UNIVERSIDADE FEDERAL DE PERNAMBUCO

CENTRO DE INFORMÁTICA

SISTEMAS DE INFORMAÇÃO

LUIS GABRIEL LIMA LOUREIRO XAVIER

Proposta de implantação de Data Warehouse para centralização dos dados e melhorias de indicadores de performance (KPI) em uma empresa de varejo

Recife

2023

UNIVERSIDADE FEDERAL DE PERNAMBUCO

CENTRO DE INFORMÁTICA

SISTEMAS DE INFORMAÇÃO

LUIS GABRIEL LIMA LOUREIRO XAVIER

Proposta de implantação de Data Warehouse para centralização dos dados e melhorias de indicadores de performance (KPI) em uma empresa de varejo

Trabalho de Conclusão de Curso apresentado no Curso de Bacharelado em Sistemas de Informação do Centro de Informática da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do grau de Bacharel em Sistemas de Informação.

Orientador(a): Jamilson Ramalho Dantas

Recife

2023

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Xavier, Luis Gabriel Lima Loureiro.
Proposta de implantação de Data Warehouse para centralização dos dados e melhorias de indicadores de performance (KPI) em uma empresa de varejo / Luis Gabriel Lima Loureiro Xavier. - Recife, 2023.
66 p : il., tab.

Orientador(a): Jamilson Ramalho Dantas
Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Pernambuco, Centro de Informática, Sistemas de Informação - Bacharelado, 2023.

1. Indicadores de Desempenho. 2. KPIs. 3. Eficiência Operacional. 4. Data Warehouse. I. Dantas, Jamilson Ramalho. (Orientação). II. Título.

000 CDD (22.ed.)

Este trabalho é dedicado à minha mãe,

minha família e todos os meus amigos,

que sempre estiveram comigo, me apoiaram e me

incentivaram a chegar até aqui

AGRADECIMENTOS

Quero deixar meus agradecimentos a todos os colegas de turma, que acompanharam minha trajetória. Em especial, Vinicius Luiz, Alexsandro Henrique, Gustavo Prazeres e Daniel Moraes, pessoas essas que sempre contribuíram para meu sucesso acadêmico.

“Você não pode gerenciar o que não pode medir." - Peter Drucker

RESUMO

Em um ambiente empresarial em constante transformação, a tomada de decisões informadas é crucial para o sucesso organizacional. Nesse contexto, os Indicadores-Chave de Desempenho (KPIs) desempenham um papel vital ao fornecer informações fundamentais sobre o desempenho operacional e estratégico. Este estudo analisa a migração de um modelo descentralizado de geração de KPIs para um Data Warehouse, em resposta aos desafios enfrentados pelo primeiro modelo. Destacamos a importância dos KPIs e sua conexão com as Métricas Empresariais, bem como o impacto desse processo na busca pela eficiência operacional. Identificamos desafios no modelo descentralizado, como falta de controle, erros frequentes e ineficiências na correção de falhas. Exploramos a mudança para um modelo centralizado em um Data Warehouse, com o objetivo de aprimorar a qualidade, confiabilidade e eficiência na geração de indicadores. Destacamos a relevância dos KPIs na busca pela eficiência operacional e como a centralização de dados pode solucionar problemas preexistentes. Este trabalho oferece valiosos insights sobre essa transição, enfatizando a importância dos Indicadores de Desempenho e das Métricas Empresariais nas organizações modernas.

Palavras-chave: Indicadores de Desempenho; KPIs; Eficiência Operacional; Data Warehouse.

ABSTRACT

In a constantly evolving business landscape, the ability to make informed decisions is of critical importance for the success of any organization. In this context, Key Performance Indicators (KPIs) play a vital role by providing essential insights into both operational and strategic performance. This study examines the transition from a decentralized model of KPI generation to a Data Warehouse in response to the challenges faced by the former model. We emphasize the significance of KPIs and their connection to Business Metrics, as well as the impact of this process on the pursuit of operational efficiency within organizations. Challenges identified in the decentralized model, such as a lack of control, frequent errors, and inefficiencies in error correction, are addressed. We delve into the shift towards a centralized model within a Data Warehouse, aiming to enhance the quality, reliability, and efficiency of KPI generation. The importance of KPIs in the pursuit of operational efficiency is underscored, along with how data centralization can resolve pre-existing issues. This work provides valuable insights into this transition, emphasizing the significance of Key Performance Indicators and Business Metrics in modern organizations.

Keywords: Key Performance Indicators; KPIs; Operational Efficiency; Data Warehouse.

LISTA DE FIGURAS

LISTA DE CÓDIGOS

LISTA DE QUADROS

LISTA DE GRÁFICOS

LISTA DE ABREVIAÇÕES

KPI	Indicadores-Chave de Desempenho
QDI	Qlik Data Integration
ETL	Extrair, Transformar e Carregar
DW	Data Warehouse
DB Link	Database Link

SUMÁRIO

1 Introdução

Nessa seção são apresentados o contexto do estudo assim como as motivações para o mesmo ser realizado.

1.1 Contexto

No cenário empresarial em constante evolução, a tomada de decisões informadas é fundamental para o sucesso de qualquer organização. Para embasar essas decisões, os Indicadores-Chave de Desempenho (KPIs) desempenham um papel fundamental, fornecendo insights sobre o desempenho operacional e estratégico. Tradicionalmente, a geração de KPIs tem sido uma tarefa descentralizada, com diferentes departamentos ou unidades de negócios gerando e monitorando seus próprios KPIs de maneira isolada. No entanto, a crescente complexidade dos ambientes de negócios e a necessidade de uma visão holística têm levado as organizações a repensar suas abordagens. [8]

Neste contexto, surge a abordagem de geração centralizada de KPIs em um ambiente de Data Warehouse. O Data Warehouse, um repositório centralizado de dados organizados e estruturados, oferece a oportunidade de consolidar informações provenientes de diversas fontes, fornecendo uma visão unificada e confiável do desempenho organizacional. A transição da geração descentralizada para centralizada de KPIs não é apenas uma mudança técnica, mas também uma transformação cultural e organizacional que busca melhorar a qualidade das decisões tomadas em todos os níveis da empresa.

Este trabalho investiga os motivos, desafios e benefícios envolvidos na transição da geração descentralizada de KPIs para uma abordagem centralizada em um Data Warehouse. Serão explorados os conceitos fundamentais relacionados à KPIs, a arquitetura de Data Warehouse, as etapas de integração de dados e transformação, bem como o impacto dessa transição na tomada de decisões organizacionais. Por meio de estudos de caso e análises comparativas, busca-se oferecer insights valiosos para empresas que consideram adotar essa abordagem.

O próximo capítulo apresenta uma revisão abrangente da literatura existente, abordando as teorias e práticas relacionadas à geração de KPIs, integração de dados, arquiteturas de Data Warehouse e os desafios enfrentados durante a transição. Posteriormente, serão discutidas metodologias e abordagens adotadas para a condução deste estudo, seguidas por uma análise detalhada dos resultados obtidos. Por fim, as conclusões derivadas dessa pesquisa serão apresentadas, juntamente com implicações práticas e possíveis direções futuras.

Em última análise, este trabalho busca contribuir para a compreensão mais profunda das vantagens e desafios associados à transição da geração descentralizada de KPIs para uma geração centralizada em um ambiente de Data

Warehouse, proporcionando uma base sólida para a tomada de decisões estratégicas no mundo empresarial moderno.

1.2 Motivação e justificativa

No atual cenário do mercado varejista, a obtenção de informações estratégicas e o monitoramento do desempenho operacional são fundamentais para se manter competitivo e alcançar o sucesso nos negócios. Entretanto, a empresa enfrenta desafios significativos, como o alto tempo de execução dos processos, a falta de controle do processo e a dificuldade na correção de erros. Para superar essas adversidades, é essencial avançar da geração descentralizada de KPIs (Indicadores-Chave de Desempenho) para a consolidação desses dados em um Data Warehouse, proporcionando ganhos efetivos em eficiência, controle e inteligência operacional.

A geração descentralizada de KPIs, caracterizada pela dispersão de dados e informações em diversos sistemas e bancos de dados, leva a um alto tempo de execução dos processos. A concorrência com outros processos e a necessidade de acessar múltiplas fontes de dados tornam as operações lentas e ineficientes, resultando em atrasos na obtenção de informações críticas para a tomada de decisão. Esse cenário afeta diretamente a capacidade da empresa de responder rapidamente às demandas do mercado, reduzindo sua agilidade e flexibilidade. [4]

Além disso, a falta de controle do processo é uma preocupação constante. A execução automática das atividades não permite um acompanhamento detalhado do que está ocorrendo em tempo real. A ausência de logs de controle impede a rastreabilidade das ações executadas, dificultando a identificação de falhas e erros nos procedimentos. A falta de visibilidade dos fluxos de trabalho compromete a capacidade dos gestores de avaliar o desempenho e a qualidade das operações, gerando insegurança na tomada de decisões estratégicas.

Outro desafio enfrentado é a dificuldade na correção de erros. A natureza automatizada dos processos torna a identificação das causas de falhas extremamente custosas e complexas. A falta de um registro centralizado dos eventos dificulta a análise dos problemas, o que resulta em demoras na correção e, consequentemente, em perdas financeiras e insatisfação dos clientes. A ausência de uma abordagem integrada para a geração de KPIs impede que a empresa realize melhorias consistentes e oportunas em seus processos. [8]

Diante desses problemas, a proposta deste Trabalho de Conclusão de Curso (TCC) é a transição da geração descentralizada de KPIs para um Data Warehouse. Esse novo modelo permitirá consolidar os dados provenientes de diversas fontes em um único repositório, promovendo uma análise unificada e coerente das informações. A implementação de um Data Warehouse proporcionará um ganho significativo no tempo de execução, uma vez que as consultas e análises serão realizadas em uma estrutura otimizada e centralizada.

Além disso, o Data Warehouse possibilita um controle detalhado dos processos, fornecendo logs de controle atualizados e acessíveis em tempo real. Isso

permitirá aos gestores acompanhar as atividades de forma mais precisa, identificar desvios e tomar decisões fundamentadas em informações confiáveis. A abordagem centralizada também simplificará a correção de erros, permitindo uma análise mais rápida e eficiente das causas de falhas, agilizando as ações corretivas. [2]

Portanto, este trabalho representa uma oportunidade significativa para explorar a importância da transição da geração descentralizada de KPIs para um Data Warehouse em uma empresa de varejo. Espera-se que essa mudança traga melhorias substanciais em termos de eficiência operacional, tomada de decisão e capacidade de resposta ao mercado, consolidando a empresa como uma referência no setor, capaz de se destacar em um ambiente de negócios cada vez mais competitivo.

1.3 Problema de Pesquisa

Diante desse contexto, sabendo que a geração de KPIs de traz resultados impressionantes, quando bem aplicada, esse trabalho pretende responder às seguintes perguntas de pesquisa:

1. Quais são os principais desafios enfrentados, pela empresa de varejo, ao tentar implantar e administrar a geração de KPIs de forma centralizada na organização?
2. É possível, por meio da consolidação dos dados em um Data Warehouse, alcançar melhorias na execução dos processos e obter ganhos de produtividade, redução de tempo e recursos dedicados ao processo de geração de KPIs?

1.4 Objetivos da pesquisa

Para responder essas perguntas, este estudo tem como objetivo principal analisar os resultados dessa transição para um Data Warehouse de uma empresa de varejo. Para conseguir atingir esse objetivo geral, os seguintes objetivos específicos foram estabelecidos:

1. Levantar métricas que permitam avaliar o impacto da transição do cálculo dos KPIs para um Data Warehouse.
2. Analisar o processo dessa implantação e as dificuldades encontradas.
3. Comparar os resultados obtidos, através das métricas, antes e depois dessa transição.

1.5 Trabalho Relacionados

Nesta seção, são discutidos os trabalhos anteriores relacionados à temática abordada neste estudo.

A survey on exploring key performance indicators

Este presente trabalho oferece uma visão abrangente das diferentes abordagens utilizadas na exploração e previsão de indicadores-chave de desempenho (KPIs), desempenhando um papel essencial na orientação das organizações em direção à realização de seus objetivos estratégicos. A pesquisa abordada neste texto inclui referências cruciais relacionadas à definição precisa de KPIs, a seleção criteriosa dos indicadores mais apropriados, a avaliação do desempenho com base em KPIs e a utilização de métodos automatizados para a interpretação de dados relacionados a esses indicadores. Além disso, são destacadas as limitações e desafios associados ao emprego de KPIs, enfatizando a necessidade crítica de conceber KPIs bem definidos como suporte fundamental à tomada de decisões nas organizações. [1]

Indicadores de desempenho: Estudo de caso na empresa net serviços

O presente trabalho constitui uma análise abrangente dos indicadores de desempenho em uma empresa de serviços, sendo uma parte fundamental do curso de graduação em Administração da Universidade Federal da Paraíba. Neste estudo de caso, são minuciosamente apresentados os indicadores-chave que foram objeto de análise, bem como a aplicação prática dos resultados pela empresa NET Serviços. Destaca-se, de maneira significativa, o papel crucial desempenhado pelo orientador do trabalho, que desempenhou um papel fundamental na concepção dos indicadores e na análise dos resultados obtidos, contribuindo assim para uma compreensão aprofundada e valiosa do desempenho da empresa e suas implicações para a gestão estratégica. [10]

Utilização de KPI – Indicadores de desempenho na cadeia de suprimentos

Este trabalho de conclusão de curso abrange um estudo de caso valioso que se concentra na aplicação de Indicadores-Chave de Desempenho (KPIs) na gestão da cadeia de suprimentos de uma indústria metalúrgica no setor da construção civil. A pesquisa tem como principal objetivo a descrição detalhada de como os indicadores de desempenho podem ser estrategicamente empregados para aprimorar as avaliações internas da organização, promovendo um monitoramento mais eficaz do desempenho organizacional. A metodologia de pesquisa adotada envolveu uma abordagem de Pesquisa Bibliográfica e uma Pesquisa de natureza descritiva, dividida em três capítulos essenciais: Supply Chain Management, Key Performance Indicator e um esclarecedor Estudo de Caso. A análise do estudo de caso oferece insights valiosos, evidenciando como a implementação adequada de KPIs pode resultar em melhorias substanciais em termos de eficiência e qualidade dos processos na cadeia de suprimentos, proporcionando um valioso recurso para aprimorar a gestão organizacional no contexto da indústria metalúrgica na construção civil. [9]

O uso de KPI’s no processo logístico em um distribuidora de medicamentos

O presente trabalho intitulado "Uso de KPI's no processo logístico de uma distribuidora de medicamentos" representa um estudo de pesquisa bibliográfica e de campo conduzido por Dailani Santos de Jesus e Rondinelle Pereira Laurindo, que focaliza a vital importância dos Indicadores-Chave de Desempenho (KPIs) para o êxito das operações logísticas. Este trabalho destaca a meticulosa seleção de artigos relevantes e publicações recentes no campo, ressaltando o uso de descritores como indicadores, gestão, qualidade e KPI para a fundamentação teórica. Além disso, o documento aborda de forma significativa o papel crucial da análise dos KPIs na identificação de pontos críticos e gargalos nos processos logísticos, destacando o envolvimento fundamental do gestor de logística na definição e acompanhamento desses indicadores. O estudo oferece uma contribuição substancial para o entendimento e a aplicação eficaz dos KPIs no contexto das operações logísticas de distribuição de medicamentos. [5]

Monotorização de KPI

No trabalho intitulado "Monotorização de KPI: Análise de resultados e proposta de novos indicadores" a autora Joana Rebelo Lopes da Silva propõe a implementação de novos indicadores financeiros como um meio eficaz para aprimorar o desempenho da empresa. Destaca-se a ênfase dada à importância de sistemas de controle de gestão estruturados e organizados como ferramentas essenciais para fornecer informações cruciais aos gestores de topo e, consequentemente, estimular o crescimento corporativo. Além disso, o documento salienta o papel fundamental dos sistemas de Planejamento de Recursos Empresariais (ERP) na gestão financeira, ao possibilitar um acesso rápido e fácil a informações valiosas que auxiliam no processo de tomada de decisão. O trabalho também ressalta a relevância da análise de resultados para embasar as decisões estratégicas e enfatiza as vantagens significativas da utilização de Indicadores-Chave de Desempenho (KPIs) no contexto da gestão financeira de uma empresa. Ele tem como objetivo apresentar uma proposta de novos indicadores financeiros para a empresa, com o objetivo de melhorar o seu desempenho. O trabalho também tem como objetivo analisar o sistema de controle de gestão da empresa, com ênfase no orçamento, e avaliar se o método utilizado para a sua elaboração é o mais adequado e com que objetivos é utilizado o orçamento. Além disso, o artigo busca destacar a importância da análise de resultados para a tomada de decisões estratégicas e as vantagens de utilizar KPIs na gestão financeira de uma empresa. [11]

2 Referencial Teórico

Nesse tópico são apresentados os referenciais teóricos importantes para o melhor entendimento do estudo.

2.1 Padronização de medição de KPIs

A geração descentralizada de KPIs enfrenta uma série de obstáculos que podem comprometer sua utilidade e confiabilidade. Uma das principais preocupações é a possível falta de padronização. Em ambientes descentralizados, diferentes unidades ou departamentos podem adotar abordagens distintas para medir e reportar indicadores, dificultando a comparação e a avaliação global do desempenho da organização.

