

## Fundamentos da inteligência de negócios: gestão da informação e de banco de dados

1. Como um banco de dados relacional organiza os dados?
2. Quais são os princípios de um sistema de gestão de banco de dados?
3. Quais são as principais ferramentas e tecnologias para extrair informações de bancos de dados?
4. Qual o papel da política de informação e da gestão de dados na organização dos recursos de dados empresariais?
5. Por que a garantia da qualidade dos dados é tão importante para um negócio?

slide 2

© 2015 Pearson. Todos os direitos reservados.

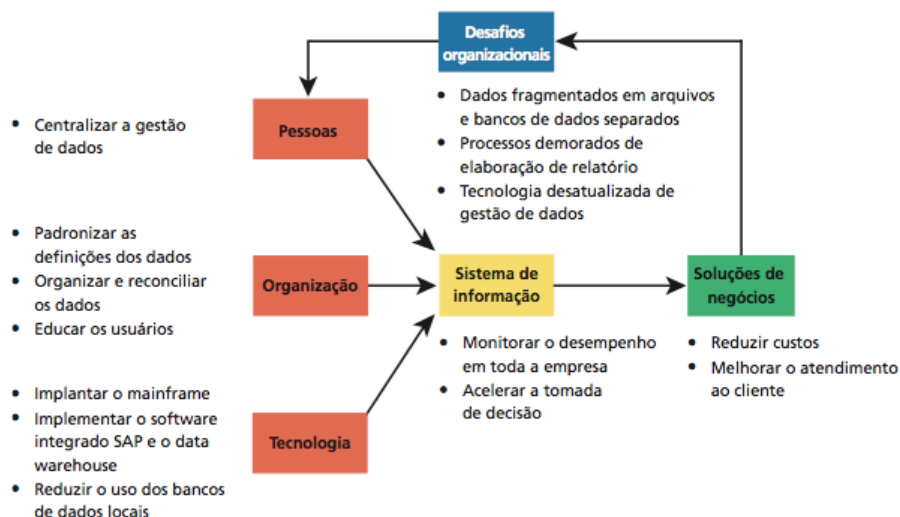
## Caso de abertura: A melhor gestão de dados ajuda o The Globe and Mail de Toronto a satisfazer seus clientes

- Contexto: The Globe and Mail, sediado em Toronto, no Canadá, impresso há 167 anos, é o maior jornal do país, com público de 3,3 milhões de leitores.
- Se usuários necessitassem de qualquer informação, eles tinham que extraí-la do mainframe e transferi-la a um dentre vários bancos de dados locais para análise, incluindo aqueles mantidos no Access, Foxbase Pro e Excel. Essa prática criou inúmeros repositórios mantidos em bancos de dados separados para fins específicos, mas nenhum repositório central onde os dados mais atualizados pudessem ser acessados em um único lugar.
- Com os dados espalhados em tantos sistemas diferentes por toda a empresa, era muito difícil efetuar uma referência cruzada entre os assinantes e os futuros clientes potenciais ao se desenvolver a mala direta.
- Problema: assinantes receberam nova oferta de assinatura de jornal !
- 2002: começou a resolver esses problemas com a implementação de um sistema integrado SAP com um data warehouse.
- Mas nem todos os usuários corporativos do migraram para o novo sistema.
- Empresa decidiu enfrentar esse novo problema educando seus usuários, especialmente seus profissionais de marketing, para que compreendessem o valor de ter todos os dados da organização em um data warehouse e as ferramentas disponíveis para acessá-los e analisá-los.

slide 3

© 2015 Pearson. Todos os direitos reservados.

## Caso de abertura: A melhor gestão de dados ajuda o The Globe and Mail de Toronto a satisfazer seus clientes



slide 4

© 2015 Pearson. Todos os direitos reservados.

