

FIAP

NBA



# Data Processing With Python

Thiago Nascimento Nogueira

- Bacharel, mestre e doutor em física pela USP
- Trabalha com ambiente Linux desde 1995
- Atuação em sistemas distribuídos desde 2007
- Big Data desde 2016
- Ciência de Dados desde 2017
- Computação Quântica desde 2021
- IA Generativa desde 2020

- Parte expositiva
  - Apresentação geral da solução
  - Introdução aos conceitos principais
  - “Caminho das pedras”
- Laboratório



# Projeto Integrado

## DS&IA - Data Processing With Python





## Projeto Integrado: Data Processing with Python

### Enunciado

A QuantumFinance possui um modelo de risco de crédito desatualizado, que tem concedido cartões para um número muito alto de mau pagadores e gerando problemas financeiros para a companhia.

O objetivo é realizar uma análise exploratória para compreender o comportamento dos dados e tentar descobrir um padrão que explique e diferencie clientes bons de maus pagadores.

Seu trabalho é:

- Explorar e analisar os dados para entender a distribuição das características e a relação com a classificação de cada cliente.
- Responder as seguintes perguntas:
  - Como as features sexo, estado civil, quantidade de dependentes e nível educacional está relacionada com a variável meta?
  - Plote um gráfico que mostre a distribuição de bons e maus pagadores por estado. Escolha o gráfico que achar mais conveniente
  - Use boxplot para explicar como o tipo de residência interfere na classificação do pagador
  - Como as features renda pessoal mensal e outras rendas se relacionam com a variável meta? (obs: não é correlação estatística)
  - A idade interfere na classificação do pagador?

### Entregáveis:

- Jupyter Notebook: O notebook deverá estar bem organizado com seções claras e textos que facilitem a compreensão da análise e decisões tomadas e que permita a obtenção do resultado final a partir do dataset disponibilizado

### Dataset:

- <https://raw.githubusercontent.com/diogenesjusto/FIAP/master/dados/credit.csv>

## Data Processing with Python

Validado  
2024-2

## DATA SET - Dicionário de dados

Var_id	Característica	Descrição
1	ID_CLIENTE	Número sequencial para o requerente (a ser usado como chave)
2	TIPO_FUNCIONARIO	Não informado
3	DIA_PAGAMENTO	Dia do mês para o pagamento da fatura, escolhido pelo requerente
4	TIPO_ENVIO_APLICACAO	Indica se a aplicação foi submetida via internet ou pessoalmente/enviada pelo correio
5	QUANT_CARTOES_ADICIONAIS	Quantidade de cartões adicionais solicitados no mesmo formulário de aplicação
6	TIPO_ENDEREÇO_POSTAL	Indica se o endereço para envio é o endereço residencial ou outro. Codificação não informada.
7	SEXO	
8	ESTADO_CIVIL	Codificação não informada
9	QUANT_DEPENDENTES	
10	NIVEL_EDUCACIONAL	Nível educacional em ordem gradual não informado
11	ESTADO_NASCIMENTO	
12	CIDADE_NASCIMENTO	
13	NACIONALIDADE	País de nascimento. Codificação não informada, mas é provável que o Brasil seja igual a 1.
14	ESTADO_RESIDENCIAL	Estado de residência
15	CIDADE_RESIDENCIAL	Cidade de residência
16	BAIRRO_RESIDENCIAL	Bairro de residência
17	FLAG_TELEFONE_RESIDENCIAL	Indica se o requerente possui um telefone residencial
18	CODIGO_AREA_TELEFONE_RESIDENCIAL	Pseudo-código de três dígitos
19	TIPO_RESIDENCIA	Codificação não informada. Em geral, existem tipos como: própria, hipotecada, alugada, dos pais, familiar, etc.
20	MESES_RESIDENCIA	Tempo na residência atual em meses
21	FLAG_TELEFONE_MOVEL	Indica se o requerente possui um telefone móvel
22	FLAG_EMAIL	Indica se o requerente possui um endereço de e-mail
23	RENTA_PESSOAL_MENSAL	Renda mensal regular pessoal do requerente em moeda brasileira (R\$)
24	OUTRAS_RENTAS	Outras rendas mensais médias do requerente em moeda brasileira (R\$)
25	FLAG_VISA	Indicador se o requerente é titular de um cartão de crédito VISA

