

# Big Data Analytics

Conceitos relacionados a Big Data e Analytics



# Agenda

- Compreender o contexto de Big Data;
- Ferramentas;
- Big Data;
- Big Data Analytics;
- Analytics e as empresas.



# Compreender o contexto de Big Data

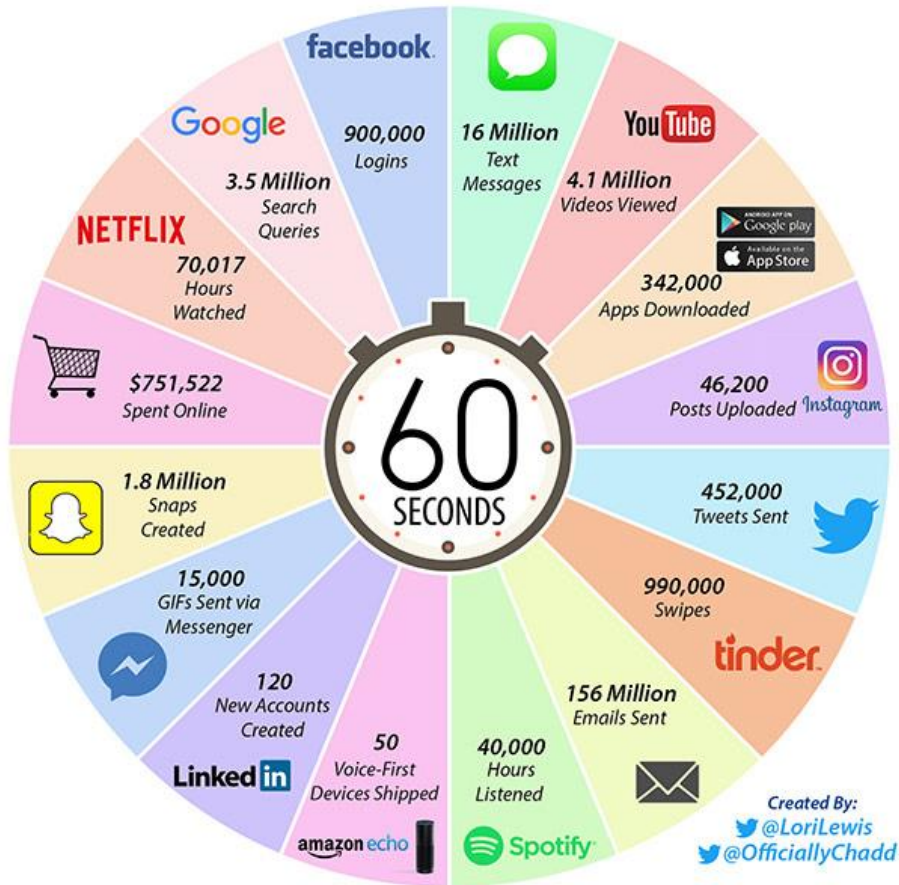


# Introdução

- Alguma vez você já pensou na quantidade de dados que geramos?



# 2017 *This Is What Happens In An Internet Minute*



# 2018 *This Is What Happens In An Internet Minute*



# Introdução

- Foi desse grande volume de dados gerados por todos nós que surgiu o que popularmente chamamos de Big Data.
- No entanto, não estamos falando apenas de mídias sociais!



# Introdução

- Diariamente as empresas geram inúmeras e diversas informações em sistemas ERP (*Enterprise Resource Planning* - Planejamento dos Recursos da Empresa) através de relatórios de vendas, financeiros, inventários, SAC, máquinas equipadas com sensores, dados de geolocalização, etc.
- A maioria desses dados foram gerados recentemente. Nos últimos 3 anos, produzimos mais dados que em todo o resto da história da humanidade.



# Introdução

- Cerca de 90% dos dados gerados pela humanidade foram gerados nesses últimos três anos.
- Você sabia que a cada dia são gerados 2,5 quintilhões de bytes (um quintilhão é o número 1 seguido de 18 zeros)?





# Introdução

- Dependendo do modo como esses dados são captados, seu formato pode variar muito.
- Dividimos os dados então em três tipos:
  - Estruturados;
  - Semiestruturados;
  - Não-estruturados.



# Introdução

- Dados estruturados
  - Dados organizados em blocos semânticos (relações).
  - Dados de um mesmo grupo possuem as mesmas descrições (atributos).
  - Descrições para todas as classes de um grupo possuem o mesmo formato (esquema).
  - Dados mantidos em um SGBD são chamados de **dados estruturados** por manterem a mesma estrutura de representação (rígida), previamente projetada (esquema).



# Introdução

- Dados semiestruturados
  - Atualmente, muitos dados não são mantidos em SGBDs.
  - Dados Web, por exemplo, apresentam uma organização bastante heterogênea e esta heterogeneidade dificulta as consultas a estes dados.
  - Assim, estes dados são atualmente classificados como **dados semiestruturados**, ou seja:
    - Não são estritamente “tipados”;
    - Não são completamente não-estruturados.



# Introdução

- Dados semiestruturados
  - São dados onde o esquema de representação está presente de forma explícita ou implícita.
  - É auto descritivo.
  - Uma análise dos dados deve ser feita para que a sua estrutura possa ser identificada e extraída.

# Introdução

- Dados semiestruturados
  - Características principais
    - Definição à posteriori
      - Esquemas são definidos após a existência dos dados;
      - Investigação de suas estruturas particulares.
    - Estrutura irregular
      - Não existe um esquema padrão para os dados;
      - Coleções de dados são definidos de maneiras diferentes, contendo informações incompletas.
    - Estrutura implícita
      - Muitas vezes existe uma estrutura implícita.
    - Estrutura parcial
      - Apenas parte dos dados disponíveis podem ter uma estrutura.



# Introdução

- Dados semiestruturados
  - Exemplos
    - XML – eXtensible Markup Language;
    - JSON – JavaScript Object Notation.



# Introdução

- Dados semiestruturados
  - Comparativo

Dados Estruturados	Dados SemiEstruturados
Esquema pré-definido	Nem sempre há um esquema
Estrutura regular	Estrutura irregular
Estrutura independente dos dados	Estrutura embutida nos dados
Estrutura reduzida	Estrutura extensa (particularidades de cada dado, visto que cada um pode ter uma organização própria)
Fracamente evolutiva	Fortemente evolutiva (estrutura modifica-se com frequencia)
Prescritiva (esquemas fechados e restrições de integridade)	Estrutura descritiva
Distinção entre estrutura e dados é clara	Distinção entre estrutura e dados não é clara

# Introdução

- Dados não-estruturados
  - São dados que não possuem uma estrutura definida.
  - Normalmente caracterizados por documentos, textos, imagens, vídeos e outros formatos.
  - As estruturas não são descritas, nem implicitamente.
  - A grande maioria dos dados atuais na Web e nas empresas seguem este formato.





