INSTITUTO FEDERAL
Minas Gerais
Campus
Ouro Branco

INSTITUTO FEDERAL DE MINAS GERAIS
CAMPUS OURO BRANCO

Trabalho Avaliativo – Governança de Dados e Plataformas de BI

Governança de Dados e sua aplicação prática em Plataformas de BI e Analytics com Governança

Aluno: Eduardo Octávio de Paula Souza
Professor: Lucas Portela Costa da Silva

Ouro Branco, 10 de Maio de 2025

1 Introdução Teórica

1.1 O que é Governança de Dados?

Governança de Dados é um conjunto de práticas, políticas, processos e padrões estabelecidos para garantir que os dados de uma organização sejam gerenciados de forma eficaz, segura e estratégica ao longo de todo o seu ciclo de vida. Trata-se de uma abordagem baseada em princípios, que envolve desde a aquisição, armazenamento, uso, compartilhamento, até o descarte dos dados, sempre com foco na integridade, qualidade e conformidade.

Em um cenário cada vez mais digital, no qual os dados se tornam o ativo mais valioso das organizações, a governança de dados torna-se essencial para empresas de todos os setores. À medida que avançam em suas jornadas de transformação digital, essas empresas dependem cada vez mais de dados confiáveis para operar com eficiência, inovar e competir. Cada área da organização é impactada por essa necessidade. Os gestores seniores precisam de dados precisos e atualizados para apoiar decisões estratégicas. Profissionais de marketing e vendas utilizam dados para entender o comportamento dos consumidores e antecipar suas demandas. As equipes de compras e cadeia de suprimentos dependem de informações confiáveis para controlar estoques e otimizar os custos operacionais. Já os agentes de conformidade precisam garantir que os dados estejam sendo tratados conforme as políticas internas e regulamentações externas, como a LGPD (Lei Geral de Proteção de Dados).

1.2 Por que é importante para a qualidade e segurança da informação?

A Governança de Dados tem como um de seus principais pilares a garantia da qualidade e da segurança da informação. Isso significa assegurar que os dados estejam corretos, íntegros, atualizados e protegidos contra acessos não autorizados, perdas acidentais ou vazamentos intencionais. O conjunto de processos, ferramentas, comportamentos e tecnologias envolvidos na governança de dados atua diretamente na proteção de informações sensíveis, promovendo um ambiente mais seguro para a operação organizacional. Ao implementar boas práticas de governança, a empresa reduz significativamente os riscos de ataques cibernéticos, além de garantir que informações estratégicas, como dados de clientes, contratos, projetos e comunicações internas, estejam devidamente resguardadas.

Para ilustrar, pense na sua vida pessoal, imagine se todos os seus dados como, documentos, fotos, senhas, conversas fossem perdidos ou acessados por pessoas mal-intencionadas. A sensação de vulnerabilidade e perda seria imediata. O mesmo acontece com as empresas, mas em uma escala que pode comprometer seu funcionamento, sua imagem no mercado e até sua existência legal, especialmente em tempos de regulamentações rigorosas como a LGPD.

Portanto, a segurança da informação, quando aliada à governança de dados, não é apenas uma medida técnica: é uma estratégia essencial para a continuidade dos negócios. Além disso, ela contribui para a qualidade da informação ao garantir que os dados utilizados nos processos decisórios estejam íntegros, rastreáveis e devidamente documentados.

1.3 Qual a relação com a tomada de decisão organizacional?

A Governança de Dados tem um papel estratégico na tomada de decisão organizacional. Decisões baseadas em dados (data-driven) exigem que as informações estejam disponíveis de forma confiável e em tempo hábil. Dados bem governados permitem análises mais precisas, identificação de oportunidades de negócio, redução de custos e resposta mais ágil às mudanças do mercado. Quando os dados são confiáveis, os gestores têm maior confiança em suas análises, o que leva a decisões mais bem embasadas e alinhadas com os objetivos organizacionais. Assim, a Governança de Dados se torna um pilar fundamental para a inteligência de negócios (BI) e para a transformação digital.

