

# Metodologias de Dados

## Tipos de Metodologias

### 1. CRISP-DM

#### a. Representação e Fases:

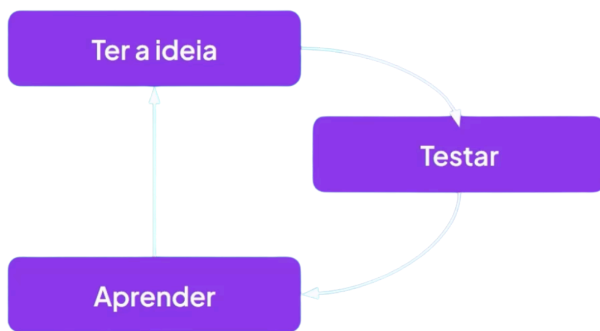


- i. Business Understanding -> Compreensão de Negócio
  - a. Definição de problemas e objetivos.
- ii. Data Understanding -> Compreensão do Dado
  - a. Que dados são necessários para responder essa pergunta de negócio;
  - b. EDA -> Exploratory Data Analysis.
- iii. Data Preparation -> Preparação dos Dados
  - a. Normalização dos dados se necessário;
- iv. Modeling -> Modelagem dos Dados;
  - a. Técnicas de Estatísticas e Machine Learning.
    - i. Inputa os dados.
  - b. Há casos que na modelagem descubra problemas nos dados e é necessário voltar para o DP -> Data Preparation.
- v. Evaluation -> Validação
  - a. Quando é validado pelos stakeholders.
- vi. Deployment -> Implementação

- a. Fazer o Deploy -> Implementar para a automação por exemplo de uma análise feita por um modelo preditivo siga nos próximos momentos sendo perene.

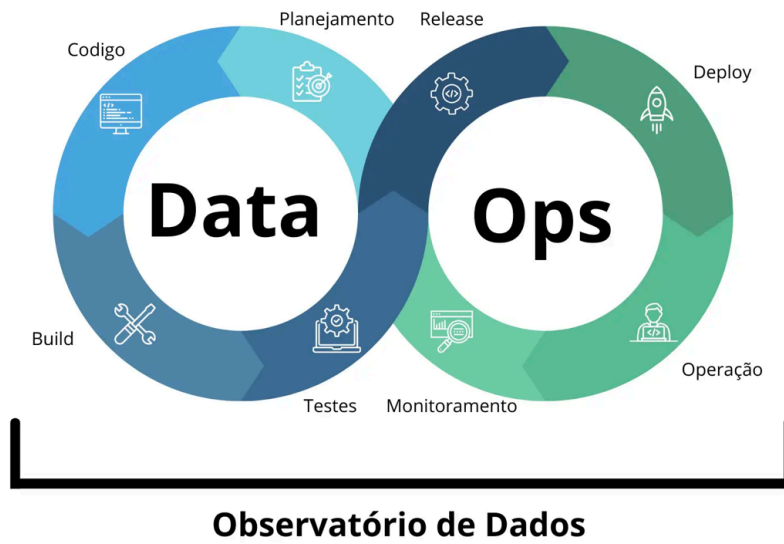
## 2. LEAN ANALYTICS

- a. Análise contínua para tomada de decisão;
- b. Representação:



- c. Normalmente usado para diminuir o tempo de execução;
  - d. Análises contínuas para tomadas de decisões mais rápidas.
  - e. Princípios :
    - 1. Definir Métricas e KPI's
    - 2. Medir constantemente os dados para hipóteses
  - f. Usos hoje: Netflix (Thumbnail), Ecommerce (Carrinho abandonado), Uber (Preços);
- ## 3. DATAOPS
- a. DevOps aplicado a dados;

b. Representação e Fases:



c. Mais usados por Engenheiros e Cientistas de Dados.

d. Componentes:

1. Integração e Entrega Contínua (CI/CD) -> Automatização de processos de desenvolvimento de software .
  - i. Objetivos:
    1. Acelerar o ciclo de desenvolvimento
    2. Melhorar a qualidade de software.
    3. Reduzir Riscos.
  - ii. Como acontece CI/CD ?
    - a. CI -> Integração Contínua.
      - i. Integrar alterações de código em um repositório compartilhado várias vezes ao dia. Cada integração é automaticamente testada por meio de testes automatizados para garantir que as alterações não introduzam erros ou conflitos no código existente.
    - b. CD -> Entrega Contínua.
      - i. Etapa que dá continuidade à CI, buscando automatizar entrega e/ou implantação das alterações de código em um ambiente de produção. Automatiza o processo de entrega, preparando o código para ser lançado, enquanto a implementação continua libera automaticamente as alterações para os usuários finais.
2. Automação nos processos: Coleta, processamento e análise dos dados;
3. Monitoramento Contínuo;

#### 4. Governança dos Dados.

#### 4. A/B TESTING

- a. Experimentação baseada em dados;
- b. Método de experimentação baseado em dados, com finalidade de testar variações para tomar decisões embasadas;
- c. Representação:



- d. Mais usado em UX (User eXperience)
- e. Orientação:
  - 1. Definir uma hipótese do teste;
  - 2. Criar duas ou mais versões;
  - 3. Testar com usuários;
  - 4. Escolher a melhor hipótese.
- f. Projetos B2C (Business to Customer) -> Empresa para o Usuário.