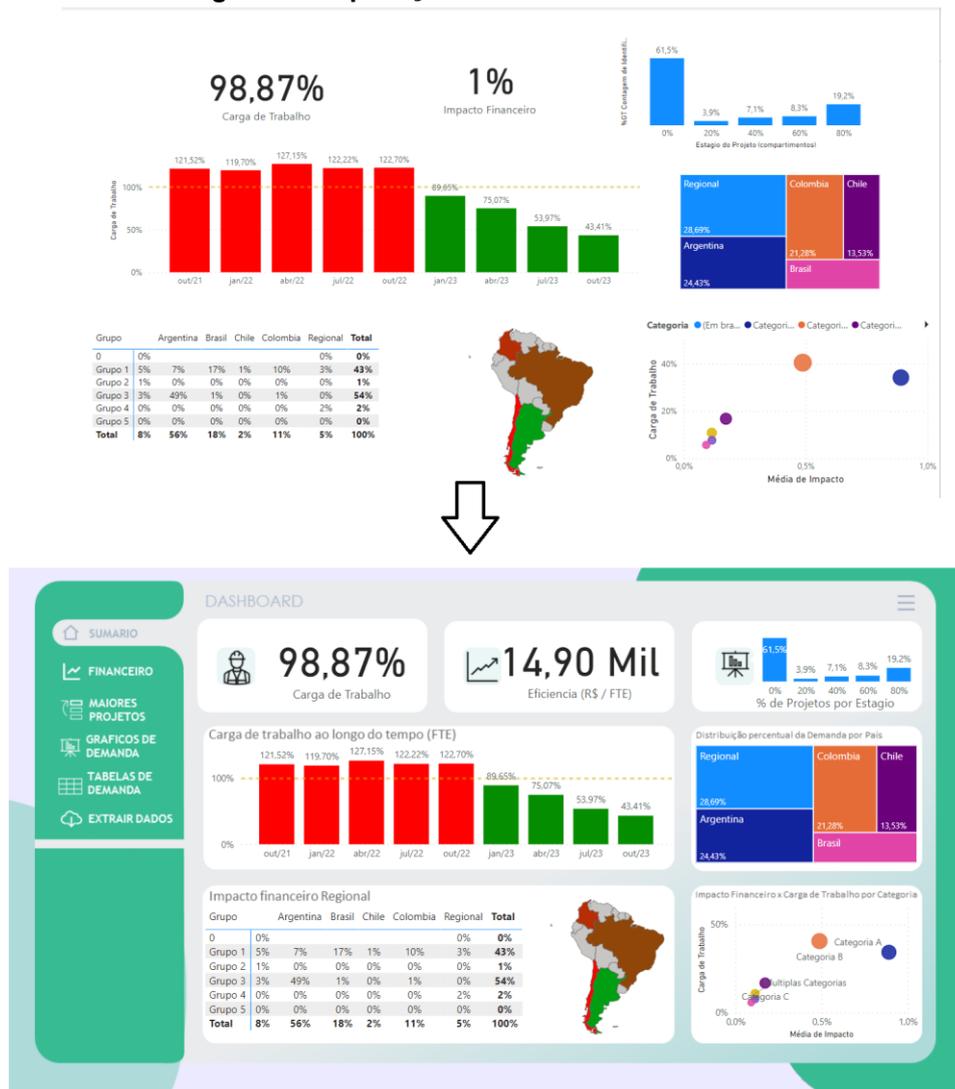


Fonte: Autoria própria (2022)

4.7.2 Posicionamento dos elementos gráficos

Com os planos de fundo definidos, eles foram inseridos no *Power BI*, e os visuais foram reposicionados para corresponder a estas novas posições, como pode ser visualizado no exemplo da Figura 36, correspondente a primeira página.

Figura 36 - Aplicação dos Modelos no Power BI



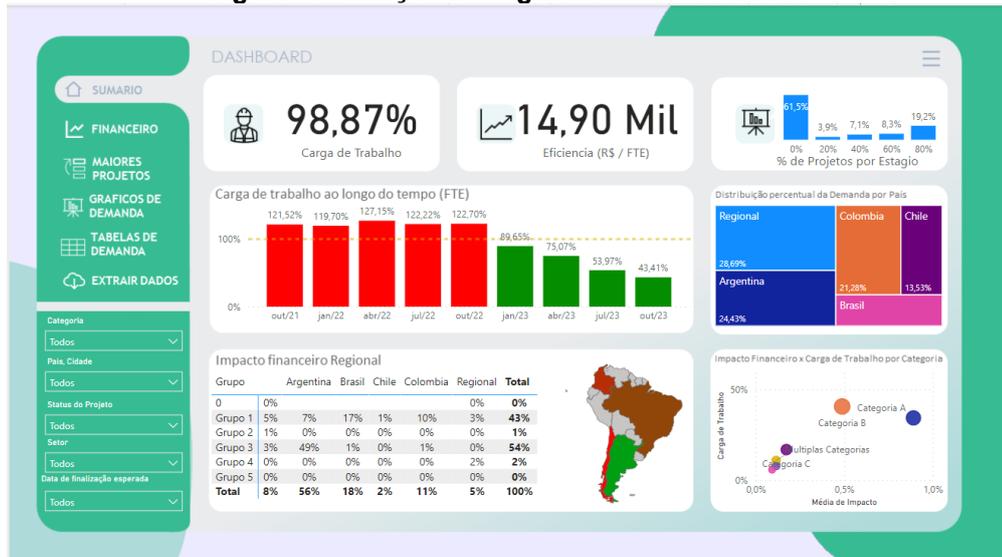
Fonte: Autoria própria (2022)

De igual maneira, o posicionamento foi replicado nas outras 5 páginas.

