

Guia Ágil

TechGuide - Alura

Business Intelligence

Nível 1

☐ Business Intelligence (BI) - Fundamentos:

- Business Intelligence é um conjunto de teorias, metodologias, processos e tecnologias que possibilitam a transformação dos dados "crus" em informações extremamente relevantes para tomada de decisão de uma empresa.
- Conhecer o processo de ETL
- Realizar a modelagem e estruturação de tabelas em um Data Warehouse
- Criar visualizações que façam sentido
- Conhecer o PowerBI

☐ Análise e Visualização de Dados (BI):

- Esses conteúdos abrangem o campo da análise de dados e visualização de informações, com foco no Power BI, Tableau, Qlik Sense e Google Sheets.
- Os vídeos e cursos oferecem introdução ao Power BI, análise de dados da Microsoft, preparação de dados, modelagem de dados, visualização de dados, análise de dados e gerenciamento de workspaces e conjuntos de dados no Power BI.
- Os cursos da Alura oferecem formação em Power BI, Data Analysis com Google Sheets, Tableau, Looker Studio e Qlik Sense, fornecendo

conhecimentos abrangentes nessas ferramentas.

- Os artigos fornecem informações sobre o que é Qlik, Tableau, Power BI e Looker Studio, bem como seus produtos e funcionalidades.
- Esses conteúdos são úteis para profissionais que desejam aprimorar suas habilidades em análise de dados, visualização de informações e utilização de ferramentas como Power BI, Tableau, Qlik Sense e Google Sheets.

☐ **Excel:**

- O Microsoft Excel é um editor de planilhas produzido pela Microsoft com ferramentas de cálculo e de construção de tabelas.
- Realizar as operações matemáticas básicas com seus operadores (soma, subtração, multiplicação e divisão)
- Conhecer as principais fórmulas, como 'MÉDIA' (AVERAGE), 'ARRED' (ROUND), 'MÁXIMO' (MAX), 'MÍNIMO' (MIN), etc
- Realizar buscas em colunas com a função 'PROCV'
- Criar gráficos

☐ **Storytelling com dados:**

- Storytelling é uma forma de contar histórias que engajam e chamam a atenção da pessoa que está ouvindo. Dentro da análise de dados, é algo muito importante para passar as informações ao receptor de modo que o mesmo compreenda não apenas os dados, mas também todo o contexto.

☐ **SQL - Fundamentos:**

- SQL (Structured Query Language, traduzindo, Linguagem de Consulta Estruturada) é uma linguagem de programação padronizada que é usada para gerenciar bancos de dados relacionais e realizar várias operações sobre os dados neles contidos.
- Conhecer os comandos mais comuns do SQL
- Usar SELECT para consultar uma tabela
- Usar INSERT para inserir dados em uma tabela
- Usar UPDATE para atualizar uma tabela

- Usar DELETE para remover dados de uma tabela
- Usar JOIN para conectar os dados de múltiplas tabelas
- Conhecer as cláusulas (FROM, ORDER BY, etc)

Nível 2

☐ Power BI:

- O Microsoft Power BI é um serviço de análise de negócios e análise de dados da Microsoft. O objetivo do Power BI é fornecer visualizações interativas e recursos de business intelligence (BI) em uma interface para que os usuários finais criem relatórios e dashboards personalizados.
- Os cursos abrangem desde uma introdução ao Power BI até tópicos avançados, como mergulho na linguagem M, modelagem de dados e aplicação de DAX ao negócio.
- Os artigos fornecem insights e exemplos práticos sobre o desenvolvimento de dimensão calendário com a linguagem M.
- Os podcasts abordam temas como inteligência aplicada no Power BI.
- Os recursos são adequados para iniciantes que desejam aprender sobre o Power BI e também para usuários avançados em busca de aprimoramento.
- O Guia oferece uma oportunidade de aprofundamento no mundo do Power BI, fornecendo conhecimentos abrangentes sobre visualização e análise de dados.

☐ Pipelines e ETL:

- Pipelines e ETL (Extract, Transform, Load) são conceitos essenciais no campo da Engenharia de Dados, desempenhando um papel fundamental na coleta, transformação e integração de dados. Compreender os fundamentos dessas áreas é essencial para aqueles que desejam trabalhar com a gestão e o processamento eficiente de grandes volumes de informações.

☐ Modelagem de dados:

- A Modelagem de Dados é uma área essencial para a compreensão e organização das informações em um sistema ou banco de dados.

- Nesse conjunto de conteúdos, você aprenderá as técnicas e conceitos fundamentais para modelar dados de forma eficiente e estruturada, visando obter informações relevantes para a tomada de decisões.