## A abordagem de banco de dados para gestão de dados

- Um **banco de dados** é um conjunto de arquivos relacionados entre si com registros sobre pessoas, lugares ou coisas.
- Em um sistema de informação, o banco de dados organiza as informações de maneira muito parecida, agrupando dados relacionados entre si.
- O **banco de dados relacional** é o tipo de banco mais comum atualmente.
- Bancos de dados relacionais organizam os dados em tabelas bidimensionais (denominadas *relações*) com colunas e linhas.

slide 5

© 2015 Pearson. Todos os direitos reservados.

## A abordagem de banco de dados para gestão de dados

- Tabela de banco de dados relacional:

FORNECEDOR

Colunas (Atributos, Campos)

Número_Fornecedor	Nome_Fornecedor	Rua_Fornecedor	Cidade_Fornecedor	Estado_Fornecedor	CEP_Fornecedor
8259	CBM Inc.	74 5th Avenue	Dayton	OH	45220
8261	B. R. Molds	1277 Gandolly Street	Cleveland	OH	49345
8263	Jackson Components	8233 Micklin Street	Lexington	KY	56723
8444	Bryant Corporation	4315 Mill Drive	Rochester	NY	11344

Linhas (Registros, Tuplas)

Campo-chave (chave primária)

slide 6

© 2015 Pearson. Todos os direitos reservados.

## A abordagem de banco de dados para gestão de dados

- Tabela de banco de dados relacional:

PEÇA

Número_Peça	Descrição_Peça	Preço_Unitário	Número_Fornecedor
137	Trinco da porta	22,00	8259
145	Retrovisor externo	12,00	8444
150	Vedação da porta	6,00	8263
152	Trava da porta	31,00	8259
155	Compressor	54,00	8261
178	Maçaneta da porta	10,00	8259

Chave primária

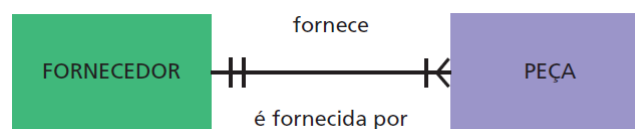
Chave estrangeira

slide 7

© 2015 Pearson. Todos os direitos reservados.

## A abordagem de banco de dados para gestão de dados

- Uma representação esquemática denominada **diagrama entidade/relacionamento** (DER) é usada para esclarecer relações entre as tabelas de um banco de dados relacional.
- Um diagrama entidade/relacionamento básico:



slide 8

© 2015 Pearson. Todos os direitos reservados.

## A abordagem de banco de dados para gestão de dados

**Figura 6.4** Amostra de relatório de pedido

As áreas sombreadas mostram quais dados vêm das tabelas PEDIDO, FORNECEDOR e ITEM. O banco de dados não mantém dados sobre Preço Total nem Pedido Total porque eles podem ser derivados a partir de outros dados das tabelas.

Número do pedido: 3502  
Data do pedido: 15/01/2014

Número do fornecedor: 8259  
Nome do fornecedor: CBM Inc.  
Endereço do fornecedor: 74 5th Avenue, Dayton, OH 45220

Número_Pedido	Número_Peça	Quantidade_Peça	Nome_Peça	Preço_Unitário	Preço Total
3502	137	10	Trinco da porta	22,00	220,00
3502	152	20	Trava da porta	31,00	620,00
3502	178	5	Maçaneta da porta	10,00	50,00
Pedido Total:					890,00

slide 9

© 2015 Pearson. Todos os direitos reservados.

## A abordagem de banco de dados para gestão de dados

- Como um pedido pode englobar muitas peças de um mesmo fornecedor e uma única peça pode ser pedida muitas vezes em diferentes pedidos, isso cria uma relação muitos para muitos entre as tabelas PEÇA e PEDIDO.

