

Introdução à Ciência de Dados

Bacharelado em Administração

Prof. Washington Santos da Silva
IFMG - Campus Formiga

26 de março de 2025

Nesta Aula

Tópicos

- Visão Geral de Ciência de Dados
 - O que é Ciência de Dados?
 - Papéis em Ciência de Dados
- Apresentação do Programa da Disciplina

O que é Ciência de Dados?

Prof. Washington Silva - Introdução à Ciência de Dados



Volume de Dados Gerados

A cada:

- Compra online
- Uso de redes sociais
- Uso do GPS do seu smartphone
- Você produz dados valiosos para a economia do séc. XXI.



Big Data



Capa de Outubro/2012

Prof. Washington Silva - Introdução à Ciência de Dados

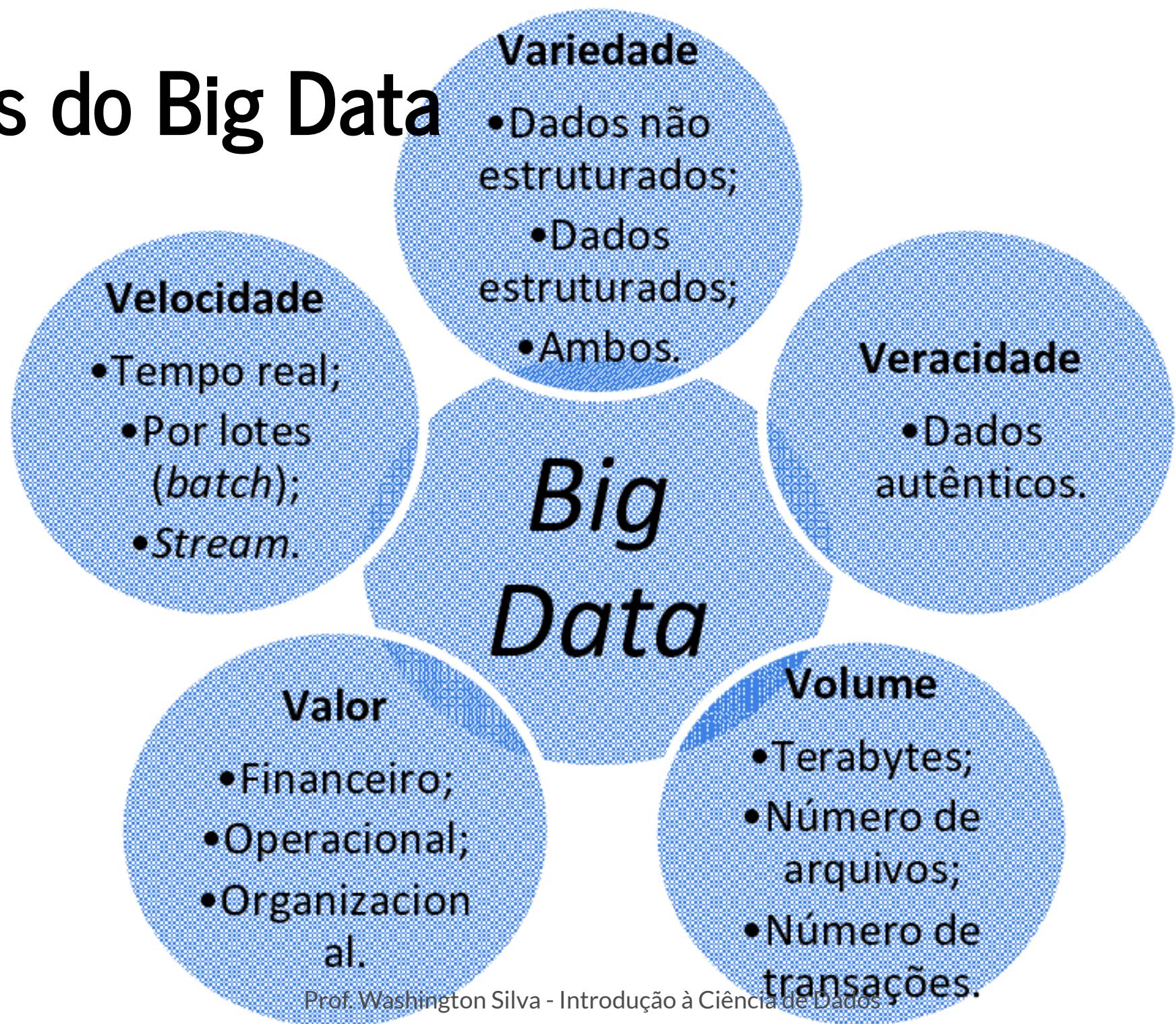
The Economist - Capas Históricas



Dilúvio de Dados - 25/02/2010

O Recurso mais Valioso do Mundo: Dados
- 06/05/2017

5 Vs do Big Data



Sobre a Disciplina

Fundamentos de Ciência de Dados para:

- Coletar, preparar
- Analisar
- e Transformar dados em *Insights*





O que é Ciência de Dados?

Pode parecer intimidador, mas em essência:

Trata-se de uma **abordagem estruturada** para extrair **significado e valor** das enormes quantidades de **dados** gerados diariamente.

Aplicações



Exemplo: compras online

- A cada toque no seu telefone;
- A cada alerta de fraude no seu cartão de crédito;
- A cada recomendação personalizada no seu serviço de streaming.

Tudo isso é Ciência de Dados trabalhando nos bastidores.

Aplicações



Exemplo: compras online

- Os varejistas analisam seu histórico de compras e hábitos de navegação para recomendar aquele presente perfeito ou decidir o melhor momento para enviar uma notificação de promoção.
- Trata-se de usar **dados** para **compreensão** e **previsão** de padrões.

O que é Ciência de Dados?

