

DESCOBERTA DE CONHECIMENTO EM BASE DE DADOS ORGANIZACIONAIS: ESTUDO DE CASO LOJA DE VENDA DE ELETRÔNICOS

ORGANIZATIONAL DATA KNOWLEDGE DISCOVERY: CASE STUDY ELECTRONICS SALE SHOP

Vanessa Coelho Ribeiro,
Renato José da Silva Camões,
Alysson da Rocha Ribeiro

RESUMO

Estar à frente no mercado é um dos desafios da maioria das empresas, mas poucas conhecem que os insumos necessários para alcançar esse objetivo circulam diariamente em seu âmbito. Com as técnicas e ferramentas de *Business Intelligence* corretamente aplicadas sobre os dados da empresa é possível identificar os requisitos necessários para oferecer serviços/produtos de qualidade tornando-se uma referência no mercado. O projeto de *Business Intelligence* consiste em coletar os dados, tratar, armazenar e analisar, dando assim, aos gestores, suporte a tomada de decisão. O presente artigo visa demonstrar as principais técnicas para estruturação de um projeto de BI, as principais ferramentas e suas aplicações colocando em prática todos os conceitos extraídos de livros e artigos abordados nesta pesquisa. Além disso, estimar o custo para implementação do projeto.

Palavras-chave: Business Intelligence, análise de dados, SQL server, Integration Services, Power Bi.

ABSTRACT

Being ahead of the market is one of the challenges of most companies, but few are aware that the inputs needed to achieve this goal are daily in scope. With the techniques and tools of Business Intelligence correctly applied on the data of the company it is possible to identify the necessary requirements to offer services / products of quality becoming a benchmark in the market. The Business Intelligence project consists of collecting data, treating, storing and analyzing, thus giving managers support for decision making. This article aims to demonstrate the main techniques for structuring a BI project, the main tools and their applications putting into practice all the concepts extracted from books and articles addressed in this research. In addition, estimate the cost to implement the project.

Keywords: Business Intelligence, Data Analysis, SQL Server, Integration Services, Power Bi.

INTRODUÇÃO

Silva (2016) em 1990 já afirmava que a tecnologia era pouco disseminada e cara. O preço de um computador pessoal era altíssimo, e todo o comércio varejista usava papéis e máquinas registradoras. Com o avanço e disseminação das tecnologias os preços dos computadores ficaram mais acessíveis e concomitante à evolução dos processos organizacionais contribuíram para que os sistemas de informações se tornasse a principal ferramenta para o gerenciamento de uma organização.

Um sistema de informação pode ser definido tecnicamente como um conjunto de componentes inter-relacionados que coletam ou recuperam, processam, armazenam e distribuem informações destinadas a apoiar a tomada de decisões, a coordenação e ao controle de uma organização. (LAUDON, 2007)

Observa-se que ao longo do tempo, o alto investimento nesses sistemas de captação e armazenamento de informações não era suficiente para obtenção dos diferenciais estratégicos desejados. A maioria das organizações que implantaram esses sistemas informacionais viram que apenas armazenar grande quantidade de dados de nada valia se essas informações se encontravam repetidas, incompletas e espalhadas em vários sistemas dentro da corporação. (PRIMAK 2008)

Desta forma, os sistemas de informação começaram a ser importantes não apenas no recebimento de dados externos provenientes de transações realizadas nas organizações, para processar os dados recebidos, armazenar as informações geradas e exibir os resultados de forma mais organizada. Por meio dessa organização de dados as empresas perceberam que ele é um fato, um valor documentado ou um resultado de medição. Quando um sentido semântico ou um significado é atribuído aos dados, gera-se informação. Quando estes significados se tornam familiares, ou seja, quando um agente os aprende, este se torna consciente e capaz de tomar decisões a partir deles, e surge o conhecimento. (SILVA, 2016)

Por isso, os dados de uma organização são um dos seus ativos mais preciosos que diariamente circulam pela empresa, e no fim, acabam armazenados em algum lugar: seja coletado por S.I. e armazenado em uma base de dados, transcrito em uma folha de papel ou apenas gerando conhecimento tácito sobre o assunto. Muitas dessas empresas não dimensionam o quão preciosos são os dados e informações armazenados de suas transações e acabam por não utilizar seu máximo potencial, gerando custos tecnológicos de armazenamento.

