

Podcast

Disciplina: Data Discovery, Olap e Visualização de Dados

Título do tema: Data Discovery

Autoria: Washington Henrique Carvalho Almeida

Leitura crítica: Henrique Salustiano Silva

Abertura:

Olá, ouvinte! No podcast de hoje vamos falar sobre o Data Discovery.

A descoberta de informações importantes, padrões e tendência que auxiliem na tomada de decisões através 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.

A descoberta de conhecimento através de bases de dados 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.

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.

Descoberta de dados, ou mais comumente conhecida como data discovery, é uma maneira de permitir que os envolvidos em algum processo de trabalho e/ou negócios possam obter as informações necessárias para realizar seus trabalhos de forma confiável, intuitiva e de alta disponibilidade.

Os dados trabalhados no contexto do data discovery têm origem diversa, diferentes instantes de coleta e em lugares distintos. Essa diversidade gera a necessidade de um esforço para consolidação e agrupamento para a geração de informação confiável e interpretável para o processo, isso até faz lembrar o termo de armazenamento de dados ou data warehouse.

Dentro do data Discovery, algumas ferramentas e formas são bem utilizadas, como as planilhas, que continuam sendo a ferramenta mais popular de business analytics para trabalhar com dados, em parte devido à sua ampla existência e disponibilidade e devido à familiaridade do usuário, dada facilidade de sua utilização.

As 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.

Dentro desse conceito de data Discovery, existe o modo de pesquisa multifacetado, na qual permite aos analistas minerar dados para obtenção de insights sem discriminar entre dados estruturados e não estruturados.

As visualizações avançadas são tudo o que as visualizações básicas de dados não são, são aquelas obtidas com ferramentas mais especializadas, com arquiteturas de dados mais estruturadas. Com visualizações avançadas, os analistas podem identificar clusters ou agregados diversos; também podem analisar dados para procurar correlações ou preditores para descobrir novos modelos analíticos.

Fechamento:

Este foi nosso podcast de hoje! Até a próxima!