

AGREGAÇÃO DE TABELAS FATOS – ENRIQUECIMENTO E NOVAS PERCEPÇÕES EM MODELO DE BUSINESS INTELLIGENCE

Silva, Klivio Rafael Nunes e ¹

Medeiros, Fábio Nicácio de ²

RESUMO

O *Business Intelligence - BI* surgiu da necessidade de responder perguntas e guiar o desenvolvimento no ambiente empresarial, utilizando e fundamentando as tomadas de decisão a partir dos dados, sendo inegável o seu sucesso e benefícios para as corporações. Este artigo discute o processo de modelagem dimensional em duas bases de dados pública, sendo aplicados conjuntos de técnicas utilizadas na construção de projetos em *data warehouse*, como também a aplicação da técnica de agregação de tabelas fatos, gerando um único objeto com dados resultantes da união. Posteriormente foram construídos 6 *dashboards*, no *Microsoft PowerBI*, a partir do trabalho realizado. Logo, conhecer tecnologias, modelos, técnicas e aplicações utilizadas no cotidiano do *BI*, tornam o profissional diferenciado, com habilidades únicas e de destaque no mercado de trabalho.

Descritores: Análise de Dados. Apresentação de Dados. *Business Intelligence*.

ABSTRACT

Business Intelligence - BI arose from the need to answer questions and guide development in the business environment, using and supporting decision-making based on data, its success and benefits for corporations being undeniable. This article discusses the dimensional modeling process in two public databases, applying sets of techniques used in the construction of projects in data warehouse, as well as the application of the fact table aggregation technique, generating a single object with data resulting from the union. Subsequently, 6 dashboards were built in Microsoft PowerBI, based on the work carried out. Therefore, getting to know technologies, models, techniques and applications used in the daily life of BI, make the professional differentiated, with unique skills and outstanding in the job market.

Descriptors: Data Analysis. Data Presentation. Business Intelligence.

1. Discente do curso de Sistemas para Internet – UNIESP. E-mail: kliviorafael@hotmail.com
2. Docente do curso de Sistemas para Internet e Sistemas para Informação – UNIESP. E-mail: fabio.medeiros@iesp.edu.br

1. Introdução

No decorrer do tempo, as sociedades passaram por transições, as quais influenciaram mudanças em diversos setores como o econômico, cultural, político e social. Neste contexto o ambiente empresarial buscou por adaptações, com o intuito de se antecipar a estas variantes, obtendo sucesso nos negócios, estabilidade no mercado e manter constante crescimento financeiro.

Deste modo, o *Business Intelligence - BI* surge com propostas de tecnologias, aplicações e processos para tratar, organizar e analisar os dados, transformando em informação, o qual melhora as práticas e performances corporativas, quando as tomadas de decisão são baseadas no conhecimento gerado pelo BI, as quais historicamente atestam seu sucesso e comprovam os benefícios da sua utilização (GAARDBOE; JONASEN, 2018).

A implementação do *BI* nas indústrias e organizações é complexa e apresenta um alto custo, devido a necessidade de tecnologias, softwares, infraestruturas, licenças, além de treinamentos visando a capacitação de mão de obra, cada vez mais especializada para o exercício da atividade (GAARDBOE; JONASEN, 2018).

Neste sentido, o *Business Intelligence* consegue ofertar uma percepção sistemática do negócio, devido a possibilidade de distribuir os dados de forma concreta, a partir de cruzamento das referências, gerando diferentes visualizações e uma ampla forma de modelar e analisar os dados, transformando em informação útil para fundamentar tomadas de decisão, e controlar o desempenho empresarial (ROSSETI *et al.*, 2008).

Sendo assim, a presente produção viaja pela história do BI, ressaltando seu relacionamento e importância com o desenvolvimento comercial, como também os principais nomes e suas contribuições para a área. Posteriormente é apresentado um estudo de caso, onde foi aplicado conceitos utilizados para modelagem e análises, além de técnicas aplicadas em tabelas, no intuito de enriquecer os dados e proporcionar melhores análises.

Tem como objetivo, abordar sobre o processo de criação de um *data warehouse* a partir da extração de quantitativos de duas massas de dados públicas, e de que conhecendo e utilizando o processo de agregação de objetos, realizar a junção

de duas tabelas fatos, resultante da aplicação das modelagens dimensionais tipo estrela, fornecendo uma visão única, em uma tabela fato agregada, com maiores possibilidades de obtenção de informação e tomadas de decisão a partir dos dados.