4.7.3 Segmentadores de dados

Após a disposição dos visuais, a próxima etapa consiste na adição de visuais de filtros, também conhecidos como segmentadores de dados, a serem posicionados na lateral esquerda inferior, para permitir que diferentes perspectivas e cenários sejam avaliados. Estes filtros podem ser visualizados na Figura 37.

Figura 37 - Adição de segmentadores de dados

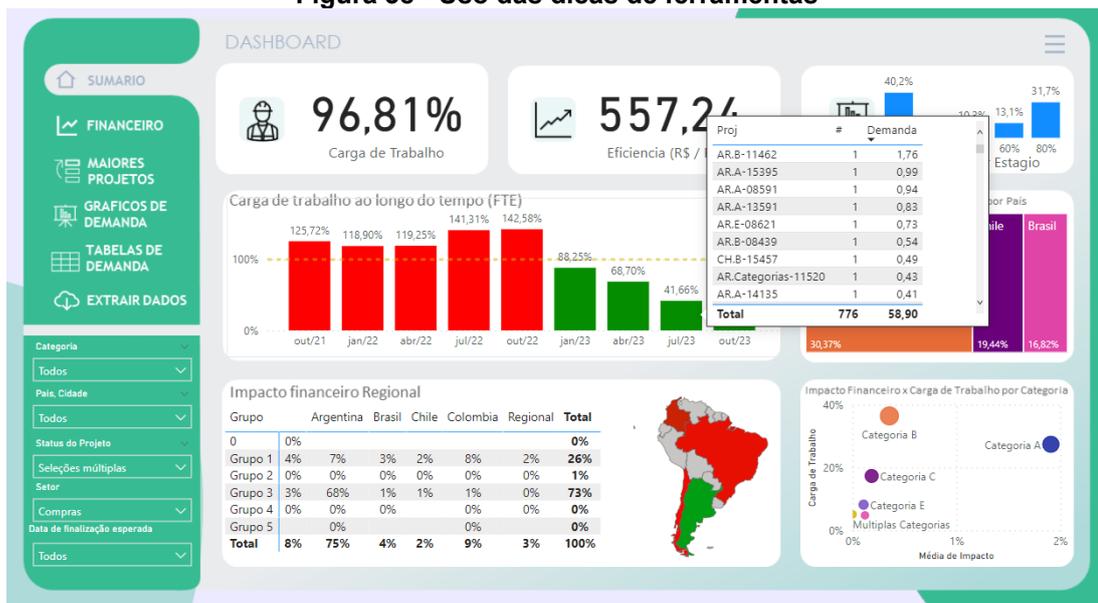


Fonte: Autoria própria (2022)

4.7.4 Dicas de ferramenta

Até agora, todos os visuais buscaram sumarizar os resultados disponíveis, de maneira a colocar em uma apresentação o resumo dos dados. A partir deste momento foram desenvolvidas algumas dicas de ferramentas, que são funções do *Power BI* que permitem a adição de mini visuais, sobrepostos a qualquer outro, que filtrem os dados selecionados pelo usuário, desta maneira, como exemplificado na Figura 38, é possível, ao passar o mouse em cima de uma coluna do gráfico, entender quais projetos impactam este valor de demanda.

Figura 38 - Uso das dicas de ferramentas



Fonte: Autoria própria (2022)

4.7.5 Conclusão do desenvolvimento

Após as etapas anteriores, o arquivo de *Power BI* está pronto para ser publicado e replicar os visuais para quaisquer novos dados adicionados ao longo do tempo.

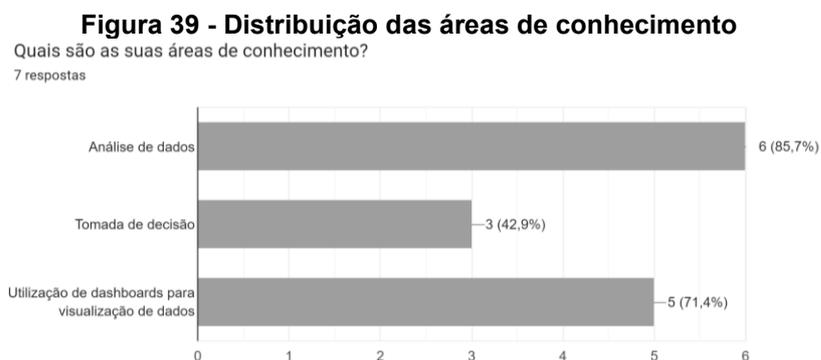
4.8 Desenvolvimento do processo

Com o conhecimento adquirido ao longo do desenvolvimento dos passos anteriores, objetiva-se neste tópico desenvolver um guia que permita que bases similares sejam modeladas com o mesmo procedimento. Através de ferramentas de arte gráfica, como o *Microsoft Power Point*, foi desenvolvido um guia passo a passo para replicar, com quaisquer outros dados similares e conhecimentos em *Power BI*, o *dashboard* desenvolvido.

O questionário desenvolvido foi compartilhado em um grupo de pessoas que trabalham com análise de dados, ou como tomadores de decisão em empresas do setor de bens de consumo, que tenham como atividade relacionada, a avaliação de projetos e carga de trabalho. Desta maneira, por se tratar de um nicho do mercado, um total de 7 respostas foram validadas, destas, apresentamos os dados.

