# PROJETO APLICADO EM BIG DATA, ANALYTICS E BI APLICADOS AOS NEGÓCIOS – CRIAÇÃO DE DASHBOARD

Raphael Caetano da Silva[[1]](#footnote-1)

**RESUMO**

Este projeto envolve a criação de um data Warehouse e a implementação de um dashboard de Business Intelligence (BI) utilizando Big Data e Analytics. Foi aplicado o Modern Data Stack, com foco na modelagem dimensional e em tecnologias de processamento de grandes volumes de dados, como o Google BigQuery. A análise foi realizada em dados da empresa fictícia Adventure Works, explorando diferentes áreas como vendas e produção.

A integração de ferramentas como DBT e Power BI permitiu a transformação de dados brutos em insights acionáveis para tomada de decisões estratégicas. Com foco em perguntas de negócio, o dashboard possibilitou a visualização de tendências, análise de vendas e desempenho de produtos e clientes, utilizando métricas como ticket médio, volume de vendas e comportamento geográfico. A solução evidencia a importância da engenharia de dados no contexto de Big Data e Analytics, oferecendo uma plataforma escalável e eficiente para análises profundas e automatizadas.

**Palavras-chave:** Big Data. Analytics. Data Warehouse. Business Inteligence (BI). DBT. Google BigQuery. PowerBI. Modelagem Dimensional. Engenharia de Dados. Modern Data Stack

**APPLIED PROJECT IN BIG DATA, ANALYTICS, AND BI APPLIED TO BUSINESS – DASHBOARD CREATION**

Raphael Caetano da Silva[[2]](#footnote-2)

**ABSTRACT**

This project involves the creation of a data Warehouse and the implementation of a Business Intelligence (BI) dashboard using Big Data and Analytics. The Modern Data Stack methodology was applied, focusing on dimensional modeling and technologies for processing large volumes of data, such as Google BigQuery. The analysis was conducted using data from the fictitious company Adventure Works, exploring different areas such as sales and production.

The integration of tools like DBT and Power BI enabled the transformation of raw data into actionable insights for strategic decision-making. Focusing on business questions, the dashboard allowed for the visualization of trends, sales analysis, and product and customer performance, utilizing metrics such as average ticket, sales volume, and geographic behavior. This solution highlights the importance of data engineering in the context of Big Data and Analytics, offering a scalable and efficient platform for deep and automated analyses.

**Keywords:** Big Data, Analytics, Data Warehouse, Business Intelligence (BI), DBT, Google BigQuery, Power BI, Dimensional Modeling, Data Engineering, Modern Data Stack

1. O PROBLEMA

## TEMA DO TRABALHO

O aumento exponencial da quantidade de dados gerados nas últimas décadas tem levado à necessidade de ferramentas mais sofisticadas para análise e tomada de decisões. Nesse contexto, o Business Intelligence (BI) tem se destacado como uma solução essencial para transformar dados brutos em informações acionáveis, permitindo que as empresas sejam mais competitivas e eficientes.

Com o crescimento da demanda por BI, ferramentas tradicionais de processamento de dados muitas vezes não são suficientes para lidar com o volume, a velocidade e a variedade dos dados gerados. Para enfrentar esse desafio, surge o conceito de Modern Data Stack (MDS), que combina tecnologias avançadas, como por exemplo DBT, Google BigQuery e Power BI, para fornecer uma solução escalável e flexível para a análise de grandes volumes de dados. O MDS se destaca por integrar pipelines de dados modernos com automação e processamento em nuvem, facilitando a transformação e visualização dos dados em tempo real. Com isso, as empresas podem tomar decisões mais rápidas e embasadas em dados, o que é vital em um ambiente de negócios cada vez mais competitivo.

## 1.2 CONTEXTUALIZAÇÃO

A Adventure Works é uma empresa fictícia criada pela Microsoft para ilustrar cenários reais de negócios. Nesse projeto, a empresa serve para ilustrar, validar e demonstrar informações que podem ser otimizados com o uso de Big Data, Analytics e Business Intelligence. A empresa atua no setor de manufatura de bicicletas, com mais de 500 produtos distintos, 20.000 clientes e 31.000 pedidos registrados. Em seu esforço para diferenciar-se da concorrência e manter o ritmo de crescimento, a Adventure Works busca se tornar uma empresa data-driven, ou seja, que baseia suas decisões estratégicas em dados e análises robustas.

A simulação de uma empresa como a Adventure Works permite explorar diversas áreas de um negócio que poderiam se beneficiar de uma plataforma de Data Warehouse e BI. O projeto abrange a análise de vendas, produção, e comportamento dos clientes, utilizando dados estruturados e não estruturados para construir relatórios que respondem a perguntas de negócios críticas, como as tendências de vendas e os produtos mais rentáveis. Ao aplicar o Modern Data Stack, a Adventure Works exemplifica como empresas de médio porte podem implementar soluções modernas de BI para otimizar seus processos e melhorar a competitividade no mercado.

## 1.3 A SITUAÇÃO PROBLEMA

A Adventure Works enfrenta o desafio de tomar decisões estratégicas baseadas em dados, porém seu sistema atual de análise e gestão de informações não é capaz de lidar com o volume crescente e a complexidade dos dados gerados por suas operações. A empresa precisa responder a uma série de perguntas críticas relacionadas às vendas, desempenho de produtos, comportamento de clientes e eficiência operacional, mas enfrenta dificuldades devido à falta de uma infraestrutura moderna para o tratamento, processamento e análise desses dados.

Atualmente, os dados da Adventure Works estão fragmentados em vários sistemas e não são processados de maneira integrada. Além disso, o processamento manual e a ausência de um data Warehouse centralizado dificultam a criação de relatórios precisos e em tempo hábil, impactando diretamente na capacidade da empresa de tomar decisões ágeis e embasadas. Isso resulta em oportunidades perdidas de otimização de vendas e operações, além de dificultar a identificação de tendências de mercado e o comportamento dos clientes.

Para solucionar esse problema, a empresa precisa implementar uma infraestrutura de Big Data e Business Intelligence que permita centralizar, transformar e visualizar os dados de maneira eficiente, permitindo uma análise em tempo real e a geração de insights que possam orientar suas estratégias de crescimento e diferenciação no mercado.

## 1.4 PERGUNTAS A SEREM RESPONDIDAS

No contexto do projeto, as seguintes perguntas de negócio foram levantadas pela diretoria da Adventure Works e precisam ser respondidas através da criação do data Warehouse e do dashboard de Business Intelligence (BI):

1. Qual o número de pedidos, quantidade comprada e valor total negociado por produto, tipo de cartão, motivo de venda, data de venda, cliente, status, cidade, estado e país?
2. Quais são os produtos com maior ticket médio por mês, ano, cidade, estado e país?  
   *(Ticket médio = Faturamento bruto - Descontos do produto / Número de pedidos no período analisado)*
3. Quem são os 10 melhores clientes em termos de valor total negociado, filtrado por produto, tipo de cartão, motivo de venda, data de venda, status, cidade, estado e país?
4. Quais são as 5 melhores cidades em termos de valor total negociado por produto, tipo de cartão, motivo de venda, data de venda, cliente, status, cidade, estado e país?
5. Qual é o número de pedidos, quantidade comprada e valor total negociado por mês e ano?
6. Qual produto tem a maior quantidade de unidades compradas para o motivo de venda "Promotion"?

