



WBA0750_v1.0

Data Discovery, Olap e Visualização de Dados



Data Discovery

Bloco 1

Washington H. C. Almeida



➤ Processo de descoberta de dados

- A descoberta de informações importantes, padrões e tendências que auxiliem na tomada de decisões por meio de processos de visualização de dados é um dos grandes desafios das empresas e dos negócios, em geral.
- O processo de descoberta de dados, conhecido como *data discovery*, tem como principal objetivo encontrar padrões e anomalias para dar suporte às decisões das empresas e interessados, com intuito de sempre estarem à frente em termos de competitividade de mercado.





Conceitos básicos de *data discovery*

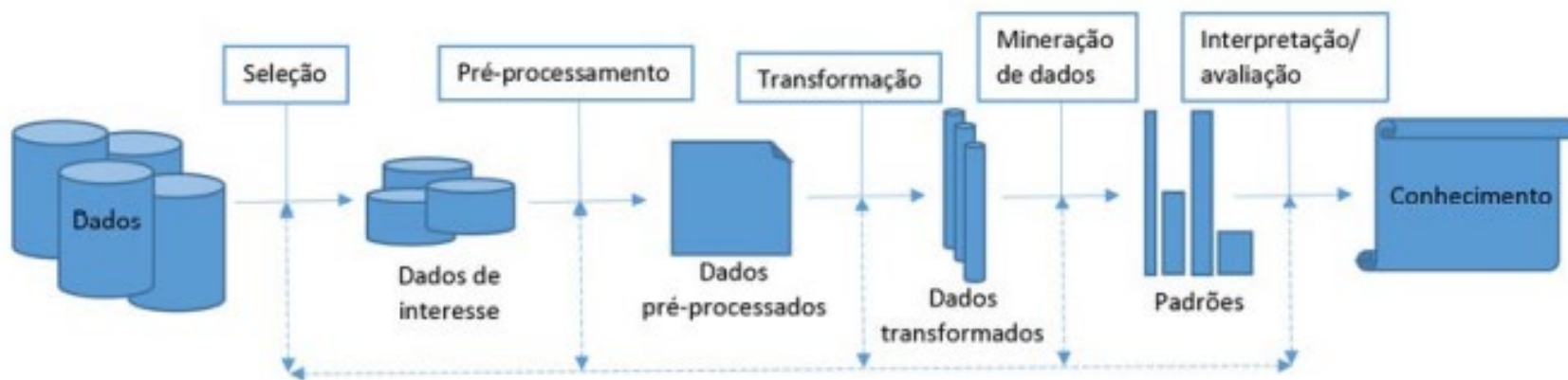
- A descoberta de conhecimento a partir de bases de dados (KDD, do inglês *knowledge Discovery in Databases*). De forma resumida, pode ser considerada como o processo de identificação de padrões novos, padrões válidos ou que façam algum sentido, que possam ser úteis e, principalmente, compreensíveis e interpretáveis.



Processo de descoberta de conhecimento

- A descoberta do conhecimento em bases de dados tem como principal objetivo a extração de conhecimento inteligível e utilizável para o apoio à tomada de decisões, e encontrar padrões e tendências. O processo de descoberta do conhecimento a partir dos dados é composto por algumas etapas.

Figura 1 – Processo de descoberta de conhecimento em bancos de dados



Fonte: Freitas, Moura e Silva (2015).



Formas tradicionais de *data discovery*

- De início, tanto as planilhas de dados, quanto de MS Excel, quanto as visualizações básicas, como gráficos básicos (barras, linhas e setores), são consideradas formas tradicionais de *data discovery*.
- As planilhas, como o MS Excel, continuam sendo a ferramenta mais popular de *business analytics* para trabalhar com dados, em parte devido à sua ampla existência e disponibilidade e, também, à familiaridade do usuário, dada a facilidade de sua utilização.





Visualizações básicas

- Visualizações básicas, como gráficos de barras, linhas ou de setores (incluindo aqueles incorporados em *dashboards*) gerados ou não no MS Excel, fornecem informações simples e diretas por representações visuais de dados que permitem aos analistas encontrar insights que podem não ser tão facilmente percebidos em um formato de texto simples.