1.4 Principais diretrizes da Governança de Dados

A Governança de Dados se baseia em diversas diretrizes, entre as quais destacam-se:

- **Qualidade dos dados**: Essa diretriz deve garantir a precisão, consistência, integridade, atualidade e completude dos dados. A má qualidade dos dados pode comprometer toda a cadeia de análise e decisão.
- **Acesso seguro**: Define regras e controles para que apenas pessoas autorizadas possam acessar ou manipular determinados dados. Isso inclui mecanismos de autenticação, controle de permissões e criptografia.
- **Responsabilidade**: Envolve a definição clara de papéis e responsabilidades sobre os dados, como os "data stewards", que são responsáveis por garantir a qualidade e a conformidade dos dados dentro de suas áreas.
- **Linhagem e rastreabilidade**: Refere-se à capacidade de rastrear a origem, as transformações e o destino dos dados ao longo de sua jornada na organização. Isso é essencial para auditoria, conformidade e compreensão do impacto de alterações nos dados.
- **Documentação e glossário**: Manter registros claros sobre as definições, estruturas, regras e contextos dos dados é fundamental para garantir um entendimento comum e padronizado entre os usuários. O glossário de dados ajuda a alinhar os termos técnicos e de negócio usados pela organização.

2 Escolha e demonstração da ferramenta de BI com governança

2.1 Ferramenta escolhida:

A ferramenta escolhida para o trabalho é o Looker, pertencente ao Google, é uma plataforma de BI e análise de dados gratuita. Projetado para uso corporativo, o Looker vai além da simples visualização, oferecendo modelagem semântica de dados e forte integração com práticas de governança, segurança e compliance.

2.1 Controle de acesso e permissões

No Looker, o controle de acesso a dados e recursos é realizado principalmente por meio da criação de grupos de usuários e da atribuição de funções. Essas funções vinculam conjuntos de permissões a modelos LookML, que determinam quais campos e dados estão disponíveis para análise. Além disso, é possível limitar o acesso com filtros personalizados, que restringem os dados exibidos com base em atributos específicos de cada usuário.

A plataforma também permite restringir o trabalho de desenvolvedores a determinados bancos de dados por meio de projetos e aplicar Access Grants para controlar o acesso a análises detalhadas, visualizações, campos ou mesclagens. Esses mecanismos garantem que cada usuário acesse apenas as informações que são pertinentes ao seu perfil, reforçando a segurança, a governança e a conformidade no uso dos dados corporativos.

2.2 Compartilhamento com segurança

O Looker permite o compartilhamento seguro de relatórios tanto com usuários internos quanto externos, oferecendo níveis configuráveis de acesso, como apenas leitura ou permissão para edição. Além disso, a plataforma disponibiliza opções avançadas de segurança, como expiração de acesso, links protegidos e controle sobre o envio de dados. A integração com o Google Cloud Identity and Access Management (IAM) fortalece ainda mais o controle sobre quem pode visualizar, editar ou compartilhar os dados, garantindo conformidade e proteção da informação.

2.3 Versionamento ou certificação de dados

O Looker utiliza um modelo de desenvolvimento baseado no LookML, uma linguagem específica para modelagem de dados que permite criar estruturas reutilizáveis e padronizadas para análise. Essa abordagem promove maior consistência e controle sobre os dados utilizados nos relatórios e dashboards da organização.

Um dos principais diferenciais do LookML é o suporte nativo ao controle de versão via Git, possibilitando que todas as alterações feitas nos modelos de dados sejam versionadas, auditadas e revertidas quando necessário. Antes da publicação, é possível validar cada modificação, garantindo que apenas alterações consistentes e testadas entrem em produção. Além disso, o Looker permite a promoção de modelos certificados, o que ajuda a padronizar o uso de dados confiáveis e evitar a criação de relatórios baseados em fontes duplicadas ou incorretas.

2.4 Criação de dicionário/glossário ou descrição dos campos

Os modelos LookML do Looker permitem a documentação detalhada de cada campo e métrica por meio de descrições integradas diretamente no código. Essas descrições são exibidas automaticamente na interface do usuário, funcionando como um dicionário de dados embutido, o que facilita a compreensão e o uso correto das informações pelos analistas e tomadores de decisão. Além disso, o Looker pode ser integrado a catálogos de dados corporativos, como o Google Data Catalog, permitindo a conexão com glossários organizacionais.

2.5 Outro: Sistema de auditoria e monitoramento de atividades

Um outro recurso avançado é o sistema de auditoria e monitoramento de atividades, que registra acessos, alterações em modelos e visualizações de relatórios. Esse recurso é essencial para manter a rastreabilidade das ações dos usuários e garantir a conformidade com normas de segurança e privacidade, como a LGPD.