# CONTEXTUALIZAÇÃO DO PROBLEMA

## 2.1 PREMISSAS E RESTRIÇÕES DO PROJETO

* **Premissas**
  + Para a execução deste projeto de Data Warehouse e Business Intelligence (BI), algumas premissas foram estabelecidas para garantir o sucesso na implementação:

1. Disponibilidade de Dados: Os dados transacionais da Adventure Works estarão acessíveis e prontos para serem extraídos dos sistemas de origem, como o banco de dados PostgreSQL, em seus respectivos schemas (HR, Sales, Production e Purchasing).
2. Ferramentas Modernas: Serão utilizadas ferramentas de última geração, como Google BigQuery para o armazenamento e processamento de grandes volumes de dados, DBT para a transformação e modelagem de dados, e Power BI para a visualização de insights e criação de dashboards.
3. Capacidade de Processamento em Nuvem: O ambiente de nuvem do Google BigQuery oferece escalabilidade para lidar com o volume de dados gerado pela Adventure Works, permitindo consultas rápidas e análises eficientes.
4. Automação e Testes de Qualidade: As transformações de dados serão automatizadas com o DBT, incluindo testes para garantir a integridade e a consistência dos dados.
5. Engajamento das Partes Interessadas: A equipe da Adventure Works, incluindo executivos e analistas de negócios, fornecerá as informações e os requisitos necessários para orientar o desenvolvimento do projeto.

* **Restrições**
  + Algumas limitações e restrições foram identificadas durante o planejamento do projeto:

1. Limitações de Tempo: O projeto deve ser concluído dentro de um prazo de 4 meses, incluindo a configuração do data Warehouse, a modelagem de dados, a criação de dashboards e a realização de testes.
2. Orçamento: O projeto possui um orçamento limitado, o que restringe a contratação de equipes adicionais e a aquisição de ferramentas de terceiros. As soluções escolhidas devem ser rentáveis e escaláveis.
3. Complexidade dos Dados: Os dados provenientes dos sistemas transacionais podem exigir transformações complexas para serem modelados em um formato analítico adequado (fatos e dimensões). Tabelas de diferentes áreas podem ter dependências cruzadas, exigindo um trabalho adicional para assegurar a qualidade e a integridade dos dados.
4. Limitações de Integração: Algumas tabelas importantes podem não estar completamente estruturadas ou disponíveis no momento inicial do projeto, impactando a velocidade de construção do data Warehouse.
5. Capacitação da Equipe: A equipe da Adventure Works pode precisar de treinamento adicional para utilizar as ferramentas de BI (Power BI) de forma eficiente, o que poderá demandar tempo e recursos além do previsto.

## 2.2 MERCADO CONSUMIDOR

A Adventure Works, uma empresa fictícia no setor de manufatura de bicicletas, atende a um mercado consumidor diversificado. Seu portfólio de mais de 500 produtos distintos, aliado a uma base de 20.000 clientes, coloca a empresa em um segmento competitivo que abrange tanto o consumidor final quanto distribuidores e revendedores.

* **Perfil dos Clientes**
  + O mercado consumidor da Adventure Works é composto por:
* Consumidores Individuais: Pessoas físicas que adquirem bicicletas e acessórios diretamente, representando uma parte significativa das vendas. Esses consumidores costumam buscar produtos de alta qualidade, variedade de modelos e customizações. A experiência de compra e o atendimento ao cliente são aspectos cruciais para manter a lealdade dessa base.
* Distribuidores e Revendedores: Representam uma parte estratégica das vendas. Esses clientes compram em grandes volumes e são responsáveis pela revenda dos produtos da Adventure Works em diferentes regiões e canais de distribuição, o que expande a presença da marca.
* **Segmentação Geográfica**
  + A Adventure Works opera em um mercado global, vendendo para diferentes regiões e países. A segmentação geográfica de suas vendas permite identificar quais regiões têm maior demanda, possibilitando a personalização das estratégias de marketing e vendas para essas áreas. Fatores como preferências regionais, sazonalidade e questões logísticas influenciam diretamente o desempenho em diferentes mercados.
* **Comportamento do Consumidor**
  + Os consumidores da Adventure Works tendem a ser sensíveis a inovações tecnológicas e à qualidade dos produtos. Com o aumento da competitividade no setor de bicicletas, especialmente no segmento de bikes elétricas e bicicletas customizadas, é fundamental que a empresa entenda as tendências de consumo, como:
    - Preferência por produtos sustentáveis: A demanda por bicicletas que utilizam materiais ecológicos ou que contribuem para a mobilidade urbana verde está em alta.
    - Crescimento do ciclismo esportivo e de lazer: O interesse por atividades ao ar livre, tanto para fins esportivos quanto recreativos, tem impulsionado as vendas, especialmente em mercados urbanos.
* **Oportunidades e Desafios**
  + A análise do mercado consumidor revela oportunidades para a Adventure Works focar em personalização e expansão de vendas internacionais. Entretanto, também há desafios, como o gerenciamento eficaz de estoques em múltiplas regiões e a adaptação rápida às mudanças nas preferências dos consumidores.

O data Warehouse e o dashboard de Business Intelligence desenvolvidos no projeto permitirão à Adventure Works monitorar de perto o comportamento dos consumidores, identificar padrões de compra e adaptar suas estratégias de vendas e marketing de maneira ágil.

## 2.3 PROPOSTA DE TRABALHO

### **2.3.1 MÉTODO DO TRABALHO**

O método utilizado neste projeto será uma abordagem de engenharia de dados aplicada, com base no Modern Data Stack (MDS), que envolve as etapas de extração, transformação e carregamento (ETL) de dados para a construção de um data Warehouse eficiente, seguido pela criação de um dashboard de Business Intelligence (BI) para análise e visualização dos dados.

A metodologia inclui as seguintes etapas:

1. **Diagnóstico e Levantamento de Requisitos**
   * Inicialmente, foi realizado um levantamento de requisitos com a equipe da Adventure Works para identificar as perguntas de negócio essenciais que precisam ser respondidas. Isso incluiu a análise das fontes de dados disponíveis e a definição das métricas e indicadores-chave de desempenho (KPIs) que guiarão o projeto. Nesta fase, também foram identificadas as principais tabelas de fatos e dimensões necessárias para o data Warehouse.
2. **Extração e Integração dos Dados**
   * Os dados transacionais da Adventure Works foram extraídos do banco de dados PostgreSQL, onde estão armazenados em diversos schemas (como vendas, produção e compras). O processo de ETL foi realizado utilizando o DBT (Data Build Tool) para automatizar as transformações necessárias e garantir a qualidade e consistência dos dados. Esta etapa envolveu a junção de diferentes fontes de dados e a normalização dos mesmos para serem carregados no Google BigQuery, o data Warehouse em nuvem escolhido para o projeto.
3. **Modelagem Dimensional**
   * Com base na metodologia de modelagem dimensional, foram criadas tabelas de fatos e dimensões que permitem uma análise detalhada dos dados. A modelagem dimensional foi aplicada para facilitar o cruzamento de informações, como análise de vendas por produto, cliente, localização geográfica, e períodos temporais. Isso garantiu a criação de uma estrutura de dados que suporta consultas rápidas e eficientes, ideal para visualizações e relatórios no Power BI.
4. **Criação de Dashboards**
   * A ferramenta Power BI foi utilizada para criar dashboards interativos que respondem diretamente às perguntas de negócio levantadas. O dashboard permite a visualização de métricas como ticket médio, volume de vendas, comportamento de clientes e regiões de maior desempenho, facilitando a análise em tempo real e a tomada de decisões estratégicas pela diretoria da Adventure Works.
5. **Testes e Validação**
   * Após a construção do data Warehouse e dos dashboards, foram realizados testes para garantir a integridade dos dados e a consistência dos relatórios gerados. Isso incluiu testes automatizados nas tabelas de dimensão e fatos, verificando a validade das chaves primárias e estrangeiras, além de garantir que as métricas estavam sendo calculadas corretamente.
6. **Entrega e Documentação**
   * A última etapa incluiu a entrega do projeto à equipe da Adventure Works, junto com a documentação detalhada de todo o processo. A documentação incluiu diagramas de modelagem, scripts de transformação de dados e instruções de uso do dashboard no Power BI. Além disso, foram realizadas sessões de treinamento com a equipe para garantir a plena utilização da solução entregue.