Conclusão

- A descoberta do conhecimento em bases de dados tem como principal objetivo a extração de conhecimento inteligível e utilizável para o apoio à tomada de decisões, e encontrar padrões e tendências.
- Uma planilha de dados pode ser uma poderosa ferramenta analítica nas mãos de um usuário experiente.



Data Discovery


Bloco 2

Washington H. C. Almeida





Formas avançadas de *data discovery*

- Um *data discovery* multifacetado (ou “modo de busca”) permite aos analistas minerar dados para obtenção de insights sem discriminar entre dados estruturados e não estruturados.
 - Podem acessar dados em documentos, e-mails, imagens, redes sociais etc., em um mecanismo de pesquisa como o Google, Yahoo! ou Bing, com a capacidade de interagir conforme suas necessidades, além de poder detalhar a fundo os dados disponíveis para descobrir novos insights.
- 

➤ Visualizações avançadas

- As visualizações avançadas são tudo o que as visualizações básicas de dados não são, ou seja, são aquelas obtidas com ferramentas mais especializadas, com arquiteturas de dados mais estruturadas.
- Elas são ferramentas para *visual discovery*, que permitem aos analistas utilizar *Big Data* para descobrir insights de maneira totalmente nova.




➤ Ferramentas de *data discovery*

- A maioria das pessoas de negócios não são programadores ETL (do inglês, *extract, transform, load*), ou seja, não detêm o conhecimento necessário para realizar integração de dados. Não conhecem a linguagem SQL ou qualquer outra linguagem, e muitos sequer têm alguma formação para realizar análise de dados.
- Os profissionais dos negócios precisam de ferramentas que trabalhem de acordo com as suas necessidades sem precisar ter que escrever códigos de programação.






Ambientes de *data discovery*

- O processo de descoberta de dados em negócios precisa ser comandado por seus usuários, ou seja, as pessoas envolvidas nos processos de negócios.
 - Tais usuários são os guias com o conhecimento necessário para identificar oportunidades e saber o que estão procurando.
 - Podem até não saber o que procuram, no entanto, quando surgir alguma pista, poderão fazer essa identificação.
- 



O cientista de dados

- O cientista de dados é alguém com conhecimento técnico vertical em estatística, NoSQL, *cloud computing*, mineração de dados (*data mining*), álgebra relacional, modelagem multidimensional, MapReduce, virtualização, entre outros.
 - O cientista de dados pode estar habilitado para atuar na instalação dos sistemas e ambientes de *data warehouse*, no gerenciamento de bancos de dados e produção de informações e atuação em *data discovery*.
- 



O analista de visualização de dados

- É um profissional com papel mais técnico em habilidades e com a necessidade de ter uma profunda compreensão de *analytics* e das ferramentas analíticas para processar dados de várias fontes. Além de extrair, transformar e carregar os dados nos diversos relatórios gerenciais e torná-los significativos dentro do contexto dos objetivos.





Conclusão

- Os profissionais de negócios buscam ferramentas de *data discovery*, que sejam autossuficientes para algumas aspectos, como agilidade, iterativas, permitam interação entre elas e os dados.
- A descoberta de dados requer um salto de fé. Essa mentalidade otimista e voltada para o futuro é uma das principais distinções entre BI e *data discovery*.



Teoria em Prática

Bloco 3

Washington H. C. Almeida



Refleta sobre a seguinte situação

- Em uma organização, não existe um ambiente clássico de DW, com processos de ETL definidos e ferramentas OLAP. Sendo assim, essa instituição teria que fazer o que para começar a fazer análise de dados para descobrir conhecimento?
- Qual seria sua proposta nesse cenário?

Norte para a resolução...

- Nesses cenários, poderiam ser utilizadas inúmeras técnicas e processos voltados a descoberta de conhecimento. Ter um DW corporativo não é a única forma de obter insights a partir dos dados.
- Muitas instituições nem procuram mais construir DW, mas sim a utilização de ferramentas de *Data Discovery* para análise de informações mais pontuais. Além disso, muitas soluções tem caminhado para o uso de *Big Data* que são soluções baseadas em uma gama nova de ferramentas.

