



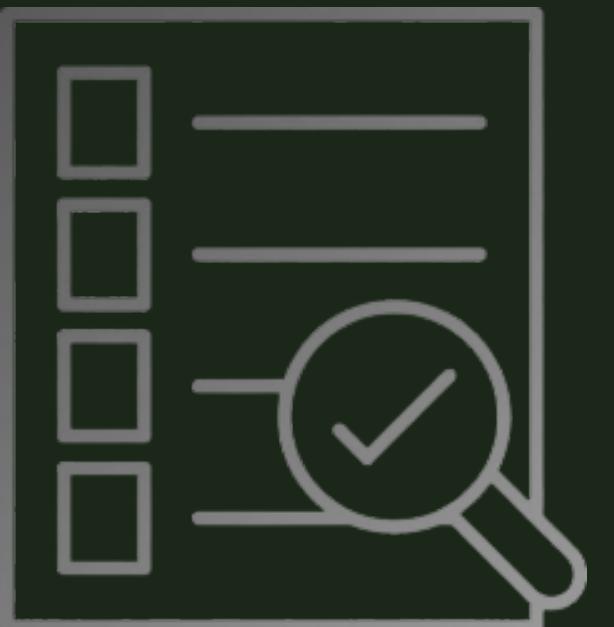
# STORYTELLING

## Ciência de Dados

MATHEUS, GABRIELA, JUAN E GIOVANNI

# ÍNDICE

- Introdução: Sobre a base de dados e seus tratamentos
- Respondendo perguntas
- Conclusão





# SOBRE A BASE DE DADOS

---

## Inadimplência de Clientes de Cartão de Crédito

- Dados de um banco em Taiwan
- Informações de titulares de cartão de crédito coletadas em outubro de 2005
- Histórico de informações de abril a setembro de 2005
- Há informações sobre o limite de crédito concedido, sexo, escolaridade, estado civil, idade, status de pagamento mensal, valores de faturas e pagamentos, e se o titular é inadimplente ou não
- Dados com foco em aprendizado de máquina



# TRATAMENTOS

- De 30.000 linhas, ficamos com 29.685 após os tratamentos
- Criamos duas colunas extras contendo a faixa das idades e de crédito dos titulares
- Convertemos alguns valores numéricos em texto, para uma análise e apresentação mais compreensiva

Valor	SEX	EDUCATION	MARRIAGE	PAY_1 ... PAY_6	default payment next month
-2	-	-	-	Desconhecido	-
-1	-	-	-	Pagamento pontual	-
0	Sem info.	Sem info.	Sem info.	Sem info.	Não
1	Masculino	Pós-graduação	Casado	Atraso de 1 mês	Sim
2	Feminino	Universidade	Solteiro	Atraso de 2 meses	-
3	-	Ensino Médio	Outro	...	-
4	-	Outro	-	...	-
5	-	Desconhecido	-	...	-
6	-	Desconhecido	-	...	-
...	-	-	-		-
9	-	-	-	Atraso de 9 meses	-