2. Fundamentação Teórica

2.1 Comércio e Desenvolvimento Social

Inicialmente, antes de abordar sobre o *Business Intelligence*, se faz necessário compreender a importância do seu objeto de trabalho para a humanidade, e como o mesmo, ainda é um valioso mecanismo que impulsiona o desenvolvimento social.

O comércio permeia o progresso de diversas sociedades no decorrer do tempo, desde as primeiras práticas de escambo ou as primeiras feiras datadas ainda na Idade Média, sendo presente na influência de novos modos de organização urbana, a exemplo dos burgos e na pluralidade social existente destes ambientes, na fundação e expansão de importantes cidades, e no contexto econômico, na criação das primeiras moedas (FREIRE, 2010).

O progresso proporcionado neste cenário foi refletido em diversos setores, servindo de alicerce para a busca de melhorias científicas a exemplo da própria ciência matemática, que ampliou e aprimorou consideravelmente a complexidade de suas operações, a medida que as demandas contábeis exigiam maiores conhecimentos. Como também o ambiente tecnológico, o qual recebeu um enorme incentivo, registrado na indústria e nas suas revoluções datadas durante o decorrer do tempo, as quais mudaram a forma de consumir e produzir (PEREIRA; LUNARDI, 2009).

2.2 História do Business

A ideia do *Business Intelligence* surgiu em meados de 1865, quando houve a primeira escrita com temática relacionada, por Richard Miller Devens, no seu livro "*Cyclopaedia of Commercial and Business Anecdotes*", o qual descreve o caso de sucesso de um banqueiro, que utilizando do conhecimento e identificando informações do comércio, conseguiu obter lucro, além de agir sempre a frente quando comparado aos seus concorrentes (BOTOS, 2018).

Esse fato levantou um importante ponto, no que tange a confiar mais em padrões e evidências originadas dos dados, para desenvolver uma estratégia de negócio, comparado a instinto ou intuição.

Após um longo período de tempo, ocorreram importantes avanços tecnológicos a exemplo do surgimento dos primeiros computadores, melhorias consideráveis de processamento e o auxílio das máquinas para realizar análises cada vez mais rápidas e complexas, o *BI* encontrou ambiente favorável para poder se desenvolver (QUEIROZ; DE BARROS CARVALHO, 2020).

Em 1958, Hans Peter Luhn, pesquisador da IBM publica um artigo intitulado “*A Business Intelligence System*”, no qual pontua sobre as capacidades de um sistema de aprender e prever com base nos interesses do usuário, permitindo obter várias tendências de inteligência de negócios (BOTOS, 2018).

De certa forma, Luhn foi responsável por construir o conceito do que entendemos hoje sobre *Business Intelligence*, e associada a essa definição, o constante desenvolvimento e aprimoramento de tecnologias, processo este que já vinha ocorrendo, permitiu maior flexibilidade e melhor inserção do BI no ambiente tecnológico (LUHN, 1958).

A otimização espacial e de processamento dos novos computadores, sendo estes menores e mais potentes, quando comparados aos seus antecessores, foi um fator considerável para o sucesso da implementação do que na época ainda era chamada de Sistemas de Suporte à Decisão - DSS (QUEIROZ; DE BARROS CARVALHO, 2020).

Contudo, o surgimento dos bancos de dados relacionais da terceira geração, os quais apresentavam enorme capacidade de armazenar e manipular dados, foi a peça chave para o sucesso, associando - se a cada vez mais casos da aplicação do *BI*, e conseqüentemente, do uso lucrativo dos dados nas tomadas de decisão (QUEIROZ; DE BARROS CARVALHO, 2020).

2.3 A Ciência do BI, entre estrelas e flocos de neve

A construção de todo um ecossistema tecnológico que favoreceu o desenvolvimento do *business intelligence*, como também a percepção executiva do

sucesso ao utilizar o conhecimento proporcionado para os negócios, o *BI* precisou passar por organizações e definições na busca por padronizar o método de trabalho.

Ralph Kimball e Bill Inmon propuseram ideias diferentes, porém convergem no objetivo de organizar e propor um ambiente exclusivo para o armazenamento de dados, os quais serão utilizados na análise do negócio, o *data warehouse* – *DWH* (YESSAD; LABIOD, 2016).

Esse ambiente “exclusivo”, é um banco de dados, que agrega os fluxos oriundos de outras fontes, necessários para a execução do trabalho, o qual permite uma ampla capacidade de referenciar diferentes origens, além de possibilitar um processo analítico mais profundo, cruzando informações e enriquecendo o resultado final (YESSAD; LABIOD, 2016).