Uma metodologia em evidência, atualmente, é o *Business Intelligence* (BI), entendido como um conjunto de tecnologias que permite o cruzamento de informações e suporte a análise dos indicadores de desempenho de um negócio. Dessa forma, as ferramentas de apoio a decisão que fazem inferências em um banco de dados histórico são também chamadas de ferramentas de BI. (MORAIS, 2010)

Uma grande vantagem da utilização de BI é a identificação dos pontos fracos da empresa podendo-se, posteriormente, aplicar as melhorias necessárias, aumentando assim a qualidade dos produtos ou serviços prestados e obtendo uma diminuição de custo, em virtude de identificar quais setores da empresa necessitam de maior investimento, qual o perfil de compra dos clientes, quais processos de negócio necessitam de modificações e etc. Tais ações geram grandes vantagens competitivas e a empresa mais atraente e com melhor posicionamento de mercado.

Este artigo apresenta um projeto que visa gerar conhecimento sobre os dados de uma empresa de vendas de eletrônicos, em que serão abordadas técnicas e ferramentas de *Business Intelligence* (BI). Para isso será necessário desenvolver um processo de extração e tratamento de dados (ETL), implementar um ambiente de armazenamento de dados históricos (DW) e desenvolver uma área de apresentação dos dados para o usuário final (OLAP).

CONCEITO DE BUSINESS INTELLIGENCE

O termo *Business Intelligence*, surgiu na década de 80 no Gartner-Group e faz referência ao processo inteligente de coleta, organização, análise, compartilhamento e monitoração de dados contidos em Data Warehouse/Data Mart, gerando informações para o suporte a tomada de decisões no ambiente de negócios. (PRIMAK, 2008)

Business Intelligence (BI), segundo Gartner (2018), é um termo abrangente que inclui aplicações, infraestrutura e ferramentas e as melhores práticas que capacitam o acesso e a análise da informação, com o intuito de melhorar e otimizar decisões de desempenho.

Já para Silva (2016), inteligência de Negócios é o nome que se atribui a iniciativas para analisar informações complexas de um contexto organizacional e de apresentar o resultado de tal análise de maneira sintetizada e simples, que facilite a interpretação pelo(s) gestor(es).

Na visão de Silberschatz (2006), os sistemas de apoio às decisões visam obter informações de alto nível a partir de informações detalhadas armazenadas nos sistemas de processamento de transação, além de usar informações de alto nível para tomar uma série de decisões.

Dessa forma, as técnicas de *Business Intelligence* visam desenvolver um conhecimento através da integração de diversas fontes de dados, possibilitando a análise de dados brutos transformando-os em conhecimento para que venha auxiliar gerentes e diretores a tomar decisões mais assertivas que consequentemente trará ganho em diversas áreas da empresa. Para que isso seja possível, o sistema de BI divide-se em quatro componentes fundamentais: sistemas de origem (OLTP), tratamento de dados (ETL), armazenamento de dados históricos (*Data Warehouse*) e apresentação dos dados (OLAP).

EXTRAÇÃO, TRANSFORMAÇÃO E CARGA DE DADOS – ETL

A ETL é uma das principais técnicas que compõem o sistema de BI. Essa técnica consiste em realizar a integração de diferentes fontes de dados, tratando-os de forma a garantir a disponibilidade e integridade dos dados que serão carregados no armazém de dados.

Do ponto de vista Gonçalves (2012), é uma técnica ou processo de banco de dados em que os dados das mais diversas origens são movidas para outros locais. Já para Silva (2016), é a transformação de dados faz-se necessária quando há a necessidade de analisar variáveis com domínios de natureza e grandezas diferentes. Isto é, após a extração dos dados dos sistemas legados que, na maioria das vezes apresentam pequenas inconsistências, passam a ser tratados, de modo que estejam

aptos para serem apresentados em consultas gerenciais, em seguida são carregados no armazém de dados, também denominado *Data Warehouse*.

DATA WAREHOUSE

Primak (2008) afirma que, na década de 90 surgiu o *Data Warehouse* que é uma grande base de dados, ou seja, um repositório único de dados considerado pelos especialistas da área de informática como uma peça essencial para a execução da prática de um projeto de *Business Intelligence*. Já Gonçalves (2012), o conceito de *data Warehouse* define-se como um repositório de uma empresa ou organização, construído para fins de relatórios e análise de negócio.

Visto que, um depósito de dados- *Data Warehouse* é um repositório (ou arquivo) de informações colhidas de várias origens, armazenadas sob um esquema unificado, em um único local. (SILBERSCHATZ, 2006). Para atingir esse objetivo essa metodologia utiliza a modelagem dimensional que tem como seus dois principais esquemas: modelo estrela (*Star schemas*) e floco de neve (*Snow flake*).