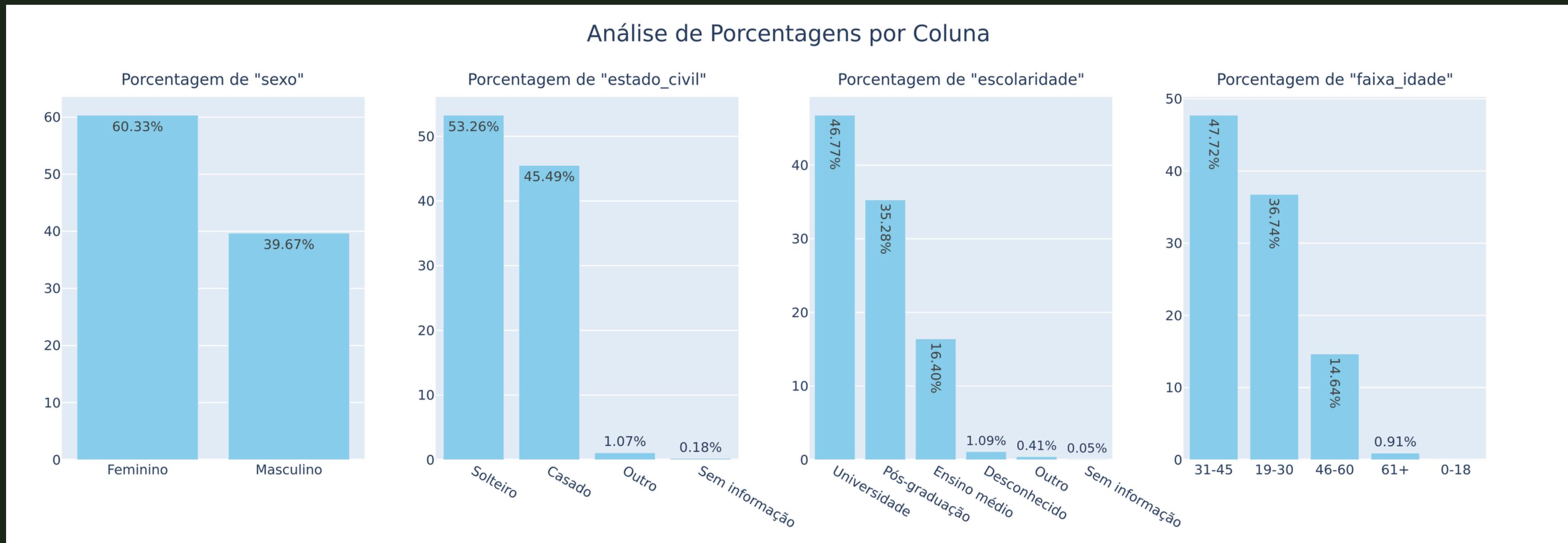
# TRATAMENTOS

- Também traduzimos os nomes de cada coluna e deixamos com um nome mais intuitivo, a fim de ter algo mais legível

	Coluna em Inglês	Coluna Traduzida
0	ID	id
1	LIMIT_BAL	qtd_credito
2	SEX	sexo
3	EDUCATION	escolaridade
4	MARRIAGE	estado_civil
5	AGE	idade
6	PAY_1	status_pagamento_setembro_2005
7	PAY_2	status_pagamento_agosto_2005
8	PAY_3	status_pagamento_julho_2005
9	PAY_4	status_pagamento_junho_2005
10	PAY_5	status_pagamento_maio_2005
11	PAY_6	status_pagamento_abril_2005
12	BILL_AMT1	fatura_setembro_2005
13	BILL_AMT2	fatura_agosto_2005
14	BILL_AMT3	fatura_julho_2005
15	BILL_AMT4	fatura_junho_2005
16	BILL_AMT5	fatura_maio_2005
17	BILL_AMT6	fatura_abril_2005
18	PAY_AMT1	pagamento_setembro_2005
19	PAY_AMT2	pagamento_agosto_2005
20	PAY_AMT3	pagamento_julho_2005
21	PAY_AMT4	pagamento_junho_2005
22	PAY_AMT5	pagamento_maio_2005
23	PAY_AMT6	pagamento_abril_2005
24	default payment next month	inadimplente

# INFORMAÇÕES GERAIS

## Distribuição das categorias dentro do dataset



# NOSSAS QUESTÕES



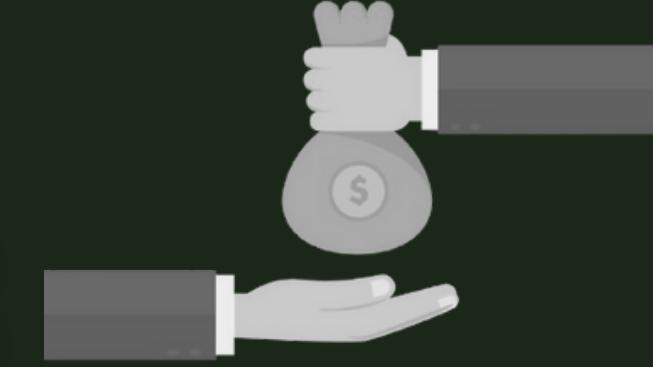
## Linha de Crédito

Como as variáveis categóricas estão relacionadas com a Linha de Crédito recebida?