A divergência existente entre Kimball e Inmon é referente a proposta de modelagem para ser aplicada no *data warehouse*, o qual tem como referência o modelo dimensional, onde as entidades são estruturadas em um modelo arquitetônico, de forma mais eficiente frente as necessidades analíticas e melhora consideravelmente o desempenho, como também o acesso ao dado pelo negócio ou cliente é facilitado, desta forma, temos uma otimização no armazenamento, como também na disponibilização do dado (YESSAD; LABIOD, 2016).

O modelo proposto por Kimball, chamado de *Star Schema* (Modelo estrela), é de longe o mais difundido e utilizado para *data warehouse* atualmente, tem como principal característica, a síntese em definir e simplificar as dimensões e fatos, e por consequência os seus relacionamentos, proporcionando diferentes perspectivas de análise para o usuário, sendo capaz de observar e filtrar os dados presente na tabela fato, de acordo de cada dimensão (SILVA; SARTORI, 2015).

Em antagonismo a Kimball, Inmon propõe um modelo um pouco mais complexo, o *Snow Flake* (Floco de Neve), que visa reduzir a redundância no armazenamento e gestão dos dados no *data warehouse*. Esse objetivo, é devido a esse tipo de modelo ainda tentar se aproximar de modelos aplicados em sistemas transacionais, porém a elevada complexidade para aplicar o mesmo, acarreta na dificuldade de administrar o ambiente, como também a consulta da informação pelo usuário (YESSAD; LABIOD, 2016).

2.4 Estrutura da Modelagem Dimensional

Independente do autor e do tipo de modelagem proposta, é possível observar uma convergência na necessidade da aplicação do modelo dimensional na arquitetura do *DWH*. Nessa nova estrutura são construídos e definidos objetos com o intuito de abrigar entradas e características do negócio analisado, como também métricas importantes para as tomadas de decisão (SILVA; SARTORI, 2015).

O primeiro objeto, e se não o de maior importância no modelo é a tabela Fato, pois tem a função de armazenar o acontecimento, o ocorrido, sendo um pouco redundante, guarda o fato propriamente dito, tem como principal característica obrigatória a existência de relações com as tabelas dimensionais, e em sua estrutura física a existência de campos com métricas e *Foreign Keys* (PITON, 2018).

Em sequência, a tabela dimensão tem a função de descrever o fato ocorrido, caracterizando o evento, com o objetivo de classificar e descrever os dados que estão presente na tabela fato, ou seja, ela que responde os questionamentos do negócio, quem foi? Quando foi? Quanto foi? Ou qualquer outra dúvida existente e que precisa ser respondida. Sua estrutura física temos a existência de 3 itens a *Surrogate Key*, a *Natural Key* e seus atributos (PITON, 2018).

Além desses dois objetos claramente estabelecidos e necessários para implementar um modelo dimensional, ainda existem outros pontos que devem ser considerados e que influênciam o conteúdo produzido e suas futuras análises. Observar a granularidade dos dados, as possíveis agregações que podem ser aplicadas, no intuito de melhorar o desempenho das consultas, como também as relações estabelecidas entre os objetos, faz parte de todo um contexto para o sucesso do *DWH*, e do projeto como um todo (BRUZAROSCO *et al.*, 2000).

3. Metodologia

Foi desenvolvido um projeto de *business intelligence* a partir de duas bases de dados distintas, ambas de origem pública, a primeira denominada “*Vacinação Brasil microdados*”, a qual apresenta dados gerados durante o processo de vacinação do Covid – 19, distribuídos por estabelecimento, população e vacina aplicada pelo território nacional, neste documento ela é identificada como Base Vacina.

A segunda base é fornecida pelo IBGE intitulada “*População residente, total e respectiva distribuição percentual, por situação do domicílio e sexo e razão de sexo, segundo os municípios e as classes de tamanho da população dos municípios – Brasil - 2010*” e similar a base anterior, neste documento será identificada como Base IBGE.

Foram aplicados os conceitos da modelagem dimensional do tipo modelo estrela nas duas bases anteriormente citadas, logo em seguida, após a construção dos dois modelos, as tabelas fatos criadas passaram por um processo de agregação, resultando em uma terceira fato, a qual é denominada durante a execução do projeto de fato agregada.

Em paralelo foi realizado um levantamento bibliográfico para auxiliar na construção de conhecimento e tomada de decisão durante a execução do projeto, fundamentando e contextualizando os procedimentos técnicos e necessários para o sucesso da atividade.

As bases de dados foram inseridas no banco de dados *Oracle SQL Developer* versão 21.2.1.204.1703, onde se encontra o *data warehouse* do projeto, como também todos os objetos criados resultantes da modelagem aplicada.