4.9 Avaliação do questionário

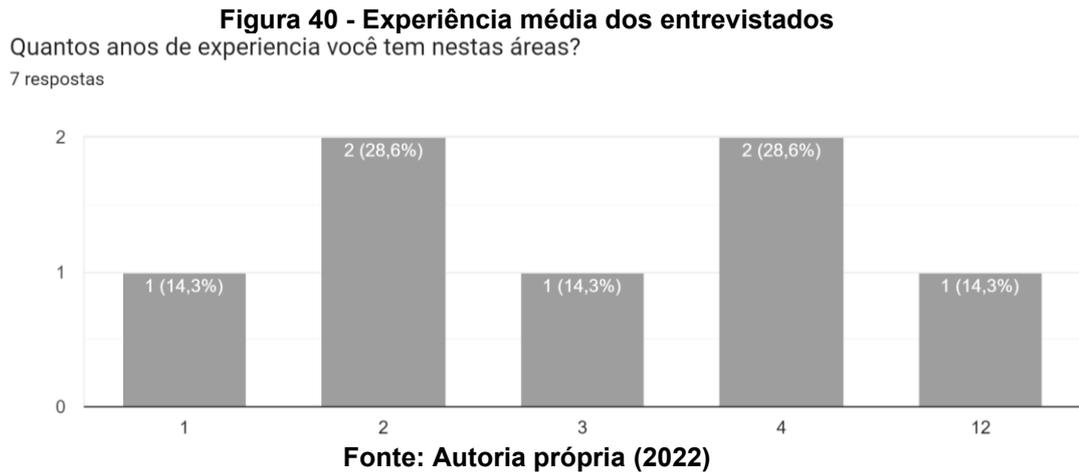
Devido ao conteúdo do trabalho, focou-se a avaliação em pessoas com capacidades e funções específicas nas áreas buscadas, e como visualizado na Figura 39, a avaliação será realizada majoritariamente por profissionais que trabalham na área de análise de dados, como também líderes responsáveis pela tomada de decisão em organizações com estes dados.



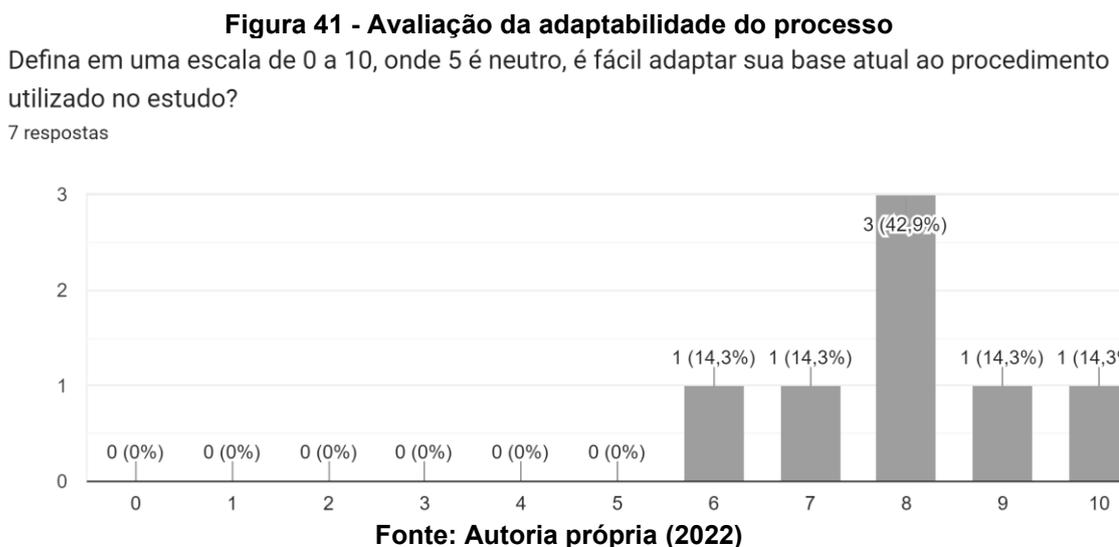
Fonte: Autoria própria (2022).

Destaca-se que a resposta do questionário da Figura 39 não era singular, pois determinados profissionais podem responder por mais de uma função.

Como visualizado na Figura 40, estes profissionais tem distintos tempos de experiencia no mercado, entre 1 e 12 anos, permitindo assim avaliar os resultados da pesquisa em distintos espectros de experiencia.



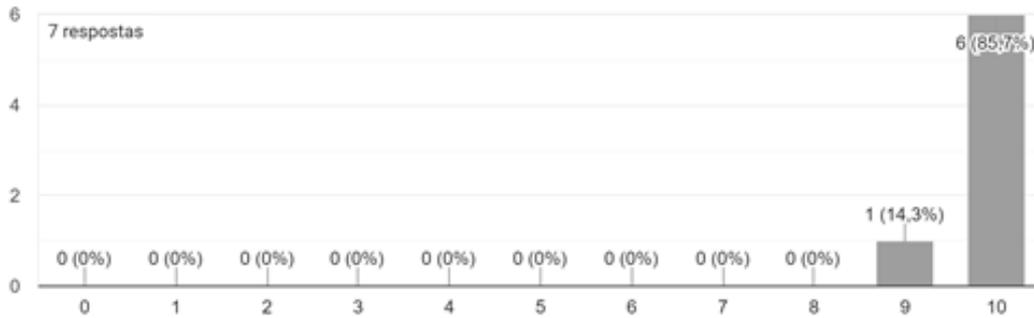
A respeito do processo gerado, conseguiu-se visualizar na Figura 41, em uma escala de 0 a 10, onde se define 5 em um ponto neutro, 0 como um resultado altamente negativo, e 10 altamente positivo, que o resultado médio avaliado pelos entrevistados é de que o processo pode ser utilizado em diferentes bases de dados, ainda que necessário certo nível de mudança.