PEÇA				ITEM		
Número_Peça	Descrição_Peça	Preço_Unitário	Número_Fornecedor	Número_Pedido	Número_Peça	Quantidade_Peça
137	Trinco da porta	22,00	8259	3502	137	10
145	Retrovisor externo	12,00	8444	3502	152	20
150	Vedação da porta	6,00	8263	3502	178	5
152	Trava da porta	31,00	8259			
155	Compressor	54,00	8261			
178	Maçaneta da porta	10,00	8259			

PEDIDO	
Número_Pedido	Data_Pedido
3502	15/01/2014
3503	16/01/2014
3504	17/01/2014

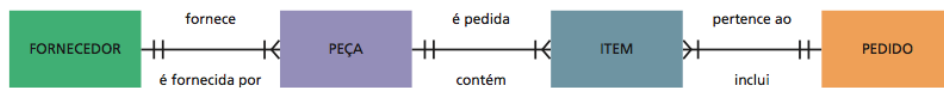
  

FORNECEDOR					
Número_Fornecedor	Nome_Fornecedor	Rua_Fornecedor	Cidade_Fornecedor	Estado_Fornecedor	CEP_Fornecedor
8259	CBM Inc.	74 5th Avenue	Dayton	OH	45220
8261	B. R. Molds	1277 Gandolly Street	Cleveland	OH	49345
8263	Jackson Components	8233 Micklin Street	Lexington	KY	56723
8444	Bryant Corporation	4315 Mill Drive	Rochester	NY	11344

slide 10

Todos os direitos reservados.

## A abordagem de banco de dados para gestão de dados



**Figura 6.6** Diagrama Entidade/Relacionamento para o banco de dados de quatro tabelas

Esse diagrama mostra o relacionamento entre as entidades FORNECEDOR, PEÇA, ITEM e PEDIDO.

slide 11

© 2015 Pearson. Todos os direitos reservados.

## Sistemas de gestão de banco de dados

- Um **sistema de gestão de banco de dados (DBMS)** é um software específico usado para criar, armazenar, organizar e acessar dados a partir de um banco de dados.
- No banco de dados relacional, três operações básicas, conforme mostrado na figura a seguir, são utilizadas para desenvolver conjuntos de dados úteis:
- select,
- project e
- join.

slide 12

© 2015 Pearson. Todos os direitos reservados.

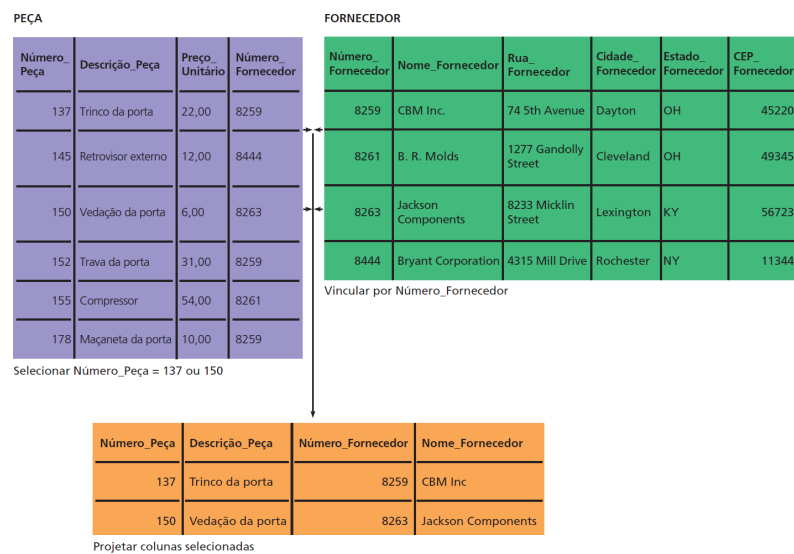
## Sistemas de gestão de banco de dados

- Conceitos:
  - Colunas: atributos
  - Linhas: tuplas ou registros
  - Chave primária
  - Chave estrangeira
  - Diagrama Entidade Relacionamento (E/R)
  - Normalização

slide 13

© 2015 Pearson. Todos os direitos reservados.