# Introdução

- Dados não-estruturados
  - Atualmente, devido a variedade de dispositivos, os dados também são variados:
    - Sensores, dispositivos inteligentes, tecnologias de colaboração, redes sociais, IOT (*Internet Of Things* – Internet das Coisas).
  - Dados não são mais relacionais, mas diversificados de paginas Web, e-mails, documentos, dados de sensores, etc.
  - Sistemas tradicionais não têm a capacidade de processar estes dados.



# Introdução

- O conceito de Big Data está ligado à essa infinidade de dados, que necessitam de ferramentas especiais para serem coletados, armazenados e analisados.
- Apesar de algumas vezes tratarmos como uma nova tecnologia, Big Data na verdade diz respeito à capacidade de se trabalhar com todos esses dados.



# Introdução

- Autonomy explains just how large 'big data' is (infographic).



# Introdução

- Big Data and the Future of Your Health.



# Introdução

- What is exactly the Internet Of Things?



# Introdução

- Understanding the Internet of Things:  
Towards a Smart Planet.



# Introdução

- The world's most valuable resource is no longer oil, but data (2017).
- Here's Why Data Is Not The New Oil (2018).



# Ferramentas





# Ferramentas

- Precisamos considerar três tipos de ferramentas no trabalho com dados.
  - Ferramentas de armazenamento e gerenciamento;
  - Ferramentas de limpeza de dados;
  - Ferramentas de mineração de dados.
- Seguem alguns exemplos.



# Ferramentas de armazenamento e gerenciamento

- Hadoop
  - Como não poderia deixar de ser, uma das primeiras coisas que vem à mente quando falamos de Big Data é o Hadoop.
  - Trata-se de um ecossistema *open source* para armazenamento distribuído em *clusters* computacionais.
  - História.



# Ferramentas de armazenamento e gerenciamento

- MongoDB
  - Uma alternativa ao banco de dados relacional, o MongoDB é um banco NoSQL que possui um esquema mais flexível para armazenamento de dados.
  - História.



# Ferramentas de limpeza de dados

- Antes de fazer qualquer trabalho com os dados é preciso fazer a limpeza dos mesmos. Isso porque muitas vezes os dados não estarão prontos para serem utilizados: principalmente dados da internet (mídias sociais, por exemplo).



# Ferramentas de limpeza de dados

- Data Cleaner
  - É uma ferramenta para análise de qualidade dos dados.
  - Permite limpeza, transformação, correspondência e fusão de dados (além de outras funcionalidades).
  - Possui acesso às diversas bases de dados, como: Oracle, MySQL, arquivos CSV, MongoDB, MongoDB, etc. Possui também integração com Pentaho e Salesforce, por exemplo.



# Ferramentas de limpeza de dados

- Open Refine
  - Ferramenta *open source* para trabalhar com limpeza e transformação de dados.
  - Permite trabalhar com grandes volumes de dados mesmo se não forem totalmente estruturados.



# Ferramentas de mineração de dados (*data mining*)

- Permitir análises a partir dos dados é o objetivo de todo projeto de Big Data Analytics.



# Ferramentas de mineração de dados (*data mining*)

- RapidMiner
  - Desenvolvida em linguagem JAVA, é uma ferramenta interessante para fazer análise preditiva, que permite que você integre seus próprios algoritmos via API (*Application Programming Interface*).
  - Proporciona uma ambiente integrado para se trabalhar também com aprendizado de máquina, *business analytics*, *text analytics*, etc.





# Ferramentas de mineração de dados (*data mining*)

- Orange
  - Permite trabalhar com data mining por meio de scripts Python.
  - Também é conhecido por seus algoritmos de aprendizado de máquina e *data analysis*.
  - Usado no trabalho com modelos preditivos, sistemas de recomendação, dentre outros.



# Ferramentas de mineração de dados (*data mining*)

- Rattle GUI
  - Software gratuito e *open source* que fornece uma interface gráfica para data mining.
  - Utiliza a linguagem estatística R.



# Ferramentas de mineração de dados (*data mining*)

- Apache Mahout
  - Projeto da Apache Software Foundation com o objetivo de criar algoritmos de aprendizagem de máquina, principalmente no que diz respeito à filtragem colaborativa, agrupamento (*clustering*) e classificação.



# Ferramentas de mineração de dados (*data mining*)

- Weka;
- Knime;
- Sci-kit (Python).



# Big Data



# Big data

- Muito se fala de Big data como uma tecnologia revolucionária, principalmente para os negócios.
- Apesar de alguns exageros, de fato o Big data tem afetado a maneira que gestores e empresas atuam.
- Se afeta os negócios, merece atenção!



# Big data

- Big data é um termo que surgiu devido à grande quantidade de dados gerados por sistemas internos de empresas, redes sociais, e muito mais.
- Todos esses dados gerados, quando analisados, podem fornecer informações jamais imaginadas.



# 5-Vs

- Para entender o que é Big data, é preciso entender o conceito dos 5vs do Big data e como essa estrutura pode auxiliar sua empresa à alavancar negócios e estabelecer procedimentos eficientes.





# 5-Vs

- A proposta de uma solução de Big data é oferecer uma abordagem consistente no tratamento do constante crescimento e da complexidade dos dados.
- Para tanto, o conceito considera os 5-Vs do Big data: o Volume, a Velocidade, a Variedade, a Veracidade e o Valor.



# 5-Vs



# 5-Vs

- **Volume**

- O conceito de volume no Big data é melhor evidenciado pelos fatos do cotidiano: troca de e-mails, transações bancárias, interações em redes sociais, registro de chamadas e tráfego de dados em linhas telefônicas.
- Todos esses servem de ponto de partida para a compreensão do volume de dados presentes no mundo atualmente.