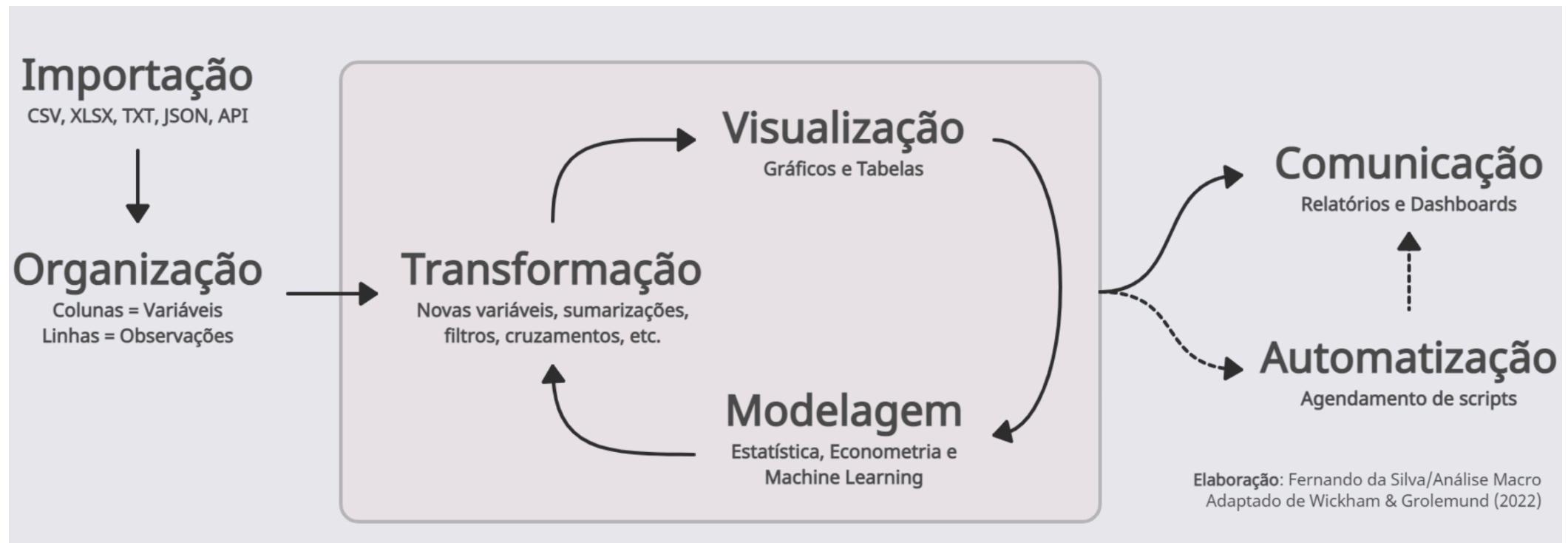


Figura 1: Ciclo de Vida da Análise de Dados

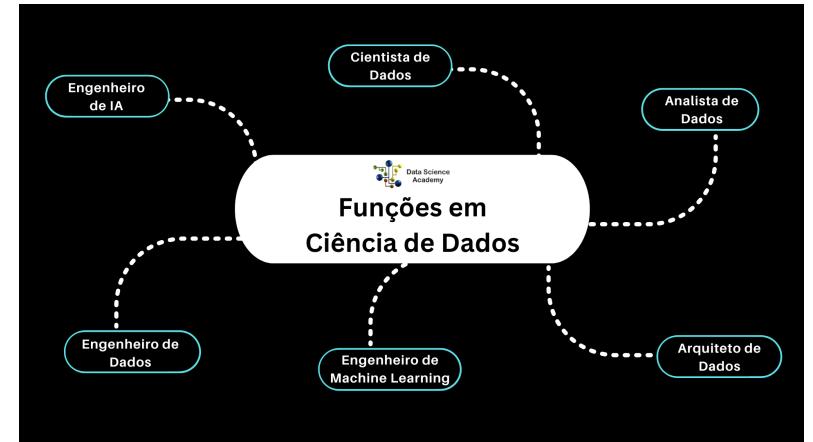
Papéis em Ciência de Dados

Prof. Washington Silva - Introdução à Ciência de Dados

Visão Geral

Papéis em Ciência de Dados

- Analista de Dados/Business Intelligence
- Cientista de Dados
- Engenheiro de Dados
- Arquiteto de Dados
- Cientista de Machine Learning
- Engenheiro de Inteligência Artificial



Analista de Dados

Habilidades Técnicas

- Domínio de ferramentas de visualização de dados como Power BI/Tableau, conhecimento avançado em planilhas eletrônicas, como MS Excel.
- Proficiência em linguagens de programação como R e/ou Python.
- Experiência básica com bancos de dados (SQL).

Habilidades Analíticas

- Capacidade de realizar análises estatísticas e interpretar grandes volumes de dados para identificar tendências e padrões.

Analista de Dados

Exemplos de Aplicações Práticas

- Análise de dados em setores como finanças, saúde e turismo, fornecendo insights para melhorar, por exemplo, a gestão de recursos e o atendimento ao cliente.
- Elaboração de análises **descritivas e diagnósticas**: a análise descritiva busca entender **o que** aconteceu, enquanto a análise diagnóstica, qual o motivo – **o porquê**.

Analista de Business Intelligence

Habilidades Técnicas

- Proficiência em ferramentas de BI como Power BI/Tableau/Qlik, com foco na criação e manutenção de dashboards.
- Conhecimento sólido em bancos de dados e linguagem SQL para extração de dados.
- Familiaridade com ETL (Extração, Transformação e Carga) e data warehousing.
- Conhecimento em Excel avançado e noções de linguagens como R, e/ou Python.

Habilidades Analíticas

- Capacidade de traduzir necessidades de negócio em requisitos técnicos para dashboards e relatórios.
- Habilidade para identificar KPIs relevantes e criar visualizações eficientes.

Analista de BI (Business Intelligence)

Exemplos de Aplicações Práticas

- Desenvolvimento de painéis interativos para acompanhamento de indicadores de desempenho em áreas como vendas, RH e operações.
- Criação de relatórios automatizados para apoiar a tomada de decisão estratégica.
- Foco em análises **preditivas e prescritivas**: identificar o que pode acontecer e sugerir o que deve ser feito.