### **2.3.2 PREVISÃO DE RECURSOS HUMANOS E MATERIAIS**

1. **Recursos Humanos**

Para a execução deste projeto, serão necessários os seguintes profissionais, cada um com responsabilidades específicas dentro do processo de construção do data Warehouse e dashboard de BI:

1. **Engenheiro de Dados**
   * Responsável pela extração, transformação e carregamento (ETL) dos dados. Utiliza ferramentas como DBT e Google BigQuery para modelar os dados de forma eficiente e garantir a qualidade da infraestrutura de dados.
   * Habilidades necessárias: SQL avançado, Python, conhecimento em Big Data e plataformas em nuvem (Google Cloud).
2. **Analytics Engineer**
   * Responsável pela criação e manutenção de pipelines de dados, garantindo a transformação eficiente dos dados brutos em modelos analíticos. Trabalha diretamente com a modelagem dos dados, utilizando DBT para preparar os dados para análise e otimização de consultas no Google BigQuery.
   * Habilidades necessárias: SQL, transformação de dados, modelagem dimensional, automação de pipelines com DBT, conhecimento em integração com plataformas de BI.
3. **Analista de Business Intelligence (BI)**
   * Responsável pela criação de visualizações e dashboards interativos no Power BI. Constrói relatórios que respondem diretamente às perguntas de negócios e permitem a exploração de métricas de vendas, clientes, produtos e regiões.
   * Habilidades necessárias: Power BI, visualização de dados, modelagem de dados e compreensão de KPIs.
4. **Gerente de Projeto**
   * Coordena as atividades da equipe, garante o cumprimento de prazos e orçamento, além de atuar como ponto de contato entre a equipe técnica e os stakeholders da Adventure Works.
   * Habilidades necessárias: gestão de projetos, comunicação, experiência com metodologias ágeis (Scrum, Kanban).
5. **Equipe de Treinamento e Suporte**
   * Responsável por fornecer treinamento aos usuários finais sobre como utilizar o dashboard de Power BI e garantir o suporte técnico inicial após a entrega do projeto.
   * Habilidades necessárias: conhecimento técnico sobre as ferramentas utilizadas e habilidades de treinamento.
6. **Recursos Materiais**
7. **Infraestrutura em Nuvem**
   * Google BigQuery: Plataforma de data Warehouse em nuvem que será utilizada para o armazenamento e processamento de grandes volumes de dados.
   * Google Cloud Platform (GCP): Para o gerenciamento da infraestrutura em nuvem, incluindo o provisionamento de servidores e ambientes de desenvolvimento.
8. **Ferramentas de Transformação de Dados**
   * DBT (Data Build Tool): Ferramenta utilizada para a transformação de dados em SQL, automação de pipelines e garantia da consistência e integridade dos dados no data Warehouse.
9. **Ferramentas de Visualização de Dados**
   * Power BI: Utilizado para a construção de dashboards interativos e relatórios visuais. Licenças de software serão necessárias para os desenvolvedores e usuários finais que acessarão o dashboard.
10. **Equipamentos de Trabalho**
    * Computadores de Alto Desempenho: Computadores com capacidade para rodar os ambientes de desenvolvimento de dados, com as seguintes especificações mínimas:
      + Processador: Intel i7 ou equivalente.
      + Memória RAM: 16 GB ou superior.
      + Armazenamento: SSD de 512 GB ou superior.
      + Sistema Operacional: Windows ou Linux, com suporte para ferramentas de desenvolvimento de dados.
11. **Software de Gerenciamento de Projeto**
    * JIRA ou Trello: Para acompanhar o progresso das tarefas e garantir o cumprimento dos prazos.
12. **Software de Documentação e Comunicação**
    * Confluence ou Google Docs: Para a criação da documentação do projeto, relatórios e comunicação entre a equipe.
    * Slack ou Microsoft Teams: Para a comunicação instantânea e reuniões entre a equipe e os stakeholders.
13. **Custo Estimado**
14. **Licenças de Software:**
    * Power BI Pro: Aproximadamente R$ 60 por usuário/mês.
    * DBT Cloud: Aproximadamente US$ 50 por desenvolvedor/mês.
15. **Infraestrutura em Nuvem:**
    * Google BigQuery: O custo dependerá do volume de dados processados e armazenados, com uma média de US$ 5 a US$ 10 por terabyte processado.
    * Google Cloud Platform (GCP): Custos variáveis com servidores e ambiente de desenvolvimento, estimados em torno de US$ 100 a US$ 200 mensais.
16. **Equipamentos:** 
    * Aproximadamente R$ 8.000 por computador de alto desempenho.
17. **Cronograma de Alocação de Recursos**
18. Fase de Diagnóstico: Engenheiro de Dados, Analista de BI, Gerente de Projeto.
19. Fase de Implementação (ETL e Modelagem): Engenheiro de Dados, Cientista de Dados (opcional), Gerente de Projeto.
20. Fase de Criação de Dashboards: Analista de BI, Gerente de Projeto.
21. Fase de Testes e Validação: Engenheiro de Dados, Analista de BI.
22. Fase de Treinamento: Equipe de Treinamento e Suporte.

### **2.3.3 DIAGRAMA DE GANTT**

Figura - Imagem gráfica do cronograma de realização do projeto

Autor, 2024

O projeto será dividido nas seguintes etapas, garantindo uma execução estruturada e eficiente para a criação do data Warehouse e dos dashboards de BI:

1. **Diagnóstico e Levantamento de Requisitos (15/09/2024 - 20/09/2024):**

A fase inicial envolve o levantamento detalhado das necessidades de negócios da Adventure Works. Nessa etapa, os stakeholders da empresa fornecem os requisitos necessários para o projeto, incluindo as principais perguntas que o data Warehouse e o dashboard de BI precisam responder. Esta fase é essencial para garantir que o projeto se alinhe com as expectativas e os objetivos estratégicos da empresa.

1. **Análise de Requisitos e Documentação (21/09/2024 - 30/09/2024):**

Nesta fase, os requisitos levantados são analisados e documentados em detalhes. A equipe de projeto cria uma documentação que descreve as perguntas de negócios, as fontes de dados a serem utilizadas e os KPIs (indicadores-chave de desempenho). A documentação serve como base para o desenvolvimento técnico, garantindo clareza e alinhamento com os objetivos do projeto.

1. **Extração dos Dados (ETL) (01/10/2024 - 10/10/2024):**

Durante essa etapa, os dados são extraídos dos sistemas transacionais da Adventure Works utilizando ferramentas de ETL. A equipe de Analytics Engineers se encarrega de configurar os pipelines para extrair os dados brutos das fontes (como PostgreSQL) e prepará-los para serem transformados no Google BigQuery. Essa fase garante que todos os dados necessários estejam acessíveis e prontos para a transformação.

1. **Transformação e Integração dos Dados (11/10/2024 - 20/10/2024):**

Após a extração dos dados, ocorre a transformação e integração no Google BigQuery. As transformações envolvem a limpeza, padronização e aplicação de regras de negócios nos dados para prepará-los para a modelagem dimensional. Esta fase também abrange a integração de diferentes fontes de dados para garantir que as informações sejam coerentes e prontas para a análise.

1. **Modelagem Dimensional Inicial (21/10/2024 - 05/11/2024):**

A modelagem dimensional é uma etapa crucial do projeto, onde as tabelas de fatos e dimensões são criadas para permitir análises eficientes. Nesta fase inicial, a equipe de Analytics Engineers constrói o esqueleto da modelagem, criando as primeiras versões das tabelas e assegurando que as chaves primárias e estrangeiras estejam corretamente configuradas para suportar consultas ágeis.