MODELO DIMENSIONAL

O modelo de dados relacional criado pelo Dr. Codd, sempre foi eficiente para os sistemas transacionais em que uns dos objetivos principais eram garantir a integridade e disponibilidade dos dados, evitando-se a redundância e garantindo o maior desempenho, que para isso, era/é aplicada uma rigorosa normalização, mas quando se trata de consultas sobre grande volume de dados, dados históricos, esse modelo é incapaz garantir a eficiência. Segundo Barbieri (2011), com o desenvolvimento de outras necessidades, alavancadas por aspectos de competitividade e busca de diferenciais de negócio e conseqüentemente tomada de decisão, esses modelos se mostraram inadequados.

Em outras palavras, na visão de Kimbal (2002), o modelo de dados dimensional consiste em dois componentes principais: Tabela fato e tabela dimensão. Uma vez que, a tabela fato armazena apenas os dados quantitativos. Podemos pensar que todo dado que pode ser contando, agrupado, sumarizado ou que possa ser utilizado em algum cálculo matemático estará armazenado nessa tabela. Já as tabelas dimensões são as tabelas de atributos, pois guardam dados que caracterizam o fato ocorrido. O fato é um acontecimento do mundo real e as dimensões são as características que compõem esse acontecimento, por exemplo: a venda de um produto é um fato e tem como dimensão o cliente, o produto, o estabelecimento e seu fornecedor.

Portanto, normalização do esquema não é obrigatória, podendo haver dados redundantes no modelo *Star Schemas* fazendo com que o espaço em disco utilizado aumente. Entretanto, para resolver esse problema utiliza-se o modelo *Snow flake* em que as dimensões são normalizadas, criando dimensões de dimensões, isso faz com que DER (Diagrama entidade relacionamento) fique parecido com um floco de neve dando origem ao seu nome. Uma desvantagem desse modelo é que por ser normalizado o desempenho das consultas venha a diminuir, o que pode se tornar um grande problema dependendo do cenário.

DATA MART

Na visão de Kimbal (2002), um *Data Mart* é uma parte do todo que compõe a área de apresentação. Em sua forma mais simples, um *Data Mart* representa dados de um único processo de negócio. Portanto, *Data Mart* são setores do *Data Warehouse* que são classificados por assuntos, visando atender a uma determinada área do negócio ou departamento. Em uma visão técnica, são diferentes schemas que cada um representa uma área de negócio da empresa. Esses schemas são compostos por diferentes tabelas de fatos e dimensões e que juntos formam o armazém de dados, também chamado de *Data Warehouse*.

ONLINE ANALYTICAL PROCESSING – OLAP

As ferramentas OLAP possibilitam com que o usuário da área de negócio possa manipular os dados armazenados de diferentes formas. Isto é, as ferramentas OLAP proporcionam condições de análise de dados on-line necessárias para responder as possíveis perguntas dos analistas, gerentes e executivos (MACHADO, 2004).

São sistemas direcionados para gestores que precisam de informações resumidas e de alto nível sobre os negócios de uma organização. Tais sistemas operam sobre bases de dados históricas e multidimensionais, comumente gerenciadas por sistemas gerenciadores de banco de dados relacionais integrados com ferramentas específicas para carga e análise de dados multidimensionais. (SILVA, 2016)

Além disso, essas ferramentas devem ser capazes de analisar um grande volume de dados em múltiplas perspectivas fazendo com que o usuário tenha uma visão mais ampla do negócio.

METODOLOGIA

A metodologia utilizada é um estudo de caso baseado em uma loja de vendas eletrônicos que atua há 18 anos no mercado realizando vendas presenciais em suas lojas físicas com o intermédio de um vendedor.

O departamento abordado no artigo será o de vendas. As análises realizadas são: Total de comissões pagas para os vendedores; Identificação dos estados com maiores vendas; Períodos do ano em que ocorreram a maior e a menor quantidade de vendas; Perfil de clientes com maiores quantidades de compras.

A obtenção dos dados foi feita através de uma extração da base de dados do sistema transacional – OLTP da empresa, responsável por controlar e gerenciar toda a entrada e saída de produtos, e disponibilizados em um arquivo do tipo CSV para aplicação das técnicas necessárias para o carregamento no *Data Mart* de vendas.

