**BIG DATA**

Neste exato momento, 2.5 quintilhões de bytes são gerados por dia para nortear indivíduos, empresas e governos, e está dobrando a cada dois anos.

Cerca de 90% de todos os dados gerados no planeta, foram gerados nos últimos 2 anos.

Aproximadamente 80% dos dados são **não-estruturados** ou estão em diferentes formatos, que dificulta a análise

Toda vez que fazemos uma compra, uma ligação ou interagimos nas redes sociais, estamos produzindo esses dados.

Com a recente conectividade em objetos, tal como relógios, carros e até geladeiras, as informações capturadas se tornam massivas e podem ser cruzadas para criar modelos preditivos cada vez mais elaborados, apontando e, até prevendo, o comportamento de empresas e clientes.

—------------------------------------------------------------------------------------------------------------------------

BIG DATA é uma coleção de conjuntos de dados, grandes e complexos, que não podem ser processados por bancos de dados ou aplicações de processamento tradicionais

**o alto volume de dados gerado em alta velocidade e com muita variedade, isso é o que define BIG DATA**

O google estima que a humanidade criou nos últimos 5 anos, o equivalente a 300 exabytes de dados ou seja: 300.000.000.000.000.000.000

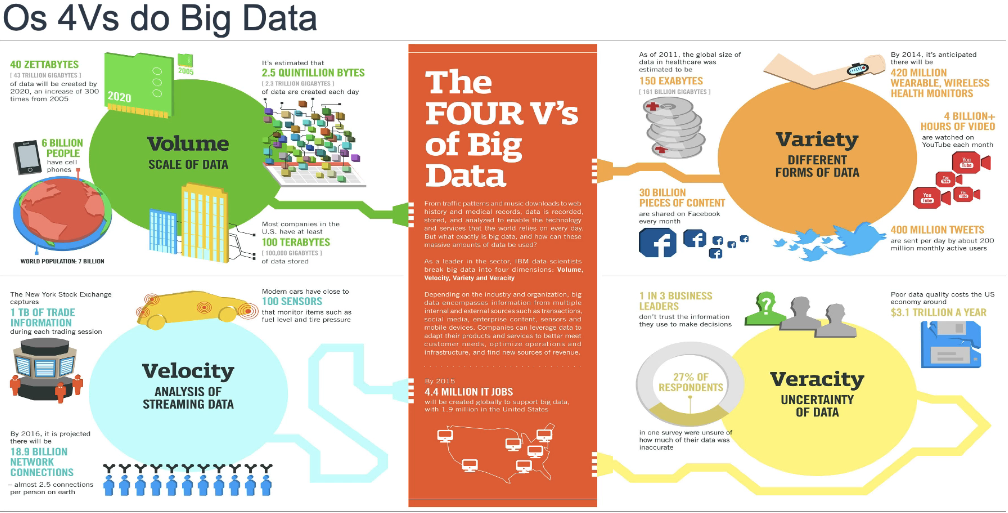
Podemos definir o conceito de BIG DATA como sendo conjuntos de dados extremamente amplos e que, por este motivo, necessitam de ferramentas especialmente preparadas para lidar com grandes volumes, velocidade e variedade, de forma que toda e qualquer informação disponível nos dados possa ser encontrada, analisada e aproveitada em tempo hábil

Análise de grades quantidades de dados para geração de resultados importantes que, em volumes menores dificilmente seriam alcançados

O big data nos dá uma visão clara do que é granular, não temos de nos fixar na causalidade, Podemos descobrir padrões e correlações nos dados que nos propiciem novas e valiosas ideias

—------------------------------------------------------------------------------------------------------------------------

**Os 4Vs do Big Data:**

****

Volume, Variedade, Velocidade e Veracidade

para ser chamado de big data é preciso conter os 4 Vs em harmonia

**VOLUME (tamanho dos dados):**

* Espera-se que 40zettabytes de dados sejam criados por ano no mundo
* Cerca de 2.5 quintillionbytes de dados são criados por dia
* Existem atualmente cerca de 6 bilhões de telefones móveis no planeta
* Cada empresa americana armazena cerca de 100 Terabytes de dados.

