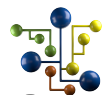
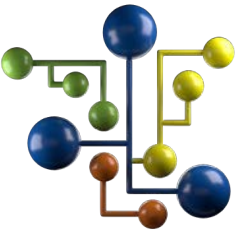


Data Science  
Academy

Data Science Academy andrewborges71@gmail.com 5b5d07a25c4cde026731456e

# Big Data Real-Time Analytics com Python e Spark





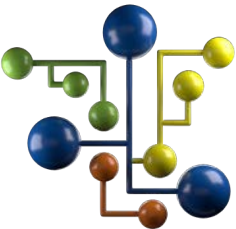
Data Science  
Academy

Data Science Academy andrewborges71@gmail.com 5b5d07a25e4cde02668b456e

# Big Data Real-Time Analytics com Python e Spark

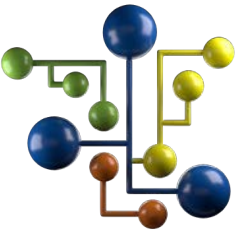
**Seja muito bem-vindo(a)!**





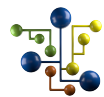
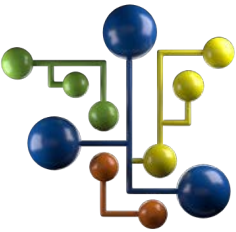
# Introdução

## Engenharia de Atributos

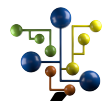
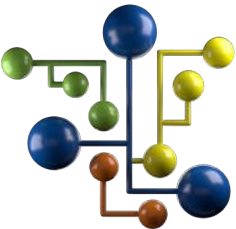


# Introdução

- O curso é inteiramente sequencial.
- Conclua os capítulos anteriores antes de avançar.
- Seguiremos trabalhando em um grande projeto distribuído ao longo de 3 capítulos com 3 Estudos de Caso:
  - Análise Exploratória de Dados
  - Engenharia de Atributos
  - Pré-Processamento de Dados



# O Que é Engenharia de Atributos?



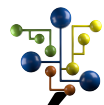
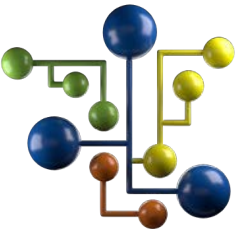
# O Que é Engenharia de Atributos?

ID_Paciente	Idade	Peso (kg)	Altura (m)	Diagnóstico
1	45	96	1.75	Saudável
2	52	110	1.67	Doente
3	29	88	1.90	Saudável

ID de Cada  
Registro

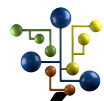
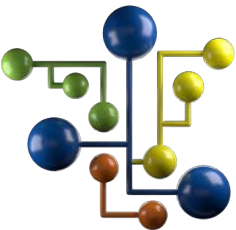
Variáveis de Entrada  
(Preditores)

Variável de Saída  
(Target/Alvo)



# O Que é Engenharia de Atributos?

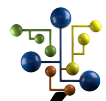
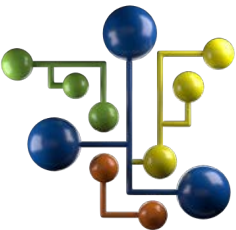
Idade	Peso (kg)	Altura (m)	Diagnóstico
45	96	1.75	Saudável
52	110	1.67	Doente
29	88	1.90	Saudável



# O Que é Engenharia de Atributos?

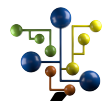
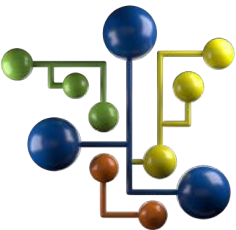
Faixa Etária	Peso (kg)	Altura (m)	Diagnóstico
De 40 a 49 Anos	96	1.75	Saudável
De 50 a 59 Anos	110	1.67	Doente
De 20 a 29 Anos	88	1.90	Saudável





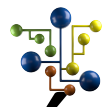
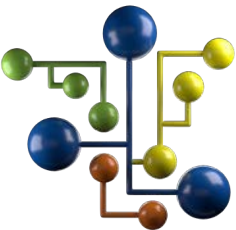
# O Que é Engenharia de Atributos?

Faixa Etária	Peso (kg)	Altura (m)	IMC (Peso / Altura * Altura)	Diagnóstico
De 40 a 49 Anos	96	1.75	31.3	Saudável
De 50 a 59 Anos	110	1.67	39.4	Doente
De 20 a 29 Anos	88	1.90	24.4	Saudável



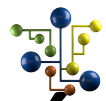
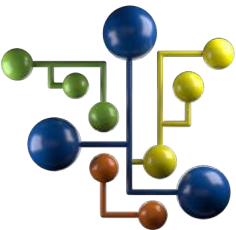
# O Que é Engenharia de Atributos?

Faixa Etária	IMC (Peso / Altura * Altura)	Diagnóstico
De 40 a 49 Anos	31.3	Saudável
De 50 a 59 Anos	39.4	Doente
De 20 a 29 Anos	24.4	Saudável



# O Que é Engenharia de Atributos?

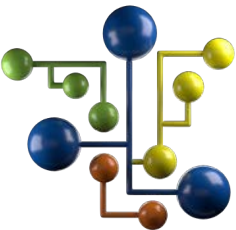
Faixa Etária	IMC (Peso / Altura * Altura)	Obesidade	Diagnóstico
De 40 a 49 Anos	31.3	Sim	Saudável
De 50 a 59 Anos	39.4	Sim	Doente
De 20 a 29 Anos	24.4	Não	Saudável



# O Que é Engenharia de Atributos?

Faixa Etária	IMC (Peso / Altura * Altura)	Obesidade	Diagnóstico
De 40 a 49 Anos	31.3	Sim	Saudável
De 50 a 59 Anos	39.4	Sim	Doente
De 20 a 29 Anos	24.4	Não	Saudável

← Isso é Engenharia de Atributos!

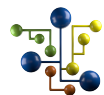


# O Que é Engenharia de Atributos?

Faixa Etária	IMC (Peso / Altura * Altura)	Obesidade	Diagnóstico
1	31.3	1	1
2	39.4	1	0
3	24.4	0	1

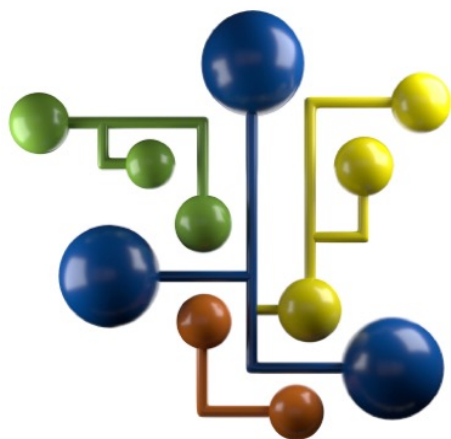
Isso é Pré-  
Processamento!





Data Science  
Academy

Data Science Academy andrewborges7@gmail.com 5b5d07a25e4cde02668b456e



Tenha uma Excelente Jornada de Aprendizagem.

Muito Obrigado por Participar!

Equipe Data Science Academy