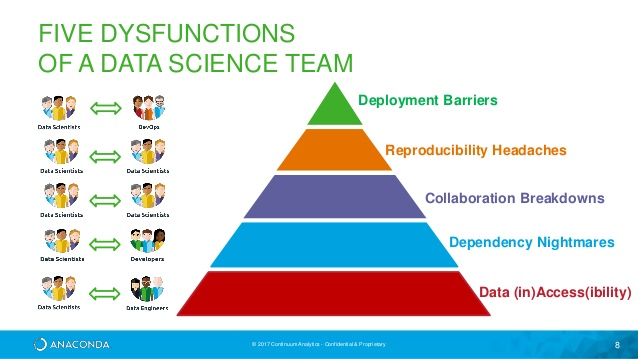
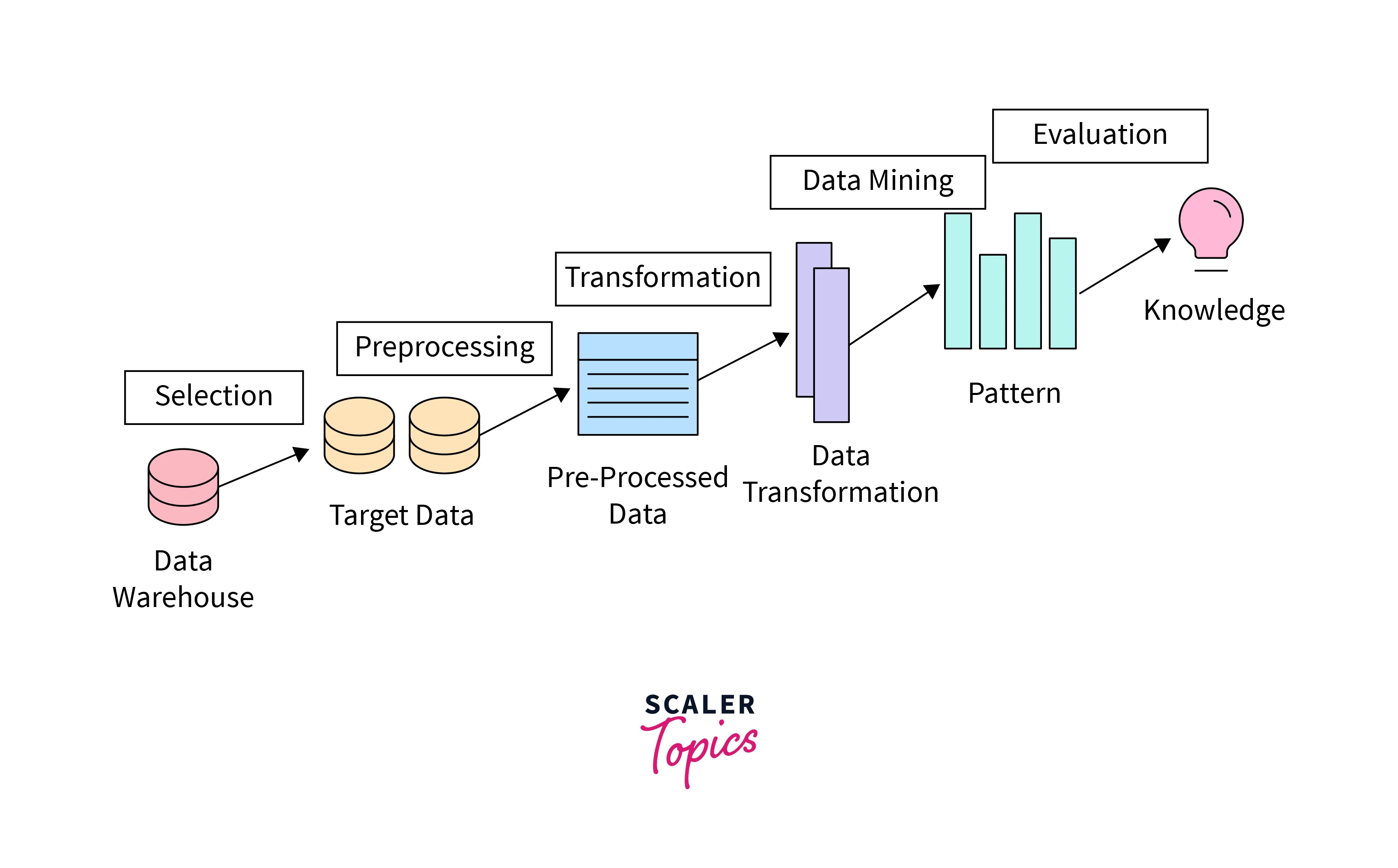
**INTRODUÇÃO À CIÊNCIA DE DADOS**