Finalizado este processo, foram elaborados *dashboards* no total de 6 painéis, utilizando como tecnologia o programa *Microsoft Power BI*, onde também foram desenvolvidas e aplicadas algumas agregações adicionais a nível de linguagem *DAX*.

3.1 Estudo de Caso

3.1.1 Modelo Estrela – Base Vacina

Inicialmente foram selecionados de forma aleatória um quantitativo de registros, para compor a Base Vacina, visto que devido a sua extensão, havia a possibilidade desenvolver problemas de performance durante a execução da modelagem, logo, a base trabalhada é composta por 856.090 linhas.

Em primeira análise, foi entendido a existência de 4 grupos em que os dados, se encontram distribuídos da seguinte forma:

- **Grupo 1** – Agrega dados referente a Vacina, com os campos: código_vacina, vacina, sistema_origem, vacina_lote, numero_dose;

- **Grupo 2** – Agrega dados referente ao Paciente, com os campos: paciente_código_grupo, paciente_código_subgrupo, paciente_grupo, paciente_subgrupo, paciente_cep, paciente_município, paciente_etnia, paciente_idade, paciente_nacionalidade, paciente_uuid, paciente_sexo_biológico, paciente_unidade_federativa, paciente_codigo_pais, paciente_pais, paciente_codigo_ibge_município, paciente_codigo_etnia;
- **Grupo 3** – Agrega dados referente ao Estabelecimento – local onde foi aplicada a vacina, com os campos: estabelecimento_razao_social, estabelecimento, estabelecimento_município, estabelecimento_codigo_ibge_município, estabelecimento_código_cnes, estabelecimento_unidade_federativa;
- **Grupo 4** – Agrega dados referente ao ato de vacinar – dados gerados durante a vacinação, com os campos: documento_uuid, data_aplicação;

3.1.2 Definindo Dimensões – Base Vacina

Após a compreensão e subdivisão da Base Vacina, já é possível definir e estruturar ainda a nível lógico as dimensões iniciais do modelo, e desta forma conseguir progredir na criação do *DWH* e no tratamento dos dados.

Foram pensadas e criadas dimensões verdadeiras, em que houve o desenvolvimento de objetos no banco, como também dimensões degeneradas, que são dados com comportamento e perfil de dimensão, mas podem ficar diluídos na tabela fato por não ter muito impacto no coletivo, devido à baixa capacidade de alterar seu conteúdo, por ter um comportamento mais estático.

Desta forma, as dimensões “verdadeiras” somam no total de 8 tabelas, além das 3 dimensões degeneradas e 2 dimensões *mockadas*, as quais vão ser explicadas, posteriormente a tabela abaixo.

| Dimensão | Campo da Base Vacina |
|--------------------|------------------------------|
| Paciente Grupo | Paciente_Codigo_Grupo |
| | Paciente_Grupo |
| Paciente Sub_Grupo | Paciente_Codigo_Subgrupo |
| | Paciente_Subgrupo |
| Estabelecimento | Estabelecimento_Codigo_Cnes |
| | Estabelecimento_Razao_Social |

| | |
|-------------------|---------------------------------------|
| | Estabelecimento |
| Estado | Estabelecimento_Unidade_Federativa |
| Município | Estabelecimento_Codigo_Ibge_Município |
| Sistema de Origem | Codigo_Sistema_Lote |
| Vacina Lote | Codigo_Lote |
| Data | Data_aplicação |
| Vacina | Codigo_Vacina |
| | Vacina |

Tabela 1 - Definição das dimensões verdadeiras da Base Vacina.

Para as dimensões Estado, Município, Sistema de Origem, Vacina Lote, foram criados campos com códigos identificadores para abrigar e ordenar os registros oriundos da base.

Em sequência e já previamente citado, foram definidas 3 dimensões degeneradas, as quais foram tratadas diretamente na tabela fato do modelo, descritas a seguir.

| Dimensão | Campo na Base Vacina | | | | | | |
|----------------|-------------------------|--------|---------|----------|---------|----------|---------|
| Sexo_Biológico | Paciente_Sexo_Biológico | | | | | | |
| | Valores em Tabela | M | | F | | | |
| | Valores Degenerados | 1 | | 2 | | | |
| Etnia | Paciente_código_etnia | | | | | | |
| | Valores em Tabela | Branca | Preta | Parda | Amarela | Índigena | (Vazio) |
| | Valores Degenerados | 1 | 2 | 3 | 4 | 5 | 99 |
| País | Paciente_código_pais | | | | | | |
| | Valores em Tabela | Brasil | Uruguai | Portugal | Ruanda | | (Vazio) |
| | Valores Degenerados | 10 | 25 | 45 | 221 | | -1 |

Tabela 2 - Definição das dimensões degeneradas da Base Vacina

Por fim, as dimensões *mockadas* foram atribuídas aos campos paciente_idade e sequência_dose, ambos na base são do tipo numérico, e se trabalhados neste

formato, poderiam ser sumarizados, e por consequência ofertando informações para o usuário de forma errônea. Portanto seus tipos foram alterados para string, impedindo qualquer tipo de agregação e mantendo a integridade do planejamento para a base.