# 5-Vs

- **Volume**

- Estima-se que o volume total de dados que circulam na internet é de 250 Exabytes por ano (Inmoment, 2014).
- Todos os dias são criados 2,5 quintilhões de bytes em forma de dados.
- É importante também compreender que o conceito de volume é relativo à variável tempo, ou seja, o que é grande hoje, pode não ser nada amanhã (Ohlhorst, 2012).
- Nos anos 90, um Terabyte era considerado Big data.
- Em 2015, teremos no mundo aproximadamente um volume de informação digital de 8 Zettabytes, um valor infinitamente maior (IBM).

# 5-Vs

- **Velocidade**

- Você cruzaria uma rua vendendo se a última informação que tivesse fosse uma fotografia tirada do tráfego circulante de 5 minutos atrás?
- Provavelmente não, pois a fotografia de 5 minutos atrás é irrelevante, você precisa saber das condições atuais para poder cruzar a rua em segurança. (Forbes, 2012)
- A mesma lógica se aplica a empresas, pois necessitam de dados em atuais sobre seu negócio, ou seja, velocidade.
- Segundo Taurion (2014) a importância da velocidade é tamanha que em algum momento deverá existir uma ferramenta capaz de analisar dados em tempo real.



# 5-Vs

- **Velocidade**

- Atualmente, os dados são analisados somente após serem armazenados, mas o tempo gasto para o armazenamento em si já desclassifica esse tipo de análise como uma análise 100% em tempo real.
- Informação é poder (The Guardian, 2010), e assim sendo a velocidade com a qual você obtém essa informação é uma vantagem competitiva das empresas.



# 5-Vs

- **Velocidade**

- A velocidade pode limitar a operação de muitos negócios, quando utilizamos o cartão de crédito por exemplo, se não obtivermos uma aprovação da compra em alguns segundos normalmente pensamos em utilizar outro método de pagamento. É a operadora perdendo uma oportunidade de negócios pela falha na velocidade de transmissão e análise dos dados do comprador.



# 5-Vs

- **Variedade**

- O volume é apenas o começo dos desafios dessa nova tecnologia. Se temos um volume enorme de dados, também obtemos a variedade dos mesmos.
- Já pensou na quantidade de informações dispersas em redes sociais? Facebook, Twitter entre outros possuem um vasto e distinto campo de informações sendo ofertadas em público a todo segundo.
- Podemos observar a variedade de dados em e-mails, redes sociais, fotografias, áudios, telefones e cartões de crédito (McAfee et al, 2012).





# 5-Vs

- **Variedade**

- Seja qual for a discussão, podemos obter infinitos pontos de vista sobre a mesma. Empresas que conseguem captar a variedade, seja de fontes ou de critérios, agregam mais valor ao negócio (Gartner).
- Dentre as três categorias de dados que já conhecemos (estruturados, semiestruturados e não-estruturados), estima-se que até 90% de todos os dados no mundo estão a forma de dados não estruturados. (ICD, 2011).



# 5-Vs

- **Veracidade**

- Um em cada três líderes não confiam nos dados que recebem (IBM).
- Para colher bons frutos do processo do Big data é necessário obter dados verídicos, de acordo com a realidade.



# 5-Vs

- **Veracidade**

- O conceito de velocidade, já descrito, é bem alinhado ao conceito de veracidade pela necessidade constante de análise em tempo real, isso significa, de dados que condizem com a realidade daquele momento, pois dados passados não podem ser considerados dados verídicos para o momento em que é analisado.
- A relevância dos dados coletados é tão importante quanto o primeiro conceito. A verificação dos dados coletados para adequação e relevância ao propósito da análise é um ponto chave para se obter dados que agreguem valor ao processo. (Hurwitz, Nugent, Halper & Marcia Kaufman).



# 5-Vs

- **Valor**

- Quanto maior a riqueza de dados, mais importante é saber realizar as perguntas certas no início de todo processo de análise (Brown, Eric, 2014).
- É necessário estar focado para a orientação do negócio, o valor que a coleta e análise dos dados trará para o negócio.



# 5-Vs

- **Valor**

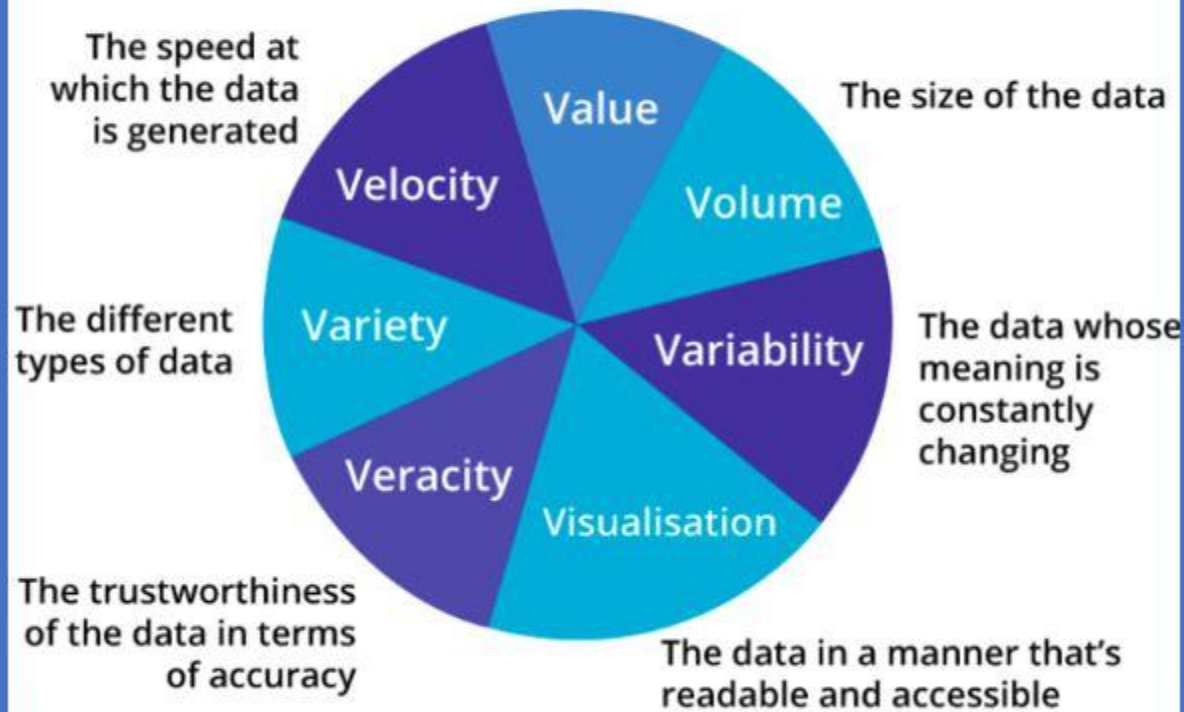
- Não é viável realizar todo o processo de Big data se não se tem questionamentos que ajudem o negócio de modo realista.
- Da mesma forma é importante estar atento aos custos envolvidos nessa operação, o valor agregado de todo esse trabalho desenvolvido, coleta, armazenamento e análise de todos esses dados tem que compensar os custos financeiros envolvidos (Taurion, 2013).



