



INTRODUÇÃO À CIÊNCIA DE DADOS



Recapitulando!!

ROTEIRO:

1

Objetivo

2

Big Data e o Pensamento
Analítico

3

Etapas de um Projeto de DS

4

Ambientes de Trabalho





Objetivos:

1. Descrever o que é Big Data;
2. Introduzir o Pensamento Analítico;
3. Definir as Etapas de um Projeto de DS;
4. Demonstrar os possíveis Ambientes de Trabalho.



A população mundial da internet

59% dos habitantes do planeta acessam a internet - em bilhões de pessoas



Fonte: Domo

www.circuitod.com.br

Data Never Sleeps 8.0

**Dados produzidos em
apenas um minuto**

<https://www.domo.com/learn/data-never-sleeps-8>

Qual o problema?



Mas o que é:

BIG DATA



Os 4Vs

1

Volume – refere-se ao tamanho dos dados.

2

Variety – refere-se ao formato dos dados.

3

Velocity – refere-se à velocidade de streaming dos dados.

4

Veracity – refere-se sobre os dados serem confiáveis ou não.





80%

NÃO ESTRUTURADO

12%

ESTRUTURADO







Pensamento Analítico.

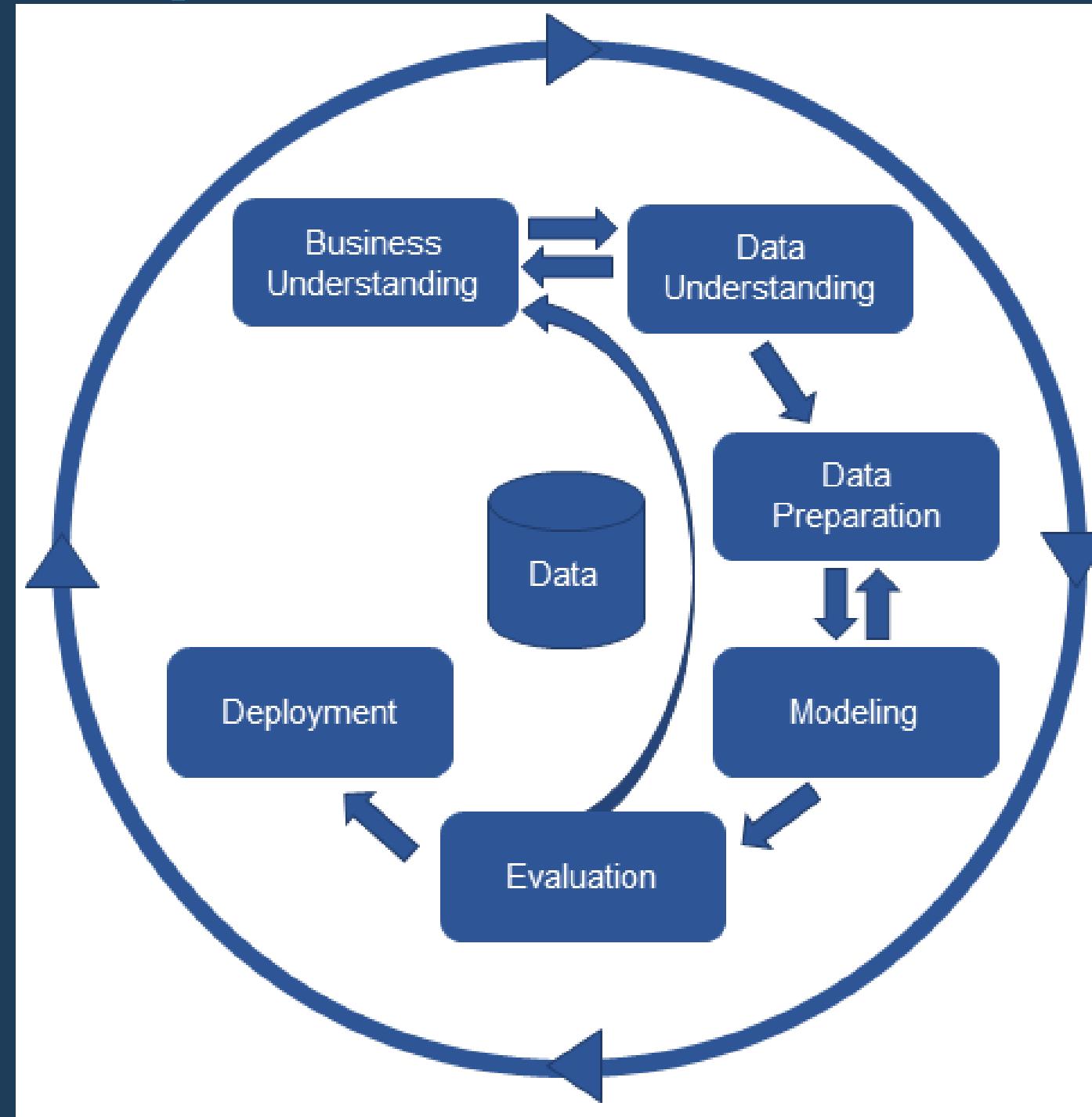
Data Science

Etapas

-
- ```
graph TD; A(()) --- B(()); B --- C(()); C --- D(()); D --- E(()); E --- F(());
```
- PROBLEMA DE NEGÓCIO
  - DADOS
  - MODELAGEM
  - TREINAMENTO DO MODELO
  - TESTE E AVALIAÇÃO DO MODELO

# Data Science

## Etapas



Modelo CRISP-DM

PROBLEMA DE NEGÓCIO

DADOS

MODELAGEM

TREINAMENTO DO MODELO

TESTE E AVALIAÇÃO DO  
MÓDELO

# Ambientes

## Linguagens



Python



R

# Ambientes

## Editor de Texto/IDE



Sublime Text



PyCharm



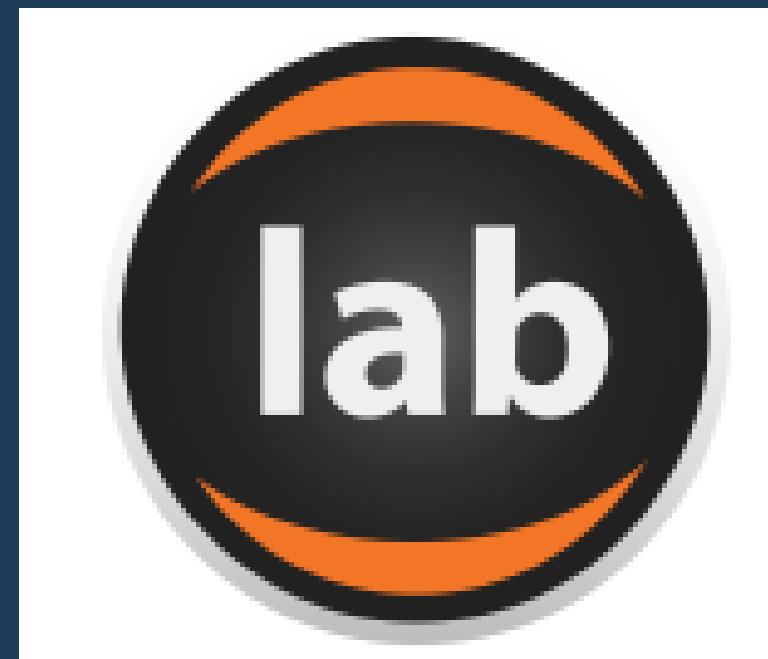
VSCode

# Ambientes

## Nosso Laboratório



Anaconda



Jupyter Lab



Jupyter Notebook

# Ambientes

## Nosso Laboratório



Google Colaboratory



Kaggle

# NÚMEROS que salvam VIDAS.



**MAURICIO ELOY**  
A ARTE DE ENCONTRAR PADRÕES