- MODELO DE CARREIRA EM PARALELO: Carreiras técnicas tinham um limite de crescimento, enquanto carreiras gerenciais não tinham uma barreira. Por isso, ótimos técnicos migravam para áreas gerenciais. Porém, a empresa perdia um técnico de elite e ganhava um gerente medíocre (não necessariamente um bom técnico será um bom gerente).

- MODELO DE CARREIRA Y: Abre dois tipos de carreiras de acordo com o perfil do profissional. Esse modelo prioriza o desempenho do profissional.

**DESCOBERTA DE CONHECIMENTO EM UMA BASE DE DADOS (KDD)**



**PROBLEMAS:** Os analistas selecionam os dados a serem minerados (podem ser enviesados). Natureza do sequenciamento e a forma do conhecimento (ao longo do trabalho, decisões diferentes podem ser tomadas).

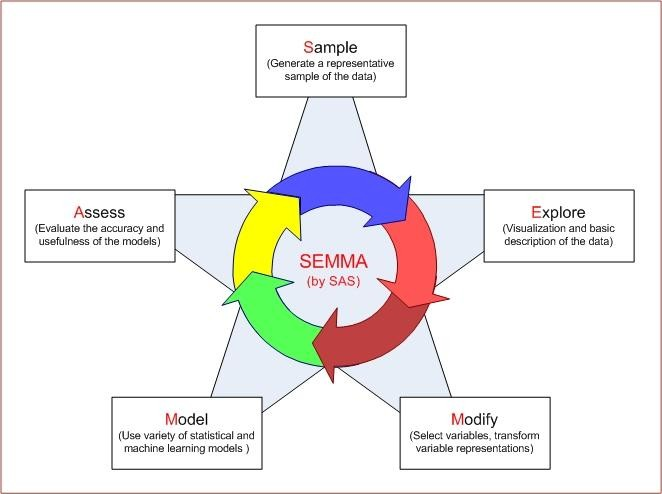
Quando o projeto começa a desviar do planejamento, usamos recursos para pressionar tal desvio ao caminho principal. No entanto, temos apenas uma vaga ideia do objetivo e podemos desprezar aprendizados importantes para uma ideia de sucesso diferente.

Podemos usar um modelo iterativo. Nesse modelo, fazemos pequenos desvios e calculamos novos resultados, sem desprezar os aprendizados e definindo novos alcances (até mesmo melhores do que o planejado).

Projeto empírico (vai aprendendo durante o desenvolvimento).