Além disso, a inconsistência nos dados coletados pode surgir como um problema central. A coleta de informações em ambientes descentralizados pode resultar em dados de qualidade variável, tornando difícil garantir a precisão e a confiabilidade dos indicadores gerados. A ausência de um método uniforme de coleta e validação de dados pode minar a capacidade de confiar plenamente nos resultados obtidos.

A complexidade da agregação de indicadores também é um desafio. Quando diversas unidades operacionais estão envolvidas na geração de indicadores , consolidar esses dados em uma visão abrangente pode ser complexo e demorado. Diferentes métricas e escalas de medição podem dificultar a criação de uma representação holística do desempenho da organização.

Outra preocupação envolve a possibilidade de desalinhamento de objetivos. Unidades descentralizadas podem ter prioridades diferentes que não necessariamente coincidem com os objetivos globais da organização. Isso pode levar à geração de indicadores que não capturam adequadamente o progresso em direção aos objetivos estratégicos.

A mitigação desses problemas requer a implementação de processos de padronização de medição, garantindo que todos os departamentos ou unidades adotem metodologias semelhantes na geração de indicadores. O estabelecimento de sistemas robustos de coleta e validação de dados também é essencial para evitar inconsistências. Além disso, é crucial manter uma comunicação eficaz entre as unidades descentralizadas para garantir o alinhamento de objetivos e a compreensão compartilhada dos indicadores e seus significados. [3]

2.2 Indicadores-Chave de Desempenho (KPIs)

Um KPI (Indicador-Chave de Desempenho) é tipicamente concebido como uma medida que se expressa predominantemente em termos de taxas, proporções, médias ou porcentagens. Ao contrário, ele não é baseado em valores numéricos brutos. Embora números brutos desempenhem um papel importante ao fornecer

informações nos relatórios de análise da web, sua eficácia é limitada pela ausência de contexto. Em contraste, os indicadores-chave de desempenho oferecem um grau mais elevado de relevância e significado, permitindo uma compreensão mais profunda das realizações e das áreas que necessitam de atenção.

Os KPIs (Indicador-Chave de Desempenho) são números projetados para transmitir informações de forma sucinta e são usados para medir o desempenho organizacional. Eles são uma ferramenta importante para simplificar a relação das pessoas com os dados da web e orientar a ação. Os bons indicadores são bem definidos, bem apresentados, criam expectativas e impulsionam ações. Eles usam taxas, razões, percentagens e médias em vez de números brutos, fornecem contexto temporal e destacam mudanças em vez de apresentar tabelas de dados. O objetivo final dos indicadores é impulsionar ações críticas para o negócio. [8]

2.3 Seleção adequada de KPIs

A seleção adequada de KPIs (Key Performance Indicators) é fundamental para orientar a tomada de decisões. Os KPIs devem ser escolhidos com base nos objetivos de negócios e nas necessidades dos usuários. Eles devem ser relevantes, mensuráveis, específicos, alcançáveis e relevantes no tempo. A escolha de indicadores inadequados pode levar a decisões equivocadas e ações ineficazes. Por outro lado, a escolha de indicadores relevantes e bem definidos pode ajudar a identificar problemas, oportunidades e tendências, permitindo que as empresas tomem decisões informadas e orientadas por dados. [8]

2.4 Geração Descentralizada de KPIs

A geração descentralizada de KPIs enfrenta uma série de obstáculos que podem comprometer sua utilidade e confiabilidade. Uma das principais preocupações é a possível falta de padronização. Em cenários descentralizados, é comum que unidades ou departamentos adotem métodos variados para medir e comunicar seus indicadores, tornando desafiadora a tarefa de comparar e avaliar o desempenho global da organização como um todo. Além disso, a inconsistência nos dados coletados pode surgir como um problema central. A coleta de informações em ambientes descentralizados pode resultar em dados de qualidade variável, tornando difícil garantir a precisão e a confiabilidade dos KPIs gerados. A ausência de um método uniforme de coleta e validação de dados pode minar a capacidade de confiar plenamente nos resultados obtidos. A complexidade da agregação de indicadores também é um desafio. Quando diversas unidades operacionais estão envolvidas na geração de KPIs, consolidar esses dados em uma visão abrangente pode ser complexo e demorado. Diferentes métricas e escalas de medição podem dificultar a criação de uma representação holística do desempenho da organização. A mitigação desses problemas requer a implementação de processos de padronização de medição, garantindo que todos os departamentos ou unidades adotem

metodologias semelhantes na geração de indicadores . O estabelecimento de sistemas robustos de coleta e validação de dados também é essencial para evitar inconsistências. Além disso, é crucial manter uma comunicação eficaz entre as unidades descentralizadas para garantir o alinhamento de objetivos e a compreensão compartilhada dos KPIs e seus significados. [4]

2.5 Inconsistência de dados

A inconsistência de dados representa um desafio significativo capaz de prejudicar a eficácia das análises e das decisões. Essa discrepância pode gerar resultados distorcidos e interpretações equivocadas da realidade, minando a confiabilidade dos insights extraídos. O impacto vai além da precisão estatística, ele afeta a confiança geral nas informações, uma vez que os usuários podem questionar a integridade dos dados fornecidos. A correção desse problema é uma prioridade essencial para organizações que buscam basear suas escolhas em informações sólidas. Uma abordagem resiliente envolve a implementação de protocolos rigorosos de verificação de dados, a adoção de padrões uniformes de coleta e armazenamento, bem como a implantação de sistemas de detecção precoce para identificar e resolver inconsistências antes que elas contaminem as análises. Investir nessa correção não é apenas uma medida técnica, mas uma estratégia de fortalecimento da confiabilidade dos dados, permitindo uma tomada de decisão mais informada e precisa. [6]

2.6 Data Warehouse: Conceitos e Benefícios

Um data warehouse, de acordo com a definição clássica, é uma infraestrutura de armazenamento de dados que desempenha um papel fundamental na gestão da informação em organizações. Ele é concebido como uma coleção meticulosamente organizada de dados, caracterizada por sua integração, orientação a assuntos, variabilidade temporal e não volatilidade. A principal finalidade de um data warehouse é fornecer um repositório centralizado e confiável de dados para facilitar a análise de informações e apoiar a tomada de decisões estratégicas e táticas em uma organização.

O data warehouse age como um hub estratégico para a gestão de dados, consolidando informações de múltiplas fontes operacionais. Os dados são extraídos dessas fontes, passam por um processo de transformação para garantir sua qualidade e relevância e, finalmente, são carregados no data warehouse. Uma das características essenciais é a organização desses dados em um modelo multidimensional, o que permite que sejam analisados de várias perspectivas, proporcionando uma visão abrangente e flexível do panorama organizacional.

Esta abordagem multidimensional torna mais fácil para que os usuários acessem rapidamente os dados necessários para responder a perguntas críticas relacionadas aos negócios. Além disso, o data warehouse é projetado com foco na

capacidade de suportar consultas complexas e análises ad hoc, permitindo que os usuários explorem os dados de maneiras profundas e diversificadas, que seriam impraticáveis em um ambiente de banco de dados operacional. [2]

2.7 Database Links (DBLinks)

O database link é uma funcionalidade que permite a criação de conexões lógicas entre bancos de dados, viabilizando consultas SQL e operações de manipulação de dados em bancos de dados remotos como se fossem locais. Sua configuração envolve a especificação de parâmetros como o nome do banco de dados remoto, credenciais de acesso e outros detalhes pertinentes. Uma vez configurado, os usuários podem executar consultas e operações diretamente no banco de dados local, com o sistema de gerenciamento de banco de dados cuidando da comunicação e do acesso aos dados do banco de dados remoto.

A principal vantagem reside na capacidade de acessar dados distribuídos em diferentes locais geográficos ou servidores sem a necessidade de duplicação física dos dados. Os database links facilitam a consolidação de informações de várias fontes, possibilitando análises e relatórios abrangentes que envolvem dados de diferentes origens. Além disso, a utilização de database links pode reduzir a necessidade de armazenamento duplicado e economizar recursos em termos de espaço em disco.

No entanto, a segurança é uma preocupação crítica ao utilizar database links. É essencial implementar medidas rigorosas de autenticação e autorização para garantir que apenas usuários autorizados acessem os dados remotos. Consultas envolvendo database links podem ser mais lentas devido à latência da rede e à comunicação entre servidores, o que requer monitoramento e otimização cuidadosos. A administração de vários database links pode ser complexa em ambientes com muitas conexões, exigindo um planejamento eficaz e uma estratégia de gerenciamento. Regulamentos de privacidade de dados podem impor restrições ao uso de database links, especialmente quando se trata de dados sensíveis ou regulamentados. [7]

3 Metodologia

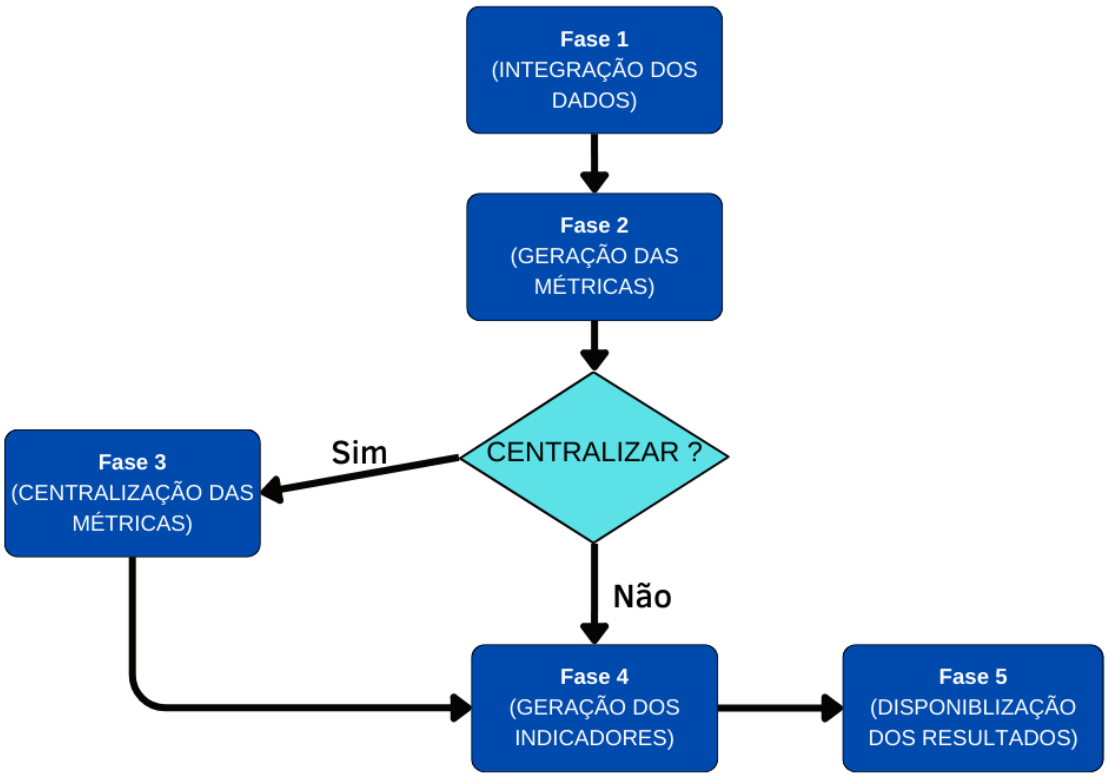
Este presente trabalho é um estudo de caso realizado num setor de TI de uma empresa de varejo, setor responsável pela criação e manutenção dos indicadores estudados, além da geração de seus resultados e disponibilização para os setores tomadores de decisão. Por lidar com diversos dados sensíveis da organização, o estudo optou por alterar os nomes de todas as informações referentes a tabelas, owners, procedures, views e materialized views, garantindo a segurança desses dados.

Visando garantir um rigoroso controle, os resultados dos indicadores são calculados de forma mensal. Isso proporciona uma análise minuciosa e detalhada do desempenho ao longo do tempo, permitindo a identificação de padrões sazonais ou tendências de crescimento de maneira mais precisa. Além disso, o cálculo mensal estabelece uma base comparativa consistente, possibilitando avaliar o progresso em relação aos objetivos estabelecidos e embasar decisões de maneira altamente informada.

Os indicadores são calculados a partir de métricas, que nada mais são do que cálculos feitos a partir de dados que foram inicialmente tratados. Porém esses resultados são apenas medidas que não fornecem informações suficientemente estruturadas para auxiliar em decisões ou planos de ação, como por exemplo o somatório das vendas realizadas em uma determinada área no mês. A partir delas são feitos outros cálculos para serem obtidos os resultados dos indicadores.

Esta seção descreve, passo a passo, as etapas para a geração de um KPI nessa empresa estudada, que parte desde dados brutos até chegar às métricas e indicadores finais. O processo é dividido em seis etapas, como descrito na Figura 1, são elas: Integração dos dados, geração das métricas, centralização das métricas, geração dos indicadores, a disponibilização desses resultados e a correção dos erros.

Figura 1. Fluxograma do processo completo da geração dos indicadores.



Fonte: Elaborado pelo autor (2023)

3.1 Integração dos Dados

Nessa etapa são identificadas as fontes de dados relevantes para a geração dos indicadores, onde são consolidados os dados brutos de diferentes fontes, como sistemas internos, plataformas de vendas, registros de treinamentos, entre outros. Os dados coletados passarão por processos de limpeza, transformação e enriquecimento, com o objetivo de garantir a integridade e a qualidade das informações. Para essa coleta de dados a empresa utiliza a ferramenta Qlik Data Integration (QDI) para integrar e replicar grandes quantidades de dados, garantindo a qualidade das análises e operações tanto na arquitetura relacional quanto no ambiente de Data Warehouse. Existem três tipos principais de dados para o processo:

- Dados descritivos: dados das características tanto das métricas quanto dos indicadores, como por exemplo quais são as métricas que o indicador vai utilizar como base para o cálculo, quais funcionários recebem quais indicadores, entre outros.
- Dados dos colaboradores: os quais serão associados aos resultados, como matrícula, nome, filial, etc.

- Dados operacionais: são os dados de valor propriamente dito utilizados para os cálculos, como a quantidade de horas trabalhadas por um funcionário, por exemplo.

Além da integração dos dados, ocorrem processos de transformação dos dados para que o cálculo seja facilitado, isso ocorre de forma procedural nos bancos de dados onde vão ocorrer os cálculos das métricas. Um exemplo de transformação é a associação das vendas ao vendedor que realizou a venda e o supervisor responsável por esse vendedor, facilitando o cálculo das métricas que utilizam os dados de vendas.

3.2 Geração das Métricas

O processo continua com a geração dos resultados das métricas, que são calculados através de um fluxo de extração, transformação e carregamento (ETL) a partir dos dados operacionais base, coletados pela etapa anterior. As métricas possuem três formas de cálculo:

- Por matrícula: onde cada funcionário só recebe o valor específico de sua matrícula, por exemplo a métrica de horas trabalhadas, a qual o valor das horas trabalhadas por um funcionário são associadas apenas para ele.
- Por filial: onde o resultado da métrica é igual para todos os funcionários da mesma filial, que estão associados a essa métrica. Por exemplo, a métrica de vendas por filial, onde só pode haver um único valor por filial.
- Geral: onde o resultado é a soma dos resultados de todas as filiais, esse tipo de métrica é gerado apenas pelo processo de centralização das métricas, que será melhor detalhado abaixo.

Vamos exemplificar esse procedimento usando o indicador de participação em treinamento, o qual é composto por duas métricas: participação em treinamento - meta e participação em treinamento - real. Os dados base dessas métricas vem um dos processos de transformação de dados citados no tópico anterior. Esse processo associa os dados de treinamento e absenteísmo dos funcionários a suas respectivas matrículas. A partir desses dados a métrica de participação em treinamento - real calcula quantos treinamentos o funcionário realizou no mês sendo analisado, já a participação em treinamento - meta calcula quantos treinamentos o funcionário deveria ter realizado.

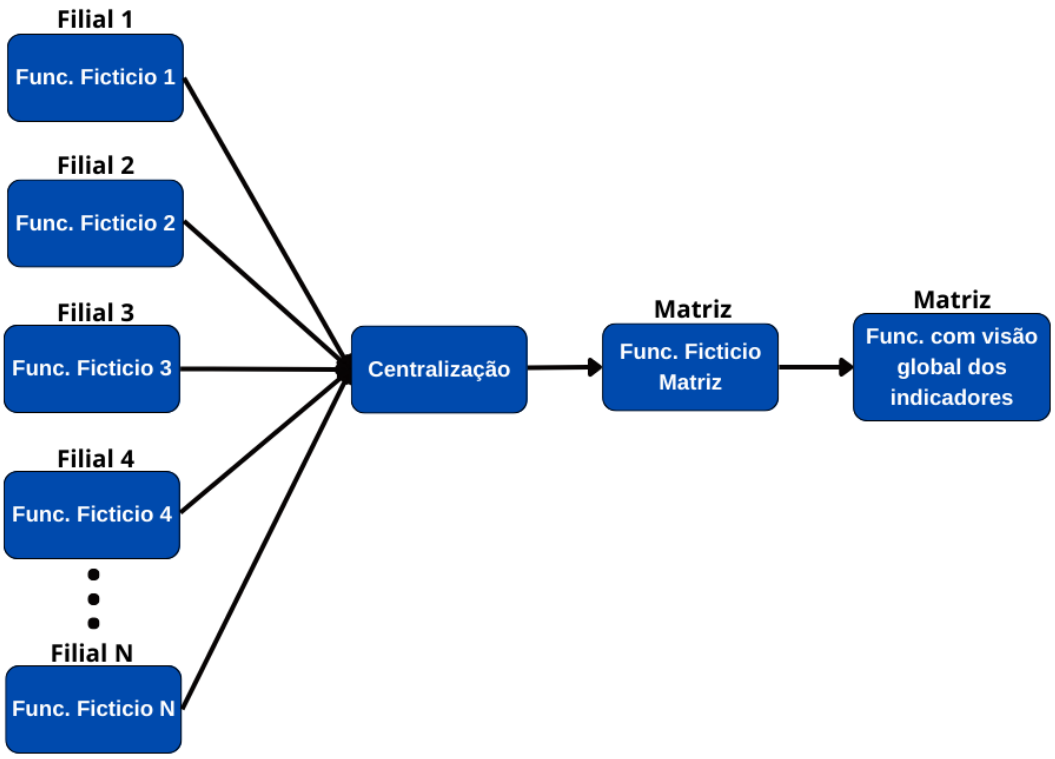
3.3 Centralização das métricas

O processo de centralização ocorre devido a existência de hierarquias de cargos na empresa, por isso é necessário que alguns funcionários da matriz ou de cargos de supervisão recebam resultados advindos de outras filiais ou de colaboradores subordinados a eles, para isso se faz necessário a centralização dos resultados.