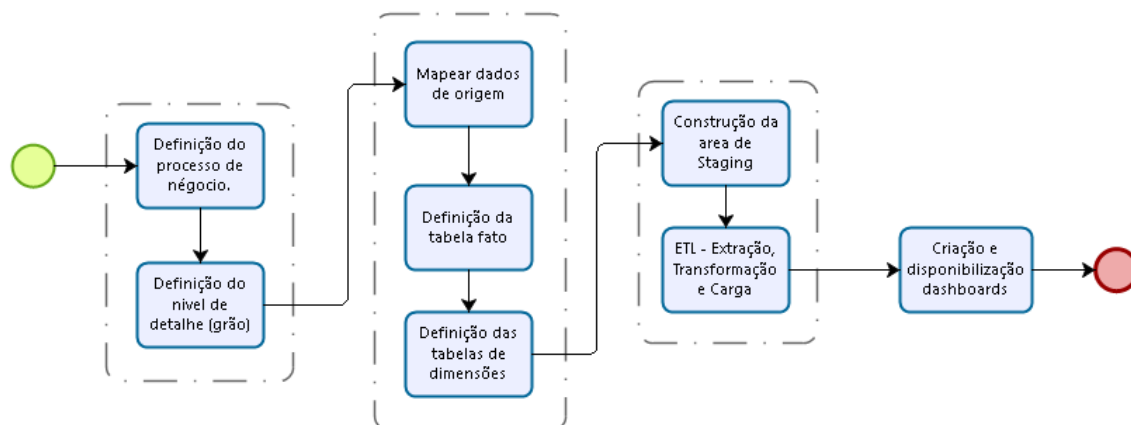
Para efetuar o tratamento dos dados, foi utilizado como servidor uma máquina com sistema operacional Windows 10, processador i7 8ª geração, 16GB de memória RAM e 1TB de HD. Nesta máquina foram instalados Microsoft SQL Server 2017 Developer, brModelo para modelagem do *Data Warehouse*, Microsoft Integration Services para a extração dos dados de origem, transformação carga no *Data Warehouse* e para visualização dos dados e criação dos dashboards será utilizado o Power Bi Desktop.

Tabela 1: Conceito e aplicação das ferramentas no estudo de caso.

Ferramentas	Descrição	Utilização
Microsoft SQL Server	<p>O Microsoft SQL Server é um sistema de gerenciamento e análise de banco de dados para soluções de comércio eletrônico, linha de negócios e Data Warehouse. (MICROSOFT, 2018)</p> <p>Acrescentando-se que, um sistema gerenciado de banco de dados é uma coleção de objetos inter-relacionados e programas, em que seu principal objetivo é recuperar informações sempre que lhe for solicitado. (SILBERSCHATZ, 2006)</p>	Gerenciamento do Data Warehouse
brModelo	<p>Ferramenta de código aberto e totalmente gratuita voltada para ensino de modelagem de banco de dados relacionais com base na metodologia defendida por Carlos A. Heuser no livro “Projeto de Banco de Dados”. Esta ferramenta foi concebida como trabalho de conclusão do curso de especialização em banco de dados pelas universidades UFSC (SC) e UNIVAG (MT), orientado pelo Professor Dr. Ronaldo dos Santos Mello, após se constatar a inexistência de uma ferramenta nacional que pudesse ser utilizada para essa finalidade. (CÂNDIDO, 2017)</p>	Desenvolvimento do modelo dimensional.
Microsoft Integration Services	<p>É um ambiente totalmente integrado ao Visual Studio, utilizando a interface e o conceito do Visual Studio para apresentar as ferramentas de extração, transformação e carga. (GONÇALVES, 2012)</p>	Para o processo de ETL – Extração, Transformação e Limpeza dos dados de origem.
PowerBI Desktop	<p>O Power BI é um conjunto de ferramentas de análise de negócios para analisar dados e compartilhar ideias. Os painéis do Power BI fornecem uma visão de 360 graus para os usuários corporativos com suas métricas mais importantes em um só lugar, atualizadas em tempo real e disponíveis em todos os seus dispositivos. (MICROSOFT, 2018)</p> <p>Além disso, o Power BI Desktop permite criar uma coleção de consultas, conexões de dados e relatórios que podem ser facilmente compartilhados com outras pessoas. O Power BI Desktop integra tecnologias comprovadas da Microsoft – o poderoso mecanismo de Consulta, visualizações e modelagem de dados – e</p>	Ferramenta de OLAP para o desenvolvimento de visualizações interativas que possibilita a análise em diversos cenários.

funciona perfeitamente com o serviço do Power BI online. (MICROSORFT, 2018).
--

O estudo de caso a seguir está dividido em 4 etapas: (1) Levantamento de requisitos para a construção do Data Warehouse, (2) definição dos dados nas origens



e (3) criação do ambiente de tratamento de dados e (4) criação da área de apresentação de dados. A imagem abaixo exhibe no fluxograma as quatro etapas do estudo de caso:

Figura 1 - Etapas do desenvolvimento de estudo de caso

ESTUDO DE CASO

Para fazer a análise do negócio, é necessário conceituar as etapas de desenvolvimento do estudo de caso para melhor elucidação.