Após essas definições, os objetos foram criados em banco e devidamente preenchidos, com valores distintos de cada campo previamente determinado da Base Vacina, para as dimensões de campo único, como anteriormente citado, foram implementados campos com códigos identificadores, que eram inseridos de forma incremental a partir do volume do campo descritivo.

Para a construção da tabela fato deste modelo, inicialmente foi realizada a seleção dos campos que se relacionam com dimensões já mencionadas, como também a substituição de todos os valores nulos por -1, identificando como “não informado” em cada objeto dimensional.

Toda tabela fato, tem como principal característica a existência de métricas, analisando comportamento e composição da Base Vacina, existem dois campos que chamam atenção, e que tem potencial para serem os responsáveis para sustentar essas medidas, o paciente_uuid e documento_uuid.

O primeiro contém dados pessoais do paciente, como o nome ou registro, já o segundo campo, contém dados gerados durante o processo de vacinação, ou seja, um documento_uuid é criado toda vez que um paciente é vacinado em um estabelecimento.

Esta lógica é perceptível, pois um registro de paciente_uuid apresenta vários registros de documento_uuid relacionados, e como o evento é a vacinação, o campo número_dose se altera para cada vez que o paciente se vacina. Logo a métrica da tabela fato deste modelo é o documento_uuid.

| Campo na Tabela Fato | Dimensão Relacionada |
|-----------------------------|-----------------------------|
| SK_Data_Aplicacao | Dimensão Data |
| SK_SEQ_ETD | Dimensão Estado |
| SK_SEQ_MNP | Dimensão Município |
| SK_SEQ_ETB | Dimensão Estabelecimento |
| SK_SEQ_SIS_ORG | Dimensão Sistema de Origem |
| SK_SEQ_VCN | Dimensão Vacina |

| | |
|----------------|-------------------------|
| SK_SEQ_VCN_LT | Dimensão Vacina Lote |
| SK_SEQ_VCN_DS | Dimensão Vacina Dose |
| SK_SEQ_AGE | Dimensão Idade |
| SK_SEQ_PAIS | Dimensão País |
| SK_SEQ_GRP | Dimensão Grupo |
| SK_SEQ_SUBBGRP | Dimensão Subgrupo |
| SK_SEQ_ETN | Dimensão Etnia |
| SK_SEQ_TIP_SEX | Dimensão Sexo Biológico |
| QTD_VAC | Métrica da Tabela |

Tabela 3 - Composição da tabela Fato Vacina

3.1.3 Modelo Estrela – Base IBGE

O processo de construção do modelo dimensional da Base IBGE, foi um pouco mais ágil, quando comparado à modelagem aplicada na Base Vacina, pois algumas dimensões foram aproveitadas, como também algumas escolhas foram replicadas, como a exemplo do processo de degenerar dimensões para este novo modelo.

Inicialmente as dimensões estado e município, ambas elaboradas na Base Vacina, foram aproveitadas nesta nova construção, logo, os objetos se tornaram genéricos para os dois cenários (Vacina e IBGE), e sendo específico ditou a quantidade de registros presentes na tabela fato deste novo modelo, pois a Base Vacina não apresenta a totalidade de cidades no país, sendo usada como critério de seleção, para que, no final ambas as bases apresentem os mesmos municípios, o resultante não existente na Base Vacina foi desprezado.

Já o campo sexo biológico passou pelo processo de degeneração, com os mesmos objetivos já citados no processo construtivo do modelo da Base Vacina para essa situação.

A tabela original, fornecida pelo IBGE, apresentava vários campos que não teriam utilidade para o projeto, e na busca de otimizar o trabalho, foram excluídos todos os campos (colunas e linhas) não necessários, restando apenas o código_município_ibge, estado, valor total e valor por sexo (homem e mulher). Estes últimos estão presentes na tabela original em valores percentuais, sendo necessário um tratamento, multiplicando o percentual pelo valor total, para que fosse possível

obter os valores inteiros do quantitativo de homens e mulheres por município e por unidade federativa, quando vistos em maior granularidade.

Além de que, para realizar o devido pareamento da tabela fato com a dimensões (estado e município) os registros relacionados foram alterados pelos códigos identificadores, para efetiva junção.