## Inadimplência

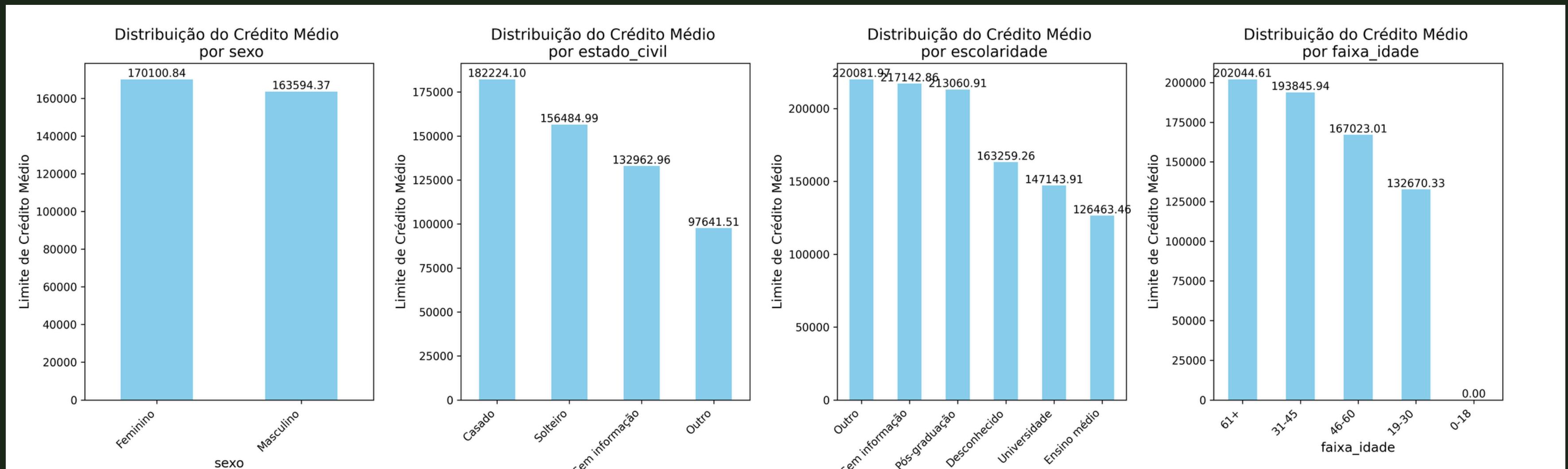
Quais as principais características que uma pessoa inadimplente possui?

# Linha de Crédito

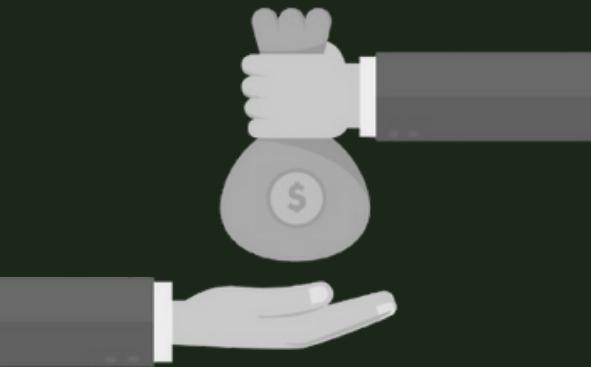


Começamos avaliando individualmente o sexo, estado civil, escolaridade e faixa etária do titular.

Primeiro, buscamos entender se a média de crédito em alguma dessas categorias era muito distinta das outras:



# Linha de Crédito



Porcentagem para a coluna 'estado\_civil':

	faixa_credito	Casado	Outro	Sem informação	Solteiro
0		0	0.00	0.00	100.00
1	1 - 100.000	39.99	1.88	0.20	<u>57.92</u>
2	100.001 - 250.000	46.52	0.69	0.23	52.56
3	250.001 - 500.000	52.22	0.41	0.08	47.29
4	500.001+	<u>55.90</u>	0.22	0.00	43.89