# 5-Vs

## The 7 Vs OF BIG DATA

Just having Big Data is of no use  
unless we can turn it into value



# 7-Vs

- **Variabilidade**

- Mudança constante no contexto e no significado dos dados.

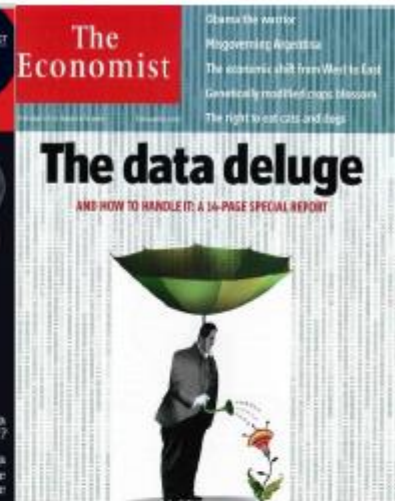
- **Visualização**

- Capacidade de apresentar resultados de forma clara, concisa e precisa.





# Big data na imprensa





# Por que Big Data?

## Why Big Data?



# Google Trends

big data  
Termo de pesquisa

hadoop  
Termo de pesquisa

+ Adicionar comparação

Estados Unidos

2004 - presente

Todas as categorias

Pesquisa na Web

Interesse ao longo do tempo



# Google Trends



Google Trends

Comparar



big data

Termo de pesquisa



hadoop

Termo de pesquisa



analytics

Termo de pesquisa



machine learning

Termo de pesquisa



Estados Unidos ▼

2004 - presente ▼

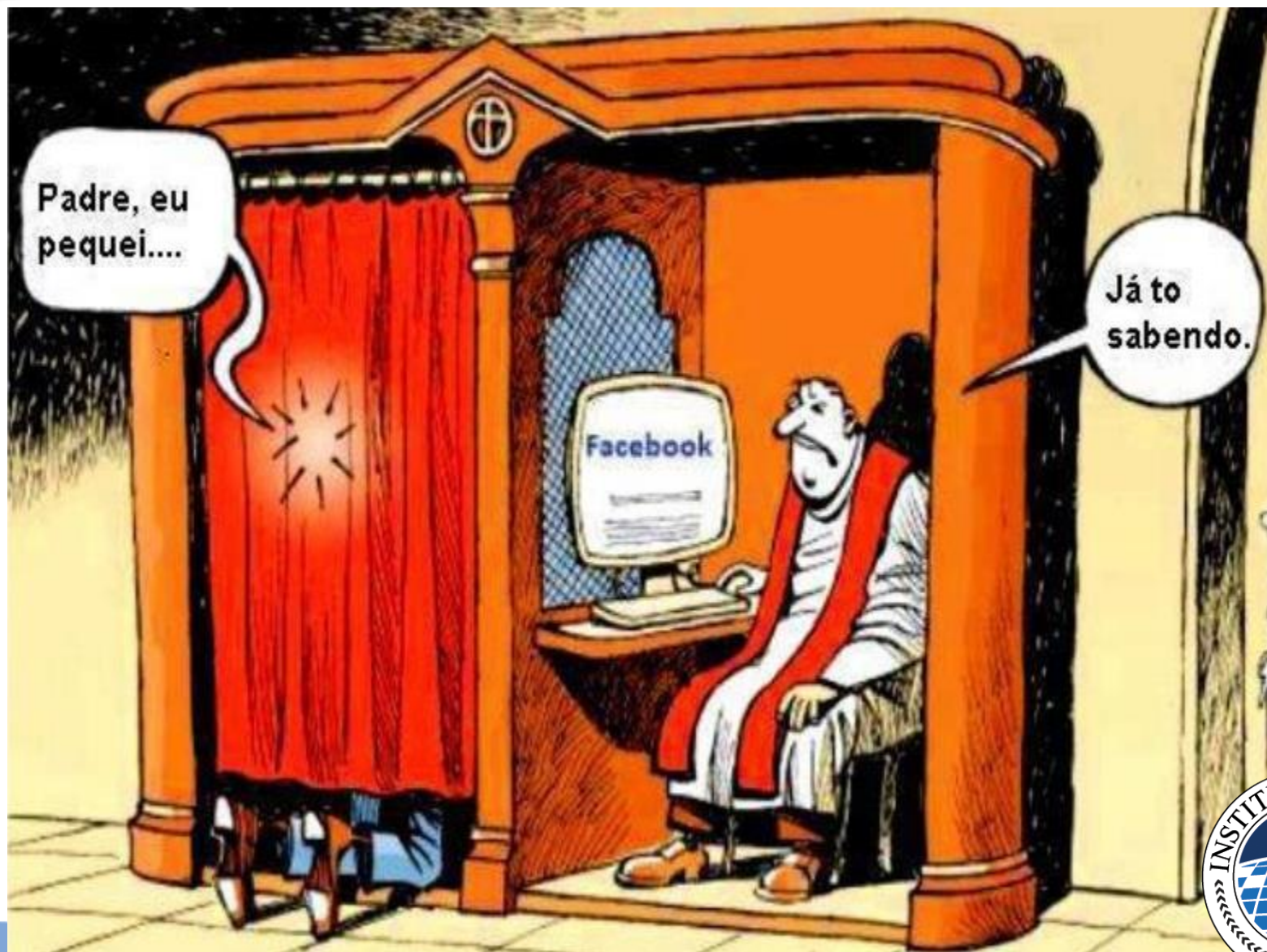
Todas as categorias ▼

Pesquisa na Web ▼

Interesse ao longo do tempo



# Disponibilidade de dados





# Um mundo mais conectado



# Revisitando a pirâmide de Maslow



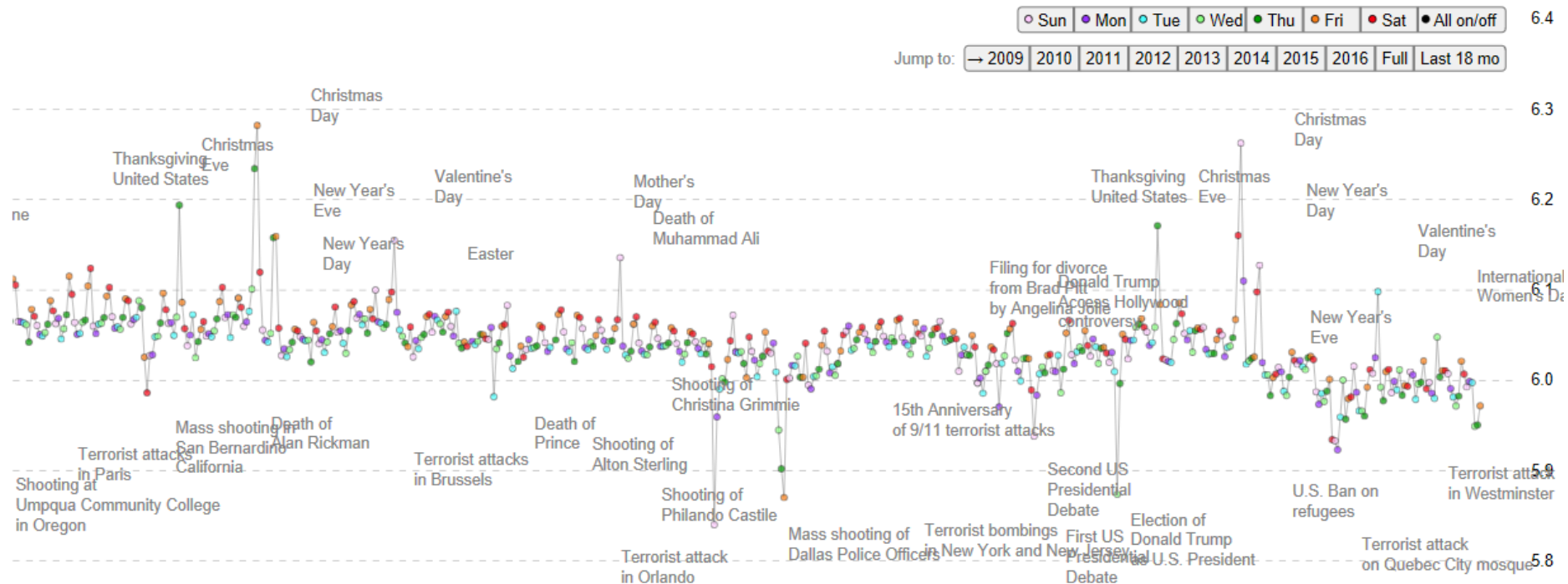


# Aumento na mobilidade dos dados



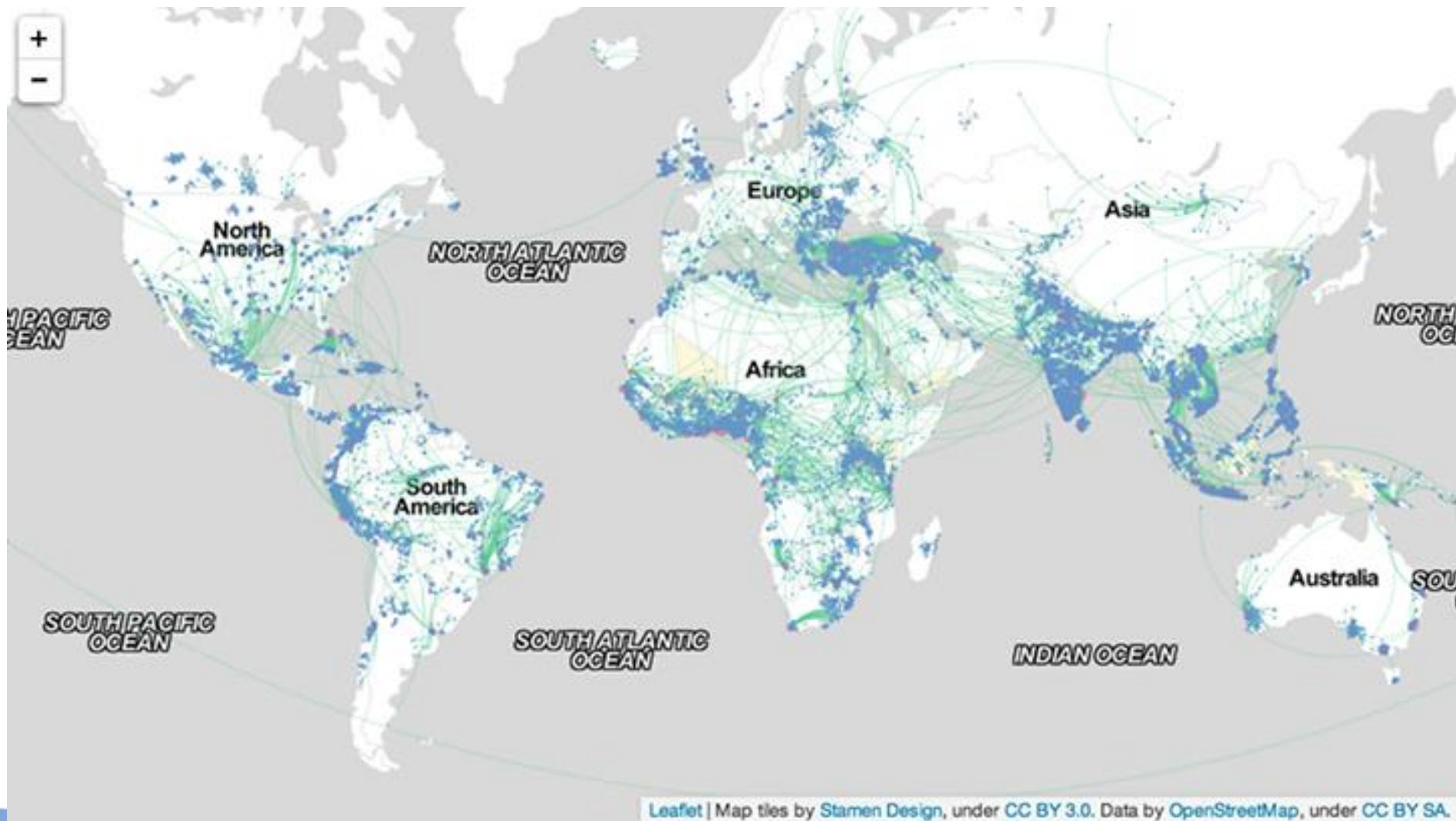