A centralização nada mais é do que a associação dos resultados das métricas para as matrículas da matriz que precisam visualizar os resultados de forma global, somando os resultados individuais de todas as filiais, esse cálculo é dito como o tipo de cálculo “geral” das métricas, citado anteriormente.

Para a realização desses cálculos, foram criados funcionários fictícios, para cada filial e para a matriz, para atuar como filial, assim eles recebem o valor total de todas as métricas para cada filial. A partir dos valores das métricas dos funcionários fictícios de cada filial, são somados esses valores para que o funcionário fictício matriz receba a soma de todas as filiais para cada métrica. A partir dos valores do funcionário fictício matriz, os funcionários dos cargos que necessitam de uma visualização global dos indicadores, recebem os mesmos valores que o funcionário fictício matriz, Figura 2.

Figura 2. Fluxograma do processo de centralização das métricas.



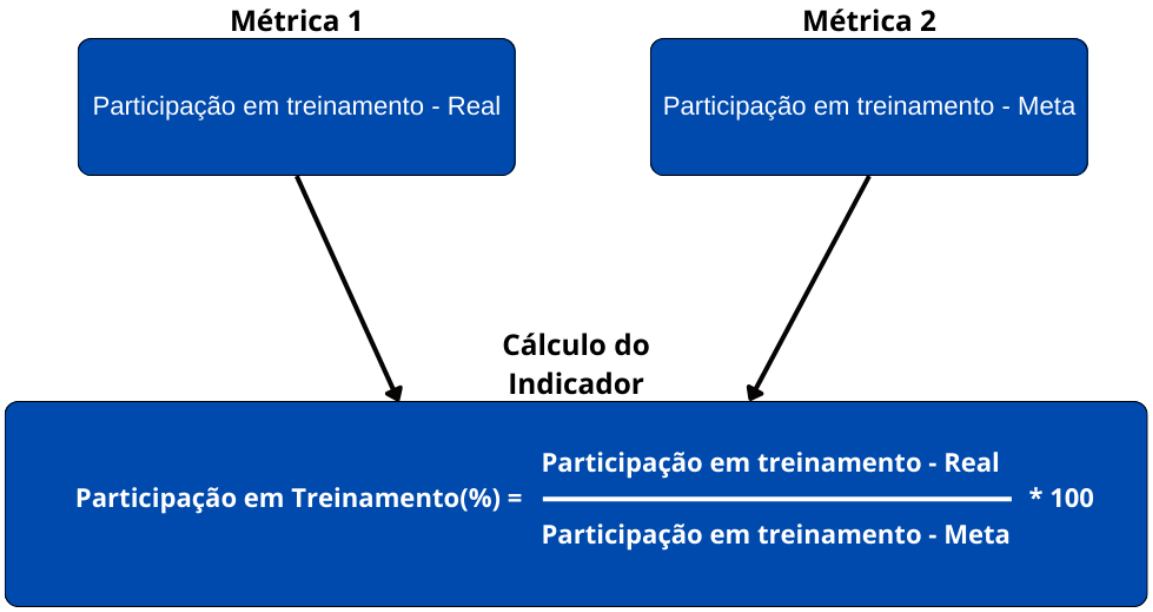
Fonte: Elaborado pelo autor (2023)

3.4 Geração dos Indicadores

Após a conclusão dos cálculos das métricas, dá-se início ao crucial processo de geração dos indicadores. Este estágio envolve a transformação dos resultados das métricas em informações estruturadas e acionáveis, que são fundamentais para orientar as decisões estratégicas e a formulação de planos de ação. Um exemplo

prático seria o uso de um KPI de participação em treinamento que mede a participação dos funcionários em treinamentos oferecidos pela empresa, visando avaliar seu envolvimento nos programas de capacitação. Para calcular os resultados do indicador para um funcionário em específico, a seguinte fórmula poderia ser utilizada:

Figura 3. Cálculo do Indicador Participação em Treinamento (%).



Fonte: Elaborado pelo autor (2023)

Nessa formulação, tanto a quantidade de cursos realizados quanto a quantidade de cursos requeridos são métricas extraídas dos dados, sendo as métricas chamadas “Participação em treinamento - Real” e “Participação em treinamento - Meta”, respectivamente (Figura 3).

O exemplo demonstra como as métricas se convertem em um indicador concreto. Essa transformação permite avaliar numericamente o nível de participação de um funcionário nos treinamentos. Importante notar que o indicador, como um KPI, possui uma meta estabelecida. Essa meta serve como um ponto de referência para avaliar e validar o desempenho do funcionário em relação aos objetivos de capacitação. A abordagem descrita reflete o processo essencial de conectar as métricas calculadas com os indicadores relevantes, traduzindo dados em informações valiosas e prontas para guiar a tomada de decisões e ações assertivas.

3.5 Disponibilização dos resultados

Após a geração dos indicadores, esses valores precisam ser disponibilizados para os tomadores de decisão em forma de um painel.A disponibilização dos resultados dos indicadores em um painel oferece uma ferramenta poderosa para monitorar o desempenho da empresa, tomar decisões mais informadas e manter todos os envolvidos atualizados sobre o progresso em relação aos objetivos estratégicos.

Essa forma de disponibilização desempenha um papel fundamental no processo de análise e tomada de decisão em uma empresa, pois ela oferece visibilidade em tempo real sobre o desempenho da empresa. Os indicadores são apresentados de forma clara e compreensível, permitindo que os líderes empresariais avaliem rapidamente como a organização está se saindo em relação aos objetivos estratégicos. Isso significa que problemas ou oportunidades podem ser identificados imediatamente, sem a necessidade de esperar por relatórios periódicos.

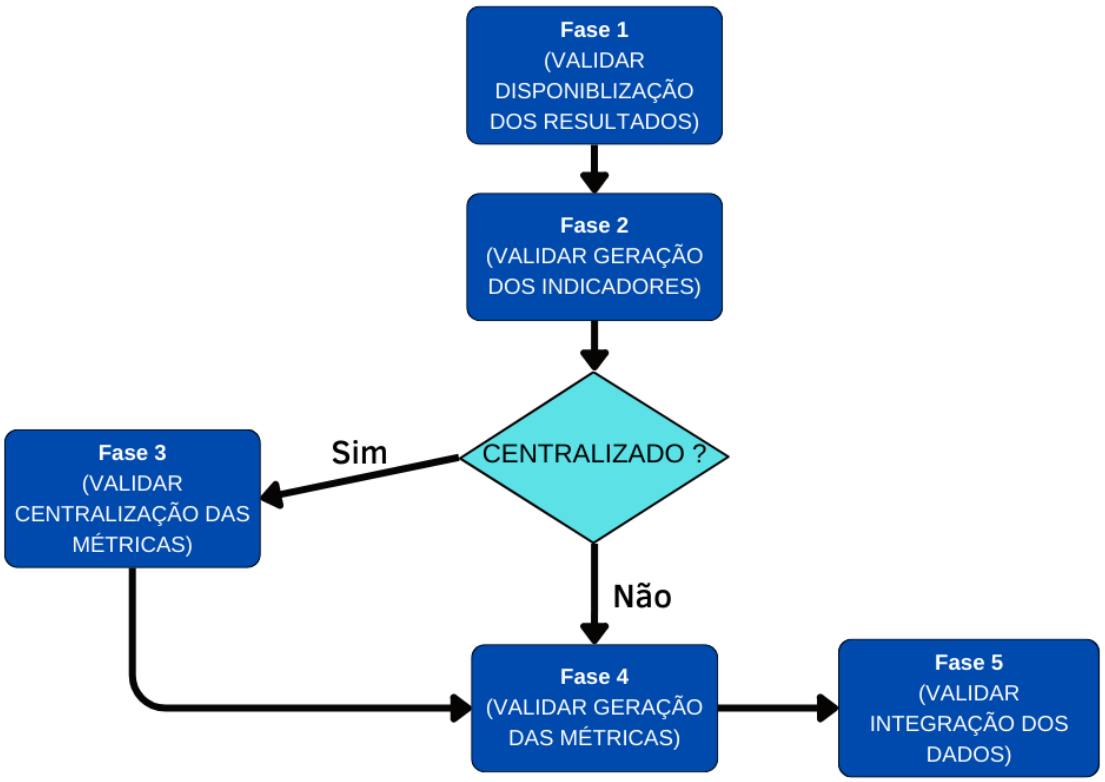
Outro benefício significativo é a capacidade de tomar decisões informadas. Com informações em tempo real e uma representação visual dos indicadores, os tomadores de decisão podem basear suas escolhas em dados sólidos, em vez de intuição ou suposições. Além disso, a presença de metas estabelecidas para cada KPI no painel ajuda a avaliar se a empresa está atingindo seus objetivos estratégicos ou se precisa fazer ajustes em sua estratégia.

Esse painel é utilizado tanto no processo descentralizado nas filiais quanto no processo de geração de indicadores em um Data Warehouse.

3.6 Correção de erros

Durante todo o processo, ocorrem, em todas as fases, validações dos resultados, porém, como a quantidade de dados é muito grande para validar, após a disponibilização dos resultados, cada setor é responsável por validar os seus resultados e, caso ocorra, apontar os erros encontrados. Após o retorno do setor responsável, são iniciados os processos de validação retroativa, ou seja, uma validação “de trás para frente”, iniciando do último processo, disponibilização dos resultados, e retornando para o primeiro, isso ocorre para identificar em qual etapa ocorreu o erro, e assim, podendo tratar mais rapidamente, para o processo seguir seu fluxo original, como é possível observar na Figura 4, abaixo.

Figura 4. Fluxograma do processo completo de validação dos resultados.



Fonte: Elaborado pelo autor (2023)

4 Cenário de teste

Neste capítulo, são apresentadas as etapas do processo, citadas anteriormente, sendo divididas em dois cenários, o processo de forma descentralizada em diferentes filiais e o processo de forma centralizada em um Data Warehouse.

4.1 Processo descentralizado

Nessa seção são apresentados cada etapa do processo para o modelo descentralizado.

4.1.1 Integração dos dados

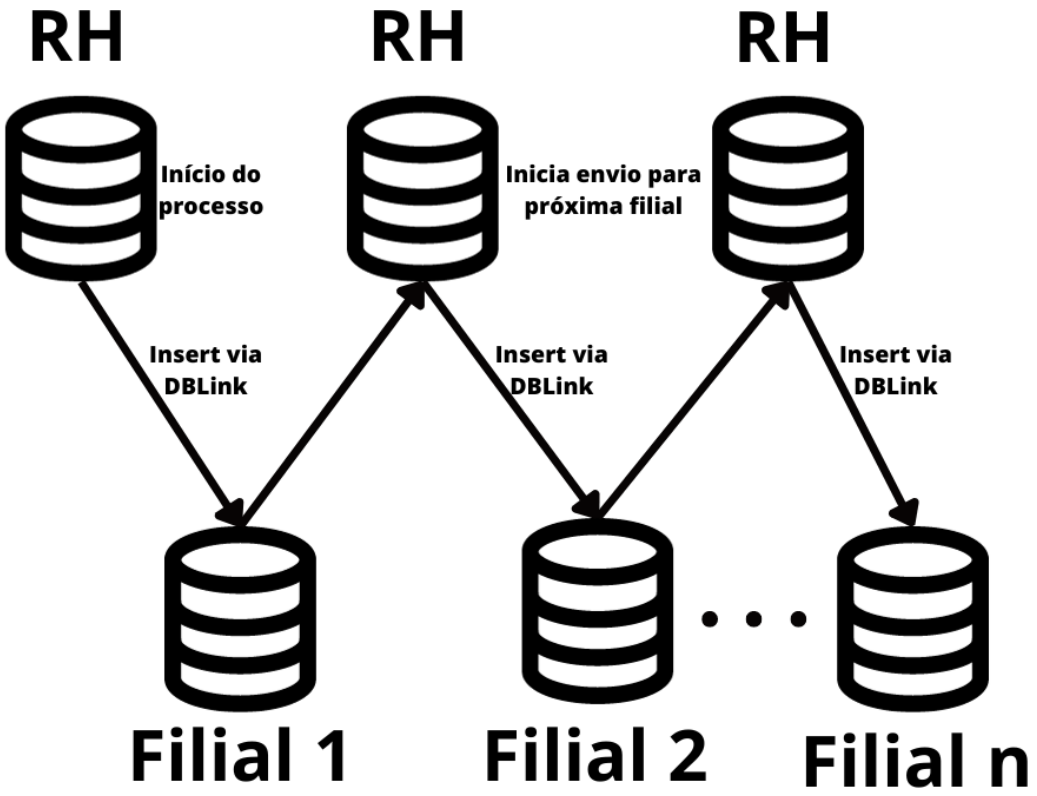
Como mencionado anteriormente, o processo de integração de dados é conduzido utilizando a ferramenta Qlik Data Integration para a transferência de dados. Entretanto, a etapa de processamento desses dados, essencial para os cálculos das métricas, envolve uma série de processos de ETL (Extração, Transformação e Carga). Vale ressaltar que esses processos não têm início nas filiais, pois cada uma delas precisa executá-los, por isso eles tem um gatilho em um único banco de dados que aciona cada filial via DBLink. Para possibilitar essa integração, são utilizados DBLinks, que se originam em outros bancos onde os sistemas geradores de dados estão em operação.

Este tipo de abordagem apresenta desafios consideráveis. Primeiramente, é suscetível a erros e perdas de dados, em parte devido à conexão via internet, tornando o processo dependente da qualidade dessa conexão. Além disso, é custoso, pois não permite a execução paralela entre as filiais, uma vez que envolve tabelas e dados que não estão localmente disponíveis nas próprias filiais.

Vamos considerar um exemplo prático para ilustrar esse processo. Suponhamos o caso do processo de associação das horas trabalhadas e não trabalhadas a cada funcionário. Este processo tem início no banco de dados utilizado pelo departamento de Recursos Humanos da empresa, o qual é distinto dos bancos de dados das filiais. Nesse cenário, um trigger (gatilho) inicia o processo, onde as tabelas referentes ao mês a ser processado têm seus dados excluídos. Posteriormente, para cada filial, é feita uma chamada via DBLink para executar uma operação de INSERT na tabela da respectiva filial. Isso é realizado com base nos dados gerados na tabela do banco de dados do RH, conforme representado na Figura 5.

Esse processo, como descrito, apresenta desafios notáveis em termos de complexidade, dependência de recursos externos e possibilidade de erros. É importante destacar que a otimização dessa etapa, bem como a minimização de erros e custos, é um dos principais objetivos deste projeto.

Figura 5. Processo de geração de dados de horas trabalhadas nas filiais.



Fonte: Elaborado pelo autor (2023)

4.1.2 Geração das métricas

No processo de geração das métricas, após a integração e transformação dos dados, é acionado o procedimento "CalculaMetrica". Este procedimento é executado remotamente em cada filial e na matriz por meio de um DBLink, como ilustrado no Código 1.

A execução desse procedimento envolve a utilização de diversos parâmetros. O parâmetro "p_pedido" é um número que é validado na tabela de pedidos (conforme mostrado no Código 1), sendo registrado para acompanhar o início e o término do processo. Essa informação é fundamental para o registro no banco de dados, conforme indicado no Código 3.

Além disso, o parâmetro "p_id_metrica" define qual métrica será executada, "p_id_banco" representa o ID da filial que será processada por meio do DBLink, "p_info" é um texto descritivo que identifica a natureza da execução, como por exemplo, "Cálculo Mensal". Por fim, "p_data_referencia_ini" e "p_data_referencia_fim" indicam, respectivamente, a data de início e término do mês que está sendo processado.

Código 1. Início da procedure de cálculo das métricas.

```
PROCEDURE CalculaMetrica
( p_pedido              NUMBER
, p_id_metrica          NUMBER
, p_id_banco            NUMBER
, p_info                VARCHAR2
, p_data_referencia_ini VARCHAR2
, p_data_referencia_fim VARCHAR2)
IS
l_pedido_pai GR_PEDIDOS%ROWTYPE;
l_sql_insert            CLOB;
l_sql_delete           CLOB;

BEGIN
    SELECT P.* INTO l_pedido_pai FROM GR_PEDIDOS P WHERE P.cd_pedido=p_pedido;
```

Fonte: Elaborado pelo autor (2023)

Essa execução é iniciada com base em um conjunto de parâmetros essenciais. O parâmetro "p_pedido" é um número que passa por uma validação na tabela de pedidos (vide Código 1). Essa validação é fundamental para registrar o início e o término do processo, tornando-se um indicador crucial, como demonstrado no Código 3.