Analista de BI vs Analista de Dados

Como as funções se confundem

- **Sobreposição de ferramentas:** ambos utilizam ferramentas de visualização como Power BI e conhecimentos em SQL.
- **Foco diferenciado:** enquanto o Analista de Dados concentra-se na exploração e interpretação aprofundada dos dados, o Analista de BI foca na criação de soluções de visualização e monitoramento contínuo.
- **Ciclo de trabalho:** o Analista de BI tende a trabalhar mais em ciclos regulares (relatórios periódicos), enquanto o Analista de Dados frequentemente trabalha em projetos específicos com prazos definidos.
- **Na prática:** em empresas menores, um mesmo profissional frequentemente assume ambos os papéis.

Cientista de Dados (Iniciante)

Habilidades Técnicas

- Proficiência em linguagens de programação como Python e/ou R, com conhecimento de bibliotecas específicas (pandas, tidyverse, scikit-learn, tidymodels).
- Conhecimento em SQL e bancos de dados relacionais e não-relacionais.
- Familiaridade com ferramentas de versionamento (Git) e ambientes de desenvolvimento.

Habilidades Analíticas

- Conhecimentos em estatística aplicada e modelagem matemática.
- Capacidade de preparar e transformar conjuntos de dados para análise e modelagem.
- Capas de Implementar os principais algoritmos de machine learning.

Cientista de Dados (Iniciante)

Exemplos de Aplicações Práticas

- Desenvolvimento de modelos de classificação e regressão para problemas como previsão, segmentação de clientes e detecção de anomalias.
- Análises **preditivas** e **prescritivas**: utilizando dados históricos para prever comportamentos futuros e recomendar ações.
- Criação de provas de conceito (POCs) para validação de hipóteses de negócio baseadas em dados.
- Apresentação de resultados técnicos em formato acessível para partes interessadas não-técnicos.

Outros Papéis em uma Equipe de Dados

Engenheiro de Dados

- Profissional especializado na construção e manutenção de infraestruturas de dados. Responsável por projetar, desenvolver e otimizar* *pipelines* de dados, data warehouses e data lakes.
- Seu foco está em garantir dados de qualidade, acessíveis e escaláveis para as equipes de análise e ciência de dados.

Cientista de *Machine Learning* (Aprendizagem de Máquina)

- Especialista focado no desenvolvimento e implementação de algoritmos de aprendizado de máquina avançados para criar modelos preditivos complexos. Domina técnicas *deep learning*, processamento de linguagem natural e visão computacional.

Habilidades Não Técnicas

Essenciais para Profissionais de Dados

- **Boa comunicação oral e escrita:** Saber explicar resultados técnicos para pessoas sem conhecimento técnico, usando linguagem simples e exemplos práticos.
- **Narrativa com dados:** Capacidade de transformar números e resultados em histórias convincentes que influenciam decisões e destacam os insights mais relevantes.
- **Capacidade analítica:** Abordar problemas de forma estruturada, fazer as perguntas certas e buscar soluções baseadas em evidências.
- **Pensamento crítico:** Questionar suposições, avaliar a qualidade de informações e considerar interpretações alternativas antes de chegar a conclusões.
- **Conhecimento do negócio:** Entender os objetivos da empresa, o mercado e como as análises se conectam às metas organizacionais.

Habilidades Não Técnicas

Essenciais para Profissionais de Dados

- **Trabalho em equipe:** Colaborar efetivamente com colegas de diferentes áreas, entender suas necessidades e integrar diferentes perspectivas.
- **Gerenciamento de projetos:** Planejar etapas de trabalho, estabelecer prazos realistas e comunicar progresso aos interessados.
- **Adaptabilidade:** Lidar bem com mudanças de requisitos, prioridades e tecnologias em um campo em constante evolução.
- **Curiosidade intelectual:** Fazer perguntas relevantes, buscar constantemente novos conhecimentos e ter interesse genuíno em explorar dados.

Cursos Online Recomendados

- Certificado Profissional Google Data Analytics
- Microsoft: Plano de carreira do Analista de Dados
- Microsoft: Treinamento para cientistas de dados

Apresentação do Programa da Disciplina

Prof. Washington Silva - Introdução à Ciência de Dados

Programa da disciplina

Site da Disciplina

https://washingtonsilva.github.io/site_bach_icd2025/