# A “felicidade média” no Twitter

Average Happiness for Twitter

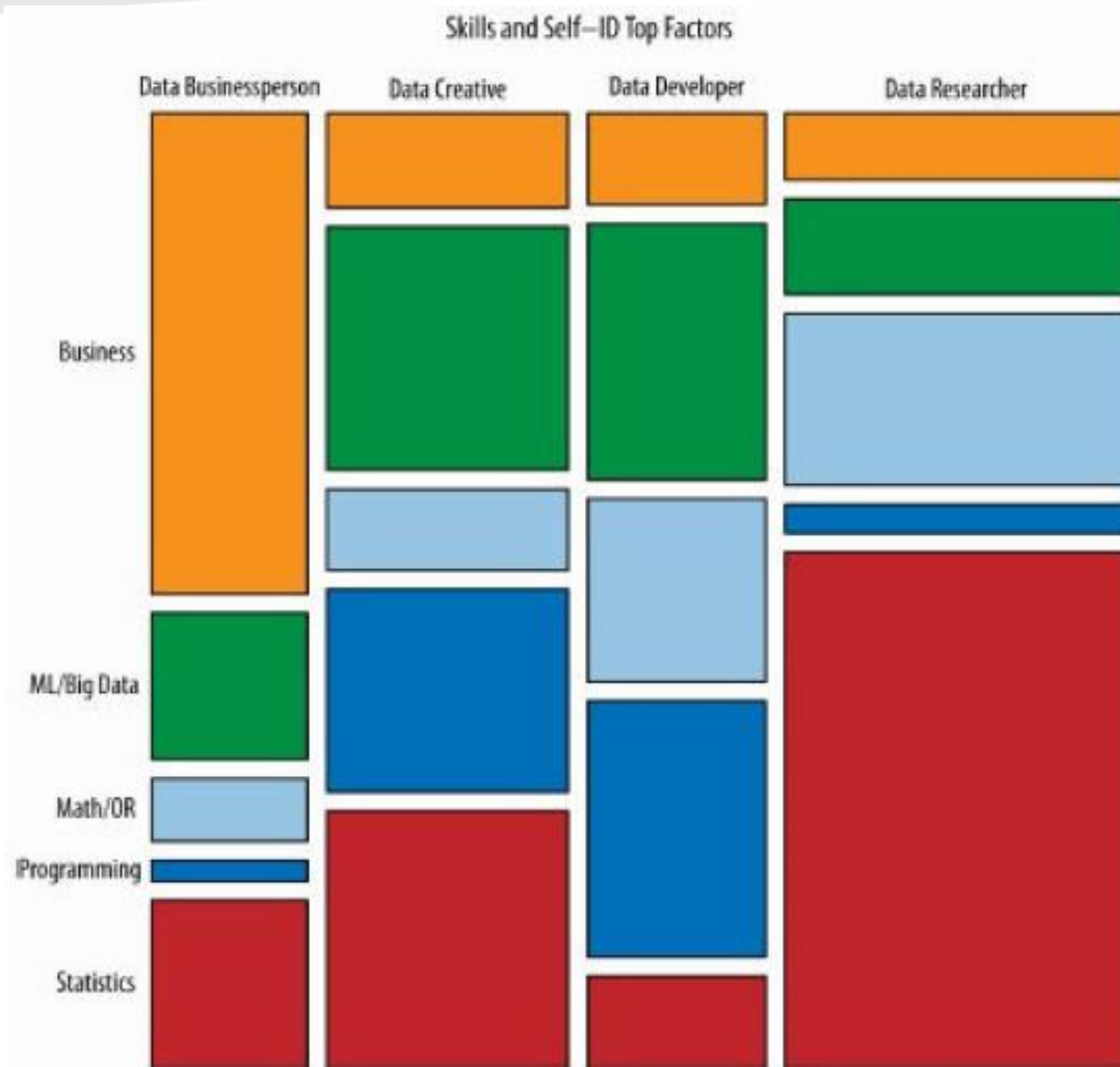




# Migração a partir de dados do Facebook (2014)



# Perfil dos analistas de dados



# Big Data Analytics



# Big Data Analytics

- **Big Data** é muito mais que um grande volume de dados;
- **Analytics** é muito mais que estatística aplicada.



# Big Data Analytics

- Tais conceitos, tanto **Big Data**, quanto **Analitics**, nem sempre são bem definidos;
- Desta forma, muitos de nós, erroneamente, nos apegamos a paradigmas tradicionais existentes e conceituamos as novas tecnologias com o “olhar” do passado.



# Big Data Analytics



# Big Data Analytics

- Big Data não é apenas uma nova forma de armazenar dados.
- Corporações ficam frustradas quando tentam simplificar uma ideia de Big Data por simplesmente comprar tecnologias do fornecedor X ou Y.



# Big Data Analytics

- Big Data Analytics traz consigo mudanças na maneira de **pensar os dados!**
- Por exemplo, pensando em relação ao volume de dados: ao sair do pensamento baseado na escassez para abundância de dados devemos pensar diferente.





# Big Data Analytics

- Por limitação e/ou dificuldade tecnológica podemos ser induzidos a construir um modelo mental de escassez de dados.
- Com isso podemos simplificar uma série de práticas.
- Vamos usar como exemplo as análises estatísticas por amostragem, ou seja, inferir comportamentos a partir de uma pequena amostra de dados.



# Big Data Analytics

- Modelos por amostragem são bastante confiáveis hoje em dia e são a base da Estatística moderna.
- Porém, com o advento da oferta de grandes massas de dados, podemos pensar em duas situações:



# Big Data Analytics

- Situação 1
  - A acurácia dos dados e, conseqüentemente, dos resultados está intimamente relacionada à amostra.
  - Pense em uma pesquisa de opinião baseada em uma amostragem aleatória de usuários de telefones fixos:
    - Se a coleta for feita no horário usual de trabalho, as respostas podem não representar a opinião das pessoas que trabalham em horários diferenciados.
    - As respostas podem nem representar a opinião dos trabalhadores pois como se trata de telefones fixos, podem ser respondidas pelas pessoas que estão em casa naquele momento.



