Big data:

1-historia

2-como funciona

3-aplicabilidade

4-comi está o mercado atual

5-futuro

Big data

Big Data descreve um imenso volume de dados, estruturados ou não, que impactam os negócios no dia a dia. A quantidade não importa, mas sim o que as empresas fazem com os dados importantes dentre esses. Big Data pode ser analisado para a obtenção de insights que levam a melhores decisões e direções estratégicas de negócio.

Um pouco da história

O termo é novo, porém o conceito de recolher e armazenar enormes quantidades de dados para avaliações posteriores é antigo. Esse nome “Big Data”, ganhou força em 2000, quando o analista Doug Laney, articulou 3 Vs: volume, velocidade e variedade.

* Volume: coleta de informações de uma variedade gigante de fontes, isso inclui redes sociais, transações, ou transmissões de maquina a máquina feitas por sensores.
* Velocidade: os dados fluem numa velocidade muito alta e sem dificuldades
* Variedade: os dados são gerados em diferentes formatos, numéricos, estruturados, e-mails, vídeos, transações financeiras.

O potencial do Big data

É quase impossível contar quantos dados estão sendo criados e armazenados no mundo todo, e esta quantidade só aumenta.

Como funciona?

Antes de descobrir como big data pode favorecer o seu negócio, você deve primeiro entender de onde isso vem. As fontes de big data geralmente caem em uma das três categorias:

Transmissão de dados (streaming data): Inclui dados que chegam aos seus sistemas de TI a partir de uma rede de dispositivos conectados. Você pode analisar esses dados ao ponto em que eles chegam e tomar decisões sobre quais dados você deve manter, não manter e quais requerem uma análise posterior mais aprofundada.

Dados de redes sociais: Os dados sobre interações sociais são um conjunto cada vez mais atraente de informação, particularmente para marketing, vendas e funções de apoio. São muitas vezes capturados de formas não estruturada ou semi-estruturada, por isso representam um desafio único quando se trata de consumo e análise.

Fontes publicamente disponíveis: Enormes quantidades de dados estão disponíveis por meio de canais públicos como o dados.gov.br do governo federal brasileiro, o livro de fatos da CIA (World Factbook) ou o Portal de Dados Abertos da União Européia

Após identificar todas as potenciais fontes de dados, considere as decisões que você precisará tomar uma vez que você começar a fazer uso destas informações. Estas incluem:

Como armazenar e gerenciar: Embora o armazenamento de dados teria sido um problema anos atrás, agora há opções de baixo custo para armazenamento de dados caso esta seja a melhor estratégia para seu negócio.

Quanto se deve analisar: Algumas organizações não excluem quaisquer dados de suas análises, o que é possível com as tecnologias atuais de alto desempenho, como a computação em grade (grid computing) ou análises em memória (in-memory analytics). Outra abordagem é determinar antecipadamente quais dados são relevantes antes de analisá-los.

Como utilizar os insights descobertos: Quanto mais conhecimento você tiver, mais confiante você estará para tomar suas decisões de negócio. É importante ter uma estratégia desenhada para quando você tiver uma abundância de informações em mãos.

O passo final para fazer big data trabalhar a favor do seu negócio é procurar tecnologias que o ajudarão a tirar melhor proveito da enorme quantidade de informações e das análises de big data. Considere:

* Armazenamento barato de grandes quantidades de dados;
* Processadores mais rápidos;
* Open source acessível e plataformas distribuidas de big data, como o Hadoop;
* Processamento paralelo, clusterização, MPP, virtualização, grandes ambientes de grid, alta conectividade e altas taxas de transferência;
* Computação em nuvem (cloud computing) e outros arranjos de alocação flexível de recursos.

Por que Big Data é importante?

A importância do big data não é quantidade de dados, mas o que você faz com eles. Quando se combina big data com a alta potência do analytics, você pode realizar tarefas relacionadas a negócios:

*Você pode analisar dados de qualquer fonte para encontrar respostas que permitam 1) redução de custos, 2) redução de tempo, 3) desenvolvimento de novos produtos e ofertas otimizadas, 4) decisões mais inteligentes.* Determinar a causa raiz de falhas, problemas e defeitos em tempo quase real;

* Gerar cupons no ponto de venda com base em hábitos de compra dos clientes;
* Recalcular carteiras de risco inteiras, em questão de minutos;
* Detectar comportamentos fraudulentos antes que eles afetem sua organização.

Quais tecnologias posso usar sem custo?

Hadoop: plataforma Open Source, que roda num servidor de internet chamado Apache, ele precisa de APIs, que se encarregam de fazer o trabalho para movimentar dados entre uma base de dados que você possui e o Hadoop.

Soluções de Big Data

Lida com volumes muito grandes de dados variados, as soluções precisam trabalhar com distribuição de processamento e elasticidade, é o que suporta aplicações com grande volume de dados e que além disso, crescem em pouco tempo.

Os bancos de dados que exploram o modelo relacional, como o MySQL e o Oracle, não se mostram adequados a estes requisitos, pois são menos flexíveis, quanto maior um banco de dados se torna, mais custoso e trabalhoso ele fica: é preciso otimizar, acrescentar novos servidores, empregar mais especialistas em sua manutenção.

. O bancos de dados relacionais se baseiam em quatro propriedades que tornam a sua adoção segura e eficiente: Atomicidade, Consistência, Isolamento e Durabilidade.

Atomicidade: toda transação deve ser atômica, isto é, só pode ser considerada efetivada se executada completamente;

Consistência: todas as regras aplicadas ao banco de dados devem ser seguidas;

Isolamento: nenhuma transação pode interferir em outra que esteja em andamento ao mesmo tempo;

Durabilidade: uma vez que a transação esteja concluída, os dados consequentes não podem ser perdidos.

Big Data no mercado America Latina

O mercado total de big data e analytics (BDA) movimentou US$ 2,48 bilhões na América Latina no ano passado. Liderado pelo Brasil e México e impulsionado pela transformação digital, o mercado deverá atingir US$ 7,41 bilhões em 2022, de acordo com projeção da.

Segundo relatório da consultoria, em 2016, o Brasil é o país mais maduro em aplicações de big data e analytics e representou 46,8% do mercado na América Latina, gerando receita de US$ 1,16 bilhão. Além disso o estudo aponta a IBM, Oracle, SAP, SAS e Teradata como líderes de mercado, sendo que as verticais que mais investem em big data e analytics são serviços financeiros, varejo e telecomunicações. O mercado latino americano representa aproximadamente 5,1% do mercado global de BDA.

A Frost & Sullivan avalia que modelos de negócios inovadores como Internet das Coisas (IoT) e computação em nuvem estão transformando o mercado e criando novas formas de coletar e melhorar os processos de armazenamento de dados. Além disso, diz que as empresas da região estão cada vez mais familiarizadas com os conceitos e benefícios da adoção de soluções de big data e analytics.

Case de sucesso: Pinterest

O Pinterest, rede social de compartilhamento de imagens, chegou recentemente a mais de 150 milhões de usuários ativos mensais. Um de seus trunfos para o sucesso é a utilização de Big Data e machine learning (aprendizado de máquina) para personalizar ao máximo a experiência de seus usuários.

A equipe técnica do Pinterest vem refinando o sistema de recomendação da rede ano após ano. Para alcançar este objetivo, seus cientistas de dados têm implementado inteligência artificial, um sistema de ranking mais rápido e conteúdo localizado.