## Sistemas de gestão de banco de dados



slide 14

© 2015 Pearson. Todos os direitos reservados.

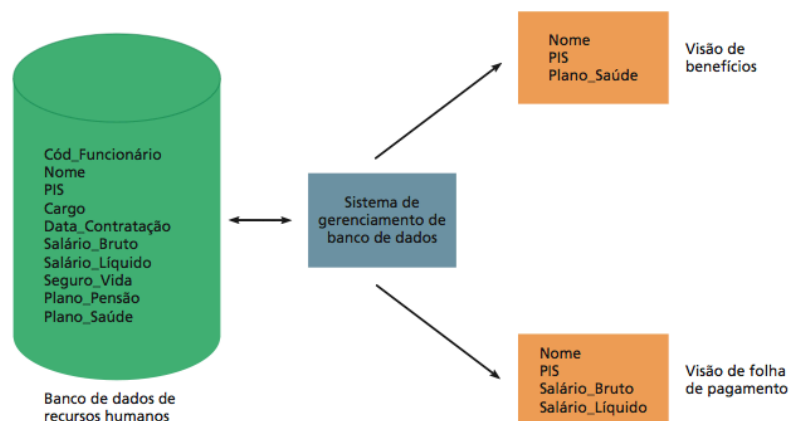
## Sistemas de gestão de banco de dados

- Os DBMS têm um recurso de **definição de dados** para especificar a estrutura do conteúdo do banco de dados.
- A maioria dos DBMS possui uma linguagem especializada, a **linguagem de manipulação de dados**, usada para acrescentar, alterar, apagar e recuperar os dados do banco.
- A linguagem de manipulação de dados mais usada atualmente é a **linguagem estruturada de consulta** ou **SQL**.
- Normalmente, os DBMS incluem recursos para geração de relatórios.

slide 15

© 2015 Pearson. Todos os direitos reservados.

## Sistemas de gestão de banco de dados



**Figura 6.7** Banco de dados de recursos humanos com múltiplas visões

Um único banco de dados de recursos humanos oferece muitas visões dos dados, conforme as necessidades de informação do usuário. Exemplificamos aqui duas visões possíveis, uma que interessa a um especialista em benefícios e outra que interessa a um funcionário do departamento responsável pela folha de pagamentos da empresa.

slide 16

© 2015 Pearson. Todos os direitos reservados.



## Sistemas de gestão de banco de dados

```
SELECT PEÇA.Número_Peça, PEÇA.Descrição_Peça,
FORNECEDOR.Número_Fornecedor,
FORNECEDOR.Nome_Fornecedor
FROM PEÇA, FORNECEDOR
WHERE PEÇA.Número_Fornecedor = FORNECEDOR.Número_
Fornecedor AND
Número_Peça = 137 OR Número_Peça = 150;
```

**Figura 6.10** Exemplo de uma consulta SQL

Vemos aqui o comando SQL para uma consulta que selecionará fornecedores das peças 137 ou 150. Ele produz uma lista com os mesmos resultados da Figura 6.8.

slide 17

© 2015 Pearson. Todos os direitos reservados.

## Bancos de dados não relacionais e bancos de dados em nuvem

- Os **sistemas de gestão de banco de dados não relacional** são projetados para gerenciar grandes conjuntos de dados em vários computadores distribuídos e para permitir que se aumente ou reduza facilmente o número de máquinas e o volume de dados.
- Eles são úteis para acelerar consultas simples a grandes volumes de dados estruturados e não estruturados, incluindo dados da Web, de mídias sociais, gráficos e outras formas difíceis de analisar com ferramentas tradicionais baseadas em SQL.
- O banco de dados NoSQL é capaz de receber informações estruturadas, semiestruturadas e não estruturadas.

slide 18

© 2015 Pearson. Todos os direitos reservados.