Verifica-se também que após a adaptação das bases, a aplicabilidade das etapas de modelagem de dados e aplicação dos elementos visuais são completamente aplicáveis, como demonstrado na Figura 42.

Figura 42 - Avaliação da Aplicabilidade do processo

Defina em uma escala de 0 a 10, onde 5 é neutro, a aplicabilidade desse processo para visualização de dados

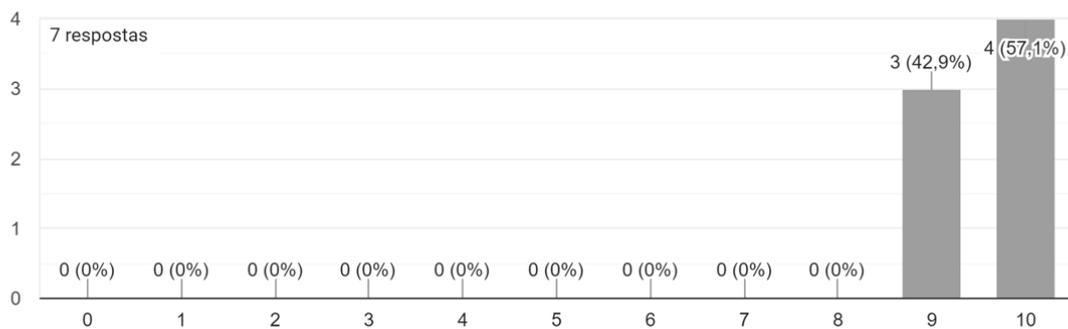


Fonte: Autoria própria (2022)

Os *dashboards* propostos são entendidos pelo público da pesquisa como positivamente significativos no impacto da qualidade das análises atuais, feitas de maneira manual, como demonstrado na Figura 43.

Figura 43 - Avaliação do resultado das análises

Defina em uma escala de 0 a 10, onde 5 é neutro, o impacto destes dashboards na qualidade das análises atuais

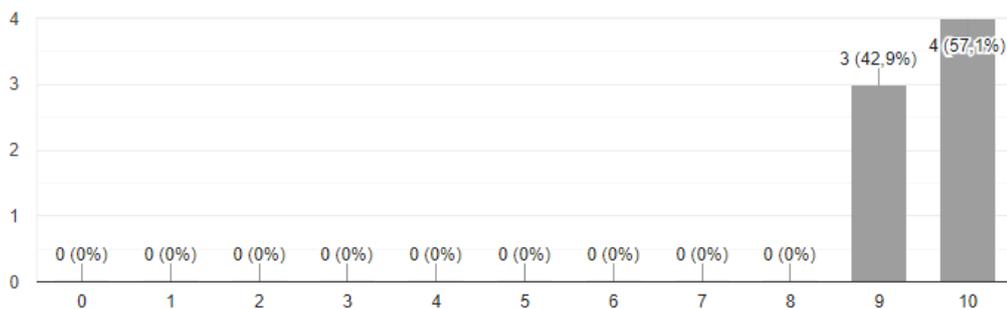


Fonte: Autoria própria (2022)

Estes *dashboards* podem ser considerados como elementos chave no aumento da eficiência no desenvolvimento destas análises, como avaliado pela Figura 44.

Figura 44 - Avaliação do impacto dos dashboards na eficiência

Defina em uma escala de 0 a 10, onde 5 é neutro, o impacto destes dashboards na eficácia das análises atuais



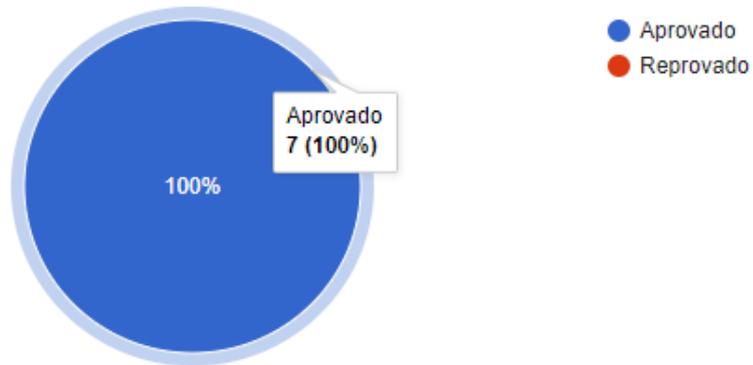
Fonte: Autoria própria (2022)

4.10 Revisão do Processo

Com base nas avaliações demonstradas no questionário da Figura 45, compreende-se que para o nicho de profissionais selecionado na pesquisa, o guia desenvolvido no Apêndice A atende aos requisitos e não necessita de mudanças consideráveis, desta maneira, validando o processo de desenvolvimento deste trabalho.

Figura 45 - Resultado geral do Procedimento
De maneira geral, o processo está aprovado para substituir o procedimento de análise atual?

7 respostas



Fonte: Autoria própria (2022)

5 CONCLUSÃO

Na indústria 4.0, a transformação de dados em conhecimento empresarial se tornou essencial para uma boa gestão de negócios. Com os avanços tecnológicos provenientes de décadas de inovação, recentemente se destaca a importância do conhecimento acerca de como tratar os dados e traduzir suas análises em informações que sejam compreensíveis para a organização.