# Big Data Analytics

- Situação 2
  - Um outro ponto é termos uma amostra de dados pequena, como realizado hoje em dia, e esta não ter representatividade estatística suficiente.
  - Pensemos agora em pesquisas de intenção de voto.
    - Geralmente estas pesquisas alcançam aproximadamente 2.000 pessoas, buscando uma visão geral da opinião do eleitorado.
    - Se quisermos mais detalhes sobre uma faixa etária em uma determinada região, a amostragem se mostra insuficiente.



# Big Data Analytics

- No entanto, quando falamos de Big Data, a ideia é outra.
- Quando todos os registros estão acessíveis, podemos ter amostras das mais diversas.
- Podemos então identificar tendências e descobrir correlações não pensadas antes.
- Podemos fazer novas perguntas e descer a novos níveis de segmentação.
- Desta forma, quebramos paradigmas, ou seja, aproveitamos oportunidades de fazer perguntas não pensadas antes de analisar os dados.



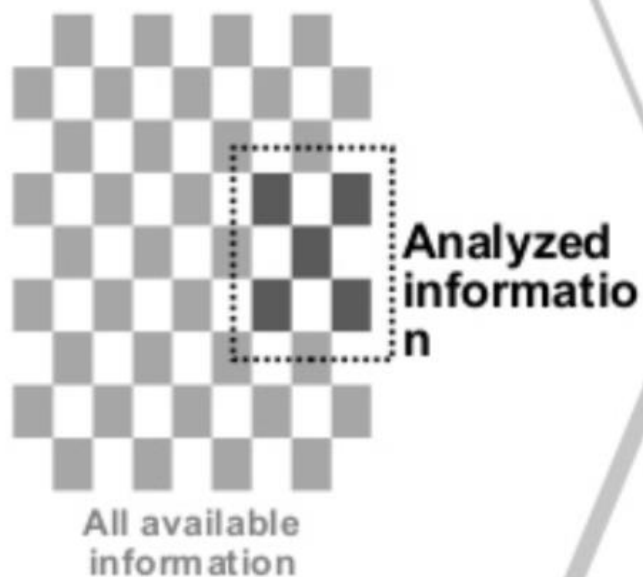
# Big Data Analytics

- As grandezas volume de dados e acurácia estão totalmente relacionadas, dada a flexibilidade de acesso aos dados e a certeza do melhor método estatístico a ser aplicado.
- A grande mudança de paradigma está em responder as perguntas específicas (que já foram preparadas anteriormente a análise) para entender os dados da maneira como eles estão correlacionados.
- Uma outra interpretação a esta nova abordagem é o desenvolvimento de algoritmos preditivos, que buscam prever comportamentos.



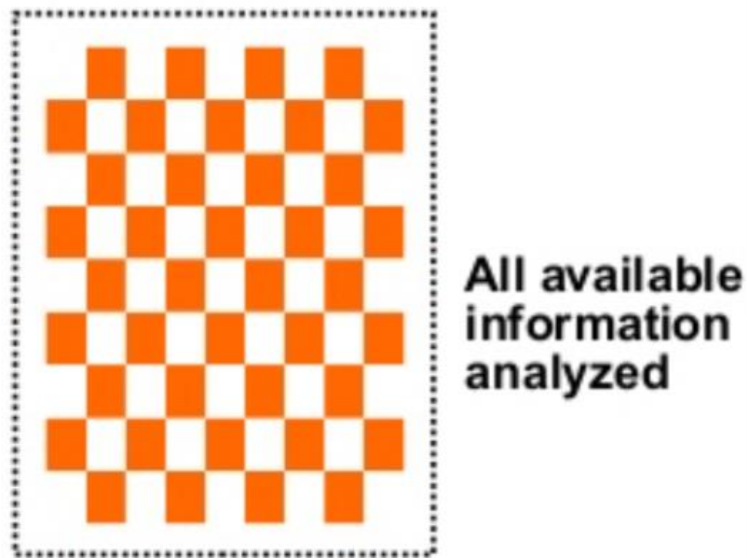
# Abordagem tradicional x Big data

## Traditional Approach



**Analyze small subsets of data**

## Big Data Approach



**Analyze *all* data**

# Exemplo

- Imaginemos a indústria de motores: com a atual capacidade de coleta de dados por meio de sensores, é grande o volume de dados gerados e as possibilidades sobre o que se pode analisar.
- A ideia básica é explicar a ocorrência de uma determinada falha/quebra de equipamento quando ela ocorre.



# Exemplo

- Dada esta nova abordagem (quebra de paradigma) pode-se fazer análises correlacionais para identificar determinados padrões que sinalizam futuros problemas.
- Quanto mais cedo uma provável anormalidade é detectada, mais eficiente é o processo de manutenção corretiva e/ou preditiva.



# Exemplo

- Na abordagem tradicional podemos reconhecer que um equipamento necessita de troca de óleo em intervalos periódicos.
- Assim, saberemos, em média, quando será a próxima troca de óleo.



# Exemplo

- Esta abordagem certamente é bastante conservadora pois ninguém, nem o fabricante, nem o consumidor, pretendem correr o risco de perder um motor por ter arriscado demais no prazo de troca.

# Exemplo

- Diferente, seria identificar uma próxima troca de óleo correlacionando dados como níveis de temperatura do motor, temperatura externa, corrente do motor, dentre outras.
- Esta nova abordagem pode levar à economia em relação ao custo da troca ou prevenir um problema iminente, quando as condições de uso foram mais severas do que o originalmente planejado para o equipamento.



# Mudança de paradigma

- Big Data Analytics é uma mudança de paradigma!
- Substituímos o modelo baseado em intuição (“eu acredito em uma coisa e, assim, explico/provo”) por direcionamento de descobertas a cada análise dos dados (“eu entendo o comportamento dos dados e suas correlações, assim, explico/provo”).



# Mudança de paradigma

- Entender a aplicabilidade de Big Data Analytics vai muito além da escolha por uma determinada tecnologia.
- A primeira abordagem é pensar e identificar as reais oportunidades que esta abordagem pode trazer, ou seja, seu valor.
- Podemos então esboçar sua aplicabilidade, o que pretende-se buscar, e, como executar esta busca de forma eficiente.



# Mudança de paradigma

- Lembre-se sempre que este é um processo evolutivo, de contínuo aprendizado.
- Novas descobertas surgirão a cada iteração, a cada exploração!