Definição do processo de negócio

Um processo é uma atividade de negócio natural, executada na sua empresa, que normalmente é aceita por um sistema de coleta de dados de origem (KIMBAL, 2002). Ainda de acordo com o autor, a melhor forma de definir o processo de negócio é ouvir os usuários, pois as medições de desempenho que eles desejam analisar resultam de medições de processos de negócio. Nesse estudo de casos mesmos foram definidos através de entrevista com os gestores e usuários solicitaram os processos: venda de produtos; pagamento de comissões e perfil de clientes.

Definição do nível de detalhe

Segundo Kimbal (2002), definir o grão, granularidade ou nível de detalhes é especificar exatamente o que uma linha da tabela de fatos representa. Para o desenvolvimento deste estudo de caso escolheu-se trabalhar com o nível de detalhe mais baixo, com o dado em seu nível atômico, pois segundo o autor, os dados em seu grão mais atômico e mais baixo faz sentido em várias frentes. Os dados atômicos são altamente dimensionais. Quanto mais detalhada e atômica a medida dos fatos, mais coisa sabemos com certeza.

Definição de fatos e dimensões

Os fatos coletados pelo sistema transicional que são utilizados para medições de desempenho são: quantidade de itens vendidos, preço unitário do produto, preço

de custo do produto e o percentual de comissão dos vendedores. Tais fatos são analisados na perspectiva das dimensões: região, produto, pessoa, tempo, vendedor. Para o caso em questão, o modelo de dados escolhido para a construção do *Data Warehouse* é o modelo estrela.

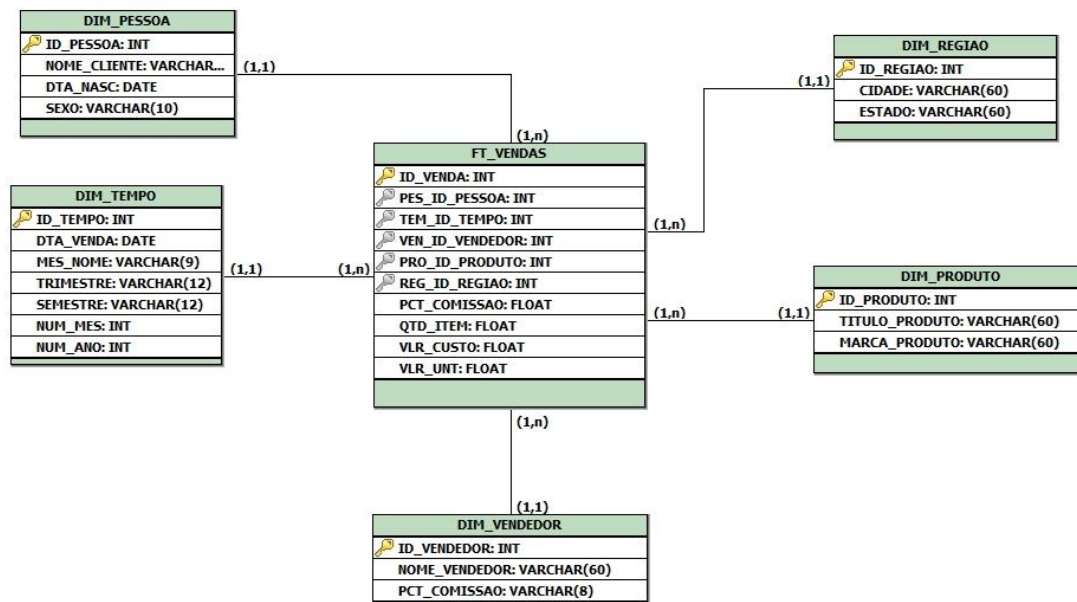


Figura 2 – Modelo de dados no formato estrela do processo de negócio de vendas.

Um modelo comumente usado para modelagem histórica e multidimensional de dados é o modelo estrela, em que uma relação central (tabela fato) armazena dados sobre os fatos que se quer analisar e relações que orbitam essa relação central armazenam dados que oferecem informação sob diferentes dimensões em relação a um fato (tabelas dimensão). (SILVA, 2016). Segundo o autor, o uso do modelo de dados estrela oferece meios de, a partir de consultas SQL ou com o uso de outros recursos, extrair informação mais refinada, atendendo à requisição, por exemplo, de extrair os perfis de consumo de clientes ou perfil de consumo por sazonalidade.