1. **Ajustes e Otimização da Modelagem (06/11/2024 - 20/11/2024):**

Com a modelagem dimensional inicial concluída, essa fase se concentra na otimização e ajustes para garantir que a estrutura dos dados seja eficiente e escalável. A equipe revisa a performance das consultas e ajusta os modelos para lidar com o volume de dados esperado. Esta fase é fundamental para garantir que o data Warehouse seja ágil e capaz de responder às consultas com rapidez, mesmo em cenários de grandes volumes de dados.

1. **Criação de Dashboards (21/11/2024 - 28/11/2024):**

Após a modelagem dos dados, o Analista de BI cria os primeiros dashboards utilizando Power BI. Esses dashboards são desenvolvidos para fornecer insights visuais que respondem diretamente às perguntas de negócios identificadas na fase de diagnóstico. O foco é permitir que os usuários possam visualizar métricas como ticket médio, volume de vendas, e performance de produtos e regiões de forma clara e intuitiva.

1. **Refinamento de Dashboards (29/11/2024 - 07/12/2024):**

Com os dashboards iniciais prontos, essa fase envolve o refinamento baseado no feedback dos stakeholders. Ajustes são feitos para melhorar a usabilidade e a clareza das visualizações, garantindo que os dashboards sejam fáceis de navegar e que todas as métricas importantes estejam corretamente representadas.

1. **Testes de Integração (08/12/2024 - 14/12/2024):**

Os testes de integração garantem que todos os componentes do projeto – desde a extração de dados até a visualização nos dashboards – estejam funcionando de maneira fluida e integrada. São realizados testes de performance e de validação dos dados para garantir que as informações estejam corretas e que o sistema responda às consultas de maneira eficiente.

1. **Validação de Resultados com Stakeholders (15/12/2024 - 25/12/2024):**

Após os testes, a equipe se reúne com os stakeholders para validar os resultados apresentados nos dashboards e relatórios. Qualquer ajuste necessário será feito com base no feedback recebido nesta fase. A validação final é crucial para assegurar que o projeto atenda a todas as expectativas e requisitos de negócio.

1. **Treinamento dos Usuários (26/12/2024 - 05/01/2025):**

Com o sistema pronto, a equipe oferece treinamento aos usuários da Adventure Works para que eles saibam como utilizar o Power BI e os dashboards criados. Este treinamento garante que os usuários finais estejam capacitados para explorar os dados de forma independente, facilitando a tomada de decisões baseada em dados.

1. **Entrega Final e Suporte Inicial (06/01/2025 - 15/01/2025):**

Na fase final, o projeto é oficialmente entregue à Adventure Works, e a equipe oferece suporte técnico inicial para garantir que qualquer problema ou dúvida seja resolvido rapidamente. Durante esse período, a equipe também fornece a documentação final do projeto, incluindo instruções detalhadas para o uso e a manutenção dos sistemas desenvolvidos.

# A SOLUÇÃO

## O SISTEMA PROPOSTO

### **3.1.1. JUSTIFICATIVAS PARA O NOVO SISTEMA**

O novo sistema proposto, que consiste em um data Warehouse integrado com um dashboard de Business Intelligence (BI), justifica-se pela crescente necessidade da Adventure Works de tomar decisões estratégicas baseadas em dados. Atualmente, a empresa enfrenta diversos desafios relacionados à gestão e análise de dados devido à falta de uma infraestrutura centralizada e eficiente. Seguem as principais justificativas para a implementação do novo sistema:

1. **Centralização e Organização dos Dados**

A Adventure Works armazena seus dados em vários sistemas transacionais, o que resulta em informações fragmentadas e dificuldade para acessar dados de forma integrada. O novo sistema de data Warehouse permitirá centralizar todos os dados relevantes da empresa em uma única plataforma, facilitando o acesso e a organização das informações de diferentes áreas, como vendas, produção, e compras.

1. **Melhoria na Tomada de Decisões**

A falta de uma estrutura robusta de análise de dados impacta diretamente a capacidade da diretoria de tomar decisões embasadas. O sistema proposto utiliza Google BigQuery para processar grandes volumes de dados em tempo real, permitindo à equipe de liderança acessar insights estratégicos por meio de dashboards interativos. Com isso, a empresa poderá identificar tendências de mercado, padrões de consumo e oportunidades de crescimento de maneira mais rápida e eficiente.

1. **Automação de Processos Analíticos**

A ausência de automação nos processos de análise leva a uma dependência de relatórios manuais e demorados. Com o novo sistema, utilizando o DBT para automação das transformações de dados, a Adventure Works poderá automatizar o fluxo de dados desde a extração até a geração de relatórios, reduzindo significativamente o tempo necessário para preparar e analisar as informações. Isso permitirá que as equipes se concentrem em análises mais profundas e decisões estratégicas.

1. **Escalabilidade e Capacidade de Processamento**

O volume de dados gerado pela Adventure Works está em constante crescimento, e os sistemas atuais não são capazes de processar essas informações de forma eficiente. O uso do Google BigQuery, uma plataforma de Big Data em nuvem, oferece uma solução escalável que se adapta às necessidades de processamento da empresa, garantindo que o sistema possa crescer conforme a demanda aumenta.

1. **Melhor Visualização e Acesso a Dados**

O uso de Power BI como ferramenta de visualização oferece uma interface amigável e poderosa para que os usuários finais da Adventure Works possam acessar, filtrar e visualizar dados de maneira intuitiva. Isso facilitará o acompanhamento de métricas e indicadores-chave de desempenho (KPIs), possibilitando a personalização de relatórios de acordo com as necessidades específicas de cada área de negócios.

1. **Aumento da Competitividade**

Em um mercado competitivo, a capacidade de uma empresa de ser ágil e tomar decisões informadas pode ser a diferença entre sucesso e estagnação. O sistema de BI proposto permitirá que a Adventure Works se torne uma empresa verdadeiramente data-driven, capaz de se adaptar rapidamente às mudanças no mercado e identificar novas oportunidades antes dos concorrentes.

1. **Redução de Erros e Confiabilidade dos Dados**

A falta de integração entre os sistemas atuais aumenta o risco de inconsistências e erros nos dados. O novo sistema, com validações automáticas implementadas através de DBT, garante que os dados carregados no data Warehouse sejam consistentes e confiáveis, resultando em análises mais precisas e seguras.

### **3.1.2. SITUAÇÃO DESEJADA: OBJETIVOS GERAIS E ESPECÍFICOS**

* **Objetivo Geral**

O objetivo geral do novo sistema é fornecer à Adventure Works uma plataforma moderna e integrada de data Warehouse e Business Intelligence (BI), permitindo que a empresa tome decisões estratégicas baseadas em dados de forma rápida, precisa e eficiente. Isso envolve a centralização e organização de seus dados, a automação de processos analíticos e a criação de dashboards interativos, resultando em uma infraestrutura de análise de dados escalável e acessível para todas as áreas da empresa.