Porcentagem para a coluna 'faixa\_idade':

	faixa_credito	0-18	19-30	31-45	46-60	61+
0		0	100.0	0.00	0.00	0.00
1	1 - 100.000	0.0	<u>47.03</u>	36.01	16.18	0.78
2	100.001 - 250.000	0.0	<u>35.15</u>	51.48	12.57	0.79
3	250.001 - 500.000	0.0	24.00	60.09	14.66	1.25
4	500.001+	0.0	13.54	<u>63.65</u>	21.18	<u>1.64</u>

Porcentagem para a coluna 'escolaridade':

	faixa_credito	Desconhecido	Ensino médio	Outro	Pós-graduação	\	Sem informação	Universidade
0		0	0.00	0.00	0.00	\	100.00	0.00
1	1 - 100.000	0.98	<u>22.99</u>	0.11	20.93	\	0.01	<u>54.99</u>
2	100.001 - 250.000	1.32	<u>14.02</u>	0.62	39.34	\	0.07	<u>44.62</u>
3	250.001 - 500.000	0.94	9.69	0.56	50.64	\	0.08	38.09
4	500.001+	0.66	8.84	0.55	<u>60.48</u>	\	0.00	29.48

# Linha de Crédito

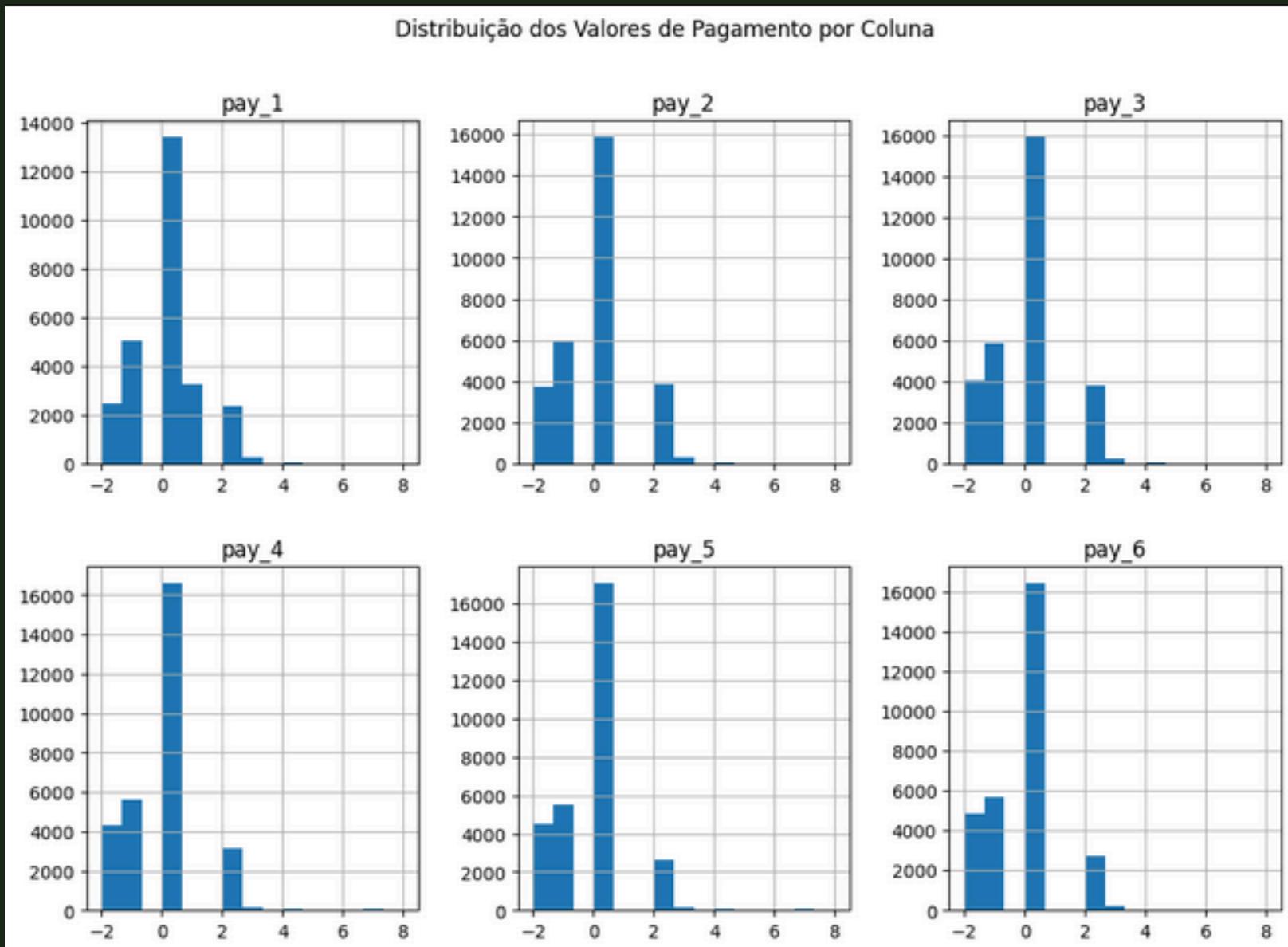