Área de staging

Como mencionado anteriormente, o *Data Warehouse* é carregado a partir de dados de diversas fontes e antes da carga os dados são verificados e tratados. Devido a limitações de processamento, rede e recursos tecnológicos torna-se inviável a retirada de dados de diversas fontes e efetuar o tratamento durante a carga no DW. Segundo Kimbal (2002), a staging área é como uma cozinha de um Data Warehouse, onde a equipe prepara os pratos e não tem contato direto com o cliente, e o cliente em nenhum momento irá consumir os pratos na cozinha.

Para o estudo de caso foi criado um banco de dados apartado com as tabelas normalizadas e foi implantado uma rotina (JOB) desenvolvido através da ferramenta SSIS (Integration Services). A função desta rotina é buscar os relatórios colocados na estrutura de pastas e carrega-los no banco de STAGING para que os dados possam ser tratados e posteriormente carregados no Data Mart de vendas. Para isso, foram utilizados os seguintes componentes na estrutura da rotina: Contêiner Foreach Loop, tarefa de execução SQL, tarefa fluxo de dados e tarefa sistema de arquivos.

Tabela 2: Aplicações dos componentes para coleta e transformação de dados

Componente	Utilização
Contêiner ForeachLoop	Varrer a estrutura de pasta selecionada para encontrar os relatórios que devem ser carregados na staging.
Tarefa de execução SQL	Gravar os dados dos arquivos carregados na tabela de controle de fluxo, como: nome do arquivo, quantidade de linhas, quantidade de linhas rejeitadas, data de carga e situação da execução. Esse componente também foi utilizado para evitar redundâncias.
Tarefa fluxo de dados	Definir o layout do arquivo, a origem e o destino do arquivo. Também foi utilizado para converter o tipo de dados de cada coluna.
Tarefa sistema de arquivos.	Mover os arquivos carregados para pasta de processado se obter êxito na execução ou para pasta de erro se houver falhas.

Por fim, a rotina foi implantada no catálogo do Integration Services no servidor SQL SERVER e agendado a execução através do SQL AGENT do SQL SERVER para execução a cada 1 hora.