Para grandes indústrias, o volume de dados envolvidos nas mais diversas operações pode ser valioso para compreensão do negócio e definição das estratégias das organizações. O uso de *dashboards* surge como um avanço tecnológico, permitindo que, com representações visuais, os gerentes e diretores tenham acesso a análises detalhadas que culminem em tomadas de decisões mais ágeis e assertivas dentro das organizações.

Sendo assim, este estudo permitiu alcançar o objetivo proposto de criação de um processo para a análise de bases de dados de projetos, utilizando os recursos da metodologia Crisp-DM, e gerando valor, ao permitir a visualização de forma consistente, precisa e rápida a indicadores importantes.

A partir da base de dados selecionada, seguindo os passos da metodologia Crisp-DM, foram definidos os dados de entrada, que após modelados, correlacionados, e atribuídos ao formato de um *dashboard*, permitiram a elaboração de um roteiro detalhado para análise de dados, que quando avaliado, através de um questionário ao público-alvo deste roteiro, permitiu averiguar a aplicabilidade do método.

Após avaliar as respostas ao roteiro desenvolvido, entendeu-se que ao estruturar as informações que se necessita em um *dashboard*, a análise se torna facilitada, e que a possibilidade de manter os visuais atualizados aplicando a mesma lógica, de maneira automática e sistemática, ao longo do tempo, permite evidenciar pontos que impactam nas decisões gerenciais em grandes indústrias. Com os dados e o conjunto de ferramentas apresentadas neste trabalho, se torna simplificado o processo de análise de dados, possibilitando tomadas de decisões mais ágeis e bem estruturadas.

Espera-se que este trabalho possa auxiliar futuros desenvolvimentos na área destacada, e propõe-se como futuros desenvolvimentos, a criação de roteiros que satisfaçam a elaboração de análises mais diversas de *dashboards*, sem a

necessidade de uma base de dados específica, bem como trabalhos que avaliem, através de estudos de caso, a eficácia do roteiro desenvolvido.

REFERÊNCIAS

AMARAL, Fernando. **Introdução à ciência de dados**: mineração de dados e big data. Rio de Janeiro: Alta Books, 2016. 320 p.

CAMILO, Cássio Oliveira; SILVA, João Carlos da. **Mineração de dados**: Conceitos, tarefas, métodos e ferramentas. Universidade Federal de Goiás (UFG), v. 1, n. 1, p. 1-29, 2009.

CAVALCANTE, Z. V.; SILVA, M. L. S. da. A importância da Revolução Industrial no mundo da Tecnologia. In: ENCONTRO INTERNACIONAL DE PRODUÇÃO CIENTÍFICA, 7. 2011. Maringá. **Anais eletrônico**. Maringá. 2011. Disponível em: https://www.unicesumar.edu.br/epcc-2011/wp-content/uploads/sites/86/2016/07/zedequias_vieira_cavalcante2.pdf. Acesso em: 12 jun. 2018.

CULOT, Giovanna; NASSIMBENI, Guido; ORZES, Guido; SARTOR, Marco. Behind the definition of Industry 4.0: analysis and open questions. **International Journal Of Production Economics**, [S.L.], v. 226, p. 107617, ago. 2020. Elsevier BV.

DATE, Christopher J. **Introdução a Sistemas de Bancos de Dados**. 8. ed. Brasil: Gen Ltc, 2004. 896 p.

DATHEIN, Ricardo. **Inovação e Revoluções Industriais**: uma apresentação das mudanças tecnológicas determinantes nos séculos XVIII e XIX. Publicações DECON Textos Didáticos 02/2003. DECON/UFRGS, Porto Alegre, Fevereiro 2003.

FUNES, Marcio Maestrello. BANCO DE DADOS MULTIDIMENSIONAIS: análise dos conceitos fundamentais e sua importância na tomada de decisão no ambiente corporativo. **Perspectivas em Ciências Tecnológicas**, São Paulo, v. 6, n. 6, p. 87-99, jun. 2017. Disponível em: <https://fatece.edu.br/arquivos/arquivos-revistas/perspectiva/volume6/6.pdf>. Acesso em: 11 nov. 2021.

GARTNER *et al* (ed.). **Critical Capabilities for Analytics and Business Intelligence Platforms**. 2021. Gartner. Disponível em: <https://www.gartner.com/en/documents/3999300>. Acesso em: 5 nov. 2021.

GESSERT, Amanda. **What's Next for Business Intelligence?**: storytelling. Storytelling. 2016. Merkle. Disponível em: <https://www.merkleinc.com/blog/whats-next-business-intelligence-storytelling>. Acesso em: 30 out. 2021.

INTERNET ENGINEERING TASK FORCE. **RFC 4180**: Common Format and MIME Type for Comma-Separated Values (CSV) Files. 1 ed. New York, 2005. Disponível em: <https://www.ietf.org/rfc/rfc4180.txt>. Acesso em: 21 out. 2022.

LASI, Heiner; FETTKE, Peter; KEMPER, Hans-Georg; FELD, Thomas; HOFFMANN, Michael. Industry 4.0. **Business & Information Systems Engineering**, [S.L.], v. 6, n. 4, p. 239-242, 19 jun. 2014. Springer Science and Business Media LLC.

LIMA, Luis Octavio. **CRISP-DM: conceitos**. Conceitos. 2020. Disponível em: <https://blog.grancursosonline.com.br/processo-de-mineracao-de-dados-com-crisp-dm/>. Acesso em: 11 nov. 2021.

LOPES, Charles Luís Rodrigues; GARCIA, Marcus Valério Rocha; DE ASSUMPÇÃO, Thiago Alexandre Alves. **AS REVOLUÇÕES INDUSTRIAIS E O SURGIMENTO DO PRO-LETARIADO URBANO**. Brasil Para Todos-Revista Internacional, v. 8, n. 1, p. 22-26, 2020.