* **Objetivos Específicos**

1. **Centralizar Dados de Diferentes Fontes**
   * Unificar os dados de múltiplas fontes (vendas, produção, compras etc.) em um único data Warehouse, facilitando a gestão e o acesso a informações relevantes para a tomada de decisões.
2. **Automatizar o Processo de ETL (Extração, Transformação e Carregamento de Dados)**
   * Implementar pipelines automatizados de ETL com o uso de DBT, permitindo a transformação de dados brutos em insights acionáveis sem necessidade de processamento manual constante.
3. **Melhorar a Visualização de Dados com Dashboards de BI**
   * Desenvolver dashboards interativos no Power BI para permitir que os gestores da Adventure Works visualizem métricas essenciais, como ticket médio, volume de vendas, desempenho de produtos e tendências regionais de maneira clara e intuitiva.
4. **Aumentar a Precisão e Confiabilidade dos Dados**
   * Implementar testes automatizados para garantir a integridade e a consistência dos dados, minimizando erros e aumentando a confiabilidade das informações para a tomada de decisões.
5. **Apoiar a Tomada de Decisões Baseadas em Dados**
   * Fornecer insights em tempo real, utilizando o poder de processamento do Google BigQuery, para permitir que a empresa identifique rapidamente oportunidades de crescimento, otimização de vendas e melhorias operacionais.
6. **Escalabilidade da Solução**
   * Criar uma infraestrutura de dados escalável que acompanhe o crescimento da Adventure Works, garantindo que o sistema possa processar grandes volumes de dados conforme a empresa expande suas operações e base de clientes.
7. **Facilitar a Análise Preditiva**
   * Preparar o ambiente de dados para suportar análises preditivas no futuro, permitindo que a empresa antecipe tendências e demandas do mercado, otimizando a produção e a gestão de estoques.
8. **Melhorar a Eficiência Operacional**
   * Reduzir o tempo e o esforço necessário para gerar relatórios e realizar análises manuais, proporcionando mais agilidade às equipes de negócios e aumentando a eficiência operacional da empresa.
9. **Capacitar a Equipe de Negócios**
   * Treinar a equipe da Adventure Works para utilizar a nova plataforma de BI, garantindo que todos os departamentos sejam capazes de explorar os dados de forma autônoma, sem a necessidade de suporte técnico constante.

## 3.2. ESCOPO DA SOLUÇÃO

O escopo da solução proposta para a Adventure Works envolve o desenvolvimento e implementação de uma plataforma integrada de data Warehouse e Business Intelligence (BI), baseada no Modern Data Stack. Esta solução abrange desde a extração de dados de sistemas transacionais até a criação de dashboards interativos que permitem a visualização de métricas essenciais para a tomada de decisões estratégicas.

Abaixo estão os componentes principais que definem o escopo da solução:

**1. Infraestrutura de Dados**

* **Data Warehouse**: Implementação de um data Warehouse em nuvem utilizando o Google BigQuery como plataforma de armazenamento e processamento de grandes volumes de dados. O data Warehouse será a base para centralizar os dados de diferentes áreas da empresa (vendas, produção, compras etc.).
* **Pipelines de ETL Automatizados**: Utilização da ferramenta DBT (Data Build Tool) para automatizar a extração, transformação e carregamento de dados (ETL). Esses pipelines garantirão que os dados sejam atualizados automaticamente e prontos para análise, minimizando a necessidade de intervenção manual.

**2. Modelagem de Dados**

* **Modelagem Dimensional**: Estruturação dos dados em um modelo dimensional que inclui tabelas de fatos e dimensões, facilitando a realização de consultas e a análise de dados de forma eficiente. Essa modelagem suportará a análise de vendas por produto, cliente, região e tempo, além de métricas financeiras e operacionais.
* **Testes e Validações**: Implementação de testes automáticos para garantir a integridade dos dados, validando chaves primárias, foreign keys, e a consistência entre diferentes tabelas e fontes de dados.

**3. Visualização e Relatórios**

* **Dashboards de Business Intelligence**: Desenvolvimento de dashboards interativos no Power BI para permitir que a Adventure Works visualize suas principais métricas e KPIs (indicadores-chave de desempenho). Os dashboards serão configurados para exibir informações como:
  + Ticket médio por produto, região e período.
  + Volume de vendas, performance de produtos e comportamento dos clientes.
  + Análise geográfica de vendas, destacando as principais regiões de atuação.
* **Customização de Relatórios**: Os dashboards serão customizáveis, permitindo que os usuários finais ajustem filtros e parâmetros conforme necessário para suas necessidades específicas.

**4. Integração de Dados**

* **Integração de Múltiplas Fontes**: O projeto contemplará a integração de dados de diferentes sistemas da Adventure Works, como o banco de dados PostgreSQL utilizado pela empresa para armazenar dados transacionais, além de outras fontes de dados relevantes.
* **Conexões com Ferramentas de Análise**: O sistema será configurado para permitir a integração futura com ferramentas de análise avançada, como machine learning, se necessário.

**5. Segurança e Controle de Acesso**

* **Segurança de Dados**: Garantir que o acesso aos dados sensíveis seja controlado através de permissões e políticas de segurança adequadas, especialmente ao utilizar uma plataforma em nuvem como o Google BigQuery.
* **Controle de Acesso no Power BI**: Configuração de níveis de acesso no Power BI, garantindo que cada usuário possa acessar apenas as informações pertinentes à sua função dentro da empresa, respeitando as políticas de governança de dados.

**6. Treinamento e Capacitação**

* **Treinamento para Usuários Finais**: O escopo incluirá o treinamento da equipe da Adventure Works para utilizar as novas ferramentas de BI, garantindo que os usuários estejam capacitados para navegar nos dashboards e gerar relatórios personalizados de forma autônoma.
* **Documentação Completa**: Entrega de documentação detalhada sobre o sistema, incluindo manuais de uso dos dashboards e orientações sobre a manutenção dos pipelines de dados.

**7. Manutenção e Suporte Inicial**

* **Suporte Pós-Implementação**: Após a entrega do sistema, será fornecido suporte inicial para garantir que quaisquer ajustes ou dúvidas sejam resolvidos durante a fase de transição para o novo ambiente.
* **Aprimoramentos Futuramente Escaláveis**: O sistema será desenhado para permitir futuras expansões, como a adição de novos conjuntos de dados ou a integração com ferramentas de análise preditiva e machine learning.

### **3.2.1. LISTA DE REQUISITOS DO SISTEMA**

* **Requisitos Funcionais (RF)**
  + **RF1**: O sistema deve permitir a integração de dados de múltiplas fontes, incluindo o banco de dados PostgreSQL da empresa.
  + **RF2:** O sistema deve permitir a centralização de dados no Google BigQuery, consolidando informações de diferentes áreas da empresa, como vendas, produção e compras.
  + **RF3:** O sistema deve realizar automaticamente a transformação dos dados utilizando o DBT, garantindo a consistência e integridade das informações.
  + **RF4:** O sistema deve permitir a criação e visualização de dashboards no Power BI, exibindo métricas e KPIs essenciais para os usuários da Adventure Works.
  + **RF5:** O sistema deve permitir que os usuários filtrem e personalizem relatórios e dashboards conforme suas necessidades específicas, incluindo filtros por região, produto e período.
  + **RF6:** O sistema deve realizar atualizações automáticas dos dashboards com base nas mudanças nos dados do data Warehouse.
  + **RF7:** O sistema deve permitir a geração de relatórios automáticos e periódicos, enviando-os por e-mail para os gestores da empresa.
  + **RF8:** O sistema deve permitir o controle de acesso, garantindo que diferentes níveis hierárquicos (como diretores, gerentes e analistas) tenham acesso apenas às informações relevantes ao seu papel.
  + **RF9:** O sistema deve permitir a realização de análises de séries temporais para a visualização de tendências de vendas e comportamento dos clientes.
  + **RF10:** O sistema deve fornecer validações automáticas para verificar a integridade dos dados durante o processo de extração e transformação.
* **Requisitos Não Funcionais (RNF)**
  + **RNF1:** O sistema deve ser capaz de escalar conforme o volume de dados da Adventure Works aumentar.
  + **RNF2:** O sistema deve garantir que as consultas no Google BigQuery sejam executadas com uma performance adequada, sem atrasos significativos.
  + **RNF3:** O sistema deve garantir que as informações sensíveis sejam protegidas por políticas de segurança, com criptografia e controle de acessos.
  + **RNF4:** O sistema deve estar disponível 24/7, com mínima interrupção, garantindo alta disponibilidade para o acesso dos dados e dashboards.
  + **RNF5:** O sistema deve ser fácil de usar e intuitivo, permitindo que os usuários finais, mesmo sem conhecimentos técnicos avançados, possam navegar e gerar insights através do Power BI.
  + **RNF6:** O sistema deve fornecer logs de auditoria detalhados para monitoramento de acessos e alterações nos dados.