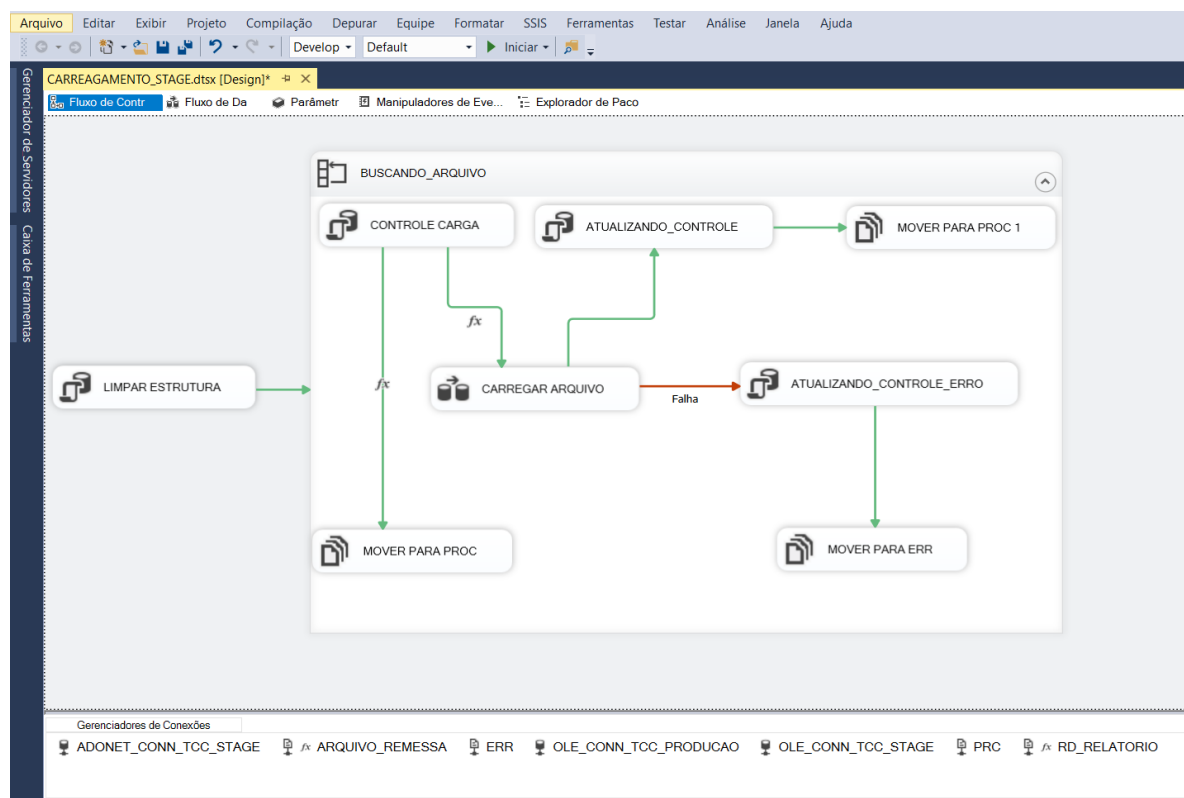


Figura 3 - Componentes do JOB de carga da área de staging.

Carregamento do data mart

Após a carga da área de staging os dados são coletados por outra rotina (JOB), também desenvolvida no SSIS. Essa rotina retira todas as inconsistências dos dados, organiza e carrega os dados dentro de suas respectivas dimensões e na fato do Data Mart no Data Warehouse. Os componentes utilizados nessa rotina foram: Contêiner de sequência, tarefa de fluxo de dados. A tabela abaixo demonstra a utilização de cada componente:

Tabela 3: Aplicação dos componentes para transformação dos dados e distribuição na estrutura do Data Mart

Componente	Utilização
Contêiner de sequência	Utilizado para agrupar e alterar a ordem de execução dos componentes.
Tarefa fluxo de dados	Além de definir a origem e destino dos dados, nesta rotina também foi utilizado componentes de fluxo de dados como: Componente de pesquisa e agregação para evitar a redundância de dados nas dimensões e componente de coluna derivada para criar colunas e efetuar transformações nos dados.

Após todo o desenvolvimento da rotina no SSIS, o pacote foi implantado nos catálogos do Integration Services no servidor do SQL SERVER e agendado na mesma

rotina do JOB anterior no SQL AGENT, porém a execução dessa rotina sempre vai ocorrer após a o termino da execução da rotina de carregar a staging.

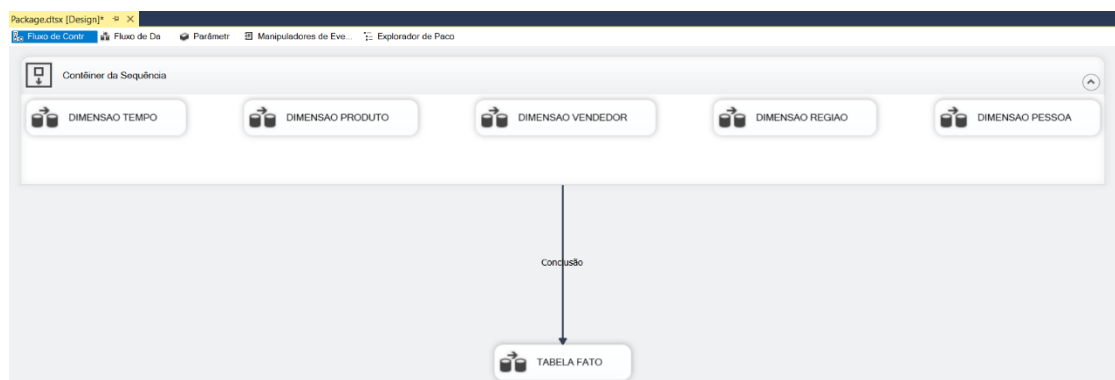


Figura 4 - Componentes do JOB de carga do Data Mart de vendas

CRIAÇÃO DOS RELATÓRIOS

Os relatórios com os indicadores de desempenho foram criados através do PowerBI desktop que foi conectado diretamente no Data Mart de vendas através da tecnologia Direct Query da própria ferramenta. Após a criação dos relatórios, eles foram armazenados no servidor de relatórios do PowerBI, para ficarem disponíveis para acesso dos usuários através do navegador WEB.

Os cálculos efetuados com as métricas foram utilizados as funções DAX, nativas do PowerBI, essas funções tem a sintaxe bem semelhantes as funções utilizadas no Excel, mas possuem suas particularidades. Elas permitem efetuar diversos tipos de cálculos matemáticos, estatísticos e ainda alterar a forma com que os cálculos são feitos a depender dos filtros selecionados nos painéis, criando um painel mais dinâmico e possibilitando com que o usuário altere o cenário analisado apenas com alguns cliques.