O parâmetro "p_id_metrica" indica qual métrica específica será calculada. Enquanto "p_id_banco" corresponde ao ID da filial que passará pelo processamento via DBLink. "p_info" é um campo de texto informativo que contextualiza a natureza da execução, por exemplo, "Cálculo Mensal". Já "p_data_referencia_ini" e "p_data_referencia_fim" estabelecem o período de início e término do mês que será processado.

Depois de estabelecidos esses parâmetros, o procedimento inicia deletando os dados que serão atualizados via DBLink (Código 2). Essa ação é baseada no código da métrica e nas datas especificadas nos parâmetros. Em seguida, o procedimento realiza a inserção dos dados das métricas nas filiais, empregando todos os parâmetros definidos (conforme descrito no Código 3).

Código 2. Delete via DBLink da procedure CalculaMetrica.

```
l_sql_delete := 'BEGIN EXECUTE IMMEDIATE ''ALTER SESSION                                '||
                '                              SET NLS_DATE_FORMAT=''''dd/mm/rrrr''''''';'||
               'EXECUTE IMMEDIATE ''ALTER SESSION SET sort_area_size = 2097152000''; '||
                GR_REP_METRICA.FC_GERA_SQL_DELETE ( p_id_metrica
                                                  , p_data_referencia_ini
                                                  , p_data_referencia_fim
                                                  ) ||'; END;';


EXECUTE IMMEDIATE 'BEGIN sinc_exec_remote_'||
                  p_id_banco                ||
                  '('''||REPLACE( l_sql_delete
                                , ''''
                                , '''''')  ||''
                  '); COMMIT; END;';

COMMIT;
```

Fonte: Elaborado pelo autor (2023)

Código 3. Insert via DBLink da procedure CalculaMetrica.

```
    l_sql_insert := 'BEGIN EXECUTE IMMEDIATE ''ALTER SESSION                                '||
                    '                              SET NLS_DATE_FORMAT=''''dd/mm/rrrr''''''';'||
                    'EXECUTE IMMEDIATE ''ALTER SESSION SET sort_area_size = 2097152000''; '||
                     GR_REP_METRICA.FC_GERA_SQL_INSERT ( p_pedido
                                                       , p_id_banco
                                                       , p_info
                                                       , p_id_metrica
                                                       , p_data_referencia_ini
                                                       , p_data_referencia_fim
                                                       ) ||'; END;';


    EXECUTE IMMEDIATE 'BEGIN sinc_exec_remote_'||
                      p_id_banco                ||
                     '('''||REPLACE( l_sql_insert
                                   , ''''
                                   , '''''')  ||''
                     '); COMMIT; END;';


    COMMIT;

END CalculaMetrica;
```

Fonte: Elaborado pelo autor (2023)

4.1.3 Centralização das métricas

A centralização desempenha um papel fundamental no contexto de geração de métricas, pois envolve a consolidação de dados de todas as filiais para proporcionar uma visão panorâmica dos indicadores da empresa na matriz. Esse procedimento é realizado por meio de um DBLink, que atua como uma ponte virtual entre as filiais e a matriz, permitindo a transferência eficiente de informações.

O cerne desse processo é a aglutinação de todos os dados provenientes das filiais. Esses dados são meticulosamente reunidos e, em seguida, encaminhados para a tabela geral de resultados de métricas, localizada na matriz. Essa tabela funciona como um repositório central onde todos os resultados convergem, criando uma visão holística dos indicadores.

Uma característica importante desse processo é a capacidade de associar os resultados às equipes ou funcionários específicos que necessitam dessa visão global. Essa associação é realizada na matriz, onde os resultados são vinculados aos respectivos destinatários, permitindo que essas partes interessadas tenham acesso direto às informações relevantes.

A Figura 2 ilustra visualmente como essa centralização ocorre, destacando o fluxo de dados das filiais para a matriz e a subsequente associação desses dados aos funcionários que precisam de uma visão global dos indicadores. Esse processo desempenha um papel crucial na tomada de decisões estratégicas, fornecendo à alta administração uma visão abrangente do desempenho da empresa em todas as suas operações.

4.1.4 Geração dos Indicadores

A geração de indicadores nas filiais é uma parte crítica do processo de monitoramento e avaliação do desempenho empresarial. Esse processo é tipicamente descentralizado, o que significa que cada filial é responsável por calcular e acompanhar seus próprios indicadores de desempenho. No entanto, para manter a integridade e a visibilidade dos indicadores em toda a organização, é essencial que esses resultados sejam consolidados e enviados para a matriz, onde podem ser analisados em conjunto e usados para tomada de decisões estratégicas.

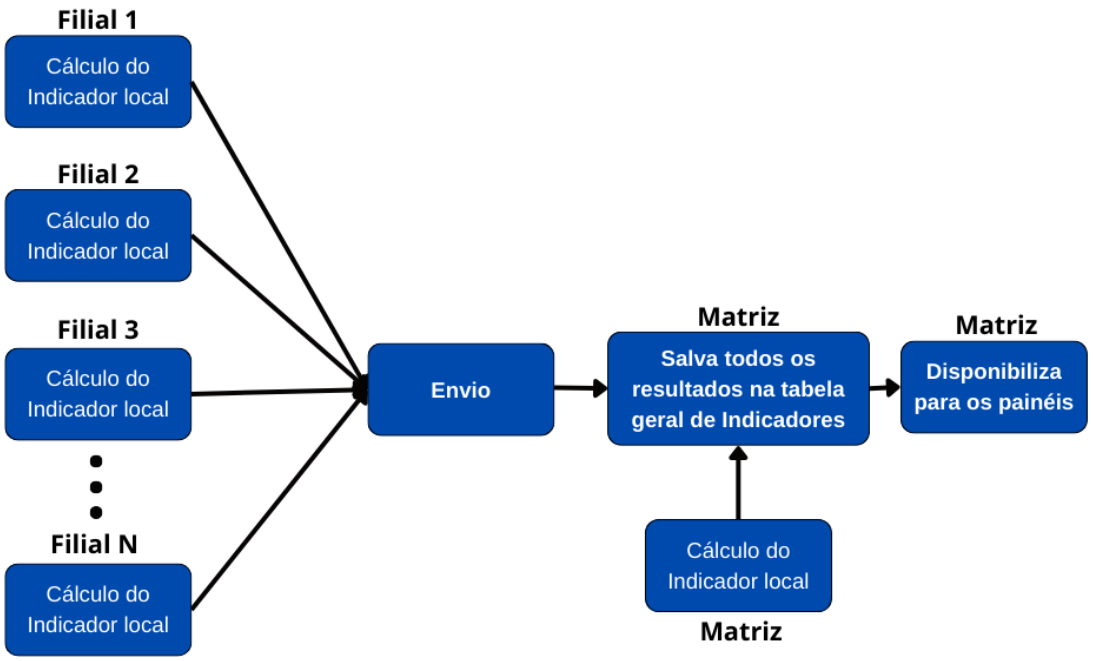
Para realizar essa tarefa de forma eficiente e precisa, é necessário o uso de um DBLink, que atua como uma ponte de comunicação segura entre as filiais e a matriz. O DBLink permite a transferência de dados de maneira automatizada e confiável, facilitando a coleta e a centralização dos resultados dos indicadores.

O processo inicia nas filiais, onde podem ser executados em paralelo entre as filiais, sendo calculados com base nas métricas locais. Esses indicadores podem variar de acordo com a natureza das operações de cada filial e seus objetivos estratégicos. Após o cálculo dos indicadores, os resultados são armazenados em no banco de dados local.

Em seguida, o DBLink é acionado para transferir esses resultados para a matriz. Os dados são enviados de maneira organizada e estruturada, geralmente seguindo um formato padronizado para facilitar a importação na matriz.

Na matriz, os resultados consolidados são recebidos e armazenados em uma tabela central de indicadores de desempenho, que a partir dela serão enviados os resultados para os painéis de visualização (Figura 6). Isso permite que a alta administração e outros tomadores de decisão tenham acesso a uma visão abrangente do desempenho de todas as filiais. Essa visão geral é essencial para a tomada de decisões estratégicas, identificação de tendências e áreas de melhoria, bem como para o acompanhamento do progresso em relação às metas estabelecidas.

Figura 6. Processo de geração e disponibilização de indicadores.



Fonte: Elaborado pelo autor (2023)

4.1.5 Disponibilização dos resultados

Para a visualização dos resultados dos indicadores utilizamos de um site criado pela própria empresa que foi projetado a partir da linguagem de programação Typescript, em conjunto com o software Node.js e o framework de desenvolvimento web Angular, site esse que disponibiliza o resultado tanto por mês quanto o acumulado anual, dados esses advindos a partir de uma materialized view de

resultado dos indicadores presente no banco de dados da matriz, porém a carga dada no site é feita sob demanda (de forma procedural) pois com a necessidade de realizar manutenções frequente no resultados dos indicadores consequentemente os resultados na view acabam por variar corriqueiramente, view essa que espelha a tabela de resultado do banco de dados onde é realizado todo o processo de cálculo mencionado anteriormente.

4.1.6 Correção de erros

O cenário em questão enfrenta desafios significativos, com destaque para a correção de erros. Devido à falta de controle do processo, é comum que os dados resultem em saídas divergentes das esperadas, o que, por sua vez, requer esforços consideráveis de retrabalho e reprocessamento das métricas e indicadores. Essa tarefa se revela dispendiosa, uma vez que não há registros que indiquem onde o processo foi interrompido ou onde ocorreram falhas durante a execução, resultando em uma busca exaustiva pelo erro.

Além disso, outra problemática enfrentada por esse modelo está relacionada à utilização de views para a geração dos dados de valor das métricas. Isso implica que, para corrigir erros quando eles ocorrem, é necessário aprofundar-se cada vez mais nas camadas das views. Um exemplo ilustrativo disso é a materialized view de vendas por vendedor, que tem um intervalo de atualização de 24 horas, o que significa que ela é atualizada apenas uma vez por dia. No entanto, essa materialized view contém duas views, cada uma com suas respectivas consultas para serem preenchidas. Como resultado, eventuais erros na atualização de qualquer tabela utilizada nas views ou falhas na atualização da materialized view podem afetar a precisão dos valores das métricas, uma vez que essas métricas se baseiam nessa view.

Esses desafios ressaltam a importância de implementar um controle mais eficaz do processo e de explorar alternativas para a geração de métricas, a fim de evitar a ocorrência de erros e otimizar o fluxo de trabalho.

4.2 Data Warehouse

Nessa seção são apresentados cada etapa do processo para o modelo centralizado.

4.2.1 Integração dos dados

Assim como no cenário anterior, é utilizada a ferramenta Qlik Data Integration para transporte de dados. Porém, os processamentos dos dados a serem utilizados nas métricas, como todos os dados são enviados para o Data Warehouse, não se faz necessário o uso de DBLINKs, já que todos os dados podem ser encontrados localmente.

Isso pode ser observado na Figura 7, abaixo, onde é mostrado como os dados de horas trabalhadas são adicionados a tabela de resultados do RH a partir de outras tabelas encontradas no próprio banco de dados do Data Warehouse.

Da mesma forma que no cenário anterior, a ferramenta Qlik Data Integration é empregada para o transporte de dados. No entanto, uma distinção fundamental se encontra no processamento desses dados para a geração das métricas. Nesse caso, uma vez que todos os dados são centralizados e armazenados no Data Warehouse, a necessidade de utilizar DBLINKs é eliminada, uma vez que todas as informações podem ser prontamente acessadas localmente.

Figura 7. Processo de geração de dados de horas trabalhadas no DW.

```
BEGIN


  DELETE
    FROM IND_RH_RESULTADOS
   WHERE DT_REFERENCIA BETWEEN p_DtRef
                           AND LAST_DAY(p_DtRef);
  COMMIT;

  INSERT INTO IND_RH_RESULTADOS (
  SELECT F01.*
       , I01.*
    FROM (SELECT *
            FROM TB_RH_INDICADOR C
           WHERE 1 = 1
             AND C.DATA BETWEEN p_DtRef
                            AND LAST_DAY(p_DtRef)
         )            I01
    JOIN TB_FUNC_HIST F01
      ON I01.MATRICULA = F01.MATRICULA
     AND LAST_DAY(TO_DATE(F01.DT_REF, 'DD/MM/RRRR')) = LAST_DAY(I01.DATA)
  );
  COMMIT;
```

Fonte: Elaborado pelo autor (2023)

4.2.2 Geração das métricas

No processo de geração de métricas, para se obter um maior controle do processo e uma melhor gestão de erros, cada métrica recebeu uma procedure para gerar seus valores. Houve também a divisão por tipo de métrica, ou seja, as métricas com mais de um tipo receberam duas procedures, por exemplo a métrica de Tempo médio de atendimento de caixa, onde cada funcionário que atua como caixa recebe o resultado dessa métrica referente a sua matrícula, em contrapartida

os gerentes responsáveis por todos os caixas recebem o valor de tempo médio de toda a filial. Precisando assim, existir duas procedures para cada forma de cálculo.

Para facilitar a gestão dos resultados, as etapas de cálculo das métricas foram organizadas em tabelas temporárias. Essa estrutura segmentada permite analisar cada fase do processo de maneira mais precisa, simplificando a compreensão do resultado final do cálculo. Para exemplificar, focaremos no cálculo do valor real do absenteísmo, o qual indica quantas faltas foram registradas por um funcionário em um determinado intervalo de tempo.

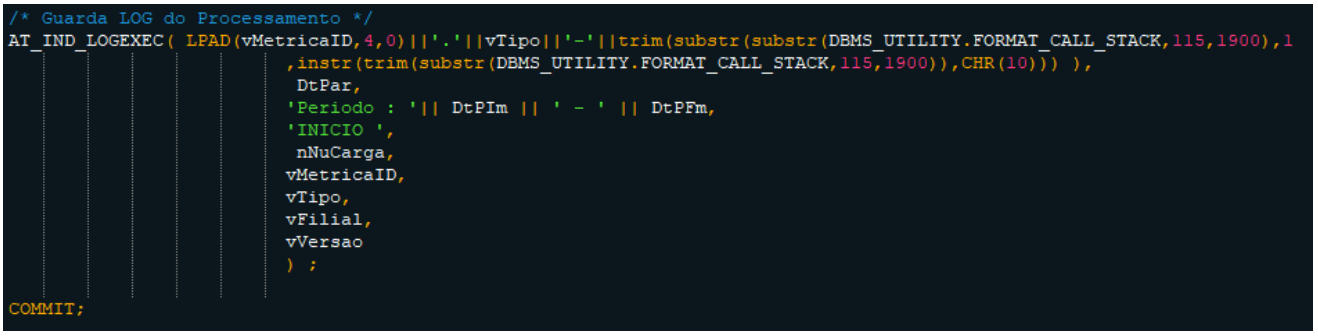
Os procedimentos de métrica são alimentados com cinco parâmetros: filial, data de referência, número de carga, perfil e matrícula. Caso algum deles não seja especificado durante a chamada da rotina, valores padrão são empregados. Por exemplo, se a data de referência não for fornecida, a data atual é utilizada. No entanto, para perfil e matrícula, todos os valores são inicialmente considerados, sendo posteriormente filtrados durante o processamento dos resultados, garantindo que apenas colaboradores relevantes contribuam para o cálculo. O número de carga, que possui um valor padrão incremental, é incorporado ao registro de execução dos procedimentos. Esse número atua como referência no log, possibilitando a identificação e rastreamento das execuções.

Os parâmetros passam por um tratamento para conversão de tipos e obtenção de informações essenciais. Por exemplo, a última data do mês é extraída com o intuito de determinar o período de cálculo (data inicial e final). Em seguida, a rotina verifica se a procedure deve ser executada para uma determinada filial, conforme o parâmetro fornecido na chamada do processamento. Caso afirmativo, o cálculo da métrica é iniciado, caso contrário, a execução é interrompida. No início do cálculo, a data e hora de início do processamento são registradas na variável "vVersao". Em seguida, a função de registro no log de execução das procedimentos é acionada para documentar o início do processamento, conforme ilustrado no Código 4.

As procedures de cálculo das métricas foram organizadas em tabelas temporárias para facilitar a manutenção dos resultados. Essa divisão permite visualizar de forma mais específica cada etapa do processo e entender o resultado final do cálculo com mais clareza, para o exemplo em questão que será o cálculo do valor Real do Absenteísmo que representa o quanto foi registrado de falta para aquele funcionário em determinado período de tempo.

É realizado um tratamento dos parâmetros a fim de trocar os tipos e obter informações que serão utilizadas posteriormente a partir deles, como por exemplo a última data do mês com o objetivo de ter o período para cálculo (Data inicial e final). Em seguida, a rotina checa se a procedure deve ser calculada para uma determinada filial recebida como parâmetro na chamada do processamento. Se sim, é dado o início do processo de cálculo da métrica, caso contrário, a execução é interrompida. No início do cálculo da métrica, a data e a hora do início do processamento são registradas na variável "vVersao". Depois disso, a função de gerar registro no log de execução das procedures é chamada para registrar o início do processamento o, como mostrado no Código 4.