## Bancos de dados não relacionais e bancos de dados em nuvem

- Além dos serviços de gestão de dados baseados em nuvem pública, as empresas têm agora a opção de usar bancos de dados em nuvens privadas.
- Ferramentas de gestão de carga de trabalho asseguram a disponibilidade dos recursos suficientes para atender as necessidades da aplicação, mesmo quando há alterações na carga de trabalho.
- A plataforma de hardware e software compartilhada reduz o número de servidores, DBMS e dispositivos de armazenamento necessários para esses projetos.

slide 19

© 2015 Pearson. Todos os direitos reservados.

## Como usar bancos de dados para melhorar o desempenho e a tomada de decisão na empresa

- Utilizamos, atualmente, o termo **big data** para descrever esses conjuntos de dados com volumes tão imensos a ponto de estarem além da capacidade que um DBMS típico tem para capturar, armazenar e analisar.
- Uma infraestrutura contemporânea para inteligência empresarial consiste em um conjunto de ferramentas para a obtenção de informações úteis.
- Esses recursos incluem *data warehouses* (armazéns de dados) e *data marts*, *Hadoop*, computação em memória e plataformas analíticas.

slide 20

© 2015 Pearson. Todos os direitos reservados.

## Como usar bancos de dados para melhorar o desempenho e a tomada de decisão na empresa

### ➤ Big data:

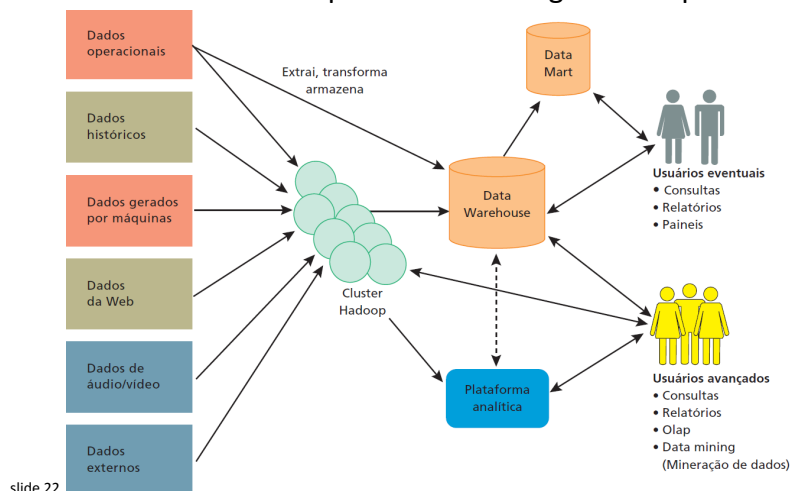
- normalmente a dados na faixa dos petabytes e exabytes — em outras palavras, bilhões a trilhões de registros, todos oriundos de fontes diversas.
- Por exemplo, um único motor a jato é capaz de gerar 10 terabytes de dados em apenas 30 minutos, sendo que existem mais de 25 mil voos por dia.
- Facebook tem mais de 250 bilhões de fotos no site, e a cada dia adiciona 350 milhões de novas fotos.
- Mesmo que os tweets sejam limitados a 140 caracteres cada, o Twitter gera mais de 8 terabytes de dados por dia.
- A informação digital está crescendo exponencialmente, de 1,8 zetabytes em 2011 para um montante esperado de 35 zetabytes em 2020.

slide 21

© 2015 Pearson. Todos os direitos reservados.

## Como usar bancos de dados para melhorar o desempenho e a tomada de decisão na empresa

### ➤ Infraestrutura contemporânea de inteligência empresarial:



slide 22

Todos os direitos reservados.

