

INTRODUÇÃO À CIENCIA DE DADOS



Por que a ciência de dados não para de crescer em todo mundo? Alguns fatos...

- Em 2020, cada pessoa gerou 1,7 megabytes de dados em apenas um segundo.
- Os usuários da internet geram cerca de 2,5 quintilhões de bytes de dados todo dia.
- Mais de 90% de todos os dados do mundo foram criados nos últimos 2 anos.
- As interações de dados aumentaram 5000% entre 2010 e 2020.
- Hoje uma pessoa levaria aproximadamente 181 milhões de anos para baixar todos os dados da internet.
- O número de dispositivos de IoT pode aumentar para 41,6 bilhões em 2025.
- As Vagas de emprego para funções em ciência de dados alcançaram cerca de 2,7 milhões em todo mundo e aumentarão em cerca de 28% até 2026.
- O mercado de análises de Big Data em Saúde pode valer US\$ 67,82 Bilhões em 2020.

O que esses fatos querem dizer?

- Temos um mercado cada vez mais competitivo e orientado a dados e as empresas precisam sobreviver nesta nova era (Big Data).
- Precisaremos mais do que nunca de processos, ferramentas, técnicas para analisar os dados, conseguindo extrair valor dessa análise e ajudar empresas no processo de tomada de decisões, aumentando seus lucros e reduzindo seus custos.

O que é Ciência de Dados?

- Ciência de Dados (Data Science) é uma ciência que visa estudar as informações, seu processo de captura, transformação, geração e, posteriormente, análise de dados.
- Combina vários campos, incluindo estatísticas, métodos científicos, inteligência artificial (IA) e análise de dados para extrair valor dos dados.
- O Cientista de dados é basicamente um profissional multidisciplinar, que tem como responsabilidade colocar em pratica o processo de Data Science.
- Como estamos vivendo na era do Big Data, a ciência de dados está se tornando cada vez mais necessária para processar e explorar grandes volumes de dados, de diferentes formatos, gerados a partir de várias fontes e em diferentes velocidades.

O que é Ciência de Dados?

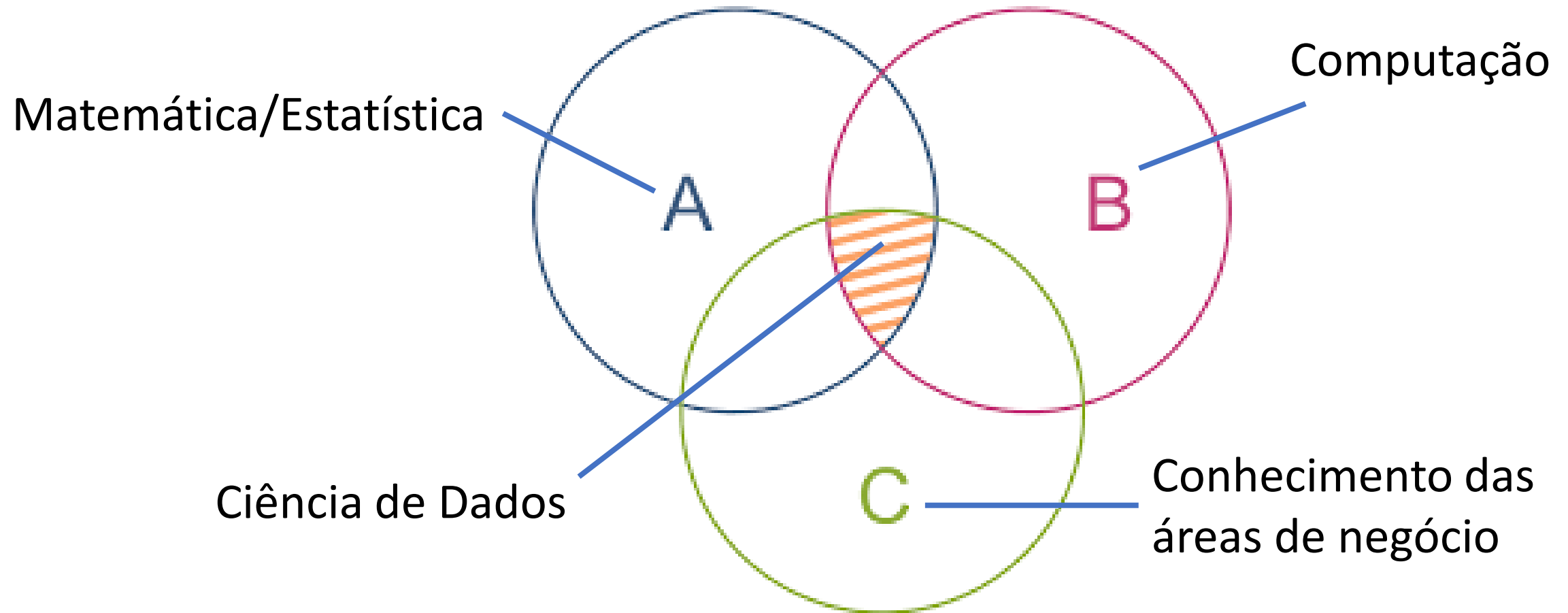
- Disciplinas/áreas do conhecimento envolvidas:
- Estatística
- Matemática
- Programação/Computação
- Conhecimento de Negócios
- Modelagem
- Análise Preditiva
- Mineração e visualização de Dados

O que é Ciência de Dados?

- Na verdade, a Ciência de Dados existe há bastante tempo.
- Somente agora seu potencial foi percebido e isso se deve em parte ao Big Data e a capacidade computacional.
- Big Data é o verdadeiro combustível para a Ciência de Dados (surgiu há mais ou menos uma década atrás)
- Objetivo: extrair e interpretar dados de forma eficaz e apresentá-los em uma linguagem simples para usuários finais e tomadores de decisão.

O que é Ciência de Dados?

- **Áreas do conhecimento**



O que é Ciência de Dados?

- Resumindo...
- A Ciência de Dados envolve o uso de métodos automatizados (Ciência da Computação) para analisar (Matemática/Estatística) enormes quantidades de dados a fim de extrair conhecimento (áreas de negócio) a partir dos dados.

Dados e Tomada de Decisão

• DADOS -> DECISÕES -> AÇÕES

EXEMPLO:

Problema Identificado: Clientes estão cancelando os planos de TV por assinatura nos últimos anos.

1- Dados e Análise: Dados Coletados e analisados, levaram a conclusão que houve um número muito grande de cancelamentos Nos planos de TV por assinatura das operadoras nos últimos 5 anos

2- Decisão: Os executivos da empresa, baseados em dados, decidiram revisar todo o conteúdo dos pacotes de TV por assinatura oferecidos em seu portfólios

3- Ação: As Operadoras de TV passaram a oferecer pacotes de Streaming em seus portfólios, criaram combos promocionais com acesso gratuito a alguns players de filmes.

Dados e Tomada de Decisão