☐ **Design de apresentações:**

- Aprenda sobre design de interfaces, criação de layouts, paletas de cores em ferramentas como Figma, PowerPoint e Canva.
- Os artigos fornecem informações e orientações sobre criação de interfaces, paletas de cores e tipografia.
- Os cursos oferecem instruções práticas sobre o uso de ferramentas como Figma, Canva e PowerPoint para criar layouts, projetar sites responsivos, editar imagens, criar apresentações e utilizar componentes da interface.
- Explore técnicas de Design Thinking.

☐ **Python para Ciência de Dados:**

- Python é uma linguagem de programação interpretada de alto nível e que suporta múltiplos paradigmas de programação, como imperativo, orientado a objetos e funcional. É uma linguagem com tipagem dinâmica e gerenciamento automático de memória.
- Aprender lógica de programação em Python
- Aprenda os fundamentos da linguagem como variáveis, funções, listas, condicionais e imports
- Criar análises de dados
- Utilizar o Matplotlib para gerar gráficos
- Usar e manipular listas para agrupar dados
- Conhecer a biblioteca NumPy
- Conhecer a biblioteca Pandas

Nível 3

☐ **Data Warehouses:**

- Um Data Warehouse é um sistema de armazenamento de dados projetado para análise de negócios, suportando processamento de transações e consultas complexas. Ele permite a consolidação de dados de várias fontes, fornecendo um ambiente unificado para análise de dados. Com recursos de normalização e limpeza de dados, garante a qualidade e a consistência dos dados armazenados. Soluções de Data Warehouse populares incluem Google BigQuery, Amazon Redshift e Snowflake, cada uma com suas próprias características e benefícios únicos.
- Data Warehouses desempenham um papel crucial na engenharia de dados. Eles fornecem uma infraestrutura que facilita a gestão, o processamento e a análise de grandes volumes de dados. Sua capacidade de armazenar e processar grandes quantidades de dados de várias fontes permite análises avançadas e suporte à tomada de decisões, o que é fundamental no atual cenário empresarial orientado por dados. Os engenheiros de dados projetam, constroem e mantêm data warehouses para garantir que os dados estejam acessíveis, seguros e prontos para análise. Além disso, soluções modernas de data warehouse, como o BigQuery, o Amazon Redshift e o Snowflake, oferecem capacidades de escalabilidade e flexibilidade para lidar com o crescente volume, velocidade e variedade de dados, tornando-se ainda mais valiosos para a engenharia de dados.

Big Data - Fundamentos:

- Big Data refere-se a conjuntos de dados extremamente grandes e complexos, que não podem ser facilmente processados ou gerenciados por métodos tradicionais. Esses conjuntos de dados são caracterizados por seu volume massivo, velocidade de geração e variedade de tipos e formatos. Com a explosão da quantidade de dados gerados diariamente, provenientes de diversas fontes como redes sociais, dispositivos IoT e transações comerciais, o Big Data apresenta desafios e oportunidades para extrair insights valiosos e tomar decisões informadas.
- A análise de Big Data envolve o uso de técnicas e tecnologias avançadas, como armazenamento distribuído, processamento em paralelo e aprendizado de máquina, para explorar e transformar esses dados em informações significativas para empresas, organizações e pesquisadores.

☐ **Processamento distribuído de dados:**

- Em processamento distribuído de dados usamos técnicas e ferramentas para realizar tarefas computacionais em um conjunto de máquinas conectadas, conhecido como cluster. Essas máquinas trabalham em conjunto para analisar e processar grandes volumes de dados de maneira mais eficiente do que seria possível em uma única máquina. O profissional pode usar várias ferramentas e frameworks para realizar esse trabalho, incluindo Apache Hadoop, Apache Spark e Databricks.

☐ **SQL - Avançado:**

- Conhecer uma variedade abrangente de conteúdos relacionados a bancos de dados e linguagem SQL.
- Uso de triggers em SQL, linguagens internas de SGBDs, modelagem de dados e manipulação de consultas avançadas.
- Utilize os principais bancos de dados relacionais (Oracle, Microsoft SQL Server, PostgreSQL e MySQL...)
- Aprenda sobre procedimentos, funções, pacotes, otimização do banco, segurança e administração dos bancos de dados.

☐ **Cloud - Fundamentos:**

- Cloud, ou computação em nuvem é a distribuição de serviços de computação pela Internet usando um modelo de preço pago conforme o uso. Uma nuvem é composta de vários recursos de computação, que abrangem desde os próprios computadores (ou instâncias, na terminologia de nuvem) até redes, armazenamento, bancos de dados e o que estiver em torno deles. Ou seja, tudo o que normalmente é necessário para montar o equivalente a uma sala de servidores, ou mesmo um data center completo, estará pronto para ser utilizado, configurado e executado.
- Conhecer a diferença entre IaaS, PaaS e SaaS
- Conhecer os maiores provedores de cloud
- Especializar-se em algum provedor

Habilidade Auxiliar: Comunicação, Negócios e Colaboração

☐ **Análise de negócios:**

- A análise de negócios auxilia os analistas de BI a coletar e analisar as necessidades empresariais, alinhando as soluções de BI com os objetivos.
- Permitir uma comunicação eficaz e tradução de requisitos entre usuários de negócios e equipes técnicas.
- Aprimora a modelagem de dados e o design de relatórios para saídas de BI precisas e perspicazes.
- Facilitar a colaboração com os stakeholders, garantindo que os projetos de BI atendam às expectativas.
- Contribuir para uma melhor tomada de decisões, analisando processos e propondo soluções baseadas em dados.