faixa_idade		0-18	19-30	31-45	46-60	61+
sexo	faixa_credito					
Feminino	0	0.0	0.00	0.00	0.00	0.00
	1 - 100.000	0.0	<u>45.47</u>	25.68	<u>41.52</u>	<u>38.52</u>
	100.001 - 250.000	0.0	<u>39.07</u>	<u>44.14</u>	<u>35.29</u>	<u>35.25</u>
	250.001 - 500.000	0.0	14.35	<u>26.22</u>	19.51	22.13
	500.001+	0.0	1.11	3.95	3.68	4.10
Masculino	0	0.0	0.00	0.00	0.00	0.00
	1 - 100.000	0.0	<u>55.12</u>	33.22	<u>42.91</u>	27.89
	100.001 - 250.000	0.0	<u>32.14</u>	<u>37.27</u>	<u>30.01</u>	<u>31.97</u>
	250.001 - 500.000	0.0	<u>11.57</u>	<u>25.15</u>	21.68	<u>33.33</u>
	500.001+	0.0	1.18	4.35	5.40	<u>6.80</u>
Sem informação	0	100.0	0.00	0.00	0.00	0.00
	1 - 100.000	0.0	0.00	0.00	0.00	0.00
	100.001 - 250.000	0.0	0.00	0.00	0.00	0.00
	250.001 - 500.000	0.0	0.00	0.00	0.00	0.00
	500.001+	0.0	0.00	0.00	0.00	0.00