Código 4. Chamada da função de registro no log para o início do cálculo.



```
/* Guarda LOG do Processamento */
AT_IND_LOGEXEC( LPAD(vMetricaID,4,0)||'.'||vTipo||'-'||trim(substr(substr(DBMS_UTILITY.FORMAT_CALL_STACK,115,1900),1
                    ,instr(trim(substr(DBMS_UTILITY.FORMAT_CALL_STACK,115,1900)),CHR(10))) ),
                     DtPar,
                    'Periodo : '|| DtPIm || ' - ' || DtPFm,
                    'INICIO ',
                     nNuCarga,
                    vMetricaID,
                    vTipo,
                    vFilial,
                    vVersao
                    ) ;

COMMIT;
```

Fonte: Elaborado pelo autor (2023)

Como mencionado anteriormente, as procedures utilizam tabelas temporárias para alcançar o resultado final. Cada etapa da procedure consiste na limpeza e inserção do próximo passo na próxima tabela temporária. Essas tabelas são posteriormente utilizadas dentro da própria procedure.

A manipulação da primeira tabela temporária (chamada de temporária 00) obtêm os perfis que devem receber o resultado para aquela métrica em específico assim como seu tipo e a descrição dela. Todas as tabelas temporárias referente a cada métrica são truncadas antes de serem carregadas.

O próximo passo, sendo a temporária 1, retorna informações dos funcionários relacionados aos perfis que recebem os resultados da matrícula, incluindo nome, matrícula, cargo e outras informações. Essas informações são importantes para a inserção dos valores na tabela de resultados de métricas. A Temporária 1 é filtrada com base no perfil e é relacionada com a Temporária 0, como mostrado no Código 5. Além disso, a tabela de dados utilizada na Temporária 1 é filtrada por datas iniciais e finais, garantindo que apenas funcionários que estavam no perfil em questão naquele período sejam incluídos no cálculo. Também há um filtro de matrícula para permitir o cálculo de resultados apenas para um funcionário específico, se necessário.

Código 5. Tratamento da tabela temporária 1 na procedure de cálculo da métrica de absenteísmo real.

```
------------TEMP1
vSql := 'TRUNCATE TABLE "IN_MGM_ABS_REAL_TMP_01"';
EXECUTE IMMEDIATE vSql;
COMMIT;

INSERT INTO IN_MGM_ABS_REAL_TMP_01 (
SELECT FU.DT_REF           AS DT_REF
     , FU.MATRICULA        AS MATRICULA
     , FU.NOME             AS NOME
     , FU.EMAIL            AS EMAIL
     , FU.SITUACAO         AS SITUACAO
     , FU.ID_PERFIL        AS ID_PERFIL
     , FU.CARGO            AS CARGO
     , FU.CARGO_ARVORE     AS CARGO_ARVORE
     , FU.LOCAL            AS LOCAL
     , FU.CCUSTO           AS CCUSTO
     , FU.CCUSTO_DESCRICAO AS CCUSTO_DESCRICAO
     , FU.FILIAL           AS FILIAL
     , FU.AREA             AS AREA
     , FU.AREA_DESCRICAO   AS AREA_DESCRICAO
     , FU.DATAADMISSAO     AS DATAADMISSAO
     , FU.AVALIADOR        AS AVALIADOR
     , FU.DT_REFER         AS DT_REFER
  FROM TB_FUNC_HIST FU
 WHERE 1 = 1
   AND TO_DATE(FU.DT_REF , 'DD/MM/RRRR') BETWEEN vDtIni AND vDtFim
   AND FU.FILIAL         = vFilial
   AND FU.MATRICULA      = DECODE(pMatricula,NULL,FU.MATRICULA,pMatricula)
   AND FU.ID_PERFIL     IN (SELECT TMP00.ID_PERFIL
                              FROM IN_MGM_ABS_REAL_TMP_00 TMP00
                             WHERE TMP00.ID_PERFIL = DECODE(pPerfil,NULL,TMP00.ID_PERFIL,pPerfil))
);
COMMIT;
```

Fonte: Elaborado pelo autor (2023)

A Temporária 2 diferente das duas anteriores não segue um padrão pois nela é realizado a consulta no banco para obter as colunas que são utilizadas para o cálculo. A temporária 3 inicia o processo de transformação dos dados buscados na tabela temporária anterior, associando informações da Temp 2 com a Temp 1 para gerar resultados de métricas. É também nesta temporária que ocorre a diferenciação do tipo de cálculo para gerar o resultado, ou seja se o resultado é específico para cada funcionário ou se ocorre um resultado geral para toda filial, por exemplo.

A Temp 4 relaciona os dados da terceira temporária com dimensões presentes no data Warehouse, como data, filial, cargo, situação do funcionário, local e centro de custo. Essas são as dimensões padrões para todas as procedures de cálculo das métricas. A temporária 4 recebe em sua cláusula select a primary key para cada dimensão a partir da relação realizada com a temporária 3. A temporária 5 calcula a métrica em si, utilizando a coluna de cálculo conduzida desde a temporária 2. Nesse caso, o cálculo é simples, apenas somando as horas da coluna absenteísmo advinda do data lake de RH e agrupando essa soma por funcionário. A tabela temporária também é responsável por gerar o ID para a tabela de fatos, combinando o número da métrica, a matrícula e o tipo. Além disso, essa tabela

temporária está associada à tabela temporária 0 devido às colunas que serão utilizadas para inserir os resultados na tabela de métricas.

As temporárias de 2-7 não necessariamente tem essa numeração, pois a temporária 2 podem ser “divididas” em diversas outras tabelas temporárias a fim de destrinchar a obtenção dessas colunas para o cálculo. Por exemplo, a tabela temporária 7 pode, na verdade, corresponder à tabela temporária 3, se considerarmos que 3 (número da tabela temporária) somado a 4 (número da tabela temporária extra da tabela temporária 2) resulta em 7.

Os resultados das métricas são inseridos na tabela de resultado das métricas. Antes da inserção, é feita uma exclusão na tabela filtrando os parâmetros passados pelo usuário a fim de os atualizar com os novos resultados, como é demonstrado no Código 6.

Código 6. Processo de deleção na tabela de resultado das métricas.

```
----------RESULTADO_IND_MET
c_TabMet := 'RESULTADO_IND_MET_'|| c_Ano || c_Mes;

v_SQL := 'DELETE FROM ' || c_TabMet || ' MET '                                  ||
         ' WHERE MET.FILIAL     = ' || vFilial                                  ||
         '   AND MET.ID_METRICA = ' || vMetricaID                               ||
         '   AND MET.TIPO       = ' || vTipo                                    ||
         '   AND MET.MATRICULA IN (SELECT TMP01.MATRICULA                     ' ||
         '                           FROM IN_MGM_ABS_REAL_TMP_01 TMP01)       ' ;

EXECUTE IMMEDIATE v_SQL;
COMMIT;
```

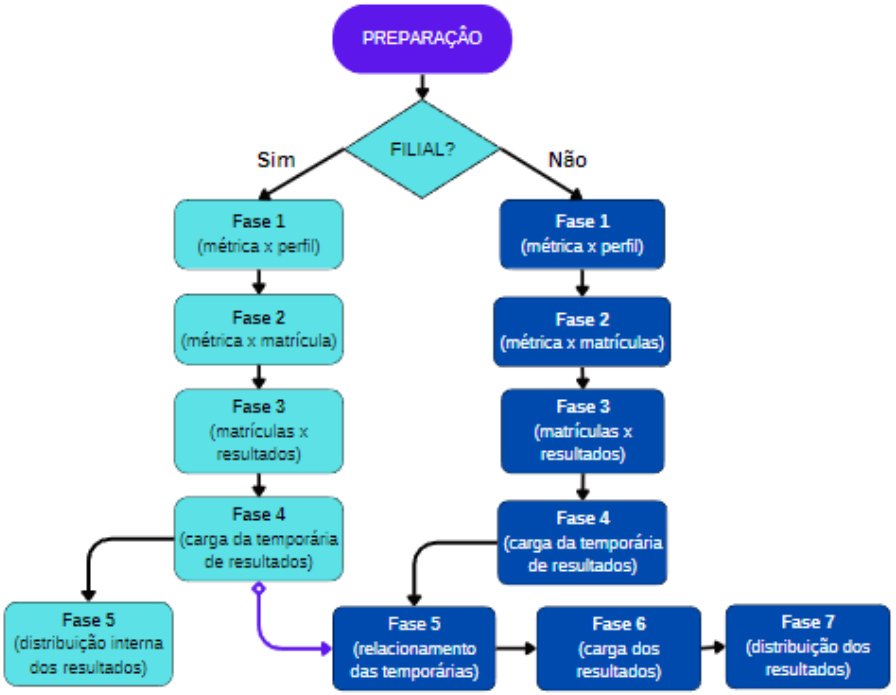
Fonte: Elaborado pelo autor (2023)

A inserção leva os dados da temporária 5 e outras variáveis calculadas durante o processo, como por exemplo a data em que foi realizada o procedimento de cálculo da métrica. Após o fim da inserção na tabela de resultado de métrica, um log é gerado indicando o fim do processo.

4.2.3 Centralização das métricas

O objetivo da centralização das métricas é enviar os resultados de cálculos do processo anterior de determinados colaboradores para outros funcionários com perfis estratégicos que os utilizarão para calcular indicadores. Por exemplo, o indicador de Absenteísmo pode ser usado pela gerente de RH para formular estratégias para melhorar o quadro geral. A centralização é feita através de uma procedure, sendo esse processamento diferente dependendo de para onde é realizado o cálculo (na filial ou na matriz) como é descrito no fluxograma da Figura 8.

Figura 8. Fluxograma do processo de centralização dos resultados das métricas.



Fonte: Elaborado pelo autor (2023)

A Rotina recebe em sua chamada um conjunto de parâmetros que inicialmente se assemelha com a procedure de cálculo de métricas (quanto o parâmetro de filial, data e número de carga) porém a mesma apresenta mais 2 parâmetros que são o tipo e o id da métrica ambos recebendo um valor padrão para o caso do usuário não os adicioná-lo na chamada da rotina. O tipo tem como valor padrão ‘D’ que representa o cálculo diário e isto está sendo utilizado como medida de segurança para não acontecer um processamento desnecessário (caso não seja diário é utilizado F), já a métrica toma como padrão ‘9999’ que não representa o id de nenhuma das métricas porém está significando que a centralização é executada para todas as métricas. Se o parâmetro da filial não for fornecido, o processo de centralização é executado para todas as filiais utilizando um cursor. Esse cursor é criado filtrando a tabela de dimensão da empresa com base no parâmetro de entrada pFilial, que é formatado como a variável numérica vFilial.

Assim como a procedure de métrica é gerado a variável para o número de carga, para o registro no log de execução das rotinas, além disso há a checagem a partir da métrica(s) para checar para quais filiais aquela métrica está sendo atribuída. A rotina, antes de começar a replicação dos resultados, envia para o log de execução (mesmo log utilizado para as procedures de métricas) para assim registrar o início da rotina. Em seguida é gerado uma temporária de resultado base para cada filial que posteriormente será buscada pela matriz.

É preciso ter em vista que ocorre dois tipos de replicação, uma voltada especificamente para as filiais e uma voltada para a matriz, com isso será apresentado o processo para cada uma. O tipo de execução que é processado é definida por uma estrutura lógica que se decide a partir do parâmetro de filial

passado na chamada da rotina, como é apresentado no fluxograma esta definição se dá logo após a preparação.

A rotina de replicação dos dados para a filial é composta por cinco fases, onde cada uma delas é registrada pelo Log de execução das rotinas, com o registro de seu início e fim. Na primeira fase, a tabela temporária que armazena informações sobre as métricas alvo é truncada e nela são armazenadas as métricas que foram passadas, juntamente com os perfis que devem receber os valores centralizados. Se um valor padrão (9999) foi passado na chamada da função, todas as métricas e seus respectivos perfis são centralizados. É importante destacar que o perfil 2 sempre recebe o todo centralizado de suas respectivas filiais, já que ele representa o perfil de diretor, sendo necessário para eles obterem informações sobre toda a empresa. Portanto, não há métricas sem pelo menos uma centralização.

A fase 2 lida com a segunda temporária da rotina a tabela temporária de métricas por matrícula, inicialmente após trunca-la, nela há a recuperação de todas métricas por matrículas para cada filial, esses dados serão advindos da dimensão de filial sendo relacionada com a tabela que é armazenado o resultado dos cálculos das métrica.

Durante a fase 3, ocorre a inserção na tabela que armazena a relação entre métrica, matrícula e perfil, onde é relacionado os resultados obtidos nas fases anteriores. Essas informações são filtradas e combinadas com dados da tabela de histórico dos funcionários, que contém o histórico dos funcionários, onde por mês é retornado as informações referentes aos colaboradores, e da tabela responsável por armazenar os relacionamentos entre as métricas e análises. De maneira estruturada, os resultados obtidos por todos os colaboradores presentes na filial são atribuídos ao perfil de diretor geral correspondente.

A fase 4 armazena na já mencionada temporária base específica para cada filial o resultado (já na estrutura correta) do cálculo de métrica que é enviado posteriormente à matriz, como está sendo mostrado no fluxograma da Figura 8, porém este processo só ocorre durante a replicação na matriz. O resultado é advindo do relacionamento da temporária da Fase 1 com a tabela de histórico dos funcionários e a tabela de resultado de métrica (filtrada apenas para o resultado do diretor como foi inserido na fase 3).

Na fase 5 ocorre o envio do resultado da centralização para os perfis estratégicos da própria filial, onde cada supervisor de setor recebe as informações referentes aos seus funcionários a partir do resultado obtido na temporária 3, esse resultado não influencia no quadro geral da empresa mas influencia no resultado e nas decisões estratégicas destes perfis que recebem tais resultados.

O processo de replicação para a matriz diferente de para a filial é realizado em 7 fases onde cada fase está sendo registrado pelo Log de execução, tanto seu início de execução quanto seu fim. As 3 primeiras fases da replicação da matriz seguem o padrão da Replicação das filiais, onde a geração das 3 temporárias segue o mesmo modelo tanto ao nome quanto a seu processo de carga (porém voltada apenas à matriz).

A quarta fase lida com uma nova temporária onde ela, após truncada é carregada com as matrículas (que não são da diretoria) que devem receber resultados centralizados, esses dados são advindos de uma relação entre as tabelas tabela que armazena as informações sobre as análises, métricas e indicadores, tabela que relaciona os perfis aos agrupamentos das métricas assim como a informação a respeito desse relacionamento, como por exemplo seu tipo, a tabela que relaciona as métrica aos seus agrupamentos e por fim a tabela de histórico dos funcionários.

A quinta fase opera com a tabela temporária que será armazenado o resultado de todas as centralizações advindas das temporárias bases das filiais, a partir de um loop utilizando dos ids dos bancos das filiais e também o nome de cada temporárias bases das filiais referente a cada uma delas, ou seja é concatenado nesta tabela temporária os resultados que devem ser centralizados para a matriz referente ao intervalo de mês que foi adquirido a partir da chamada da rotina, como é demonstrado no Código 7.

Código 7: Fase 5 da Centralização das métricas na matriz.

```
EXECUTE IMMEDIATE 'TRUNCATE TABLE "REP_METRICA_TMP"';
COMMIT;

FOR C in (SELECT NVL(L.CD_EMPRESA, 99)                              AS CD_EMPRESA
               , B.ID                                               AS BANCO
               , 'REP_METRICA_TMP_' || SUBSTR(L.DS_EMPRESA, 3) AS REPMETRICA
            FROM BD_INF        B
               , BD_FILIAIS L
           WHERE B.ID                        = L.ID_BANCO (+)
             AND NVL(L.CD_EMPRESA, 99) NOT IN (0, 99)
             AND B.ATIVO                     = 1
           ORDER BY NVL(L.CD_EMPRESA, 99)
         )
  LOOP

    vSql := 'INSERT INTO INDICADORES.REP_METRICA_TMP              ' ||
            ' (SELECT R.*                                          ' ||
            '    FROM '||C.RepMetrica||' R                         ' ||
            '   WHERE 1 = 1                                        ' ||
            '     AND R.DATA                                       ' ||
            '         BETWEEN TO_DATE('''||vDtIni||''',''DD/MM/RRRR'')' ||
            '             AND TO_DATE('''||vDtFim||''',''DD/MM/RRRR'')' ||
            ' )                                                    ' ;

    EXECUTE IMMEDIATE vSql;
    COMMIT;

  END LOOP;
COMMIT;
```

Fonte: Elaborado pelo autor (2023)

A sexta fase deleta da tabela de resultado das métricas os dados referente ao que foi centralizado, ou seja todos aqueles dados que são advindos da filial para a centralização na matriz, pois ainda na fase 6 estes dados são inseridos na tabela de resultado das métricas o que foi obtido na Fase 5 para o perfil de diretor geral (que tem seu resultado baseado em todos os funcionários).