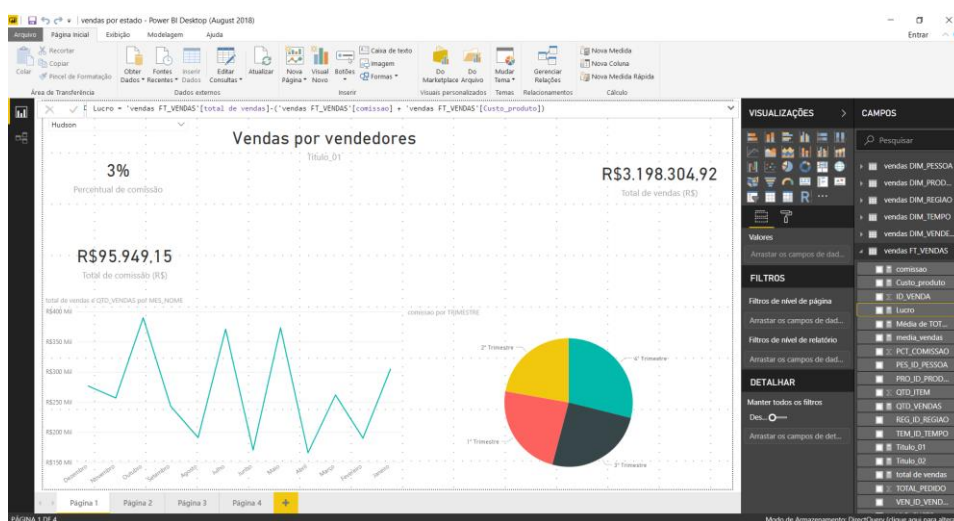


Figura 5 - Interface de desenvolvimento do PowerBI Desktop

RESULTADOS

usam os valores dos atributos descritivos para prever valores futuros ou desconhecidos de outros atributos de interesse.

Essa melhoria possibilitará aos gestores preverem possíveis problemas, antecipando-se as ações necessárias para suas correções, evitando-se prejuízos maiores ou perda de mercado.

REFERÊNCIAS

BUSINESS Intelligence (BI). Gartner, 2018. Disponível em: <<https://www.gartner.com/it-glossary/business-intelligence-bi/>>. Acesso em: 24 ago. 2018.

CÂNDIDO, C. H. Projeto brModelo 2.0 (Versão antiga). **Modelagem ER**, 2017. Disponível em: <<http://www.sis4.com/brModelo/antigo.html>>. Acesso em: 22 set. 2018.

GARTNER. **Business Intelligence (BI).** Gartner, 2018. Disponível em: <<https://www.gartner.com/it-glossary/business-intelligence-bi/>>. Acesso em: 24 ago. 2018.

GONÇALVES, R. R. **Integração de dados na prática.** 1ª. ed. São Paulo: Érica LTDA, 2012.

KIMBAL, R.; ROSS, M. **The Data Warehouse Toolkit.** Rio de Janeiro: Elsevier, 2002.

LAUDON, K.; LAUDON, J. **Sistemas de informação Gerenciais.** São Paulo: Person, 2011.

MACHADO, F. N. R. **Tecnologia e projeto de data warehouse.** 6ª. ed. São Paulo: Erica, 2004.

MICROSOFT. **Documentação do SQL Server.** Microsoft, 2018. Disponível em: <<https://docs.microsoft.com/pt-br/sql/sql-server/sql-server-technical-documentation?view=sql-server-2017>>. Acesso em: 22 set. 2018.

MICROSOFT. **Tutorial: Introdução ao serviço do Power BI.** PowerBi, 2018. Disponível em: <<https://docs.microsoft.com/pt-br/power-bi/service-get-started>>. Acesso em: 23 set. 2018.

MORAIS, E.; SIDENÉIA, S.; CARITÁ, E. **Business Intelligence utilizando tecnologias Web para análise de fatores de risco na ocorrência de doença arterial coronariana.** São Paulo: [s.n.], 2010.

PRIMAK, F. V. **Decisões com B.I.** 1ª. ed. Rio de Janeiro: Ciência Moderna, 2008.

SILBERSCHATZ, A.; KORTH, H. F.; SUDARSHAN, S. **Sistemas de banco de dados.** 5ª Edição. ed. Rio de Janeiro: Elsevier, 2006.

SILVA, L. A. D.; PERES, S. M.; BOSCARIOLI, C. **Introdução a mineração de dados.** Rio de Janeiro: Elsevier, 2016.