Todos esses tratamentos acima citados foram realizados ainda a nível de ambiente *Microsoft Excel*, os quais posteriormente foram inseridos no ambiente Oracle, desta forma o desenho final da tabela Fato da Base IBGE ficou neste formato:

| Campo da Tabela Fato | Dimensão Relacionada |
|-----------------------------|-----------------------------|
| SK_SEQ_ETD | Dimensão Estado |
| SK_SEQ_MNP | Dimensão Município |
| TIP_SEXO | Dimensão Sexo Biológico |
| QTD_TOTAL | Métrica da Fato |
| QTD_SEXO | Métrica da Fato |
| DAT_INCLUSAO_BI | Dimensão Data |

Tabela 4 - Composição da tabela Fato Agregada

Após a elaboração e implementação dos dois modelos no *DWH*, foi iniciado o processo de definição do último objeto do projeto, a tabela fato agregada, a qual é o resultado da união da tabela fato da Base IBGE, com a tabela fato da Base Vacina.

A regra para a agregação é que a tabela resultante da união, deve conter os mesmos campos, ou seja, apenas é possível unir a uma única tabela quando estas forem compatíveis em igualdade de campos.

Ao olhar as tabelas fatos dos modelos implementados, é possível perceber que a tabela do modelo IBGE é mais limitada do que a tabela do modelo Vacina, logo ela que irá ditar os campos que irão ser transportados para a tabela agregada.

Vale ressaltar, que para essa tabela, além dos campos iguais, foram adicionadas a métrica da tabela fato Vacina, para que a quantidade de pessoas vacinadas, possam ser comparadas, pelo desenvolvimento de métricas com a quantidade de pessoas total por município fornecido pelo IBGE.

Como também, o tempo é herdado da Base Vacina, visto que o processo de vacinação ocorre diariamente, e foi objetivado ver o desempenho e distribuição do

evento pelas granularidades temporais, não sendo capaz de realizar pela Base IBGE, por sua característica estática dos dados.

Sendo assim, o modelo final da tabela Fato Vacinação Agregada consiste nessa estrutura:

| Campo da Tabela Fato | Dimensão Relacionada | Base de Origem |
|-----------------------------|-----------------------------|-------------------------|
| SK_SEQ_ETD | Dimensão Estado | Base Vacina e Base IBGE |
| SK_SEQ_MNP | Dimensão Município | Base Vacina e Base IBGE |
| TIP_SEXO | Dimensão Sexo Biológico | Base Vacina e Base IBGE |
| QTD_PSS | Métrica da Fato Agregada | Base IBGE |
| QTD_PSS_VCN | Métrica da Fato Agregada | Base Vacina |
| DATA_VCN | Dimensão Data | Base Vacina |

Tabela 5 - Relação de Origens da tabela Fato Agregada

Ao fim de todo esse tratamento, os dados foram transportados para o *Microsoft Power BI*, onde foram elaborados o total de 6 painéis, sendo que destes, 2 estão associados ao *drill down*, tornando secundários aos principais.

Houve também a sumarização de alguns campos durante o desenvolvimento dos *dashboards*, a nível de linguagem *DAX*, desta forma foi possível melhorar a apresentação gráfica dos dados, e por consequência um melhor entendimento da informação.

4. Resultados e Discussões

Para a distribuição dos dados para a construção dos *dashboards*, inicialmente foram idealizados 3 segmentos, a Visão Vacina a qual agrega todo conteúdo sobre a vacina (Fabricante, Lote, Sistema de Origem, Quantidade de Vacinas por Município e Estado, Dose) e o primeiro *drill* visualizando a distribuição das vacinas por Lote, Município e a Quantidade por Estado.