# Analytics e as empresas





# Panorama de mercado

- Após compreendemos os paradigmas existentes no contexto sobre análise de dados devemos tentar entender como isso está afetando as empresas!



# Panorama de mercado

- As empresas sempre usaram dados para fazer análises internas.
- O que mudou é que no contexto de Big Data Analytics, os gestores estão munidos com mais dados.



# Panorama de mercado

- Além disso, novas tecnologias permitem extrair valor mais facilmente a partir dessas análises.
- Para muitas pessoas, Big data é um termo que se refere apenas ao acúmulo de um grande volume de dados, porém é muito mais que isso, como já vimos antes.



# Panorama de mercado

- Porém, dados apenas não são suficientes para criar valor para seu cliente e muito menos para sustentar alguma vantagem competitiva para sua empresa.
- É preciso saber o que se busca responder (boas perguntas são o primeiro passo para uma estratégia bem-sucedida) e entender como interpretar e usar esses dados.
- Além de compreender os tipos de análises que podem ser usadas, é preciso entender no que investir.



# Oportunidades

- Identificar tendências e/ou padrões por meio da análise de dados em relação a comportamento do usuário ou consumidor.
- Desta forma podemos identificar tendências de consumo, detectar fraudes (ações que fujam muito do padrão de comportamento), etc.



# Oportunidades

- Identificar perfis de clientes.



# Oportunidades

- Desenvolver e/ou otimizar produtos e/ou serviços ou ainda processos a partir da identificação de tendências e peculiaridades do mercado e da organização.



# Oportunidades

- Otimizar estoques e promoções com análises preditivas mais precisas.





# Oportunidades

- Construir cenários e simulações mais próximos da realidade antes de agir.



# Exemplo

- Conteúdo extraído da Revista Exame, em 15/05/2015, sobre como o Big Data está mudando o mercado:
- “...A procura por gerentes de projeto com experiência em big data mais do que dobrou (123%) em 2014, segundo a Wanted Analytics, empresa que analisa sites de emprego no mundo todo. Um impulsionador desse mercado é o presidente americano, Barack Obama, que criou uma secretaria de serviços digitais com status de ministério para trabalhar com a quantidade massiva de dados que o governo produz.



# Exemplo

- Em um evento de computação, Obama fez uma reverência pessoal ao profissional de big data escolhido para chefiar a secretaria, DJ Patil, ex-LinkedIn. “Ajude-nos a construir serviços digitais melhores para o povo americano, ajude-nos a liberar inovações em áreas como saúde e mudança climática”, disse Obama em seu pedido a DJ.



# Exemplo

- O mais empolgante é que o Big Data tem efeito sobre profissionais de todos os departamentos. Analistas do banco de investimento UBS usaram vigilância por satélite de 100 estacionamentos do Walmart e recolheram dados sobre o número de carros estacionados em cada um, todos os meses...



# Técnicas e ferramentas

- Diversas empresas optam por técnicas e ferramentas que apoiam suas decisões, diminuindo as incertezas e, por consequência, prevenindo a tomada de decisões equivocadas.
- Dentre elas podemos destacar as ferramentas de BI, que facilitam a coleta, organização, análise, compartilhamento e o controle de dados internos e externos à organização.
- É através de ferramentas de BI que organizações de todo o mundo são capazes de criar relatórios e painéis analíticos que permitem a identificação e visualização de informações relevantes para o processo de tomada de decisão.



# Técnicas e ferramentas

- A primeira questão a ser considerada deve levar em conta se a tecnologia escolhida é capaz de responder às perguntas levantadas.
- Este questionamento pode ser considerado básico, mas no momento da escolha diversas características são observadas (onde a mais básica pode ser esquecida ou ter um peso inferior as demais) fugindo assim do real propósito para a escolha da solução.



# Outras características importantes

- Conexão à fonte de dados
  - Permitir que os usuários visualizem dados de maneira interativa, sem impor dimensões, padrões, estruturas ou condições predefinidas.
- Junção de dados
  - Combinar rapidamente fontes de dados diferentes, dispensando a elaboração de *scripts* específicos, e permitindo a integração dinâmica de tabelas, linhas e colunas em uma análise visual simples, sem assistência do serviço de TI.
- Modelos estatísticos
  - Integrar modelos preditivos e fluxos de eventos, superando as limitações da análise histórica de dados.



# Outras características importantes

- Colaboração Contextual
  - Capturar anotações e compartilhar com segurança essas informações com equipes internas e externas, a fim de embasar decisões colaborativas
  - Publicação de *dashboards* em portais corporativos e plataformas sociais, fornecendo um contexto visualmente incontestável para a aceleração do processo de tomada de decisão.
- Classe empresarial
  - Ter desempenho, performance e escalabilidade a grande volumes de dados.
  - Suportar arquitetura robusta que possa assegurar os quesitos de governança e segurança da informação.





# Qualificação de ferramentas

- Existem empresas que qualificam softwares analíticos baseado em contextos dos mais diversos.
- As duas principais empresas que atuam no mercado e são requisitadas por outras organizações para ajudá-las na escolha de uma solução que mais se adeque a suas necessidades, são:
  - Forrester Research;
  - Gartner Inc.



# IMPORTANTE!

- Ambas empresas, Forrester Research e Gartner Inc., divulgam seus próprios relatórios, muito respeitados no mercado.
- Em nenhuma hipótese devemos consumir as informações sobre ferramentas ali contidas sem a leitura prévia dos seus respectivos relatórios como um todo!



# The Forrester Wave™

- The Forrester Wave™: Multimodal Predictive Analytics and Machine Learning (PAML) Solutions, Q3 2018
- The Forrester Wave™: Enterprise Insight Platforms, Q1 2019



# Gardner Magic Quadrant

- Magic Quadrant for Business Intelligence and Analytics Platforms (Q1 2018).
- GARTNER MAGIC QUADRANT FOR ANALYTICS AND BUSINESS INTELLIGENCE PLATFORMS 2018: WHAT'S NEW?
- What's Changed: 2018 Gartner Magic Quadrant for Analytics and Business Intelligence Platforms
- Gainers and Losers in Gartner 2018 Magic Quadrant for Data Science and Machine Learning Platforms



# Gardner Magic Quadrant

- 2019 Gartner Magic Quadrant for Data Science and Machine Learning Platforms.
- 2019 Gartner Magic Quadrant for Analytics and Business Intelligence Platforms.