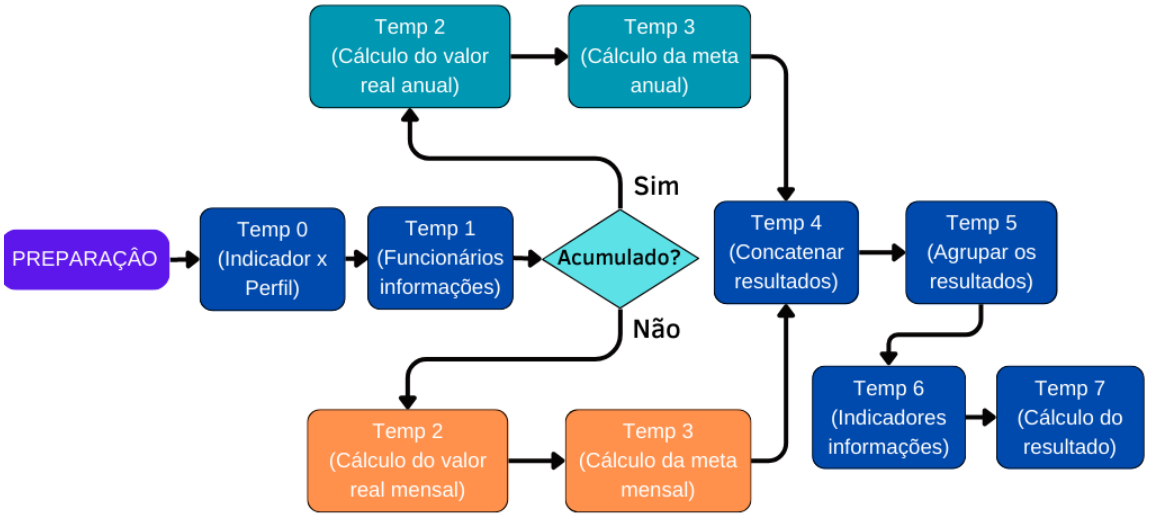
Por fim a fase 7 distribui os resultados que foram centralizados para o diretor para os outros colaboradores que também devem receber estes resultados, essa inserção se dá utilizando a própria tabela de resultado de métrica filtradas pela matrícula do diretor geral (contendo os resultados que são distribuídos), a tabela da temporária 4 (contendo as matrículas e as métricas para onde os resultados são distribuídos) e a View que relaciona os perfis com as métricas e o seu tipo, finalizando assim a centralização onde os resultados já estão distribuídos para seus respectivos perfis.

A centralização deve ser realizada sequencialmente inicialmente nas filiais e posteriormente na matriz, pois a rotina está sempre truncando a temporária contendo o resultado centralizado das filiais e apenas quando essa rotina é chamada na matriz que ocorre a inserção de resultados na tabela de resultado das métricas.

4.2.4 Geração dos indicadores

A realização do cálculo dos indicadores envolve o uso dos resultados das métricas relacionadas, bem como da centralização de seus resultados. Para esse fim, tanto o cálculo do indicador quanto o da métrica são acompanhados por tabelas temporárias que ajudam a organizar a geração dos resultados. No entanto, diferentemente das métricas, o número de tabelas temporárias para os indicadores é padrão para todos, além disso o indicador não irá buscar os dados de um data lake, ela busca no data mart dos resultados das métricas. Para ilustrar esse processo, vamos considerar o exemplo do indicador de absenteísmo, que utiliza três métricas em seu cálculo. Foi criado um fluxograma para descrever o processo de geração desse resultado na procedure correspondente que pode ser visualizado na Figura 9.

Figura 9. Fluxograma do fluxo do ETL do cálculo dos indicadores.



Fonte: Elaborado pelo autor (2023)

A rotina de cálculo dos indicadores está seguindo um padrão semelhante à rotina de cálculo das métricas, também sendo um processo de ETL. No entanto, há uma diferença importante: a chamada da função recebe um parâmetro adicional que indica o tipo de cálculo do indicador. O valor "1" é usado para o cálculo padrão, enquanto o valor "2" é usado para o cálculo anual, que não faz parte da análise desse estudo.

Além disso, ao contrário da rotina de processamento das métricas, a rotina de cálculo de indicador altera a data para o intervalo do ano inteiro caso o cálculo do acumulado seja solicitado na chamada do processo, e não apenas para o mês solicitado como ocorre no processamento das métricas.

Em seguida a rotina gera um número de carga, assim como é gerado para o número de carga da métrica e registra o início do processamento no log. Depois disso, verifica se a rotina deve ser executada para a filial solicitada e transforma os parâmetros, se necessário, como por exemplo, formatando as datas para se obter a data inicial e final para o processamento. O cálculo dos indicadores, assim como as métricas, utilizam de tabelas temporárias para acompanhar o processamento e organizar as fases da execução. No entanto, o número de temporárias para os indicadores, diferente da métrica, é padrão, sendo nove.

A temporária 0 contém perfis e métricas relacionadas ao indicador e passa por uma limpeza antes de ser carregada. Para melhorar sua performance, a função Gather_table_stats é utilizada, ela coleta informações sobre a tabela recebida como parâmetros, incluindo índices e colunas, otimizando as consultas do banco de dados para tomar decisões mais eficientes na execução das consultas, como escolher o

melhor plano de execução. O mesmo comando é executado para todas as tabelas temporárias manipuladas no cálculo do indicador.

A temporária 1, após truncada, armazena informações dos funcionários que contêm os perfis contidos na temporaria 0, filtrando aos colaboradores que contenham os perfis/matrículas passados na chamada da procedure, e da tabela de histórico dos funcionários além de só manter os resultados dos colaboradores pertencentes a filial alvo do cálculo.. O Código 8 demonstra esse processo.

Código 8. Deleção e carga na temporária 1 no cálculo do indicador de absenteísmo

```
------------TEMP1
vSql := 'TRUNCATE TABLE "IN_IND_ABS_TMP_01"';
EXECUTE IMMEDIATE vSql;
COMMIT;

INSERT INTO IN_IND_ABS_TMP_01 (
SELECT FU.DT_REF           AS DT_REF
     , FU.MATRICULA        AS MATRICULA
     , FU.NOME             AS NOME
     , FU.EMAIL            AS EMAIL
     , FU.SITUACAO         AS SITUACAO
     , FU.ID_PERFIL        AS ID_PERFIL
     , FU.CARGO            AS CARGO
     , FU.CARGO_ARVORE     AS CARGO_ARVORE
     , FU.LOCAL            AS LOCAL
     , FU.CCUSTO           AS CCUSTO
     , FU.CCUSTO_DESCRICAO AS CCUSTO_DESCRICAO
     , FU.FILIAL           AS FILIAL
     , FU.AREA             AS AREA
     , FU.AREA_DESCRICAO   AS AREA_DESCRICAO
     , FU.DATAADMISSAO     AS DATAADMISSAO
     , FU.AVALIADOR        AS AVALIADOR
     , FU.DT_REFER         AS DT_REFER
  FROM TB_FUNC_HIST FU
 WHERE 1 = 1
   AND TO_DATE(FU.DT_REF, 'DD/MM/RRRR') BETWEEN vDtIni AND vDtFim
   AND FU.FILIAL = vFilial
   AND FU.ID_PERFIL IN (SELECT ID_PERFIL
                          FROM IN_IND_ABS_TMP_00 TMP00
                         WHERE TMP00.Id_Perfil = DECODE(pPerfil,Null,TMP00.Id_Perfil,pPerfil))
   AND FU.Matricula = DECODE(pMatricula,Null,FU.Matricula,pMatricula)
);
COMMIT;
```

Fonte: Elaborado pelo autor (2023)

A temporária seguinte recebe a carga com o resultado das métricas (tanto os calculados quanto os centralizados) relacionadas ao indicador a partir da tabela de resultado de métricas utilizando a data passada como parâmetro na chamada da função, para obter os resultados que são referentes ao período que deve ser calculado, como pode ser visto no Código 9. Essa temporária foi nomeada de forma diferente pois ela possui uma função distinta das temporárias de cálculo tendo maior enfoque na preparação que no próprio cálculo.

Código 9.Trecho de código em que está sendo Truncada e carregada a Temporária de extração dos resultados das métricas no indicador.

```
------------ METRICA_RESUL
vSql := 'TRUNCATE TABLE "IN_IND_ABS_MET_RESUL"';
EXECUTE IMMEDIATE vSql;
COMMIT;


vSql := 'INSERT INTO IN_IND_ABS_MET_RESUL (                                    ' ||
        'SELECT *                                                              ' ||
        '  FROM VW_GR_METRICA_' || c_Ano || ' VW                               ' ||
        ' WHERE VW.DATA BETWEEN ''' || vDtIni || ''' AND ''' || vDtFim  ||'''  ' ||
        '   AND VW.ID_METRICA IN (SELECT DISTINCT ID_METRICA FROM IN_IND_ABS_TMP_00))          ' ;

EXECUTE IMMEDIATE vSql;
COMMIT;
```

Fonte: Elaborado pelo autor (2023)

A Temporária 2 e 3 possuem uma peculiaridade pois ambas são distintas dependendo do tipo de cálculo de indicador que foi passado como parâmetro, ou seja, há uma maneira de carregar a temporária para o cálculo corrente e outra para o cálculo do acumulado. Porém é padrão que essas temps recebem a fórmula de cálculo entre os resultados obtidos pelas métricas relacionadas a ele, sendo a tabela temporária 2 responsável pelo valor real e a temporária 3 responsável pela meta.

Para o cálculo padrão é apenas utilizado o mês recebido na chamada da função, sendo o tipo de cálculo é mensal, a temp 2 no caso do exemplo que utilizamos (Absenteísmo) utiliza da soma dos resultados da métrica absenteísmo real sobre o valor da soma dos resultados métrica de absenteísmo total para cada matrícula, a coluna ‘Desc_fato’ traz a descrição da métrica sendo ela chamada na geração da coluna de ‘Valor_Real’, é utilizado o decode devido ao resultado ser obtido de diferentes métricas advindas da mesma tabela (no caso a temporária carregada com o resultado da tabela sendo filtrada pelo período solicitado) sendo assim caso não atenda o que foi solicitado o 0 é irrelevante para o resultado da soma, assim como é demonstrado no Código 10.

Código 10. Carga da Temporária 2 para o cálculo padrão do indicador de Absenteísmo.

```
INSERT INTO IN_IND_ABS_TMP_02 (
SELECT MIN(M.DATA)                                                    AS ORDER_D
     , TO_CHAR(M.DATA, 'MM/RR')                                       AS DATA
     , M.MATRICULA                                                    AS MATRICULA
     , Max(Tmp01.Nome)                                                AS Nm_Func
     , 'R'                                                            AS TIPO
     , Decode(SUM(Decode(m.Desc_Fato,
                         'Absenteismo Total',
                         m.Valor_Real,
                         0)),
              0,
              0,
              (SUM(Decode(m.Desc_Fato,
                          'Absenteismo Real',
                          m.Valor_Real,
                          0))) /
              (SUM(Decode(m.Desc_Fato,
                          'Absenteismo Total',
                          m.Valor_Real,
                          0))) * 100)                                 AS Valor_Real
     , TMP01.ID_PERFIL                                                AS ID_PERFIL
     , COUNT(DISTINCT LAST_DAY(M.DATA))                               AS QTD_MESES
  FROM IN_IND_ABS_MET_RESUL M
  JOIN (SELECT DISTINCT
               VAT1.MATRICULA
             , VAT1.ID_PERFIL
             , VAT1.NOME
          FROM IN_IND_ABS_TMP_01 VAT1) TMP01
    ON TMP01.MATRICULA = M.MATRICULA
   AND TMP01.ID_PERFIL = M.ID_PERFIL
 WHERE M.ID_METRICA IN (SELECT DISTINCT ID_METRICA FROM IN_IND_ABS_TMP_00)
   AND M.DATA BETWEEN vDtIni AND vDtFim
 GROUP BY TO_CHAR(M.DATA, 'MM/RR'), M.MATRICULA, TMP01.ID_PERFIL
 );
COMMIT;
```

Fonte: Elaborado pelo autor (2023)

Já a Temporária 3 segue a mesma estrutura da temporária 2 porém é voltada ao resultado da meta. No caso do indicador que está sendo exemplificado, a meta apenas necessita buscar da tabela temporária de resultado das métricas o resultado referente às metas como é possível ver no Código 11, sem necessitar realizar um cálculo como foi realizado no real (temporária 2). Ambas as temporárias foram truncadas antes de realizar as novas inserções.

Código 11. Carga da Temporária 3 para o cálculo padrão do indicador de Absenteísmo.

```
INSERT INTO IN_IND_ABS_TMP_03 (
SELECT MIN(M.DATA)                                          AS ORDER_D
     , TO_CHAR(M.DATA, 'MM/RR')                             AS DATA
     , M.MATRICULA                                          AS MATRICULA
     , Max(Tmp01.Nome)                                      AS Nm_Func
     , 'M'                                                  AS TIPO
     , MAX(Decode(m.Desc_Fato,
                  'Absenteismo Meta',
                  m.Valor_Real,
                  0))                                       AS Valor_Real
     , COUNT(*)                                             AS QTD
     , TMP01.ID_PERFIL                                      AS ID_PERFIL
     , COUNT(DISTINCT LAST_DAY(M.DATA))                     AS QTD_MESES
  FROM IN_IND_ABS_MET_RESUL M
  JOIN (SELECT DISTINCT
               VAT1.MATRICULA
             , VAT1.ID_PERFIL
             , VAT1.NOME
          FROM IN_IND_ABS_TMP_01 VAT1) TMP01
    ON TMP01.MATRICULA = M.MATRICULA
   AND TMP01.ID_PERFIL = M.ID_PERFIL
 WHERE M.ID_METRICA IN (SELECT DISTINCT ID_METRICA FROM IN_IND_ABS_TMP_00)
   AND M.DATA BETWEEN vDtIni AND vDtFim
 GROUP BY TO_CHAR(M.DATA, 'MM/RR'), M.MATRICULA, TMP01.ID_PERFIL
 );
COMMIT;
```

Fonte: Elaborado pelo autor (2023)

O cálculo do acumulado anual leva em consideração todo o período em que o colaborador recebeu o indicador ao longo do ano, ao contrário do cálculo corrente, que considera apenas o mês atual. É importante destacar que, em caso de troca de perfil durante o ano, o colaborador só recebe o acumulado para o período em que atuou no perfil correspondente ao indicador. Se o funcionário atuar em dois perfis que possuem o mesmo indicador, ele receberá dois acumulados, um para cada perfil.

As temporárias 4 e seguintes seguem o mesmo processo, como demonstrado no fluxograma da Figura 9, tanto para o acumulado quanto para o padrão. Após o truncamento, a temporária 4 é carregada com a união das temporárias 2 e 3. A temporária 5 separa o VALOR REAL (coluna com o resultado do cálculo da métrica) em duas colunas (REAL e META) usando a função DECODE e o parâmetro TIPO (‘R’ para Real e ‘M’ para Meta) a fim de separá los.

A Temporária 6, após ser truncada, contém informações sobre o indicador, incluindo a forma como ele deve ser exibido (como o número de casas decimais) e dados importantes, como o valor que aplicado na meta se obtêm o nível mínimo (o valor aceitável, mas abaixo da meta) e a super meta (o valor acima do esperado). Esses dados são obtidos de uma view externa filtrada pelo indicador passado na função.

A Temporária 7 é responsável por relacionar as tabelas temporária 5 e 6, utilizando os resultados de ambas para calcular o resultado final, como pode ser visto no Código 12.

Código 12. Cláusula SELECT da Temporária 7 para o cálculo do indicador de Absenteísmo.

```
INSERT INTO IN_IND_ABS_TMP_07 (
SELECT IR.ORDER_D                                                    AS ORDER_D
     , IR.DATA                                                       AS DATA
     , DECODE(R.ORIENTACAO, 'MaiorMelhor', 1, 'MenorMelhor', 2, 0)   AS ORIENTACAO
     , ROUND((R.META_MINIMA * IR.META) / 100, R.CASAS_DECIMAIS)      AS NIVELMINIMO
     , ROUND(IR.META, R.CASAS_DECIMAIS)                              AS META
     , ROUND((R.SUPER_META * IR.META) / 100, R.CASAS_DECIMAIS)       AS SUPER_META
     , ROUND(IR.REAL, CASAS_DECIMAIS)                                AS REAL
     , R.NOME                                                        AS NOME
     , R.ID                                                          AS ID_INDICADOR
     , IR.MATRICULA                                                  AS MATRICULA
     , IR.NM_FUNC                                                    AS NM_FUNC
     , CASE
         WHEN R.ORIENTACAO = 'MaiorMelhor' AND
              IR.REAL < (R.META_MINIMA * IR.META) / 100 THEN
          'VERMELHO'
         WHEN R.ORIENTACAO = 'MenorMelhor' AND
              IR.REAL > (R.META_MINIMA * IR.META) / 100 THEN
          'VERMELHO'
         WHEN R.ORIENTACAO = 'MaiorMelhor' AND
              IR.REAL >= (R.META_MINIMA * IR.META) / 100 AND
              IR.REAL < IR.META THEN
          'AMARELO'
         WHEN R.ORIENTACAO = 'MenorMelhor' AND
              IR.REAL <= (R.META_MINIMA * IR.META) / 100 AND
              IR.REAL > IR.META THEN
          'AMARELO'
         WHEN R.ORIENTACAO = 'MaiorMelhor' AND IR.REAL >= IR.META AND
              IR.REAL < (R.SUPER_META * IR.META) / 100 THEN
          'VERDE'
         WHEN R.ORIENTACAO = 'MenorMelhor' AND IR.REAL <= IR.META AND
              IR.REAL > (R.SUPER_META * IR.META) / 100 THEN
          'VERDE'
         WHEN R.ORIENTACAO = 'MaiorMelhor' AND
              IR.REAL >= (R.SUPER_META * IR.META) / 100 THEN
          'AZUL'
         WHEN R.ORIENTACAO = 'MenorMelhor' AND
              IR.REAL <= (R.SUPER_META * IR.META) / 100 THEN
          'AZUL'
       END                                                           AS COR
```

Fonte: Elaborado pelo autor (2023)

A tabela temporária gera a coluna "cor" sendo ela um farol para visualizar o desempenho do indicador de forma rápida. A cor do farol indica se o resultado está abaixo, dentro ou acima das metas estabelecidas: vermelho indica abaixo do nível mínimo, amarelo indica entre o mínimo e a meta, verde indica entre a meta e a super meta, e azul indica acima da super meta. A orientação indica se é melhor que o indicador seja maior ou menor, pois alguns indicadores têm um melhor desempenho quando o resultado é menor, enquanto outros têm um melhor desempenho quando o resultado é maior.