3 Conjunto de dados simulados

3.1 Tema: Controle de estoque e compras de peças de PC gamer por fornecedor

A planilha está disponível em:

https://docs.google.com/spreadsheets/d/13glLSJQm9benk6vCr2sdBze6o4qUCfwOOmrDOCQs6RI/edit?usp=sharing

3.2 Explicação da tabela:

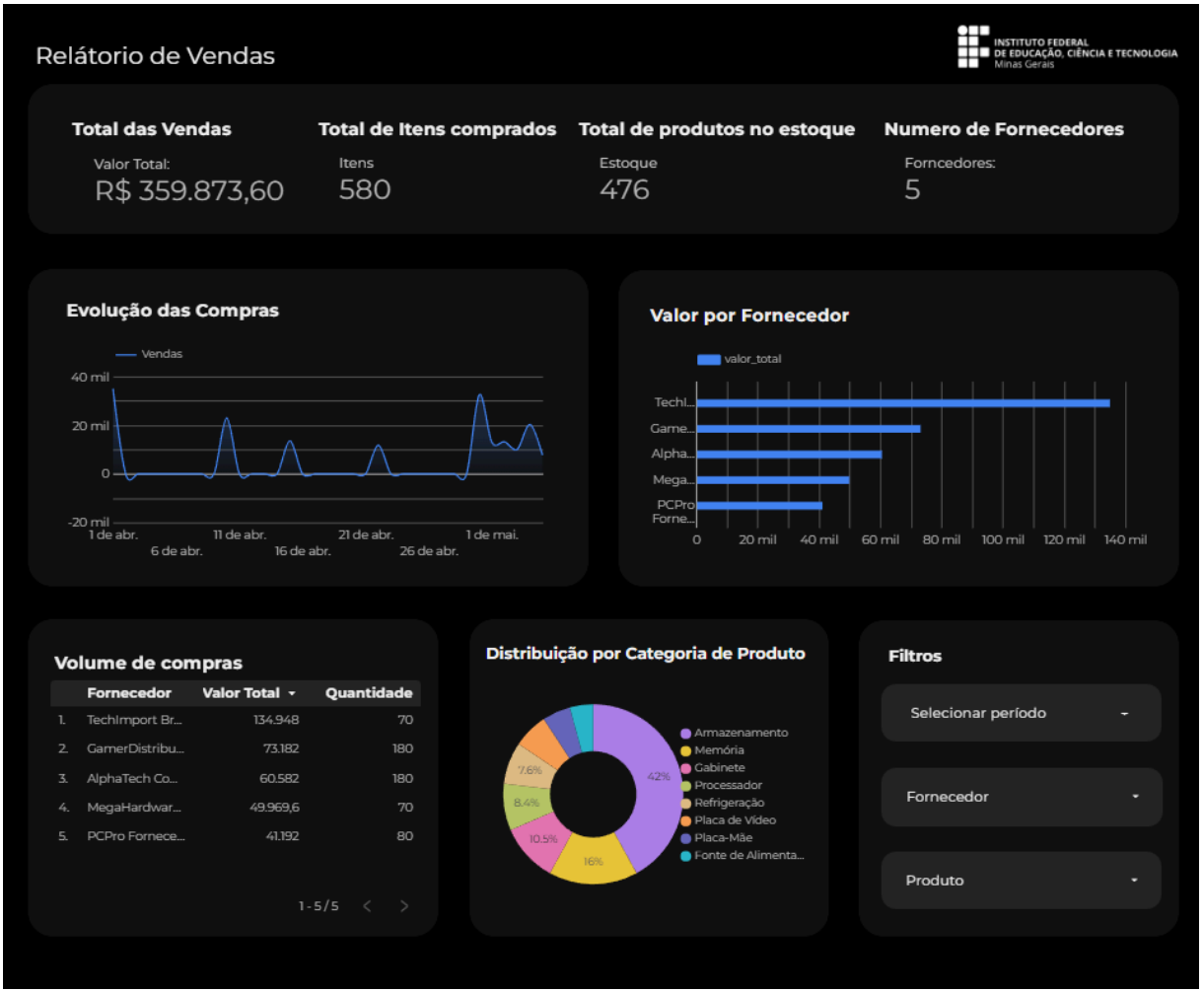
Este conjunto de dados simula o controle de estoque e compras de uma loja de peças de PC gamer, semelhante à Kabum!, Terabyte, etc. Inclui:

- **Fornecedores**: Contém informações como nome, CNPJ e contato das empresas fornecedoras.
- **Produtos**: Lista de peças vendidas, com preços, categorias e estoques atualizados.
- **Compras**: Registro das compras realizadas pela loja, com data, quantidade e valor total.