☐ **Habilidades de comunicação:**

- Um bom nível de comunicação facilita o atingimento de objetivos, resolução de problemas, além de aumentar a produtividade, porque cada profissional saberá exatamente o que se espera dele e transmitir com clareza suas ideias.

☐ **Git e GitHub - Fundamentos:**

- Git é um sistema de controle de versão distribuído gratuito e de código aberto projetado para lidar com tudo, desde projetos pequenos a muito grandes com velocidade e eficiência.
- GitHub é um serviço de hospedagem para desenvolvimento de software e controle de versão usando Git.
- Criar um repositório
- Clonar um repositório
- Fazer commit, push e pull de e para o repositório
- Reverter um commit

- Criar branches e pull requests
- Lidar com merge e conflitos

☐ **Inglês técnico:**

- Um bom nível de inglês técnico facilita o atingimento de objetivos, resolução de problemas, além de aumentar a produtividade, e também permite o consumo de materiais de diferentes fontes, principalmente de documentações oficiais.

Habilidade Auxiliar: Estatística e Engenharia de Dados

☐ **Engenharia de Dados - Fundamentos:**

- Uma pessoa Engenheira de Dados desempenha um papel crucial ao projetar, implementar e manter as infraestruturas de dados que permitem que uma organização trabalhe de maneira eficiente com suas informações. A pessoa engenheira de dados cria pipelines de dados para integrar, limpar e transformar dados provenientes de várias fontes e formatos. Isso possibilita a geração de insights valiosos para o negócio e melhora a tomada de decisões estratégicas.

☐ **Bancos de dados não relacionais:**

- Os bancos de dados não relacionais, também conhecidos como bancos de dados NoSQL, são uma alternativa aos bancos de dados relacionais tradicionais. Eles são particularmente adequados para lidar com grandes volumes de dados estruturados e não estruturados e são capazes de escalar horizontalmente em vários servidores para suportar cargas de trabalho pesadas. Entre os bancos de dados NoSQL mais populares estão o MongoDB e o Datomic. Neste segmento, você aprenderá os fundamentos dos bancos de dados NoSQL, suas características e usos típicos.

☐ **Estatística e Matemática - Fundamentos:**

- Equações, funções e limites
- Logaritmos

- Matrizes, determinantes, vetores e espaço vetorial
- Derivadas e integrais
- Diferença entre média, mediana e moda
- Distribuição de frequência
- Variância e desvio padrão
- Distribuição binomial, poisson e normal
- Nível e intervalo de confiança
- Técnicas de amostragem
- Introdução à Regressão linear
- Séries temporais

☐ **Governança de Dados:**

- Governança de dados é um conjunto de políticas, processos e práticas que estabelecem a responsabilidade, a integridade, a qualidade e o uso adequado dos dados em uma organização. Ela envolve a definição de regras e padrões para a coleta, armazenamento, gerenciamento, compartilhamento e uso dos dados, garantindo que sejam confiáveis, consistentes e seguros ao longo de sua vida útil.
- A governança de dados desempenha um papel fundamental na engenharia de dados, pois garante que os dados sejam gerenciados de forma eficiente, confiável e em conformidade com as regulamentações e políticas internas. Ela estabelece diretrizes para a gestão dos dados, incluindo a definição de metadados, a identificação de proprietários de dados, a documentação de políticas de acesso e privacidade, a implementação de medidas de segurança e a garantia da qualidade dos dados.
- A governança de dados também promove a colaboração entre as equipes, estabelecendo processos de tomada de decisão baseados em dados confiáveis e padronizados. Isso resulta em uma melhor qualidade das análises, uma base sólida para tomada de decisões estratégicas e uma maior confiança nos dados utilizados para impulsionar as iniciativas de engenharia de dados.

☐ **Proteção de dados:**

- A proteção de dados é o processo de proteger informações importantes de forma que garanta a confidencialidade, integridade e a disponibilidade destes dados.

☐ **R para Ciência de Dados:**

- R é uma linguagem de programação comumente usada em estatística computacional e análise de dados.
- Aprender a analisar, limpar e visualizar dados
- Elaborar gráficos
- Juntar bancos de dados
- Tirar sumários estatísticos
- Aprender modelos preditivos no R