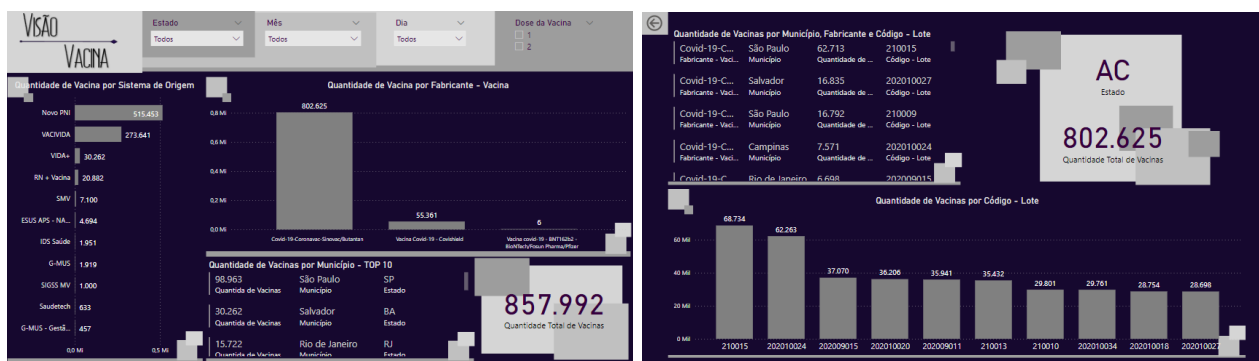


Figura 1 - Painel Visão Vacina e Drill Down complementando visão

O segundo segmento é o ato de vacinar, a Visão Vacinação, que tem como objetivo ofertar informações sobre o andamento da vacinação pelos estados, como também pelos estabelecimentos, em associação, o segundo *drill* revela o processo de vacinação pelas características da população (Etnia, Idade, Subgrupo, Grupo).

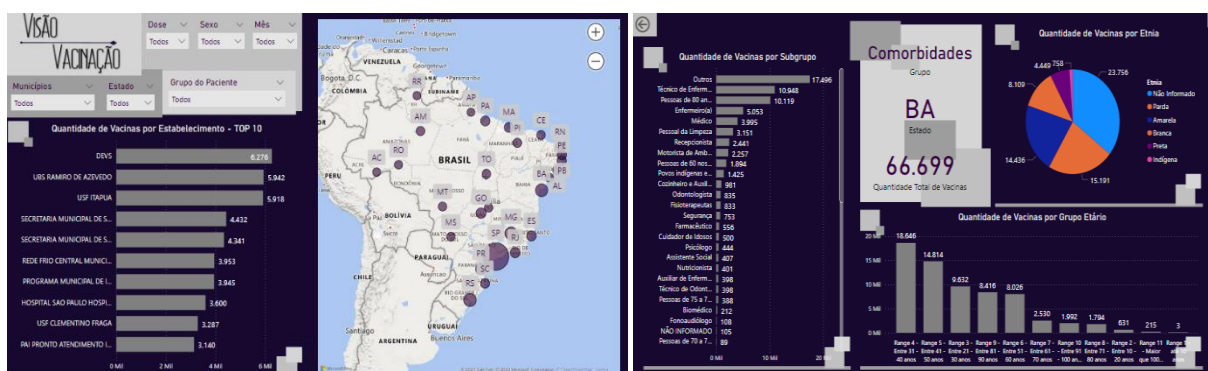


Figura 2 - Painel Visão Vacinação e Drill Down complementando visão

Por fim, a última visão ofertada, é oriunda da agregação dos modelos, a Visão População, onde foi possível propor uma visão gráfica comparativa, da quantidade da população vacinada pelo total demográfico por estado e município, como também a eficiência do processo de vacinação, sugerido pela cobertura vacinal, podendo ver pela totalidade, ou por sexo biológico e a distribuição mensal.

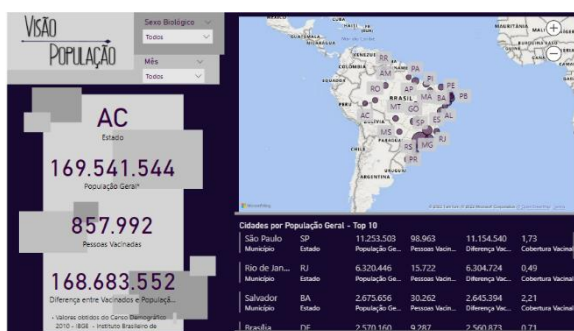


Figura 3 - Painel Visão População

Sendo assim, torna-se perceptível a importância e a forte contribuição da Base IBGE no enriquecimento dos dados, como também na ampliação das possibilidades de informação gerada, devido a agregação dos modelos dimensionais.

Apesar de ser uma base rica e com sua complexidade, as informações do processo de vacinação ainda apresentavam certas limitações, que foram solucionadas ou amenizadas, depois da aplicação dos processos anteriormente descritos.

Além de que, a presença e utilização dos dados oriundos do IBGE, permitiram realizar comparações, criar métricas e gerar análises com os dados oriundos apenas pelo processo de agregação, não sendo possível quando as bases estavam dissociadas.

Associado a este cenário, a escolha do modelo estrela, facilitou o desenvolvimento e entendimento das relações entre os atores e o evento analisado, como também, permitiu uma agregação mais suave e fluída das tabelas fatos, devido a simplicidade dos objetos e das relações de dependências criadas, e também da existência de objetos dimensionais similares nas duas bases.