Tipo de Dado	Classificação	Justificativa
Nome do fornecedor	Público	Informação comum em catálogos ou sites comerciais.
CNPJ e e-mails	Confidencial	Dados empresariais que devem ser protegidos contra uso indevido.
Dados de compras (valores, quantidade)	Interno restrito	Dados estratégicos que revelam volumes e negociações comerciais.
Estoque atual	Interno	Informação útil para tomada de decisão interna, não deve ser exposta a terceiros.

4 Dashboard na ferramenta escolhida

4.1 Dashboard:



O dashboard está disponível em:

https://lookerstudio.google.com/reporting/c55c7cc8-e34e-4020-a68e-1db73b6f294a

4.2 Perfis Simulados:

Perfil	Acesso permitido	Restrições
Gerente de Compras	Acesso total aos KPIs, relatórios de compras, gráficos por fornecedor e produto	Nenhuma restrição
Equipe de Estoque	Acesso apenas aos dados de estoque atual, categorias de produtos e movimentações internas	Sem acesso a valores financeiros ou gráficos de gastos
Financeiro	Visualização de indicadores monetários e relatórios de valor total por fornecedor	Sem acesso ao estoque físico ou detalhes técnicos dos produtos
TI	Acesso completo com permissões de edição no painel para manutenção e evolução	Nenhuma restrição, mas com responsabilidade técnica e auditoria
Fornecedores	Acesso restrito aos seus próprios dados (via filtros por id_fornecedor)	Sem acesso aos dados de outros fornecedores ou análises estratégicas

5 Conclusão

5.1 Como a ferramenta escolhida apoia a governança dos dados

O Looker apoia a governança dos dados por meio de recursos como controle de acesso baseado em funções e atributos de usuário, versionamento de modelos via Git, filtros de acesso personalizados e documentação integrada com o LookML.

5.2 Como isso influencia a confiabilidade das decisões

Com dados bem documentados, controlados e acessíveis apenas por usuários autorizados, as decisões tomadas com base em análises feitas no Looker tornam-se mais confiáveis. Os gestores passam a contar com informações consistentes, auditáveis e extraídas de fontes oficiais, reduzindo erros e aumentando a precisão das estratégias adotadas.

5.3 Quais riscos seriam evitados com o uso correto da ferramenta?

O uso correto do Looker ajuda a evitar riscos como: acesso não autorizado a informações sensíveis, uso de dados desatualizados ou inconsistentes, duplicação de relatórios, falhas de segurança e descumprimento de normas como a LGPD. Além disso, reduz a dependência de planilhas manuais e análises paralelas fora dos padrões da empresa.

5.4 Quais boas práticas foram aplicadas no trabalho

Durante a construção do painel de BI, foram aplicadas diversas boas práticas relacionadas à governança de dados, com o objetivo de garantir o uso responsável, seguro e confiável das informações da empresa. As principais práticas foram:

- **Controle simulado de acesso por perfil de usuário:** O trabalho definiu perfis com diferentes níveis de acesso, respeitando o princípio do menor privilégio, onde cada perfil tem acesso apenas às informações necessárias à sua função.
- **Uso de filtros por dimensão relevante:** Filtros como fornecedor, categoria de produto e período de compra foram configurados para permitir que os dados sejam analisados de forma segmentada, possibilitando simular restrições por perfil e facilitar auditorias.
- **Padronização e documentação dos dados:** As tabelas utilizadas foram organizadas com campos consistentes e descrições claras. Os nomes dos campos foram normalizados e as unidades de medida padronizadas, garantindo integridade semântica.
- **Transparência e rastreabilidade**: As visualizações criadas representam fielmente os dados de origem, e as fontes utilizadas (como planilhas estruturadas) permitem a rastreabilidade dos registros. Isso facilita auditorias e aumenta a confiabilidade nas decisões baseadas nesses dados.

6 Bibliografia

GOOGLE CLOUD. O que é governança de dados? Disponível em: https://cloud.google.com/learn/what-is-data-governance?hl=pt-BR.

ZENDESK. A importância da segurança da informação. Disponível em: https://www.zendesk.com.br/blog/importancia-da-seguranca-da-informacao/.

GOOGLE CLOUD. Controle de acesso e gerenciamento de permissões no Looker. Disponível em: https://cloud.google.com/looker/docs/access-control-and-permission-management?hl=pt-br.