MARQUESONE, Rosangela. **Big Data: Técnicas e tecnologias para extração de valor dos dados**. Editora Casa do Código, 2016.

MICROSOFT (Redmond). **Documentação do Power BI**. 2021. Disponível em: <https://docs.microsoft.com/pt-br/power-bi/>. Acesso em: 15 nov. 2021.

MOHAJAN, Haradhan Kumar. The First Industrial Revolution: creation of a new global human era. **Journal Of Social Sciences And Humanities**. Bangladesh, p. 377-387. out. 2019.

NEGASH, Solomon; GRAY, Paul. Business Intelligence. **Handbook On Decision Support Systems 2**, [S.L.], p. 175-193, 2008. Springer Berlin Heidelberg.

NONAKA, Ikujiro; TAKEUCHI, Hirotaka. **Criação do conhecimento na empresa: como as empresas japonesas geram o conhecimento na empresa**. 7. ed. Rio de Janeiro: Campus, 1997. 358 p.

NUMERRO (org.). **The Complete Guide to Power BI Visuals + Custom Visuals**. 2021. Disponível em: <https://www.numerro.io/guides/power-bi-visuals-guide>. Acesso em: 3 nov. 2021.

PAIXÃO, Maik. **Good graphics tell good stories: how do you create powerful, compelling stories from cold, raw digital data?**. 2017. Disponível em: <https://bootcamp.uxdesign.cc/good-graphics-tell-good-stories-9d41cf12dd51>. Acesso em: 12 nov. 2021.

PONTOTEL (Brasil). **Full-time Equivalent FTE: para que serve este indicador, como calcular e quais os benefícios!**. para que serve este indicador, como calcular e quais os benefícios!. 2021. Disponível em: <https://www.pontotel.com.br/full-time-equivalent-fe>. Acesso em: 21 out. 2022.

RAMM, Saskia; KOPF, Eva-Maria; DINTER, Barbara; HÖNIGSBERG, Sarah. What Makes a Good Story - The Use and Acceptance of Storytelling in Business Intelligence. **Proceedings Of The 54Th Hawaii International Conference On System Sciences**, [S.L.], v. 54, n. 1, p. 5706-5715, jan. 2021. Hawaii International Conference on System Sciences.

RANJAN, Jayanthi. BUSINESS INTELLIGENCE: concepts, components, techniques and benefits. **Journal Of Theoretical And Applied Information Technology**. Ghaziabad, p. 60-70. jan. 2009.

SACOMANO, José Benedito; SILVA, Márcia Terra da; GONÇALVES, Rodrigo Franco; BONILLA, Sílvia Helena; SÁTYRO, Walter Cardoso (org.). **Indústria 4.0: conceitos e fundamentos**. São Paulo: Blucher, 2018.

SENA, Jhonatan Kayo Silva. **Um estudo sobre a importância da tecnologia big data e perspectivas futuras de sua aplicação na indústria**. 2019.

SCHERMANN, Michael; HEMSEN, Holmer; BUCHMÜLLER, Christoph; BITTER, Till; KRCMAR, Helmut; MARKL, Volker; HOEREN, Thomas. Big Data. **Business & Information Systems Engineering**, [S.L.], v. 6, n. 5, p. 261-266, 4 set. 2014. Springer Science and Business Media LLC.

SCHWAB, Klaus. **The fourth industrial revolution**. Currency, 2017.

STEARNS, Peter N. Introduction: defining the industrial revolution. In: STEARNS, Peter N. **The Industrial Revolution in World History**. 5. ed. New York: Routledge, 2020. Cap. 1. p. 1.

TERRA, José Cláudio Cyrineu. **Gestão do conhecimento: o grande desafio empresarial**. Terra Fórum Consultores, 2005. Acesso em: 14 out. 2022.

TETILA, Everton Castelão. **Modelagem de dados com Data Warehouse e OLAP: um estudo de caso**. EaD & Tecnologias Digitais na Educação, Dourados, v. 2, n. 3, p. 20-29, nov. 2014. ISSN 2318-4051. Disponível em: <https://ojs.ufgd.edu.br/index.php/ead/article/view/3514>. Acesso em: 12 nov. 2021.

WOOD Junior, T. (1992). **Fordismo, Toyotismo e Volvismo: os caminhos da indústria em busca do tempo perdido**. Revista de Administração de Empresas, 32(4), 6–18.

APÊNDICE A - Procedimento para análise de projetos

PROCEDIMENTO PARA ANÁLISE DE PROJETOS

DASHBOARDS DE ANÁLISE DE PROJETOS

Baseado na metodologia Crisp-DM esse passo a passo tem como objetivo principal, guiá-lo na construção de um dashboard na ferramenta Power BI da Microsoft que possibilite análises e tomadas de decisões rápidas, acerca de bases de dados de projetos.

1. Você já utilizou o Power BI?

NÃO?

O que você acha de aprender um pouco sobre Power BI primeiro?

Você pode acessar tutoriais práticos e rápidos no site da Microsoft: [Power BI no Microsoft Learn](#)

E também baixar o software nesse link: [Power BI Desktop](#)

Agora que você já sabe como funciona, vamos modelar os dados?

2. Mapeie e entenda os dados que você tem disponível

1. Tabela de Projetos: Deve conter 1 linha por projeto, apenas informações únicas.
Exemplo de Colunas: Data de criação, Responsável, Nome, Status

2. Tabela de Demandas: Pode conter múltiplas linhas para o mesmo projeto em um período.
Exemplo de Colunas: País, Cidade, Time, Categoria

3. Tabela de Capacidades: Contem uma única linha por time em um período de tempo. Não deve ser associada a nenhum projeto.