## 3.3 Modelagem de Dados:

A modelagem de dados desempenha um papel central na estruturação e organização das informações dentro de um ambiente de data Warehouse. O processo envolve a transformação de dados brutos, normalmente armazenados em ambientes OLTP (Online Transaction Processing), em um modelo otimizado para consultas analíticas e geração de relatórios. Para a Adventure Works, foi implementado o modelo Star Schema, que melhora a performance das consultas e facilita a criação de dashboards interativos no Power BI.

### **3.3.1 Situação Atual (OLTP)**

No esquema atual representado na **Figura 2**, observamos o seguinte conjunto de informações:

1. **Estrutura de Dados Normalizada**: O esquema OLTP é altamente normalizado, o que significa que os dados estão organizados em várias tabelas relacionadas para minimizar a redundância e garantir a integridade dos dados. Isso é ideal para operações transacionais diárias (inserções, atualizações, exclusões), mas pode ser ineficiente para consultas analíticas complexas que exigem a junção de várias tabelas.
2. **Foco em Transações**: O esquema é projetado para capturar e processar transações, como vendas, pedidos, informações de clientes e fornecedores. Cada área de negócios, como vendas, produção, compras, e recursos humanos, tem suas próprias tabelas relacionadas.
3. **Consultas Complexas**: Para gerar relatórios ou realizar análises, é necessário consultar e unir diversas tabelas, o que pode ser demorado e ineficiente para análises em larga escala.
4. **Falta de Organização para BI**: O modelo OLTP não é otimizado para relatórios de Business Intelligence (BI), pois as tabelas e relações são complexas e voltadas para o registro de transações, dificultando a agregação e análise de dados.

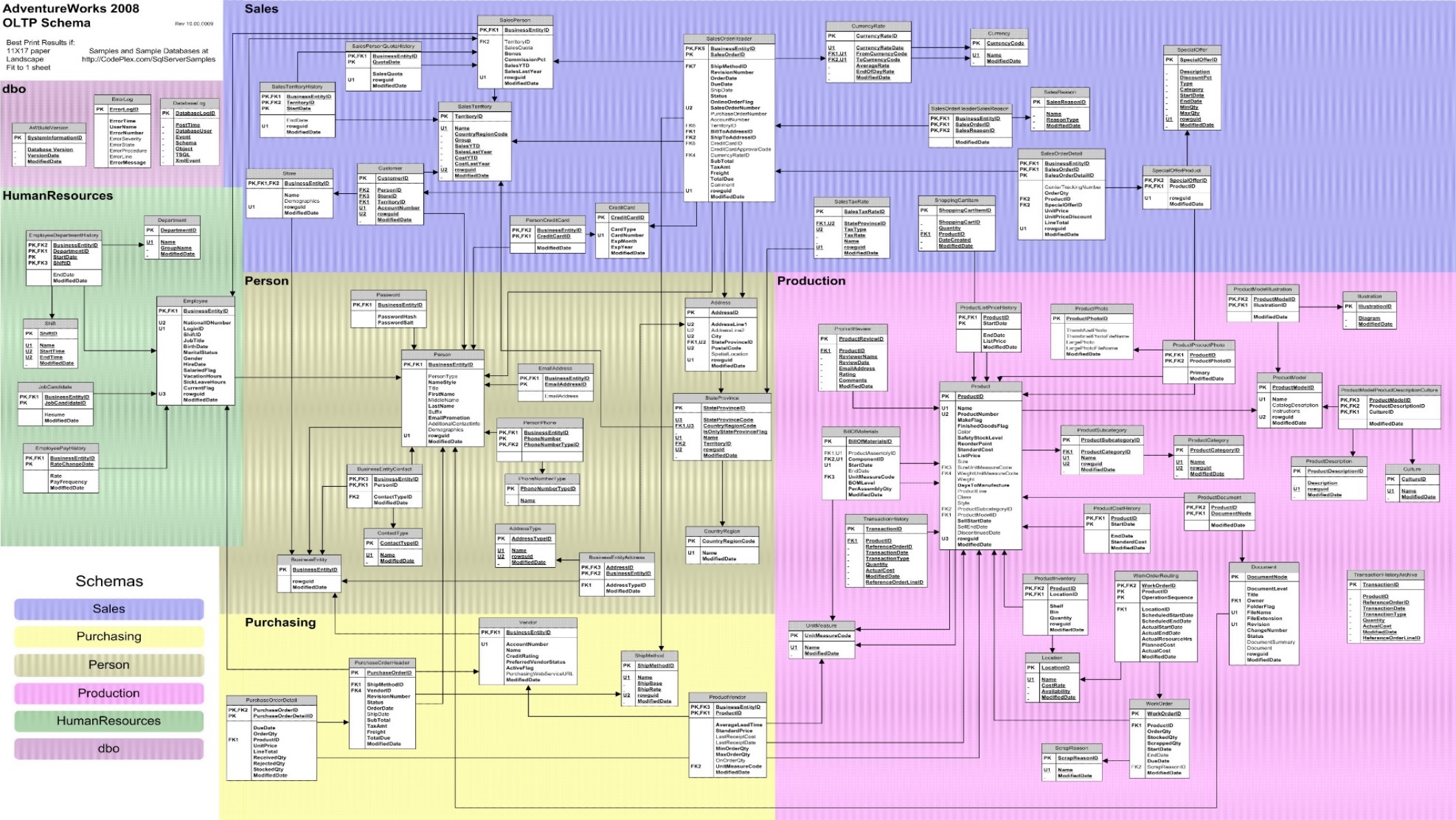


Figura : Modelo conceitual de dados - OLTP

Fonte: Microsoft

Com base no modelo atual, a Adventure Works encontra dificuldades para realizar consultas rápidas e eficientes. A falta de otimização para análise de dados cria desafios para a criação de relatórios interativos e dashboards. A solução proposta de modelagem dimensional transformará esse ambiente, trazendo melhorias em termos de eficiência, performance e facilidade de uso para análises e geração de relatórios

### **3.3.2 Modelagem Dimensional – Star Schema**

Após a transição do modelo OLTP para o Star Schema, a Adventure Works adotou uma estrutura de data Warehouse que permite uma análise rápida e eficiente dos dados.