**VARIEDADE (formato dos dados):**

* 150 exabytes é a estimativa de dados que foram gerados especificamente para tratamento de casos de doença em todo o mundo por ano desde 2011.
* Mais de 4 bilhões de horas por mês são usadas para assistir vídeos no YOUTUBE
* 30 bilhões de imagens são publicadas por mês no Facebook
* 200 milhões de usuários ativos por mês, publicam 400 milhões de tweets por dia

**VELOCIDADE (geração de dados):**

* 1 Terabyte de informação é criada durante uma única sessão da bolsa de valores Americana, a New York Stock Exchange (NYSE)
* Aproximadamente 100 sensores estão instalados nos carros modernos para monitorar nível de combustível, pressão dos pneus e muitos outros aspectos do veículo
* 18.9 bilhões de conexões de rede já existem no mundo

**VERACIDADE (confiabilidade dos dados):**

* Atualmente 1 em cada 3 gestores tem experimentado problemas relacionados a veracidade dos dados para tomar decisões de negócios
* Além disso, estima-se que 3.1 trilhões de dólares por ano sejam desperdiçados devido a problemas de qualidade dos dados.

—------------------------------------------------------------------------------------------------------------------------

**BIG DATA x CIÊNCIA DE DADOS**

* Big Data é a matéria-prima, ou seja, dados.
* Ciência de dados é um conjunto de técnicas para análise de dados.
* Quando aplicamos ciência de dados ao big data extraímos valor e então temos o que é chamado de **Big Data Anaytics**.

—------------------------------------------------------------------------------------------------------------------------

**EXEMPLOS DE APLICAÇÃO DO BIG DATA ANALYTICS**

* Uma companhia Aérea pode extrair, armazenar, processar e analisar dados de viagens dos passageiros a fim de oferecer rotas com maior probabilidade de venda.
* Uma rede de supermercados pode extrair, armazenar, processar e analisar dados de compras a fim de detectar padrões e organizar os produtos de forma a aumentar as vendas
* Uma rede de hotéis pode extrair, armazenar, processar e analisar dados de comentários de clientes em redes sociais para customizar seus serviços, aumentar as vendas e reduzir custos.
* Uma rede de hospitais pode extrair, armazenar, processar e analisar dados de exames médicos a fim de personalizar e otimizar o atendimento dos pacientes

A parte mais relevante é o valor que voce extrai dos dados

1024 Kilobytes = 1 Megabyte

1024 Megabytes = 1 Gigabyte

1024 Gigabytes = 1 Terabyte

1024 Terabytes = 1 Petabyte

1024 Petabytes = 1 Exabyte

1024 Exabytes = 1 Zettabyte

1024 Zettabytes = 1 Yottabyte [Facebook e Google estão aqui]

1024 Yottabytes = 1 Brontobyte

1024 Brontobytes = 1 Geobyte

Referência e Bibliografia

Big Data

[https://www.amazon.com.br/Big-Data-Cezar-Taurionebook/dp/B00ZQ0O5DC/ref=sr\_1\_fkmr0\_1?ie=UTF8&qid=1458517512&sr=8-1- fkmr0&keywords=big+data+cesar+taurion](https://www.amazon.com.br/Big-Data-Cezar-Taurionebook/dp/B00ZQ0O5DC/ref=sr_1_fkmr0_1?ie=UTF8&qid=1458517512&sr=8-1-)

Big Data. Como Extrair Volume, Variedade, Velocidade e Valor da Avalanche de Informação Cotidiana

[https://www.amazon.com.br/Variedade-Velocidade-Avalanche-Informa%C3%A7%C3%A3oCotidiana/dp/8535270906/ref=sr\_1\_1?ie=UTF8&qid=1458517546&sr=8- 1&keywords=big+data+como+extrair+volume+variedade+velocidade+e+valor+da+avalanche+d e+informa%C3%A7%C3%A3o+cotidiana](https://www.amazon.com.br/Variedade-Velocidade-Avalanche-Informa%C3%A7%C3%A3oCotidiana/dp/8535270906/ref=sr_1_1?ie=UTF8&qid=1458517546&sr=8-)

4 Princípios Para Extrair Valor do Big Data

<https://blog.dsacademy.com.br/4-principios-para-extrair-valor-do-big-data>