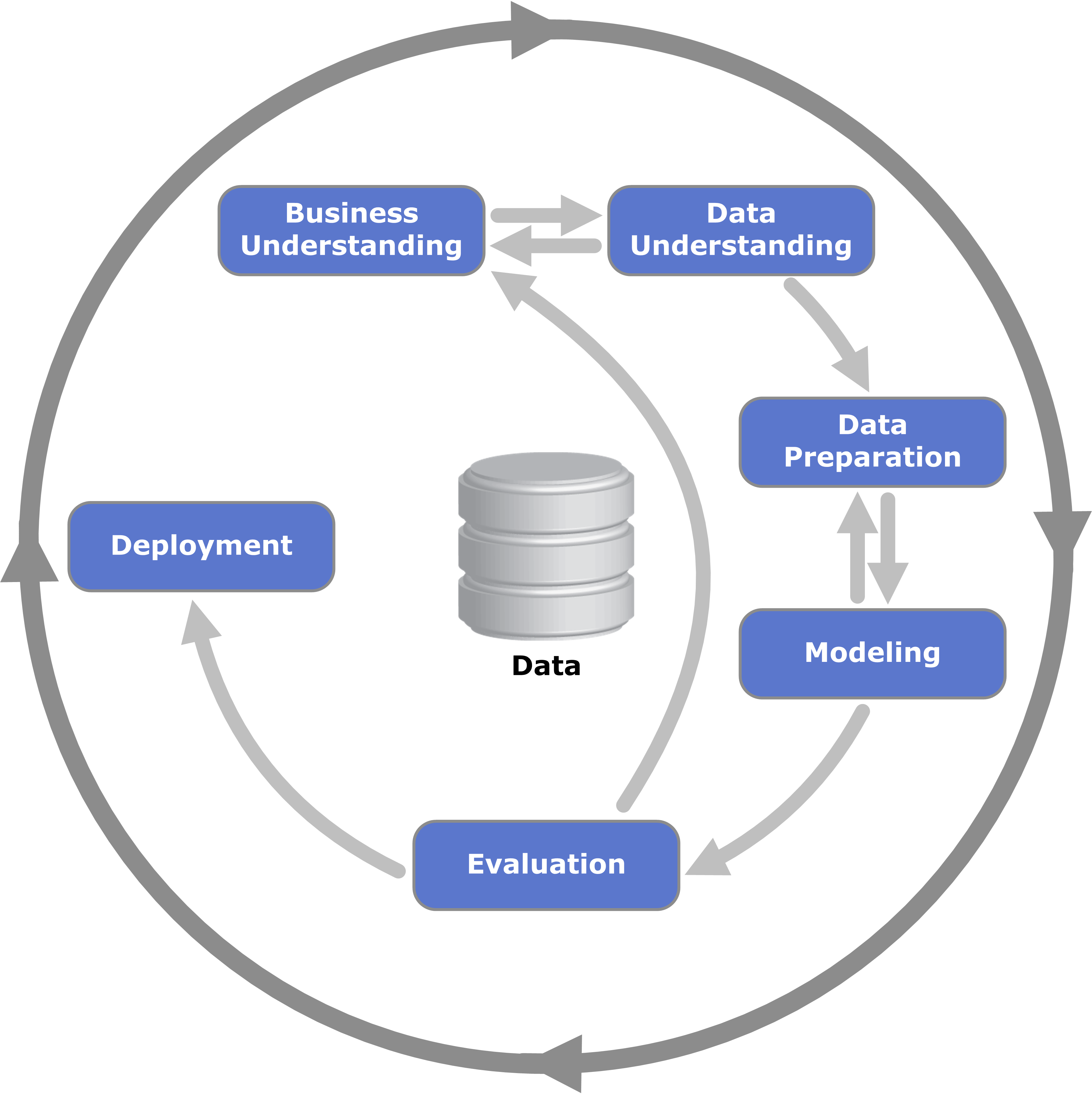
Gastamos mais dinheiro num ambiente determinístico (o campo de dados é volátil a mudanças).

**CICLO DE REFINAMENTO:**



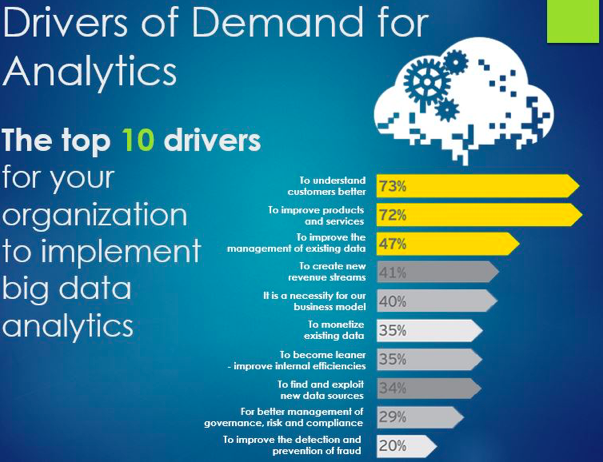
CRISP-DM, semma refinado

SEMMA



**MOTIVOS PARA IMPLEMENTAR ANÁLISE DE DADOS:**

Prever o comportamento do grupo (a média é sempre previsível).



MEDO: Dar problema (vazamento de informações); empresas pequenas (sem dinheiro para investimento); sem contato com dados; medo de novos padrões;

DESAFIOS: Não sabe lidar com o valor dos dados ou não conhece (sem competência) para fazer a transição; Segurança e análise (qualquer erro leva a empresa à falência); Converter dados em valor de mercado; Profissionais não qualificados (resistência para contratar cientistas de dados); Alta demanda e poucos profissionais; Sem investimento.