

# Crisp-DM in a Nutshell

## O Caos!



**80% dos projetos de DS não são colocados em produção!**  
[Gartner]

Causas mais comuns:

Não atende a uma necessidade do cliente



Problemas nos dados descobertos tardiamente

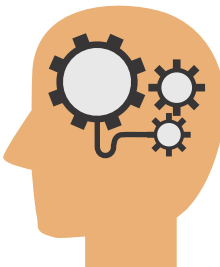
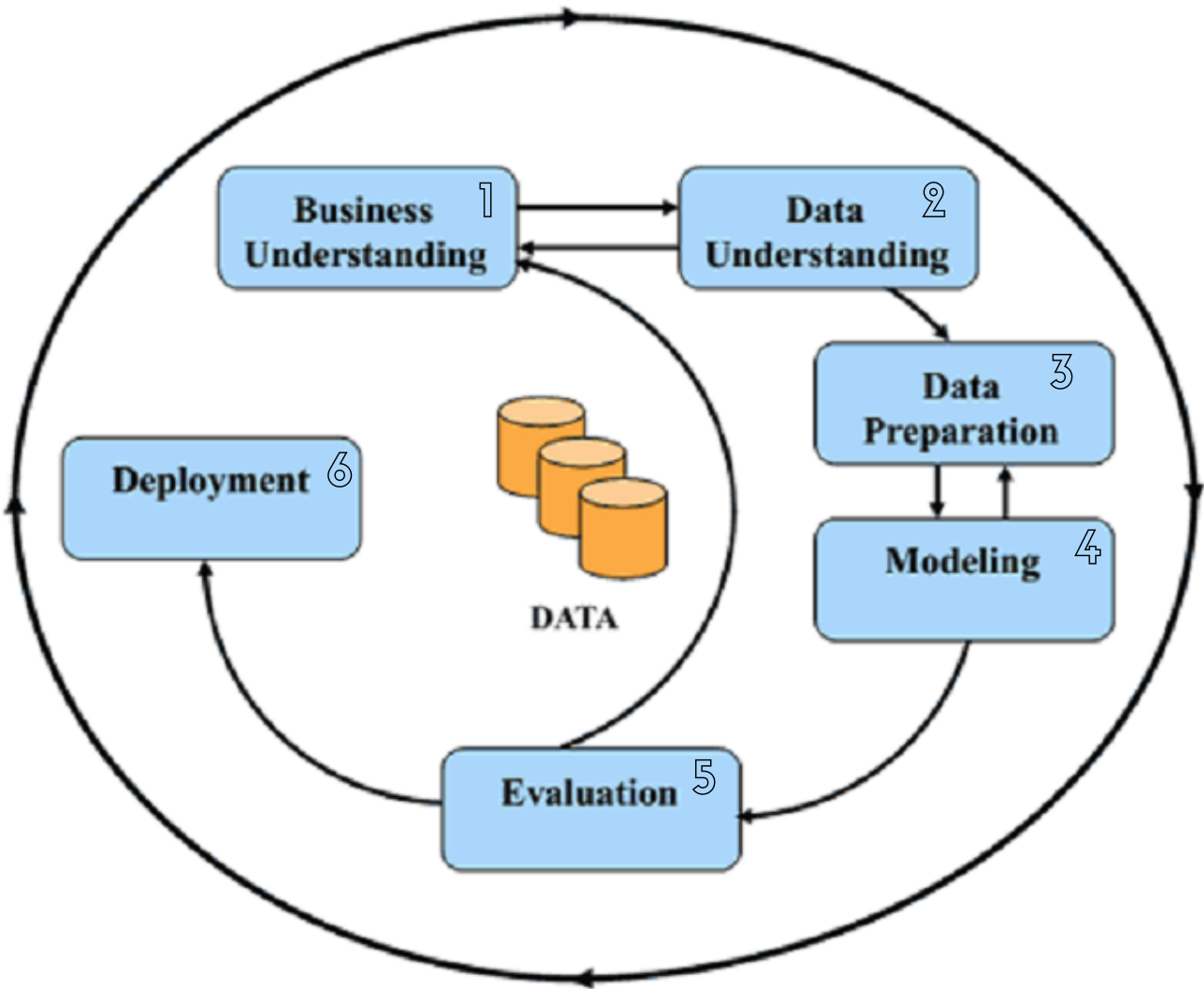
Modelos muito complexos

## A solução!

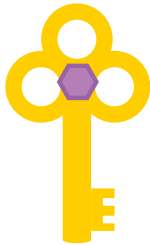
Utilizar uma **metodologia** para o desenvolvimento do projeto. As metodologias mais conhecidas são o KDD e o Crisp-DM. O primeiro não foca diretamente no negócio nem na criação de um modelo. Além disso não fomenta a iteração entre as etapas. Por isso vamos de **CRISP-DM**!

## Cross-Industry Standard Process for Data Mining

Guia em 6 passos para planejar, organizar, executar e analisar um projeto de data science



Manter a organização e ter uma linha de raciocínio clara, evitando surpresas desagradáveis ao final do projeto



**Iteração entre as etapas!**

1 - entender profundamente o negócio, as áreas envolvidas, o objetivo do uso de DS nesse negócio e as pessoas envolvidas



2 - intimamente ligado ao entendimento do negócio e a proposta de aplicação.

- Quais dados são necessários?
- Coletar, explorar e verificar qualidade

3 - preparar o dado para a modelagem:

- Limpar os dados
- Gerar novos atributos
- Integrar dados



4 - Encontrar respostas para o nosso problema!

- Aplicar a técnica que pode responder o problema do negócio

5 - Avaliar do ponto de vista de DS (quais métricas fazem sentido aqui?) e principalmente avaliar do ponto de vista do negócio



6 - Se a avaliação estiver satisfatória o modelo está pronto para ser utilizado! Caso contrário o ciclo se repete.