4. Ao lado está um exemplo de dados. Eles não são exclusivos, você pode usar outros. O importante é seguir a lógica.

5. Temos 3 tabelas, nas quais:

- **Projetos:** É a tabela principal e contém todos os detalhes únicos a respeito de um projeto. A relação entre essa e as demais é pela 'identificação do projetos'.
- **Demandas:** Mapeie todas as demandas de um projeto para cada time envolvido sendo assim, diversas linhas pertencem ao mesmo projeto.
- **Capacidades:** Similar a tabela de demandas, devendo conter a capacidade dos mesmos times

3. Defina a correlação entre as tabelas

Você pode fazer essas relações até mesmo manualmente. A dica é entender quais são os itens obrigatórios.

Projetos	Obrigatorio
Nome	X
Identificação do Projeto	X
Status do Projeto	X
País	X
Categoria	X
Grupo	X
Time	X
Impacto Financeiro	X
Data de Finalização Esperada	X
Cidade	X
Cadeado	X

Demanda	Obrigatorio
Time	X
Demanda	X
Trimestre	X
Identificação do Projeto	X

Capacidade	Obrigatorio
Time	X
Demanda	X
Trimestre	X

4. Importe os dados para o Power BI

Crie uma página em branco e importe seus arquivos um por um, seguindo os passos da figura

5. Abrir em 'transformar dados' para modelagem

Essa função abrirá o Power Query

6. No Power Query

Tabelas importadas

Todas as etapas de transformação aplicadas a tabela

Quando um arquivo é importado, o Power BI automaticamente tenta fazer alguns passos, como transformar a primeira linha em cabeçalho, e alterar as colunas para como números inteiros, decimais, textos ou datas

Você pode desfazer qualquer uma destas etapas a selecionando com o botão direito e deletando

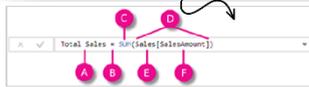
Nas tabelas do exemplo, as principais alterações envolvem:

- Promover a primeira linha para cabeçalho das colunas;
- Alterar o tipo de dado da coluna e definir colunas numéricas;
- Remover valores duplicados na tabela de projetos;
- Renomear novas colunas.

7. Criação de tabelas de índice

VOCÊ SABE O QUE É DAX? VEM APRENDER!!

Para evitar situações de referências circulares entre as tabelas, não é aconselhável que se conectem duas colunas contendo múltiplas linhas com o mesmo dado, por isto, se faz necessário a criação de **tabelas virtuais**, a serem utilizadas como **índices**, ou **tabelas de dimensões**, que filtrem as **tabelas de fatos**.



- Esta fórmula inclui os seguintes elementos de sintaxe:
- R.: O nome de medida. Total Sales.
 - B. O operador de sinal de igual (=), que indica o início da fórmula. Quando calculada, ela retornará um resultado.
 - C. A função DAX SUM, que soma todos os números da coluna Sales[SalesAmount]. Você aprenderá mais sobre as funções mais tarde.
 - D. Os parênteses (), que envolvem uma expressão que contém um ou mais argumentos. A maioria das funções exige pelo menos um argumento. Um argumento transmite um valor para uma função.
 - E. A tabela referenciada, Sales.
 - F. A coluna referenciada, [SalesAmount], na tabela Sales. Como agregar uma SUM.

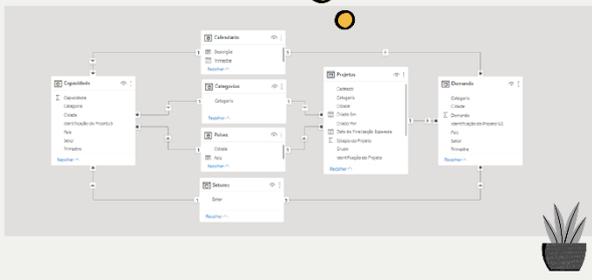
A linguagem DAX (Data Analysis Expressions) é uma coleção de funções para calcular e retornar um ou mais valores. Resumindo, o DAX ajuda a criar informações de dados já presentes em seu modelo, como uma função de soma.

Exemplos de expressões DAX utilizadas no projeto

1. Categorias = DISTINCT(Capacidade(Categoria))

1. Setores = DISTINCT(Capacidade(Grupo de Trabalho))

8. Criar as correlações entre as tabelas igual as mapeados no passo 3



9. Desenvolver medidas para os KPI's

A partir dos dados das tabelas, 3 KPIs foram definidos como principais medidas de análise:

- Carga de trabalho;
- Eficiência;
- E ranking dos maiores projetos.

$$\text{Carga de trabalho (\%)} = \frac{\sum \text{Demanda (FTE)}}{\sum \text{Capacidade (FTE)}}$$

$$\text{Eficiência (R$/FTE)} = \frac{\sum \text{Impacto Financeiro (R\$)}}{\sum \text{Demanda (FTE)}}$$

$$\text{Impacto top 20 (\%)} = \frac{\sum_{\text{20 Maiores Proj.}} \text{Impacto Financeiro}}{\sum_{\text{Todos os projetos}} \text{Impacto Financeiro}}$$

As equações são baseadas em FTEs (Full Time Employee), com isso é analisado o impacto financeiro e demandas pelo percentual de colaboradores. Você deve usar os KPIs que fazem sentido para seus objetivos de análise