## Data Processing with Python

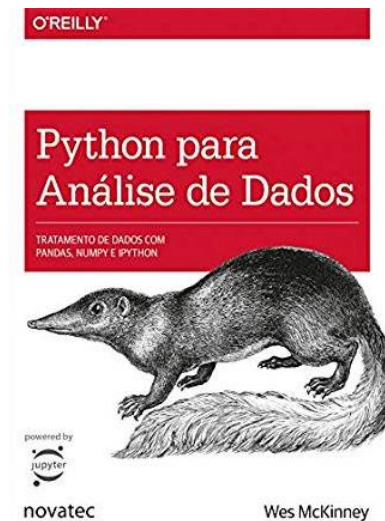
### DATA SET - Dicionário de dados

Validado  
2024-2

Var_id	Característica	Descrição
26	FLAG_MASTERCARD	Indicador se o requerente é titular de um cartão de crédito MASTERCARD
27	FLAG_DINERS	Indicador se o requerente é titular de um cartão de crédito DINERS
28	FLAG_AMERICAN_EXPRESS	Indicador se o requerente é titular de um cartão de crédito AMERICAN EXPRESS
29	FLAG_OUTROS_CARTOES	Apesar de ser rotulado como "FLAG", esse campo apresenta três valores não explicados
30	QUANT_CONTAS_BANCARIAS	
31	QUANT_CONTAS_BANCARIAS_ESPECIAIS	
32	VALOR_PATRIMONIO_PESSOAL	Valor total dos bens pessoais, como casas, carros, etc., em moeda brasileira (R\$)
33	QUANT_CARROS	Quantidade de carros que o requerente possui
34	EMPRESA	Se o requerente forneceu o nome da empresa onde ele/ela trabalha formalmente
35	ESTADO_PROFISSIONAL	Estado onde o requerente trabalha
36	CIDADE_PROFISSIONAL	Cidade onde o requerente trabalha
37	BAIRRO_PROFISSIONAL	Bairro onde o requerente trabalha
38	FLAG_TELEFONE_PROFISSIONAL	Indica se o número de telefone profissional foi fornecido
39	CODIGO_AREA_TELEFONE_PROFISSIONAL	Pseudo-código de três dígitos
40	MESES_NO_TRABALHO	Tempo no emprego atual em meses
41	CODIGO_PROFISSAO	Código profissional do requerente. Codificação não informada
42	TIPO_OCUPACAO	Codificação não informada
43	CODIGO_PROFISSAO_CONJUGE	Código profissional do cônjuge. Codificação não informada
44	NIVEL_EDUCACIONAL_CONJUGE	Nível educacional do cônjuge em ordem gradual não informado
45	FLAG_DOCUMENTO_RESIDENCIAL	Indicador de confirmação documental do endereço residencial
46	FLAG_RG	Indicador de confirmação documental do número do cartão de cidadão
47	FLAG_CPF	Indicador de confirmação documental do status de contribuinte
48	FLAG_COMPROVANTE_RENDA	Indicador de confirmação documental da renda
49	PRODUTO	Tipo de produto de crédito aplicado. Codificação não informada
50	FLAG_REGISTRO_ACSP	Indicador se o requerente possui alguma inadimplência de crédito anterior
51	IDADE	Idade do requerente no momento da inscrição
52	CEP_RESIDENCIAL_3	Três dígitos mais significativos do CEP residencial atual
53	CEP_PROFISSIONAL_3	Três dígitos mais significativos do CEP do emprego atual
54	ROTULO_ALVO_MAU=1	Variável Alvo: BAD=1, BOM=0



## Referência: Python



# Contatos



/vagner-dos-santos

profvagner.macedo@fiap.com.br



Copyright © 2019 | Professor (a) Vagner S. Macedo

Todos os direitos reservados. Reprodução ou divulgação total ou parcial deste documento, é expressamente proibido sem consentimento formal, por escrito, do professor/autor.

FIAP