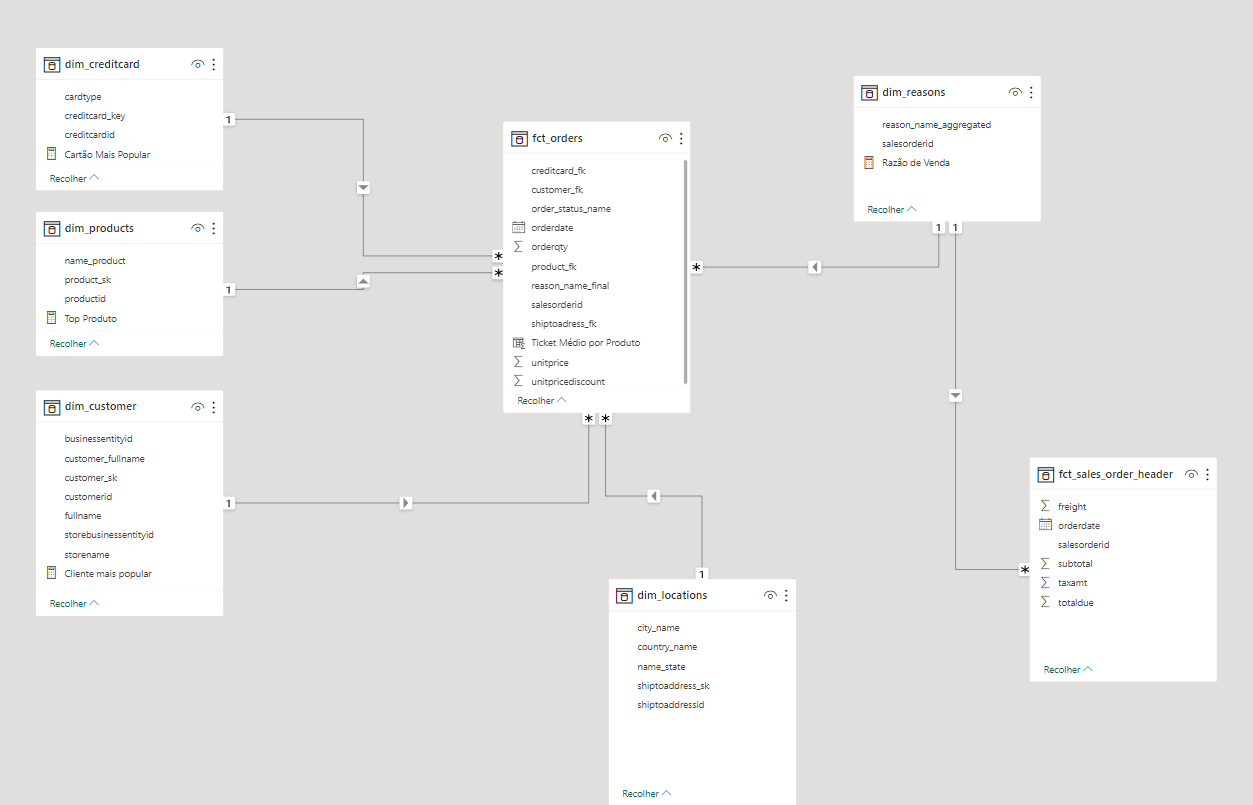


Figura : Modelo conceitual de dados (Fato x Dimensão)

*Fonte: Autor, 2024*

**Estrutura de Modelagem Dimensional**

1. **Tabelas de Fatos e Dimensões Otimizadas**:
   * **Fatos**: As tabelas de fatos, agora otimizadas para agregação e consulta, armazenam dados transacionais de grande volume (como vendas e produção), oferecendo melhor performance e acessibilidade. Essas tabelas estarão integradas com as dimensões para facilitar as análises.
   * **Dimensões**: As tabelas de dimensões serão estruturadas para suportar múltiplos filtros e classificações. Isso permitirá que os usuários explorem as informações com mais flexibilidade, resultando em análises mais detalhadas e relevantes.
2. **Desnormalização para Melhor Performance**:
   * A desnormalização dos dados na modelagem dimensional melhora a performance das consultas analíticas. Agora, as informações relevantes para as análises estarão prontamente acessíveis em poucas tabelas, eliminando a necessidade de múltiplas junções complexas, o que resultava em lentidão no ambiente OLTP.
3. **Facilidade de Uso para Business Intelligence**:
   * Com a modelagem dimensional, os dados estão organizados de uma forma intuitiva para os usuários finais. O processo de criar relatórios e dashboards tornou-se mais fácil, uma vez que as tabelas de fatos e dimensões fornecem a base necessária para gerar visualizações dinâmicas e responder a perguntas de negócios sem a necessidade de consultas complexas.
   * Melhoria: Agora os gestores podem explorar as informações rapidamente, aplicando filtros por produto, região, e tempo de maneira simples, e obtendo respostas em tempo real.
4. **Consultas Mais Rápidas e Eficientes**:
   * Com a modelagem dimensional, as consultas que antes eram lentas e demoradas no ambiente OLTP foram drasticamente otimizadas. O novo sistema é capaz de processar grandes volumes de dados em segundos, permitindo que a Adventure Works realize análises avançadas de forma rápida, mesmo com um aumento no volume de dados.
   * Melhoria: As consultas de vendas, comportamento do cliente e performance de produtos, que antes exigiam um tempo significativo para serem processadas, agora podem ser realizadas de forma ágil, proporcionando insights imediatos.
5. **Análises Preditivas e Avançadas**:
   * A nova modelagem oferece suporte para análises preditivas. Ao integrar dados de diferentes áreas em um modelo centralizado, a Adventure Works poderá identificar padrões e antecipar tendências de mercado com maior precisão.
   * Melhoria: Isso possibilita que a empresa tome decisões estratégicas antecipadas, como ajustes na produção ou estratégias de vendas focadas em áreas de maior potencial.
6. **Atualizações Automáticas e Consistentes**:
   * Com a automatização dos pipelines de dados através do DBT, as tabelas de fatos e dimensões serão atualizadas regularmente, garantindo que os relatórios sempre utilizem as informações mais recentes.
   * Melhoria: Isso elimina a necessidade de processos manuais de extração e preparação de dados, resultando em maior consistência e confiabilidade nas análises.
7. **Dashboards Interativos Melhorados no Power BI**:
   * Os dashboards agora são alimentados diretamente pelo Google BigQuery, permitindo visualizações interativas, como gráficos de séries temporais, mapas de calor e análises geográficas de vendas. Com a modelagem dimensional, esses dashboards são mais rápidos e precisos.
   * Melhoria: A velocidade de atualização e a capacidade de customização dos dashboards aumentaram significativamente, melhorando a experiência dos usuários finais

### **3.3.2 Comparativo: Melhorias após a Modelagem Dimensional**

|  |  |  |
| --- | --- | --- |
| Característica | OLTP | Modelagem Dimensional |
| Estrutura de Dados | Normalizada, várias tabelas relacionadas | Desnormalizada, com tabelas de fatos e dimensões otimizadas |
| Foco | Transações operacionais | Análises e relatórios de BI com alta performance |
| Consultas | Complexas e demoradas | Simples, rápidas e eficientes |
| Facilidade de Análise | Difícil de unir dados de várias áreas | Dados otimizados para BI, facilitando a análise e visualização |
| Performance | Lento para grandes volumes de dados | Alta performance para grandes volumes de dados |
| Atualizações de Dados | Manuais e demoradas | Automáticas, consistentes e em tempo real |
| Análises | Limitadas a relatórios básicos | Suporte a análises avançadas e preditivas |

## 3.4 Extração, Transformação e Carregamento (ETL)

O processo ETL (Extract, Transform, Load) é fundamental para a integração de dados em um data Warehouse, permitindo que dados brutos de diferentes fontes sejam convertidos em informações valiosas e facilmente acessíveis para análise. Na Adventure Works, a implementação de um fluxo de ETL automatizado foi crucial para garantir que os dados fossem preparados de maneira eficiente e consistente, facilitando as análises no Power BI e o suporte à tomada de decisões estratégicas.

### **3.4.1 Fluxo do ETL na Adventure Works**

1. **Extração**: Os dados são extraídos do banco de dados transacional PostgreSQL, onde estão armazenadas as informações de várias áreas operacionais.
   1. *Nota: No projeto da Adventure Works, o processo de ETL (Extract, Transform, Load) será emulado utilizando arquivos CSV como fonte de dados. Esses arquivos serão extraídos de diferentes sistemas operacionais e processados através de ferramentas específicas para garantir que sejam transformados e carregados corretamente no data Warehouse. O uso do DBT (Data Build Tool) será essencial para automatizar as transformações e realizar o carregamento dos dados, utilizando a funcionalidade de seed do DBT para alimentar o sistema.*
2. **Transformação**: Utilizando o DBT, as transformações necessárias são aplicadas aos dados, convertendo-os para o formato correto e aplicando regras de negócios relevantes para a Adventure Works. As transformações garantem que os dados estejam prontos para análise e sejam facilmente integrados ao Star Schema.
   1. **Fontes de dados Extraídas**: 12

Dados de diferentes sistemas operacionais são extraídos e disponibilizados em arquivos **CSV**.