Para inserir novos resultados em um indicador, primeiro é necessário excluir os resultados anteriores da tabela correspondente, resultados estes que serão substituídos (caso existam) pelos novos calculados. Em seguida, é verificado por meio de um operador lógico se o usuário passou uma matrícula ou perfil específico na rotina. Se sim, apenas os resultados relacionados a esses parâmetros são excluídos.

Após a deleção, é realizada a inserção do resultado da procedure de cálculo do indicador na tabela de resultado dos indicadores a partir da temporária 7

relacionada com a tabela informacional dos indicadores (para verificar se o indicador está válido para o ano em que está sendo inserido). Com isso é gerado no log o indicativo que a execução da procedure foi finalizada, terminando assim o processamento do indicador.

4.2.5 Disponibilização dos resultados

Para a visualização dos resultados dos indicadores utilizamos o mesmo site citado anteriormente, no modelo anterior.

4.2.6 Correção de erros

O processo de "Reparação" dos resultados tem como objetivo assegurar que todas as matrículas recebam algum tipo de resultado de um indicador. Isso é alcançado adicionando um valor nulo (não podendo ser 0, devido à possibilidade de indicadores terem esse valor) para as matrículas que inicialmente não obtiveram resultados. Em alguns casos, em vez de adicionar um valor nulo, é incluído um valor padrão que não influencia na média global da empresa, denominado valor de complemento. Essas duas modalidades de inserção correspondem aos tipos "2" e "3" mencionados na procedure de cálculo dos indicadores. Esse procedimento é fundamental para facilitar a manutenção dos indicadores, permitindo a identificação das matrículas que não estão recebendo seus resultados.

É importante ressaltar que a ausência de resultados para algumas matrículas não necessariamente indica um erro no cálculo ou falta de dados relacionados a um colaborador. Existem indicadores que são adicionados apenas no final do mês (ou mesmo meses após), enquanto o cálculo é realizado diariamente, o que explica essa ocorrência frequente. Além disso, situações como férias ou licença médica de um funcionário podem resultar na ausência de resultados, dependendo do indicador em questão.

A reparação dos resultados é executada por meio de uma procedure que recebe os mesmos parâmetros utilizados para o indicador, com exceção do tipo e do novo parâmetro passado na procedure de reparação, que é o indicador que precisa ser "reparado". A procedure analisa os valores resultantes do indicador na tabela de resultados e compara-os com as matrículas que deveriam estar recebendo esses dados. Para os valores ausentes, a procedure insere um valor nulo ou o valor de complemento, com base nas informações da relação entre o indicador e o funcionário na tabela de resultados de indicadores. Adicionalmente, as informações sobre o valor e as metas (nível mínimo, meta e supermeta) são incorporadas à tabela, juntamente com o valor "reparado". Esse processo garante que todas as matrículas sejam atendidas e que os indicadores estejam completos e prontos para análise e avaliação.

5 Resultados

Neste capítulo, são apresentados os resultados obtidos após a execução do processo de geração de indicadores. Isso envolve todas as etapas, desde a integração dos dados até a disponibilização dos resultados, que foram repetidas trinta vezes para cada cenário. Dentro deste contexto, serão realizadas comparações detalhadas entre os dois modelos, focando especialmente no tempo de execução global do processo dos 55 indicadores ativos. Essas análises oferecerão insights fundamentais sobre a eficiência e o impacto de cada cenário, orientando as conclusões finais deste estudo.

5.1 Tempo de execução

Sobre o tempo de execução de cada processo, ambos foram executados 30 vezes, sendo mostrados na Quadro 1, onde a coluna “FILIAIS (h)” representa o modelo descentralizado nas filiais e a coluna “DW (h)” representa o modelo centralizado no Data Warehouse (DW).

Quadro 1. Resultados das trinta execuções para cada modelo.

EXECUÇÃO	FILIAIS (h)	DW (h)
1º	16:12	06:07
2º	16:12	06:07
3º	16:13	06:08
4º	16:16	06:13
5º	16:19	06:16
6º	16:55	06:21
7º	17:00	06:29
8º	17:22	06:31
9º	17:33	06:39
10º	17:37	06:41
11º	17:39	06:43
12º	17:50	06:48
13º	17:59	06:59
14º	18:24	07:01
15º	18:57	07:04
16º	19:43	07:06
17º	20:20	07:07
18º	20:25	07:07
19º	20:37	07:13
20º	20:44	07:13
21º	20:57	07:16
22º	20:57	07:17
23º	21:07	07:31
24º	21:35	07:35
25º	21:40	07:37
26º	21:55	07:41
27º	22:10	07:42
28º	22:30	07:48
29º	22:33	07:50
30º	22:55	07:51

Fonte: Elaborado pelo autor (2023)

Após as trinta iterações de cada processo, foram criados sumários estatísticos, conforme apresentado no Quadro 2. Esses sumários incluem métricas como média, mediana, moda, variância e desvio padrão para cada cenário. Além disso, foram produzidos gráficos de boxplot e histograma para ambos os casos (Gráficos 1-4).

Quadro 2. Sumário estatístico obtido a partir das execuções de cada modelo.

	FILIAIS (h)	DW (h)
Média	19:17	07:00
Desvio Padrão	02:16	00:33
Variância	00:13	00:01
Coeficiente de Variação	11,78%	7,97%
Mediana	19:20	07:05
Moda	16:12	06:07
	20:57	07:07
		07:13

Fonte: Elaborado pelo autor (2023)

A partir dos resultados obtidos no Quadro 2, é possível inferir que o modelo de data warehouse demonstrou um desempenho notavelmente superior em termos de tempo de execução em comparação com o modelo descentralizado em filiais.

O modelo descentralizado em filiais apresenta uma média de tempo de execução de 19 horas e 17 minutos, com uma dispersão relativamente alta, conforme indicado pelo desvio padrão de 2 horas e 16 minutos e no coeficiente de variação de mais de 11%. A mediana e a moda estão próximas à média, indicando uma distribuição de tempo de execução relativamente simétrica em torno desse valor central.

Já para o modelo de data warehouse, a média de tempo de execução é significativamente menor, totalizando 7 horas. Além disso, o desvio padrão é de apenas 33 minutos, indicando uma dispersão muito menor em relação à média. A variância é extremamente baixa, com apenas 1 minuto, sugerindo que os tempos de execução estão altamente concentrados em torno da média. A mediana, que é de 7 horas e 5 minutos, e a moda, que varia entre 6 horas e 7 minutos a 7 horas e 13 minutos, também estão próximas da média, indicando uma distribuição mais simétrica e consistente dos tempos de execução.

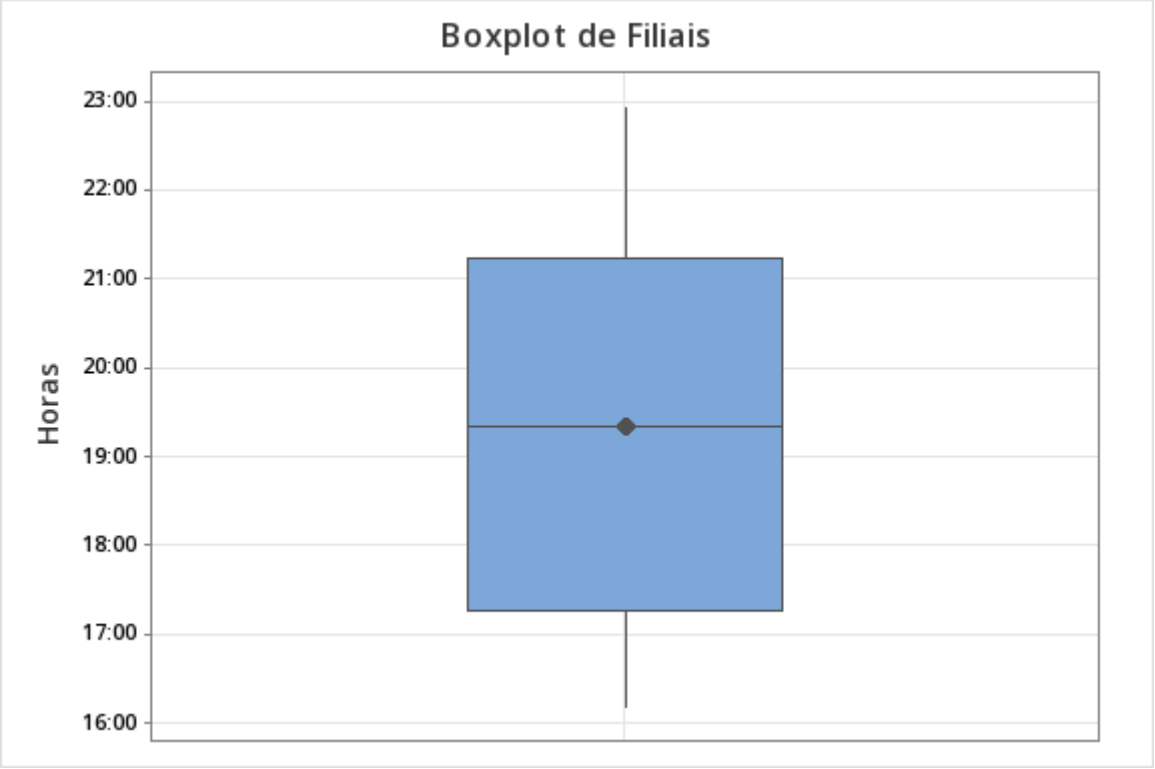
Esses resultados sugerem que o modelo de data warehouse é mais eficiente em termos de tempo de execução, oferecendo uma maior consistência e previsibilidade em comparação com o modelo descentralizado. A redução significativa no tempo médio de execução no modelo de data warehouse pode ter implicações importantes em termos de eficiência operacional e capacidade de

resposta aos processos de geração de indicadores, tornando-o uma escolha mais vantajosa do ponto de vista de desempenho.

No gráfico de boxplot para o modelo descentralizado em filiais (Gráfico 1), podemos observar que o tempo de execução mediano (linha no meio da caixa) está em torno de 19:20. Não há valores atípicos (outliers) além das hastes, sugerindo que os tempos de execução permaneceram dentro de um intervalo relativamente consistente.

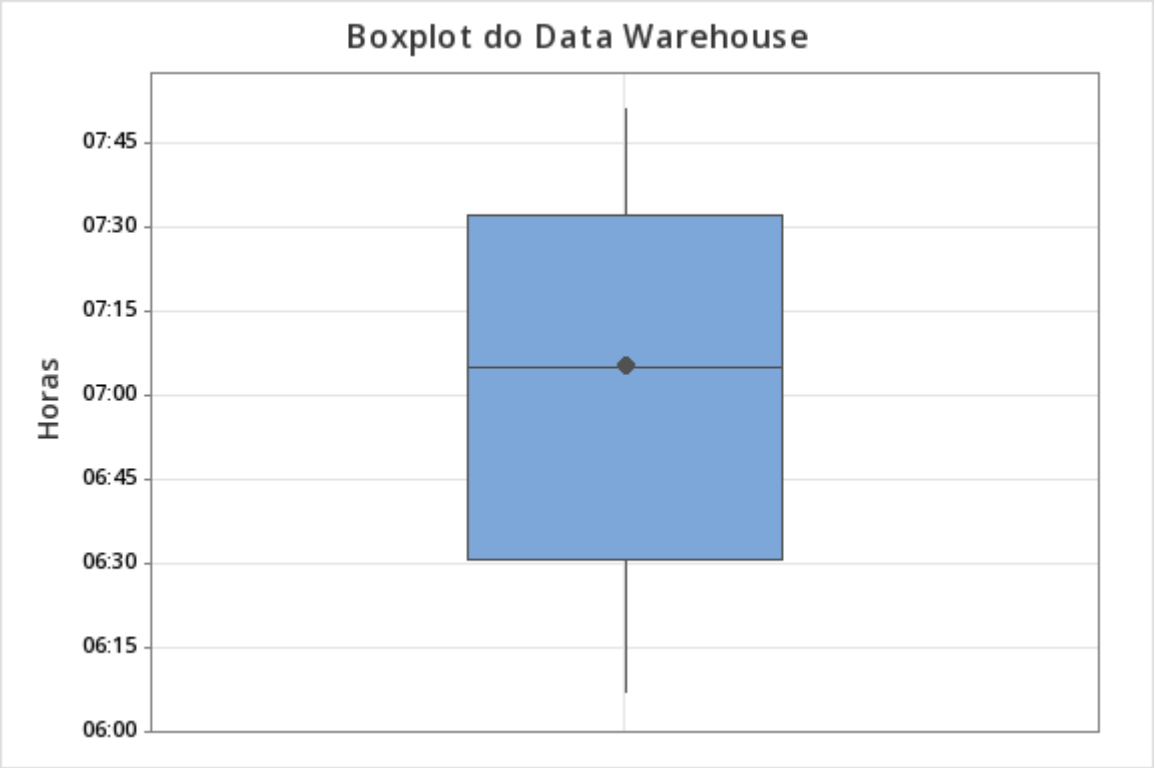
Já no gráfico de Boxplot para o Modelo Centralizado em um Data Warehouse (Gráfico 2), o tempo de execução mediano (linha no meio da caixa) está em torno de 07:07. A dispersão dos dados é menor em comparação com o modelo descentralizado, indicado pelo comprimento mais curto da caixa. Não há valores atípicos (outliers) além das hastes, sugerindo que os tempos de execução permaneceram dentro de um intervalo relativamente consistente.

Gráfico 1. Boxplot do modelo descentralizado em filiais.



Fonte: Elaborado pelo autor (2023)

Gráfico 2. Boxplot do modelo centralizado em Data Warehouse.



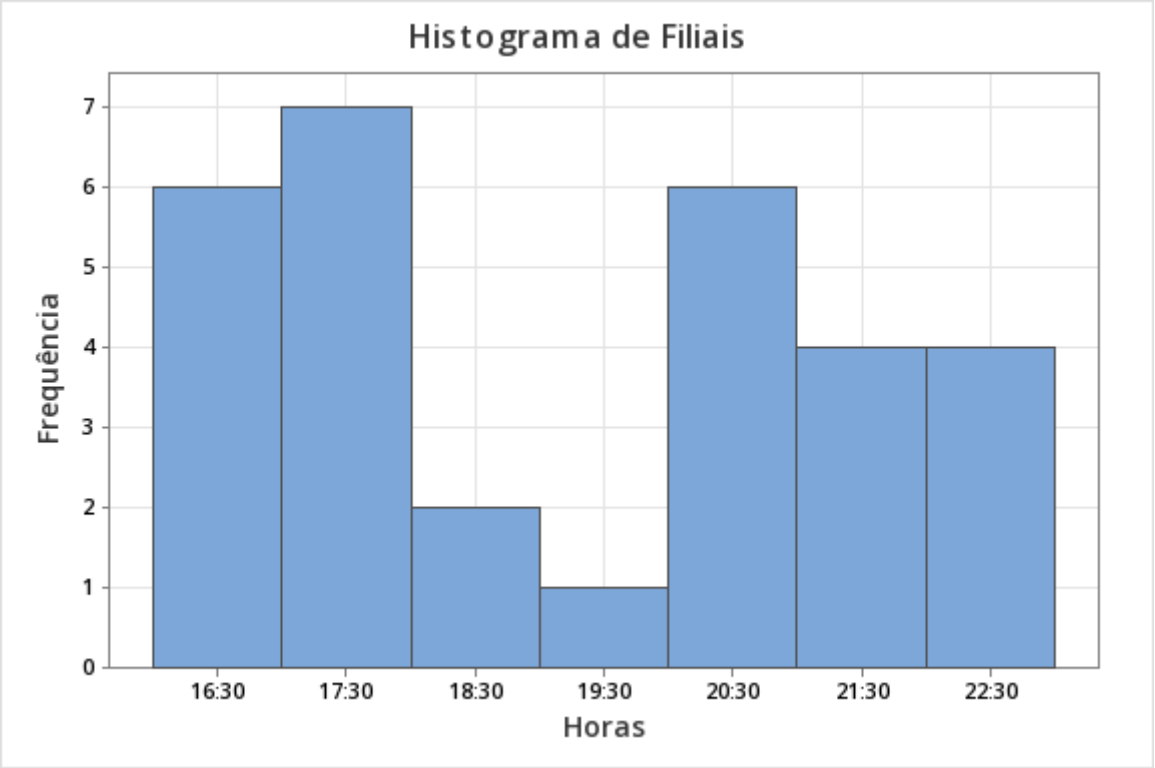
Fonte: Elaborado pelo autor (2023)

Os gráficos de histograma são divididos em intervalos (ou bins) no eixo horizontal e mostram a frequência com que os tempos de execução caem dentro de cada intervalo. Cada barra vertical no histograma representa um intervalo específico de tempos de execução e a altura da barra indica a frequência com que os tempos de execução caem nesse intervalo.

Analisando os histogramas, é possível inferir que, para o histograma do modelo descentralizado em filiais (Gráfico 3), os tempos de execução variam significativamente. Existem intervalos onde os tempos de execução são mais frequentes (as barras são mais altas), indicando que algumas execuções foram mais rápidas, enquanto outros intervalos têm menos frequência, sugerindo execuções mais lentas. A forma geral do histograma é assimétrica, isso mostra que as interações têm uma grande variação nos tempos de execução.