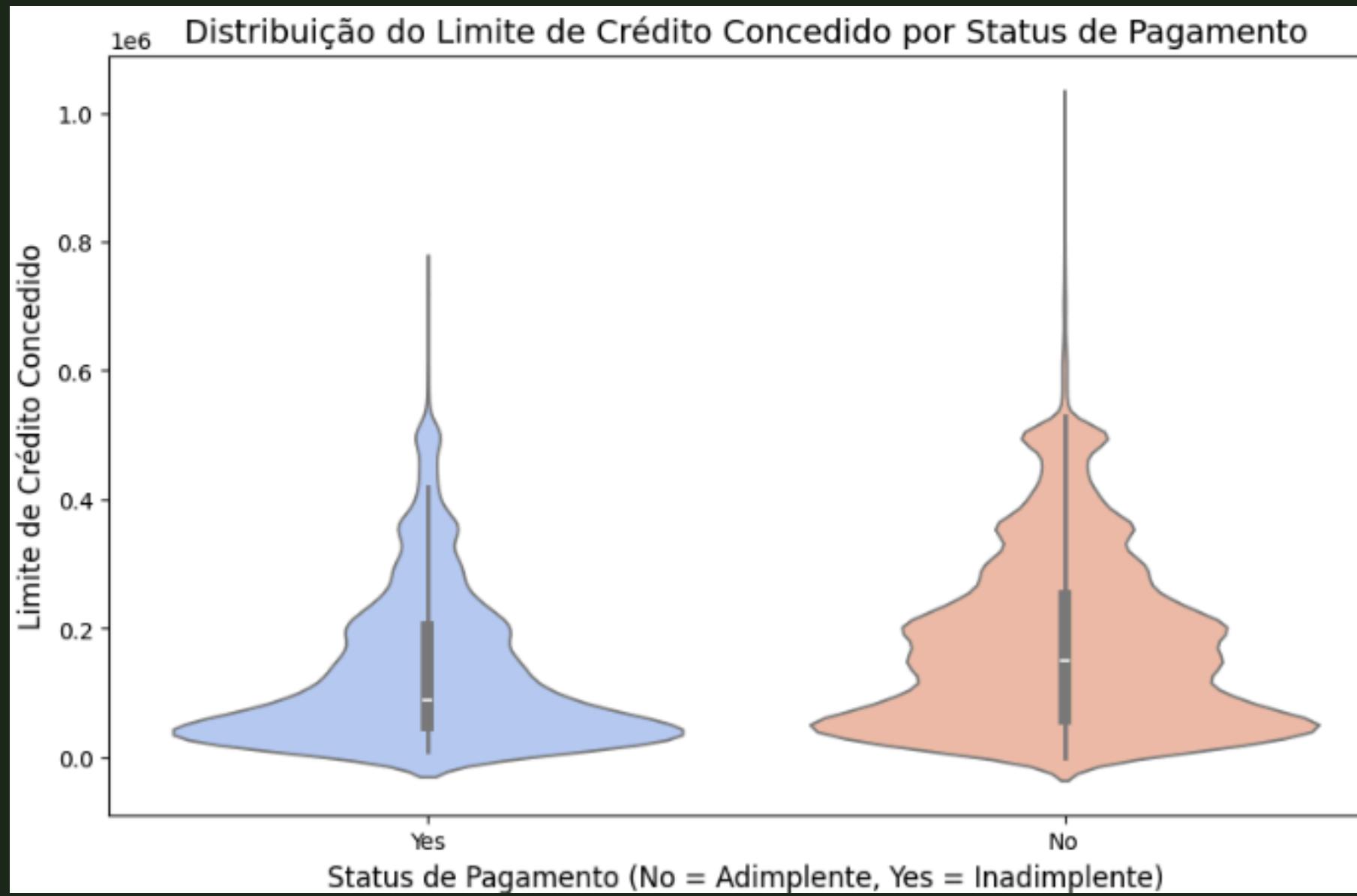
# Inadimplência



A análise revelou que **78,13%** dos indivíduos são adimplentes e **21,87%** inadimplentes. O estado civil, mais que o sexo, se mostrou relevante: solteiros tendem à inadimplência, enquanto casados, independentemente de sexo ou escolaridade, são mais adimplentes.



# Inadimplência



O atraso médio dos inadimplentes foi de dois meses, com a maioria recebendo crédito entre 0 e 200 mil dólares taiwaneses, e valores acima disso não influenciaram na inadimplência.

Dessa forma, a persona para o produto de crédito seria composta por homens e mulheres, com escolaridade a partir de "Graduate School" e estado civil casado, pois esse perfil demonstrou maior propensão à adimplência, independentemente do limite de crédito.

# Conclusões



- **Estado Civil:** Pessoas casadas tendem a ter uma faixa de crédito mais alta, enquanto solteiros geralmente têm uma faixa de crédito menor
- **Faixa Etária:** Indivíduos de 46-60 anos e acima de 61 anos são mais propensos a ter faixas de crédito mais altas.
- **Escolaridade:** A escolaridade parece estar relacionada com faixas de crédito mais altas, especialmente para aqueles com pós-graduação.

# Conclusões



- **Adimplência:** A maioria dos indivíduos (78,13%) é adimidente, enquanto 21,87% são inadimplentes.
- **Relevância do estado civil:** O estado civil teve mais impacto na inadimplência do que o sexo. Solteiros tendem a ser mais inadimplentes, enquanto casados, independentemente de sexo ou escolaridade, são mais adimplentes.
- **Atraso médio:** O atraso médio dos inadimplentes foi de dois meses.
- **Faixa de crédito:** A maioria dos inadimplentes recebeu crédito entre 0 e 200 mil dólares taiwaneses. Créditos acima desse valor não influenciaram a inadimplência.
- **Persona ideal:** A persona ideal para o produto de crédito são homens e mulheres casados, com escolaridade a partir de "Graduate School", pois esse perfil demonstrou maior adimplência, independentemente do limite de crédito.



MUITO OBRIGADO