* 1. **Tabelas de Fato**: 2

As tabelas de fatos são as principais responsáveis por armazenar os dados transacionais, como vendas e pedidos, e representam os eventos de negócios.

* 1. **Tabelas de Dimensão**: 5

As tabelas de dimensões armazenam informações descritivas (como produto, cliente, localização) que auxiliam nas análises, categorizando e filtrando os dados de acordo com diferentes perspectivas.

* 1. **Testes de Integridade de Dados**: 35

O sistema realiza 35 testes de integridade automatizados para garantir que os dados carregados estejam corretos, completos e coerentes com as regras de negócios da empresa.

1. **Carregamento**: Os dados transformados são carregados no Google BigQuery, onde são organizados em tabelas de fatos e dimensões, permitindo consultas rápidas e análises avançadas por meio do Power BI.

## 3.5 Business Intelligence - Dashboard

Na Adventure Works, o Power BI foi a ferramenta escolhida para implementar o Business Intelligence, conectando-se diretamente ao Google BigQuery, onde os dados são armazenados após o processo de ETL.

* **Google BigQuery:** Como o data Warehouse central do projeto, o Google BigQuery permite a armazenagem e consulta de grandes volumes de dados de forma escalável e eficiente. Ele oferece suporte para consultas rápidas e avançadas, facilitando a análise de dados diretamente no Power BI.
* **Power BI:** O Power BI é a principal ferramenta de visualização de dados, responsável por criar dashboards interativos que facilitam a interpretação das informações armazenadas no data Warehouse. Com uma interface intuitiva, os gestores da Adventure Works podem personalizar relatórios, aplicar filtros e explorar métricas de desempenho detalhadas.
  + *Nota: O dashboard criado no Power BI, utilizado para visualização das principais métricas de desempenho e relatórios da Adventure Works, será enviado em anexo juntamente com este documento.*

# 4 CONCLUSÕES

## 4.1. REFLEXÕES E COMPARAÇÃO ENTRE OBJETIVOS INICIAIS X ALCANÇADOS

Os objetivos iniciais deste projeto eram claros: criar uma solução robusta de Business Intelligence (BI) que permitisse à Adventure Works transformar seus dados em insights acionáveis, otimizando a tomada de decisões e melhorando a eficiência operacional. Desde o início, esperava-se a implementação de um data Warehouse eficiente, a automação do processo de ETL, e a criação de dashboards interativos para monitorar o desempenho da empresa.

Ao comparar os objetivos iniciais com os resultados alcançados, observa-se que:

* **Objetivos Iniciais**: Pretendia-se centralizar os dados em um data Warehouse, automatizar as transformações e o carregamento dos dados, e implementar um sistema de dashboards interativos para visualização em tempo real.
* **Resultados Alcançados**: O sistema de data Warehouse foi implementado com sucesso utilizando o Google BigQuery. O processo de ETL foi automatizado com o DBT, garantindo a consistência e a atualização regular dos dados. Por fim, os dashboards criados no Power BI fornecem visualizações detalhadas e personalizáveis das principais métricas de desempenho da empresa.

Portanto, os objetivos traçados foram integralmente atingidos, com o sistema permitindo análises de dados rápidas, precisas e acessíveis.

## 4.2. VANTAGENS E DESVANTAGENS DO SISTEMA

**Vantagens**:

1. **Centralização dos Dados**:  
   O sistema centralizou os dados em um data Warehouse no Google BigQuery, facilitando o acesso e a gestão de grandes volumes de informações de diferentes áreas da empresa.
2. **Automação e Consistência no Processo de ETL**:  
   A automatização do ETL com o DBT garantiu a consistência dos dados e eliminou a necessidade de processos manuais, assegurando que as informações estejam sempre atualizadas para análises e relatórios.
3. **Dashboards Interativos e Personalizáveis**:  
   A implementação dos dashboards no Power BI trouxe uma visualização clara e interativa dos dados. Com relatórios personalizáveis e atualizados em tempo real, os gestores podem explorar e filtrar as informações com facilidade, o que facilita a tomada de decisões.
4. **Escalabilidade**:  
   O uso de Google BigQuery como data Warehouse garante que o sistema possa escalar conforme o volume de dados da Adventure Works aumenta, sem comprometer a performance das consultas e análises.

**Desvantagens**:

1. **Dependência de Conectividade**:  
   A utilização de ferramentas em nuvem como o Google BigQuery e o Power BI requer uma conexão estável com a internet. Qualquer interrupção na conectividade pode impactar o acesso às análises e aos dashboards.
2. **Curva de Aprendizado**:  
   A integração com ferramentas avançadas como o Power BI e o DBT pode exigir uma curva de aprendizado para os usuários menos familiarizados com esses sistemas. Treinamentos adicionais podem ser necessários para que todos os usuários possam utilizar plenamente o sistema.

## 4.3. CONSIDERAÇÕES FINAIS

Este projeto trouxe à Adventure Works uma solução completa e integrada de Business Intelligence (BI), com um sistema automatizado e escalável para análise de dados em tempo real. A implementação do data Warehouse em Google BigQuery, associada ao processo de ETL automatizado com DBT e à visualização de dados através dos dashboards no Power BI, proporcionou à empresa uma ferramenta poderosa para monitorar suas operações, analisar o desempenho e tomar decisões mais informadas.

Além de entregar uma solução prática e eficaz para a Adventure Works, o projeto também desempenhou um papel crucial na consolidação dos conhecimentos adquiridos ao longo do curso de pós-graduação em Big Data, Analytics e BI Aplicados aos Negócios. A criação do sistema de BI permitiu aplicar na prática conceitos como modelagem dimensional (Star Schema), ETL automatizado e visualização de dados interativa, além de explorar tecnologias avançadas como o Google BigQuery, DBT e Power BI. Este projeto funcionou como uma importante checagem de aprendizado, comprovando a capacidade de integrar teorias e ferramentas tecnológicas em um cenário realista de resolução de problemas de negócios.

### **4.3.1 Importância para o Aprendizado na Pós-Graduação**

O projeto foi especialmente importante para consolidar os conhecimentos teóricos abordados ao longo da pós-graduação. A sua execução proporcionou a oportunidade de aplicar na prática os conceitos de engenharia de dados, análise de grandes volumes de dados (Big Data) e a criação de dashboards eficientes. Esta experiência não apenas fortaleceu a compreensão dos conceitos técnicos, mas também demonstrou como essas tecnologias podem ser utilizadas para gerar valor real para uma empresa, facilitando a tomada de decisões baseada em dados.

### **4.3.2 Conclusão**

As vantagens do sistema são notáveis, oferecendo uma visão consolidada e detalhada dos dados, uma redução significativa no tempo de geração de relatórios, e a capacidade de realizar análises. Embora existam desafios, como a dependência de conectividade e a necessidade de treinamento, os benefícios superam essas desvantagens. Em resumo, os objetivos foram alcançados com sucesso, tanto no aspecto técnico quanto educacional, equipando a empresa com uma solução moderna e escalável de BI, ao mesmo tempo em que provou a eficácia do aprendizado obtido durante a pós-graduação. Este projeto não apenas contribuiu para o desenvolvimento de uma ferramenta estratégica para a empresa, mas também comprovou a importância do aprendizado aplicado, integrando teoria e prática de forma eficaz.

1. Aluno concludente do curso de Pós-Graduação em Big Data, Analytics E Bi Aplicados Aos Negócios da Universidade Estácio de Sá [↑](#footnote-ref-1)
2. Student completing the Postgraduate course in Big Data, Analytics, and BI Applied to Business at Universidade Estácio de Sá [↑](#footnote-ref-2)