Dica do(a) Professor(a)

Bloco 4

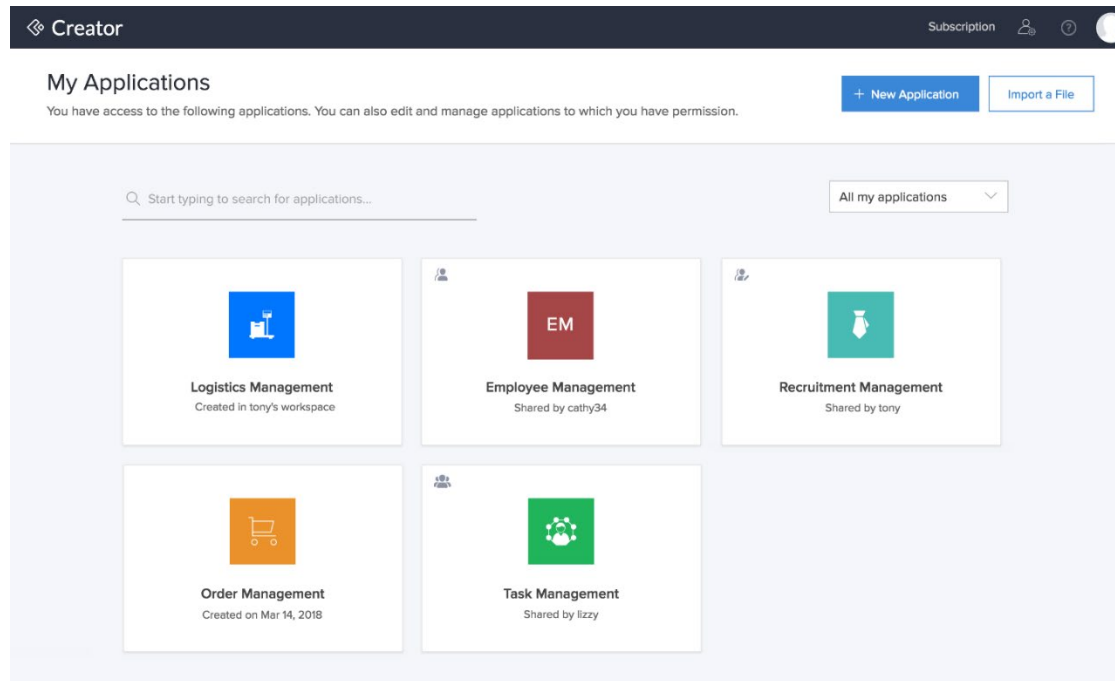
Washington H. C. Almeida



Ferramenta – Zoho Creator

- Zoho Creator é uma ferramenta que permite representar visualmente seus dados para que você possa analisá-los facilmente.

Figura 2 – Tela inicial do Zoho Creator



Fonte: captura de tela de Zoho Creator.



Referências

AMARAL, Fernando. **Introdução a ciência de dados**: mineração de dados e Big Data. (versão Kindle). Rio de Janeiro: Alta Books, 2016.

DAVENPORT, T. H., PATIL, D. J. Data scientist: the sexiest job of the 21st century. **Harvard Bus. Ver.**, Brighton, 90 (10), p. 70–76, 2012. Disponível em: <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century>. Acesso em: 25 jul. 2019.

FREITAS, N.; MOURA, C.; SILVA, M. Sistema multiagente para mineração de imagens de satélite. **Anais [...]**, João Pessoa-PB, 25 a 29 de abril de 2015. Disponível em: https://www.researchgate.net/publication/283716349_Sistema_multiagente_para_mineracao_de_imagens_de_satelite. Acesso em: 25 jul. 2019.

MATTHEW, N. Why visual analytics? **Tableau**, Seattle, 2019. Disponível em: https://www.tableau.com/sites/default/files/whitepapers/752750_core_why_visual_analytics_whitepaper_0.pdf. Acesso em: 24 jul. 2019.

RYAN, L. **The visual imperative**: creating a visual culture of data discovery. Cambridge: Elsevier, 2016.

TRINDADE, R. A máquina no lugar do homem: A inteligência artificial eliminará empregos, mas novas profissões surgirão; qualificação será fundamental. **UOL**, São Paulo, 17 de agosto de 2018. Disponível em: <https://www.uol/tecnologia/especiais/inteligencia-artificial-vai-acabar-com-empregos-.htm>. Acesso em: 19 ago. 2019.

Bons estudos!