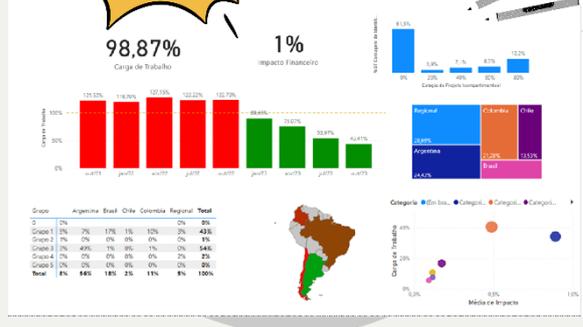
10. Como representar os dados

Esses são os visuais que a plataforma disponibiliza

Nesse momento é importante pensar como criar um **storytelling** com os dados

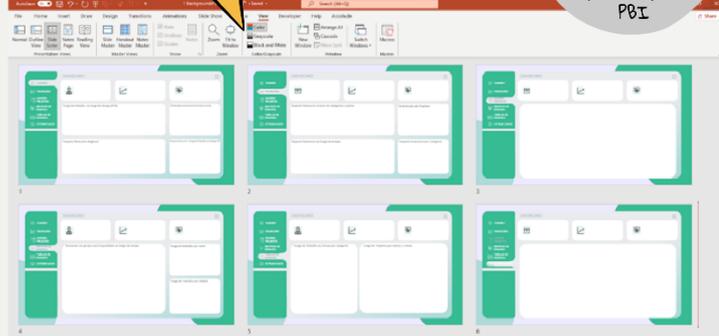
Storytelling é estruturado em técnicas que visam destacar pontos que agregarão mais valor. O procedimento é fundamentado em estratégia de visualização.

11. Criar drafts dos dados e seus visuais



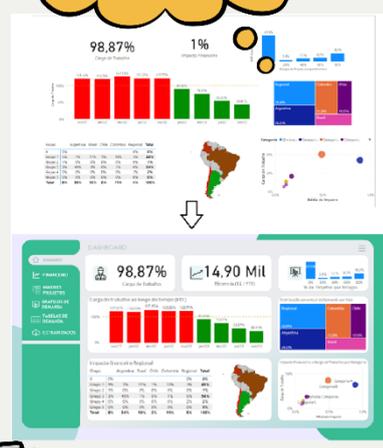
12. Criar um plano de fundo para adequar os visuais

O Power Point é uma ótima ferramenta para isso. Depois é só exportar para o PBI



13. Posicionar os dados nos visuais no Power BI

14. Criar segmentadores de dados



15. Criar dicas de ferramentas



Agora está pronto para você fazer as melhores análises do seu negócio!



Publique e defina uma frequência e forma de atualização dos dados



Não se esqueça de dar a sua opinião no forms! É anônimo, clique no link:

<https://forms.gle/USEtsPo4iWEmzXwyZ>

Para mais informações, acesse o projeto:

[TCC2_Eduardo e Ellen V2.docx](#)



APÊNDICE B - Questionário de validação do procedimento

QUESTIONÁRIO DE VALIDAÇÃO DO PROCEDIMENTO

Avaliação do Roteiro de **Análise de Dados de Projetos**

Olá!! Este questionário faz parte do trabalho de conclusão do curso de Engenharia Mecânica dos acadêmicos Ellen C. Lestensky Pereira e Eduardo Rodrigues de Oliveira. A finalidade dele é **qualificar como o roteiro de análise de dados é eficaz e aplicável** nas avaliações de informações referentes a projetos. Esse procedimento faz parte dos resultados esperados do nosso projeto e seu feedback é muito importante! A sua avaliação é anônima e, antes de começarmos, conte sobre sua experiência:

Disponível neste link: ... Fazer no Canva

dudu.r.oliveira@gmail.com (não compartilhado) [Alternar conta](#)

***Obrigatório**

Quais são as suas áreas de conhecimento? *

Dentro de empresas de bens de consumo

Análise de dados

Tomada de decisão

Utilização de dashboards para visualização de dados

Outro: _____

Quantos anos de experiencia você tem nestas áreas? *

Em anos

3 _____

Você confirma ter lido todo o documento disponibilizado? *

Sim

Próxima Página 1 de 2 [Limpar formulário](#)

Avaliação do Roteiro de **Análise de Dados de Projetos**

dudu.r.oliveira@gmail.com (não compartilhado) [Alternar conta](#)

***Obrigatório**

Avaliação do Roteiro de análise de dados de projetos

Defina em uma escala de 0 a 10, onde 5 é neutro, o quão fácil é adaptar sua base atual a utilizada no estudo *

0 1 2 3 4 5 6 7 8 9 10

Muito difícil Muito Fácil

Defina em uma escala de 0 a 10, onde 5 é neutro, a aplicabilidade desse processo para visualização de dados *

0 1 2 3 4 5 6 7 8 9 10

Pouco aplicável Totalmente aplicável

Defina em uma escala de 0 a 10, onde 5 é neutro, o impacto destes dashboards na qualidade das análises atuais *

0 1 2 3 4 5 6 7 8 9 10

Impacto altamente negativo Impacto altamente positivo

Defina em uma escala de 0 a 10, onde 5 é neutro, o impacto destes dashboards na eficácia das análises atuais *

0 1 2 3 4 5 6 7 8 9 10

Pouco eficaz Totalmente eficaz

De maneira geral, o processo está aprovado para substituir o procedimento de análise atual? *

Aprovado

Reprovado

Se a resposta for negativa, justifique

Sua resposta _____

Voltar **Enviar** Página 2 de 2 [Limpar formulário](#)