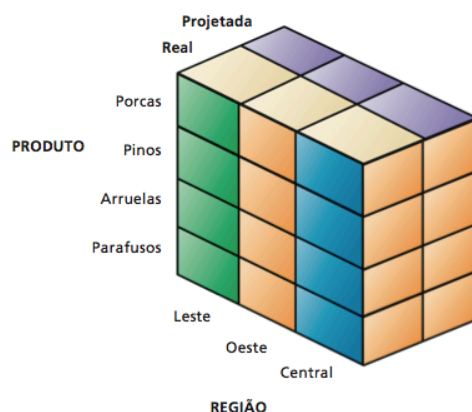
## Como usar bancos de dados para melhorar o desempenho e a tomada de decisão na empresa

- Ferramentas Analíticas:
  - OLAP (Online Analytical Processing — Olap)
    - Permite análise multidimensional de dados, de forma que os usuários vejam os mesmos dados de diferentes maneiras, pois usa múltiplas dimensões.
    - Ex: Compare as vendas do produto 403 com o planejado, por trimestre e por região de vendas, nos dois últimos anos. (Diferente da consulta tradicional: quantas unidades do produto número 403 foram expedidas em fev 2014?)
  - Data mining
    - Mais orientado por descoberta: associações, sequências, classificações, aglomerações e prognósticos.
    - Fornece percepções dos dados corporativos que não podem ser obtidas com o Olap, descobrindo padrões e relacionamentos ocultos em grandes bancos de dados e inferindo regras a partir deles para prever comportamentos futuros.
  - Text mining e Web mining
    - Dados não estruturados: e-mails, memorandos, transcrições de chamadas à central de atendimento ao cliente, respostas para pesquisas, casos judiciais, descrições de patentes e relatórios de serviços.
    - Google Trends e Insight for Search, rastreiam a popularidade de diversas palavras e frases usadas nas consultas feitas pelo Google para descobrir no que as pessoas estão interessadas e o que elas desejam comprar.

slide 23

© 2015 Pearson. Todos os direitos reservados.

## Como usar bancos de dados para melhorar o desempenho e a tomada de decisão na empresa



**Figura 6.13** Modelo multidimensional de dados

Essa visão mostra produto versus região. Girando o cubo 90 graus, a face que aparecer mostrará produto versus vendas projetadas e reais. Girando novamente o cubo 90 graus, pode-se ver região versus vendas reais e projetadas. Outras visões também são possíveis.

slide 24

© 2015 Pearson. Todos os direitos reservados.

## Seção Interativa: Organizações

### Inteligência Empresarial ajuda Jardim Zoológico de Cincinnati a conhecer seus clientes

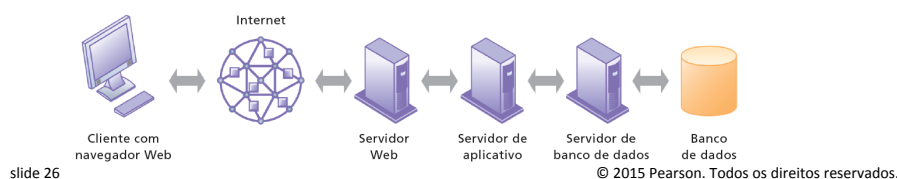
- Contexto: É uma das instituições zoológicas mais bem classificadas do mundo; está entre os 10 melhores zoológicos norte-americanos pela Zagat, e como o melhor zoológico para crianças pela revista Parents Magazine.
- Para aumentar a receita e melhorar o desempenho, a equipe de gerência sênior do Jardim Zoológico realizou uma avaliação que concluiu que a gestão tinha limitados conhecimento e compreensão do que estava realmente acontecendo no Zoo. Quem vem ao Zoo? Com qual frequência costuma vir? O que faz e o que compra? A administração não tinha ideia.
- Cada um dos quatro fluxos de receita — admissões, filiação, varejo e serviços de alimentação — possuía plataformas diferentes de sistema.
- O Zoo substituiu seus quatro sistemas legados de ponto de venda por uma única plataforma.
- Zoo agora atualiza seu sistema de inteligência empresarial com dados de previsão do tempo, tomar decisões mais acertadas sobre o planejamento de horários de trabalho dos funcionários e sobre o planejamento de estoque.
- Segmento de clientes que só gasta com ingresso foi identificado. Marketing direcionado ajudou a economizar 40m\$ do orçamento anual.
- Análise de dados mostrou que vendas de alimentos caem após 15hs e de sorvete sobem 1 hr antes do parque fechar. Adaptação: lojas de alimentos fecham, sorveterias ficam abertas o dia inteiro.
- 6 meses após a implantação da solução de inteligência empresarial, o Zoo conseguiu aumento de 30,7% nas vendas de alimentos, e aumento de 5,9% nas vendas no varejo comparado ao ano anterior.