5. Conclusão

Conhecer a história do *Business Intelligence*, e a sua contribuição no desenvolvimento e progresso do mundo empresarial, torna o profissional mais consciente da sua atuação no ambiente de trabalho, e da importância dos resultados gerados pelos dados.

Ter conhecimento e saber aplicar processos técnicos, desde os mais consolidados e legados, até ideias mais recentes e inovadoras, moldam o profissional em sua completude, deixando mais preparado para os desafios do cotidiano.

Durante toda a atuação deste projeto, optou-se em respeitar as etapas e escolhas realizadas, sempre se embasando em conhecimento científico, com o objetivo de extrair e obter a melhor informação possível a partir dos dados, definir quantas e quais as características das dimensões, as métricas das tabelas fatos, quantas fatos implementar, e por fim o design final dos *dashboard*, é resultante do domínio, escolha e execução das melhores metodologias, no momento adequado em cada etapa do trabalho.

Como também, uma visão ampla do negócio, e uma busca constante de ofertar o melhor para o cliente, sabendo da importância e a necessidade daquela informação, molda um perfil de profissional proativo, que procura melhorar e enriquecer os dados trabalhados, usando metodologias e técnicas adequadas.

Portanto, todo esse conjunto de características, contribuem na formação de profissionais de *business intelligence* mais completos para o mercado de trabalho, atentos e capazes de aplicar melhorias na rotina, as quais associados ao conhecimento histórico da área, fundamentando decisões, e que por consequência reflete em projetos ricos tecnicamente e que satisfazem as necessidades da corporação.

Referências Bibliográficas

BOTOS, Horia Mircea. Business Intelligence and Competitive Intelligence: The Evolution of the Term. **Res. & Sci. Today**, v. 16, p. 56, 2018.

BRUZAROSCO, Donizete Carlos; CASTOLDI, André Vinicius; DOS SANTOS PACHECO, Roberto Carlos. Criando data warehouse com o modelo dimensional. **Acta Scientiarum. Technology**, v. 22, p. 1389-1397, 2000.

FREIRE, Ana Lucy Oliveira. O desenvolvimento do comércio e a produção do espaço urbano. **GeoTextos**, 2010.

GAARDBOE, Rikke; JONASEN, Tanja Svarre. Business intelligence success factors: a literature review. **Journal of Information Technology Management**, v. 29, n. 1, p. 1-15, 2018.

População residente, total e respectiva distribuição percentual, por situação do domicílio e sexo, e razão de sexo, segundo os municípios e as classes de tamanho da população dos municípios - Brasil – 2010. **IBGE**, Disponível em: <https://www.ibge.gov.br/estatisticas/sociais/saude/9662-censo-demografico-2010.html?=&t=destaques>. Acessado em: 30/09/2022

LUHN, Hans Peter. A business intelligence system. **IBM Journal of research and development**, v. 2, n. 4, p. 314-319, 1958.

PEREIRA, Agostinho Oli Koppe; LUNARDI, Tarciso. A Evolução da Indústria e do Comércio e a Responsabilidade Sócio-Ambiental no Ato de Consumo. **Diritto & Diritti**. Disponível em: <https://www.diritto.it/a-evolucao-da-industria-e-do-comercio-e-a-responsabilidade-socio-ambiental-no-ato-de-consumo/>. Acessado em: 25/08/2022.

PITON, Rafael. **Data Warehouse Passo a Passo**, O guia prático de como construir um data warehouse do zero, Porto Alegre, Raizzer, 2018.

QUEIROZ, Rosa Maria Diekn; DE BARROS CARVALHO, Lucas Filipe. Business Intelligence Uma visão teórica sobre BI e suas estratégias. **TECNOLOGIAS EM PROJEÇÃO**, v. 11, n. 2, p. 44-58, 2020.

ROSSETTI, Adroaldo et al. A organização baseada no conhecimento: novas estruturas, estratégias e redes de relacionamento. **Ciência da Informação**, v. 37, p. 61-72, 2008.

SILVA, Marcio; SARTORI, Marco. DATA WAREHOUSE A VANTAGEM DA MODELAGEM DIMENSIONAL DE DADOS.

VACINAÇÃO BRASIL microdados. Disponível em:
<https://www.kaggle.com/datasets/guilhermegoncalves/vacinao-brasil-microdados>.
Acessado em: 30/09/2022.

YESSAD, Lamia; LABIOD, Aissa. Comparative study of data warehouses modeling approaches: Inmon, Kimball and Data Vault. In: **2016 International Conference on System Reliability and Science (ICSRS)**. IEEE, 2016. p. 95-99.