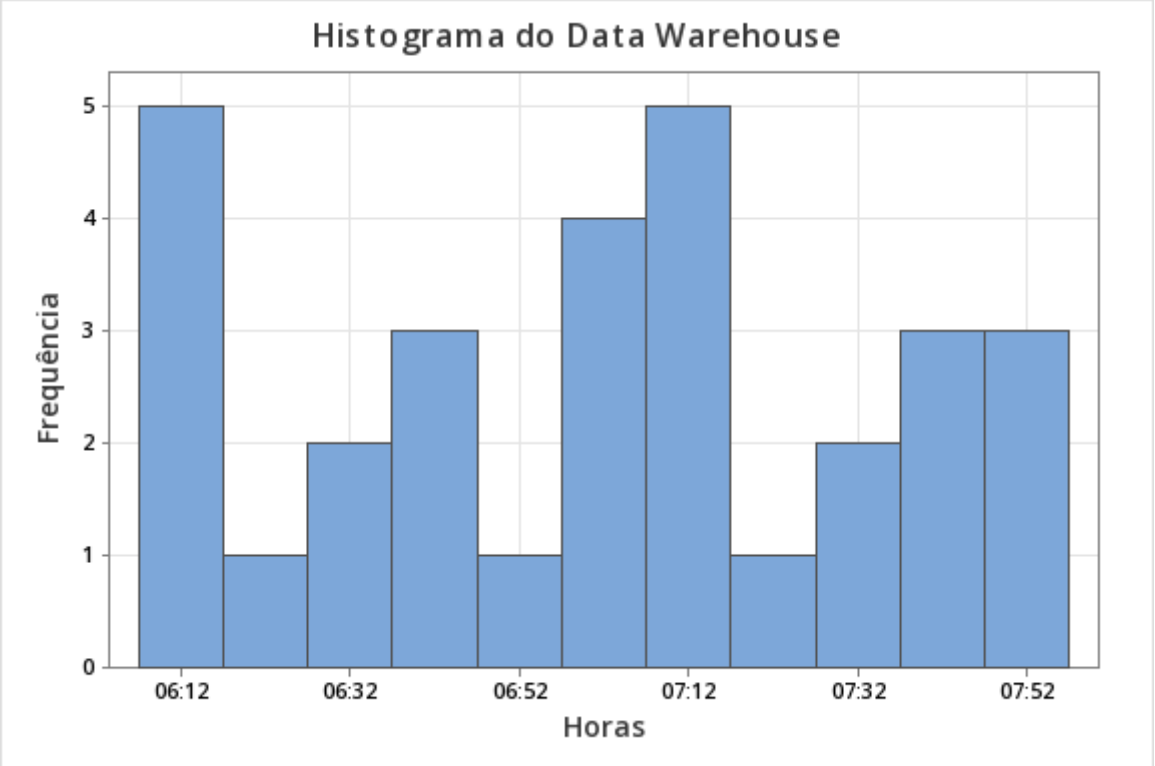
Para o histograma para o modelo centralizado em um data warehouse (Gráfico 4), podemos observar que os tempos de execução estão concentrados em uma faixa mais estreita de valores. Isso é indicado pelo fato de que a maioria das barras no histograma são próximas umas das outras. A forma do histograma é mais simétrica em comparação com o modelo descentralizado. Isso sugere que os tempos de execução são mais consistentes, com menos variação entre as diferentes iterações. Não há valores atípicos evidentes à direita ou à esquerda do histograma, indicando que os tempos de execução permaneceram dentro de um intervalo relativamente consistente.

Gráfico 3. Histograma do modelo descentralizado em filiais.



Fonte: Elaborado pelo autor (2023)

Gráfico 4. Histograma do modelo centralizado em Data Warehouse.



Fonte: Elaborado pelo autor (2023)

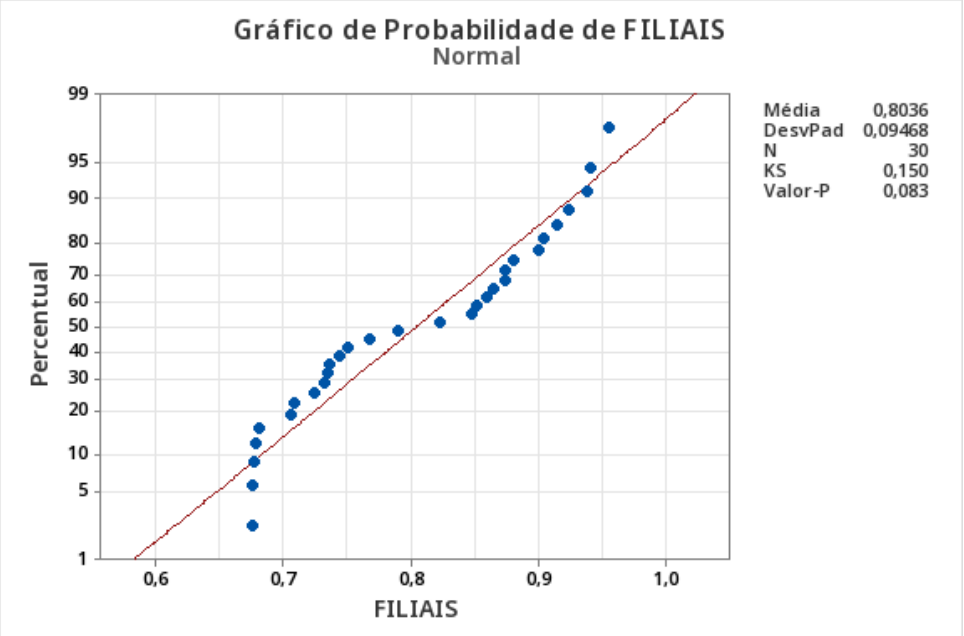
Em resumo, os histogramas e gráficos de boxplot oferecem uma representação visual e clara da distribuição dos tempos de execução para ambos os modelos. No modelo descentralizado, observa-se uma ampla variação nos tempos de execução, com algumas execuções mais rápidas e outras mais lentas. Em contraste, o modelo centralizado exibe tempos de execução mais consistentes, concentrados em uma faixa mais estreita de valores, o que indica uma maior previsibilidade no processo de geração de indicadores. Esses resultados sugerem que o modelo centralizado é mais eficiente em termos de tempo de execução, proporcionando uma operação mais consistente e estável em comparação com o modelo descentralizado.

5.2 Teste de Hipótese

Neste estudo, foi aplicado um teste de hipótese T pareado para avaliar as diferenças estatisticamente significativas nos tempos de execução entre dois modelos de geração de KPIs: o modelo descentralizado (representado pela amostra "FILIAIS") e o modelo centralizado (representado pela amostra "DW") (Quadro 1).

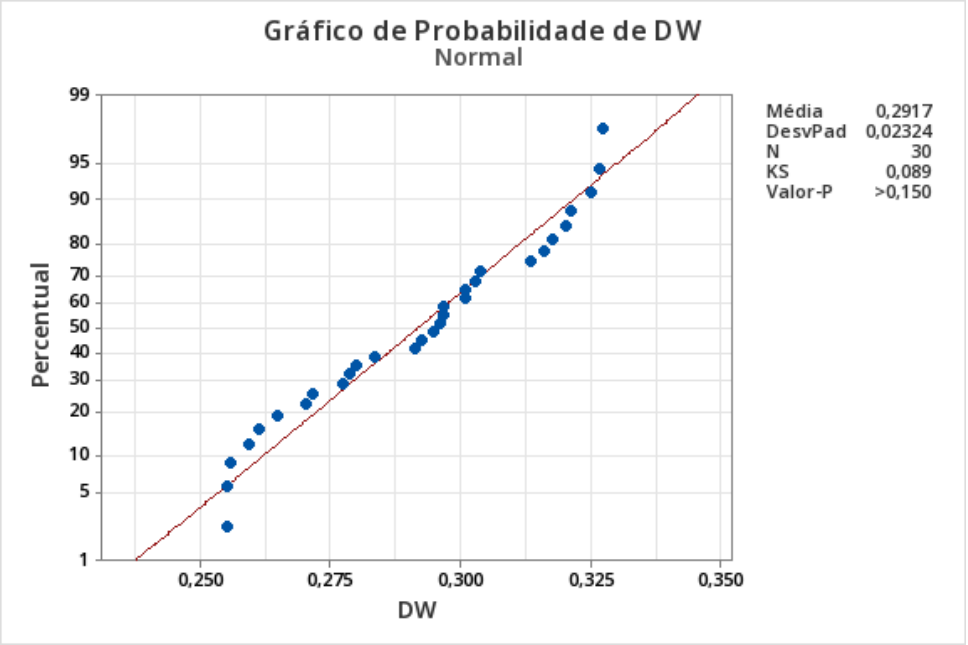
Foram realizados testes de normalidade para cada amostra (Figuras 10 e 11) e ambas apresentaram um valor_p maior que 0,05 (alfa), o que nos indica que é possível considerar as amostras como próximas a uma normal.

Figura 10. Resultados do teste de normalidade para Filiais.



Fonte: Elaborado pelo autor (2023)

Figura 11. Resultados do teste de normalidade para DW.



Fonte: Elaborado pelo autor (2023)

O teste de hipótese T pareado é uma ferramenta estatística utilizada para determinar se existe uma diferença significativa entre as médias de duas amostras relacionadas, ou seja, quando os dados são coletados de pares de observações em circunstâncias semelhantes. Ele é especialmente útil quando se deseja comparar o desempenho de dois métodos ou modelos diferentes.

No contexto deste estudo, a hipótese nula (H_0) considerada foi que não existe diferença significativa nos tempos de execução médios entre os modelos descentralizado e centralizado. A hipótese alternativa (H_1) foi que existe uma diferença significativa nos tempos de execução médios entre os dois modelos. Após realizar trinta execuções para cada modelo e calcular os tempos de execução, foram obtidos os seguintes resultados:

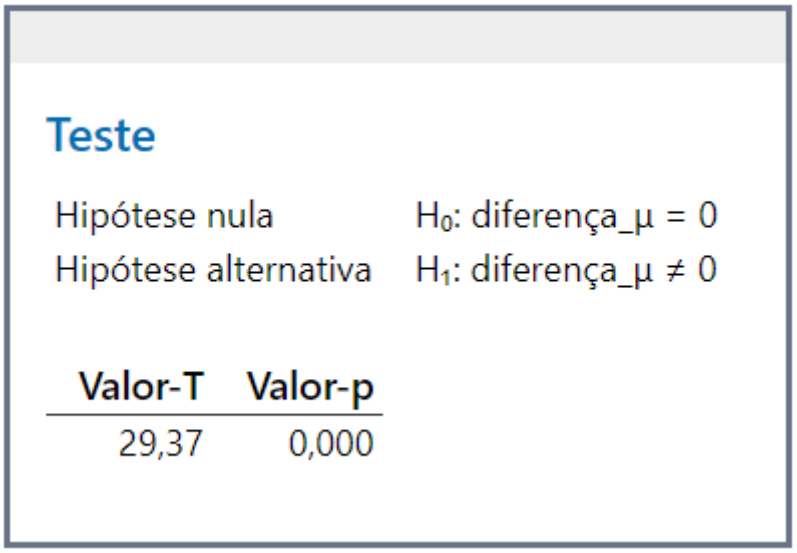
Para o modelo descentralizado (FILIAIS), a média dos tempos de execução foi de 19 horas e 17 minutos. Para o modelo centralizado (DW), a média dos tempos de execução foi notavelmente menor, registrando 7 horas e 0 minutos.

Além disso, o teste revelou um valor-p extremamente baixo, igual a zero (Figura 12). Isso significa que há uma diferença estatisticamente significativa nos tempos de execução entre os modelos descentralizado e centralizado. Com base nos resultados, podemos rejeitar a hipótese nula (H_0) e concluir que o modelo centralizado é mais eficiente em termos de tempo de execução do que o modelo descentralizado.

Essa análise estatística fornece uma base sólida para a tomada de decisões informadas sobre a escolha do modelo mais eficaz na geração de KPIs. Os resultados indicam que o modelo centralizado é estatisticamente superior em termos de tempo de execução, o que pode ter implicações significativas para a eficiência

operacional e o desempenho da empresa em um ambiente de negócios cada vez mais competitivo.

Figura 12. Resultados do teste de hipótese T pareado.



Teste

Hipótese nula	H_0: diferença_μ = 0
Hipótese alternativa	H_1: diferença_μ ≠ 0

Valor-T	Valor-p
29,37	0,000

Fonte: Elaborado pelo autor (2023)

6 Conclusão

Peterson et al. (2006) expõem uma série de benefícios e perspectivas sobre a definição de indicadores de desempenho, destacando que o propósito fundamental desses indicadores é a capacidade de sintetizar dados empresariais de maneira significativa. Eles servem para efetuar comparações eficazes, economizar tempo e fornecer apoio à gestão na tomada de decisões.

Tendo isso em vista, o presente estudo abordou a transição da geração de Indicadores-Chave de Desempenho (KPIs) de um modelo descentralizado para um modelo centralizado em um Data Warehouse (DW) em uma empresa de varejo. A motivação para esta transição foi impulsionada por desafios encontrados no modelo descentralizado, incluindo alto tempo de execução, falta de controle do processo e dificuldade na correção de erros devido à complexidade do sistema.

A metodologia empregada envolveu a coleta de dados, preparação de dados, geração de métricas, centralização das métricas e finalmente a geração dos KPIs. Este processo permitiu uma avaliação abrangente das métricas e indicadores associados aos processos de negócios da empresa.

Os resultados obtidos por meio de análises estatísticas, incluindo um teste de hipótese T pareado, revelaram que o modelo centralizado em um Data Warehouse oferece uma significativa melhoria no tempo de execução em comparação com o modelo descentralizado. A diferença média nos tempos de execução entre os dois modelos foi estatisticamente significativa, indicando claramente a superioridade do modelo centralizado em termos de eficiência de tempo.

Além disso, a análise dos resultados mostrou que o modelo centralizado proporciona uma maior consistência nos tempos de execução, com menor variação, tornando-o mais previsível e confiável em comparação com o modelo descentralizado.

Essas descobertas têm implicações significativas para a empresa de varejo. A transição para um modelo centralizado em um Data Warehouse pode resultar em economia de tempo significativa, melhorando a eficiência operacional e permitindo uma tomada de decisão mais ágil e informada. Além disso, a maior previsibilidade nos tempos de execução pode melhorar a confiabilidade dos processos de geração de KPIs.

Em resumo, este estudo demonstra a importância da análise de processos de geração de indicadores e a vantagem de considerar a transição para um modelo centralizado em um Data Warehouse. A eficiência aprimorada e a consistência nos resultados podem contribuir para o sucesso contínuo da empresa no mercado competitivo de varejo, fornecendo informações de qualidade de forma mais rápida e eficaz para apoiar a tomada de decisões estratégicas.

7. Referências

[1] BADAWY, M. et al. A survey on exploring key performance indicators. Future Computing and Informatics Journal, v. 1, n. 1, p. 47–52, 2016. ISSN 2314-7288. Available at: <https://www.sciencedirect.com/science/article/pii/S2314728816300034>.

[2] CHAUDHURI, S.; DAYAL, U. An overview of data warehousing and olap technology. SIGMOD Rec., Association for Computing Machinery, New York, NY, USA, v. 26, n. 1, p. 65–74, mar 1997. ISSN 0163-5808. Available at: <https://doi.org/10.1145/248603.248616>.

[3] FERNANDES, H. M. D. S. Melhoria dos Processos de um Armazém com Base na Medição e Padronização dos KPI. Dissertação (Mestrado), 2023. Available at: <https://hdl.handle.net/1822/84675>.

[4] Yang Gang, Tian Qing, Zhou Xingshe, Wang Tao, Li Wei Chao e Guan Tao. Rule engine based kpi (key performance indicator) generation method in business activity monitoring. 2012.

[5] JESUS, D. S. d.; LAURINDO, R. P. O uso de kpi's no processo logístico em uma distribuidora de medicamentos. 2019. Available at: <http://hdl.handle.net/123456789/1999>.

[6] LI, P.; DAI, C.; WANG, W. Inconsistent data cleaning based on the maximum dependency set and attribute correlation. Symmetry, v. 10, n. 10, p. 516, 2018. Available at: <https://doi.org/10.3390/sym10100516>.

[7] Oracle Corporation. Oracle Database 19c Documentation - CREATE DATABASE LINK. 2019. Accessed on August 26, 2023. Available at: <https://docs.oracle.com/en/database/oracle/oracle-database/19/sqlrf/CREATEDATABASE-LINK.htmlGUID-D966642A-B19E-449D-9968-1121AF06D793>.

[8] PETERSON, E. T. The Big Book of Key Performance Indicators. [S.l.]: Web analytics demystified, 2006.

[9] RODRIGUES, A. C.; CANELADA, M. Utilização de kpi – indicadores de desempenho na cadeia de suprimentos. um estudo de caso em indústria metalúrgica no setor da construção civil. 2015. Available at: <http://hdl.handle.net/11077/1418>.

[10] SILVA, A. F. d. Indicadores de desempenho: Estudo de caso na empresa net serviços. 2013. Available at: <https://repositorio.ufpb.br/jspui/handle/123456789/1419>.

[11] SILVA, J. R. L. d. Monotorização de KPI: Análise de Resultados e Proposta de Novos Indicadores - Estudo de Caso: Garcia Garcia, SA. Dissertação (Mestrado), 2021.Available at: <http://hdl.handle.net/10400.14/35163>.