slide 25

© 2015 Pearson. Todos os direitos reservados.

## Bancos de dados e a Web

- Muitas empresas estão usando a Web para disponibilizar algumas informações de seus bancos de dados internos a clientes e parceiros de negócios.
- Como muitos bancos de dados *back-end* não conseguem interpretar comandos escritos em HTML, o servidor Web repassaria essas solicitações para o software que traduz comandos HTML em SQL, de maneira que elas pudessem ser processadas pelo DBMS que trabalha com o banco de dados.



## Gestão dos recursos de dados

- Uma **política de informação** especifica as regras para compartilhar, disseminar, adquirir, padronizar, classificar e inventariar a informação.
- A **administração de dados** é responsável pelas políticas e procedimentos específicos pelos quais as informações podem ser gerenciadas como recurso organizacional.
- As funções que o grupo desempenha são chamadas de **gestão de banco de dados**.
- O **data cleansing** consiste em atividades para detectar e corrigir, dentro do banco de dados, informações incorretas, incompletas, formatadas inadequadamente ou redundantes.

slide 27

© 2015 Pearson. Todos os direitos reservados.

## Resolvendo problemas organizacionais:

### Será que o Big Data traz grandes recompensas?

- Em 2012 foi gerado um número estimado de 988 exabytes de informações digitais, equivalente a duas pilhas de livros colocadas a partir do sol até Plutão.
- Como as empresas atuais estão tirando proveito do “big data”?
- Agências policiais estão usando big data para descobrir padrões ocultos na atividade criminal, como as correlações entre tempo, oportunidade e organizações.
- As informações sobre os criminosos, como a foto de um suspeito com detalhes de delitos cometidos no passado ou endereços com mapas, podem ser visualizadas em segundos em uma parede de vídeo ou retransmitidas imediatamente para os oficiais em uma cena de crime.
- Vestas: maior empresa de energia eólica do mundo, com mais de 43 mil turbinas eólicas em 66 países. Sua biblioteca eólica armazena atualmente 2,8 petabytes de dados e inclui aproximadamente 178 parâmetros, como pressão barométrica, umidade, direção do vento, temperatura, velocidade do vento e outros dados históricos.
- É possível prever em 15 mins, e não em três semanas, a colocação ideal de turbinas, economizando um mês do tempo de determinação do local de instalação de uma turbina e possibilitando que seus clientes alcancem seu ROI muito mais rápido.
- Autozone e Hertz
- Limitações: Google x gripe = erro!
- Sears e Kmart: perdeu terreno devido a preços altos. Atualmente programa de fidelidade e Hadoop. Mas não há certeza se ela conseguirá se reenguer.

slide 28

© 2015 Pearson. Todos os direitos reservados.

## Resumo

1. Como um banco de dados relacional organiza os dados?
2. Quais os princípios de um sistema de gestão de banco de dados?
3. Quais as principais ferramentas e tecnologias para extrair informações de bancos de dados e, assim, melhorar o desempenho da empresa e a tomada de decisão?
4. Qual o papel da política de informação e da gestão de dados na gestão dos recursos de dados organizacionais?
5. Por que é importante para a empresa garantir a qualidade dos dados?

slide 29

© 2015 Pearson. Todos os direitos reservados.